

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA



DIVISIÓN DE INGENIERÍA MECÁNICA E INDUSTRIAL



Serie de Calidad y Estadística Industrial

# Fundamentos de Probabilidad y Aplicaciones, con R, Minitab y Excel

SEGUNDA EDICIÓN

OCTAVIO ESTRADA CASTILLO

Serie de Calidad y Estadística Industrial.

# Fundamentos de Probabilidad y Aplicaciones, con R, Minitab y Excel. Segunda Edición

Octavio Estrada Castillo

*Columna de Ingeniería,  
de Sebastián y  
vitral del maestro  
Narciso Quagliata*



Acrobat Reader  
Haz Click

ESTRADA CASTILLO, Octavio.  
*Fundamentos de Probabilidad y Aplicaciones,  
con R, Minitab y Excel. Segunda Edición*  
México, Universidad Nacional Autónoma de México,  
Facultad de Ingeniería, 361 p.  
ISBN 978-607-30-7586-2

---

FUNDAMENTOS DE PROBABILIDAD  
Y APLICACIONES, CON R, MINITAB Y EXCEL. SEGUNDA EDICIÓN

Segunda edición electrónica  
de un ejemplar (10 MB) en formato PDF  
Publicado en línea: 8 de mayo de 2023

D.R. © 2023, UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO  
Avenida Universidad 3000, Col. Universidad Nacional Autónoma de México,  
Ciudad Universitaria, Delegación Coyoacán, Ciudad de México, C.P. 04510

FACULTAD DE INGENIERÍA  
<http://www.ingenieria.unam.mx/>

ISBN 978-607-30-7586-2

Esta edición y sus características son propiedad de la Universidad Nacional Autónoma de México. Prohibida la reproducción o transmisión total o parcial por cualquier medio sin la autorización escrita del titular de los derechos patrimoniales.

Hecho en México.

---

UNIDAD DE APOYO EDITORIAL  
Cuidado de la edición: Patricia Eugenia García Naranjo  
Diseño editorial y formación: Nismet Díaz Ferro

Imagen de portada:  
Columna de Ingeniería, de Sebastián y vitral del maestro  
Narciso Quagliata, Fotografía: Luis Enrique Vite Rangel

*«Conserva celosamente tu derecho a reflexionar, porque incluso el hecho de pensar erróneamente es mejor que no pensar en absoluto».*

*Hipatia de Alejandría  
(370?-416?)*

## Prólogo

Desde 1988 he estado inmerso en la práctica profesional de la Ingeniería, particularmente, en la industria manufacturera, y en específico, en el área de Calidad y Estadística Industrial. Paralelamente, he hecho una trayectoria académica como profesor en la Facultad de Ingeniería de la UNAM desde hace más de 35 años. He impartido más de 25 asignaturas diferentes de matemáticas, física, computación, probabilidad, estadística, investigación de operaciones y calidad. Me ha tocado laborar en el sector público federal también.

Siempre he tenido la inquietud de escribir textos sobre las asignaturas en las que he participado, pero generalmente he tenido cargos académico-administrativos en la administración central de la UNAM o en la propia Facultad, que no me dejaban dedicarme a esta noble y gratificante labor. Ahora que me integro completamente como profesor de carrera en el área de calidad, investigación de operaciones y estadística industrial, he llevado a la praxis este deseo.

Este es el tercer libro de una serie sobre Calidad y Estadística Industrial, que tendrá, al menos, los siguientes títulos:

- I. Desarrollo histórico de la calidad.
- II. Metodología y herramientas para la solución de problemas y para la mejora continua.
- III. Fundamentos de probabilidad y aplicaciones con R, Minitab y Excel.
- IV. Fundamentos de estadística y aplicaciones con R, Minitab y Excel.
- V. Muestreo de aceptación y aplicaciones con R, Minitab y Excel.
- VI. Control estadístico de procesos y aplicaciones con R, Minitab y Excel.
- VII. Normatividad vigente sobre sistemas de calidad.
- VIII. Metrología, certificación de producto y certificación de software.
- IX. Teoría del muestreo.
- X. Estadística no paramétrica.
- XI. Diseño de experimentos.
- XII. Regresión y correlación.

- XIII. Confiabilidad.
- XIV. Estadística Multivariable.
- XV. Procesos Estocásticos.

El propósito de estos volúmenes es proporcionar la teoría necesaria, la metodología, las herramientas, ejemplos y aplicaciones prácticas de cada uno de los temas, de una manera formal, dinámica, amena y didáctica. Quisiera remarcar que en estos volúmenes hablo de mis conocimientos y experiencia en el apasionante tema de la calidad, la probabilidad y la estadística y que traté de apegarme lo más posible a citar a los autores originales de estas ideas, pero no debe olvidarse que se trata de un texto dirigido a alumnos, por lo cual no lleno de citas el texto para hacerlo más didáctico.



#### **OBJETIVO DE ESTE LIBRO:**

El alumno aplicará los conceptos, la metodología, las herramientas y técnicas de la probabilidad para analizar algunos fenómenos aleatorios que ocurren en la naturaleza, la sociedad y la industria, así como modelar y resolver problemas sujetos a incertidumbre.



## Índice temático

Prólogo.....	IV
Objetivo de este libro.....	V
Índice temático.....	VI
1 Desarrollo Histórico de la Probabilidad y la Estadística .....	1
Ejercicios propuestos del capítulo 1.....	7
2 Teoría Matemática de la Probabilidad .....	8
2.1 Conceptos básicos.....	8
2.1.1 Relación entre la probabilidad y la estadística, el método científico y el papel de la probabilidad y la estadística en la investigación básica y aplicada y sus etapas .....	15
2.2 Teoría de Conjuntos .....	18
2.3 Principios Fundamentales del Conteo y Análisis Combinatorio .....	24
2.4 Enfoques, Interpretaciones y Escuelas de Probabilidad.....	34
2.4.1 Escuela Clásica o de Laplace .....	34
2.4.2 Escuela Frecuentista o de Von Mises .....	37
2.4.3 Escuela Subjetivista o de Savage .....	40
2.4.4 Escuela Axiomática, Constructivista o de Kolmogorov.....	41
2.5 Probabilidad condicional, probabilidad de eventos independientes, probabilidad total y teorema de Bayes .....	44
Teorema o Ley de Probabilidad Total .....	52
Ejercicios propuestos del capítulo 2.....	57
3 Variable aleatoria, funciones de probabilidad de una variable aleatoria y sus propiedades.....	69

3.1	Variable Aleatoria Discreta, Función de Probabilidad y Función de Probabilidad Acumulada para una variable aleatoria discreta y sus propiedades . . . . .	70
3.2	Variable Aleatoria Continua, Función de Probabilidad y Función de Probabilidad Acumulada para una variable aleatoria continua y sus propiedades . . . . .	81
3.3	Percentiles, Deciles, Cuartiles, Medidas de una Función de Probabilidad y sus Propiedades . . . . .	85
3.4	Esperanza Matemática, Momentos, Valor Esperado, Varianza y sus Propiedades . . . . .	90
3.5	Función Generatriz de Momentos y Función Característica de una variable aleatoria. . . . .	99
3.6	Teorema de la Desigualdad de Chevshev. . . . .	115
3.7	Funciones de probabilidad de variables dependientes de variables aleatorias . . . . .	117
	Ejercicios propuestos del capítulo 3. . . . .	123
4.	Variables aleatorias conjuntas . . . . .	127
4.1.	Funciones de probabilidad de variables aleatorias conjuntas discretas . . . . .	129
	Propiedades de la Covarianza . . . . .	131
	Propiedades del Coeficiente de Correlación . . . . .	132
4.2.	Funciones de probabilidad de variables aleatorias conjuntas continuas. . . . .	140
4.3	Funciones de probabilidad de variables dependientes de variables aleatorias conjuntas . . . . .	146
	Procedimiento para deducir una función de probabilidad derivada de otras funciones de probabilidad conjuntas. . . . .	147
4.4.	Funciones de probabilidad con $n$ variables aleatorias conjuntas . . . . .	151
4.4.1	Funciones de probabilidad con $n$ variables aleatorias conjuntas discretas. . . . .	151
4.4.2	Funciones de probabilidad con $n$ variables aleatorias conjuntas continuas. . . . .	156
4.5	Características de combinaciones lineales de variables aleatorias . . . . .	161
	Ejercicios propuestos del capítulo 4. . . . .	162



5.	Modelos probabilísticos de Fenómenos Aleatorios Discretos. . . . .	168
5.1	Modelo probabilístico de Bernoulli . . . . .	170
5.2	Modelo Probabilístico Hipergeométrico $h(x;n,D,N)$ . . . . .	173
5.3	Modelo Probabilístico Binomial . . . . .	187
5.3.1	Aproximación de la distribución hipergeométrica a través de la binomial . . . . .	195
5.4	Modelo Probabilístico de Pascal o Binomial Negativo $bn(x; p, r)$ . . .	198
5.5	Modelo Probabilístico Geométrico $g(x; p)$ . . . . .	206
5.5.1	Propiedad de Olvido o de Falta de Memoria de la Distribución Geométrica . . . . .	210
5.6	Modelo Probabilístico de Poisson, $po(x; c)$ . . . . .	213
5.7	Modelo Probabilístico Multinomial, $p_m(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k)$ . . .	223
	Ejercicios propuestos del capítulo 5. . . . .	226
6	Modelos probabilísticos continuos de fenómenos aleatorios . . . . .	232
6.1	Distribución uniforme continua . . . . .	234
6.2	Distribución Triangular . . . . .	236
6.3	Distribución Exponencial Negativa . . . . .	239
6.3.1	Propiedad de Olvido o de Falta de Memoria de la Distribución Exponencial Negativa. . . . .	244
6.3.2	Distribución de probabilidad exponencial negativa desplazada. . . .	249
6.3.3	Función de Probabilidad de Laplace o Función de Probabilidad Exponencial de doble lado . . . . .	250
6.4.	Función de Probabilidad Gamma. . . . .	251
6.5.	Función de Probabilidad Tipo Beta . . . . .	260
6.6.	Función de Probabilidad de Weibull . . . . .	268
6.6.1	Función de Probabilidad de Rayleigh. . . . .	276
6.6.2	Función de Probabilidad de Rice . . . . .	277
6.7	Función de Probabilidad de Pareto. . . . .	278
6.8	Función de Probabilidad de Cauchy–Lorentz. . . . .	279
6.9	Función de Probabilidad Logística . . . . .	285
	Ejercicios propuestos del capítulo 6. . . . .	286
7	La Función de Probabilidad Normal y las Funciones de Probabilidad Muestrales . . . . .	290

7.1	Distribución Lognormal. ....	304
7.2	Propiedad reproductiva o aditiva de la distribución normal .....	309
	Corolario de normalidad de la media muestral .....	312
7.3	Teoremas del Límite Central .....	315
	Teorema de De Moivre-Laplace .....	315
7.4	Distribución de probabilidad normal de la media aritmética (promedio) de una muestra .....	323
7.5	Distribución de probabilidad de la varianza de una muestra .....	327
	Distribución de probabilidad ji cuadrada .....	327
7.5.1	Teorema de aditividad de la distribución ji cuadrada .....	349
7.6	Distribución de probabilidad de la media de una muestra para una población con varianza desconocida .....	337
7.7	Distribución de probabilidad del cociente entre varianzas de dos poblaciones.....	346
	Ejercicios propuestos del capítulo 7.....	354
	Referencias bibliográficas .....	359
	Referencias mesográficas .....	360

## 1. Desarrollo Histórico de la Probabilidad y la Estadística

La probabilidad y la estadística nacieron como dos ramas del conocimiento diferentes. La probabilidad se derivó como una rama de las matemáticas y conserva un enfoque deductivo en el cual se trata de establecer modelos matemáticos generales de fenómenos sujetos al azar o a la incertidumbre. La Estadística mantiene un enfoque inductivo, ya que parte de la observación, experimentación, recolección de datos y procesamiento para tratar de conocer o entender un fenómeno hasta su modelación matemática. Actualmente ambas ramas del conocimiento son casi indivisibles y coinciden en el propósito de analizar y resolver problemas sujetos a incertidumbre.

La probabilidad y el azar siempre han estado en la mente del ser humano. Por ejemplo: sumerios y asirios utilizaban un hueso extraído del talón de animales como ovejas, ciervos o caballos, denominado astrágalo o talus, que tallaban para que pudieran caer en cuatro posiciones distintas, por lo que son considerados como los precursores de los dados. En el caso de la civilización egipcia, algunas pinturas encontradas en las tumbas de los faraones muestran tanto astrágalos como tableros para el registro de los resultados. Por su parte, los juegos con dados se practicaron ininterrumpidamente desde los tiempos del Imperio Romano hasta el Renacimiento. Uno de estos juegos, denominado "hazard", palabra que en inglés y francés significa riesgo o peligro, fue introducido en Europa con la Tercera Cruzada. Las raíces etimológicas del término provienen de la palabra árabe "al-azar", que significa "dado". Posteriormente, en el "Purgatorio" de Dante el término aparece ya como "azar".

El uso de los métodos estadísticos se remonta al menos al siglo V aC, el historiador Tucídides, en su Historia de la Guerra del Peloponeso, describe cómo los atenienses calculaban la altura de la muralla de Platea, contando el número de ladrillos de una sección expuesta de la muralla que estuviera lo suficientemente cerca como para contarlos. El conteo era repetido varias veces por diferentes soldados. El valor más frecuente (la moda en términos más modernos) era tomado como el valor del número de ladrillos más probable. Multiplicando

este valor por la altura de los ladrillos usados en la muralla les permitía a los atenienses determinar la altura de las escaleras necesarias para trepar las murallas. En el poema épico hindú Majabhárata (libro 3: la historia del rey Nala), del siglo III aC, el rey Ritupama estimaba el número de frutas y hojas (2095 frutas y 50,000,000 hojas (5 crores)) a partir de dos grandes hojas de un árbol Vibhitaka, contándolos en un solo vástago. Este número era luego multiplicado por el número de vástagos en las ramas. Este estimado fue posteriormente verificado y se halló que estaba muy cerca del número verdadero.



**FIGURA 1.1.** Fragmento del Mahabhárata

**NOTA:** Vishnu y Brahman. (sf). [https://upload.wikimedia.org/wikipedia/commons/0/05/Disguised\\_Vishnu\\_and\\_a\\_Brahman%2C\\_Folio\\_from\\_a\\_Razmnama%2C.jpg](https://upload.wikimedia.org/wikipedia/commons/0/05/Disguised_Vishnu_and_a_Brahman%2C_Folio_from_a_Razmnama%2C.jpg)

El primer escrito de estadística fue encontrado en un libro del siglo IX titulado Manuscrito sobre el descifrado de mensajes criptográficos, de Al-Kindi (801-873). En su libro, Al-Kindi da una descripción detallada sobre el uso de las estadísticas y análisis de frecuencias en el descifrado de mensajes, este fue el nacimiento tanto de la estadística como del criptoanálisis.

La palabra probabilidad viene del latín *probabilitas*, *posibilitatis*, formada del verbo *probare* (comprobar, probar), el sufijo *-bilis* (indica posibilidad), y

el sufijo -tat- que indica cualidad. Entonces, el término probabilidad indica la cualidad -dad) de poder (-able) probar.

La historia de la probabilidad formalmente comienza cuando Gerolamo Cardano (1501-1576) escribió su Libro *Liber de Ludo Aleae*, escrito en el año de 1560, pero publicado póstumamente hasta 1663, el cual constituye el primer tratado serio de probabilidad abordando métodos de cierta efectividad sobre los juegos de azar.

En el siglo XVII Pierre de Fermat (1607-1665) y Blaise Pascal (1629-1695) tratan de resolver algunos problemas relacionados con los juegos de azar, hasta entonces comienza a elaborarse una teoría aceptable sobre estos juegos.

**FIGURA 1.2.** Pierre de Fermat (1607-1665) y Blaise Pascal (1623-1662)



**NOTA:** Pierre de Fermat. (2020). [https://es.wikipedia.org/wiki/Pierre\\_de\\_Fermat](https://es.wikipedia.org/wiki/Pierre_de_Fermat)  
Blaise Pascal. (2020). [https://es.wikipedia.org/wiki/Blaise\\_Pascal](https://es.wikipedia.org/wiki/Blaise_Pascal)

Christian Huygens (1629-1695) conoció la correspondencia entre Blaise Pascal y Pierre Fermat suscitada por Antoine Gombaud, Caballero De Méré (1607-1684), se planteó el debate de determinar la probabilidad de ganar una partida, y publicó (en 1657) el primer libro sobre probabilidad *De Ratiociniis in Ludo Aleae*, un tratado sobre juegos de azar en donde introdujo algunos conceptos importantes en este campo, como la esperanza matemática, además de resolver algunos de los problemas propuestos por Pascal, Fermat y De Méré.

John Graunt (1620-1674) en su libro *Natural and Political Observations Made upon the Bills of Mortality*, estimó la población de Londres en 1662 a través de registros parroquiales. Él sabía que había cerca de 13,000 funerales al año en

Londres y que de cada once familias tres personas morían por año. Estimó de los registros parroquiales que el tamaño promedio de las familias era de ocho miembros y calculó que la población de Londres era de cerca de 384,000 habitantes.

Durante el siglo XVIII, debido muy particularmente a la popularidad de los juegos de azar, el cálculo de probabilidades tuvo un notable desarrollo. Destacan el teorema de Jacob Bernoulli (1654-1705) en su obra póstuma *Ars Conjectandi* de 1713, y en 1738 el primer caso particular estudiado por Abraham De Moivre (1667-1754) del teorema del límite central.

Es muy recomendable hacer un seguimiento de toda la familia Bernoulli por sus notables contribuciones en todos los campos de la ciencia ([https://es.wikipedia.org/wiki/Familia\\_Bernoulli](https://es.wikipedia.org/wiki/Familia_Bernoulli)).

**FIGURA 1.3.** Jacob Bernoulli (1654-1705) y Abraham De Moivre (1667-1754)



**NOTA:** Jakob Bernoulli. (2019). [https://es.wikipedia.org/wiki/Jakob\\_Bernoulli](https://es.wikipedia.org/wiki/Jakob_Bernoulli)  
Abraham de Moivre. (2020). [https://es.wikipedia.org/wiki/Abraham\\_de\\_Moivre](https://es.wikipedia.org/wiki/Abraham_de_Moivre)

En 1809 Carl Friedrich Gauss (1777-1855) inició el estudio de la teoría de errores y en 1810 Pierre-Simon Laplace (1749-1827), que ya había considerado anteriormente el tema, completó el desarrollo de esta teoría. En 1812 Laplace publicó *Théorie Analytique des Probabilités* en el que expone un análisis matemático sobre los juegos de azar, y en 1814, *Essai Philosophique sur la Probabilité*.

El término "bayesiano" se refiere a Thomas Bayes (1702-1761), quien probó un caso especial de lo que se conoce hoy como Teorema de Bayes. Sin embargo, fue Pierre-Simon Laplace (1749-1827), quien introdujo una visión general del teorema y lo aplicó a mecánica celeste, estadísticas médicas, confiabilidad y jurisprudencia.



**FIGURA 1.4.** Carl Friedrich Gauss (1777-1855) y Pierre-Simon Laplace (1749-1827)

**NOTA:** Carl Friedrich Gauss. (2020). [https://es.wikipedia.org/wiki/Carl\\_Friedrich\\_Gauss](https://es.wikipedia.org/wiki/Carl_Friedrich_Gauss)  
 Pierre-Simon Laplace. (2019). [https://es.wikipedia.org/wiki/Pierre-Simon\\_Laplace](https://es.wikipedia.org/wiki/Pierre-Simon_Laplace)

Pierre-Simon Laplace afirmó: "Es notable que una ciencia que comenzó con consideraciones sobre juegos de azar haya llegado a ser el objeto más importante del conocimiento humano". Comprender y estudiar el azar es indispensable, porque la probabilidad es un soporte necesario para tomar decisiones en cualquier ámbito.

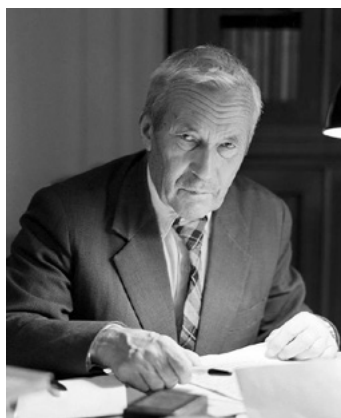
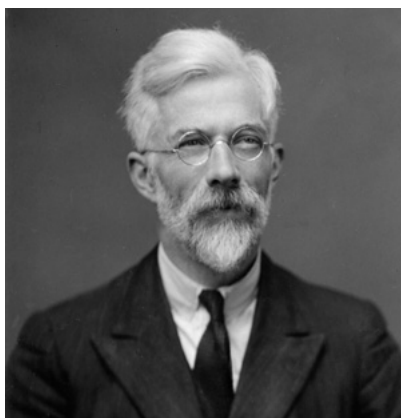
El término "estadística" deriva de la palabra del neolatín *statisticum collegium* (consejo de estado) y la palabra italiana *statista* ('hombre de estado' o político). La palabra alemana *statistik*, introducida primeramente en 1749 por Gottfried Achenwall (1719-1772), originalmente designaba el análisis de datos acerca del estado, significando la 'ciencia del estado' (llamado posteriormente «aritmética política» en idioma inglés). El término fue introducido en Inglaterra en 1792 por sir John Sinclair (1754-1835) cuando publicó el primero de los 21 volúmenes titulados *Statistical Account of Scotland*.

Joseph-Louis Lagrange (1736-1813) propuso una distribución parabólica de errores en 1776 y sugirió en 1781 otras dos distribuciones para errores: una distribución coseno y una distribución logarítmica. El método de mínimos cuadrados se debe a Adrien-Marie Legendre (1752-1833), que lo introdujo en 1805 en su *Nouvelles méthodes pour la détermination des orbites des comètes*. El término "error probable" (*der wahrscheinliche Fehler*) —la desviación media— fue introducido en 1815 por el astrónomo alemán Friedrich Wilhelm Bessel (1784-1846).

Otros personajes que contribuyeron fueron Augustus De Morgan (1806-1871), conocido por formular las leyes que llevan su nombre y establecer un concepto riguroso del procedimiento, inducción matemática; así como George Boole (1815-1864), quien en 1854 publicó *An Investigation of the Laws of Thought on Which are Founded the Mathematical Theories of Logic and Probabilities*.

Karl Pearson (1857-1936) estableció la disciplina de la estadística matemática. Desarrolló una intensa investigación sobre la aplicación de los métodos estadísticos en la biología y fue el fundador de la bioestadística.

**FIGURA 1.5.** Sir Ronald Aylmer Fisher (1890-1962) y Andréi Nikoláyevich Kolmogórov (1903-1987)



**NOTA:** Ronald Fisher. (2019). <https://psicologiaymente.com/biografias/ronald-fisher>  
 Andréi Kolmogórov. (2019). <https://www.timetoast.com/timelines/estadistica-2cce55a4-2558-4968-926a-5333d097a0e0>

Sir Ronald Aylmer Fisher (1890-1962) desarrolló el análisis de la varianza y está considerado como el padre de la estadística moderna y del diseño experimental.

Jerzy Neyman (1894-1981) desarrolló los intervalos de confianza utilizados en la estadística inferencial. La idea de que la “probabilidad” debería ser interpretada como el “grado de creencia subjetivo en una proposición” fue propuesto por John Maynard Keynes (1883-1946) a comienzos de la década de 1920. El método subjetivo bayesiano fue sucesivamente desarrollado y popularizado en los años cincuenta por Leonard Jimmie Savage (1917-1971). La inferencia objetiva bayesiana fue desarrollada con posterioridad por Harold Jeffreys (1891-1989), cuyo libro *Theory of probability* apareció en 1939. Edwin Thompson Jaynes (1922-1998) se centró en la física estadística y en los fundamentos de la probabilidad y la inferencia estadística.

## Ejercicios propuestos del capítulo 1

1. Elaborar fichas biográficas sobre algunos de los principales contribuyentes de la Probabilidad y de la Estadística: Thomas Bayes, George E. P. Box, Pafnuti Chebyshev, David R. Cox, Gertrude Cox, Harald Cramér, Francis Ysidro Edgeworth, Bradley Efron, Bruno de Finetti, Ronald A. Fisher, Francis Galton, Carl Friedrich Gauss, William Sealey Gosset (Student), Andréi Kolmogórov, Pierre-Simon Laplace, Erich L. Lehmann, Aleksandr Liapunov, Abraham de Moivre, Jerzy Neyman, Blaise Pascal, Karl Pearson, Charles S. Peirce, Adolphe Quetelet, C. R. Rao, Walter A. Shewhart, Charles Spearman, Thorvald N. Thiele, John W. Tukey, Abraham Wald, Al-Kindi.
2. Investigar quiénes fueron los principales contribuyentes de la Ley de los Grandes Números.
3. Investigar quiénes fueron los principales contribuyentes del Teorema del Límite Central.
4. Indagar sobre algunas de las principales metodologías de investigación científica que se basan en el enfoque de la probabilidad y la estadística.

## 2. Teoría Matemática de la Probabilidad

### 2.1. Conceptos básicos

Un fenómeno es una manifestación de una actividad que se produce en la naturaleza y que se percibe a través de los sentidos. En filosofía, el fenómeno (del griego φαινόμενον: 'aparición, manifestación', en plural: φαινόμενα) es el aspecto que las cosas ofrecen ante nuestros sentidos; es decir, el primer contacto que se tiene con las cosas, es lo que se denomina experiencia. El estudio de los fenómenos se lleva a cabo a través del método científico, mediante la observación, experimentación, teorización y comprobación. Para llevar a efecto lo anterior, es necesario conocer y aplicar las bases de la probabilidad y la estadística.

Un fenómeno o experimento es determinista si se puede predecir con certeza lo que va a ocurrir antes de llevarlo a efecto, por ejemplo, al soltar un objeto en el aire en un punto de la superficie terrestre, se puede asegurar que este se va a proyectar hacia abajo, por efecto de la gravedad, a menos que haya una fuerza que contrarreste a la gravedad. Si al experimentar no se sabe con certeza lo que va a ocurrir previamente, se trata de un experimento aleatorio. Una vaca que se encuentra amamantando a una cría se sabe que da leche, pero no se puede asegurar con certeza, cuántos litros de leche va a producir diariamente, esta es una variable sujeta a incertidumbre y que depende de muchas características como el alimento, el cuidado que se le provea, salud, etcétera.

### Escalas de medición

Las características que se van a analizar de un fenómeno pueden ser medidas a través de escalas muy diversas:

1. **Escala Clasificatoria.** Cuando la característica a analizar solo es clasificable; por ejemplo, en un lote un producto puede ser defectuoso o no defectuoso. La apariencia de un objeto solo se mide si gusta o no gusta. El código que se utiliza en las placas de un carro es de tipo clasificatorio, solo existe para identificarlo, pero no mide nada del vehículo. De la misma forma, el número de cuenta de un alumno, solo lo identifica. El tipo de sangre de cada persona utiliza una escala clasificatoria. En la prueba del embarazo, una mujer está embarazada o no lo está, no hay escalas intermedias.
2. **Escala Ordinal.** En esta escala, la característica que se observa puede servir para ordenar objetos o entes; sin un instrumento de medición, se puede ordenar a un grupo de personas por su estatura, simplemente comparándose unos con otros. Igualmente, un grupo de personas puede ordenarse de acuerdo con su color de piel. Con una escala ordinal se puede clasificar, por lo que la escala ordinal contiene a la clasificatoria.
3. **Escala Comparativa.** Este tipo de escalas ya permite clasificar, ordenar y comparar a un objeto o ente contra otros y medir diferencias entre ellos. Al ordenar a un grupo de personas por estatura, se puede establecer que una persona es más alta que otra, pero no se puede saber qué tanto es más alta. Con la escala comparativa ya es posible hacer esto. La temperatura de una persona es un ejemplo de escala comparativa, se puede saber si un objeto o ente está más caliente que otro y saber cuántos grados más caliente está, pero no se tiene un cero u origen absoluto de referencia.
4. **Escala Absoluta.** Esta escala contiene a todas las demás, puede clasificar, ordenar, comparar y determinar qué tan cerca o qué tan lejos está de un origen o cero absoluto, como los ejes coordenados en un sistema cartesiano. La diferencia con la escala comparativa es que la escala absoluta sí tiene un origen o cero de referencia fijo, lo que no sucede en el caso de la comparativa. La mayoría de las magnitudes físicas que se analizan en ingeniería se miden con escala absoluta: la distancia, tiempo, presión, resistencia mecánica, resistencia eléctrica, peso, etcétera, son ejemplos.

Un evento es todo posible suceso de un fenómeno (un hecho observable en un momento dado) o de un acontecimiento que ocurre en una posición y momento determinados (por lo tanto, puede especificarse como un punto en el espacio-tiempo). Se trata de los posibles resultados que pueden obtenerse de un experimento. Un evento puede ser determinista, cuando se sabe con certeza qué resultado se va a obtener, o puede ser aleatorio, cuando no se puede afirmar con certeza el resultado. La hora de llegada a un punto específico es un evento

aleatorio; así como la resistencia mecánica de un cable o la temperatura de un cuerpo a cierta hora en un punto abierto, pero si estuviera en un refrigerador a temperatura controlada sería determinística la medición.

El espacio muestral o espacio de muestreo (denotado  $E$ ,  $S$ ,  $\Omega$  o  $U$ ) consiste en el conjunto completo de todos los posibles resultados de un experimento aleatorio, junto con una estructura sobre el mismo. Los espacios muestrales pueden ser discretos o continuos.

- a) Los espacios muestrales continuos por definición siempre son infinitos no numerables. Por ejemplo, la hora de llegada a una reunión, entre las 20:00 y las 21:00 horas; o la resistencia rockwell de una lámina de acero, entre 70 y 76.

Los espacios muestrales discretos siempre son numerables y pueden ser:

- b) Espacios muestrales discretos finitos, como por ejemplo, al lanzar una moneda al aire y ver qué lado de la moneda cae hacia arriba, solo puede haber dos posibles resultados {águila o sol}; al escoger un artículo en un lote de  $N$  de ellos, para determinar si es defectuoso o no defectuoso, solo puede haber dos posibles resultados, defectuoso o no defectuoso; al lanzar un dado homogéneo con seis caras, la cara que cae hacia arriba solo puede ser: A, R, Q, J, 9 rojo y 7 negro, por lo que solo puede haber esos posibles resultados y su espacio muestral sería:  $S = \{A, R, Q, J, 9 \text{ rojo}, 7 \text{ negro}\}$ , cada cara tiene cierto grado de importancia, peso o jerarquía y en este caso están ordenadas de mayor a menor. Lanzar dos dados numéricos podría dar como resultado el siguiente espacio de eventos:  $S = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\}$ ; en el anterior ejercicio, si lo que se busca como resultado es la suma de las dos caras superiores, el espacio de eventos sería  $S = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$ .
- c) Espacios muestrales discretos infinitos numerables. El número de piezas que tiene un artículo complejo, como un coche, puede afirmarse que es finito, pero es tan grande que se considera infinito; en este caso, contar el número de defectos que presenta un coche puede ser  $S = \{0, 1, 2, 3, \dots\}$ .

Un universo es una colección completa de entes u objetos que cumplen ciertas propiedades de tal manera que se identifican como elementos del mismo. Una población es la colección completa de valores o lecturas de una cierta



característica de un universo. Una muestra es un subconjunto de una población seleccionada de acuerdo con una regla o plan. Para ilustrar lo anterior, se abordarán diversos ejemplos.

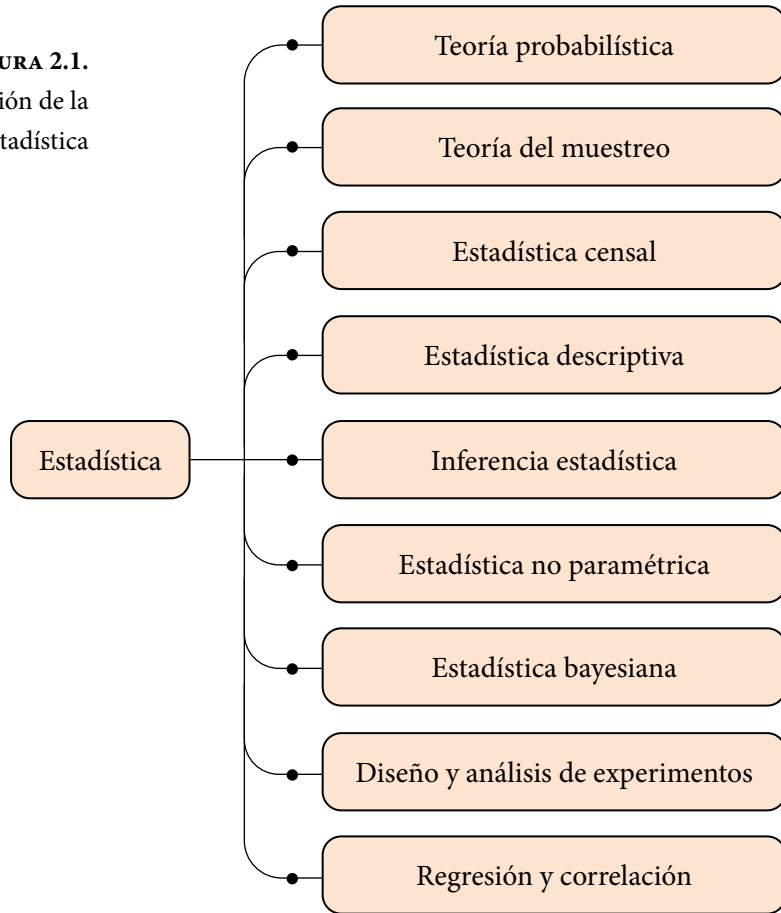
Suponga que el universo es el conjunto de estudiantes de la UNAM. Una población sería el conjunto completo de estaturas de cada uno de ellos; otra población sería el conjunto de pesos; otra, el conjunto de tipos de sangre; otra el promedio de avance que llevan en sus estudios. Una muestra sería del conjunto de promedios de avance de los estudiantes de la Facultad de Ingeniería de la UNAM; otra muestra podría ser del conjunto de estudiantes que mide entre 1.50 y 1.80 de estatura. Por definición se puede establecer un universo dentro de otro universo, una población dentro de otra población o una muestra dentro de otra muestra.

La Estadística es la rama de la ciencia que estudia las reglas para recolectar, capturar, organizar, presentar, procesar y analizar los datos obtenidos al realizar varios ensayos de un experimento y para inferir conclusiones acerca de este último. Proporciona, además, los métodos para el diseño estadístico de experimentos y para tomar decisiones cuando aparecen situaciones de incertidumbre. Algunos autores establecen que la estadística no es ciencia, ya que algunas de las reglas que emplea son empíricas, como el hecho de realizar una tabla de frecuencias como más adelante se verá.

Al realizar un análisis estadístico, el estudio puede llevarse a efecto considerando una sola variable, por ejemplo, las ganancias, el rendimiento, los costos, índice de masa corporal, etcétera. A este tipo de estadística se le conoce como de una variable o univariada. Si el análisis comprende dos o más variables, como son peso y estatura; temperatura y presión; temperatura, presión y tiempo; etcétera, se le denomina estadística de varias variables o multivariable.

Las partes en las que se descompone la estadística se muestran en la figura 2.1 y se definen a continuación.

**FIGURA 2.1.**  
Clasificación de la  
Estadística



**Teoría probabilística:** rama de la matemática que proporciona los fundamentos, modelos matemáticos y el lenguaje que se usa en la estadística.

**Teoría del muestreo:** Es la rama de la estadística que se encarga de definir las reglas para tomar muestras de una población específica, el tamaño de dichas muestras, el método a seguir para tomarlas y los parámetros que indicarán la representatividad de estas.

La estadística censal se refiere al estudio estadístico de las poblaciones en su conjunto, su dimensión, estructura, parámetros y su dinámica. La Demografía (del griego δῆμος *dēmos* 'pueblo' y γραφία *grafía* 'trazo, descripción' —estudio de la población—) es una rama del conocimiento que estudia las poblaciones humanas, su dimensión, estructura, evolución y características generales. La demografía estudia estadísticamente la estructura y la dinámica de las poblaciones,

así como los procesos concretos que determinan su formación, conservación y desaparición. Tales procesos son los de fecundidad, mortalidad y migración (emigración e inmigración).

La Estadística descriptiva es la rama de la estadística que se encarga de analizar las reglas para recolectar, presentar y procesar los datos obtenidos al hacer una medición u observación de una característica particular de un objeto, con la finalidad de conocer su comportamiento. Si se conocen con certeza los valores que tomará la característica particular en cuestión, previamente al experimento, a dicha característica se le denomina determinística. En este caso se puede conocer su comportamiento sin necesidad de hacer el experimento; si es el caso, el experimento se realizará con la finalidad de comprobar los resultados esperados. Si los valores que tomará la característica no pueden predecirse con certeza, antes del experimento, a dicha variable se le denomina aleatoria. Por otra parte, dentro del estudio de características aleatorias, se puede ver que existen dos tipos:

- a) Aquella cuyos resultados pueden ser medidos en forma discreta; por ejemplo: el número de llegadas a una cola, el número de defectos en un lote, el número de ases que se obtienen en un juego de póker, etcétera, a estas se les denomina variables aleatorias discretas.
- b) Variables aleatorias continuas, son aquellas que tienen unidades de medida continua; por ejemplo: la cantidad de leche que produce una vaca diariamente, el tiempo de vida de un producto, el tiempo de espera en una cola, etcétera.

La inferencia estadística es la rama de la estadística que proporciona las reglas para estimar ciertos valores de una población, con base en los resultados de una muestra, formular hipótesis sobre la verdad de estas estimaciones y tomar decisiones con base en estos resultados. La estimación de parámetros poblacionales a partir de muestras se lleva a efecto a través de estimadores puntuales, de los cuales hay que determinar qué cualidades deben reunir para ser válidos y representativos, o a través de estimadores por intervalos de confianza, de los cuales hay que determinar la distribución de probabilidad que presentan y el nivel de confianza que se desea tener. Las hipótesis estadísticas que se formulan deben ser probadas para comprobar su validez y representatividad. Tanto la estimación por intervalos de confianza como las pruebas de hipótesis se pueden clasificar para una población o para dos o más poblaciones. Para una población, generalmente los parámetros que se estiman son la media, varianza,

desviación estándar y fracción de éxitos. Para dos poblaciones generalmente lo que se estima es el cociente entre varianzas, la diferencia de medias o la diferencia de proporciones. También, dentro de la estadística inferencial, se ven otro tipo de pruebas estadísticas como las de bondad de ajuste, que permiten probar la adecuación de un conjunto de datos obtenidos empíricamente a un modelo probabilístico específico.

La estadística paramétrica es la rama de la estadística inferencial que comprende los procedimientos estadísticos y de decisión que están basados en las distribuciones de los datos reales, los cuales generalmente se suponen normales. Estas son determinadas usando un número finito de parámetros. Para aplicar la estadística paramétrica se requiere conocer la distribución de probabilidad que siguen los datos o en su defecto, con base en el Teorema del Límite Central, tomar muestras “grandes”. Cuando se desconoce la distribución de probabilidad que siguen los datos y no se pueden tomar muestras “grandes” por restricciones económicas o de otro tipo, entonces se debe aplicar la estadística no paramétrica. Por ejemplo, los datos categorizados en: niños, jóvenes, adultos y ancianos no pueden ser interpretados mediante la estadística paramétrica ya que no se puede hallar un parámetro numérico (como por ejemplo la media de edad) cuando los datos no son numéricos.

La estadística bayesiana es una rama de la estadística en la que la evidencia sobre el verdadero estado del mundo se expresa en términos de estimaciones extraídas de datos históricos duros, así como de estimaciones basadas en la opinión de expertos, con ciertos grados de creencia o, más específicamente, las probabilidades bayesianas.

El diseño y análisis de experimentos es una rama de la estadística que permite plantear si una o más variables conocidas como efectos, dependen a su vez de otras variables llamadas posibles causas o factores. Esta rama de la estadística permite diseñar, realizar y cuantificar, a través de experimentos el grado de dependencia que existe entre las causas y los efectos. En un diseño experimental se manejan deliberadamente una o más variables vinculadas a las causas, para medir la respuesta que tienen sobre otras variables de interés llamadas efectos. El diseño experimental establece la serie de pautas a seguir para determinar qué variables hay que considerar, de qué manera, cuántas veces hay que repetir el experimento y en qué orden, para poder establecer con un grado de confianza predefinido la necesidad de una presunta relación de causa-efecto. El diseño experimental encuentra aplicaciones muy diversas en la industria, agricultura, mercadotecnia, medicina, ecología, ciencias de la conducta, etcétera, constituyendo una fase esencial en el desarrollo de un estudio experimental.

Una regresión es un ajuste de un conjunto de puntos obtenidos empíricamente a un modelo matemático en particular. Si el ajuste es entre dos variables  $\{(x,y) | f: x \in \mathbb{R} \rightarrow y \in \mathbb{R}, y = f(x)\}$ , la regresión puede ser lineal, polinomial, exponencial, logarítmica, trigonométrica, etcétera. Si el ajuste es a un modelo matemático de varias variables  $\{(x_1, x_2, \dots, x_n, y) | f: (x_1, x_2, \dots, x_n) \in \mathbb{R}^n \rightarrow y \in \mathbb{R}, y = f(x_1, x_2, \dots, x_n)\}$ , se denomina regresión múltiple.

Una correlación corresponde con la definición de indicadores que permitan determinar qué tan bueno es el ajuste entre el conjunto de puntos dados y el modelo matemático usado.

### 2.1.1. Relación entre la probabilidad y la estadística, el método científico y el papel de la probabilidad y la estadística en la investigación básica y aplicada y sus etapas

Como ya se mencionó antes, la relación entre probabilidad y estadística es muy estrecha y viene dada por el hecho de que ambas áreas del conocimiento tienen como objeto de estudio los fenómenos aleatorios. La probabilidad y la estadística nacieron como dos ramas del conocimiento diferentes. La probabilidad se derivó como una rama de las matemáticas y conserva un enfoque deductivo en el cual se trata de establecer modelos matemáticos generales de fenómenos sujetos al azar o a la incertidumbre. La estadística mantiene un enfoque inductivo, ya que parte de la observación, experimentación, recolección de datos y procesamiento para tratar de conocer o entender un fenómeno hasta su modelación matemática. Actualmente ambas ramas del conocimiento son casi indivisibles y coinciden en el propósito de analizar y resolver problemas sujetos al azar o a la incertidumbre.

El enfoque de la probabilidad y la estadística representa el nexo común que existe en la mayoría de las investigaciones científicas en las que interviene el tratamiento de datos y la interpretación y predicción de los mismos. El análisis estadístico cobra una especial importancia cuando se realizan investigaciones en campos críticos, donde una mala interpretación de los datos puede provocar consecuencias nefastas para la sociedad.

La probabilidad proporciona los modelos matemáticos y el lenguaje a la estadística, y esta a su vez aporta las herramientas que van, desde el análisis e

interpretación de datos hasta el proceso de predicción y toma de decisiones. Es tal la importancia que ha adquirido la estadística que aparece en prácticamente todas las áreas del conocimiento, siendo un factor clave en las predicciones y toma de decisiones a partir de datos observados.

La tarea más importante de la estadística es la de proporcionar alternativas cuantitativas al mero juicio personal, de forma que se realicen conclusiones objetivas. Las pruebas de hipótesis son consideradas como una de las herramientas más importantes en este tipo de toma de decisiones subjetivas, pues, mediante ellas, se pueden tomar decisiones que, en principio, no dependen de la persona que las realiza. En el enfoque tradicional de la investigación una fase sustancial es la experimentación y no se concibe esta sin el enfoque de la probabilidad y la estadística.

Tal como se describió en el volumen anterior, la investigación, según Ezequiel Ander-Egg, es un procedimiento reflexivo, sistemático, controlado y crítico, que permite descubrir nuevos hechos o datos, relaciones o leyes, en cualquier campo del conocimiento humano.

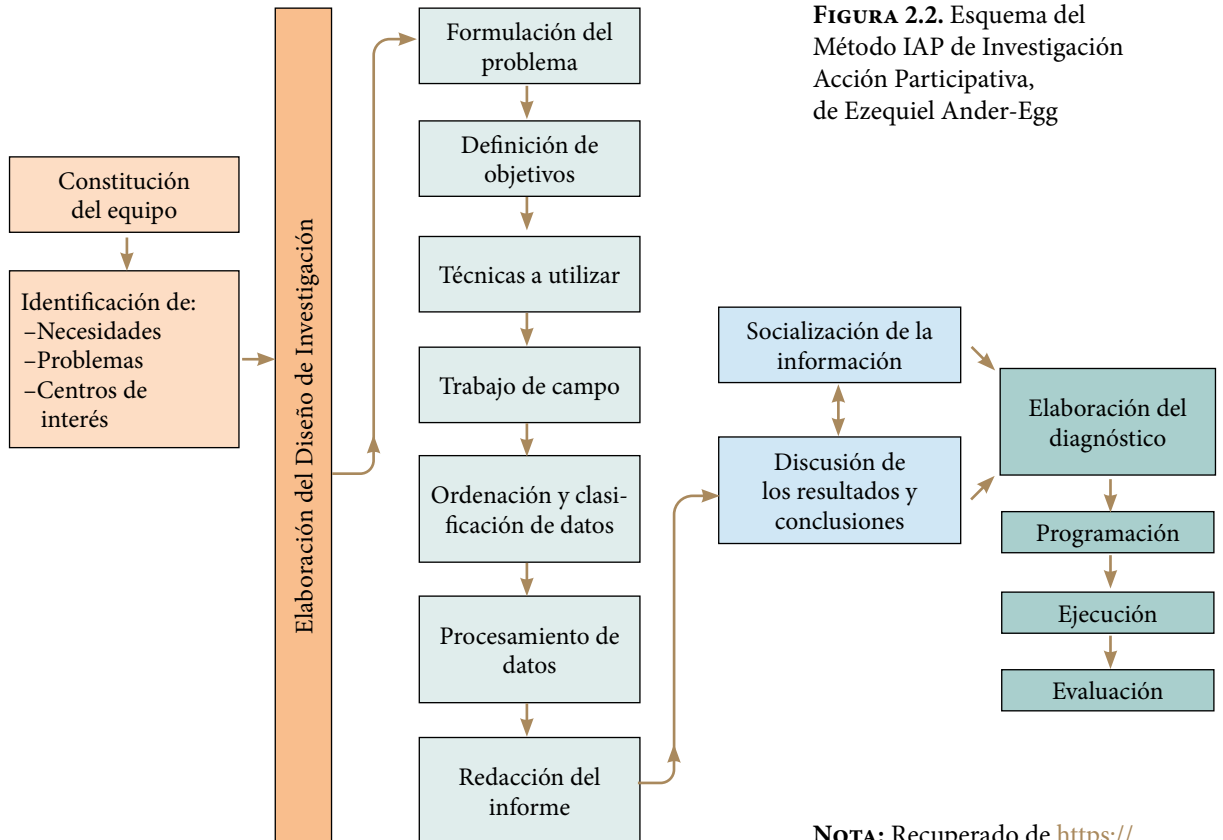
La investigación social es el proceso que, utilizando la metodología científica, permite obtener nuevos conocimientos en el campo de la realidad social. Ezequiel Ander-Egg, establece que el método IAP consta de las siguientes fases:

- A) Constitución del equipo.
- B) Diseño de la investigación.
- C) Proceso de investigación.
  - 1. Identificación de necesidades básicas, problemas y centros de interés.
  - 2. Formulación del problema.
  - 3. Técnicas para recoger datos y obtener información.
  - 4. Trabajo de campo, recolección de datos.
  - 5. Ordenación y clasificación de la información.
  - 6. Análisis e interpretación de datos.
  - 7. Redacción del informe preliminar.
  - 8. Socialización de la información.
  - 9. Elaboración del diagnóstico.
  - 10. Elaboración del programa o proyecto.
  - 11. Formación de equipos y/o grupos de trabajo.
  - 12. Desarrollo de actividades.
  - 13. Control operativo.



Esta secuencia de pasos se muestra en la figura 2.2.

Como se puede apreciar, la probabilidad y la estadística intervienen directamente en las etapas de la tres a la siete del proceso de investigación.



**FIGURA 2.2.** Esquema del Método IAP de Investigación Acción Participativa, de Ezequiel Ander-Egg

**NOTA:** Recuperado de [https://www.sis.net/documentos/Digitalizados/22939\\_Repensando%20la%20investigacion.pdf](https://www.sis.net/documentos/Digitalizados/22939_Repensando%20la%20investigacion.pdf)

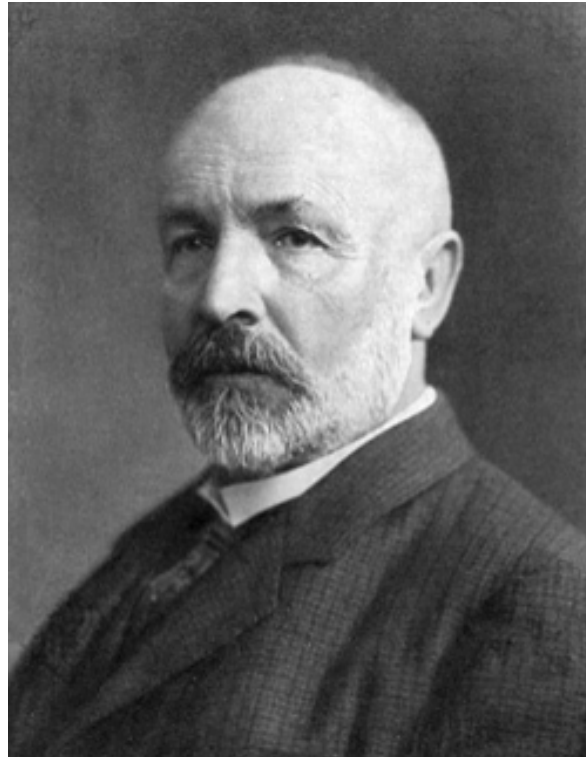
## 2.2. Teoría de Conjuntos

En matemáticas, un conjunto es una colección de elementos considerada en sí misma como un objeto, cuya correcta y precisa definición permite dilucidar si un elemento pertenece o no al conjunto. Los elementos de un conjunto pueden ser muy disímbolos, por ejemplo: nombres, países, personas, números, colores, letras, imágenes, etcétera. La teoría de conjuntos es una rama de la lógica matemática que estudia las propiedades y relaciones entre conjuntos.

El desarrollo histórico de la teoría de conjuntos se atribuye a Georg Ferdinand Ludwing Philipp Cantor (1845-1918), que comenzó a investigar este tema en la segunda mitad del siglo XIX, precedido por algunas ideas de Bernhard Placidus Johann Gonzal Nepomuk Bolzano (1781-1848) e influido por Julius Wilhelm Richard Dedekind (1831-1916).

**FIGURA 2.3.** Georg Ferdinand Ludwing Philipp Cantor (1845-1918)

**NOTA:** Georg Cantor. (2020). [https://es.wikipedia.org/wiki/Georg\\_Cantor](https://es.wikipedia.org/wiki/Georg_Cantor)



La pertenencia de un elemento  $a$  a un conjunto  $A$  se indica como  $a \in A$ . La no pertenencia se indica como  $a \notin A$ . Dos conjuntos  $A$  y  $B$  son iguales si y solo si contienen los mismos elementos.

Se dice que un conjunto  $A$  es subconjunto de  $B$  si todos los elementos de  $A$  pertenecen al conjunto  $B$ , lo cual se representa matemáticamente como  $A \subset B$ .

Si  $A$  es subconjunto de  $B$  y  $B$  es subconjunto de  $A$ , entonces  $A = B$ . Si  $A$  es un subconjunto de  $B$  o  $A$  es igual a  $B$ , se representa como  $A \subseteq B$ .

Los conjuntos numéricos usuales en matemáticas son: el conjunto de los números naturales  $N = \{1, 2, 3, 4, 5, \dots\}$ , el de los números enteros  $I = \{\dots, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, \dots\}$ , el de los números racionales como cociente de dos enteros  $Q = \{p/q \mid q \neq 0; p, q \in I\}$ , el de los números reales  $R$  y el de los números complejos  $C = \{a+ib \mid a, b \in R; i = \sqrt{-1}\}$ . Cada uno es un subconjunto del siguiente:  $N \subset I \subset R \subset C$ .

Se denotará al número de elementos en un conjunto arbitrario  $A$  con  $n(A)$ .

Se define el conjunto vacío  $\phi = \{ \}$  como aquel que NO contiene elementos. Un conjunto universal es un conjunto formado por todos los objetos de estudio en un contexto dado.

El complemento de un conjunto  $A$  es el conjunto de elementos que pertenecen al conjunto universal, pero que no pertenecen a  $A$ . Al complemento de  $A$  por costumbre se le representa como  $\bar{A}$ ,  $A^c$ ,  $A^-$ . Cabe señalar que el complemento del conjunto universal es el conjunto vacío y el complemento del conjunto vacío es el conjunto universal.

Existen tres operaciones básicas entre dos conjuntos:

1. La unión de dos conjuntos  $A$  y  $B$ , lo cual se representa como  $A \cup B$ , es el conjunto completo de elementos contenidos en  $A$ , en  $B$  o en ambos.
2. La intersección de  $A$  y  $B$  que se representa como  $A \cap B$  es el conjunto de elementos contenidos simultáneamente en  $A$  y en  $B$ .
3. La diferencia entre dos conjuntos  $A$  y  $B$  es el conjunto formado por los elementos que pertenecen a  $A$ , pero no pertenecen a  $A \cap B$ , lo cual se representa por:  $A - B = A - A \cap B$ .

### Ejercicio 2.1

Suponga que el conjunto universal  $U$  es el conjunto de puntos naturales del plano  $XY$ , para los cuales  $x \leq 4$  y  $y \leq 5$ . Sea  $A$  el conjunto de puntos naturales, contenido en  $U$ , para los cuales  $5x + 4y \leq 20$ ; sea  $B$  el conjunto de puntos naturales, contenido en  $U$ , para los cuales  $y < x$ .

- a) Obtenga la expresión matemática del conjunto universal, del conjunto  $A$ , del conjunto  $B$ , y enumere sus elementos.

En la figura 2.4 se muestran los puntos del conjunto universal en color gris, los puntos de A y de B.

$$U = \{(x, y) \mid x \leq 4; y \leq 5; x, y \in \mathbb{N}\}$$

$$A = \{(x, y) \mid 5x + 4y \leq 20; x \leq 4; y \leq 5; x, y \in \mathbb{N}\}$$

$$B = \{(x, y) \mid y < x; x \leq 4; y \leq 5; x, y \in \mathbb{N}\}$$

$$U = \{(1,1), (2,1), (3,1), (4,1), (1,2), (2,2), (3,2), (4,2), (1,3), (2,3), (3,3), (4,3), (1,4), (2,4), (3,4), (4,4), (1,5), (2,5), (3,5), (4,5)\}$$

Nótese que  $n(U) = 20$ .

$$A = \{(1,1), (2,1), (3,1), (1,2), (2,2)\} \quad n(A) = 5$$

$$B = \{(2,1), (3,1), (4,1), (3,2), (4,2), (4,3)\} \quad n(B) = 6$$

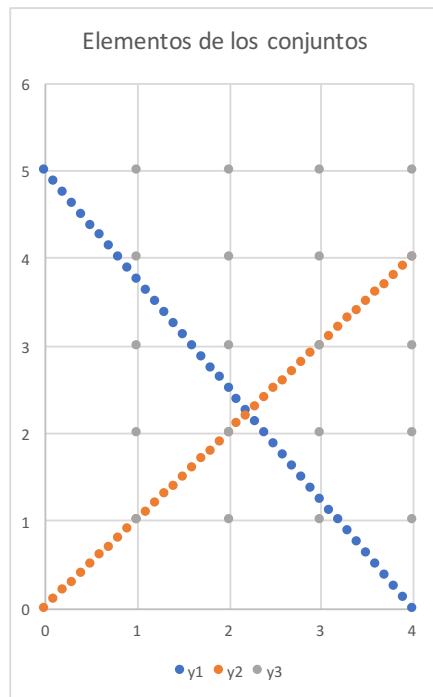


FIGURA 2.4.

b) Obtenga el complemento de A y el complemento de B.

$$A^c = \{(4,1), (3,2), (4,2), (1,3), (2,3), (3,3), (4,3), (1,4), (2,4), (3,4), (4,4), (1,5), (2,5), (3,5), (4,5)\}$$

$$B^c = \{(1,1), (1,2), (2,2), (1,3), (2,3), (3,3), (1,4), (2,4), (3,4), (4,4), (1,5), (2,5), (3,5), (4,5)\}$$

- c) Determine la unión de  $A$  y  $B$ , su intersección, la diferencia  $A-B$  y la diferencia  $B-A$ .

$$A \cup B = \{(1,1), (2,1), (3,1), (1,2), (2,2), (4,1), (3,2), (4,2), (4,3)\}$$

$$A \cap B = \{(2,1), (3,1)\}$$

$$A - B = \{(1,1), (1,2), (2,2)\}$$

$$B - A = \{(4,1), (3,2), (4,2), (4,3)\}$$

- d) Compruebe que  $(A \cup B)^c = A^c \cap B^c$ .

$$(A \cup B)^c = \{(1,3), (2,3), (3,3), (1,4), (2,4), (3,4), (4,4), (1,5), (2,5), (3,5), (4,5)\}$$

$$A^c \cap B^c = \{(1,3), (2,3), (3,3), (1,4), (2,4), (3,4), (4,4), (1,5), (2,5), (3,5), (4,5)\}$$

- e) Compruebe que  $(A \cap B)^c = A^c \cup B^c$ .

$$(A \cup B)^c = \{(1,1), (4,1), (1,2), (2,2), (3,2), (4,2), (1,3), (2,3), (3,3), (4,3), (1,4), (2,4), (3,4), (4,4), (1,5), (2,5), (3,5), (4,5)\}$$

$$A^c \cup B^c = \{(1,1), (4,1), (1,2), (2,2), (3,2), (4,2), (1,3), (2,3), (3,3), (4,3), (1,4), (2,4), (3,4), (4,4), (1,5), (2,5), (3,5), (4,5)\}$$

A los conjuntos se les puede representar gráficamente a través de los diagramas de Venn, los cuales fueron ideados alrededor de 1880 por John Venn (1834–1923). Al conjunto universal  $U$  se le representa como un rectángulo y a los subconjuntos  $A$  y  $B$  con elipses o círculos, como se ilustra en la figura 2.6.

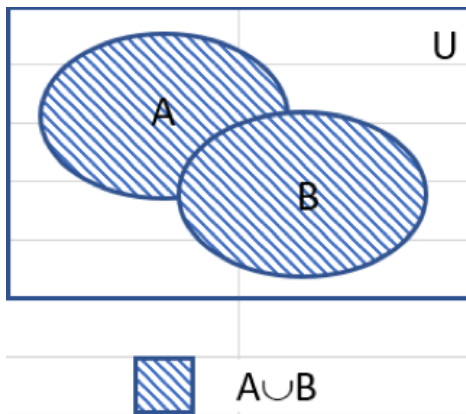
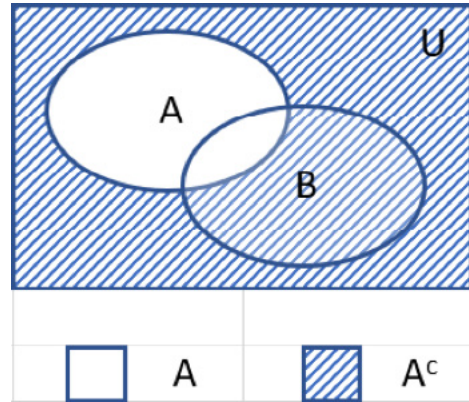
**FIGURA 2.5.** John Venn  
(1834 -1923)

**NOTA:** John Venn. (2020). [https://es.wikipedia.org/wiki/John\\_Venn](https://es.wikipedia.org/wiki/John_Venn)

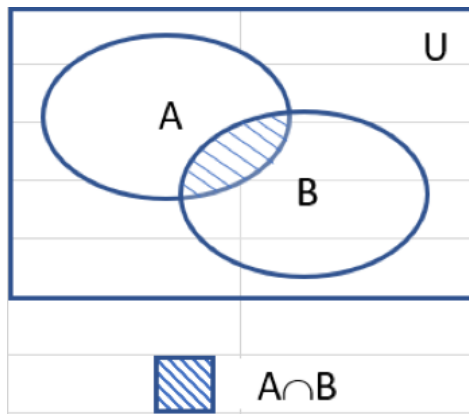


FIGURA 2.6. Diagramas de Venn

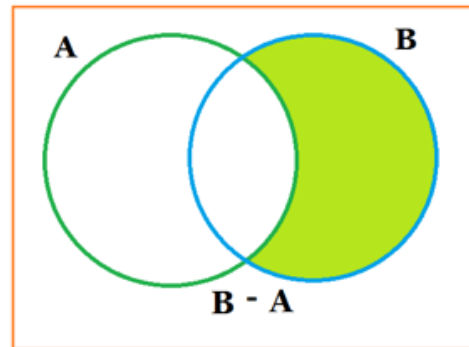
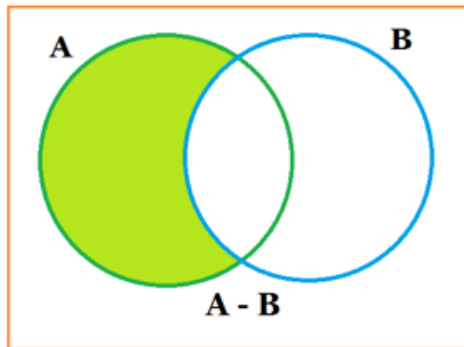
2.6.a.  
Complemento de A



2.6.b. Unión de A y B



2.6.c. Intersección de A y B



2.5.d. Diferencia entre A y B

Dos conjuntos  $A$  y  $B$  son mutuamente excluyentes si se cumple que  $A \cap B = \phi$ .  
 Dos conjuntos  $A$  y  $B$  son colectivamente exhaustivos si se cumple que  $A \cup B = U$ .  
 Nótese que  $A$  y  $A^c$  son mutuamente excluyentes y colectivamente exhaustivos.

Una colección de conjuntos  $A_1, A_2, A_3, \dots, A_k$  son mutuamente excluyentes y colectivamente exhaustivos si se cumple que:

Para todo  $i \neq j \rightarrow A_i \cap A_j = \phi$  Conjuntos mutuamente excluyentes.

$A_1 \cup A_2 \cup A_3 \cup \dots \cup A_k = U$  Conjuntos colectivamente exhaustivos.

Propiedades de las operaciones entre conjuntos.

1.  $U^c = \phi$ .
2.  $\phi^c = U$ .
3.  $(A^c)^c = A$ .
4.  $A \cup A = A$ .
5.  $A \cap A = A$ .
6.  $A \cup (B \cap C) = (A \cup B) \cap C$ . Ley asociativa de la unión de conjuntos.
7.  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ . Ley asociativa de la intersección de conjuntos.
8.  $A \cup B = B \cup A$ . Ley conmutativa de la unión de conjuntos.
9.  $A \cap B = B \cap A$ . Ley conmutativa de la intersección de conjuntos.
10.  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ . Ley distributiva de la intersección sobre la unión.
11.  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$ . Ley distributiva de la unión sobre la intersección.
12.  $A \cup \phi = A$ . Elemento neutro de la unión de conjuntos.
13.  $A \cap \phi = \phi$ .
14.  $A \cap A^c = \phi$ .  $A$  y  $A^c$  son mutuamente excluyentes.
15.  $A \cup A^c = U$ .  $A$  y  $A^c$  son colectivamente exhaustivos.
16.  $A^c = U - A$ .
17. Si  $A \subseteq B$  y  $B \subseteq C \rightarrow A \subseteq C$ . Ley transitiva de los subconjuntos.
18. Si  $A \subseteq B \rightarrow B^c \subseteq A^c$ .
19.  $(A \cup B)^c = A^c \cap B^c$ . Primera Ley de De Morgan.
20.  $(A \cap B)^c = A^c \cup B^c$ . Segunda Ley de De Morgan.
21. Si  $A$  y  $B$  no tienen elementos en común,  $n(A \cup B) = n(A) + n(B)$ .
22. Si  $A$  y  $B$  no tienen elementos en común  $n(A \cap B) = 0$ .
23. Para dos conjuntos cualesquiera  $A$  y  $B$ , se cumple que:  
 $n(A \cup B) = n(A) + n(B) - n(A \cap B)$ .



### 2.3. Principios Fundamentales del Conteo y Análisis Combinatorio

El análisis combinatorio es la parte de las matemáticas discretas que estudia la forma en la que se pueden enumerar los grupos que se consiguen formar con elementos de conjuntos, teniendo en cuenta, bien los objetos tomados, o bien el orden en que se toman, o ambos valores a la vez.

El análisis combinatorio también conocido como técnicas de conteo permite determinar el número total de posibles resultados lógicos que cabe esperar al realizar algún experimento, sin tener necesidad de enumerar cada uno de ellos. En esta definición se entiende por experimento cualquier acto que se realice o presente con el fin de observar su resultado.

Antes de proceder a analizar las técnicas de conteo que se utilizan, se definirá una función muy importante en la matemática aplicada, denominada Función Gamma.

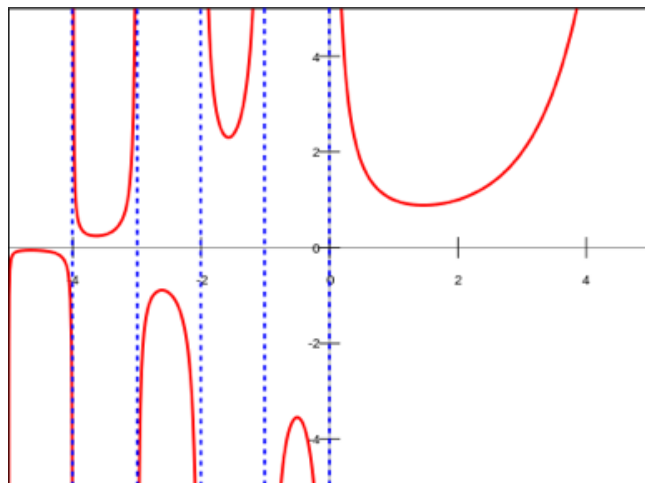
La importancia de esta función radica en que se utiliza para el desarrollo de otras funciones. Aparece ocasionalmente en los problemas físicos tal como la normalización de las funciones de onda de Coulomb y en el cómputo de las probabilidades en la mecánica estadística. La función Gamma se define por la siguiente expresión integral:

$$\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt \quad ; z > 0 \quad (2.1)$$

A esta forma de definirla se le llama la Forma de Euler. Su gráfica, extendida para valores negativos, se presenta en la figura 2.7 a continuación.

**FIGURA 2.7.** Gráfica de la Función Gamma

**NOTA:** Función gamma. (2019). [https://es.wikipedia.org/wiki/Funci%C3%B3n\\_gamma](https://es.wikipedia.org/wiki/Funci%C3%B3n_gamma)



Si se hace  $z=1/2$  se obtiene que:

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} e^{-t} \sqrt{t} dt = \sqrt{\pi} \quad (2.2)$$

La Función Gamma puede ser escrita como:

$$\Gamma(z+1) = \int_0^{\infty} e^{-t} t^z dt \quad (2.3)$$

Fácilmente se comprueba con esta última expresión que si  $z=0$  entonces

$$\Gamma(1) = \int_0^{\infty} e^{-t} dt = 1 \quad (2.4)$$

El resultado más importante que se puede deducir de esta expresión se obtiene al considerar a  $z$  como un número natural, es decir,  $z=n$ ,  $n \in \mathbb{N}$ , y resolver la integral (2.3) por partes:

$$\Gamma(n+1) = \int_0^{\infty} e^{-t} t^n dt = n(n-1)(n-2)(n-3) \cdots 1 = n! \quad (2.5)$$

Por esta propiedad, a la Función Gamma se le denomina también Función Factorial Generalizada.

*Principio de Adición:* Si un acto, evento o suceso  $A$  puede efectuarse de  $m$  maneras diferentes y si otro suceso  $B$  distinto del anterior, puede llevarse a cabo de  $n$  maneras diferentes, entonces, uno cualquiera de los eventos  $A$  o  $B$  puede ocurrir de  $m+n$  diferentes formas, siempre y cuando  $A$  y  $B$  no puedan ocurrir simultáneamente.

### Ejercicio 2.2

Para ir hacia el norte de una ciudad hay cuatro caminos diferentes y para ir hacia el sur de la misma ciudad hay tres caminos diferentes, de tal manera que hay siete formas de salir de dicha ciudad.

### Ejercicio 2.3

Sean  $A = \{a, b, c, d, e, f\}$  y  $B = \{c, e, g, h\}$ .

$$\begin{aligned} n(A) &= 6; n(B) = 4; n(A \cap B) = 2, \text{ entonces:} \\ n(A \cup B) &= n(A) + n(B) - n(A \cap B) = 6 + 4 - 2 = 8. \end{aligned}$$

*Principio de Multiplicación:* Si un acto, evento o suceso puede realizarse de  $m$  maneras diferentes y, efectuado este, un segundo acto puede efectuarse de  $n$  maneras diferentes, entonces ambos actos pueden efectuarse en el orden indicado de  $mn$  maneras distintas.

## Ejercicio 2.4

Para ir del punto  $A$  al punto  $B$  existen tres caminos y para ir del punto  $B$  al punto  $C$  existen cuatro caminos diferentes.

- ¿De cuántas maneras diferentes se puede ir de  $A$  a  $C$ , pasando por  $B$ ?  
Por el Teorema Fundamental del Conteo, existen  $3 \cdot 4 = 12$  maneras diferentes de ir de  $A$  a  $C$  pasando por  $B$ .
- ¿De cuántas maneras diferentes se puede ir de  $A$  a  $C$ , pasando por  $B$  y regresar a  $C$ ?  
Por el Teorema Fundamental del Conteo, existen  $3 \cdot 4 \cdot 3 = 36$  maneras diferentes de ir de  $A$  a  $C$  pasando por  $B$ .
- ¿De cuántas maneras diferentes se puede hacer un viaje redondo de  $A$  a  $C$  y de  $C$  a  $A$ , pasando por  $B$ , si no se permite usar cada camino más de una vez?  
Por el Teorema Fundamental del Conteo, existen  $3 \cdot 4 \cdot 3 \cdot 2 = 72$  maneras diferentes de ir de  $A$  a  $C$  y de  $C$  a  $A$  pasando por  $B$ .

El principio fundamental del conteo puede generalizarse de la siguiente forma:

*Teorema Fundamental del Conteo Generalizado.* Si el evento  $A_1$  puede realizarse de  $n_1$  maneras diferentes y una vez efectuado este, el evento  $A_2$  puede llevarse a efecto de  $n_2$  maneras diferentes, y luego el evento  $A_3$  puede efectuarse de  $n_3$  formas diferentes y así sucesivamente, el evento  $A_k$  puede efectuarse de  $n_k$  maneras diferentes, entonces los  $k$  eventos pueden efectuarse uno tras otro en el orden establecido de  $n_1 n_2 \cdots n_k$  maneras diferentes.

Se llaman ordenaciones de  $n$  objetos de orden  $r$ , a los diferentes grupos ordenados que se pueden formar al escoger  $r$  objetos de un grupo de  $n$  objetos dados,  $r \leq n$ , de tal forma que dos ordenaciones se consideran distintas si difieren en alguno de sus objetos o en el orden de ellos.

Por el Teorema Fundamental del Conteo, si se tiene un conjunto con  $n$  elementos, el primer elemento a seleccionar se puede escoger de  $n$  maneras diferentes, el segundo elemento puede escogerse de  $n-1$  elementos diferentes, el

tercero puede seleccionarse de  $n-2$  formas diferentes y así sucesivamente, por lo que el  $r$ -ésimo elemento puede ser seleccionado de  $n-r+1$  maneras diferentes, lo que conlleva a la siguiente fórmula:

$$O(r,n) = O_n^r = n(n-1)(n-2) \cdots (n-r+1)$$

$$O_n^r = \frac{n(n-1)(n-2) \cdots (n-r+1)(n-r)(n-r-1) \cdots 1}{(n-r)(n-r-1) \cdots 1}$$

$$O_n^r = \frac{n!}{(n-r)!} \quad (2.6)$$

Si se permite la repetición de objetos en cada grupo, entonces, la fórmula es:

$$OR_n^r = n^r \quad (2.7)$$

### Ejercicio 2.5

Calcular el número de placas diferentes que se pueden formar con las letras A, B, C, D, E y con los dígitos 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 si cada placa contiene en forma ordenada tres letras y tres números si:

a) No se permite la repetición.

El número de ordenaciones diferentes que se pueden hacer con cinco letras de orden tres es  $(5)(4)(3) = 60$ . El número de ordenaciones diferentes que se pueden hacer con diez dígitos de orden tres es  $(10)(9)(8) = 720$ , por lo que el número de placas diferentes que se pueden formar es  $(60)(720) = 43200$ .

b) Se permite la repetición.

En este caso:  $(5^3)(10^3) = 125\,000$ .

c) No se permite que el cero esté al principio.

$(5)(5)(5)(9)(10)(10) = 112\,500$ .

Una permutación es el número de maneras diferentes en que un grupo de  $n$  objetos dados pueden ordenarse. Se trata de una ordenación de  $n$  objetos de orden  $n$ , la fórmula de cálculo es:

$$P_n = O_n^n = n! \quad (2.8)$$

Si se permite la repetición, la fórmula es:

$$PR_n = OR_n^n = n^n \quad (2.9)$$

Se llaman combinaciones de  $n$  objetos de orden  $r$  a los diversos grupos que pueden formarse al elegir  $r$  objetos de  $n$  objetos dados, de tal manera que dos combinaciones se consideran distintas en uno de sus objetos por lo menos. En las combinaciones el orden no importa. Para este caso se considera que  $r \leq n$ . La fórmula se deduce al dividir las ordenaciones de  $n$  objetos de orden  $r$  entre el número de permutaciones de  $r$  objetos para evitar la repetición. La fórmula de las combinaciones es la siguiente:

$$C(r, n) = C_n^r = \frac{n!}{r!(n-r)!} \quad (2.10)$$

A diferencia de las ordenaciones, en las combinaciones no interesa el orden de los objetos, sino únicamente la clase de los mismos.

En el caso de que se permitan las repeticiones se obtiene la fórmula:

$$CR(r, n) = CR_{n+r-1}^r = \frac{(n+r-1)!}{r!(n-1)!} \quad (2.11)$$

### Ejercicio 2.6

De un grupo de diez personas debe elegirse un comité formado por cinco. Calcular el número de comités diferentes que se pueden elegir si:

- a) Las diez personas son elegibles libremente.

$$C(5, 10) = 10!/(5!5!) = 10(9)(8)(7)(6)/5! = 252.$$

- b) En el comité debe haber un presidente.

Existen 10 maneras de elegir un presidente y  $C(4, 9)$  maneras de obtener los demás integrantes, por lo que:

$$10C(4, 9) = 10(9!/(4!5!)) = 1260.$$

## Ejercicio 2.7

Se escogen dos cartas de una baraja normal de 52 cartas.

FIGURA 2.8. Baraja



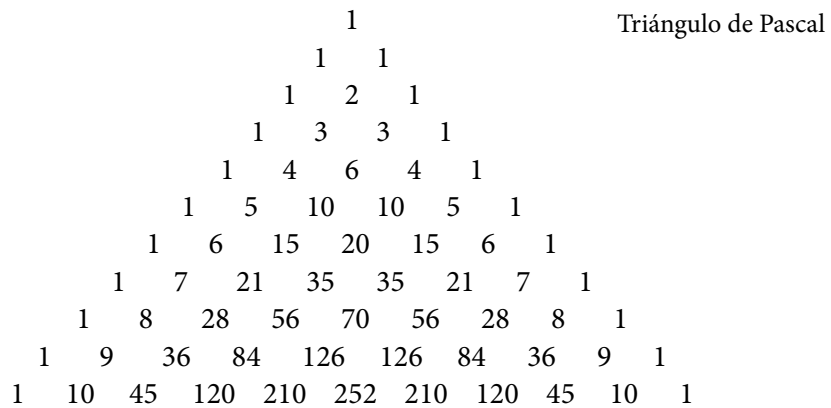
a) ¿De cuántas maneras diferentes se pueden escoger?

$$C(2, 52) = (52)(51)/2 = 1326.$$

b) ¿De cuántas maneras diferentes se pueden escoger un As y una figura?

$$4 \text{ ases} * 12 \text{ figuras} = 48.$$

A los números  $C(r, n)$  también se les conoce como números combinatorios o coeficientes binomiales y la razón se explicará a continuación. Observe los números que se muestran en la siguiente figura a la cual se le conoce como Triángulo de Pascal. Observe que dicho triángulo es simétrico, que la primera fila corresponde con el valor de  $n=0$ , la segunda fila es para  $n=1$ , la tercera fila para  $n=2$  y así sucesivamente. El primer número de cada fila corresponde con 1, el segundo y el penúltimo número corresponde con  $n$ , el tercer y el antepenúltimo número corresponde con  $n(n-1)/2!$ , el cuarto y el número anterior al antepenúltimo corresponde con  $n(n-1)(n-2)/3!$ , y así sucesivamente. También, nótese que el valor situado en la fila  $i$  y en la columna  $j$  es resultado de la suma de los números que se encuentran centrados inmediatamente en la fila arriba de él.



Observe qué ocurre al calcular las combinaciones  $C(r, n)$  como se muestra en la figura 2.10.

**FIGURA 2.10.** Números combinatorios o coeficientes binomiales

$$\binom{0}{0} = 1$$

$$\binom{1}{0} = 1 \quad \binom{1}{1} = 1$$

$$\binom{2}{0} = 1 \quad \binom{2}{1} = 2 \quad \binom{2}{2} = 1$$

$$\binom{3}{0} = 1 \quad \binom{3}{1} = 3 \quad \binom{3}{2} = 3 \quad \binom{3}{3} = 1$$

$$\binom{4}{0} = 1 \quad \binom{4}{1} = 4 \quad \binom{4}{2} = 6 \quad \binom{4}{3} = 4 \quad \binom{4}{4} = 1$$

$$\binom{5}{0} = 1 \quad \binom{5}{1} = 5 \quad \binom{5}{2} = 10 \quad \binom{5}{3} = 10 \quad \binom{5}{4} = 5 \quad \binom{5}{5} = 1$$

$$\binom{6}{0} = 1 \quad \binom{6}{1} = 6 \quad \binom{6}{2} = 15 \quad \binom{6}{3} = 20 \quad \binom{6}{4} = 15 \quad \binom{6}{5} = 6 \quad \binom{6}{6} = 1$$

$$\binom{7}{0} = 1 \quad \binom{7}{1} = 7 \quad \binom{7}{2} = 21 \quad \binom{7}{3} = 35 \quad \binom{7}{4} = 35 \quad \binom{7}{5} = 21 \quad \binom{7}{6} = 7 \quad \binom{7}{7} = 1$$

Los números combinatorios de la figura 2.10 y los números que componen el triángulo de Pascal de la figura 2.9 son iguales y dan origen a los coeficientes de los términos en el binomio de Newton, como se aprecia en su desarrollo:



$$(a+b)^n = a^n b^0 + n a^{n-1} b^1 + \frac{n(n-1)}{2!} a^{n-2} b^2 + \frac{n(n-1)(n-2)}{3!} a^{n-3} b^3 + \dots + \frac{n(n-1)(n-2)}{3!} a^3 b^{n-3} + \frac{n(n-1)}{2!} a^2 b^{n-2} + n a^1 b^{n-1} + a^0 b^n \quad (2.12)$$

Lo que también puede ser escrito como:

$$(a+b)^n = \binom{n}{0} a^n b^0 + \binom{n}{1} a^{n-1} b^1 + \binom{n}{2} a^{n-2} b^2 + \binom{n}{3} a^{n-3} b^3 + \dots + \binom{n}{n-3} a^3 b^{n-3} + \binom{n}{n-2} a^2 b^{n-2} + \binom{n}{n-1} a^1 b^{n-1} + \binom{n}{n} a^0 b^n \quad (2.13)$$

Nótese qué ocurre si en el binomio de Newton de la ecuación 2.13 se hace  $a=b=1$ . Se obtiene que la suma de las combinaciones de  $n$  números de todos los órdenes  $r$  está dado por la siguiente expresión:

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \binom{n}{3} + \dots + \binom{n}{n-3} + \binom{n}{n-2} + \binom{n}{n-1} + \binom{n}{n} = 2^n \quad (2.14)$$

### Ejercicio 2.8

Suponga que tiene un conjunto con  $N$  elementos, ¿cuántas muestras diferentes se pueden obtener de esta población?

Para ilustrar lo anterior, suponga que  $N = \{a, b, c, d, e\}$ :

Muestras de tamaño uno:	$\{a\}, \{b\}, \{c\}, \{d\}, \{e\}$ .
Muestras de tamaño dos:	$\{a, b\}, \{a, c\}, \{a, d\}, \{a, e\}, \{b, c\}, \{b, d\}, \{b, e\}, \{c, d\}, \{c, e\}, \{d, e\}$ .
Muestras de tamaño tres:	$\{a, b, c\}, \{a, b, d\}, \{a, b, e\}, \{a, c, d\}, \{a, c, e\}, \{a, d, e\}, \{b, c, d\}, \{b, c, e\}, \{b, d, e\}, \{c, d, e\}$ .
Muestras de tamaño cuatro:	$\{a, b, c, d\}, \{a, b, c, e\}, \{a, b, d, e\}, \{a, c, d, e\}, \{b, c, d, e\}$ .
Muestras de tamaño cinco:	$\{a, b, c, d, e\}$ .

Como se puede apreciar, existen  $2^5 - 1 = 31$  muestras diferentes que se pueden extraer de una población de cinco elementos.

En una población de 10 elementos existirían  $2^{10} - 1 = 1,023$  muestras.

En una muestra de 20 elementos existirían  $2^{20} - 1 = 1,048,575$  muestras.

En una muestra de 50 elementos existirían  $2^{50} - 1 = 1,125,899,906,842,620$  muestras.

En general, en una muestra finita de  $N$  elementos existirían  $2^N - 1$  muestras.

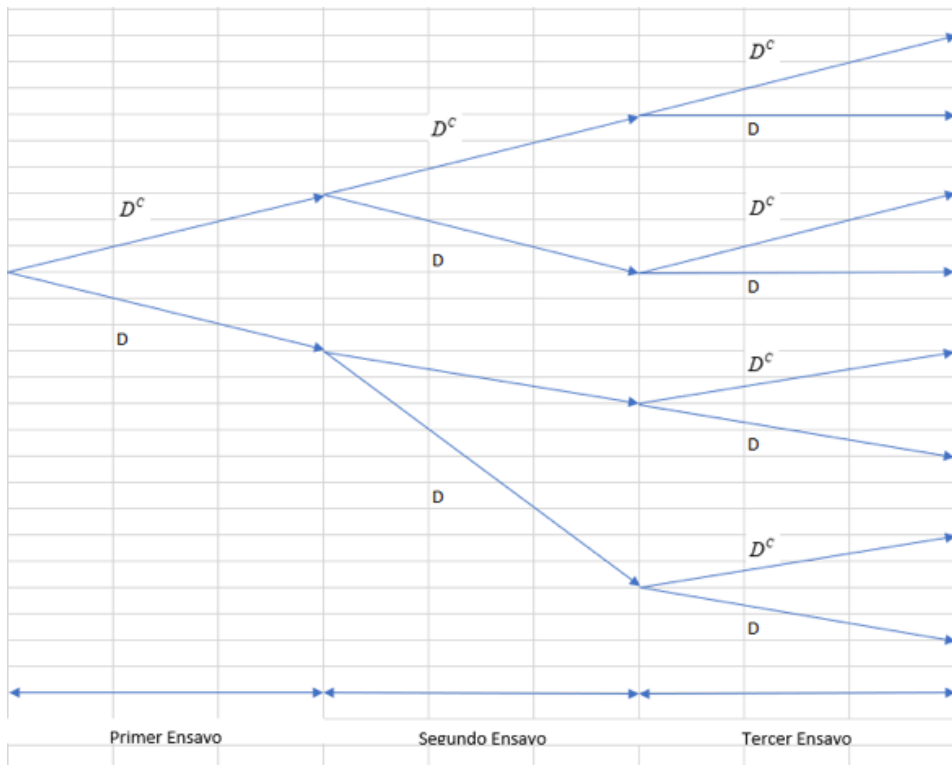
Un diagrama de árbol o árbol de decisión es una herramienta gráfica que se usa para enumerar todas las posibilidades lógicas de una secuencia de actos o eventos que ocurran de un número finito de maneras.

### Ejercicio 2.9

Un experimento consiste en escoger un artículo de un lote de  $N$  de ellos y determinar si este es defectuoso o no defectuoso. Se realizan tres ensayos. Determine el diagrama de árbol correspondiente.

En la figura 2.11 se muestra dicho árbol.  $D$  significa que se extrajo un artículo defectuoso y  $D^c$  expresa que se extrajo un producto No Defectuoso.

FIGURA 2.11. Diagrama de Árbol



A partir del diagrama de árbol de la figura 3.11 se obtienen los posibles resultados, cada rama representa un posible resultado.

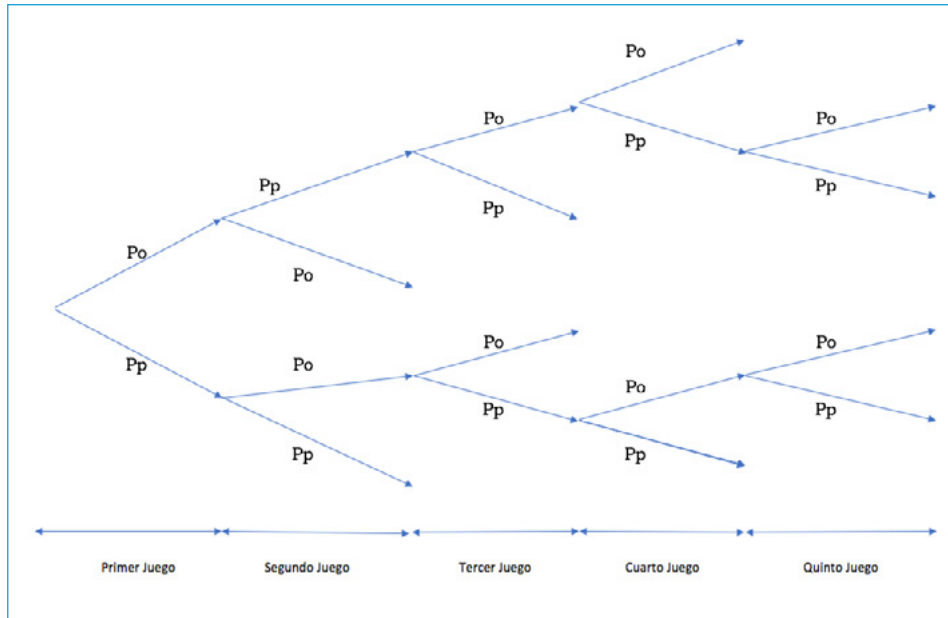
Espacio de resultados =  $\{D^c D^c D^c, D^c D^c D, D^c D D^c, D^c D D, D D^c D^c, D D^c D, D D D^c, D D D\}$ .

## Ejercicio 2.10

Un experimento podría ser: Polo y Pepe participan en un torneo de Ping-Pong. El primero que gane dos juegos seguidos o un total de tres será el vencedor del torneo. ¿De cuántas maneras puede terminar el torneo?

En la figura 2.12 se muestran los posibles resultados. *Po* significa que el juego lo ganó Polo y *Pp* que lo ganó Pepe.

FIGURA 2.12. Diagrama de Árbol del Juego de Ping-Pong



Como se puede apreciar en la figura 2.12 el torneo tiene 10 formas diferentes de terminar.

El número total de maneras de dividir un conjunto de  $n$  elementos en  $k$  subconjuntos mutuamente excluyentes y colectivamente exhaustivos de  $n_1, n_2, n_3, \dots, n_k$  elementos para los cuales se cumple que  $n_1 + n_2 + n_3 + \dots + n_k = n$  es

$$\binom{n}{n_1 \ n_2 \ n_3 \ \dots \ n_k} = \frac{n!}{n_1! \ n_2! \ n_3! \ \dots \ n_k!}; \quad n_1 + n_2 + n_3 + \dots + n_k = n \quad (2.15)$$

## 2.4. Enfoques, Interpretaciones y Escuelas de Probabilidad

Históricamente, las distintas interpretaciones de probabilidad han dado origen a cuatro escuelas de probabilidad que se pueden clasificar en:

1. Clásica o de Laplace.
2. Frecuentista o de Von Misses.
3. Subjetivista o de Savage.
4. Axiomática, Constructivista o de Kolmogorov.

En el contexto de la ingeniería, cada una es valiosa en sí misma y las cuatro son aplicables en diferentes situaciones o contextos.

### 2.4.1. Escuela Clásica o de Laplace

Si un evento  $A$  contenido en el espacio muestral finito  $S$  de un experimento, está formado por  $N(A)$  puntos muestrales y el espacio muestral por  $N(S)$  puntos muestrales, igualmente verosímiles o que tienen la misma posibilidad de ocurrir (equiprobables), se dice que la probabilidad de que el evento  $A$  ocurra, está dada por la relación

$$p(A) = \frac{N(A)}{N(S)} \quad (2.16)$$

Se dice que este valor es una probabilidad a priori, porque se asigna sin depender de ninguna experiencia previa a la realización del experimento, o sin que este se realice; solo se deben observar dos requisitos: que el número de casos posibles sea finito, y que todos y cada uno de esos casos posibles sean igualmente verosímiles. La teoría clásica conformó el concepto de probabilidad a partir de un pensamiento determinista. Ante la falta de conocimiento con relación al experimento aleatorio del que se derivan los eventos, se parte del supuesto de que existe una simetría recíproca entre los distintos resultados posibles, simetría que permite considerar que todos los resultados posibles pueden ser considerados equivalentes desde el punto de vista de la probabilidad.

Desde la perspectiva matemática, la teoría laplaciana no presenta contradicción alguna; cumple con las exigencias del rigor matemático, a pesar de que no fue construida sobre una base axiomática. En cuanto a la definición misma,

resulta tautológica, ya que en ella se hace alusión al concepto que se pretende definir; cuando se habla de casos posibles, significa igualmente posibles, lo cual equivale a suponer equiprobabilidad. La definición clásica resulta endeble, basta con que el resultado se sesgue ligeramente a favor de un resultado, por ejemplo, que en un nacimiento sea más factible un niño que una niña, para que la definición ya no sea aplicable, porque deja de haber simetría.

## Ejercicio 2.11

Un dado numérico de seis caras, hecho de un material homogéneo, con caras idénticas y redondeado de las esquinas, al lanzarlo tiene seis posibilidades igualmente verosímiles de caer. Su espacio muestral es  $S = \{1, 2, 3, 4, 5, 6\}$ , por lo que la probabilidad de que caiga el 3 es  $p(x=3) = \frac{1}{6}$ . La probabilidad de que caiga un número impar es  $p(x \text{ sea impar}) = \frac{3}{6} = \frac{1}{2}$ . La probabilidad de que caiga un múltiplo de 3 es  $p(x \text{ múltiplo de } 3) = \frac{2}{6} = \frac{1}{3}$ . La probabilidad de que caiga el 9 es  $p(x=9) = \frac{0}{6} = 0$ . La estimación que se hace de estas probabilidades parte del supuesto de que el dado es homogéneo, si el dado está cargado todas las estimaciones anteriores serían falsas.

Otro inconveniente de la definición clásica es que no considera la probabilidad de que el número total de posibles resultados sea infinito. Paradójicamente existen muchos casos de aplicación, en los que es evidente la equiprobabilidad en espacios muestrales continuos. El ejercicio que se muestra a continuación tiene un número infinito de posibles resultados, pero hay una manera de estimar la probabilidad desde un punto de vista geométrico.

## Ejercicio 2.12

Se escogen dos puntos  $x$  y  $y$  aleatoria e independientemente uno de otro, sobre un segmento de recta  $[A, B]$ . Obtenga la probabilidad de que la distancia de  $y$  a  $x$  sea por lo menos el doble de la que hay desde  $A$  hasta  $x$ .

$$x \in [A, B], y \in [A, B] \quad \rightarrow \quad |y-x| \geq 2(x-A)$$

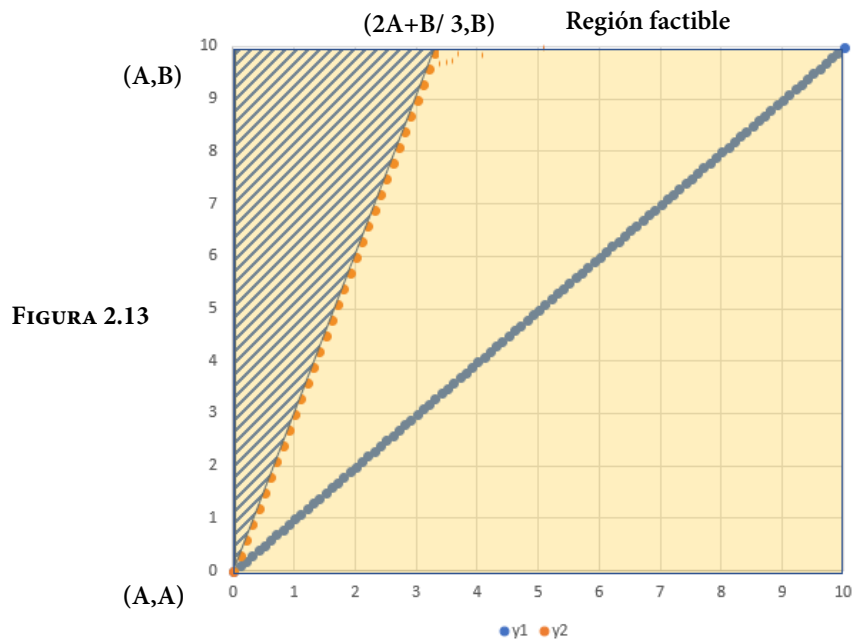
$$\begin{aligned} \text{Primer caso: } y-x \geq 0 \quad &\rightarrow \quad y \geq x \\ y-x \geq 2(x-A) \quad &\rightarrow \quad y \geq 3x-2A \end{aligned}$$

$$\begin{aligned} \text{Segundo caso: } y-x < 0 \quad &\rightarrow \quad y < x \\ -(y-x) \geq 2(x-A) \quad &\rightarrow \quad y \leq -x+2A \end{aligned}$$

Para trazar las regiones se tomará como intervalo  $[A, B] = [0, 10]$ . En la figura 2.16 se muestra el conjunto universal  $S = \{(x, y) \mid A \leq x \leq B; A \leq y \leq B; x, y \in R\}$ . La región factible que cumple las desigualdades es la que se muestra con trama en la figura 2.13.

La probabilidad de que la distancia de  $y$  a  $x$  sea por lo menos el doble de la distancia de  $x$  a  $A$  estaría dada como el cociente del área de la región factible entre el área total.

Área de la región factible = área del triángulo formado por los puntos  $(A, A)$ ,  $(A, B)$ ,  $((2A+B)/3, B)$



$$\text{Área de la región factible} = \left[\frac{(B-A)}{3}\right] \left[\frac{(B-A)}{2}\right] = \frac{(B-A)^2}{6}$$

$$\text{Área completa} = (B-A)^2$$

$$p(|y-x| > x-A) = \frac{1}{6} = 0.1667$$

### 2.4.2. Escuela Frecuentista o de Von Mises

Suponga que se realiza un experimento  $n$  veces y se observa que un evento  $A$  ocurre  $n(A)$  veces, el cociente  $f_A = n(A)/n$  se denomina frecuencia relativa del evento  $A$ , la cual satisface las siguientes propiedades:

- » Los resultados favorables son menores o cuando más iguales a la cantidad de resultados posibles que arroja el juego:  $n(A) \leq n$ .
- » El valor de esa fracción siempre estará ubicado entre 0 y 1:  $0 \leq f_A \leq 1$ .
- » Los valores más cercanos a 1 indican una alta posibilidad de ocurrencia de un resultado. Si  $f_A = 1$  entonces  $A$  ocurrió en las  $n$  repeticiones. Los valores más cercanos a 0 indican una baja o nula posibilidad de ocurrencia. Si  $f_A = 0$  entonces  $A$  no ocurrió en ninguna de las  $n$  repeticiones.
- » Si dos eventos son mutuamente exclusivos, la frecuencia relativa de al menos uno de ellos es la suma de sus frecuencias relativas:

$$\text{Si } A \cap B = \phi \rightarrow f_{A \cup B} = f_A + f_B \quad (2.17)$$

Como resultado de la observación del comportamiento de numerosas repeticiones de experimentos reales, tras una larga serie de realizaciones del experimento, la frecuencia relativa observada de un evento se aproxima a la probabilidad del evento; es decir, la probabilidad a priori, puede ser corroborada a posteriori, empíricamente. La frecuencia relativa correspondiente a un evento, puede ser considerada como una buena aproximación de la probabilidad de ese evento. Si la probabilidad de un evento simple  $A$ , de un espacio muestral asociado a un experimento, tiene un valor  $P(A)$  conocido o desconocido, al realizarse repetidamente el experimento, una y otra vez,  $n$  veces; si se cuenta el número de veces  $n(A)$  que el evento  $A$  fue observado, la relación  $n(A)/n$  es tan cercana al valor  $P(A)$ , que puede considerarse como una buena aproximación de  $P(A)$ :

$$p(A) \approx \frac{n(A)}{n} \quad (2.18)$$

En muchos problemas, la probabilidad de obtener algún resultado específico se puede interpretar como la frecuencia relativa con la que se obtendría ese resultado, si el experimento se repitiera un número grande de veces, en condiciones similares. No hay claridad de cuál podría ser un número suficientemente grande de veces, ni de qué se debe entender por condiciones similares, pues



desde luego no se trata de condiciones idénticas, ya que entonces el experimento ya no sería aleatorio sino determinista. Tales condiciones son demasiado vagas para servir como base de una definición rigurosa de probabilidad. Por otra parte, si se realizaran otras secuencias de ensayos del mismo experimento, para un evento en particular, se obtendrían diferentes valores de su frecuencia relativa, quizá muy similares entre sí, pero finalmente diferentes; no se sabría cuál de ellos es el más cercano a la verdadera probabilidad.

Cuando se usa la frecuencia relativa  $f_A$  como aproximación de un valor establecido de la probabilidad  $P(A)$ , se debe tener consciencia de que  $f_A \neq P(A)$  y de que se está cometiendo un error de estimación. Se puede intuir esa propiedad de los fenómenos aleatorios, conocida como regularidad estadística, que consiste en que, a medida que se aumenta el número de repeticiones del experimento, en condiciones similares, varía cada vez menos la frecuencia relativa  $f_A$  y tiende a estabilizarse en un valor fijo  $P(A)$ , o dicho de otra manera, la frecuencia relativa  $f_A$  converge a la probabilidad  $P(A)$ , excepto que esta convergencia en probabilidad debe ser entendida como un hecho empírico y no como un resultado matemático. La estabilidad de las frecuencias en ensayos repetidos es un hecho objetivo de naturaleza independiente del conocimiento de cualquier persona sobre ello. La interpretación frecuencial de probabilidad consiste en asignar como probabilidad del evento  $A$  a la relación  $n(A)/n$ , que es su frecuencia relativa. La probabilidad concebida como proporción hace que la interpretación frecuencial sea fácilmente asimilada. Debe ser establecida como criterio formal de asignación (se trata de una interpretación a posteriori) pues la probabilidad solo puede asignarse luego de haber realizado el experimento en forma repetida. Es de esperar que, mientras mayor sea el número de veces que se realice el experimento, esta aproximación será mejor y, en el límite, se obtendrá el valor preciso:

$$p(A) = \lim_{n \rightarrow \infty} \frac{n(A)}{n} \quad (2.19)$$

Este resultado es atribuido a Jacob Bernoulli y es conocido como Ley de los Grandes Números, en su forma más básica e intuitiva: la frecuencia relativa tiende a variar cada vez menos, a medida que el número de observaciones va creciendo, estabilizándose alrededor de un valor definido. El principio de estabilidad de las frecuencias permite obtener la probabilidad de un evento siempre que se dispone de un número suficiente de observaciones para calcular la frecuencia relativa. Cuando una muestra es tomada correctamente, la frecuencia relativa de un evento supone la probabilidad de dicho evento en el total de la población.

La Ley de los Grandes Números establece que si  $x_1, x_2, x_3, \dots$  es una sucesión infinita de variables aleatorias independientes e idénticamente distribuidas, que cumplen que su media o esperanza matemática  $E(|X_i|) < \infty$  y tienen el valor esperado  $\mu$ , entonces:

$$p(\lim_{n \rightarrow \infty} \bar{x}_n = \mu_x) = 1 \quad (2.20)$$

es decir, el promedio de las variables aleatorias converge a  $\mu_x$  casi seguramente (en un conjunto de probabilidad 1).

Para una demostración completa de la Ley de los Grandes Números consulte la demostración de la Ley fuerte de los Grandes Números de Kolmogorov, en la página: [https://es.wikipedia.org/wiki/Ley\\_de\\_los\\_grandes\\_n%C3%BAmeros](https://es.wikipedia.org/wiki/Ley_de_los_grandes_n%C3%BAmeros) o más formalmente, consulte el trabajo de Yu V. Prokhorov, “On the strong law of large numbers” *Izv. Akad. Navk SSSR Ser. Mat.*, 14(1950) pp. 523-536.

Esta ley justifica la interpretación intuitiva del valor esperado de una variable aleatoria como el “promedio a largo plazo al hacer un muestreo repetitivo”.

La interpretación frecuencial no es aplicable para asignar probabilidades a eventos aislados que no son susceptibles de experimentación. Los frequentistas hablan de probabilidades solo cuando se trata de experimentos aleatorios bien definidos, de modo que este criterio está limitado a los casos en que los resultados presentan regularidad estadística, consecuencia de la observación de un experimento que es considerado aleatorio, porque es repetible bajo las mismas condiciones.

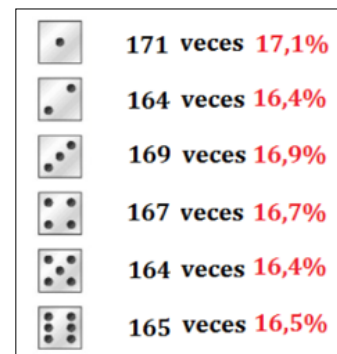
### Ejercicio 2.13

En un local de apuestas una persona llevó su propio dado y, ante la desconfianza de que estuviera cargado, lo probaron lanzándolo  $N=1000$  veces. Los resultados se muestran en la figura 2.14. Estime la probabilidad de que caiga el 3.

**FIGURA 2.14.** Resultados de lanzar un dado 1000 veces

**NOTA:** Dado. (2015). <http://ilcolibri.altervista.org/webtraining/wp-content/uploads/2015/12/Estadisticas-dado.png>

Probabilidad de que  $p(x=3) = 169/1000 = 0.169$ .



### 2.4.3. Escuela Subjetivista o de Savage

La asignación de probabilidades a eventos no equiprobables, correspondientes a experimentos no repetibles, puede hacerse a través de lo que se conoce como interpretación subjetivista de probabilidad. Bajo esta escuela, la probabilidad de un evento  $A$  puede ser simplemente una medida del grado de credibilidad que se tiene sobre la ocurrencia del evento  $A$ , expresado en términos numéricos. Se ha desarrollado una serie de teorías subjetivistas, que consideran que *la probabilidad de un evento es el grado de creencia que un sujeto tiene sobre la ocurrencia de ese evento, determinada a partir de su intuición, sentimiento, sentido común, experiencia o conocimiento*. Otra persona puede tener diferente opinión o información distinta y asignar una probabilidad diferente al mismo resultado. Se requiere de un individuo idealmente racional que enjuicie la fuerza de la evidencia, desde la más absoluta imparcialidad, haciéndola equivaler a un determinado grado de probabilidad.

Cuando se trabaja con probabilidad subjetiva, lo más importante es la consistencia de las asignaciones; el mismo sujeto, ante dos hipótesis lógicamente equivalentes, dadas las evidencias para ambas, lógicamente equivalentes también, asignará grados de creencia iguales; por lo general, la asignación subjetiva de probabilidades lleva consigo emociones, optimismo, pesimismo, deseos de que ocurra o de que no ocurra; todo eso se debe eliminar, de ser posible.

#### Ejercicio 2.14

En un juego de fútbol americano, los Cóndores de la UNAM tienen tiempo para una jugada más y necesitan anotar en esa jugada para ganar el partido. El entrenador del equipo puma siente que hay solamente tres jugadas que se pueden usar: un pase largo, un pase pantalla o una corrida. La probabilidad de que el entrenador mande un pase largo es 0.35; de que mande un pase pantalla, 0.25; y de que mande una corrida, 0.40. Basado en su experiencia pasada sobre situaciones similares, el entrenador siente que las probabilidades de lograr una anotación con tales jugadas son: pase largo 0.5, pase pantalla 0.3 y corrida 0.1.

### 2.4.4. Escuela Axiomática, Constructivista o de Kolmogórov

Sean  $S$  el espacio muestral y  $A$  cualquier evento de  $S$ . Se llamará función de probabilidad sobre el espacio muestral  $S$  a  $p(A)$  si satisface los siguientes axiomas:

$$1. P(A) \geq 0 \quad (2.21)$$

$$2. P(S) = 1 \quad (2.22)$$

$$3. \text{ Si } A \cap B = \phi \rightarrow p(A \cup B) = p(A) + p(B) \quad (2.23)$$

Teoremas subsecuentes:

$$a) p(\phi) = 0 \quad (2.24)$$

Demostración: De las propiedades del conjunto vacío, se sabe que  $A \cap \phi = \phi$  y que  $A \cup \phi = A$ , por lo que si se aplican los axiomas 2.21 y 2.23:

$$p(A \cup \phi) = p(A) = p(A) + p(\phi) \quad \rightarrow \quad p(\phi) = 0$$

$$b) \text{ Si } A \subset S \rightarrow 0 \leq p(A) \leq 1 \quad (2.25)$$

Demostración: De las propiedades del conjunto vacío, se sabe que:

$\phi \subseteq A \subseteq S$  por lo que aplicando el teorema 2.24 y el axioma de la expresión 2.23, se cumple la propiedad 2.25.

$$c) p(A^c) = 1 - p(A) \quad (2.26)$$

Demostración:  $A \cap A^c = \phi$  y  $A \cup A^c = S$ , por lo que aplicando los axiomas 2.23 y 2.24, se obtiene que  $p(A \cup A^c) = p(S) \rightarrow p(A) + p(A^c) = 1 \rightarrow p(A^c) = 1 - p(A)$

$$d) \text{ Sean } A \subset S \text{ y } B \subset S \rightarrow p(A \cup B) = p(A) + p(B) - p(A \cap B) \quad (2.27)$$

Demostración:

Por una parte  $A \cap (B \cap A^c) = \phi$ , por otra parte,  $A \cup B = A \cup (B \cap A^c)$

De la misma forma  $(A \cap B) \cap (B \cap A^c) = \phi$  y  $B = (A \cap B) \cup (B \cap A^c)$

De esta forma  $p(A \cup B) = p(A) + p(B \cap A^c)$  y  $p(B) = p(A \cap B) + p(B \cap A^c)$

Si se restan las expresiones anteriores:

$$p(A \cup B) - p(B) = p(A) - p(A \cap B) \quad \rightarrow \quad p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

e) El teorema anterior se puede generalizar a tres o más conjuntos:

$$p(A \cup B \cup C) = p(A) + p(B) + p(C) - p(A \cap B) - p(A \cap C) - p(B \cap C) + p(A \cap B \cap C) \quad (2.28)$$

$$p(A_1 \cup A_2 \cup \dots \cup A_k) = \sum_{i=1}^k p(A_i) - \sum_{i < j=2}^k p(A_i \cap A_j) + \sum_{i < j < r=3}^k p(A_i \cap A_j \cap A_r) + \dots + (-1)^{k-1} p(A_1 \cap A_2 \cap \dots \cap A_k) \quad (2.29)$$

### Ejercicio 2.15

(Fuente: problema 1-1, página 33. Hines, Montgomery, et al. Probabilidad y Estadística para Ingeniería. Cuarta Edición. Editorial Patria). Después de ensamblados se realiza una inspección final en aparatos de televisión. Tres tipos de defectos se identifican como críticos, mayores y menores, los cuales se codifican como  $A$ ,  $B$  y  $C$  respectivamente, por una empresa de ventas por correo. Los datos se analizan obteniéndose los siguientes resultados:

Aparatos con defectos críticos únicamente	:	2%
“ “ “ mayores “	:	5%
“ “ “ menores “	:	7%
“ “ “ críticos y mayores	:	3%
“ “ “ “ “ menores	:	4%
“ “ “ mayores y menores	:	3%
“ “ los tres tipos de defectos	:	1%

a) ¿Qué porcentaje de los aparatos tiene algún tipo de defecto?

$$p(A \cup B \cup C) = p(A) + p(B) + p(C) - p(A \cap B) - p(A \cap C) - p(B \cap C) + p(A \cap B \cap C)$$

$$p(A \cup B \cup C) = 2 + 5 + 7 - 3 - 4 - 3 + 1 = 5\%$$

b) ¿Qué porcentaje de los aparatos no tiene defectos?

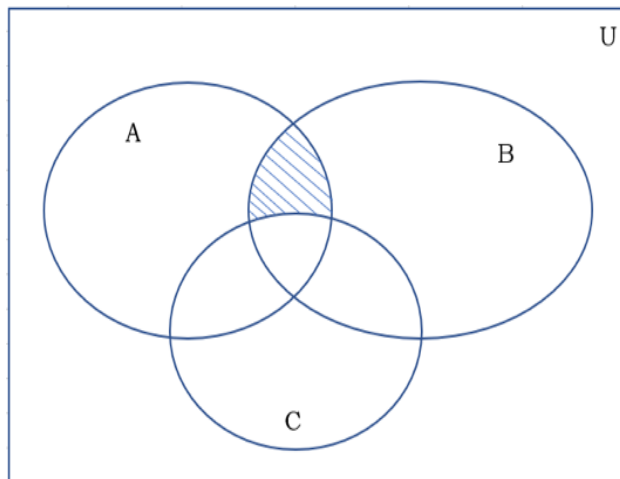
$$\% \text{ sin defectos} = 100\% - 5\% = 95\%$$

c) Los aparatos con defectos críticos o mayores deben desecharse. ¿Qué porcentaje corresponde con esa categoría?

$$p(A \cup B) = p(A) + p(B) - p(A \cap B) = 2\% + 5\% - 3\% = 4\%$$

d) ¿Qué porcentaje de los aparatos tiene defectos críticos y mayores pero no menores?

Para ilustrar lo que solicitan observe el diagrama de Venn de la figura 2.15.



**FIGURA 2.15.**  
Diagrama de Venn

$$p(A \cap B \cap C^c) = p(A \cap B) - p(A \cap B \cap C) = 3\% - 1\% = 2\%$$

## 2.5. Probabilidad Condicional, Probabilidad de Eventos Independientes, Probabilidad Total y Teorema de Bayes

En la baraja americana, como se muestra en la figura 2.11, se tiene un lote de  $N=52$  cartas, con dos colores (negro y rojo) y cuatro íconos (en inglés se denominan suit) diferentes: picas negras, corazones rojos, tréboles negros y diamantes rojos; existen 13 números de cada uno: 2, 3, 4, 5, 6, 7, 8, 9, 10, J (Jack), Q (Queen),  $k$  (King) y A (As=1), el orden en que se muestran es el de menor a mayor importancia individualmente.

Suponga que a un jugador en particular, se le asigna una carta aleatoriamente y posteriormente se le asigna una segunda carta, sin derecho a descartar ninguna de ellas. Observe que existen 52 cartas diferentes, de las cuales cuatro son ases. La probabilidad de que la primera carta que se le dio sea As, por probabilidad laplaciana, es  $p(x=As)=4/52$ ; la probabilidad de que la primera carta no sea As es su complemento  $p(x=no\ As)=48/52$ . Sin embargo, ¿qué sucede con la segunda carta? La probabilidad de que la segunda carta sea As, depende del resultado de la primera carta; es decir, si la primera carta fue as, la probabilidad de que la segunda carta también sea As es  $3/51$ ; si la primera carta no fue As, entonces la probabilidad de que la segunda sea As es  $4/51$ . A este tipo de probabilidades que están condicionadas a un resultado de un ensayo previo, se les denomina probabilidades condicionales.

Sea  $A$  un evento previo a un segundo evento  $B$ , la probabilidad condicional de  $B$  dado  $A$ , lo cual se escribe como  $p(B|A)$ , se define como:

$$p(B|A) = \frac{p(A \cap B)}{p(A)} \quad (2.30)$$

De la misma forma se define

$$p(A|B) = \frac{p(A \cap B)}{p(B)} \quad (2.31)$$

Nótese que no existe conmutatividad en las dos definiciones anteriores, son conceptos diferentes.

Una propiedad importante del concepto de probabilidad condicional, parte de despejar la probabilidad de la intersección de  $A$  y  $B$  en las ecuaciones 2.30 y 2.31 e igualar:

$$P(A \cap B) = P(B|A)P(A) = P(A|B)P(B) \quad (2.32)$$



Ejercicio 2.16

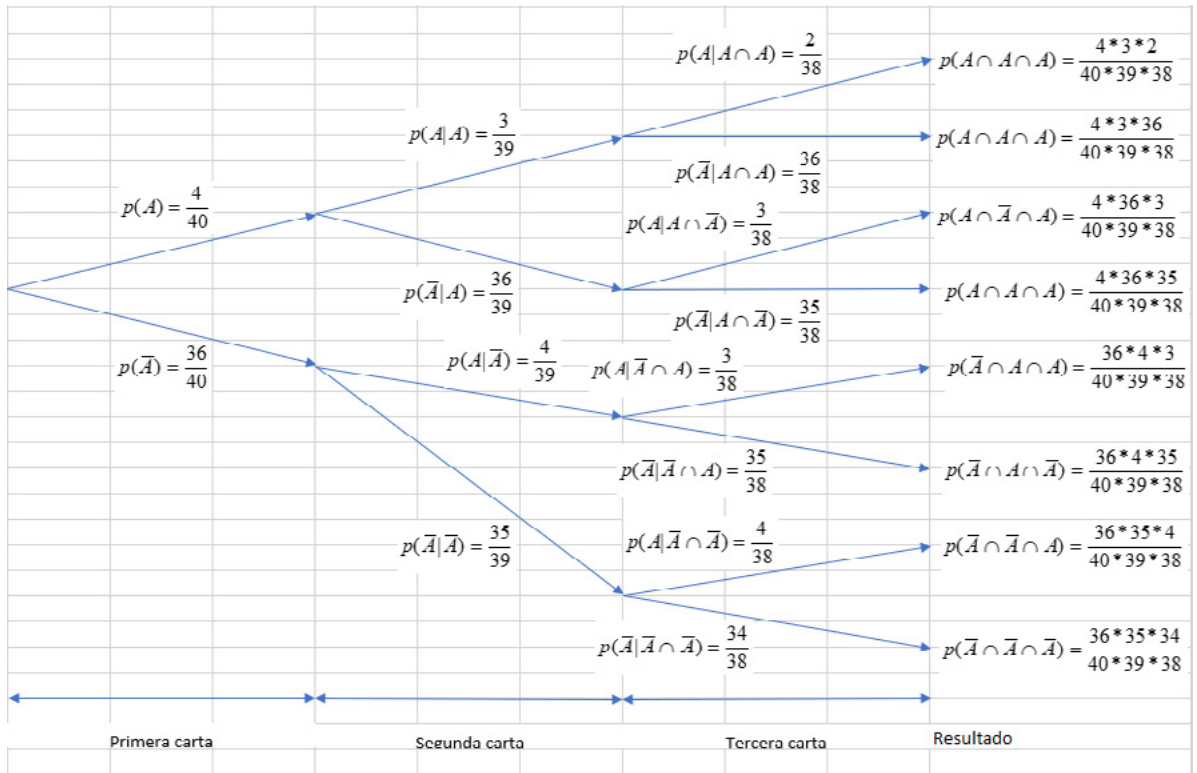
Sea un juego con baraja española, en la cual se le reparten aleatoriamente a cada jugador tres cartas.

La baraja española a diferencia de la baraja inglesa, no tiene los números 8, 9 y 10. Es decir, existen, cuatro íconos: espada, copa, oro y mazo y de cada una de ellas se tienen las figuras: 2, 3, 4, 5, 6, 7, sota, caballo, rey y as, con lo cual se tienen  $N=40$  cartas, de las cuales, cuatro son ases.

- a) Trace el diagrama de árbol de los eventos que le pueden ocurrir a una persona al realizar el juego.

Los posibles resultados se muestran en el diagrama de árbol de la figura 2.16.

FIGURA 2.16. Diagrama de árbol de posibles resultados



- b) Calcule la probabilidad de que la segunda carta sea as dado que la primera también fue as:

$$p(A|A) = (3/39)(4/40)/(4/40) = 3/39 = 1/13 = 0.076923$$

- c) Obtenga la probabilidad de que la primera y la segunda sean ases:

$$P(A \cap A) = p(A|A)p(A) = (3/39)(4/40) = (1/13)(1/10) = 1/130 = 0.007692$$

- d) Determine la probabilidad de que las tres cartas sean ases:

$$\begin{aligned} P(A \cap A \cap A) &= p(A|A \cap A)p(A \cap A) = p(A|A \cap A)p(A|A)p(A) \\ &= (2/38)(3/39)(4/40) = (1/19)(1/13)(1/10) = 1/2470 = 0.000405 \end{aligned}$$

- e) Calcule la probabilidad de obtener dos ases en tres cartas.

La probabilidad de cada rama se obtiene multiplicando la probabilidad de cada carta obtenida. Cada rama es mutuamente excluyente con cada una de las otras, por lo que la probabilidad de la unión de dos ramas se obtiene como la suma de las probabilidades de cada rama. Para este caso particular, dos ases se obtienen en los casos de tres ramas: segunda:  $A \cap A \cap A^{\bar{}}$ ; tercera:  $A \cap A^{\bar} \cap A$ ; quinta:  $A^{\bar} \cap A \cap A$ , de la figura 3.19.

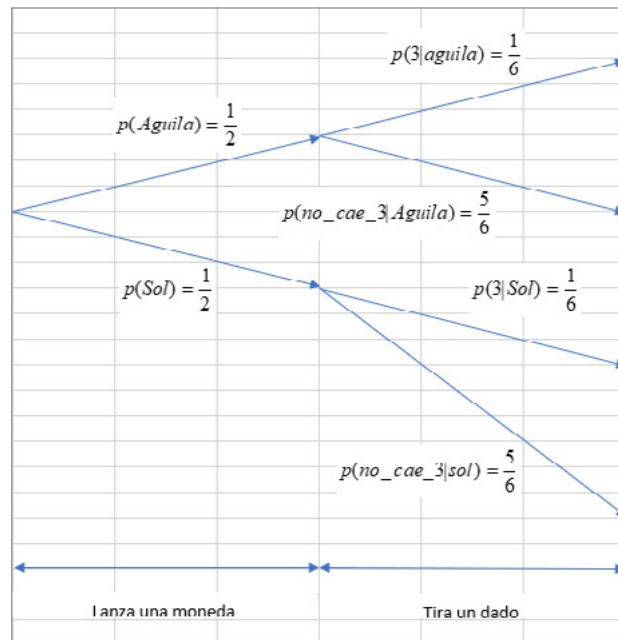
$$p(\text{dos ases en tres cartas}) = 3 * (4 * 3 * 36) / (40 * 39 * 38) = 0.021862$$

### Ejercicio 2.17

Suponga que se lanza una moneda homogénea y después se lanza un dado homogéneo de seis caras.

En la figura 2.17 se muestran los posibles resultados.

FIGURA 2.17.



a) Calcule la probabilidad de que caiga el tres dado que cayó el águila.

$$p(3|\text{águila}) = 1/6$$

b) Calcule la probabilidad de que no caiga el cinco dado que cayó el sol.

$$p(\text{no\_cae } 5|\text{sol}) = 5/6$$

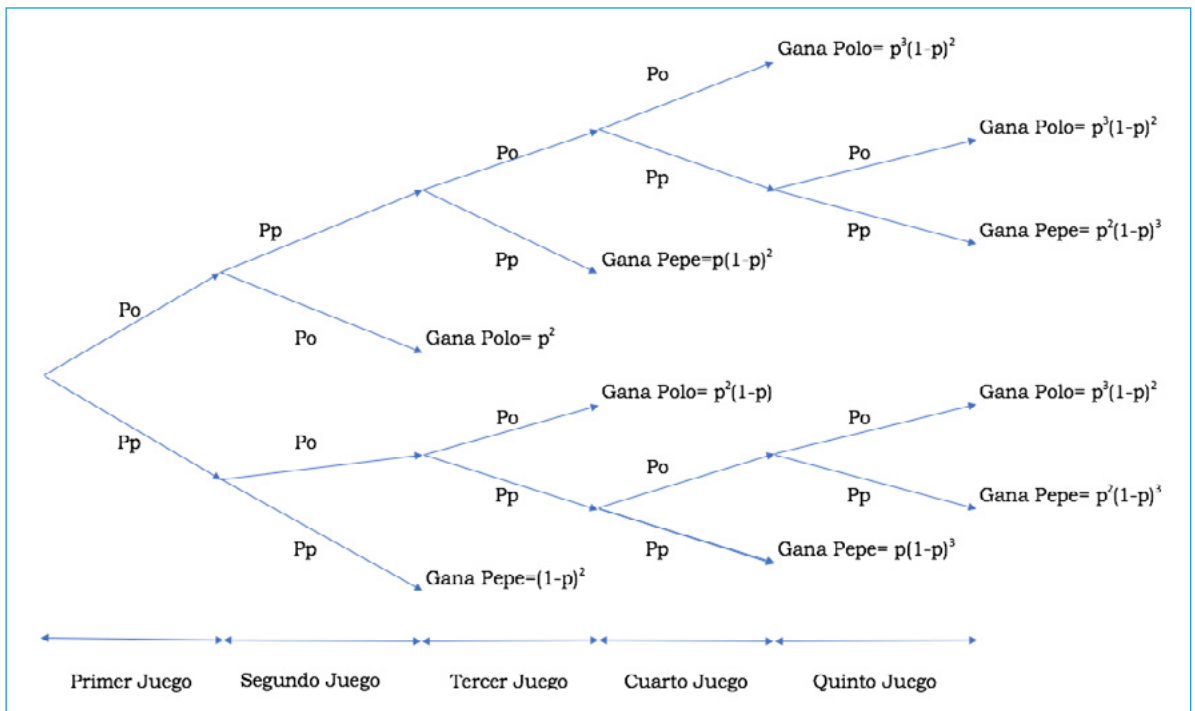
c)  $P(3 \cap \text{águila}) = (1/6)(1/2) = 1/12 = 0.083333$

### Ejercicio 2.18

En el ejercicio 2.10, suponga que la probabilidad de que Polo gane un partido de Ping-Pong es constante e igual a  $p$  y la probabilidad de que Pepe gane un partido es  $1-p$ . Los diferentes resultados se muestran en el diagrama de árbol de la figura 2.18

- Obtenga la probabilidad de que gane Polo.
- Obtenga la probabilidad de que gane Pepe.
- Determine quién tiene ventaja de ganar.

FIGURA 2.18



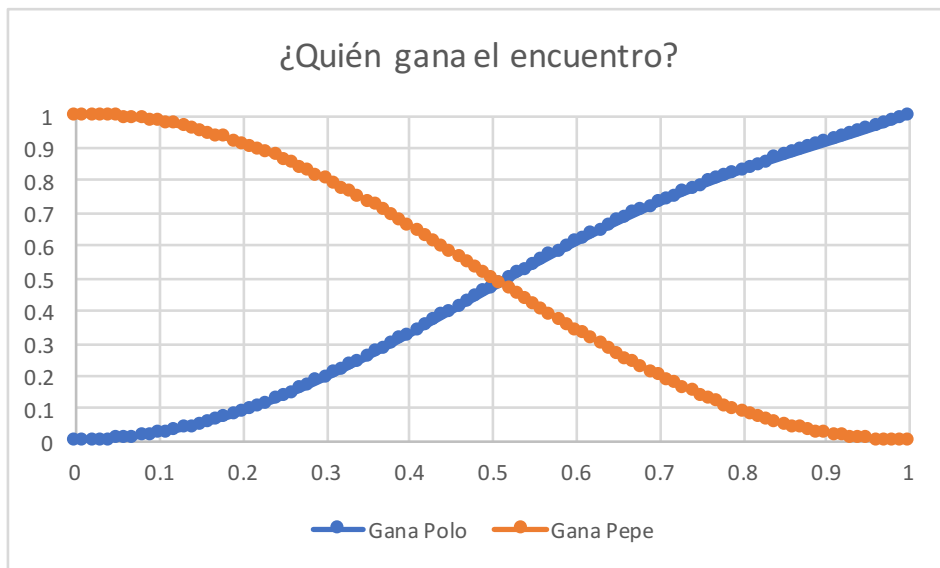
En este caso, nótese que  $p(\text{Polo gane un partido}) = p = \text{constante}$  y que este valor no depende de los ensayos previos, es decir, la probabilidad condicional es la misma de evento a evento.

$$P(\text{gane Polo}) = 3p^3(1-p)^2 + p^2(2-p)$$

$$P(\text{gane Pepe}) = p(2p+1)(1-p)^3 + (1+p)(1-p)^2$$

Si se grafica el valor de  $p$  contra quien gana el encuentro, se puede saber quién tiene mayor probabilidad de ganar, como se muestra en la figura 2.19.

FIGURA 2.19.



Nótese que si  $p < 0.5$  Pepe tiene ventaja sobre Polo. Si  $p > 0.5$  Polo tiene ventaja sobre Pepe y si  $p = 0.5$  tienen la misma probabilidad de ganar cada uno de ellos.

En los ejercicios 2.17 y 2.18 resueltos anteriormente surgió una propiedad de una importancia capital para la probabilidad y la estadística.

Se pudo apreciar en el caso del ejercicio 2.17 que lanzar un dado no depende en lo más mínimo de lanzar una moneda previamente. De la misma forma, en el ejercicio 2.18 se pudo apreciar que la probabilidad de que Polo gane un partido no depende de cómo le vaya en un partido previo.

Suponga que se realiza un evento  $A$  y posteriormente un evento  $B$ , si al calcular la probabilidad de  $B$  dado  $A$ ,  $P(B|A) = P(B)$  o en forma inversa  $P(A|B) = P(A)$ , se dice que los eventos  $A$  y  $B$  son estadísticamente independientes.

De otra forma, fácilmente se puede demostrar que dos eventos  $A$  y  $B$  son estadísticamente independientes, si se cumple que  $P(A \cap B) = P(A)P(B)$ .

La confiabilidad se define como la probabilidad, o el grado de certeza que se tiene, de que cierto producto o servicio se desempeñe adecuadamente, bajo ciertas condiciones dadas durante un cierto período de tiempo.

Ahora se ilustrará cómo se calcula la confiabilidad de un sistema en serie y de un sistema en paralelo. Para ello, se usará el concepto de caja negra. Suponga que un sistema está constituido por dos componentes  $A$  y  $B$  en serie, como se ilustra en la figura 2.20. Suponga que la confiabilidad del componente  $A$  es  $p_A$  y la confiabilidad del componente  $B$  es  $p_B$ .

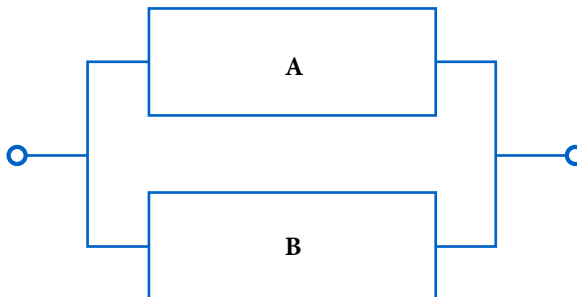
**FIGURA 2.20.** Sistema en serie



Para que este sistema opere, deben funcionar  $A$  y  $B$  simultáneamente, por lo que la confiabilidad del sistema en serie está dada como la intersección de las confiabilidades de cada una de sus componentes, y como se supone que la confiabilidad de una de sus componentes no depende de la confiabilidad de la otra, se puede suponer que existe independencia estadística entre ambas. En este caso:

$$P(A \cap B) = p(B|A)p(A) = p(B)p(A) = p_A p_B \quad (2.33)$$

Por otra parte, suponga que un sistema está constituido por dos componentes  $A$  y  $B$  en paralelo, como se ilustra en la figura 2.21.



**FIGURA 2.21.**  
Sistema en paralelo

Para que este sistema se desempeñe adecuadamente, basta que funcione una sola de las componentes, ya sea A o B, por lo que la confiabilidad del sistema en paralelo está dada como la unión de las confiabilidades de cada una de sus componentes, y como se supone que la confiabilidad de una de sus componentes no depende de la confiabilidad de la otra, se puede suponer que existe independencia estadística entre ambas. En este caso:

$$P(A \cup B) = p(A) + p(B) - P(A \cap B) = p_A + p_B - p_A p_B \quad (2.34)$$

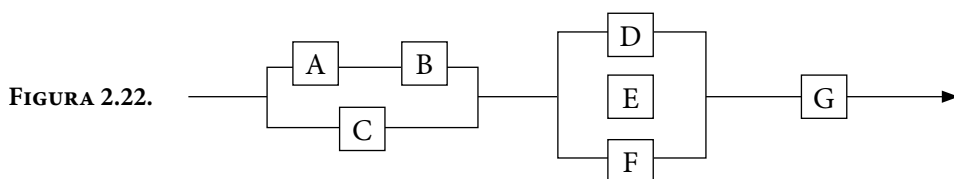
Para ilustrar el significado de constituir un sistema en serie o en paralelo, suponga que la confiabilidad del componente A es 0.9 y la confiabilidad del componente B es 0.95. Entonces, la confiabilidad del sistema en serie es  $(0.9)(0.95) = 0.855$ ; nótese que, en un sistema en serie, la confiabilidad del sistema completo siempre es menor a la confiabilidad de cada una de sus componentes. En el caso de la confiabilidad de un sistema en paralelo, esta es  $(0.9) + (0.95) - (0.9)(0.95) = 0.995$ ; obsérvese que, la confiabilidad del sistema en paralelo completo siempre es mayor que la confiabilidad de cada una de sus componentes.

De lo anterior se desprende que es mejor desde el punto de vista de la confiabilidad de un sistema colocar sistemas en paralelo que sistemas en serie. Sin embargo, no siempre es posible colocar todas las componentes en paralelo. Por ejemplo, en un sistema productivo, su proceso productivo está constituido por un conjunto de etapas de proceso consecutivas, es decir, en serie. En algunas etapas, que se consideren críticas o muy lentas, sí sería conveniente colocar otra componente similar en paralelo, lo cual duplicaría la velocidad de esa etapa y al mismo tiempo elevaría su confiabilidad, pero el costo de comprar otro equipo, para ponerlo en paralelo, podría ser incosteable.

Se procederá a calcular la confiabilidad de un sistema más complejo, aprovechando los conceptos de confiabilidad en serie y confiabilidad en paralelo.

### Ejercicio 2.19

Un sistema tiene tres subsistemas I, II y III, como se muestra en la figura 2.22. Los tres subsistemas, así como sus componentes, operan independientemente uno del otro y la confiabilidad de cada componente A, B, C, D, E, F y G es 0.95.



- a) Obtenga la confiabilidad de cada subsistema y la confiabilidad del sistema completo.

Las componentes  $A$  y  $B$  están en serie y las componentes  $A$  y  $B$  están en paralelo con  $C$ , por lo que la confiabilidad del primer subsistema es:

$$p_I = p[(A \cap B) \cup C] = p_A p_B + p_C - p_A p_B p_C = (0.95)^2 + 0.95 - (0.95)^3 = 0.995125$$

Las componentes  $D$ ,  $E$  y  $F$  del subsistema II se encuentran en paralelo, por lo que su confiabilidad se puede calcular de dos formas diferentes:

$$\begin{aligned} p_{II} &= p(D \cup E \cup F) = p_D + p_E + p_F - p_D p_E - p_D p_F - p_E p_F + p_D p_E p_F \\ &= 3(0.95) - 3(0.95)^2 + (0.95)^3 = 0.999875 \end{aligned}$$

El subsistema III solo tiene una sola componente, por lo que su confiabilidad es la confiabilidad de la componente, que en este caso es  $p_{III} = 0.95$ .

Nótese que los subsistemas I, II y III están en serie, por lo que la confiabilidad del sistema completo es igual al producto de las confiabilidades de los subsistemas, es decir

$$p_{sist} = p(I \cap II \cap III) = (0.995125)(0.999875)(0.95) = 0.945251$$

- b) ¿Qué componentes cambiaría para mejorar la confiabilidad del sistema completo?

Como se puede apreciar en el inciso anterior, el subsistema con menor confiabilidad es el tercero, por lo que conviene cambiarlo por un componente de mayor confiabilidad, por ejemplo, uno de 0.99 de confiabilidad; de esta forma, la confiabilidad del sistema completo sería

$$p_{sist} = p(I \cap II \cap III) = (0.995125)(0.999875)(0.99) = 0.985051$$

Si se requiere aumentar aún más la confiabilidad, convendría cambiar la componente  $C$ , observe que ocurre si se cambia por una con una confiabilidad de 0.99.

$$\begin{aligned} p_I &= p[(A \cap B) \cup C] = p_A p_B + p_C - p_A p_B p_C = (0.95)^2 + 0.99 - (0.9)^2(0.99) = 0.999025 \\ p_{sist} &= p(I \cap II \cap III) = (0.999025)(0.999875)(0.99) = 0.988911 \end{aligned}$$

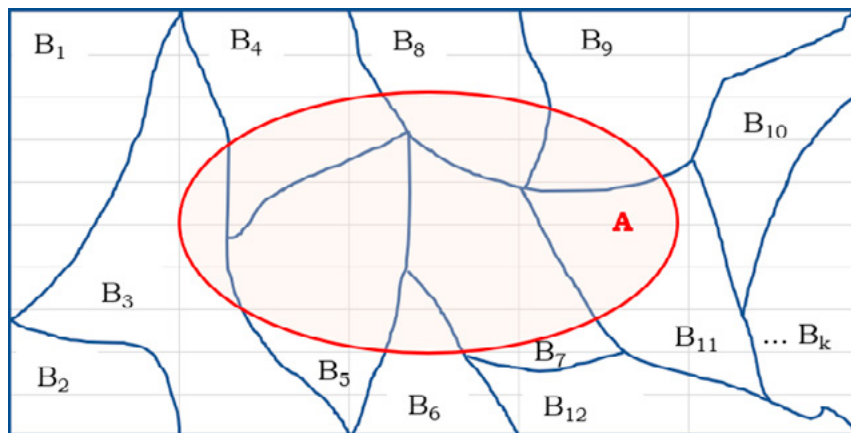
## Teorema o Ley de Probabilidad Total

Suponga que se tiene un conjunto  $B$ , dividido o particionado en una serie de subconjuntos  $B_1, B_2, B_3, \dots, B_k$ , tales que cada uno de ellos es mutuamente excluyente con respecto a los otros, es decir,  $B_i \cap B_j = \emptyset \forall i \neq j$ , de tal forma que son colectivamente exhaustivos, es decir  $B_1 \cup B_2 \cup B_3 \cup \dots \cup B_k = B$ . A esta forma de dividir un conjunto se le llama una PARTICIÓN, como se aprecia en la figura 2.23. En este caso, se representará al conjunto  $B$  como si se tratara del espacio total de eventos. El símbolo  $\forall$  es un símbolo matemático que significa “para todo”. Suponga que  $A \subseteq B$ , como se muestra en la figura 2.23.

Obsérvese que:

$$A = (A \cap B_1) \cup (A \cap B_2) \cup (A \cap B_3) \cup \dots \cup (A \cap B_k)$$

FIGURA 2.23. Partición de un conjunto  $B$



Dado que se trata de eventos mutuamente excluyentes

$$p(A) = p(A \cap B_1) + p(A \cap B_2) + p(A \cap B_3) + \dots + p(A \cap B_k)$$

Pero, por definición de probabilidad condicional,  $p(A \cap B_i) = p(A|B_i)p(B_i)$ , de tal forma que:

$$p(A) = p(A|B_1)p(B_1) + p(A|B_2)p(B_2) + p(A|B_3)p(B_3) + \dots + p(A|B_k)p(B_k) \quad (2.35)$$

A este concepto se le conoce como *Teorema o Ley de Probabilidad Total*.



Por otra parte, de la definición de probabilidad condicional

$$p(A \cap B_i) = p(A|B_i)p(B_i) = p(B_i|A)p(A) \text{ lo que implica que:}$$

$$p(A \cap B_i) = p(B_i|A)p(A) = p(A|B_i)p(B_i)$$

$$p(B_i|A) = \frac{p(A|B_i)p(B_i)}{p(A)}$$

En la expresión anterior, si se sustituye la expresión 2.35, se obtiene lo que se denomina el *Teorema de Bayes*:

$$p(B_i|A) = \frac{p(A|B_i)p(B_i)}{\sum_{j=1}^{j=k} p(A|B_j)p(B_j)}$$

$$p(B_i|A) = \frac{p(A|B_i)p(B_i)}{p(A|B_1)p(B_1) + p(A|B_2)p(B_2) + \dots + p(A|B_k)p(B_k)} \quad (2.36)$$

El Teorema de Bayes, es una proposición planteada por el filósofo inglés Thomas Bayes (1702-1761), que vincula un conjunto de probabilidades a priori (causas), que pueden ser estimaciones basadas en hechos históricos previos o suposiciones basadas en la escuela subjetivista con probabilidades a posteriori (efectos), cuando se posee nueva información, que permiten la comprobación de las probabilidades a priori. No es fácil de asimilar pero con varios ejemplos se ilustrará su utilidad.



**FIGURA 2.24.** Thomas Bayes (1702-1761)

**NOTA:** Thomas Bayes. (2020). [https://es.wikipedia.org/wiki/Thomas\\_Bayes](https://es.wikipedia.org/wiki/Thomas_Bayes)

## Ejercicio 2.20

En el ejercicio 2.14 se dijo que en un juego de futbol americano, los Cóndores de la UNAM tienen tiempo para una jugada más y necesitan anotar en esa jugada para ganar el partido. El entrenador del equipo puma siente que hay solamente tres jugadas que se pueden usar: un pase largo, un pase pantalla o una corrida. La probabilidad de que el entrenador mande un pase largo es 0.35; de que mande un pase pantalla, 0.25; y de que mande una corrida, 0.40. Basado en su experiencia pasada sobre situaciones similares, el entrenador siente que las probabilidades de lograr una anotación con tales jugadas son: pase largo 0.5, pase pantalla 0.3 y corrida 0.1.

a) ¿Qué probabilidad de anotar tienen los Cóndores?

Sean:

$A$ : Anotar en la siguiente jugada

$B_1$ : Pase largo

$B_2$ : Pase pantalla

$B_3$ : Corrida

$$p(B_1) = 0.35$$

$$p(B_2) = 0.25$$

$$p(B_3) = 0.40$$

$$p(A|B_1) = 0.5$$

$$p(A|B_2) = 0.3$$

$$p(A|B_3) = 0.1$$

Por el Teorema de la Probabilidad Total:

$$P(A) = p(A|B_1)p(B_1) + p(A|B_2)p(B_2) + p(A|B_3)p(B_3)$$

$$P(A) = 0.5 * 0.35 + 0.3 * 0.25 + 0.1 * 0.40 = 0.29$$

b) Dado que anotaron, ¿cuál es la probabilidad de que haya sido por una corrida?

$$p(B_3|A) = \frac{p(A|B_3)p(B_3)}{P(A)} = \frac{(0.1) * (0.40)}{0.29} = 0.137931$$

c) Dado que anotaron, ¿cuál fue la jugada más probable que realizaron?

$$p(B_1|A) = \frac{p(A|B_1)p(B_1)}{P(A)} = \frac{(0.5) * (0.35)}{0.29} = 0.603448$$

$$p(B_2|A) = \frac{p(A|B_2)p(B_2)}{P(A)} = \frac{(0.3) * (0.25)}{0.29} = 0.258621$$

$$p(B_3|A) = \frac{p(A|B_3)p(B_3)}{P(A)} = \frac{(0.1) * (0.40)}{0.29} = 0.137931$$

Lo más probable fue que hayan mandado un pase largo.

### Ejercicio 2.21

El narcotráfico en Chacaltianguis Veracruz tiene tres medios para transportar precursores para droga: por tierra, por río o por aire. Por experiencias previas se ha estimado que la mitad del material es transportado por tierra; 30%, por río y el resto, por aire. La transportación por tierra puede ser 60% por brecha y el resto por autopista. La probabilidad de librar los retenes que la Policía Federal ha impuesto a lo largo de los diversos trayectos es 10% por autopista, 5% por brecha, 6% por mar y 2% por aire.

En la figura 2.25 se muestran las posibles opciones:

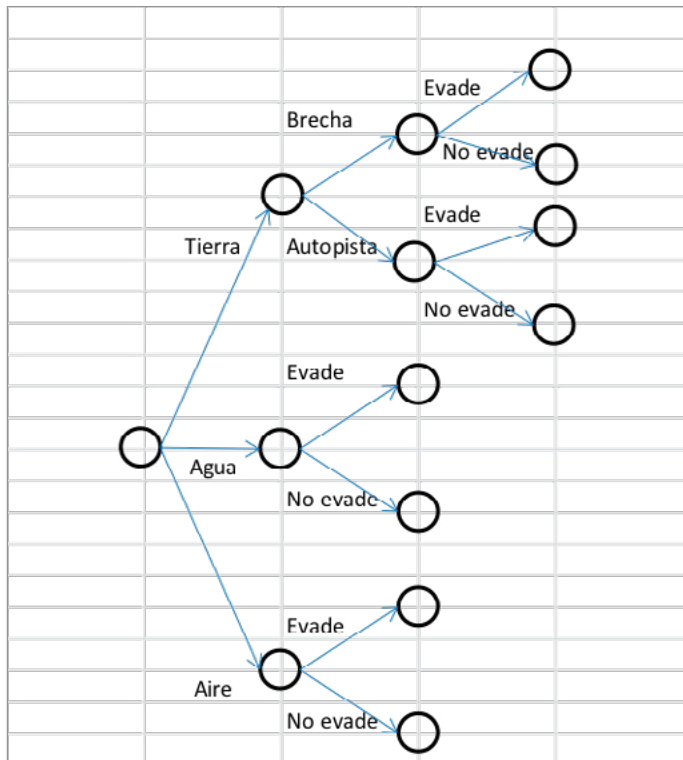


FIGURA 2.25.

a) ¿Qué porcentaje de los precursores logrará evadir la acción de la justicia?

Sea

$E$ : Evadir la acción de la justicia

$B|T$ : Transporte terrestre por Brecha

$Au|T$ : Transporte Terrestre por Autopista

$R$ : Transporte por Río

$Ai$ : Aire

$$\begin{aligned} P(E) &= P(E|(B|T)) P(B|T)P(T) + P(E|(Au|T)) P(Au|T)P(T) + P(E|R)P(R) + P(E|Ai)P(Ai) \\ &= (0.05)(0.60)(0.50) + (0.10)(0.40)(0.50) + (0.06)(0.30) + (0.02)(0.20) \\ &= 0.057 \end{aligned}$$

b) Si los precursores no fueron interceptados, ¿cuál es la probabilidad de que hayan sido transportados por río?

$$P(R|E) = \frac{P(E|R) P(R)}{P(E)} = \frac{(0.06)(0.30)}{(0.057)} = 0.3158$$

## Ejercicios propuestos del capítulo 2

- 2.1 En las siguientes características determine cuál es la escala de medición más adecuada para cada una de ellas.
- |                             |                                    |
|-----------------------------|------------------------------------|
| a) Placas de un coche.      | g) Temperatura.                    |
| b) Número de cuenta.        | h) Dureza de un metal.             |
| c) Apariencia de un objeto. | i) Resistencia mecánica            |
| d) Color de piel.           | al torque de un perno.             |
| e) Tipo de sangre.          | j) Luminosidad del amanecer.       |
| f) Sabor.                   | k) Número de defectos en un coche. |
- 2.2. Explique y ejemplifique las diferencias que existen entre universo, población, muestra y unidad muestral.
- 2.3. Explique con sus propias palabras qué es la investigación.
- 2.4. Elabore un cuadro comparativo entre los métodos de investigación de Russell L. Ackoff, Checkland y Ezequiel Ander-Egg.
- 2.5. Elabore un glosario de términos básicos en Probabilidad y Estadística.
- 2.6. Describa detalladamente qué diferencias, similitudes y complementariedades existen entre la Probabilidad y la Estadística.
- 2.7. Explique la relación que existe entre la probabilidad, la estadística, el método científico, el papel de la probabilidad y la estadística en la investigación básica y aplicada, y sus etapas.
- 2.8. Sea el universo el conjunto de números naturales hasta el número 20, incluyendo al cero. Sean  $A = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$ ,  $B = \{1, 3, 5, 7, 9\}$ ,  $C = \{\text{números divisibles entre 3 del universo}\}$ ,  $D = \{\text{números primos en el universo}\}$ .

Realice las siguientes operaciones e ilústrelas con diagramas de Ven:

- |               |  |
|---------------|--|
| a) $A \cup B$ | f) $(A \cup B) - (A \cap B)$                   |
| b) $A \cap B$ | g) Compruebe que $(C \cup D)^c = C^c \cap D^c$ |
| c) $D^c$      | h) Compruebe que $(C \cap D)^c = C^c \cup D^c$ |
| d) $C \cup D$ | i) $(D - A)^c$                                 |
| e) $C \cap D$ | j) $U - (D - A)^c$                             |

- 2.9. Demuestre las Leyes de De Morgan
- 2.10. Si  $A$ ,  $B$  y  $C$  son eventos mutuamente excluyentes,  $P(A) = 0.2$ ,  $P(B) = 0.3$  y  $P(C) = 0.2$ , calcule:
- $P(A \cup B \cup C)$
  - $P[A' \cap (B \cup C)]$
  - $P(B \cup C)'$
- 2.11. En una generación de 100 estudiantes de nivel medio superior, 54 estudian matemáticas, 69 historia y 35 ambas materias, si se elige al azar uno de estos estudiantes, encuentre la probabilidad de que:
- Estudie matemáticas o historia.
  - No estudie ninguna de esas materias.
  - Estudie historia, pero no matemáticas.
- 2.12. En el último año de preparatoria de un grupo de 100 estudiantes, 42 estudiaron matemáticas, 68 psicología, 54 historia, 22 matemáticas e historia, 25 matemáticas y psicología, 7 estudiaron historia pero no estudiaron matemáticas ni psicología, 10 estudiaron las tres materias y 8 no estudiaron ninguna de las tres. Si se elige al azar a un estudiante, determine la probabilidad de que:
- Una persona inscrita en psicología estudie las tres materias.
  - Una persona que no estudia psicología esté tomando tanto historia como matemáticas.
- 2.13. Si cada uno de los artículos incluidos en un catálogo comienza con tres letras distintas, seguidas de 4 dígitos distintos y diferentes de cero, encuentre la probabilidad de elegir al azar un artículo que tenga como primera letra una vocal y como último dígito un número par.

- 2.14. Se construye un dado de manera que el 1 y el 2 ocurran con el doble de la frecuencia que se presenta el 5, el cual ocurre con una frecuencia 3 veces superior al 3, al 4, o al 6. Si se lanza el dado una sola vez, obtenga la probabilidad de que:
- El número sea par.
  - El número sea un cuadrado perfecto.
  - El número sea mayor que 4.
- 2.15. Se extraen dos cartas sucesivas de un mazo de baraja americana, sin reemplazo, ¿cuál es la probabilidad de que ambas cartas sean mayores que 2 y menores que 8?
- 2.16. Supóngase que, entre seis pernos, dos son más cortos que una longitud específica. Si se esconden dos pernos al azar, ¿cuál es la probabilidad de que los dos más cortos sean escogidos?
- 2.17. Durante un examen final el profesor sospecha que un alumno ha copiado por lo que procede a anular su examen, el alumno protesta y entonces el profesor le propone lo siguiente: “Vamos a lanzar una moneda (justa) cuatro veces, si se obtienen cuatro águilas o cuatro soles, te anulo el examen”. El alumno vuelve a protestar y propone mejor lo siguiente: “Vamos a lanzar una moneda (justa) una vez, si se obtiene un sol me anula el examen”. ¿Considera usted que el alumno estaba copiando?, ¿Sí?, ¿No?, ¿Por qué?
- 2.18. Un examen de opción múltiple consta de 10 preguntas, en cada pregunta aparecen cuatro posibles respuestas de las cuales solo una es la correcta. Considerando que la persona que presenta este examen no sabe cuál es la respuesta a ninguna de las diez preguntas, calcular la probabilidad de que:
- Responda incorrectamente todas las preguntas.
  - Apruebe el examen, sabiendo que para ello se necesita responder correctamente por lo menos ocho de las diez preguntas.
- 2.19. En cierta prisión,  $\frac{3}{5}$  de los internos son menores de 25 años,  $\frac{3}{5}$  son hombres y  $\frac{5}{8}$  son mujeres o de 25 o más años, ¿Cuál es la probabilidad de que un interno elegido al azar sea una mujer de cuando menos 25 años?
- 2.20. Una bolsa A contiene 2 botellas de tabletas de aspirina y 3 botellas de tabletas para la tiroides. Una segunda bolsa B contiene 3 botellas de tabletas

de aspirina, 2 botellas de tabletas para la tiroides y una botella de tabletas laxantes. Si se extrae al azar una botella de tabletas de cada bolsa, encuentre la posibilidad de que:

- a) Ambas botellas contengan tabletas para la tiroides.
- b) Ninguna botella contenga tabletas para la tiroides.
- c) Las dos botellas contengan tabletas diferentes.

2.21. En un proyecto de construcción de edificios, para finalizar la obra se requiere la terminación sucesiva de una serie de actividades. Sean

$E$  = Excavación completa a tiempo y  $P(E) = 0.8$

$F$  = Cimentación terminada a tiempo y  $P(F) = 0.7$

$S$  = Super estructura terminada a tiempo y  $P(S) = 0.9$

Asuma independencia estadística entre estos eventos.

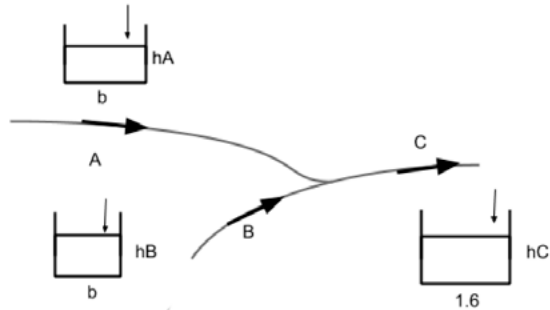
- a) Defina el evento (proyecto terminado a tiempo) en términos de  $E, F, S$ .
- b) Calcule la probabilidad de terminación a tiempo.
- c) Defina en términos de  $E, F, S$  y sus complementos el siguiente evento.  
 $G$ : la excavación estará a tiempo y al menos 1 de las otras 2 operaciones no estará a tiempo.
- d) Defina el evento  $H$ : solo una de las 3 operaciones estará terminada a tiempo.

2.22. Un contratista está sometiendo a prueba dos trabajos  $A$  y  $B$ . La probabilidad de que gane el trabajo  $A$  es  $P(A) = \frac{1}{4}$ , y de que gane el trabajo  $B$  es de  $P(B) = \frac{1}{3}$ .

- a) Asumiendo que ganar el trabajo  $A$  y ganar el trabajo  $B$  son eventos independientes ¿cuál es la probabilidad de que el contratista obtendrá al menos un trabajo?
- b) ¿Cuál es la probabilidad de que el contratista obtenga el trabajo  $A$ , si ganó al menos un trabajo?
- c) Si también sometió a prueba un trabajo  $C$ , con probabilidad de ganarlo de  $P(C) = \frac{1}{4}$ , ¿Cuál es la probabilidad de que obtenga al menos un trabajo?
- d) Asuma otra vez que existe independencia estadística entre los eventos  $A, B$  y  $C$ .
- e) ¿Cuál es la probabilidad de que el contratista no obtenga ningún trabajo?



- 2.23. El cruce de secciones de los ríos A, B y C son mostrados en la figura y el nivel de desbordamiento de A y B sobre los niveles de flujo medio, son como sigue.



Nivel de desbordamiento en A		Nivel de desbordamiento en B	
ft	Probabilidad	ft	Probabilidad
0	0.25	0	0.20
2	0.25	2	0.20
4	0.25	4	0.20
6	0.25	6	0.20
		8	0.20

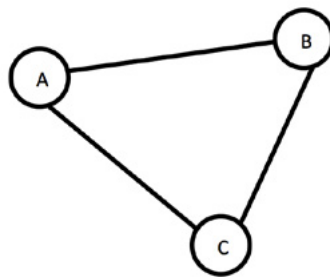
Asuma que las velocidades de flujo de A, B y C son las mismas. ¿Cuál es la probabilidad de que el desbordamiento en C será superior a 6 ft sobre el nivel medio?

- 2.24. Un par de dados se tiran una vez. Sea "A" el evento en el que el primer dado muestra un 1, "B" el evento en el que el segundo dado muestra un 6 y "C" el evento en el que la suma de ambos dados es igual a 7. ¿Los eventos A, B y C son independientes?
- 2.25. Una prueba de opción múltiple contiene 25 preguntas cinco opciones cada una. ¿Cuál es la probabilidad de que al contestar de manera aleatoria cada pregunta, más de la mitad de las respuestas sean incorrectas?
- 2.26. Dibuje cinco círculos concéntricos de radio  $Kr$ , donde  $K=1,2,3,4,5$  respectivamente. Sombree el círculo de radio  $r$  y dos anillos con radio exterior correspondientes  $3r$  y  $5r$ . Entonces seleccione aleatoriamente un punto en el círculo de radio  $5r$ .

Obtenga la probabilidad de que tal punto este en:

- a) El círculo de radio  $2r$ .
- b) La región sombreada.

- 2.27. En un cierto examen el estudiante recibe  $n$  preguntas y  $n$  respuestas que deben ser asignadas a las preguntas. Si el estudiante las asigna aleatoriamente, encontrar:
- a) La probabilidad de que la primera pregunta sea correctamente respondida.
  - b) La probabilidad de que ninguna sea respondida correctamente.
  - c) A qué valor tiende la respuesta del inciso b) si  $n$  es un número muy grande.
- 2.28. Durante el primer año de uso de un amplificador de radio, se pueden requerir tres tipos de reparaciones y las probabilidades correspondientes son 0.05, 0.04 y 0.02. ¿Cuál es la probabilidad de que un amplificador, seleccionado al azar, requiera reparación durante su primer año de uso? Suponga cada tipo de reparación independiente de las otras dos.
- 2.29. Se sabe que el 20% de los relojes producidos por una compañía son defectuosos. Calcular la probabilidad de que en cuatro relojes tomados al azar:
- a) Ninguno resulte defectuoso
  - b) Más de dos sean defectuosos
- 2.30. El sistema de supercarreteras entre las ciudades  $A$ ,  $B$  y  $C$  se muestra en la figura siguiente. Viajar entre  $A$ ,  $B$  y  $C$  durante los meses de invierno no siempre es posible porque algunas partes de la súper carretera pueden no estar abiertas al tráfico, debido a las extremas condiciones del tiempo.



Sea  $E_1, E_2, E_3$ , eventos que denotan que las supercarreteras  $AB, AC$  y  $CB$  están abiertas respetivamente. En un día cualquiera se asume que:

$$P(E_1) = 2/5$$

$$P(E_3/E_2) = 4/5$$

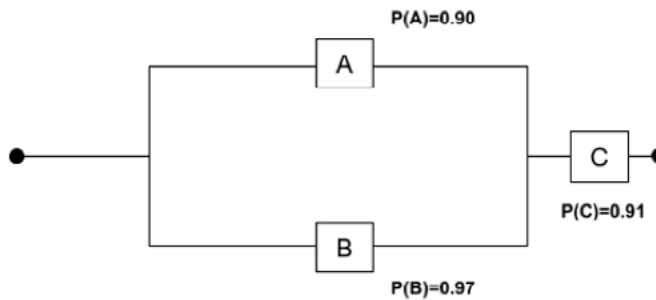
$$P(E_2) = 2/4$$

$$P(E_1/E_2E_3) = 1/2$$

$$P(E_3) = 2/3$$

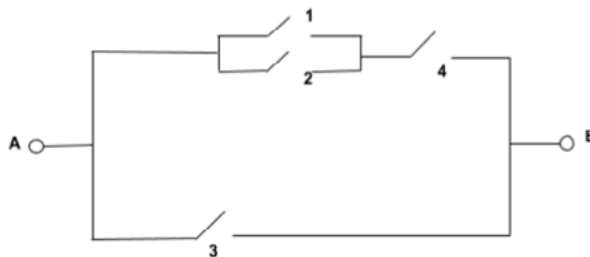
- ¿Cuál es la probabilidad de que un viajero pueda hacer un viaje de  $A$  a  $B$  si tiene que pasar por la ciudad  $C$ ?
- ¿Cuál es la probabilidad de que el viajero pueda llegar a la ciudad  $B$ ?
- ¿Cuál ruta tomaría primero para maximizar su elección para llegar a  $B$ ?

2.31. Un sistema contiene tres componentes que se encuentran conectados entre sí, como se muestra en la siguiente figura:

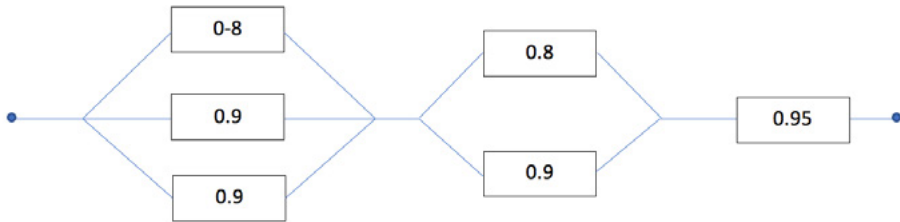


Las probabilidades indican la seguridad de que la componente funcione adecuadamente. Si se supone que el funcionamiento de una componente en particular es independiente del de los demás, determinar la probabilidad de que el sistema trabaje adecuadamente.

2.32. En la red eléctrica de la siguiente figura los cuatro interruptores son cerrados en un cierto instante, si es igualmente probable que un interruptor trabaje o no, ¿cuál es la probabilidad de que un camino cerrado exista de izquierda a derecha?



- 2.33. Considere el diagrama de un sistema electrónico en el que se muestran las probabilidades de que los componentes del sistema funcionen adecuadamente. ¿Cuál es la probabilidad de que todo el sistema opere si el ensamble III y al menos unos de los componentes de los ensambles I y II deben operar para que funcione el ensamble? Considere que los componentes de cada ensamble operan en forma independiente y que los ensambles operan independientes unos de otros.



- 2.34. Efectúe lo que se pide:
- Demuestre que  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - ¿Cuáles son las hipótesis en el teorema de Bayes?
  - ¿Cuáles son los axiomas fundamentales de la probabilidad?
- 2.35. Supóngase que se puede llevar un automóvil a reparar a uno de dos talleres que están cerca de la oficina; en uno la probabilidad de que lo entreguen a tiempo es de  $\frac{2}{3}$ , y en el otro, es de  $\frac{1}{3}$ . Si se eligió al azar uno de los talleres y entregaron a tiempo el automóvil, ¿cuál es la probabilidad de que el taller elegido haya sido el segundo?
- 2.36. Dos proveedores,  $A$  y  $B$ , suministran una misma pieza a un fabricante. Los antecedentes demuestran que el 5% de las piezas entregadas por  $A$  estaban defectuosas, y el 9% de las piezas entregadas por  $B$  también estaban defectuosas. Además,  $A$  entrega 4 veces más piezas que  $B$ . Si se extrae al azar una pieza y se observa que no está defectuosa, ¿Cuál es la probabilidad de que la haya fabricado  $A$ ?
- 2.37. En cierta región del país se sabe, por experiencias pasadas, que la probabilidad de elegir a un adulto de más de 40 años con cáncer es de 0.02. Si la probabilidad de que un doctor diagnostique en forma correcta que una persona tiene cáncer es de 0.78 y la probabilidad de que diagnostique de manera incorrecta que una persona tiene cáncer cuando de hecho no lo tiene es de 0.06

- a) ¿Cuál es la probabilidad de que diagnostique que una persona tiene cáncer?
- b) ¿Cuál es la probabilidad de que una persona a la que se le ha diagnosticado cáncer esté en realidad enfermo?
- 2.38. Un ingeniero debe seleccionar un proyecto para ejecutarlo de entre dos posibles: el  $A$  y el  $B$ . La probabilidad de que el proyecto  $A$  sea redituable es 0.7 y que lo sea el  $B$  es 0.6. El ingeniero por su experiencia se decide en el 55% de los casos por los proyectos del tipo  $A$ . Bajo estas condiciones, el ingeniero se decide por uno de los proyectos y no resulta redituable. ¿Cuál es la probabilidad de que el proyecto seleccionado sea del tipo  $A$ ?
- 2.39. La probabilidad de que un hombre casado vea cierto programa de televisión es de 0.4; la probabilidad de que una mujer casada vea ese programa es de 0.5; y la probabilidad de que un hombre vea el programa, dado que su mujer lo ve, es de 0.7. Encuentre la probabilidad de que:
- a) Una pareja casada vea el programa
- b) Una esposa vea el programa, dado que su esposo lo ve
- c) Cuando menos una persona de una pareja de casados vea el programa
- 2.40. En un almacén se encuentran 80 cajas con 100 fusibles cada una. Veinte cajas contienen fusibles producidos por la máquina  $A$ , 30 cajas contienen fusibles producidos por la máquina  $B$  y las últimas 30 tienen fusibles producidos por la máquina  $C$ . Las cajas están almacenadas al azar, sin que importe la máquina de procedencia. La máquina  $A$  produce 5% de fusibles defectuosos; la máquina  $B$ , 3% y la máquina  $C$ , 2%. Si se selecciona una de estas cajas al azar, se toma uno de sus fusibles y se encuentra que es defectuoso, ¿cuál es la probabilidad de que haya sido producido por la máquina  $B$ ?
- 2.41. Tres máquinas  $A1$ ,  $A2$  y  $A3$  producen respectivamente la tercera parte de la producción total de remaches de una fábrica. El 4% de los remaches producidos por la máquina  $A1$  tiene algún defecto, el 2% de la producción de la máquina  $A2$  también tiene algún defecto y el 3% de la producción de  $A3$  es también defectuosa. Se seleccionó un remache al azar que resultó estar defectuoso. Hallar la probabilidad de que el remache seleccionado haya sido producido por la máquina  $A3$ .

- 2.42. Un hombre ebrio tiene  $n$  llaves en la mano y ensaya abrir la puerta de su casa probando de una en una aleatoriamente sin descartar las que no necesita (tales ensayos son aleatorios e independientes).
- ¿Cuál es la probabilidad de que abra la puerta en el primer intento? ¿Y en el segundo?
  - ¿Cuál es la probabilidad de que abra la puerta en el “ $r$ -ésimo” intento?
  - Si dicha persona debe abrir “ $n$ ” puertas con las “ $n$ ” llaves; ¿cuál es la probabilidad de que por lo menos abra una?
  - Determine la probabilidad de que el número de intentos infructuosos que realice sean 1 o un múltiplo de 4.
  - Calcule las probabilidades anteriores suponiendo que sí hace el descarte de cada llave que prueba.
- 2.43. Una forma de transporte entre dos ciudades que están apartadas 200 Km puede ser a través de tres alternativas: por carretera ( $C$ ), por tren ( $T$ ) y por avión ( $A$ ). Debido a las necesidades y costos, las probabilidades de decidirse por  $C$ ,  $T$  o  $A$  están dadas por:  $p(T)/p(C) = \frac{1}{2}$ ,  $P(C)/P(A) = \frac{2}{3}$ . Solo uno de estos tres medios de transporte puede ser construido; sin embargo, si el comité se decide por la construcción de la vía ferroviaria, la probabilidad de que sea terminada en un año es del 50%; si se decide por la carretera, la probabilidad es del 75%; y si se decide por la vía aérea, tiene una probabilidad del 90% de que el aeropuerto sea terminado en un año.
- ¿Cuál es la probabilidad de que las dos ciudades tengan un medio de transporte en un año?
  - Si algún medio de transporte es terminado en un año entre las dos ciudades, ¿cuál es la probabilidad de que este sea por vía aérea?
  - Si el comité se decide en favor de las facilidades que existen por tierra, cuál es la probabilidad de que la decisión final haya sido por carretera.
- 2.44. Antes de diseñar un túnel a través de una región rocosa, una exploración geológica es llevada a cabo para investigar las fracturas y las superficies de deslizamiento que existen entre los estratos de la roca. Por motivos económicos, solo porciones de los estratos son explorados. Entonces los geólogos solo pueden clasificar a los estratos como altamente fisurados ( $H$ ), medianamente fisurados ( $M$ ), ligeramente fisurados ( $L$ ).

También se sabe que:  $\frac{P(H)}{P(M)} = 1$ ,  $\frac{P(M)}{P(L)} = \frac{1}{8}$

Basados en esta información los ingenieros diseñan el túnel y estiman que si la roca es de condición  $L$ , la seguridad del diseño propuesto es del 99%; sin embargo, si la condición es  $M$ , la probabilidad de falla es el doble que la de la condición  $L$ ; si la condición es  $H$ , la probabilidad de falla es 10 veces mayor que para la condición  $L$ .

- a) ¿Cuál es la seguridad esperada del diseño del túnel propuesto?
- b) Un instrumento más confiable es usado para aumentar la predicción de las condiciones de la roca. Si los resultados indican que el alto fisuramiento en la roca es prácticamente imposible, pero no puede dar más información para la relativa ocurrencia de las condiciones  $M$  y  $L$ . A la luz de esta nueva información, ¿cuál debería ser la seguridad del diseño del túnel?
- c) Si el túnel colapsó, ¿cuál es la probabilidad de que hayan existido las condiciones  $M$  y  $L$  respectivamente?

2.45. Con base en varios estudios una compañía ha clasificado, de acuerdo con la posibilidad de descubrir petróleo, las formaciones geológicas en tres tipos. La compañía pretende perforar un pozo en un determinado sitio, al que se le asignan las probabilidades de 0.35, 0.40 y 0.25 para los tres tipos de formaciones respectivamente. De acuerdo con la experiencia, se sabe que el petróleo se encuentra en un 40% de formaciones del tipo I, en un 20% de formaciones del tipo II y en un 30% de formaciones del tipo III. Si la compañía no descubre petróleo en ese lugar, determinar la probabilidad de que exista una formación del tipo II.

2.46. Tres líneas aéreas (1, 2 y 3) hacen vuelos a Beirut. La experiencia ha demostrado que sufren accidentes: el 0.4% de la línea 1, el 0.5% de la línea 2 y el 0.7% de la línea 3. Pasado mañana el profesor desea ir a Beirut y elige aleatoriamente una línea aérea (suponga que no hay preferencia por alguna de ellas). Estime la probabilidad de que:

- a) Seleccione la línea 1 y que sufra un accidente.
- b) Sufra un accidente.
- c) Habiendo sufrido el accidente ¿Cuál fue la línea más probable que tomó?

2.47. Tres líneas de autobuses  $A$ ,  $B$  y  $C$  efectúan viajes diarios sin escalas entre Guadalajara y Morelia. Por experiencia se sabe que: el 60% de los viajes

de la línea A sale con retraso, el 50% de la línea B hace lo propio y el 45% de la línea C también. Si una persona selecciona al azar cualquiera de las tres líneas y su autobús sale retrasado, determinar la probabilidad de que haya seleccionado la línea B.

- 2.48. Mañana tendrá lugar el encuentro de fútbol entre los equipos: América y Cruz Azul. Nuestro comentarista deportivo ve la situación así: “Estoy seguro de que se abrirá el marcador y cualquiera de los equipos tiene iguales posibilidades de hacerlo; si América anota el primer gol su moral se levantará y la probabilidad de que el próximo gol también sea suyo es de  $\frac{2}{3}$  contra  $\frac{1}{3}$  de que sea de Cruz Azul; en cambio si es de Cruz Azul quien anota primero, la reacción no se hará esperar y habrá un segundo gol que puede ser con igual probabilidad para cualquier bando. Si el marcador llega a ponerse 2–0 a favor de cualquier equipo, la desmoralización del uno y la apatía del otro impedirán que haya más goles; en cambio, si llega a ponerse 1–1 pueden ocurrir tres cosas igual de probables: que América anote y gane 2–1; que Cruz Azul anote y gane 2–1 o que no haya más goles”.

Cuál es, de acuerdo con nuestro comentarista, la probabilidad de que:

- Cruz Azul gane.
- Cruz Azul gane dado que abrió el marcador.
- América haya abierto el marcador dado que Cruz Azul ganó.

Se sugiere construir un diagrama de árbol definiendo los siguientes eventos:

A: Gana el equipo América. Aa: El equipo América abre el marcador.

C: Gana el equipo Cruz Azul. Cc: El equipo Cruz Azul abre el marcador.

Los eventos A y Aa son independientes.

- 2.49. La policía planea hacer respetar los límites de velocidad, utilizando radares en 4 ubicaciones diferentes dentro de los límites de la ciudad. Se operan radares en cada una de las ubicaciones  $L_1$ ,  $L_2$ ,  $L_3$  y  $L_4$  en 40%, 30%, 20% y 10% del tiempo, y si una persona que rebasa los límites de velocidad en su camino al trabajo tiene probabilidades de 0.2, 0.1, 0.5 y 0.2 respectivamente, de pasar por estos lugares.

- ¿Cuál es la probabilidad de que reciba una multa?
- Si una persona recibió una multa por exceso de velocidad en su camino al trabajo, ¿Cuál es la probabilidad de que haya pasado por el radar ubicado en  $L_2$ ?



### 3. Variable Aleatoria, Funciones de Probabilidad de una Variable Aleatoria y sus Propiedades

Se denomina variable aleatoria (o estocástica) a una relación funcional del tipo  $f: S \rightarrow R$  que le asigna a cada evento simple de un espacio muestral  $S$  un número, cuyo valor puede ser discreto o continuo en los números reales. Estos valores posibles representan los resultados de experimentos que todavía no se han llevado a cabo o cantidades inciertas.

Por ejemplo, lanzar un dado numérico de seis caras y registrar la cara que cae hacia arriba, en este caso, su espacio muestral es  $\{1, 2, 3, 4, 5, 6\}$ ; nótese que a cada tirada le corresponde un número de los seis citados anteriormente. En este caso particular, la variable aleatoria es discreta. En la naturaleza existen muchos ejemplos de variables aleatorias discretas, como el número de llegadas a una cola, el número de artículos defectuosos en un lote, el número de defectos que presenta un automóvil, el número de vehículos que atraviesan por un puente, entre otros.

De la misma forma, existen muchos ejemplos de variables aleatorias continuas; al medir la carga de un resorte con un medidor de cargas, el valor que toma se encuentra en un rango continuo, lo mismo si se midiera la resistividad dieléctrica de un material, la dureza mecánica de una lámina, la temperatura de un cuerpo, el tiempo de atención en una estación de trabajo, el tiempo de vida de un producto, por mencionar algunos.

### 3.1. Variable Aleatoria Discreta, Función de Probabilidad y Función de Probabilidad Acumulada para una variable aleatoria discreta y sus propiedades

#### Ejercicio 3.1

Suponga que se lanzan dos dados numéricos de seis caras, uno rojo y otro azul, y se registran cuatro variables de este experimento: la primera variable será la suma de los números de las caras superiores de cada uno de los dados; la segunda variable será la diferencia del número de la cara superior del dado azul menos el número de la cara superior del dado rojo; la tercera variable será el producto de los números de las caras superiores de cada uno de los dados; y la cuarta variable será el tiempo transcurrido a partir de que se lanzan los dados y hasta que se quedan fijos ambos dados. Las primeras tres variables son discretas y la cuarta es continua; las cuatro variables coexisten en el mismo experimento. Se registrará a continuación el espacio muestral de los resultados de las primeras tres variables. Para ello, se aplicará el concepto de probabilidad de Laplace suponiendo que los dados son homogéneos y cada cara tiene la misma posibilidad de quedar hacia arriba. De esta forma, la probabilidad de que un dado en particular caiga con una de sus caras hacia arriba será  $\frac{1}{6}$ . La probabilidad de que caigan los dos dados en una sola posición es  $(\frac{1}{6})(\frac{1}{6}) = 1/36$ .

Para ilustrar los diferentes resultados que ocurren, se supondrá que un dado es rojo y el otro azul. Los números del dado de color rojo se etiquetarán con la letra  $x$  y los números del dado de color azul se etiquetarán con la letra  $y$ , de tal forma que el resultado de cada evento simple se representará como una pareja ordenada  $(x,y)$ . La primera variable se etiquetará con la letra  $u = x+y$ ; la segunda variable, con la letra  $v = y-x$ ; la tercera, con la letra  $w = xy$ . La tabla de resultados se muestra a continuación para cada una de las variables definidas. En cada tabla se colocan cuatro columnas: la primera es el conjunto de eventos simples que pueden ocurrir; la segunda es el valor numérico que toma cada variable; la tercera es la probabilidad asociada a cada valor que toma la variable respectiva; y la cuarta columna es la probabilidad acumulada de la variable correspondiente, desde el valor inicial hasta el valor que toma en la fila donde se encuentre.

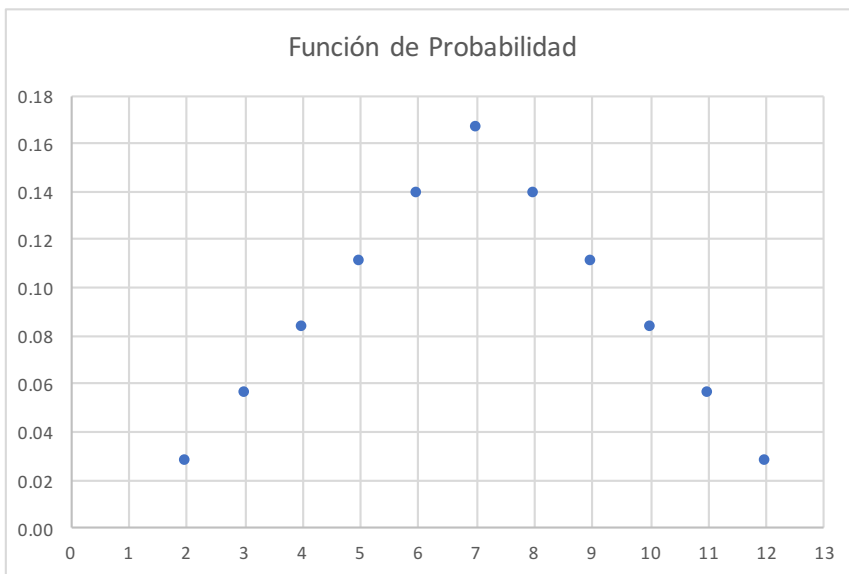
La relación  $p: u \rightarrow p(u)$  se conoce como función de probabilidad y la función  $P: u \rightarrow P(u)$  se denomina función de probabilidad acumulada. La gráfica de ambas funciones se muestra en las figuras 3.1 y 3.2.

De acuerdo con la figura 3.3, nótese que la función está compuesta por dos ramas que representan rectas con valores discretos en su abscisa  $u$ , las cuales pueden escribirse en términos matemáticos, con la ecuación de una recta dados dos de sus puntos.

FIGURA 3.1.

Evento	$u = x+y$	$p(u)$	$P(u)$
(1,1)	2	1/36	1/36
(1,2), (2,1)	3	2/36	3/36
(1,3), (3,1), (2,2)	4	3/36	6/36
(1,4), (4,1), (2,3), (3,2)	5	4/36	10/36
(1,5), (5,1), (4,2), (2,4), (3,3)	6	5/36	15/36
(1,6), (6,1), (2,5), (5,2), (3,4), (4,3)	7	6/36	21/36
(2,6), (6,2), (3,5), (5,3), (4,4)	8	5/36	26/36
(3,6), (6,3), (4,5), (5,4)	9	4/36	30/36
(4,6), (6,4), (5,5)	10	3/36	33/36
(5,6), (6,5)	11	2/36	35/36
(6,6)	12	1/36	36/36
Suma de resultados=		36/36=1	

FIGURA 3.2.



Nótese en la figura 3.2 que los puntos caen sobre dos líneas rectas, las cuáles se calcularán a continuación, con la ecuación de una recta dados dos de sus puntos:

$$p(u) = \frac{p_2(u_2) - p_1(u_1)}{(u_2 - u_1)} (u - u_1) + p_1(u_1) \quad (3.1)$$

Para la primera rama, la recta pasa por los puntos (2, 1/36) y (7, 6/36), por lo que su expresión matemática es:

$$p(u) = \frac{6/36 - 1/36}{(7-2)} (u-2) + \frac{1}{36}$$

$$p(u) = \frac{\frac{5}{36}}{5} (u-2) + \frac{1}{36}$$

$$p(u) = \frac{u}{36} - \frac{1}{36}$$

Para la segunda rama, la recta pasa por los puntos (7, 6/36) y (12, 1/36), por lo que su expresión matemática es:

$$p(u) = \frac{1/36 - 6/36}{(12-7)} (u-7) + \frac{6}{36}$$

$$p(u) = \frac{\frac{-5}{36}}{5} (u-7) + \frac{6}{36}$$

$$p(u) = \frac{-u}{36} + \frac{8}{36}$$

Por lo que la expresión matemática que representa a la función de probabilidad está dada por:

$$p(u) = \left\{ \begin{array}{ll} \frac{u}{36} + \frac{1}{36} & u=2, 3, 4, 5, 6, 7 \\ \frac{-u}{36} + \frac{8}{36} & u=8, 9, 10, 11, 12 \\ 0 & \text{para cualquier otro valor} \end{array} \right\}$$

FIGURA 3.3.

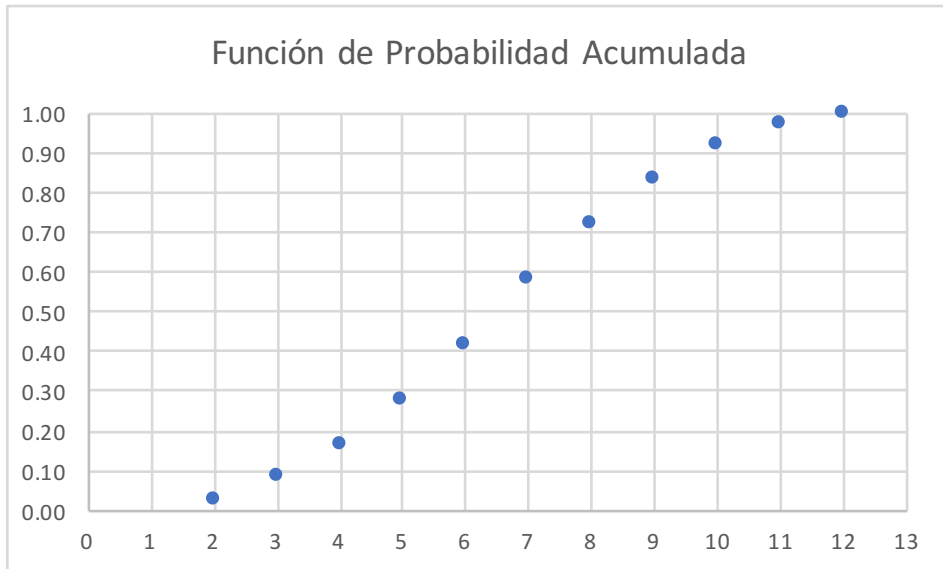


FIGURA 3.4.

Evento	$v=y-x$	$p(v)$	$P(v)$
(6,1)	-5	1/36	1/36
(6,2), (5,1)	-4	2/36	3/36
(6,3), (5,2), (4,1)	-3	3/36	6/36
(6,4), (5,3), (4,2), (3,1)	-2	4/36	10/36
(6,5), (5,4), (4,3), (3,2), (2,1)	-1	5/36	15/36
(6,6), (5,5), (4,4), (3,3), (2,2), (1,1)	0	6/36	21/36
(5,6), (4,5), (3,4), (2,3), (1,2)	1	5/36	26/36
(4,6), (3,5), (2,4), (1,3)	2	4/36	30/36
(3,6), (2,5), (1,4)	3	3/36	33/36
(2,6), (1,5)	4	2/36	35/36
(1,6)	5	1/36	36/36
Suma de resultados=		36/36=1	

Para el caso de la variable aleatoria  $v=y-x$ , su función de probabilidad se muestra en la figura 3.5 y su función de probabilidad acumulada se muestra en la figura 3.6.

FIGURA 3.5.

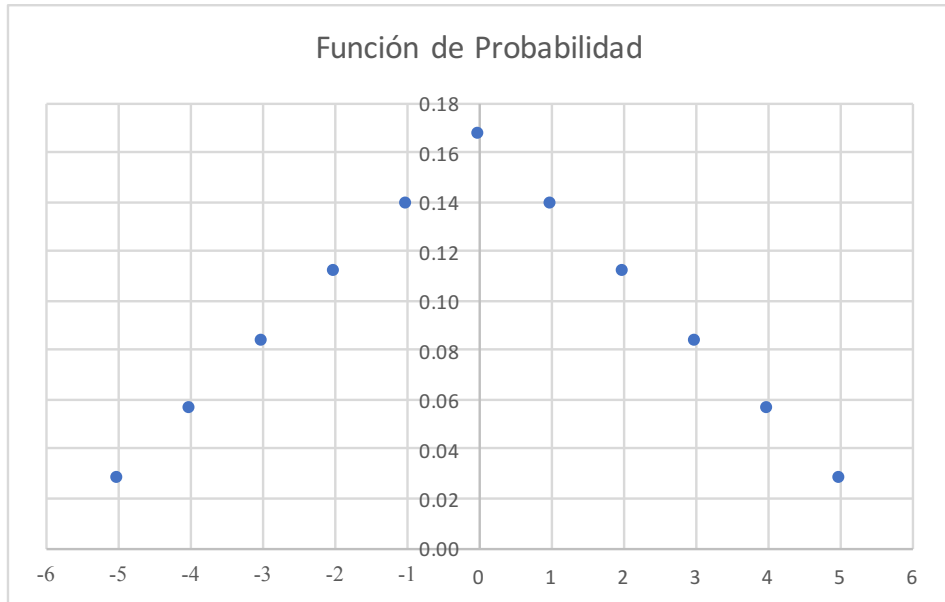
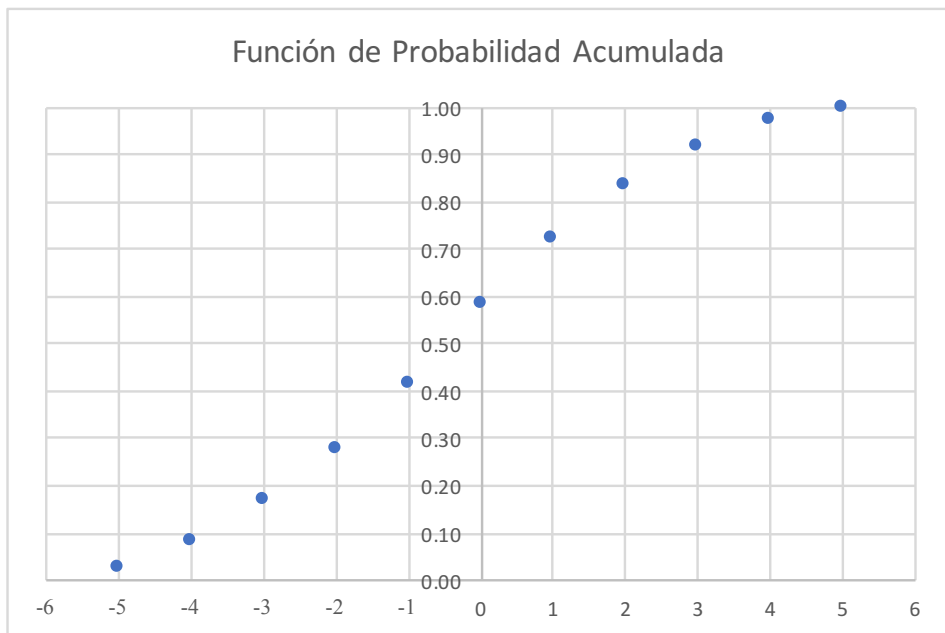


FIGURA 3.6.



Nótese de la figura 3.6, que la función de probabilidad  $p(v)$  también tiene una expresión matemática dada por la siguiente expresión:

$$p(v) = \left\{ \begin{array}{ll} \frac{v}{36} + \frac{1}{36} & v = -5, -4, -3, -2, -1, 0 \\ \frac{-v}{36} + \frac{1}{36} & v = 1, 2, 3, 4, 5 \\ 0 & \text{para cualquier otro valor} \end{array} \right\}$$

Observe la similitud de la función de probabilidad  $p(u)$  con la función de probabilidad  $p(v)$ ; tienen la misma forma, solo difieren en el dominio de definición de cada una de ellas.

FIGURA 3.7.

Evento	$w=xy$	$p(w)$	$P(w)$
(1,1)	1	1/36	1/36
(1,2), (2,1)	2	2/36	3/36
(1,3), (3,1)	3	2/36	5/36
(1,4), (4,1), (2,2)	4	3/36	8/36
(1,5), (5,1)	5	2/36	10/36
(2,3), (3,2), (1,6), (6,1)	6	4/36	14/36
(2,4), (4,2)	8	2/36	16/36
(3,3)	9	1/36	17/36
(2,5), (5,2)	10	2/36	19/36
(3,4), (4,3), (2,6), (6,2)	12	4/36	23/36
(3,5), (5,3)	15	2/36	25/36
(4,4)	16	1/36	26/36
(3,6), (6,3)	18	2/36	28/36
(4,5), (5,4)	20	2/36	30/36
(4,6), (6,4)	24	2/36	32/36
(5,5)	25	1/36	33/36
(5,6), (6,5)	30	2/36	35/36
(6,6)	36	1/36	36/36
Suma de resultados=		36/36=1	

Para el caso de la variable aleatoria  $w=xy$ , su función de probabilidad se muestra en la figura 3.8 y su función de probabilidad acumulada se muestra en la figura 3.9.

FIGURA 3.8.

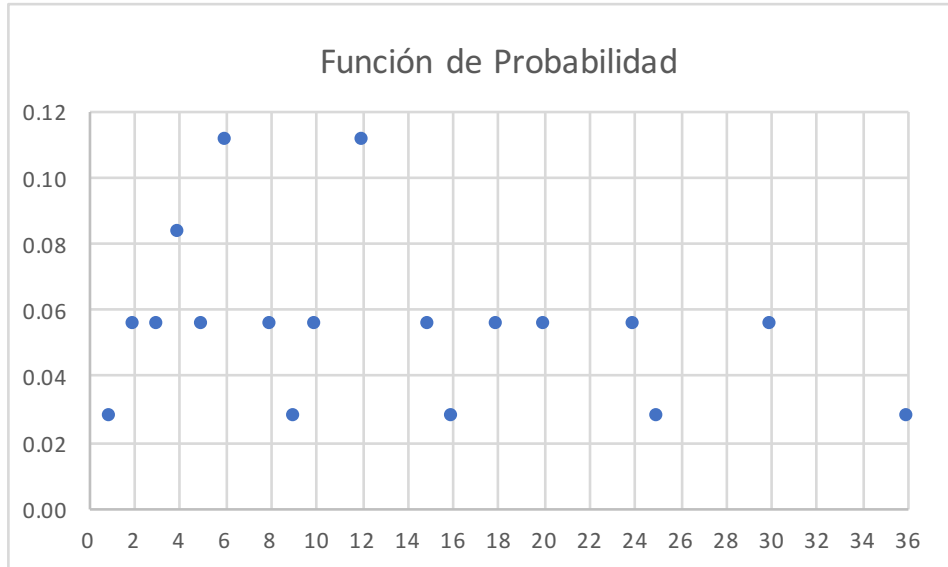
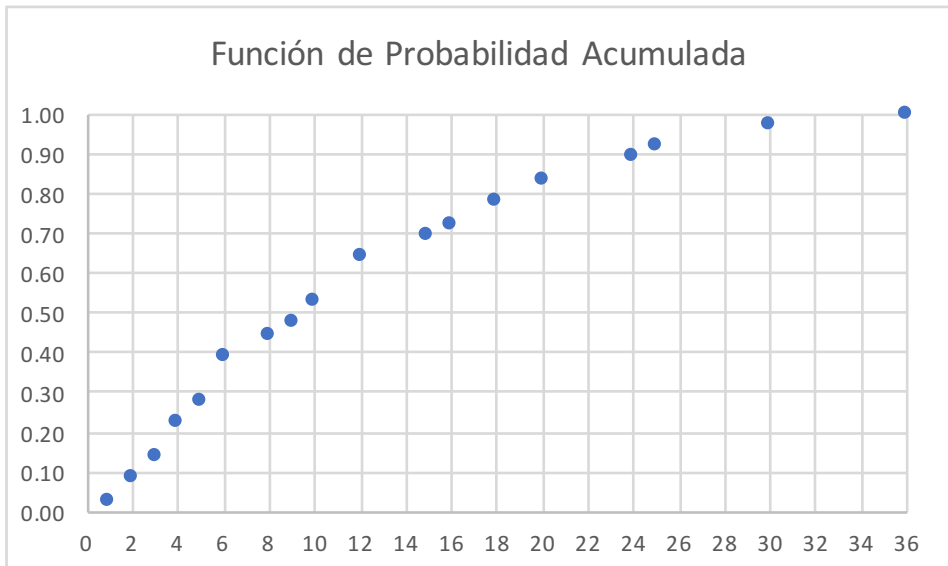


FIGURA 3.9.



Para cada una de las variables aleatorias  $u$ ,  $v$  y  $w$  se obtuvieron las funciones  $p(u)$ ,  $p(v)$  y  $p(w)$  conocidas como funciones de probabilidad y las funciones  $P(u)$ ,  $P(v)$  y  $P(w)$  denominadas funciones de probabilidad acumuladas. A las funciones de probabilidad acumuladas de variables aleatorias discretas se les conoce como funciones masa de probabilidad. Se mencionó que una variable



aleatoria como  $u$ ,  $v$  y  $w$  son funciones, el dominio de cada una de las tres funciones o variables aleatorias, en este caso proviene del mismo evento de tirar dos dados, pero la imagen de cada una de ellas es diferente en las tres. A su vez, la imagen de cada variable aleatoria  $u$ ,  $v$  y  $w$  es el dominio de su función de probabilidad y también es el dominio de su función de probabilidad acumulada.

Si observa cada una de las tres funciones de probabilidad,  $p(u)$ ,  $p(v)$  y  $p(w)$  se dará cuenta que las tres cumplen las mismas dos propiedades:

- i.  $0 \leq p(u) \leq 1$
- ii.  $\sum_{\forall u} p(u) = 1$  (3.2)

De la misma forma, las funciones de probabilidad acumuladas  $P(u)$ ,  $P(v)$  y  $P(w)$  cumplen las siguientes propiedades:

- i.  $0 \leq P(u) \leq 1$
- ii.  $P(u)$  es no decreciente, es decir,

$$\text{Si } u_1 \leq u_2 \quad \text{entonces } P(u_1) \leq P(u_2) \quad (3.3)$$

- iii.  $P(u_{\max}) = 1$

Existe una relación directa entre la función de probabilidad acumulada y la función de probabilidad, nótese que:

$$\text{i. } P(u) = \sum_{\forall t \leq u} p(t) \quad (3.4)$$

$$\text{ii. } p(u_i) = P(u_i) - P(u_{i-1}) \quad (3.5)$$

Observe que para el cálculo de probabilidades con las funciones de probabilidad acumuladas discretas:

$$p(t \leq u) = F(u) = \sum_{t=a}^{t=u} p(t)$$

$$p(c < t \leq d) = F(d) - F(c)$$

$$p(c \leq t \leq d) = F(d) - F(\text{anterior\_a\_c}) \quad (3.6)$$

$$p(c < t < d) = F(\text{anterior\_a\_d}) - F(c)$$

$$p(c \leq t < d) = F(\text{anterior\_a\_d}) - F(\text{anterior\_a\_c})$$

## Ejercicio 3.2

Sea  $t$  una variable aleatoria y su supuesta función de probabilidad definida en la tabla de la figura 3.10. Esta función se obtuvo del cociente entre los números de las caras superiores de los dos dados citados en el ejemplo 3.1, es decir,  $t = y/x$ .

FIGURA 3.10.

$t$	$p(t)$	$t$	$p(t)$	$t$	$p(t)$
0.1667	0.0278	0.7500	0.0278	1.6667	0.0278
0.2000	0.0278	0.8000	0.0278	2.0000	0.0833
0.2500	0.0278	0.8333	0.0278	2.5000	0.0278
0.3333	0.0556	1.0000	0.1667	3.0000	0.0556
0.4000	0.0278	1.2000	0.0278	4.0000	0.0278
0.5000	0.0833	1.2500	0.0278	5.0000	0.0278
0.6000	0.0278	1.3333	0.0278	6.0000	0.0278
0.6667	0.0556	1.5000	0.0556		

- a) Verifique si cumple las propiedades que debe cumplir una función de probabilidad discreta.

La primera propiedad de las expresiones 3.2 la cumple, ya que todas las probabilidades dadas para cada valor de  $t$  son mayores que cero y menores que uno.

La segunda propiedad también la cumple, ya que como se aprecia en la tabla de la figura 3.11 la suma de todas las probabilidades es uno.

- b) Obtenga la función de probabilidad acumulada. Se muestra en la tabla de la figura 3.11.

FIGURA 3.11.

$t$	$p(t)$	$P(t)$
0.1667	0.0278	0.0278
0.2000	0.0278	0.0556
0.2500	0.0278	0.0833
0.3333	0.0556	0.1389
0.4000	0.0278	0.1667
0.5000	0.0833	0.2500
0.6000	0.0278	0.2778

$t$	$p(t)$	$P(t)$
0.6667	0.0556	0.3333
0.7500	0.0278	0.3611
0.8000	0.0278	0.3889
0.8333	0.0278	0.4167
1.0000	0.1667	0.5833
1.2000	0.0278	0.6111
1.2500	0.0278	0.6389
1.3333	0.0278	0.6667
1.5000	0.0556	0.7222
1.6667	0.0278	0.7500
2.0000	0.0833	0.8333
2.5000	0.0278	0.8611
3.0000	0.0556	0.9167
4.0000	0.0278	0.9444
5.0000	0.0278	0.9722
6.0000	0.0278	1.0000

- c) Grafique la función de probabilidad y la función de probabilidad acumulada. Se muestran en las figuras 3.12 y 3.13.

**FIGURA 3.12.**

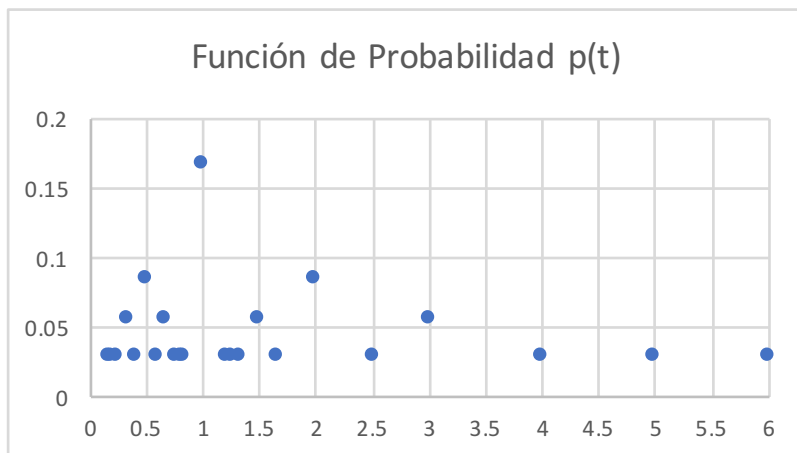
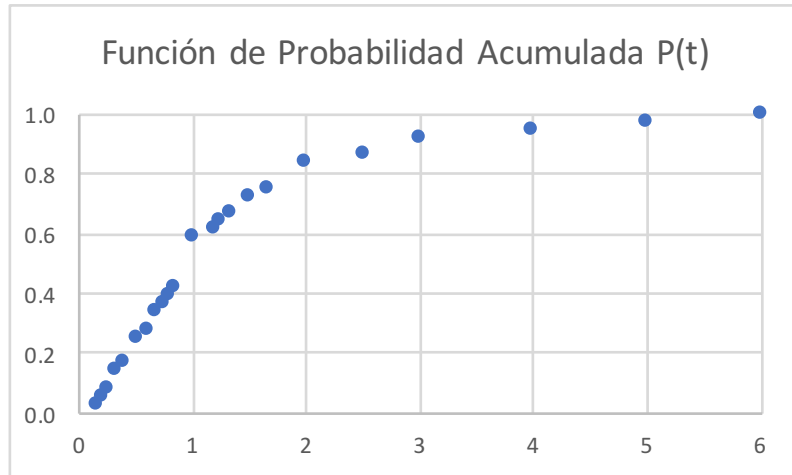


FIGURA 3.13.



d) Calcule las siguientes probabilidades:

i.  $p(t=1) = 0.1667$

ii.  $p(t \neq 1) = 1 - 0.1667 = 0.8333$

iii.  $p(t < 1) = P(t=1) - p(t=1) = 0.5833 - 0.1667 = 0.4167$

iv.  $p(t \leq 1) = P(t=1) = 0.5833$

v.  $p(t > 1) = 1 - p(t \leq 1) = 1 - 0.5833 = 0.4167$

vi.  $p(1 < t < 2) = p(t < 2) - p(t \leq 1) = p(t \leq 1.6667) - p(t \leq 1) = 0.75 - 0.5833 = 0.1667$

vii.  $p(1 \leq t < 2) = p(t < 2) - p(t < 1) = p(t \leq 1.6667) - p(t \leq 0.8333) = 0.75 - 0.4167 = 0.3333$

viii.  $p(1 < t \leq 2) = p(t \leq 2) - p(t \leq 1) = 0.8333 - 0.5833 = 0.25$

ix.  $p(1 \leq t \leq 2) = p(t \leq 2) - p(t < 1) = 0.8333 - 0.4166 = 0.6167$

### 3.2. Variable Aleatoria Continua, Función de Probabilidad y Función de Probabilidad Acumulada para una variable aleatoria continua y sus propiedades

De la misma forma que para el caso discreto, para el caso de las variables aleatorias continuas, se analiza un fenómeno en el que se define una función denominada variable aleatoria continua para la cual el dominio es un conjunto de eventos simples y la imagen es un intervalo real  $x \in [a, b]$ . Para cada variable continua se definen dos funciones denominadas función de probabilidad  $f(x)$  y función de probabilidad acumulada  $F(x)$ .

A la función de probabilidad de una variable aleatoria continua  $x$  también se le denomina función de densidad de probabilidad  $f(x)$  en el intervalo real  $x \in [a, b]$ , la cual cumple las siguientes propiedades:

$$\begin{aligned} \text{i. } & f(x) \geq 0 \\ \text{ii. } & \int_a^b f(x) dx = 1 \end{aligned} \quad (3.7)$$

Asimismo, la función de probabilidad acumulada  $F(x)$ , de una variable aleatoria continua  $x \in [a, b]$ , cumple las siguientes propiedades:

$$\begin{aligned} \text{i. } & 0 \leq F(x) \leq 1 \\ \text{ii. } & F(x) \text{ es no decreciente, es decir,} \\ & \text{Si } x_1 \leq x_2 \text{ entonces } P(x_1) \leq P(x_2) \end{aligned} \quad (3.8)$$

$$\text{iii. } F(a) = 0$$

$$\text{iv. } F(b) = 1$$

Existe una relación directa entre la función de probabilidad acumulada y la función de densidad de probabilidad, nótese que:

$$\text{i. } F(x) = \int_a^x f(x) dt \quad (3.9)$$

$$\text{ii. } p(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} f(x) dx = \int_a^{x_2} f(x) dx - \int_a^{x_1} f(x) dx = F(x_2) - F(x_1) \quad (3.10)$$

Nótese que si  $x_1$  es un valor real puntual la probabilidad

$$p(x_1) = \int_{x_1}^{x_1} f(t) dt = 0$$

Por lo que:

$$p(c \leq x \leq d) = p(c \leq x < d) = p(c < x \leq d) = p(c < x < d) = F(d) - F(c) \quad (3.11)$$

Aunque lo anterior no se aplica cuando se utiliza una función de probabilidad continua para estimar como aproximación la probabilidad de una variable aleatoria discreta. Esto se hace frecuentemente cuando se utiliza un discretizador de una variable continua, por ejemplo, el reloj es un instrumento discretizador del tiempo, no se puede medir exactamente porque el tiempo nunca se detiene, lo que se hace en estos casos es sumar y restar una cantidad para calcularlo como un intervalo. Para el caso de pretender estimar la probabilidad de que  $x$  tome un valor natural, lo que se hace generalmente es sumar y restar 0.5, es decir:

$$\begin{aligned} p(x=x_1) &= p(x_1 - 0.5 \leq x \leq x_1 + 0.5) = F(x_1 + 0.5) - F(x_1 - 0.5) \\ p(x \leq x_1) &= F(x_1 + 0.5) \\ p(x \leq x_1) &= F(x_1 - 0.5) \\ p(c \leq x \leq d) &= F(d) - F(c - 0.5) \\ p(c \leq x < d) &= F(d - 0.5) - F(c - 0.5) \\ p(c < x \leq d) &= F(d) - F(c) \\ p(c < x < d) &= F(d - 0.5) - F(c) \end{aligned} \quad (3.12)$$

### Ejercicio 3.3

Sea la función  $f(x)$  definida por la siguiente regla de correspondencia:

$$f(x) = \left\{ \begin{array}{ll} \frac{1}{2} \cos(x) & x \in \left[ -\frac{\pi}{2}, \frac{\pi}{2} \right] \\ 0 & \text{cualquier otro caso} \end{array} \right\}$$

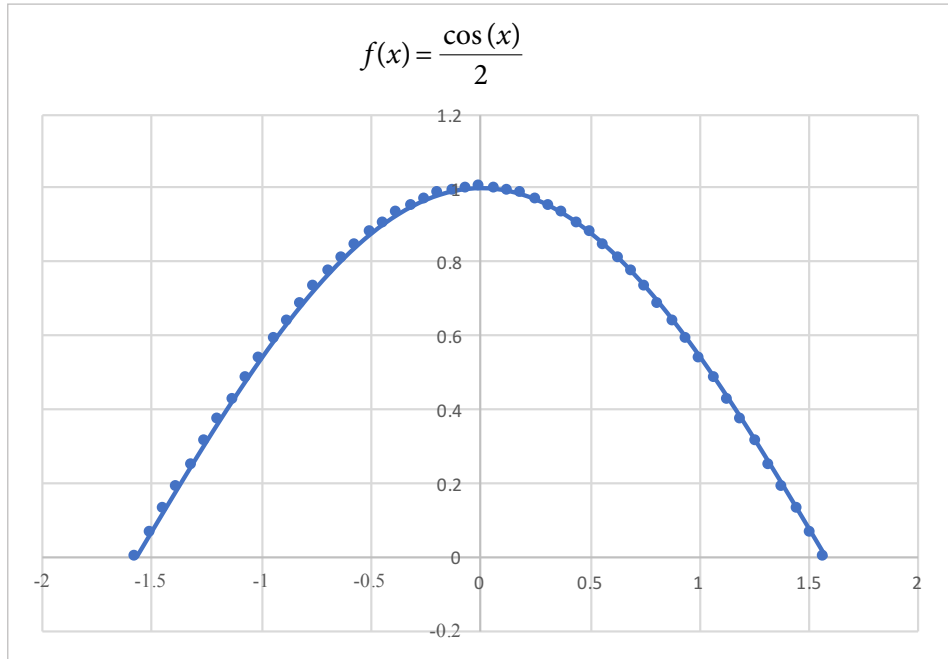
- a) Determine si se trata de una función de probabilidad continua. La gráfica de esta función se muestra en la figura 3.14. Como se puede apreciar en la figura 3.14

$$f(x) = \frac{1}{2} \cos(x) \geq 0 \quad \forall x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

Además

$$\frac{1}{2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos(x) dx = \left[ \frac{1}{2} \operatorname{sen}(x) \right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = 1$$

FIGURA 3.14.



b) Obtenga la función de probabilidad acumulada.

$$F(x) = \frac{1}{2} \int_{-\frac{\pi}{2}}^x \cos(t) dt = \left[ \frac{1}{2} \operatorname{sen}(t) \right]_{-\frac{\pi}{2}}^x = \frac{1}{2} \operatorname{sen}(x) + \frac{1}{2}$$

$$F(x) = \left\{ \begin{array}{ll} 0 & x < -\frac{\pi}{2} \\ \frac{1}{2} \operatorname{sen}(x) + \frac{1}{2} & x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \\ 1 & x > \frac{\pi}{2} \end{array} \right\}$$

La gráfica de la función de probabilidad acumulada se muestra en la figura 3.15.

c) Calcule las siguientes probabilidades.

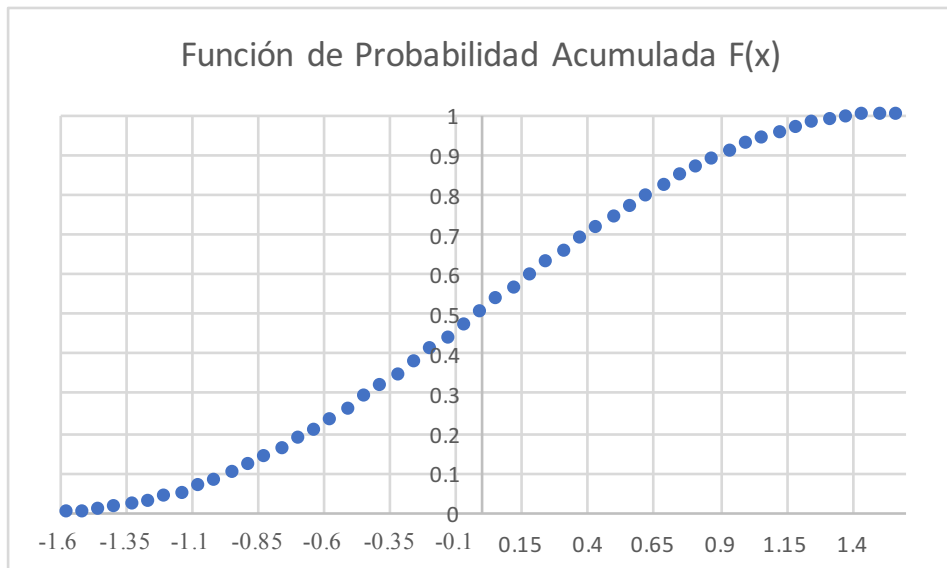
i.  $p(x \leq 1) = F(x=1) = \frac{\sin(1)}{2} + \frac{1}{2} = 0.920735$

ii.  $p(-0.5 < x < 0.5)$ . Observe que  $x$  es simétrica con respecto al origen, por lo que  $p(-a < x < a) = 2p(0 < x < a) = 2p(x < a) - 1$

$$p(-0.5 < x < 0.5) = 2p(x < 0.5) - 1 = 2\left[\frac{\sin(0.5)}{2} + \frac{1}{2}\right] - 1 = 0.479426$$

iii.  $p(x > 1.0) = 1 - p(x \leq 1) = 1 - 0.920735 = 0.079265$

FIGURA 3.15.





### 3.3. Percentiles, deciles, cuartiles, medidas de una función de probabilidad y sus propiedades

Una variable aleatoria y su función de probabilidad presentan ciertas propiedades que tienen una importancia capital para conocerla. Estas propiedades se clasifican en Medidas de Tendencia Central: Media, Mediana, Moda, Semirango. Medidas de Dispersión: Rango, Varianza, Desviación Estándar, Coeficiente de Variación. Medidas de Forma: Coeficiente de Asimetría y Coeficiente de Curtosis o de Aplanamiento.

Por el momento, solo se explicarán algunas de ellas y se retomará el tema una vez que se aborde el tema de Esperanza Matemática.

El Rango ( $R$ ) de una variable aleatoria es una medida de dispersión y se obtiene de la diferencia de los valores extremos del dominio de definición de la función de probabilidad. El Semirango es una medida de tendencia central cuyo valor es el valor central del dominio.

Para ilustrar lo relacionado con el rango de una variable aleatoria discreta, observe que para el caso del ejemplo 3.1 del lanzamiento de dos dados cuyas caras superiores se representan con las letras  $x$  y  $y$ , y las variables aleatorias asociadas al evento  $u = x+y$ ,  $v = y-x$  y  $w = xy$ , nótese que el dominio de definición de la función de probabilidad es:  $u \in [2, 12]$ ,  $v \in [-5, 5]$  y  $w \in [1, 36]$ , por lo que el rango de cada una de ellas es el siguiente:  $R_u = 11$  (recuerde que el dominio incluye los valores extremos, por lo que los puntos son once),  $R_v = 11$  y  $R_w = 36$ . El semirango ( $SR$ ) en el caso de las variables anteriores es el punto central de cada dominio, en este caso:  $SR_u = 7$ ,  $SR_v = 0$ ,  $SR_w = 18$ .

Para el ejemplo 3.3, que representa a una variable aleatoria continua, cuya gráfica se presenta en la figura 3.14, su dominio es  $x \in [-\pi/2, \pi/2]$ ; su rango se obtiene directamente de la diferencia del valor mayor del intervalo menos el valor mínimo del intervalo, en este caso  $R_x = \pi$ ; y su semirango  $SR$  es cero.

Los percentiles son los valores que parten al dominio en cien partes, de tal manera que cada percentil representa la probabilidad acumulada en porcentaje hasta ese valor. El primer percentil ( $p_1$ ) representa el valor del dominio para el cual se tiene el 1% de la probabilidad acumulada hasta ese valor. El segundo percentil ( $p_2$ ) representa el valor del dominio para el cual se tiene el 2% de la probabilidad acumulada hasta ese valor, y así sucesivamente; el percentil  $p_n$  representa el valor para el cual su probabilidad acumulada es del  $p_n\%$ . El intervalo interpercentil (IIP) es la diferencia entre el percentil 99 y el percentil 1. Este intervalo contiene al 98% de probabilidad.

Para el caso de una variable aleatoria discreta, existe una forma de estimar los percentiles a través de interpolación lineal, usando la expresión 3.13 de la ecuación de una recta dados dos de sus puntos. Sea  $x$  una variable aleatoria discreta, con función de probabilidad  $p(x)$ . Suponga que se desea conocer el percentil  $x_p$  para el cual, la probabilidad acumulada hasta ese punto es  $p$ , es decir,  $p = p(x_p)$ . Suponga también que  $x_1 < x_p < x_2$  y que se sabe que  $(x_1, p_1)$  y  $(x_2, p_2)$ , en donde,  $p_1 = p(x_1)$  y  $p_2 = p(x_2)$ , entonces, el percentil  $x_p$  estaría dado por la expresión:

$$x_p = \frac{x_2 - x_1}{p_2 - p_1} (p - p_1) + x_1 \quad (3.13)$$

En la anterior expresión, se debe buscar que los puntos  $x_1$  y  $x_2$  estén lo más cercanos posibles a  $x_p$ , para que la interpolación lineal se ajuste adecuadamente.

#### Ejercicio 3.4

Estime el percentil 70 para el caso de la suma de las caras de los dados del ejercicio 3.1.

Observe que la curva correspondiente a la función de probabilidad acumulada en su segunda rama pasa por los puntos  $(7, 0.5833)$  y  $(8, 0.7222)$ , por lo que para calcular la abscisa para una ordenada de 0.70, se utilizará la interpolación lineal para una probabilidad  $p(u) = 0.70$ , usando la fórmula de una recta dados dos de sus puntos:

De la expresión 3.13, si se despeja la  $u$ :

$$u = \frac{u_2 - u_1}{(p(u_2) - p(u_1))} (p(u) - p(u_1)) + u_1$$

$$u = \frac{8 - 7}{0.7222 - 0.5833} (0.70 - 0.5833) + 7$$

$$u_{70\%} = 7.8402$$

El valor de  $u_{70\%} = 7.8402$  significa que la probabilidad acumulada hasta este valor es del 70%, lo que significa que más del 70% de las veces la suma dará un valor menor a 8.

Los deciles son los valores que parten al dominio en diez partes de tal manera que cada decil representa un porcentaje del 10%, por lo cual también

son percentiles. Existen diez deciles en todo el dominio; el primer decil representa el valor del dominio para el cual la probabilidad acumulada hasta ese valor es del 10%; el segundo decil representa el valor del dominio para el cual la probabilidad acumulada hasta ese valor es del 20%, y así sucesivamente. A la diferencia entre el noveno decil y el primer decil se le denomina Intervalo Interdecil (*IID*). El intervalo interdecil contiene el 90% de probabilidad.

### Ejercicio 3.5

Para el ejercicio 3.2, calcule el decil 9, el decil 1 y el intervalo interdecil.

Observe que la curva correspondiente a la función de probabilidad acumulada en su primera rama pasa por los puntos (0.25, 0.0833) y (0.3333, 0.1389), por lo que para calcular la abscisa para una ordenada de 0.10, se utilizará la interpolación lineal para una probabilidad  $p(t)=0.10$ , usando la fórmula de una recta dados dos de sus puntos:

De la expresión 3.12, si se despeja la  $u$  y se sustituye por  $t$ :

$$t = \frac{t_2 - t_1}{p(t_2) - p(t_1)} [p(t) - p(t_1)] + t_1$$

$$t = \frac{0.3333 - 0.25}{0.1389 - 0.0833} (0.10 - 0.0833) + 0.25$$

$$t = 0.248$$

De la misma forma, la curva correspondiente a la segunda rama de la función de probabilidad acumulada pasa por los puntos (2.5, 0.8611) y (3, 0.9167), por lo que interpolando linealmente para  $p(t)=0.90$ :

$$t = \frac{3 - 2.5}{(0.9167 - 0.8611)} [0.90 - 0.8611] + 2.5$$

$$t = 2.85$$

El intervalo interdecil sería  $IID = 2.85 - 0.248 = 2.602$ .

Esto significa que en el intervalo [0.248, 2.85] cae el 80% de la probabilidad.

Nótese que para el caso del ejercicio 3.2, en la figura 3.11 se puede observar que el primer cuartil o percentil 25 es  $q_{t1} = 0.5$  y representa el valor de  $t$  para el cual, la probabilidad acumulada hasta ese valor es 25%; el tercer cuartil  $q_{t3} = 1.6667$  representa el valor de  $t$  para el cual, la probabilidad acumulada hasta ese valor es 75%. Se le llama intervalo intercuartil a la diferencia entre el tercer cuartil menos el primer cuartil, por lo que el intervalo intercuartil en este ejemplo 3.2 es  $IIC = 1.6667 - 0.5 = 1.1667$ , lo que significa que entre el punto  $t1 = 0.5$  y el punto  $t2 = 1.6667$  se acumula el 50% de probabilidad.

Para el caso de una variable aleatoria continua, el proceso de obtener los percentiles, deciles o cuartiles puede ser más sencillo, si se obtiene la función de probabilidad acumulada, se iguala al percentil que se requiere y se despeja la variable aleatoria independiente. En términos matemáticos significa que si  $p = F(x_p)$ , entonces  $x_p = F^{-1}(p)$ , en donde  $F^{-1}$  es la función inversa de  $F$ .

## Ejercicio 3.6

Para el ejercicio 3.3 calcule:

- a) El percentil 67.

En el caso analizado la función de probabilidad acumulada fue:

$$F(x) = \frac{\text{sen}(x)}{2} + \frac{1}{2}$$

Por lo que

$$\frac{\text{sen}(x)}{2} + \frac{1}{2} = 0.67$$

$$\frac{\text{sen}(x)}{2} = 0.67 - 0.5$$

$$\text{sen}(x) = 0.34$$

$$x = \text{angsen}(0.34)$$

$$x = 0.3469$$

- b) Obtenga el decil 4.

$$\frac{\text{sen}(x)}{2} + \frac{1}{2} = 0.4$$

$$\frac{\text{sen}(x)}{2} = 0.4 - 0.5$$

$$\text{sen}(x) = -0.1$$

$$x = \text{angsen}(-0.1)$$

$$x = -0.1002$$

c) Calcule el intervalo intercuartil

$$\frac{\text{sen}(x)}{2} + \frac{1}{2} = 0.25 \quad \frac{\text{sen}(x)}{2} + \frac{1}{2} = 0.75$$

$$\text{sen}(x) = -0.5 \quad \text{sen}(x) = 0.5$$

$$x_{0.25} = -0.5236 \quad x_{0.75} = 0.5236$$

Por lo que el intervalo intercuartil es  $IIC = 1.0472$

La mediana de una variable aleatoria es aquel valor del dominio que parte a la mitad a la probabilidad acumulada, el 50% cae a la izquierda de este valor y el otro 50% a la derecha. La mediana es el percentil 50, el decil 5 o el segundo cuartil. Para una variable aleatoria discreta, puede ser obtenida con la expresión 3.12, es decir:

$$m_e = \frac{x_2 - x_1}{p_2 - p_1} (0.5 - p_1) + x_1 \quad (3.14)$$

La moda, cuando es única, se define como el valor de  $X$  que proporciona el máximo valor que alcanza la función de probabilidad.

Como se puede comprobar por interpolación lineal, en el caso del ejemplo 3.1, para  $u = x + y$ , la mediana es  $m_e = 6.5$  y la moda es  $m_o = 7$ . Para  $v = y - x$ , la mediana es  $-0.5$  y la moda es  $m_o = 0$ . Para  $w = xy$ , la mediana es  $m_e = 9.5$ , y de acuerdo con la figura 3.8, existen dos valores de  $x$  para los cuales se logra la altura más alta en su función de probabilidad, en  $x = 6$  y en  $x = 12$ ; algunos autores definen a este tipo de funciones como bimodal, otros autores sostienen que, al no ser única, se dice que no existe moda.

Para el caso del ejercicio 3.2, en donde se definió  $t = y/x$ , la mediana, por interpolación lineal, es  $m_e = 0.91665$  y la moda, al observar la figura 3.12, es  $m_o = 1$ .

Para el caso del ejercicio 3.3, dado que se trata de una curva continua, y simétrica, como se muestra en la figura 3.14, la mediana y la moda toman el mismo valor  $m_e = m_o = 0$ .

### 3.4. Esperanza matemática, momentos, valor esperado, varianza y sus propiedades

Sea  $x$  una variable aleatoria, con función de probabilidad  $p(x)$  para el caso discreto (en el caso discreto no toma todos los valores del intervalo, sólo los que se definen en el dominio) y  $f(x)$  para el caso continuo, en el intervalo  $x \in [a, b]$ . Sea  $g(x)$  una función de  $x$ . Se define la Esperanza Matemática de  $g(x)$  como:

$$E\{g(x)\} = \left\{ \begin{array}{ll} \sum_{\forall x \in [a, b]} g(x)p(x) & \text{caso discreto} \\ \int_{\forall x \in [a, b]} g(x)f(x)dx & \text{caso continuo} \end{array} \right\} \quad (3.15)$$

Propiedades de la Esperanza Matemática:

- i.  $E\{k\} = k$
  - ii.  $E\{kg(x)\} = kE\{g(x)\}$
  - iii.  $E\{g_1(x) + g_2(x)\} = E\{g_1(x)\} + E\{g_2(x)\}$
  - iv.  $E\{k_1g_1(x) + k_2g_2(x)\} = k_1E\{g_1(x)\} + k_2E\{g_2(x)\}$
  - v.  $E\{k_1g_1(x) + k_2g_2(x) + \dots + k_n g_n(x)\} = k_1E\{g_1(x)\} + k_2E\{g_2(x)\} + \dots + k_nE\{g_n(x)\}$
- (3.16)

Demostrar estas propiedades es relativamente sencillo, ya que de Cálculo se sabe que tanto una sumatoria como una integral son operadores lineales. Por ello, solo se demostrará la primera y la quinta, la demostración se hará para variables aleatorias continuas, para las discretas es similar la demostración; asimismo, el resto de las propiedades se demostrarían en forma similar.

Para la primera propiedad:

Recuerde que  $\int_{\forall x \in [a,b]} f(x)dx = 1$

$$E\{k\} = \int_{\forall x \in [a,b]} kf(x)dx = k \int_{\forall x \in [a,b]} f(x)dx = k$$

Para la quinta propiedad:

$$\begin{aligned} E\{k_1g_1(x) + k_2g_2(x) + \dots + k_n g_n(x)\} &= \\ \int_{\forall x \in [a,b]} [k_1g_1(x) + k_2g_2(x) + \dots + k_n g_n(x)] f(x)dx &= \\ k_1 \int_{\forall x \in [a,b]} g_1(x) f(x)dx + k_2 \int_{\forall x \in [a,b]} g_2(x) f(x)dx + \dots + k_n \int_{\forall x \in [a,b]} g_n(x) f(x)dx &= \\ k_1 E\{g_1(x)\} + k_2 E\{g_2(x)\} + \dots + k_n E\{g_n(x)\} & \end{aligned}$$

El concepto de Esperanza Matemática es fundamental en la probabilidad y la estadística como se verá en las definiciones siguientes:

Sea  $x$  una variable aleatoria con función de probabilidad  $p(x)$ , si esta es discreta; y  $f(x)$ , si es continua en el intervalo  $x \in [a, b]$ ; se define el momento de orden  $k$  con respecto al origen, como:

$$\mu'_k = \left\{ \begin{array}{ll} \sum_{\forall x \in [a,b]} x^k p(x) & \text{caso discreto} \\ \int_{\forall x \in [a,b]} x^k f(x)dx & \text{caso continuo} \end{array} \right\} \quad (3.17)$$

Al momento de orden uno de la variable aleatoria  $x$ , se le conoce como media o valor esperado de  $x$ , y representa una medida de la tendencia central que presenta la gráfica de la función de probabilidad de la variable aleatoria  $x$ :

$$\mu'_k = \left\{ \begin{array}{ll} \sum_{\forall x \in [a,b]} xp(x) & \text{caso discreto} \\ \int_{\forall x \in [a,b]} xf(x)dx & \text{caso continuo} \end{array} \right\} \quad (3.18)$$

El concepto de media de una variable aleatoria es fundamental para entender la tendencia central de su función de probabilidad, por lo que conviene analizar algunas de sus propiedades:

- i. La media de una constante es la constante misma:

$$\mu_a = E\{a\} = a \quad (3.19)$$

- ii. La media de una constante por una variable aleatoria es la constante por la media de esa variable aleatoria:

$$\mu_{ax} = E\{ax\} = aE\{x\} = a \mu_x \quad (3.20)$$

- iii. La media de una suma de dos variables aleatorias es igual a la suma de las medias de cada una de ellas:

$$\mu_{x_1+x_2} = E\{x_1+x_2\} = E\{x_1\} + E\{x_2\} = \mu_{x_1} + \mu_{x_2} \quad (3.21)$$

- iv. La media de una combinación lineal de dos variables aleatorias es igual a la misma combinación lineal de las medias de cada una de las dos variables:

$$\mu_{a_1x_1+a_2x_2} = E\{a_1x_1 + a_2x_2\} = a_1E\{x_1\} + a_2E\{x_2\} = a_1\mu_{x_1} + a_2\mu_{x_2} \quad (3.22)$$

- v. La media de una combinación lineal de  $n$  variables aleatorias es igual a la misma combinación lineal de las medias de cada una de las  $n$  variables aleatorias:

$$\begin{aligned} \mu_{a_1x_1+a_2x_2+\dots+a_nx_n} &= E\{a_1x_1 + a_2x_2 + \dots + a_nx_n\} = \\ \mu_{a_1x_1+a_2x_2+\dots+a_nx_n} &= a_1E\{x_1\} + a_2E\{x_2\} + \dots + a_nE\{x_n\} \\ \mu_{a_1x_1+a_2x_2+\dots+a_nx_n} &= a_1 \mu_{x_1} + a_2 \mu_{x_2} + \dots + a_n \mu_{x_n} \end{aligned} \quad (3.23)$$

Sea  $x$  una variable aleatoria con función de probabilidad  $p(x)$ , si esta es discreta; y  $f(x)$  si es continua en el intervalo  $x \in [a, b]$ ; se define el momento de orden  $k$  con respecto a la media, como:



$$\mu_k = \left\{ \begin{array}{ll} \sum_{\forall x \in [a,b]} (x-\mu_x)^k p(x) & \text{caso discreto} \\ \int_{\forall x \in [a,b]} (x-\mu_x)^k f(x) dx & \text{caso continuo} \end{array} \right\} \quad (3.24)$$

Observe que existe una relación directa entre los momentos de orden  $k$  con respecto a la media y los momentos de orden  $k$  con respecto al origen. Se mostrarán a continuación las relaciones para los cuatro primeros momentos con respecto a la media y se proporcionará una fórmula para cualquier valor de  $k$ .

El momento de orden uno con respecto a la media es cero, como se observa a continuación:

$$\mu_1 = E\{(x-\mu_x)\} = E\{x\} - \mu_x = \mu_x - \mu_x = 0 \quad (3.25)$$

El momento de orden dos con respecto a la media recibe el nombre de Varianza y se representa por  $\sigma^2$ , la cual es una medida del grado de dispersión de la gráfica de la función de probabilidad de la variable aleatoria  $x$  con respecto a su media:

$$\text{var}\{x\} = \sigma_x^2 = \mu_2 = \left\{ \begin{array}{ll} \sum_{\forall x \in [a,b]} (x-\mu_x)^2 p(x) & \text{caso discreto} \\ \int_{\forall x \in [a,b]} (x-\mu_x)^2 f(x) dx & \text{caso continuo} \end{array} \right\} \quad (3.26)$$

Nótese que

$$\begin{aligned} \text{var}(x) &= E\{(x-\mu_x)^2\} \\ \text{var}(x) &= E\{x^2\} - 2\mu_x E\{x\} + E\{\mu_x^2\} \\ \text{var}(x) &= E\{x^2\} - 2\mu_x \mu_x + \mu_x^2 \\ \text{var}(x) &= E\{x^2\} - \mu_x^2 \end{aligned}$$

Lo que implica que la varianza también se puede calcular como el momento de segundo orden con respecto al origen, menos el momento de primer orden con respecto al origen al cuadrado, es decir:

$$\sigma_x^2 = \text{var}(x) = \mu_2' - \mu_1'^2 = E\{x^2\} - \mu_x^2 \quad (3.27)$$

El concepto de varianza de una variable aleatoria es fundamental para entender la dispersión de la variable y de su función de probabilidad, por lo que conviene analizar algunas de sus propiedades:

- i. La varianza de una constante es cero:

$$\sigma_a^2 = \text{var}(a) = E\{(a-a)^2\} = 0 \quad (3.28)$$

- ii. La varianza de una constante por una variable aleatoria es la constante al cuadrado por la varianza de esa variable aleatoria:

$$\sigma_x^2 = \text{var}\{ax\} = E\{(ax-a\mu_x)^2\} = E\{a^2(x-\mu_x)^2\} = a^2\sigma_x^2 \quad (3.29)$$

- iii. La varianza de una suma de dos variables aleatorias es igual a la suma de las varianzas de cada una de ellas menos dos veces la covarianza entre ellas, es decir:

$$\begin{aligned} \sigma_{x_1+x_2}^2 &= \text{var}\{x_1+x_2\} = E\{[(x_1+x_2)-(\mu_{x_1}+\mu_{x_2})]^2\} \\ \sigma_{x_1+x_2}^2 &= E\{[(x_1-\mu_{x_1})-(x_2-\mu_{x_2})]^2\} \\ \sigma_{x_1+x_2}^2 &= E\{(x_1-\mu_{x_1})^2\} + E\{(x_2-\mu_{x_2})^2\} + 2E\{(x_1-\mu_{x_1})(x_2-\mu_{x_2})\} \\ \sigma_{x_1+x_2}^2 &= \text{var}\{x_1\} + \text{var}\{x_2\} + 2\text{cov}\{x_1, x_2\} \\ \sigma_{x_1+x_2}^2 &= \sigma_{x_1}^2 + \sigma_{x_2}^2 + \sigma_{x_1, x_2}^2 \end{aligned} \quad (3.30)$$

En la expresión 3.30 anterior, aparece un nuevo concepto, al cual se le denomina covarianza de dos variables aleatorias la cual se define de la siguiente forma:

Sean dos variables aleatorias  $x_1$  y  $x_2$ , con funciones de probabilidad  $f_1(x_1)$  y  $f_2(x_2)$ , se define la covarianza entre ellas como:

$$\sigma_{x_1, x_2}^2 = \text{cov}\{x_1, x_2\} = E\{(x_1-\mu_1)(x_2-\mu_2)\}$$

$$\sigma_{x_1, x_2}^2 = \left\{ \begin{array}{ll} \sum_{\forall x_1} \sum_{\forall x_2} (x_1 - \mu_1)(x_2 - \mu_2) f_1(x_1) f_2(x_2) & x_1, x_2 \text{ discretas} \\ \int_{\forall x_1} \int_{\forall x_2} (x_1 - \mu_1)(x_2 - \mu_2) f_1(x_1) f_2(x_2) dx_1 dx_2 & x_1, x_2 \text{ continuas} \end{array} \right\} \quad (3.31)$$

Más adelante se analizarán las propiedades de la covarianza, por lo pronto, cabe remarcar que, si la variable aleatoria  $x_1$  no depende de la variable aleatoria  $x_2$ , es decir, si ambas son estadísticamente independientes, la covarianza de ambas es cero, por lo que se puede afirmar que:

$$\text{Si } x_1 \text{ es independiente de } x_2 \rightarrow \sigma_{x_1, x_2}^2 = \text{cov}\{x_1, x_2\} = 0 \quad (3.32)$$

Lo que implica que:

$$\text{Si } x_1 \text{ es independiente de } x_2 \text{ entonces } \sigma^2(x_1 + x_2) = \sigma_{x_1}^2 + \sigma_{x_2}^2 \quad (3.33)$$

- iv. La varianza de una combinación lineal de dos variables aleatorias es igual a la combinación lineal de las varianzas de cada una de las dos variables, pero con los coeficientes al cuadrado más dos veces el producto de los dos coeficientes por la covarianza entre ellas:

$$\begin{aligned} \sigma_{a_1x_1+a_2x_2}^2 &= E\{[(a_1x_1 + a_2x_2) - (a_1\mu_{x_1} + a_2\mu_{x_2})]^2\} \\ \sigma_{a_1x_1+a_2x_2}^2 &= E\{[(a_1x_1 - a_1\mu_{x_1}) + (a_2x_2 - a_2\mu_{x_2})]^2\} \\ \sigma_{a_1x_1+a_2x_2}^2 &= a_1^2 E\{(x_1 - \mu_{x_1})^2\} + a_2^2 E\{(x_2 - \mu_{x_2})^2\} + 2a_1a_2 E\{(x_1 - \mu_{x_1})(x_2 - \mu_{x_2})\} \\ \sigma_{a_1x_1+a_2x_2}^2 &= a_1^2 \text{var}\{x_1\} + a_2^2 \text{var}\{x_2\} + 2a_1a_2 \text{cov}\{x_1, x_2\} \\ \sigma_{a_1x_1+a_2x_2}^2 &= a_1^2 \sigma_{x_1}^2 + a_2^2 \sigma_{x_2}^2 + 2a_1a_2 \sigma_{x_1, x_2}^2 \end{aligned} \quad (3.34)$$

Nuevamente:

$$\text{Si } x_1 \text{ es independiente de } x_2 \text{ entonces } \sigma_{a_1x_1+a_2x_2}^2 = a_1^2 \sigma_{x_1}^2 + a_2^2 \sigma_{x_2}^2 \quad (3.35)$$

- v. La varianza de una combinación lineal de  $n$  variables aleatorias está dada por la siguiente expresión:

$$\begin{aligned}\sigma_{a_1x_1+a_2x_2+\dots+a_nx_n}^2 &= a_1^2 \text{var}\{x_1\} + a_2^2 \text{var}\{x_2\} + \dots + a_n^2 \text{var}\{x_n\} \\ &+ 2a_1a_2 \text{cov}\{x_1,x_2\} + 2a_1a_3 \text{cov}\{x_1,x_3\} + \dots + 2a_{n-1}a_n \text{cov}\{x_{n-1},x_n\} \\ \sigma_{a_1x_1+a_2x_2+\dots+a_nx_n}^2 &= a_1^2 \sigma_{x_1}^2 + a_2^2 \sigma_{x_2}^2 + \dots + a_n^2 \sigma_{x_n}^2 \\ &+ 2a_1a_2 \sigma_{x_1,x_2}^2 + 2a_1a_3 \sigma_{x_1,x_3}^2 + \dots + 2a_{n-1}a_n \sigma_{x_{n-1},x_n}^2\end{aligned}\quad (3.36)$$

De la misma forma:

Si  $x_i$  es independiente de  $x_j \forall i \neq j$  entonces

$$\sigma_{a_1x_1+a_2x_2+\dots+a_nx_n}^2 = a_1^2 \sigma_{x_1}^2 + a_2^2 \sigma_{x_2}^2 + \dots + a_n^2 \sigma_{x_n}^2 \quad (3.37)$$

A la raíz cuadrada de la varianza de una variable aleatoria  $x$  se le conoce como Desviación Estándar de  $x$ , y se acostumbra representarla como  $\sigma_x$ :

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{\text{var}\{x\}} = \sqrt{E\{(x-\mu_x)^2\}} \quad (3.38)$$

Un concepto que se utiliza como alternativa al concepto de desviación estándar es el denominado desviación media:

$$\text{Desv\_Media} = E\{|x-\mu_x|\} \quad (3.39)$$

El valor absoluto en la expresión anterior es necesario, ya que de no hacerlo, en términos de la expresión 3.23, se haría cero.

Se le llama coeficiente de variación (CV) al cociente de la desviación estándar entre la media de la variable aleatoria:

$$CV = \frac{\sigma_x}{\mu_x} \quad (3.40)$$

El momento de orden tres con respecto a la media y su relación con los momentos con respecto al origen:

$$\mu_3 = E\{(x - \mu_x)^3\} = E\{x^3 - 3\mu_x x^2 + 3\mu_x^2 x - \mu_x^3\} = E\{x^3\} - 3\mu_x E\{x^2\} + 3\mu_x^2 E\{x\} - \mu_x^3$$

$$\mu_3 = E\{x^3\} - 3\mu_x [E\{x^2\} - \mu_x^2] - \mu_x^3$$

$$\mu_3 = \mu'_3 - 3\mu'_1 \sigma_x^2 - \mu_1^3 \quad (3.41)$$

Un concepto importante que se define en términos de los momentos de orden tres y de orden dos, con respecto a la media, es el coeficiente de asimetría, el cual representa la asimetría que presenta la gráfica de la función de probabilidad de la variable aleatoria  $x$  con respecto a su media:

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} \quad (3.42)$$

Las medidas de asimetría son indicadores que permiten establecer el grado de simetría (o asimetría) que presenta una distribución de probabilidad de una variable aleatoria, sin tener que hacer su representación gráfica. En la figura 3.16 se muestra el significado de la asimetría negativa y de la asimetría positiva.

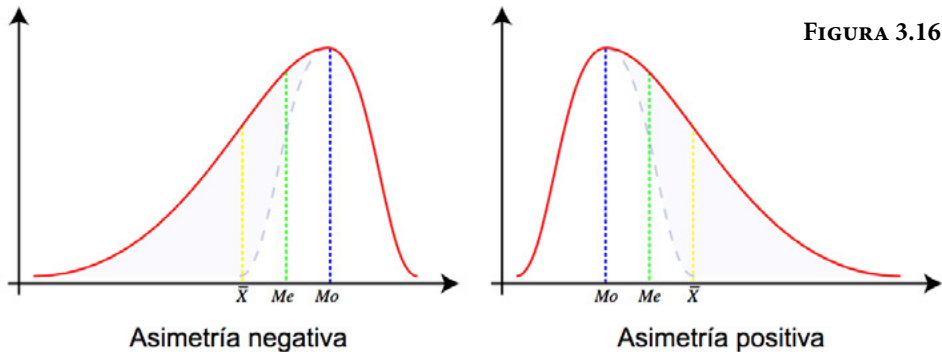


FIGURA 3.16

NOTA: Asimetría estadística. (2020).

[https://es.wikipedia.org/wiki/Asimetr%C3%ADa\\_estad%C3%ADstica](https://es.wikipedia.org/wiki/Asimetr%C3%ADa_estad%C3%ADstica)

El momento de orden cuatro con respecto a la media y su relación con los momentos con respecto al origen:

$$\mu_4 = E\{(x - \mu_x)^4\} = E\{x^4 - 4\mu_x x^3 + 6\mu_x^2 x^2 - 4\mu_x^3 x + \mu_x^4\}$$

$$\mu_4 = E\{x^4\} - 4\mu_x E\{x^3\} + 6\mu_x^2 E\{x^2\} - 4\mu_x^3 E\{x\} + \mu_x^4$$

$$\mu_4 = \mu_4' - 4\mu_1' \mu_3' + 6\mu_1'^2 \mu_2' - 3\mu_1'^4$$

$$\mu_4 = \mu_4' - 4\mu_1' \mu_3' + 6\mu_1'^2 \mu_2' - 3\mu_1'^4$$

$$\mu_4 = \mu_4' - 4\mu_1' \mu_3' + 6\mu_1'^2 [\mu_2' - \mu_1'^2] + 3\mu_1'^4$$

$$\mu_4 = \mu_4' - 4\mu_1' \mu_3' + 6\mu_1'^2 \sigma_x^2 + 3\mu_1'^4 \quad (3.43)$$

Otro concepto importante que se define en términos de los momentos de orden cuatro y dos con respecto a la media es el coeficiente de curtosis, el cual representa el grado de aplanamiento o “picudez” que presenta la gráfica de la función de probabilidad de la variable aleatoria  $x$ :

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 \quad (3.44)$$

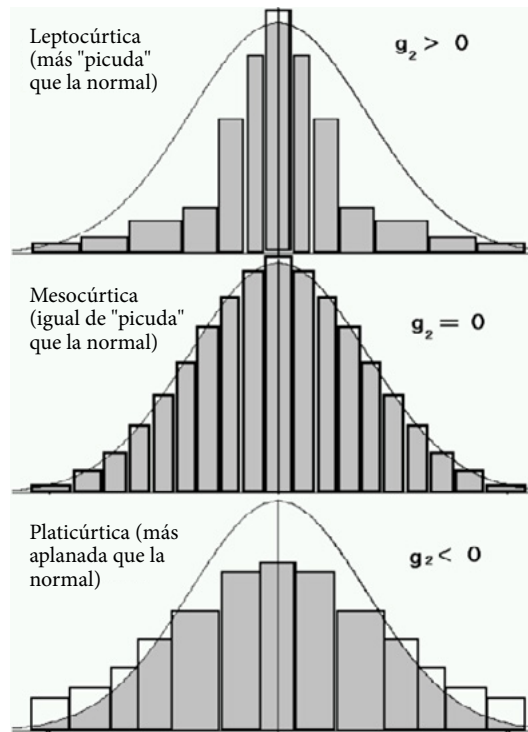


FIGURA 3.17.

**NOTA:** Medidas de curtosis (coeficiente de curtosis) (sf). <https://www.uv.es/ceaces/base/descriptiva/curtosis.htm>

En general, los momentos de orden  $k$  con respecto a la media pueden expresarse en términos de los momentos con respecto al origen, utilizando la siguiente expresión:

$$\mu_k = \sum_{j=0}^{j=k} (-1)^j \binom{k}{j} \mu^j \mu_{k-j}' \quad k=0, 1, 2, \dots \quad (3.45)$$

### 3.5. Función Generatriz de Momentos y Función Característica de una variable aleatoria

Calcular los momentos definidos anteriormente puede llegar a ser muy complicado, por lo cual, existe una forma alterna para determinarlos, a través de lo que se conoce como la Función Generatriz de Momentos:

$$FGM(t) = E\{e^{tx}\} = \left\{ \begin{array}{ll} \sum_{\forall x \in [a,b]} e^{tx} p(x) & \text{caso discreto} \\ \int_{\forall x \in [a,b]} e^{tx} f(x) dx & \text{caso continuo} \end{array} \right\} \quad (3.46)$$

Se le denomina Función Generatriz de Momentos a esta expresión porque al derivarla con respecto a  $t$  y evaluar en  $t=0$ , se obtienen los diversos momentos con respecto al origen:

$$\frac{d^k}{dt^k} FGM(t)|_{t=0} = E\{x^k e^{tx}\}_{t=0} = \mu'_k \quad (3.47)$$

Cabe señalar que de existir, la Función Generatriz de Momentos de una variable aleatoria es única, de conformidad con las propiedades de unicidad de una sumatoria y de una integral.

Sea  $x$  una variable aleatoria, se le denomina Función Característica de  $x$  a la expresión:

$$FC_x(t) = E\{e^{itx}\} = \left\{ \begin{array}{ll} \sum_{\forall x \in [a,b]} e^{itx} p(x) & \text{caso discreto} \\ \int_{\forall x \in [a,b]} e^{itx} f(x) dx & \text{caso continuo} \end{array} \right\} \quad (3.48)$$

Cabe señalar que de existir, la Función Característica de una variable aleatoria es única, de conformidad con las propiedades de unicidad de una sumatoria y de una integral.

## Ejercicio 3.7

Para el caso del ejercicio 3.1, la variable aleatoria definida como  $u = x_1 + x_2$ , la suma de los números de las caras superiores de dos dados que aparecen al ser lanzados, calcule:

- La media de  $x$ , es decir, el momento de primer orden con respecto al origen.
- La varianza, es decir, el momento de segundo orden con respecto a la media y la desviación estándar de  $x$ .
- La desviación media.
- El coeficiente de variación.
- El momento de tercer orden con respecto a la media y el coeficiente de asimetría.
- El momento de cuarto orden con respecto a la media y el coeficiente de curtosis.

Para obtener los valores solicitados, se usará Excel, con la siguiente tabla:

**FIGURA 3.18.** Resultados de lanzar dos dados y sus probabilidades

Evento	$u=x+y$	$p(u)$	$up(u)$	$ u-uprom $ $* p(u)$	$(u-uprom)^2$ $* p(u)$	$(u-uprom)^3$ $* p(u)$	$(u-uprom)^4$ $* p(u)$	
(1,1)	2	0.0278	0.0556	0.1389	0.6944	-3.4722	17.3611	
(1,2), (2,1)	3	0.0556	0.1667	0.2222	0.8889	-3.5556	14.2222	
(1,3), (3,1), (2,2)	4	0.0833	0.3333	0.2500	0.7500	-2.2500	6.7500	
(1,4), (4,1), (2,3), (3,2)	5	0.1111	0.5556	0.2222	0.4444	-0.8889	1.7778	
(1,5), (5,1), (4,2), (2,4), (3,3)	6	0.1389	0.8333	0.1389	0.1389	-0.1389	0.1389	
(1,6), (6,1), (2,5), (5,2), (3,4), (4,3)	7	0.1667	1.1667	0.0000	0.0000	0.0000	0.0000	
(2,6), (6,2), (3,5), (5,3), (4,4)	8	0.1389	1.1111	0.1389	0.1389	0.1389	0.1389	
(3,6), (6,3), (4,5), (5,4)	9	0.1111	1.0000	0.2222	0.4444	0.8889	1.7778	
(4,6), (6,4), (5,5)	10	0.0833	0.8333	0.2500	0.7500	2.2500	6.7500	
(5,6), (6,5)	11	0.0556	0.6111	0.2222	0.8889	3.5556	14.2222	
(6,6)	12	0.0278	0.3333	0.1389	0.6944	3.4722	17.3611	
			uprom=	7.0000	1.9444	5.8333	0.0000	80.5000

- Como se puede apreciar en la figura 3.18, la media es 7.
- La varianza es 5.8333, por lo que la desviación estándar es su raíz, es decir, la desviación estándar es 2.4152.



- c) La desviación media es 1.9444, la cual difiere de la desviación estándar 2.4152 como se puede apreciar.
- d) El coeficiente de variación sería  $CV = 2.4152/7 = 0.345$
- e) El momento de tercer orden con respecto a la media es 0. Note en la figura 3.29 la simetría de la función, es lógico que el tercer momento sea cero.  
Coeficiente de asimetría =  $0/(5.8333)^{(3/2)} = 0$
- f) El momento de cuarto orden es 80.5  
Coeficiente de curtosis =  $80.5/(5.8333)^2 - 3 = -0.634286$

Como se puede apreciar el coeficiente de curtosis es negativo lo que implica que la curva es platicúrtica o platocúrtica (un poco más plana que la curva normal estándar que más adelante se definirá).

### Ejercicio 3.8

Para el ejercicio 3.2 obtenga:

- a) La media de  $x$ , es decir, el momento de primer orden con respecto al origen.
- b) La varianza, es decir, el momento de segundo orden con respecto a la media y la desviación estándar de  $x$ .
- c) La desviación promedio.
- d) El coeficiente de variación.
- e) El momento de tercer orden con respecto a la media y el coeficiente de asimetría.
- f) El momento de cuarto orden con respecto a la media y el coeficiente de curtosis.

Para obtener los valores solicitados, se usará Excel, con la tabla que se muestra en la figura 3.19.

- a) Como se puede apreciar en la figura 3.19, la media es 1.4292.
- b) La varianza es 1.7274, por lo que su desviación estándar es 1.3143.
- c) La desviación promedio es 0.9454, que difiere de la desviación estándar 1.3143.
- d) El coeficiente de variación  $CV = 1.3143/1.4292 = 0.9196$ .
- e) El momento de tercer orden con respecto a la media es 4.2558.  
Coeficiente de asimetría =  $4.2558/(1.7274)^{(3/2)} = 1.8745$   
La curva de la función de probabilidad tiene asimetría positiva.
- f) El momento de cuarto orden es 18.6507  
Coeficiente de curtosis =  $18.6507/(1.7274)^2 - 3 = 3.2504$   
La curva de la función de probabilidad es leptocúrtica.

FIGURA 3.19.

$t$	$p(t)$	$tp(t)$	$(t-tprom)*p(t)$	$(t-tprom)^2p(t)$	$(t-tprom)^3p(t)$	$(t-tprom)^4p(t)$
0.1667	0.0278	0.0046	0.0351	0.0443	-0.0559	0.0706
0.2000	0.0278	0.0056	0.0341	0.0420	-0.0516	0.0634
0.2500	0.0278	0.0069	0.0328	0.0386	-0.0455	0.0537
0.3333	0.0556	0.0185	0.0609	0.0667	-0.0731	0.0801
0.4000	0.0278	0.0111	0.0286	0.0294	-0.0303	0.0312
0.5000	0.0833	0.0417	0.0774	0.0719	-0.0668	0.0621
0.6000	0.0278	0.0167	0.0230	0.0191	-0.0158	0.0131
0.6667	0.0556	0.0370	0.0424	0.0323	-0.0246	0.0188
0.7500	0.0278	0.0208	0.0189	0.0128	-0.0087	0.0059
0.8000	0.0278	0.0222	0.0175	0.0110	-0.0069	0.0044
0.8333	0.0278	0.0231	0.0166	0.0099	-0.0059	0.0035
1.0000	0.1667	0.1667	0.0715	0.0307	-0.0132	0.0057
1.2000	0.0278	0.0333	0.0064	0.0015	-0.0003	0.0001
1.2500	0.0278	0.0347	0.0050	0.0009	-0.0002	0.0000
1.3333	0.0278	0.0370	0.0027	0.0003	0.0000	0.0000
1.5000	0.0556	0.0833	0.0039	0.0003	0.0000	0.0000
1.6667	0.0278	0.0463	0.0066	0.0016	0.0004	0.0001
2.0000	0.0833	0.1667	0.0476	0.0272	0.0155	0.0088
2.5000	0.0278	0.0694	0.0297	0.0319	0.0341	0.0365
3.0000	0.0556	0.1667	0.0873	0.1371	0.2153	0.3383
4.0000	0.0278	0.1111	0.0714	0.1836	0.4720	1.2134
5.0000	0.0278	0.1389	0.0992	0.3542	1.2648	4.5162
6.0000	0.0278	0.1667	0.1270	0.5803	2.6527	12.1249
		1.4292	0.9454	1.7274	4.2558	18.6507

## Ejercicio 3.9

Sea  $x$  la variable aleatoria discreta, cuya función de probabilidad es:

$$p(x) = \left\{ \begin{array}{ll} \frac{1}{2^{x+1}} & x= 0, 1, 2, 3, \dots \\ 0 & \text{cualquier otro caso} \end{array} \right\}$$

- i. Verifique que cumple las propiedades de una función de probabilidad.

Observe que  $1/2^{x+1} > 0$

Por otra parte, su función de probabilidad acumulada es:

$$P(x) = p(t \leq x) = \sum_{t=0}^{t=x} \frac{1}{2^{x+1}}$$

$$P(x) = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots + \frac{1}{2^{x+1}}$$

Si esta última expresión se divide en ambos lados entre dos y se restan entre sí  $(P(x) - P(x)/2)$ , se obtiene que:

$$\frac{1}{2}P(x) = \frac{1}{2} - \frac{1}{2^{x+1}}$$

$$P(x) = 1 - \frac{1}{2^{x+1}}$$

Observe en esta última expresión que si  $x$  tiende a infinito,  $P(x)$  tiende a uno con lo cual se cumple que la suma de todas las probabilidades debe ser uno.

- ii. Calcule las siguientes probabilidades:

a)  $p(x=2) = 1/2^3 = 1/8 = 0.125$

b)  $p(x < 2) = p(x=0) + p(x=1) = 0.5 + 0.25 = 0.75$

De otra forma,  $p(x < 2) = P(x=1) = 1 - 1/2^2 = 0.75$

c)  $p(x > 2) = 1 - P(x=2) = 1 - (1 - 1/2^3) = 0.125$

- iii. Calcule la Función Generatriz de Momentos.

$$FGM(t) = E\{e^{tx}\} = \sum_{x=0}^{x \rightarrow \infty} \frac{e^{tx}}{2^{x+1}} = \frac{1}{2} \sum_{x=0}^{x \rightarrow \infty} \left(\frac{e^t}{2}\right)^x$$

$$2FGM(t) = 1 + \frac{e^t}{2} + \left(\frac{e^t}{2}\right)^2 + \left(\frac{e^t}{2}\right)^3 + \dots$$

Observe que:

$$S_n = 1 + \frac{e^t}{2} + \left(\frac{e^t}{2}\right)^2 + \left(\frac{e^t}{2}\right)^3 + \dots + \left(\frac{e^t}{2}\right)^n$$

Si esta última expresión se multiplica por  $(e^t/2)$  de ambos lados y se restan entre sí:

$$\left(1 - \frac{e^t}{2}\right) S_n = 1 - \left(\frac{e^t}{2}\right)^{n+1}$$

$$S_n = \frac{1 - \left(\frac{e^t}{2}\right)^{n+1}}{\left(1 - \frac{e^t}{2}\right)}$$

Observe que si  $t < 0$ , el límite de  $S_n$  cuando  $n$  tiende a infinito es  $1/(1 - e^t/2)$ , por lo que la función generatriz de momentos queda definida como:

$$FGM(t) = \frac{1}{2 - e^t} \quad \forall t < 0$$

iv. La media de  $x$ , es decir, el momento de primer orden con respecto al origen.

$$\mu_x = \frac{d}{dt} FGM(t) = \frac{d}{dt} \left( \frac{1}{2 - e^t} \right) \Big|_{t=0} = \frac{e^t}{(2 - e^t)^2} \Big|_{t=0} = 1 \quad \forall t < 0$$

v. La varianza, es decir, el momento de segundo orden con respecto a la media y la desviación estándar:

$$E\{x^2\} = \frac{d^2}{dt^2} FGM(t) = \frac{d^2}{dt^2} \left( \frac{1}{2 - e^t} \right) \Big|_{t=0} = \frac{(2 + e^t)e^t}{(2 - e^t)^3} \Big|_{t=0} = 3 \quad \forall t < 0$$

$$\sigma_x^2 = E\{x^2\} - E^2\{x\} = 3 - 1 = 2$$

La desviación estándar sería raíz(2).

vi. La desviación media sería:

$$desv\_media = E\{|x-\mu_x|\} = \sum_{x=0}^{x \rightarrow \infty} |x-1| / 2^{(x+1)}$$

$$desv\_media = \frac{1}{2} + \frac{0}{2^2} + \frac{1}{2^3} + \frac{2}{2^4} + \frac{3}{2^5} + \dots$$

$$desv\_media = \frac{1}{2} + \frac{1}{2^2} \left[ \frac{0}{2^0} + \frac{1}{2^1} + \frac{2}{2^2} + \frac{3}{2^3} + \frac{4}{2^4} + \dots \right]$$

Pero

$$\frac{1}{2^1} + \frac{2}{2^2} + \frac{3}{2^3} + \frac{4}{2^4} + \dots = 2$$

Por lo que

$$desv\_media = \frac{1}{2} + \frac{1}{2^2} [2] = 1$$

Nótese que la desviación media es diferente a la desviación estándar.

vii. El coeficiente de variación sería  $CV = \text{raíz}(2)/1 = \text{raíz}(2) = 1.4142$ .

viii. El momento de tercer orden con respecto a la media y el coeficiente de asimetría.

$$E\{x^3\} = \frac{d^3}{dt^3} FGM(t) = \frac{d^3}{dt^3} \left( \frac{1}{2-e^t} \right) \Big|_{t=0} = \frac{(8 + 4e^t - e^{2t})e^t}{(2-e^t)^4} \Big|_{t=0} = 13 \quad \forall t < 0$$

$$\gamma_1 = \frac{\mu'_3 - 3\mu'_1\sigma_x^2 - \mu_1^3}{\mu_2^{3/2}} = \frac{13 - 3(1)(2) - 1}{2^{3/2}} = \frac{6}{2^{3/2}} = 2.12132$$

ix. El momento de cuarto orden con respecto a la media y el coeficiente de curtosis.

$$E\{x^4\} = \frac{d^4}{dt^4} FGM(t) = \frac{d^4}{dt^4} \left( \frac{1}{2-e^t} \right) \Big|_{t=0} = \frac{8e^{2t} - 44e^{3t} + 22e^{4t} + 1e^{5t}}{(2-e^t)^4} \Big|_{t=0} = 75 \quad \forall t < 0$$

$$g_2 = \frac{\mu'_4}{\mu_2^2} - 3 = \frac{\mu'_4 - 4\mu'_1\mu'_3 + 4\mu_1^2\mu'_2 + 3\mu_1^4}{\mu_2^2} - 3 = \frac{75 - 4(1)(13) + 4(1)(2) + 3(1)}{4} - 3 = 5.5$$

x. La función característica de momentos de la variable  $x$ .

$$FC_x(t) = E\{e^{itx}\} = \sum_{x=0}^{x \rightarrow \infty} \frac{e^{itx}}{2^{x+1}} = \frac{1}{2} \sum_{x=0}^{x \rightarrow \infty} \left(\frac{e^{it}}{2}\right)^x$$

$$2FC_x(t) = \frac{e^{it}}{2} + \left(\frac{e^{it}}{2}\right)^2 + \left(\frac{e^{it}}{2}\right)^3 + \dots$$

Observe que:

$$Sn = 1 + \frac{e^{it}}{2} + \left(\frac{e^{it}}{2}\right)^2 + \left(\frac{e^{it}}{2}\right)^3 + \dots + \left(\frac{e^{it}}{2}\right)^n$$

Si esta última expresión se multiplica por  $(e^{it}/2)$  de ambos lados y se restan entre sí:

$$\left(1 - \frac{e^{it}}{2}\right) Sn = 1 - \left(\frac{e^{it}}{2}\right)^{n+1}$$

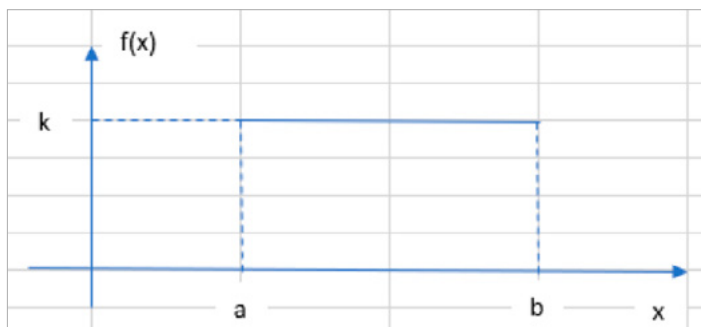
$$Sn = \frac{1 - \left(\frac{e^{it}}{2}\right)^{n+1}}{\left(1 - \frac{e^{it}}{2}\right)}$$

Observe que si  $t < 0$ , el límite de  $Sn$  cuando  $n$  tiende a infinito es  $1/(1 - e^{it}/2)$ , por lo que la función generatriz de momentos queda definida como:

$$FGM(t) = \frac{1}{2 - e^{it}} \quad \forall t < 0$$

### Ejercicio 3.10

Sea  $x$  la variable aleatoria uniforme continua, cuya función de probabilidad se muestra en la figura 3.20.



**FIGURA 3.20.**  
Función de Probabilidad Uniforme

- i. Determine el valor de  $k$  para que se considere una función de probabilidad.

Para que una función  $f(x)$  se considere función de probabilidad debe ser no negativa, lo que implica que  $k \geq 0$ . Además, el área bajo la curva en el intervalo de definición siempre debe ser uno, por lo que:

$$k(b-a) = 1 \quad \rightarrow \quad k = 1/(b-a)$$

- ii. Obtenga la expresión matemática de la función de probabilidad uniforme.

$$f(x) = \left\{ \begin{array}{ll} \frac{1}{b-a} & x \in [a, b] \\ 0 & \text{cualquier otro caso} \end{array} \right\} \quad (3.49)$$

- iii. Calcule la función de probabilidad acumulada.

$$F(x) = \int_a^x f(t) dt$$

$$F(x) = \left\{ \begin{array}{ll} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a, b] \\ 1 & x < b \end{array} \right\} \quad (3.50)$$

Nótese que  $F(x)$  es el área de un rectángulo de base  $(x-a)$  y de altura  $1/(b-a)$

- iv. Obtenga su media, mediana y moda

$$\mu_x = \int_a^b x f(x) dx = \frac{1}{b-a} \int_a^b x dx = \frac{x^2}{2(b-a)} \Big|_a^b$$

$$\mu_x = \frac{b^2 - a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(b-a)}$$

$$\mu_x = \frac{b+a}{2} \quad (3.51)$$

Nótese en la figura 3.20 que la mediana es igual a la media. La función es multimodal, cada punto del intervalo tiene la misma altura, por lo que no existe moda.

v. Determine su varianza y su desviación estándar.

$$E\{x^2\} = \int_a^b xf(x)dx = \frac{1}{b-a} \int_a^b x^2 dx = \frac{x^2|_a^b}{3(b-a)}$$

$$E\{x^2\} = \frac{b^3-a^3}{3(b-a)} = \frac{b^2+ab+a^2}{3}$$

$$\sigma_x^2 = \frac{b^2+ab+a^2}{3} - \frac{(b+a)^2}{4}$$

$$\sigma_x^2 = \frac{4b^2+4ab+4a^2-3b^2-6ab-3a^2}{12}$$

Por lo que su varianza es:

$$\sigma_x^2 = \frac{(b-a)^2}{12} \quad (3.52)$$

Su desviación estándar sería:

$$\sigma_x = \frac{(b-a)}{\sqrt{12}} \quad (3.53)$$

vi. Su coeficiente de variación sería:

$$CV = \frac{\sigma_x}{\mu_x} = \frac{\frac{(b-a)}{\sqrt{12}}}{\frac{(b+a)}{2}} = \frac{(b-a)}{\sqrt{3}(b+a)} \quad (3.54)$$

vii. Calcule su coeficiente de asimetría.

$$\mu_3 = E\{(x-\mu_x)^3\} = \frac{1}{b-a} \int (x-\mu_x)^3 dx$$

$$\mu_3 = \frac{1}{3(b-a)} (x-\mu_x)^4 \Big|_a^b = \frac{[(b-\mu_x)^4 - (a-\mu_x)^4]}{3(b-a)}$$

$$\text{Pero } b - \frac{b+a}{2} = \frac{b-a}{2} \quad a - \frac{b+a}{2} = -\frac{b-a}{2}$$



$$\mu_3 = 0$$

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = 0 \quad (3.55)$$

viii. Obtenga su coeficiente de curtosis.

$$\mu_4 = E\{(x-\mu_x)^4\} = \frac{1}{b-a} \int (x-\mu_x)^4 dx$$

$$\mu_4 = \frac{1}{5(b-a)} (x-\mu_x)^5 \Big|_a^b = \frac{[(b-\mu_x)^5 - (a-\mu_x)^5]}{5(b-a)}$$

$$\text{Pero } b - \frac{b+a}{2} = \frac{b-a}{2} \quad a - \frac{b+a}{2} = -\frac{b-a}{2}$$

$$\mu_4 = \frac{(b-a)^5}{5(2^4)}$$

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} = \frac{\frac{(b-a)^5}{5(2^4)}}{\frac{(b-a)^4}{144}} - 3 = \frac{9}{5}(b-a) - 3 \quad (3.56)$$

ix. Calcule su función generatriz de momentos.

$$FGM(t) = E\{e^{tx}\} = \frac{1}{(b-a)} \int_a^b e^{tx} dx = \left[ \frac{e^{tx}}{t(b-a)} \right]_a^b$$

$$FGM(t) = \frac{e^{bt} - e^{at}}{t(b-a)} \quad (3.57)$$

x. Obtenga su Función Característica

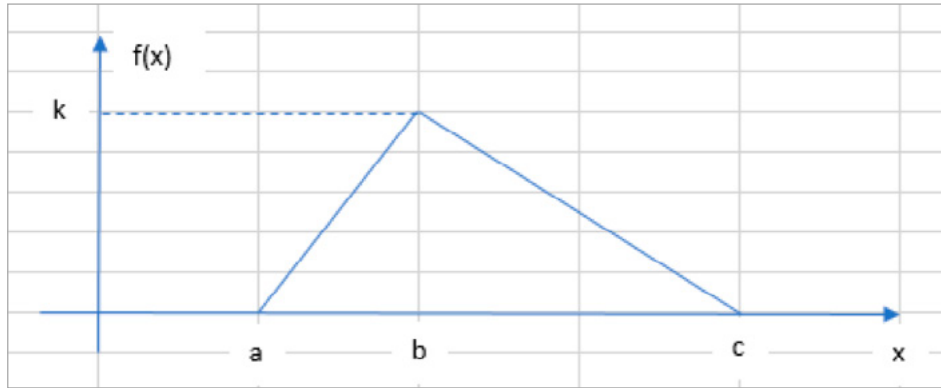
$$FC_x(t) = E\{e^{itx}\} = \frac{1}{(b-a)} \int_a^b e^{itx} dx = \left[ \frac{e^{itx}}{it(b-a)} \right]_a^b$$

$$FC_x(t) = \frac{e^{ibt} - e^{iat}}{it(b-a)} \quad (3.58)$$

## Ejercicio 3.11

Sea  $x$  la variable aleatoria con distribución triangular, cuya función de probabilidad se muestra en la figura 3.21.

FIGURA 3.21. Distribución de probabilidad triangular



- a) Determine el valor de  $k$  para que se considere una función de probabilidad.

Para que una función  $f(x)$  se considere función de probabilidad debe ser no negativa, lo que implica que  $k \geq 0$ . Además, el área bajo la curva en el intervalo de definición siempre debe ser uno, por lo que, calculando el área del triángulo e igualando a uno:

$$k(c-a)/2 = 1 \quad \rightarrow \quad k = 2/(c-a)$$

- b) Obtenga la expresión matemática de la función de probabilidad triangular.

Como se aprecia en la figura 3.21 la función presenta dos ramas que son rectas. La primera recta pasa por los puntos  $(a, 0)$  y  $(b, k)$ . La segunda recta pasa por los puntos  $(b, k)$  y  $(c, 0)$ , por lo que para obtener su función de probabilidad se utilizará la fórmula de la ecuación de una recta dados dos de sus puntos:

$$y = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) + y_1$$

Por lo que la primera rama está dada por:

$$y = \frac{k - 0}{b - a}(x - a) + 0$$

$$y = \frac{2(x-a)}{(c-a)(b-a)}$$

La segunda rama está dada por:

$$y = \frac{0-k}{(c-b)}(x-b) + k$$

$$y = \frac{2}{(c-a)} - \frac{2(x-b)}{(c-a)(c-b)} = \frac{2(c-x)}{(c-a)(c-b)}$$

Por lo que la función de probabilidad triangular está dada por la siguiente expresión:

$$f(x) = \left\{ \begin{array}{ll} \frac{2(x-a)}{(c-a)(b-a)} & a \leq x \leq b \\ \frac{2(c-x)}{(c-a)(c-b)} & b < x \leq c \\ 0 & \text{para cualquier otro valor} \end{array} \right\} \quad (3.59)$$

- i. Calcule la función de probabilidad acumulada.

$$F(x) = \left\{ \begin{array}{ll} 0 & x < a \\ \frac{(x-a)^2}{(c-a)(b-a)} & a \leq x \leq b \\ 1 - \frac{(c-x)^2}{(c-a)(c-b)} & b < x \leq c \\ 1 & x > c \end{array} \right\} \quad (3.60)$$

Nótese que la segunda rama de  $F(x)$  es el área del triángulo formado por los puntos  $(a, 0)$ ,  $(x, 0)$  y  $(x, h)$ , donde  $h = 2(c-x)/[(c-a)(c-b)]$ , la cual se obtuvo por triángulos semejantes  $h/H = b/B$ . La tercera rama se obtuvo como uno menos el área del triángulo  $(x, 0)$ ,  $(x, h)$ ,  $(c, 0)$  en la segunda rama y también aplicando la relación entre triángulos semejantes.

ii. Calcule la media de la distribución triangular.

$$\mu_x = E\{x\} = \frac{2}{(c-a)(b-a)} \int_a^b x(x-a) dx + \frac{2}{(c-a)(c-b)} \int_b^c x(c-x) dx$$

$$\mu_x = \frac{2}{(c-a)(b-a)} \left[ \frac{x^3}{3} - a \frac{x^2}{2} \right]_a^b + \frac{2}{(c-a)(c-b)} \left[ c \frac{x^2}{2} - \frac{x^3}{3} \right]_b^c$$

$$\mu_x = \frac{2}{(c-a)(b-a)} \left[ \frac{b^3}{3} - \frac{ab^2}{2} - \frac{a^3}{3} + \frac{a^3}{2} \right] + \frac{2}{(c-a)(c-b)} \left[ \frac{c^3}{2} - \frac{c^3}{3} - \frac{b^2c}{2} + \frac{b^3}{3} \right]$$

$$\mu_x = \frac{2}{(c-a)(c-b)} \left[ \frac{(b-a)(b^2+ab+a^2)}{3} - \frac{a}{2}(b-a)(b+a) \right] +$$

$$\frac{2}{(c-a)(c-b)} \left[ \frac{c}{2}(c-b)(c+b) - \frac{(c-b)(c^2+bc+b^2)}{3} \right]$$

$$\mu_x = \frac{1}{3(c-a)} \left[ 2b^2 + 2ab + 2a^2 - 3ab - 3a^2 + 3c^2 + 3bc - 2c^2 - 2bc - 2b^2 \right]$$

$$\mu_x = \frac{1}{3(c-a)} \left[ (c-a)(a+b+c) \right]$$

La media de la distribución triangular es:

$$\mu_x = \frac{a+b+c}{3} \tag{3.61}$$

iii. Calcule la mediana y la moda de la distribución triangular.

En este caso la mediana dependerá de la asimetría del triángulo definido en la figura 3.21, y existirán dos casos diferentes:

Si la primera rama corresponde con un triángulo mayor que el de la segunda rama, entonces la mediana estaría situada en la primera rama, y para obtener la mediana se iguala la función de probabilidad acumulada de la primera rama a 0.5, y se despeja el valor de  $x$ , es decir:

$$\frac{2(x-a)}{(c-a)(b-a)} = 0.5$$

$$x = \frac{(c-a)(b-a)}{4} + a$$

Si la segunda rama corresponde con un triángulo mayor que el de la primera rama, entonces la mediana estaría situada en la segunda rama, y para obtener la mediana se iguala la función de probabilidad acumulada de la segunda rama a 0.5, y se despeja el valor de  $x$ . es decir:

$$1 - \frac{(c-x)^2}{(c-a)(c-b)} = 0.5$$

$$(c-x)^2 = 0.5(c-a)(c-b)$$

$$x = c - \sqrt{0.5(c-a)(c-b)}$$

De tal forma que la mediana sería:

$$m_e = \begin{cases} a + \sqrt{0.5(c-a)(c-b)} & (b-a) \geq (c-b) \\ b + \sqrt{0.5(c-a)(c-b)} & (b-a) < (c-b) \end{cases} \quad (3.62)$$

La moda de la distribución triangular sería el punto más alto de la función de probabilidad, la cual se encuentra exactamente en

$$m_o = b \quad (3.63)$$

iv. Calcule la varianza y la desviación estándar de la distribución triangular.

$$\sigma_x^2 = E\{x^2\} - E^2\{x\}$$

$$E\{x^2\} = \frac{2}{(c-a)(b-a)} \int_a^b x^2(x-a)dx + \frac{2}{(c-a)(c-b)} \int_b^c x^2(c-x)dx$$

$$E\{x^2\} = \frac{2}{(c-a)(b-a)} \left[ \frac{x^4}{4} - a \frac{x^3}{3} \right]_a^b + \frac{2}{(c-a)(c-b)} \left[ c \frac{x^3}{3} - \frac{x^4}{4} \right]_b^c$$

$$E\{x^2\} = \frac{2}{(c-a)(b-a)} \left[ \frac{b^4}{4} - \frac{ab^3}{3} - \frac{a^4}{4} + \frac{a^4}{3} \right] + \frac{2}{(c-a)(c-b)} \left[ \frac{c^4}{3} - \frac{c^4}{4} - \frac{bc^3}{3} + \frac{b^4}{4} \right]$$

$$E\{x^2\} = \left[ \frac{3b^4 - 4ab^3 - 3a^4 + 4a^4}{6(c-a)(b-a)} \right] + \left[ \frac{4c^4 - 3c^4 - 4bc^3 + 3b^4}{6(c-a)(c-b)} \right]$$

$$E\{x^2\} = \left[ \frac{3b^3 - ab^2 - a^2b - a^3 + c^3 - 3bc^2 - 3b^2c - 3b^3}{6(c-a)} \right]$$

$$E\{x^2\} = \left[ \frac{-ab^2 - a^2b - a^3 + c^3 - 3bc^2 - 3b^2c}{6(c-a)} \right]$$

$$\sigma_x^2 = E\{x^2\} - E^2\{x\} = \left[ \frac{-ab^2 - a^2b - a^3 + c^3 - 3bc^2 - 3b^2c}{6(c-a)} \right] - \left[ \frac{a+b+c}{3} \right]^2$$

$$\sigma_x^2 = \frac{a^2 + b^2 + c^2 - ab - ac - bc}{18}$$

La varianza de la distribución triangular es:

$$\sigma_x^2 = \frac{a^2 + b^2 + c^2 - ab - ac - bc}{18} \quad (3.64)$$

Su desviación estándar sería:

$$\sigma_x = \frac{\sqrt{a^2 + b^2 + c^2 - ab - ac - bc}}{3\sqrt{2}} \quad (3.65)$$

v. Calcule el coeficiente de variación de la distribución triangular:

$$CV = \frac{\sigma_x}{\mu_x} = \frac{\sqrt{a^2 + b^2 + c^2 - ab - ac - bc}}{\sqrt{2}(a+b+c)} \quad (3.66)$$

### 3.6. Teorema de la Desigualdad de Cheyvshev

Sea  $x$  una variable aleatoria con distribución de probabilidad discreta o continua, con media  $\mu_x$  y desviación estándar  $\sigma_x$ , y sea  $k$  un número positivo. Entonces:

$$p(|x - \mu_x| \geq k\sigma_x) \leq \frac{1}{k^2} \quad (3.67)$$

Se hará la demostración para el caso de una variable aleatoria continua, el caso discreto se resuelve de forma similar.

Demostración:

Por definición de varianza:

$$\begin{aligned} \sigma_x^2 &= \int_{-\infty}^{\infty} (x - \mu_x)^2 f(x) dx \\ \sigma_x^2 &= \int_{-\infty}^{\mu_x - \sqrt{k}} (x - \mu_x)^2 f(x) dx + \int_{\mu_x - \sqrt{k}}^{\mu_x + \sqrt{k}} (x - \mu_x)^2 f(x) dx + \int_{\mu_x + \sqrt{k}}^{\infty} (x - \mu_x)^2 f(x) dx \end{aligned}$$

Nótese que:

$$\int_{\mu_x - \sqrt{k}}^{\mu_x + \sqrt{k}} (x - \mu_x)^2 f(x) dx \geq 0$$

Por lo que:

$$\sigma_x^2 \geq \int_{-\infty}^{\mu_x - \sqrt{k}} (x - \mu_x)^2 f(x) dx + \int_{\mu_x + \sqrt{k}}^{\infty} (x - \mu_x)^2 f(x) dx$$

Por otra parte, por propiedades de las desigualdades,  $(x - \mu_x)^2 \geq K$  si y solo si  $|x - \mu_x| \geq \text{raíz}(K)$ , por lo que se mantiene la desigualdad si se sustituye  $(x - \mu_x)^2$  por  $K$  en la expresión anterior:

$$\begin{aligned} \sigma_x^2 &\geq \int_{-\infty}^{\mu_x - \sqrt{k}} K f(x) dx + \int_{\mu_x + \sqrt{k}}^{\infty} K f(x) dx \\ \sigma_x^2 &\geq K [ p(x \leq \mu_x - \sqrt{K}) + p(x \leq \mu_x + \sqrt{K}) ] \end{aligned}$$

Expresión que también puede escribirse como:

$$p(|x - \mu_x| \geq \sqrt{K}) \leq \frac{\sigma_x^2}{K}$$

Si se hace  $k=\text{raíz}(K)/\sigma_x$ , entonces:

$$p(|x - \mu_x| \geq k\sigma_x) \leq \frac{1}{k^2}$$

Esta desigualdad también puede escribirse de la siguiente forma:

$$p(|x - \mu_x| < k\sigma_x) \geq 1 - \frac{1}{k^2}$$

$$p(\mu_x - k\sigma_x \leq x \leq \mu_x + k\sigma) \geq 1 - \frac{1}{k^2} \quad (3.68)$$

### Ejercicio 3.12

La revisión de una convocatoria de concurso de oposición abierto (COA), en la UNAM, la realiza la Dirección General de Asuntos del Personal Académico (DGAPA), la cual requiere en promedio dos días. Por experiencias previas, la varianza se estima en 0.25 días. Si se desea que el 95% de las convocatorias se revise a tiempo, ¿con qué anticipación deben enviarse a la DGAPA?

Nótese que, en este caso en particular, se desconoce la distribución de probabilidad del tiempo de revisión, solo se conoce su media y su varianza. En estos casos es cuando es aplicable la desigualdad de Chevyshev.

De la ecuación 3.95

$$1 - 1/k^2 = 0.95 \rightarrow k > 4.4721$$

Según la expresión 3.95, el tiempo de envío está entre  $2 - 4.47 \cdot \text{raíz}(0.25)$  y  $2 + 4.47 \cdot \text{raíz}(0.25) \rightarrow 0 < t < 4.236$

Se debe enviar a la DGAPA con un mínimo de 4.236 días antes.



### 3.7. Funciones de probabilidad de variables dependientes de variables aleatorias

La mayoría de los problemas en Ingeniería involucran el estudio de relaciones funcionales de una variable dependiente con una o más variables independientes. Si una o más de las variables independientes son a su vez aleatorias esto repercutirá en que a su vez la variable dependiente sea aleatoria. Un punto importante es conocer el modelo probabilístico de la variable dependiente a partir del modelo probabilístico de las variables independientes. Se abordará el caso, comenzando con funciones de una sola variable independiente.

Sea  $x$  una variable aleatoria discreta con función de probabilidad  $p_x(x)$  y sea  $y=h(x)$  una relación funcional entre la variable aleatoria  $x$  y la variable dependiente  $y$ . Dado que  $y$  depende de  $x$ , esto implicará que  $y$  también sea aleatoria con cierto tipo de función de probabilidad desconocida  $p_y(y)$ . La pregunta es cómo deducir la función de probabilidad  $p_y(y)$  a partir de la función  $p_x(x)$ .

El análisis para el caso de una variable aleatoria discreta se ilustrará con un ejemplo:

#### Ejercicio 3.13

Sea  $x$  una variable aleatoria discreta cuya función de probabilidad se muestra a continuación. Sea  $y=4x^2$ . Obtenga la función de probabilidad de  $y$ , su media, su desviación estándar, su desviación promedio, su coeficiente de asimetría y su coeficiente de curtosis.

$x$	$p(x)$
0	0.5729292
1	0.35808075
2	0.06510559
3	0.00382974
4	5.4711E-05

Se construye la tabla de la figura 3.22

FIGURA 3.22.

$x$	$p(x)$	$y=4x^2$	$py(y)$	$yp(y)$	$ y-y_{prom} p(y)$	$(y-y_{prom})^2p(y)$	$(y-y_{prom})^3p(y)$	$(y-y_{prom})^4p(y)$
0	0.5729	0	0.5729	0.0000	1.4984	3.9190	-10.2496	26.8067
1	0.3581	4	0.3581	1.4323	0.4958	0.6865	0.9505	1.3161
2	0.0651	16	0.0651	1.0417	0.8714	11.6635	156.1119	2,089.4976
3	0.0038	36	0.0038	0.1379	0.1279	4.2684	142.4979	4,757.2380
4	0.0001	64	0.0001	0.0035	0.0034	0.2062	12.6546	776.8005
	Suma=		1	2.6154	2.9969	20.7435	301.9654	7,651.6588

La tercera y cuarta columna de la tabla de la figura 3.22 conforman la función de probabilidad de  $y$ .

Media	=	2.6154
Varianza	=	20.7435
Desv Est	=	4.5545
Desv Media	=	2.9969
Coef Var	=	1.7414
Coef Asim	=	3.1962
Coef Curt	=	14.7824

Para el caso de una variable aleatoria continua  $x$ , cuya función de probabilidad acumulada es  $F(x) = P_x(t \leq x)$ , considere por el momento que  $y = h(x)$  es una función monótona (creciente o decreciente), de tal manera que la inversa  $x = h^{-1}(y)$  también es función. La función de probabilidad acumulada de  $y$  está dada por  $F(y) = P_y(t \leq y) = P_x(t \leq x) = P_x(t \leq h^{-1}(y))$ , de tal manera que la función de densidad de  $y$ , está dada por la derivada de esta expresión, lo que implica que:

$$f_y(y) = f_x(x) \left| \frac{dx}{dy} \right| \quad (3.69)$$

#### Ejercicio 3.14

Un ingeniero industrial de la Facultad de Ingeniería de la UNAM ofrece llevar a cabo el diseño de un proceso productivo, para lo cual, con base en su experiencia previa en proyectos semejantes, sabe que el tiempo requerido para su terminación sigue la distribución de probabilidad definida en la tabla siguiente:

$t$	10	11	12	13	14
$p(t)$	0.1	0.3	0.4	0.1	0.1

La ganancia o utilidad por la realización de este proyecto depende del tiempo que se demore en entregar el diseño y está modelado por  $u = 40000(12 - t)$ . Obtenga la distribución de probabilidad de la utilidad y calcule la utilidad promedio necesaria para emprender este proyecto.

$t$	$p(t)$	$u$	$p(u)$	$up(u)$
10	0.1	80000	0.1	8000
11	0.3	40000	0.3	12000

$t$	$p(t)$	$u$	$p(u)$	$up(u)$
12	0.4	0	0.4	0
13	0.1	-40000	0.1	-4000
14	0.1	-80000	0.1	-8000
		Suma=	1	8000

La función de probabilidad está definida en las columnas tres y cuatro. La utilidad media que él recibe es de \$8,000.00.

## Ejercicio 3.15

Sea  $x$  una variable aleatoria con función de probabilidad:

$$f_x(x) = 2x \cos(x^2) \quad 0 \leq x \leq \sqrt{\frac{\pi}{2}}$$

Determine la función de probabilidad de  $y=x^2/2$ .

La función de probabilidad acumulada de  $f_x(x)$  se obtiene como una integral:

$$F(x) = \int_0^x 2t \cos(t^2) dt = [\text{sen}(t^2)]_0^x = \text{sen}(x^2)$$

Despejando  $x$  de la expresión  $y=x^2/2$  y sustituyendo en  $f_x(x)$ :

$$x = \pm \sqrt{2y}$$

$$f_x(x) = f_x(\pm \sqrt{2y}) = 2(\sqrt{2y}) = \cos(2y)$$

Nótese que la  $x$  es positiva en el intervalo dado. También observe que para  $x=0 \rightarrow y=0$ ; y que para  $x=\text{raíz}(\pi/2) \rightarrow y=\pi/4$

Derivando  $x$  con respecto a  $y$ :

$$x = \pm \sqrt{2y}$$

$$\frac{dx}{dy} = \pm \frac{1}{\sqrt{2y}}$$

$$\left| \frac{dx}{dy} \right| = \frac{1}{\sqrt{2y}}$$

Por lo que la función de probabilidad de  $y$  es:

$$f_y(y) = 2(\sqrt{2y}) \cos(2y) \frac{1}{\sqrt{2y}}$$

$$f_y(y) = 2\cos(2y) \quad 0 \leq y \leq \frac{\pi}{4}$$

### Ejercicio 3.16

(Problema 3–8, página 82 y problema 7–15, página 203, Hines, Montgomery, Goldsman y Borror, *Probabilidad y Estadística para Ingeniería*, Cuarta Edición, Grupo Patria, 2013). Las especificaciones para un resistor indican que la vida útil del mismo debe ser mayor de 1000 horas. Se sabe que la vida útil del resistor se distribuye con una función del tipo

$$f(x) = \frac{1}{\theta} e^{-\left(\frac{x-\beta}{\theta}\right)} \quad x \geq \beta, \theta > 0 \quad (3.70)$$

A la cual se le conoce como la distribución exponencial de dos parámetros. La utilidad estimada por cada resistor es de \$9.00; sin embargo, una unidad con fallas debe reemplazarse a un costo de \$3.00 para la compañía. Dos procesos de manufactura pueden producir resistores que tengan una media de vida útil satisfactoria. El proceso  $A$  presenta un parámetro  $\theta_A=2000$  con una  $\beta_A=750$ ; el proceso  $B$  presenta un parámetro  $\theta_B=2500$  con una  $\beta_B=500$ ; sin embargo, los costos de manufactura del proceso  $A$  son solo la mitad de los del proceso  $B$ . ¿Qué valor del costo del proceso de manufactura es crítico, al grado de imponer el uso del proceso  $A$  o el  $B$ ?

Lo primero por realizar es formular la función de utilidad por cada resistor:

Si el resistor se fabrica con el proceso  $A$ , su función de utilidad neta sería:

$$u_A = \begin{cases} 6 - c_A & t < 1000 \\ 9 - c_A & t \geq 1000 \end{cases}$$

Si el resistor se fabrica con el proceso  $B$ , su función de utilidad neta sería:

$$u_B = \begin{cases} 6 - 2c_A & t < 1000 \\ 9 - 2c_A & t \geq 1000 \end{cases}$$

La utilidad neta esperada, con el proceso  $A$  sería:

$$u_A = (6 - c_A) p(t_A < 1000) + (9 - c_A) p(t_A \geq 1000)$$

La utilidad neta esperada, con el proceso  $B$  sería:

$$u_B = (6 - c_A) p(t_B < 1000) + (9 - 2c_A) p(t_B \geq 1000)$$

Por otra parte, el cálculo de probabilidades con la función de probabilidad exponencial de dos parámetros se realiza a través de su función de probabilidad acumulada, la cual se obtiene al integrar la función de densidad de la expresión (3.70):

$$F(t) = p(y \leq t) = \int_{\beta}^t \frac{1}{\theta} e^{-\frac{(y-\beta)}{\theta}} dy$$

$$u = \frac{(y-\beta)}{\theta} \rightarrow du = -\frac{dy}{\theta}$$

$$F(t) = p(y \leq t) = - \left[ e^{-\frac{(y-\beta)}{\theta}} \right]_{\beta}^t = 1 - e^{-\frac{(t-\beta)}{\theta}} \quad (3.71)$$

Con la expresión anterior, ya se pueden calcular las probabilidades para cada proceso de manufactura:

$$p(t_A \leq 1000) = 1 - e^{-\left(\frac{1000-\beta_A}{\theta_A}\right)} \quad p(t_B \leq 1000) = 1 - e^{-\left(\frac{1000-\beta_B}{\theta_B}\right)}$$

$$p(t_A \leq 1000) = 1 - e^{-\left(\frac{1000-500}{2000}\right)} \quad p(t_B \leq 1000) = 1 - e^{-\left(\frac{1000-750}{2500}\right)}$$

$$p(t_A \leq 1000) = 0.2212 \quad p(t_B \leq 1000) = 0.0952$$

$$p(t_A \leq 1000) = 1 - 0.2212 = 0.7788 \quad p(t_B \leq 1000) = 1 - 0.0952 = 0.9048$$

De tal manera que las utilidades netas esperadas estarían dadas por:

$$u_A = (6 - c_A) 0.2212 + (9 - c_A) 0.7788$$

$$u_B = (6 - 2c_A) 0.0952 + (9 - 2c_A) 0.9048$$

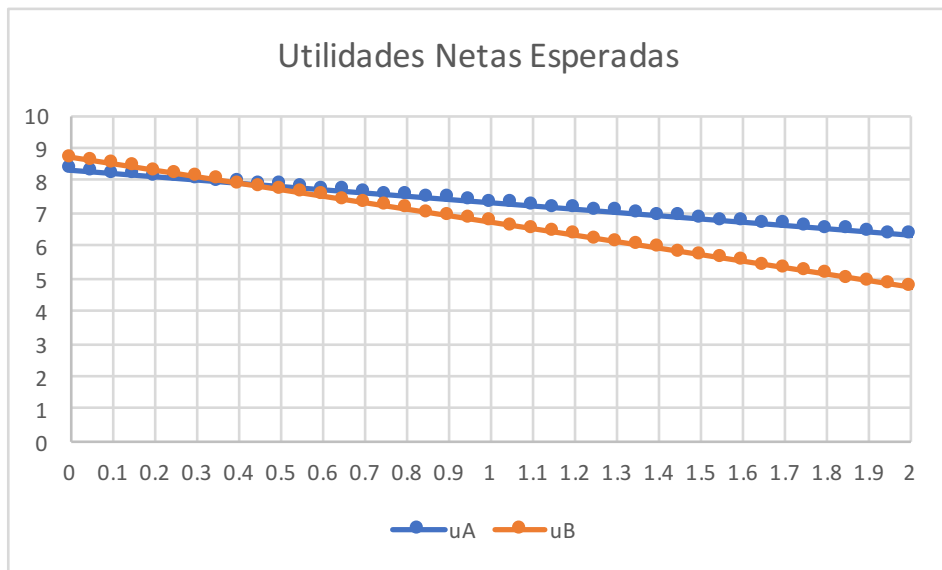
Es decir:

$$u_A = 8.3364 - u_A$$

$$u_B = 8.7145 - 2c_A$$

Como se puede apreciar, se trata de dos rectas, las cuales pueden ser graficadas en un sistema cartesiano ( $c$ ,  $u$ ), como se muestra en la gráfica de la figura 3.23.

**FIGURA 3.23.**



Como se puede apreciar en la figura 3.23 el proceso  $B$  es más rentable que el proceso  $A$  hasta un costo de 38 centavos aproximadamente, pero después de ese valor, es mejor el proceso  $A$ .

## Ejercicios propuestos del capítulo 3

- 3.1. Un lote consta de 16 artículos, de los cuales 10 son buenos, 4 con pequeños defectos y 2 con defectos graves. Se eligen al azar 2 artículos (sin sustitución).
- Obtener la distribución de probabilidad de la variable aleatoria  $x$ = número de artículos buenos y dibujarla.
  - Determinar la media, la varianza y la distribución de probabilidad acumulada  $F(x)$  y dibujarla.
  - Calcular la probabilidad de que el segundo sea bueno dado que el primero también lo fue (usando la definición de probabilidad condicional).
- 3.2. Una moneda confiable es lanzada tres veces y la variable aleatoria  $x$  es igual al número total de águilas. Determine y dibuje  $f_x(x)$  y  $F_x(X)$ .
- 3.3. En una estación de trenes se tienen dos máquinas de patio distintas que se usan para formarlos. La primera puede acomodar de 10 a 14 carros por hora; al ser su distribución simétrica, la probabilidad de colocar 12 carros es el doble de colocar 11 y la de 11 es el doble de la de 10. La segunda máquina mueve de 10 a 12 carros por hora con media de 11 carros por hora y varianza de  $2/3$ . Se sabe que si se forman trenes a razón de 20 carros por hora es imposible que los trenes salgan a tiempo; si se colocan 21 carros por hora, la probabilidad de salir a tiempo es de 0.2; si se forman 22, la probabilidad es de 0.4; 23 es 0.6; 24 es 0.8. Si se toman 25 o más carros por hora la probabilidad de salir a tiempo es 1.
- Calcular la probabilidad de que el total de carros formados por hora, por las dos máquinas, sea mayor de 23.
  - Calcular la probabilidad de que el total de carros formados por hora sea 23, dado que se sabe que alguna de las 2 máquinas colocó 11 carros.
  - Calcular la probabilidad de que se formen carros a razón de 23 por hora en un día, sabiendo que ese día los trenes no salieron a tiempo.

- 3.4. Los equipos de fútbol Universidad y América van a jugar una serie de tres partidos. Un universitario estima que la probabilidad de que gane su equipo en un juego individual es 0.6, sin que importen los resultados anteriores. Si  $x$  es la variable aleatoria que representa el número de juegos que puede ganar el equipo Universidad:
- Construya en forma tabular y gráfica la distribución de probabilidad de la variable aleatoria  $x$ .
  - Obtenga el segundo momento con respecto a la media.
- 3.5. La duración en horas de una componente electrónica es una variable aleatoria cuya  $F_x(X)$  es:

$$F_x(x) = 1 - e^{-\frac{x}{100}} \quad \forall x > 0$$

Calcule:

- La función de densidad de probabilidad  $f_x(X)$
  - $P(x \geq 200 \text{ horas})$
- 3.6. Sea  $x$  una v.a., demuestre que:  $\text{Var}(aX+b) = a^2 \text{Var}(X)$ , donde  $a$  y  $b \in R$
- 3.7. Sea la variable aleatoria con función densidad de probabilidad

$$f(x) = \begin{cases} 2e^{-2x} & x \geq 0 \\ 0 & x \leq 0 \end{cases}$$

- Demuestre que se trata de una función de probabilidad.
- Obtenga su función de probabilidad acumulada.
- Determine su función generadora de momentos y su función característica.
- Calcule sus parámetros estadísticos básicos como media, mediana, moda, varianza, desviación estándar, coeficiente de variación, coeficiente de asimetría y coeficiente de curtosis.
- Calcule las siguientes probabilidades:
  - $p(x < 3)$
  - $p(x > 2)$
  - $p(0.5 < x < 1.5)$
  - Obtenga el valor de  $k$ , tal que  $p(\mu_x - k\sigma_x < x < \mu_x + k\sigma_x) = 0.90$



3.8. Sea la función de densidad:

$$f_x(x) \left\{ \begin{array}{ll} 0 & x < 0 \\ -\frac{x}{18} + \frac{1}{3} & 0 \leq x \leq 6 \\ 0 & x > 6 \end{array} \right\}$$

- a) Demuestre que se trata de una función de probabilidad.
- b) Obtenga su función de probabilidad acumulada.
- c) Determine su función generadora de momentos y su función característica.
- d) Calcule sus parámetros estadísticos básicos como media, mediana, moda, varianza, desviación estándar, coeficiente de variación, coeficiente de asimetría y coeficiente de curtosis.
- e) Calcule las siguientes probabilidades:
  - i.  $p(x < 3)$
  - ii.  $p(x > 2)$
  - iii.  $p(0.5 < x < 1.5)$
  - iv. Obtenga el valor de  $k$ , tal que  $p(\mu_x - k\sigma_x < x < \mu_x + k\sigma_x) = 0.90$

3.9. Se tiene una variable aleatoria “ $x$ ” y una función  $f(x)$  de ella, tal que  $1 \leq x \leq 2$

$$f(x) = \left\{ \begin{array}{ll} x & 0 \leq x \leq 1 \\ 2-x & 1 \leq x \leq 2 \\ 0 & \text{cualquier otro caso} \end{array} \right\}$$

- a) Demuestre que se trata de una función de probabilidad.
- b) Obtenga su función de probabilidad acumulada.
- c) Determine su función generadora de momentos y su función característica.
- d) Calcule sus parámetros estadísticos básicos como media, mediana, moda, varianza, desviación estándar, coeficiente de variación, coeficiente de asimetría y coeficiente de curtosis.
- e) Calcule las siguientes probabilidades:
  - i.  $p(x < 1)$
  - ii.  $p(x > 1.5)$
  - iii.  $p(0.5 < x < 1.5)$
  - iv. Obtenga el valor de  $k$ , tal que  $p(\mu_x - k\sigma_x < x < \mu_x + k\sigma_x) = 0.90$

3.10. La función de densidad de probabilidad de una variable aleatoria  $x$  está determinada por:

$$f(x) \left\{ \begin{array}{ll} \frac{1}{16} x e^{-\left(\frac{x}{4}\right)} & x > 0 \\ 0 & \text{para cualquier otro valor} \end{array} \right\}$$

- Calcule la función de distribución acumulada.
- Obtenga la función generatriz de momentos.
- Determine la media y la desviación estándar de  $X$ .

3.11. A partir de la definición de esperanza matemática  $E\{t\} = \int_0^{\infty} t f(t) dt$  demuestre que:  $E\{t\} = \int_0^{\infty} R(t) dt$  donde:  $R(t) + F(t) = 1$

3.12. Obtenga la media y la varianza de la distribución  $F$ .

## 4. Variables aleatorias conjuntas

En la mayoría de los fenómenos naturales o de los problemas que enfrenta la Ingeniería, se debe trabajar simultáneamente con dos o más variables aleatorias. Por ejemplo, el número de personas que llegan a una cola y el tipo de problema que llegan a tratar. Ambas variables afectan al tiempo de servicio. En este caso, el número de llegadas a la cola es discreto y el tipo de problema a tratar también es discreto. Otro ejemplo diferente se presenta en el área de recibo de una empresa manufacturera, donde se reciben lotes de lámina de acero, a los cuáles se les inspecciona el calibre (espesor de la hoja) y la dureza mecánica, ya que afectan el desempeño del troquel que se usa para doblar y cortar la lámina; nótese que, en este segundo caso, ambas variables, el calibre y la dureza, son variables aleatorias continuas.

Otro ejemplo sería llevar a efecto un estudio del Índice de Masa Corporal (IMC) de los estudiantes de la Facultad de Ingeniería de la UNAM. El índice de masa corporal de una persona se define como:

$$\text{IMC} = \text{Peso} / \text{Estatura}^2$$

Este tipo de análisis requiere el estudio de dos variables aleatorias de la comunidad de estudiantes de la Facultad de Ingeniería: su peso y su estatura. En este caso, ambas variables son continuas en un cierto rango.

Existen ocasiones en que una de las variables es discreta y la otra continua, por ejemplo, el número de llegadas a una cola, el cual es discreto y el tiempo entre llegadas, que es continuo. Si se pretende analizar un modelo de líneas de espera se tendrá que emplear una función de probabilidad conjunta, en la cual se trata un modelo mixto, es decir, una de sus variables es discreta y la otra es continua.

El objetivo de este tema es estudiar las denominadas funciones de probabilidad conjunta para dos o más variables aleatorias. Asimismo, definir y calcular otro tipo de funciones de probabilidad ligadas a ellas, como lo son las funciones

de probabilidad marginales, las funciones de probabilidad condicionales, extender el concepto de esperanza matemática para varias variables aleatorias y obtener parámetros o medidas de tendencia central, de dispersión y de forma para funciones de varias variables, así como definir y aplicar los conceptos de regresión y correlación.

#### 4.1. Funciones de probabilidad de variables aleatorias conjuntas discretas

Se dice que  $p(x, y)$  es una función de probabilidad discreta, si para cada pareja ordenada  $(x, y)$  del dominio de definición de  $x$  y  $y$  existe un número  $p(x, y)$  que cumple las siguientes propiedades:

- i.  $p(x, y) \geq 0 \quad \forall x \in \text{Dominio de } x$   
 $\quad \quad \quad \forall y \in \text{Dominio de } y$
- ii.  $\sum_{\forall x} \sum_{\forall y} p(x, y) = 1$  (4.1)

A las funciones:

$$p_x(x) = \sum_{\forall y} p(x, y)$$

$$p_y(y) = \sum_{\forall x} p(x, y) \quad (4.2)$$

se les denomina *funciones de probabilidad marginales discretas* de  $x$  y de  $y$  respectivamente.

Cada una de estas distribuciones de probabilidad cumple con las propiedades de una función de probabilidad discreta y posee medidas de tendencia central como media, mediana, moda; medidas de dispersión varianza, desviación estándar, coeficiente de variación; medidas de forma como coeficiente de asimetría, coeficiente de curtosis, etcétera. Asimismo, cada una de las funciones de probabilidad marginales tiene una función generatriz de momentos que la define.

A las funciones de probabilidad:

$$p_{y|x}(y) = \frac{p(x, y)}{p_x(x)} \quad \forall x, y$$

$$p_{x|y}(x) = \frac{p(x, y)}{p_y(y)} \quad \forall x, y \quad (4.3)$$

se les denomina *funciones de probabilidad discretas condicionales de y dado x, así como de x dado y*, respectivamente.

A las expresiones:

$$\begin{aligned} E\{y|x\} &= \sum_{\forall y} y p_{y|x}(y) \\ E\{x|y\} &= \sum_{\forall x} x p_{x|y}(x) \end{aligned} \quad (4.4)$$

se les denomina *esperanzas condicionales discretas de y dado x, así como de x dado y*, respectivamente. A las gráficas de estas esperanzas condicionales se les denomina *regresión de y dado x, así como de x dado y*, respectivamente.

Se dice que *x es estadísticamente independiente de y*, si y solo si

$$p_{xy}(x,y) = p_x(x) p_y(y) \quad (4.5)$$

Se define la covarianza de *x* con *y* de la siguiente forma:

$$\sigma_{x,y}^2 = \text{cov}(x,y) = E \{ (x-\mu_x)(y-\mu_y) \} \quad (4.6)$$

La covarianza también puede calcularse como:

$$\sigma_{x,y}^2 = \text{cov}(x,y) = E \{ xy \} - \mu_x \mu_y \quad (4.7)$$

Nótese, de la expresión anterior, que si *x* y *y* son estadísticamente independientes entonces la covarianza es cero.

En probabilidad y estadística, la covarianza es un valor que indica el grado de variación conjunta de dos variables aleatorias respecto a sus medias. Es el dato básico para determinar si existe una dependencia entre ambas variables, además es el dato necesario para estimar otros parámetros básicos como el coeficiente de correlación.

Cuando los valores altos de una de las variables suelen mayoritariamente corresponderse con los valores altos de la otra, y lo mismo se verifica para los pequeños valores de una con los de la otra, se corrobora que tienden a mostrar comportamiento similar, lo que se refleja en un valor positivo de la covarianza; por el contrario, cuando a los mayores valores de una variable suelen corresponder en general los menores de la otra, expresando un comportamiento opuesto, la covarianza es negativa. El signo de la covarianza, por lo tanto, expresa la tendencia en la relación entre las variables.

## Propiedades de la Covarianza

Estas propiedades se deducen directamente de las propiedades de la Esperanza Matemática (de la expresión 3.15 y de sus propiedades 3.16), así como de las propiedades de la varianza (expresiones 3.29 a 3.38).

Sean  $x$ ,  $y$ ,  $v$  y  $w$  variables aleatorias y  $a$ ,  $b$ ,  $c$  y  $d$  constantes no aleatorias. Entonces:

- i.  $\sigma_{a,x}^2 = \text{cov}(a, x) = E\{(a - \mu_a)(x - \mu_x)\} = 0$
- ii.  $\sigma_{x,x}^2 = \text{cov}(x, x) = E\{(x - \mu_x)(x - \mu_x)\} = E\{(x - \mu_x)^2\} = \sigma_x^2$
- iii.  $\sigma_{x,y}^2 = \text{cov}(x, y) = E\{(x - \mu_x)(y - \mu_y)\} = \text{cov}(y, x) = \sigma_{y,x}^2$
- iv.  $\sigma_{ax,by}^2 = \text{cov}(ax, by) = E\{(ax - a\mu_x)(by - b\mu_y)\} = ab \text{cov}(x, y) = ab\sigma_{x,y}^2$
- v.  $\sigma_{x+a,y+b}^2 = \text{cov}(x+a, y+b) = E\{(x+a - \mu_x - a)(y+b - \mu_y - b)\} = \text{cov}(x, y) = \sigma_{x,y}^2$
- vi.  $\sigma_{ax+by,cu+dv}^2 = \text{cov}\{ax+by, cu+dv\} =$   
 $= ac \text{cov}\{x, u\} + ad \text{cov}\{x, v\} + bc \text{cov}\{y, u\} + bd \text{cov}\{y, v\}$  (4.8)

La última propiedad de las expresiones 4.8 se puede generalizar a una combinación lineal de  $n$  variables.

La versión normalizada de la covarianza, el coeficiente de correlación, indica la magnitud de la especificidad de la relación lineal.

Se define el coeficiente de correlación, denotado por  $\rho$ , a la siguiente expresión:

$$\rho = \frac{\sigma_{x,y}^2}{\sigma_x \sigma_y} = \frac{\text{cov}(x,y)}{\sqrt{\text{var}(x)} \sqrt{\text{var}(y)}} \quad (4.9)$$

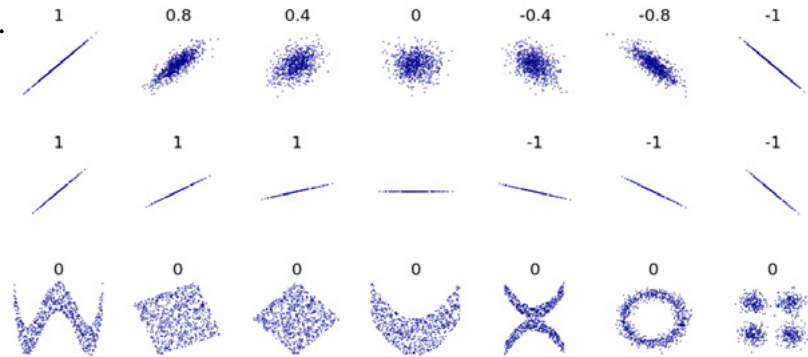
De nueva cuenta observe que si  $x$  es estadísticamente independiente de  $y$ , entonces su coeficiente de correlación es cero, lo cual es una condición necesaria, más no suficiente.

## Propiedades del Coeficiente de Correlación

El valor del índice de correlación varía en el intervalo  $[-1,1]$ , indicando el signo el sentido de la relación:

- i. Si  $\rho=1$ , existe una correlación positiva perfecta. El índice indica una dependencia total entre las dos variables, denominada relación directa: cuando una de ellas aumenta, la otra también lo hace en proporción constante.
- ii. Si  $0 < \rho < 1$ , existe una correlación positiva.
- iii. Si  $\rho=0$ , no existe relación lineal. Pero esto no necesariamente implica que las variables son independientes: pueden existir todavía relaciones no lineales entre las dos variables.
- iv. Si  $-1 < \rho < 0$ , existe una correlación negativa.
- v. Si  $\rho=-1$ , existe una correlación negativa perfecta. El índice indica una dependencia total entre las dos variables llamada relación inversa: cuando una de ellas aumenta, la otra disminuye en proporción constante.

FIGURA 4.1.



Varios grupos de puntos  $(x,y)$ , con el coeficiente de correlación para cada grupo. Nótese que la correlación refleja la no linealidad y la dirección de la relación lineal. En la figura del centro, la varianza de  $y$  es nula, por lo que la correlación es indeterminada.

**NOTA:** Coeficiente de correlación de Pearson. (2019).

[https://es.wikipedia.org/wiki/Coeficiente\\_de\\_correlaci%C3%B3n\\_de\\_Pearson](https://es.wikipedia.org/wiki/Coeficiente_de_correlaci%C3%B3n_de_Pearson)



## Ejercicio 4.1

El acabado superficial de una pieza metálica pulida en una fresadora se afecta por la velocidad de alimentación  $x$  y la profundidad de corte  $y$ . La máquina herramienta que se utiliza tiene tres velocidades de alimentación y cuatro profundidades de corte. Por experiencias pasadas en el uso se sabe que la probabilidad de que ocurra cada uno de los eventos compuestos está dada de acuerdo a la siguiente tabla:

FIGURA 4.2. Función de probabilidad conjunta discreta

Vel Alim (pulg/min)	Profundidad de corte (pulg)			
	0.15	0.18	0.2	0.25
0.2	0.0583	0.0648	0.0771	0.0880
0.25	0.0783	0.0854	0.0889	0.0922
0.3	0.0880	0.0878	0.0933	0.0978

- a) Determine si se trata de una función de probabilidad conjunta discreta.

Como se observa en el cuadro de la figura 4.2, todos los valores de probabilidad son positivos. Además, si se suman todas las probabilidades que aparecen en cada celda, se obtiene uno, por lo cual, sí se trata de una función de probabilidad discreta.

- b) Grafique la función de probabilidad  $p(x,y)$ . Se muestra en la figura 4.6. Para graficarla se usó Minitab 17 (<http://www.minitab.com/es-mx/>).

Para realizar la graficación de esta función, usando Minitab se llevan a cabo los siguientes pasos:

- i. Primero se capturan los datos de la función que se va a graficar en la hoja de entrada de Minitab, como se muestra en la figura 4.3. Otra forma más sencilla es, si los datos ya se tienen en una hoja de Excel, solo se copian los encabezados y se pegan en la hoja de entrada de Minitab.

Worksheet 1 ***				
↓	C1	C2	C3	C4
	x	y	p(x,y)	
1	0.20	0.15	0.0583039	
2	0.20	0.18	0.0647821	
3	0.20	0.20	0.0771496	
4	0.20	0.25	0.0880448	
5	0.25	0.15	0.0783274	
6	0.25	0.18	0.0853946	
7	0.25	0.20	0.0889282	
8	0.25	0.25	0.0921673	
9	0.30	0.15	0.0880448	
10	0.30	0.18	0.0877503	
11	0.30	0.20	0.0933451	
12	0.30	0.25	0.0977621	
13				

FIGURA 4.3.

- ii. Se selecciona en el menú principal Graph posteriormente 3D Scatterplot y al dar enter aparece la pantalla de la figura 4.4.
- iii. En la pantalla de la figura 4.4 se selecciona la opción Simple (la opción With Groups se elige si se hicieran varias lecturas para un mismo valor del dominio). Aparece la pantalla de la figura 4.5.

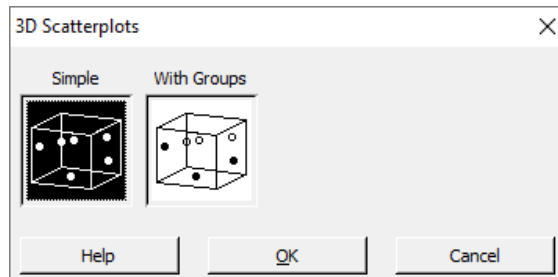


FIGURA 4.4

- iv. En la pantalla de la figura 4.5 se ubica en la celda que dice Z variable, posteriormente se selecciona con un click la opción C3  $p(x,y)$  y se selecciona *select* para que la escoja; lo mismo se hace con Y variable y con x variable, al final se da click en *Ok*, apareciendo la gráfica de la función de probabilidad discreta bivariada como se muestra en la figura 4.6.

FIGURA 4.5.

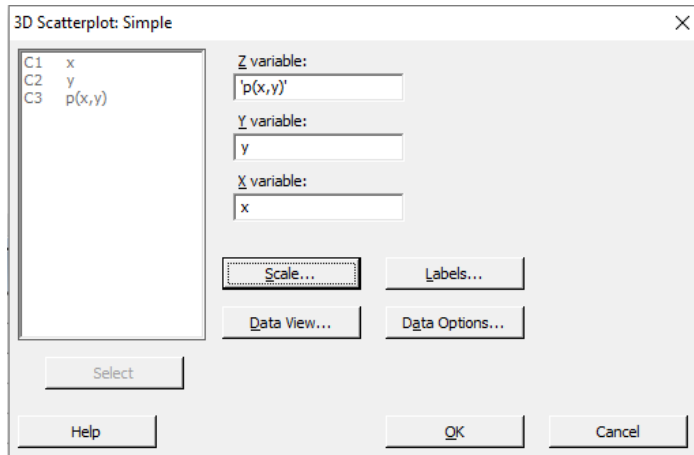
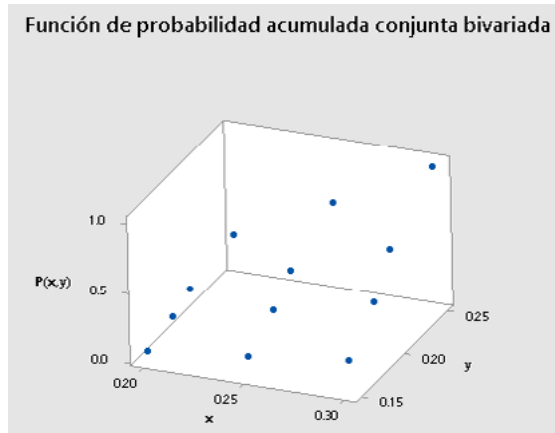


FIGURA 4.6.

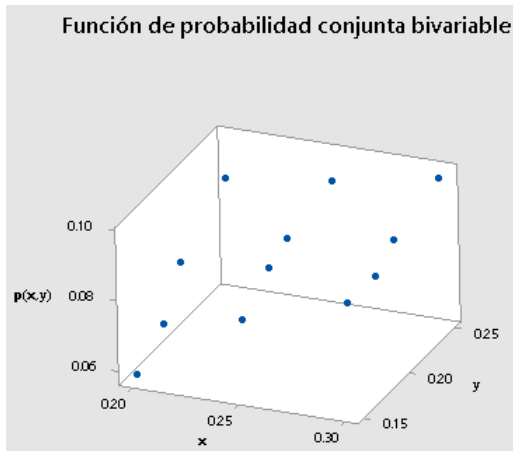


c) Obtenga la función de probabilidad acumulada  $P(x,y)$ . Se muestra en la tabla de la figura 4.7.

FIGURA 4.7.

Vel Alim (pulg/min)	Profundidad de corte (pulg)			
	0.15	0.18	0.2	0.25
0.2	0.0583	0.1231	0.2002	0.2883
0.25	0.1366	0.2868	0.4529	0.6331
0.3	0.2247	0.4626	0.7220	1.0000

d) Grafique la función de probabilidad acumulada  $P(x,y)$ . Se muestra en la figura 4.8. Para graficarla se usó Minitab 17 (<http://www.minitab.com/es-mx/>).

**FIGURA 4.8.**

- e) Obtenga las funciones de probabilidad marginales. Se muestran en el costado lateral derecho y en el costado inferior de la tabla de la figura 4.9.

**FIGURA 4.9.**

Vel Alim (pulg/min)	Profundidad de corte (pulg)				$p_x(x)$
	0.15	0.18	0.2	0.25	
0.2	0.0583	0.0648	0.0771	0.0880	0.2883
0.25	0.0783	0.0854	0.0889	0.0922	0.34482
0.3	0.0880	0.0878	0.0933	0.0978	0.3669
$p_y(y)$	0.22468	0.23793	0.25942	0.27797	

- f) Calcule las funciones de probabilidad condicionales de  $y$  dado  $x$  y de  $x$  dado  $y$ . Se ilustran en la tabla de las figuras 4.10 y 4.11.

**FIGURA 4.10.**

$p_{y x}(y)$	0.15	0.18	0.2	0.25
0.2	0.2022	0.2247	0.2676	0.3054
0.25	0.2272	0.2477	0.2579	0.2673
0.3	0.2400	0.2392	0.2544	0.2665

**FIGURA 4.11.**

$p_{x y}(x)$	0.15	0.18	0.2	0.25
0.2	0.2595	0.2723	0.2974	0.3167
0.25	0.3486	0.3589	0.3428	0.3316
0.3	0.3919	0.3688	0.3598	0.3517

- g) Calcule las esperanzas matemáticas condicionales de  $y$  dado  $x$  y de  $x$  dado  $y$  respectivamente. Se muestran en el costado lateral derecho y en el costado inferior de las figuras 4.12 y 4.13 respectivamente.
- h) Grafique las regresiones de  $y$  dado  $x$  y de  $x$  dado  $y$ . Las gráficas se muestran en las figuras 4.14.a y 4.14.b.

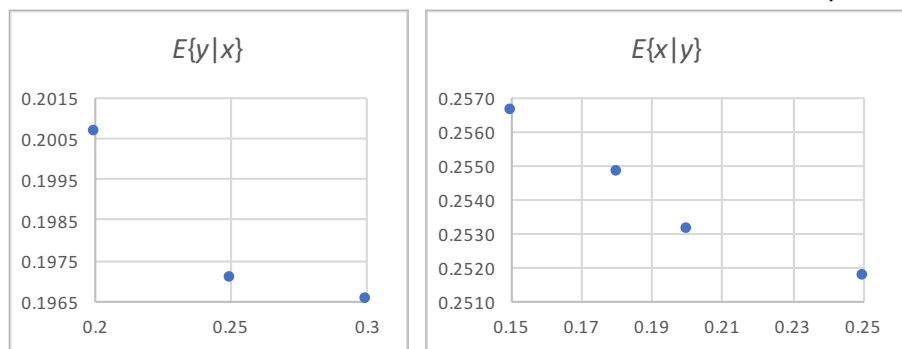
FIGURA 4.12.

$p_{y x}(y)$	0.15	0.18	0.2	0.25	$E\{y x\}$
0.2	0.2022	0.2247	0.2676	0.3054	0.2007
0.25	0.2272	0.2477	0.2579	0.2673	0.1971
0.3	0.2400	0.2392	0.2544	0.2665	0.1965

FIGURA 4.13.

$p_{x y}(y)$	0.15	0.18	0.2	0.25	$E\{x y\}$
0.2	0.2595	0.2723	0.2974	0.3167	
0.25	0.3486	0.3589	0.3428	0.3316	
0.3	0.3919	0.3688	0.3598	0.3517	
$E\{x y\}$	0.2566	0.2548	0.2531	0.2517	

FIGURA 4.14.a y 4.14.b



- i) Compruebe que las variables  $x$  y  $y$  no son estadísticamente independientes.

Para que  $x$  y  $y$  sean estadísticamente independientes, se debe cumplir que para cualquier valor de  $x$  y cualquier valor de  $y$ ,  $p(x,y) = p(x)p(y)$ . Con una sola celda que no cumpla esta propiedad se demuestra que no son independientes. Se prueba con la primera, la celda (0.2,0.15), a partir de los datos de la tabla de la figura 3.25, se observa que no se cumple la igualdad.

$$p(0.2,0.15) = 0.0583$$

$$p_x(0.2) p_y(0.15) = (0.2883)(0.2247) = 0.0648$$

j) Calcule la covarianza de  $x$  con  $y$ .

FIGURA 4.15.

Se forma  
la siguiente tabla:

$x$	$y$	$p(x,y)$	$xyp(x,y)$
0.2	0.15	0.0583039	0.0017
0.2	0.18	0.0647821	0.0023
0.2	0.2	0.0771496	0.0031
0.2	0.25	0.0880448	0.0044
0.25	0.15	0.0783274	0.0029
0.25	0.18	0.0853946	0.0038
0.25	0.2	0.0889282	0.0044
0.25	0.25	0.0921673	0.0058
0.3	0.15	0.0880448	0.0040
0.3	0.18	0.0877503	0.0047
0.3	0.2	0.0933451	0.0056
0.3	0.25	0.0977621	0.0073
Suma=			0.0502

Las medias de cada una de las variables se calculan como:

$$\mu_x = \sum_{\forall x} x p_x(x) = 0.2539$$

$$\mu_y = \sum_{\forall y} y p_y(y) = 0.1979$$

Utilizando la tabla de la figura 4.14

$$\sigma_{x,y}^2 = \text{cov}(x,y) = E\{(x-\mu_x)(y-\mu_y)\}$$

$$\sigma_{x,y}^2 = \text{cov}(x,y) = E\{xy\} - \mu_x \mu_y$$

$$\sigma_{x,y}^2 = \text{cov}(x,y) = \sum_{\forall x} \sum_{\forall y} xyp(x,y) - \mu_x \mu_y$$

$$\sigma_{x,y}^2 = \text{cov}(x,y) = 0.0502 - (0.2539)(0.1979)$$

$$\sigma_{x,y}^2 = \text{cov}(x,y) = -5.7E-05$$

Como se puede apreciar el coeficiente de correlación es cercano a cero y es negativo. El que sea cero significa que no hay correlación lineal, pero puede haber correlación no lineal. El signo negativo indica que para valores altos de una de las variables, la otra presenta valores bajos y viceversa.

FIGURA 4.16.

Vel Alim (pulg/min)	Profundidad de corte (pulg)						
	0.15	0.18	0.2	0.25	$p_x(x)$	$xp_x(x)$	$x^2p_x(x)$
0.2	0.0583	0.0648	0.0771	0.0880	0.2883	0.0577	0.0115
0.25	0.0783	0.0854	0.0889	0.0922	0.3448	0.0862	0.0216
0.3	0.0880	0.0878	0.0933	0.0978	0.3669	0.1101	0.0330
$p_y(y)$	0.2247	0.2379	0.2594	0.2780	Suma=	0.2539	0.0661
$yp_y(y)$	0.0337	0.04283	0.05188	0.06949	0.19791	var(x)=	0.0016
$y^2p_y(y)$	0.00506	0.00771	0.01038	0.01737	0.04051		
				var(y)=	0.00135		

k) Obtenga el coeficiente de correlación.

Usando la misma tabla anterior 4.15:

$$\sigma_x^2 = E\{x^2\} - \mu_x^2 = \sum_{v_x} x^2 p_x(x) - \mu_x^2 = 0.0661 - 0.2539^2 = 0.0016$$

$$\sigma_y^2 = E\{y^2\} - \mu_y^2 = \sum_{v_y} y^2 p_y(y) - \mu_y^2 = 0.0405 - 0.1979^2 = 0.00135$$

$$\rho = \frac{\sigma_{x,y}^2}{\sigma_x \sigma_y} = \frac{-5.7E-05}{\sqrt{0.0016} \sqrt{0.00135}} = -0.0388$$

Observe que el coeficiente de correlación es negativo y cercano a cero. El que sea cero significa que no hay correlación lineal, pero puede haber correlación no lineal. El signo negativo indica que para valores altos de una de las variables, la otra presenta valores bajos y viceversa.

## 4.2. Funciones de probabilidad de variables aleatorias conjuntas continuas

Se dice que  $f(x, y)$  es una función de densidad de probabilidad bivariada continua, si para cada pareja ordenada  $(x, y)$  del dominio de definición de  $x$  y  $y$ , existe un número  $f(x, y)$  que cumple las siguientes propiedades:

$$\text{i. } f(x, y) \geq 0 \quad \begin{array}{l} \forall x \in \text{Dominio de } x \\ \forall y \in \text{Dominio de } y \end{array} \quad (4.10)$$

$$\text{ii. } \int_{\forall x} \int_{\forall y} f(x, y) \, dy dx = \int_{\forall y} \int_{\forall x} f(x, y) \, dx dy$$

A las funciones:

$$\begin{aligned} f_x(x) &= \int_{\forall y} f(x, y) dy \\ f_y(y) &= \int_{\forall x} f(x, y) dx \end{aligned} \quad (4.11)$$

Se les denomina *funciones de densidad de probabilidad marginales continuas* de  $x$  y de  $y$  respectivamente. Cada una de estas distribuciones de probabilidad cumple con las propiedades de una función de probabilidad continua y posee medidas de tendencia central como: media, mediana, moda; medidas de dispersión: varianza, desviación estándar, coeficiente de variación; medidas de forma como: coeficiente de asimetría, coeficiente de curtosis, etcétera. Asimismo, cada una de las funciones de densidad de probabilidad marginales tiene una función generatriz de momentos que la define.

A las funciones de probabilidad:

$$\begin{aligned} f_{y|x}(y) &= \frac{f(x, y)}{f_x(x)} \quad \forall x, y \\ f_{x|y}(x) &= \frac{f(x, y)}{f_y(y)} \quad \forall x, y \end{aligned} \quad (4.12)$$

Se les denomina *funciones de densidad de probabilidad condicionales continuas bivariadas* de  $y$  dado  $x$ , así como de  $x$  dado  $y$ , respectivamente.



A las expresiones:

$$E\{y|x\} = \int_{\forall y} y f_{y|x}(y) dy$$

$$E\{x|y\} = \int_{\forall x} x f_{x|y}(x) dx \quad (4.13)$$

Se les denomina *esperanzas condicionales continuas* de  $y$  dado  $x$ , así como de  $x$  dado  $y$ , respectivamente. A las gráficas de estas esperanzas condicionales se les denomina *regresión* de  $y$  dado  $x$ , así como de  $x$  dado  $y$ , respectivamente.

La variable aleatoria  $x$  es *estadísticamente independiente* de la variable aleatoria  $y$ , si y solo si,

$$f_{xy}(x, y) = f_x(x) f_y(y) \quad (4.14)$$

Este resultado se puede generalizar para el caso de  $n$  variables estadísticamente independientes entre sí,  $x_1, x_2, \dots, x_n$ :

$$f_{x_1 x_2 \dots x_n}(x_1, x_2, \dots, x_n) = f_{x_1}(x_1) f_{x_2}(x_2) \dots f_{x_n}(x_n) \quad (4.15)$$

Sean las variables aleatorias  $x$  y  $y$ , estadísticamente independientes entre sí, con funciones generatrices de momentos  $FGM_x(t)$  y  $FGM_y(t)$ , y sea  $z = g(x, y)$ . La función generatriz de momentos de  $z$  se obtiene como el producto de las funciones generatrices de  $x$  y  $y$ , es decir,

$$FGM_z(t) = FGM_x(t) FGM_y(t) \quad (4.16)$$

También se puede generalizar esta expresión:

$$FGM_z(t) = FGM_{x_1}(t) FGM_{x_2}(t) \dots FGM_{x_n}(t) \quad (4.17)$$

La covarianza para el caso continuo se define de la misma forma que para el discreto según las expresiones 4.6, 4.7; además, también cumple las propiedades de las expresiones 4.8.

El coeficiente de correlación se define de la misma forma que para el caso discreto, según la expresión 4.9 y satisface las mismas propiedades citadas en 4.10.

## Ejercicio 4.2

(Problema 4-9, página 124, Hines, Montgomery, Goldsman y Borror, *Probabilidad y Estadística para Ingeniería*, Cuarta Edición, Grupo Patria, 2013). Suponga que  $x$  y  $y$  son las calificaciones codificadas del primer y segundo final de la asignatura de probabilidad de la División de Ciencias Básicas de la Facultad de Ingeniería de la UNAM. Suponga que su función de densidad de probabilidad conjunta está dada por la expresión:

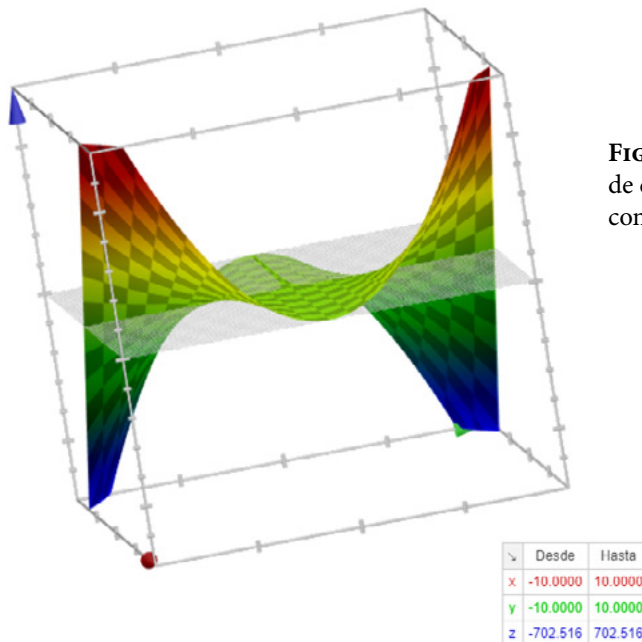
$$f(x,y) = \begin{cases} 6xy^2 & 0 \leq x \leq 1; 0 \leq y \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

- a) Determine si se trata de una función de densidad de probabilidad conjunta continua.

Como se puede apreciar en la definición de la función, en el dominio de definición,  $f(x,y) \geq 0$ , todos los valores de densidad de probabilidad son positivos. Además, al calcular la integral doble se obtiene uno, por lo cual, sí se trata de una función de densidad de probabilidad continua.

$$\int_0^1 \int_0^1 6xy^2 \, dx dy = \int_0^1 3[x^2]_0^1 y^2 dy = \int_0^1 3y^2 dy = [y^3]_0^1 = 1$$

- b) Grafique la función de densidad de probabilidad  $f(x,y)$ . Se muestra en la figura 4.17. Para graficarla se usó Google Graph (<https://www.google.com.mx/>).



**FIGURA 4.17.** Función de densidad de probabilidad conjunta bivariada continua

c) Obtenga la función de densidad de probabilidad acumulada  $F(x, y)$ .

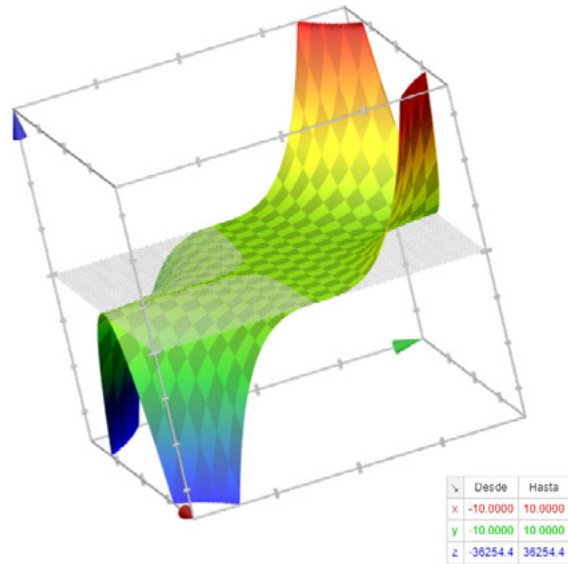
$$F(x, y) = \int_0^y \int_0^x 6uv^2 \, dudv = \int_0^y 3[u^2]_0^x v^2 \, dv$$

$$F(x, y) = \int_0^y 3x^2 v^2 \, dv = x^2 [v^3]_0^y$$

$$F(x, y) = x^2 y^3$$

d) Grafique la función de probabilidad acumulada  $F(x, y)$ . Se muestra en la figura 4.18. Para graficarla se usó Google Graph (<https://www.google.com.mx/>).

**FIGURA 4.18.** Función de Probabilidad Acumulada Conjunta Bivariada Continua



e) Obtenga las funciones de densidad de probabilidad marginales.

$$f_x(x) = \int_{\forall y} f(x, y) \, dy = \int_0^1 6xy^2 \, dy = 2x[y^3]_0^1 = 2x$$

$$f_y(y) = \int_{\forall x} f(x, y) \, dx = \int_0^1 6xy^2 \, dx = 3[x^2]_0^1 y^2 = 3y^2$$

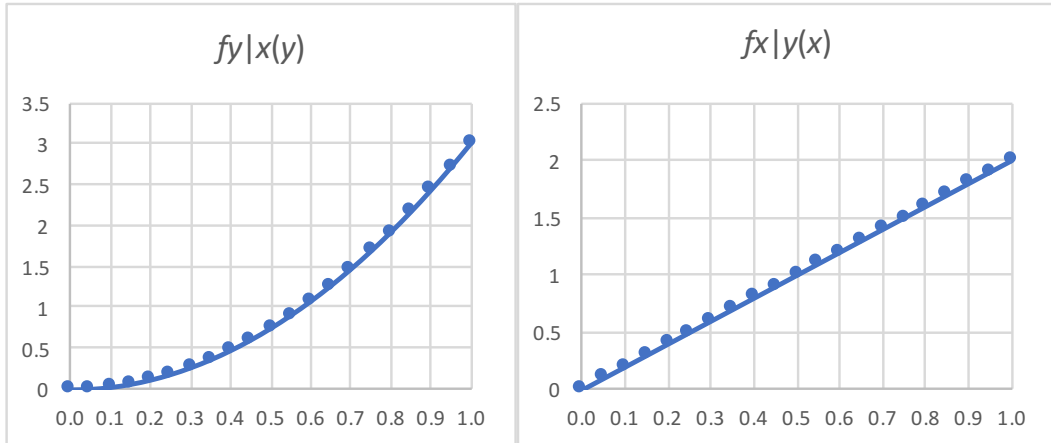
f) Calcule las funciones de densidad de probabilidad condicionales de  $y$  dado  $x$  y de  $x$  dado  $y$ .

$$f_{y|x}(y) = \frac{f(x, y)}{f_x(x)} = \frac{6xy^2}{2x} = 3y^2$$

$$f_{x|y}(x) = \frac{f(x, y)}{f_y(y)} = \frac{6xy^2}{3y^2} = 2x$$

- i. Grafique las regresiones de  $y$  dado  $x$  y de  $x$  dado  $y$ . Las gráficas se muestran en las figuras 4.19.a y 4.19.b.

FIGURAS 4.19.a Y 4.19b.



- ii. Demuestre que las variables  $x$  y  $y$  son estadísticamente independientes.

Observe que para cualquier valor de  $x$  y  $y$  en el dominio de definición, se cumple que:

$$f(x,y) = 6xy^2 = (2x)(3y^2) = f_x(x) f_y(y) \quad \text{l.q.q.d.}$$

- i) Calcule la covarianza de  $x$  con  $y$ .

$$\sigma_{x,y}^2 = \text{cov}(x,y) = E\{xy\} - \mu_x \mu_y$$

Pero como son independientes:

$$E\{xy\} = E\{x\} E\{y\} = \mu_x \mu_y$$

$$\sigma_{x,y}^2 = \text{cov}(x,y) = 0$$

- j) Calcule el coeficiente de correlación.

$$\rho = \frac{\sigma_{x,y}^2}{\sigma_x \sigma_y} = 0$$

## Ejercicio 4.3

Un taladro de precisión, colocado sobre un punto deseado, hará un agujero aceptable si queda dentro de una distancia de cinco micras del punto deseado. Usando el punto deseado como origen del sistema de coordenadas rectangulares y suponiendo que las coordenadas  $(x, y)$  del punto de contacto son variables aleatorias con la distribución conjunta:

$$f_{xy}(x, y) = \frac{1}{8\pi} e^{-\frac{1}{8}(x^2+y^2)} \quad x \in \mathfrak{R} \quad y \in \mathfrak{R}$$

Obtener la probabilidad de que el barrenado sea aceptable.

Las coordenadas del centro del barrenado son  $(x, y)$  y debe estar a cinco micras del origen por lo que se pide calcular:

$$p(\sqrt{x^2+y^2} \leq 5)$$

También:

$$p(x^2+y^2 \leq 25)$$

$$p(y \leq \sqrt{25-x^2}) = \frac{2 \cdot 2}{8\pi} \int_0^5 \int_0^{\sqrt{25-x^2}} e^{-\frac{1}{8}(x^2+y^2)} dy dx$$

Esta integral es más fácil integrarla en coordenadas polares, sólo falta considerar el diferencial de área en polares:

$$p(r \leq 5) = \frac{1}{8\pi} \int_0^{2\pi} \int_0^5 e^{-\frac{1}{8}r^2} r dr d\theta$$

$$p(r \leq 5) = -\frac{1}{2\pi} \int_0^{2\pi} \left[ e^{-\frac{r^2}{8}} \right]_0^5 d\theta$$

$$p(r \leq 5) = \frac{1}{2\pi} \left[ 1 - e^{-\frac{25}{8}} \right] [\theta]_0^{2\pi} = 1 - e^{-\frac{25}{8}}$$

$$p(r \leq 5) = 0.956063$$

### 4.3. Funciones de probabilidad de variables dependientes de variables aleatorias conjuntas

Tal como se mencionó anteriormente, muchos problemas en Ingeniería involucran el estudio de relaciones funcionales de una variable dependiente con dos o más variables independientes. Si una o más de las variables independientes son a su vez aleatorias, esto repercutirá en que a su vez la variable dependiente sea aleatoria. Un punto importante es conocer el modelo probabilístico de la variable dependiente, a partir del modelo probabilístico de las variables independientes.

Sean  $x$  y  $y$  dos variables aleatorias discretas con función de probabilidad  $p_{x,y}(x,y)$  y sea  $z=h(x,y)$  una relación funcional entre las variables aleatorias  $x$  y  $y$ , así como la variable dependiente  $z$ . Dado que  $z$  depende de  $x$  y  $y$ , esto implicará que  $z$  también sea aleatoria con cierto tipo de función de probabilidad desconocida  $p_z(z)$ . La pregunta es cómo deducir la función de probabilidad  $p_z(z)$  a partir de la función  $p_{x,y}(x,y)$ .

#### Ejercicio 4.4

Para la función de probabilidad conjunta que se muestra en la figura 4.2 del ejercicio 4.1, obtenga la función de probabilidad de  $z$  suponiendo:

- a)  $z=x+y$ . En la tabla de la figura 4.20 se muestran los resultados obtenidos de  $z=x+y$ .

**FIGURA 4.20.**

$x$	$y$	$z$	$p_{x,y}(x,y) = p_z(z)$
0.2	0.15	0.35	0.0583
0.2	0.18	0.38	0.0648
0.2	0.2	0.4	0.0771
0.2	0.25	0.45	0.0880
0.25	0.15	0.4	0.0783
0.25	0.18	0.43	0.0854
0.25	0.2	0.45	0.0889
0.25	0.25	0.5	0.0922
0.3	0.15	0.45	0.0880
0.3	0.18	0.48	0.0878
0.3	0.2	0.5	0.0933
0.3	0.25	0.55	0.0978

b)  $z = x/y$ . En la tabla de la figura 4.21 se muestran los resultados obtenidos de  $z = x/y$ .

FIGURA 4.21

$x$	$y$	$z$	$p_z(z)$
0.2	0.25	0.80	0.0880
0.2	0.2	1.00	0.0771
0.25	0.25	1.00	0.0922
0.2	0.18	1.11	0.0648
0.3	0.25	1.20	0.0978
0.25	0.2	1.25	0.0889
0.2	0.15	1.33	0.0583
0.25	0.18	1.39	0.0854
0.3	0.2	1.50	0.0933
0.25	0.15	1.67	0.0783
0.3	0.18	1.67	0.0878
0.3	0.15	2.00	0.0880

### Procedimiento para deducir una función de probabilidad derivada de otras funciones de probabilidad conjunta

Para el caso de variables aleatorias continuas, se sugiere aplicar el siguiente procedimiento:

Sean  $x$  y  $y$  dos variables aleatorias continuas con función de probabilidad conjunta  $f_{x,y}(x,y)$  y sea  $z = g(x,y)$  una relación funcional entre las variables aleatorias  $x$ ,  $y$  y la variable dependiente  $z$ .

- i. Se define una nueva variable aleatoria  $w = h(x,y)$  que sea fácil de despejar y que sea derivable.
- ii. Se obtiene la inversa de  $x = g^{-1}(z,w)$  y  $y = h^{-1}(z,w)$  en función de  $z$  y  $w$ .
- iii. Se calcula el Jacobiano de  $x$  y  $y$  en función de  $z$  y  $w$ , es decir:

$$J \begin{pmatrix} x & y \\ z & w \end{pmatrix} = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix} \quad (4.18)$$

iv. La densidad conjunta  $l(z,w)$  se obtiene con la siguiente expresión:

$$l(z,w) = f(h^{-1}, g^{-1}) \left| j \begin{pmatrix} x & y \\ z & w \end{pmatrix} \right| \quad (4.19)$$

v. La función de densidad que se busca de  $z$ , es la marginal de  $l(z,w)$  y se obtiene como:

$$g_z(z) = \int_{\forall w} l(z,w) dw \quad (4.20)$$

#### Ejercicio 4.5

Para la función de probabilidad conjunta dada por la expresión:

$$f(x,y) = \begin{cases} 6xy^2 & 0 \leq x \leq 1; 0 \leq y \leq 1 \\ 0 & \text{en otro caso} \end{cases}$$

Obtenga la función de densidad  $f_z(z)$  si se define  $z$  como:

a)  $z = x + y$ .

i. Se aplicará el procedimiento antes mencionado usando como función alternativa  $w = x - y$  por facilidad.

ii. Por suma y resta:

$$z = x + y$$

$$w = x - y$$

→

$$x = \frac{1}{2}(z+w)$$

$$y = \frac{1}{2}(z-w)$$



$$\text{iii. } J\left(\frac{x}{z} \frac{y}{w}\right) = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix} = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{vmatrix} = \frac{1}{2} \left(-\frac{1}{2}\right) - \frac{1}{2} \left(\frac{1}{2}\right) = -\frac{1}{2}$$

$$\text{iv. } l(z, w) = f(h^{-1}, g^{-1}) \left| j\left(\frac{x}{z} \frac{y}{w}\right) \right| = 6 \left[ \frac{1}{2}(z+w) \right] \left[ \frac{1}{2}(z-w) \right]^2 \left| -\frac{1}{2} \right|$$

$$l(z, w) = \frac{3}{8} (z+w)(z-w)^2 = \frac{3}{8} [z^3 - wz^2 - w^2z + w^3]$$

- v. Nótese que si  $x=0$  y  $y=0$ , entonces  $z=0$ . Asimismo, si  $x=1$  y  $y=1$ , entonces  $z=2$ . De la misma forma, si  $x=0$  y  $y=1$ ,  $w=-1$  y si  $x=1$  y  $y=0$ ,  $w=1$ . De esta forma, se sabe que  $0 \leq z \leq 2$  y que  $-1 \leq w \leq 1$ , por lo que:

$$g_z(z) = \int_{-\infty}^{\infty} l(z, w) dw = \frac{3}{8} \int_{-1}^1 [z^3 - wz^2 - w^2z + w^3] dw$$

$$g_z(z) = \frac{3}{8} \left[ z^3w - \frac{w^2}{2} z^2 - \frac{w^3}{3} z + \frac{w^4}{4} \right]_{-1}^1$$

$$g_z(z) = \frac{3}{8} \left[ 2z^3 - \frac{2}{3} z \right] = \frac{z}{4} (3z^2 - 1)$$

b)  $z = x^2 + y^2$ .

- i. Por facilidad, recordando las ecuaciones de transformación de un sistema cartesiano a un sistema polar, se definirá  $w = \text{ang tan}(x/y)$ .
- ii. Se despejan las variables  $x$  y  $y$  en función de  $z$  y  $w$ .

$$z = x^2 + y^2$$

$$w = \text{ang tan} \left( \frac{y}{x} \right)$$

$$\frac{y}{x} = \tan(w) \rightarrow y = x \tan(w)$$

$$z = x^2 + x^2 \tan^2(w) = x^2 (1 + \tan^2(w)) = x^2 \sec^2(w)$$

$$x^2 = \frac{z}{\sec^2(w)} = z \cos^2(w)$$

$$x = \sqrt{z} \cos(w)$$

$$y = \sqrt{z} \cos(w) \tan(w) = \sqrt{z} \operatorname{sen}(w)$$

$$\text{iii. } J \left( \begin{matrix} x & y \\ z & w \end{matrix} \right) = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix} = \begin{vmatrix} \frac{1}{2\sqrt{z}} \cos(w) & \frac{1}{2\sqrt{z}} \operatorname{sen}(w) \\ -\sqrt{z} \operatorname{sen}(w) & -\sqrt{z} \cos(w) \end{vmatrix} = \frac{1}{2} \cos^2(w) + \frac{1}{2} \operatorname{sen}^2(w) = \frac{1}{2}$$

$$\text{iv. } l(z, w) = 6 (\sqrt{z} \cos(w)) (\sqrt{z} \operatorname{sen}(w))^2 \left| \frac{1}{2} \right|$$

$$l(z, w) = 3z^{\frac{3}{2}} \cos(w) \operatorname{sen}^2(w)$$

$$g_z(z) = \int_0^{\frac{\pi}{2}} 3z^{\frac{3}{2}} \cos(w) \operatorname{sen}^2(w) dw$$

$$\text{v. } g_z(z) = z^{\frac{3}{2}} \left[ \operatorname{sen}^3(w) \right]_0^{\frac{\pi}{2}} = z^{\frac{3}{2}}$$

#### 4.4. Funciones de probabilidad con $n$ variables aleatorias conjuntas

En caso de que se tengan  $n$  variables aleatorias, la variable aleatoria será un vector aleatorio con  $n$  componentes:  $\vec{x} = (x_1, x_2, \dots, x_n)$ . La función de probabilidad estará representada de la siguiente forma  $f(\vec{x}) = f(x_1, x_2, \dots, x_n)$ . Para que sea una función de probabilidad debe cumplir dos requisitos:

##### 4.4.1. Funciones de probabilidad con $n$ variables aleatorias conjuntas discretas

i.  $p(x_1, x_2, \dots, x_n) \geq 0$

ii. 
$$\sum_{\forall x_1} \sum_{\forall x_2} \dots \sum_{\forall x_n} p(x_1, x_2, \dots, x_n) = 1 \quad (4.21)$$

Además, se cumple que:

iii. 
$$p(a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2, \dots, a_n \leq x_n \leq b_n) = \sum_{x_1=a_1}^{x_1=b_1} \sum_{x_2=a_2}^{x_2=b_2} \dots \sum_{x_n=a_n}^{x_n=b_n} p(x_1, x_2, \dots, x_n) \quad (4.22)$$

Las funciones marginales están dadas como:

$$p_1(x_1) = \sum_{x_2=a_2}^{x_2=b_2} \sum_{x_3=a_3}^{x_3=b_3} \dots \sum_{x_n=a_n}^{x_n=b_n} p(x_1, x_2, \dots, x_n)$$

iv. 
$$p_2(x_2) = \sum_{x_1=a_1}^{x_1=b_1} \sum_{x_3=a_3}^{x_3=b_3} \dots \sum_{x_n=a_n}^{x_n=b_n} p(x_1, x_2, \dots, x_n) \quad (4.23)$$

⋮

$$p_n(x_n) = \sum_{x_1=a_1}^{x_1=b_1} \sum_{x_3=a_3}^{x_3=b_3} \dots \sum_{x_{n-1}=a_{n-1}}^{x_{n-1}=b_{n-1}} p(x_1, x_2, \dots, x_n)$$

v. La función de probabilidad condicional de  $x_1$  y su valor esperado están dadas por las siguientes expresiones:

$$p_{x_1|x_2, \dots, x_n}(x_1) = \frac{p(x_1, x_2, \dots, x_n)}{\sum_{\forall x_1} p(x_1, x_2, \dots, x_n)} \quad (4.24)$$

$$E \{x_1 | x_2, x_3, \dots, x_n\} = \sum_{\forall x_1} x_1 p_{x_1 | x_2, \dots, x_n}(x_1) \quad (4.25)$$

A la gráfica de la esperanza de la expresión 4.25 se le denomina Regresión de  $x_1$  sobre  $(x_2, x_3, \dots, x_n)$ .

Las variables  $(x_1, x_2, \dots, x_n)$  son variables aleatorias independientes, sí y sólo sí, para toda  $(x_1, x_2, \dots, x_n)$  se cumple que:

$$p(x_1, x_2, \dots, x_n) = p(x_1) p(x_2) \dots p(x_n) \quad (4.26)$$

#### Ejercicio 4.6

El porcentaje de concentración de madera dura en la pulpa bruta  $x$ , la presión de la cuba  $y$ , y el tiempo de cocción de la pulpa  $z$ , son tres variables aleatorias que afectan la resistencia del papel que se fabrica en la Fábrica de Papel Tuxtepec Oaxaca, FaPaTux. La función de probabilidad conjunta de las tres variables está dada por la siguiente tabla:

**FIGURA 4.22.**

$x$	$y$	$z$	$p(x,y,z)$
2	3	400	0.05506
2	3	500	0.05522
2	3	650	0.05599
2	4	400	0.05568
2	4	500	0.05611
2	4	650	0.05631
4	3	400	0.0555
4	3	500	0.05511
4	3	650	0.05554
4	4	400	0.05548
4	4	500	0.05564
4	4	650	0.0559
8	3	400	0.05527
8	3	500	0.05495
8	3	650	0.05547
8	4	400	0.05554
8	4	500	0.05537
8	4	650	0.05586
Suma=			1

- a) Compruebe que cumple las características propias de una función de probabilidad discreta.

Como se puede apreciar en la tabla 4.22,  $p(x,y,z) > 0$  para todos los valores de  $x$ ,  $y$  y  $z$ . Además, se muestra que la suma total de todos los valores de  $p(x,y,z)$  es uno.

- b) Calcule la probabilidad de que  $p(x \leq 4, y=4, z \leq 500)$

$$p(x \leq 4, y=4, z \leq 500) = \sum_{x=2}^{x=4} \sum_{y=4}^{y=4} \sum_{z=400}^{z=500} p(x, y, z)$$

$$p(x \leq 4, y=4, z \leq 500) = p(2,4,400) + p(2,4,500) + p(4,4,400) + p(4,4,500)$$

$$p(x \leq 4, y=4, z \leq 500) = 0.056106 + 0.056311 + 0.055483 + 0.055638$$

$$p(x \leq 4, y=4, z \leq 500) = 0.223538$$

- c) Obtenga su función de probabilidad conjunta acumulada:

$x$	$y$	$z$	$p(x,y,z)$	$P(x,y,z)$
2	3	400	0.055063	0.055063
2	3	500	0.055217	0.110280
2	3	650	0.055988	0.166268
2	4	400	0.055681	0.110744
2	4	500	0.056106	0.222067
2	4	650	0.056311	0.334366
4	3	400	0.055497	0.110560
4	3	500	0.055105	0.220882
4	3	650	0.055539	0.332409
4	4	400	0.055483	0.221724
4	4	500	0.055638	0.443790
4	4	650	0.055904	0.667532
8	3	400	0.055273	0.165833
8	3	500	0.054953	0.331108
8	3	650	0.055469	0.498104
8	4	400	0.055539	0.332536
8	4	500	0.055372	0.664927
8	4	650	0.055862	1.000000

d) Determine sus funciones marginales de probabilidad de  $x$ ,  $y$  y  $z$ :

$x$	$p_x(x)$	$y$	$p_y(y)$	$z$	$p_z(z)$
2	0.334366	3	0.498104	400	0.332536
4	0.333166	4	0.501896	500	0.332391
8	0.332468			650	0.335073

e) Calcule las marginales  $p(x,y)$ ,  $p(x,z)$  y  $p(y,z)$

$x$	$y$	$p_{xy}(x,y)$	$x$	$z$	$p_{xz}(x,z)$	$y$	$z$	$p_{yz}(y,z)$
2	3	0.16627	2	400	0.11074	3	400	0.16583
2	4	0.1681	2	500	0.11132	3	500	0.16528
4	3	0.16614	2	650	0.1123	3	650	0.167
4	4	0.16703	4	400	0.11098	4	400	0.1667
8	3	0.1657	4	500	0.11074	4	500	0.16712
8	4	0.16677	4	650	0.11144	4	650	0.16808
			8	400	0.11081			
			8	500	0.11033			
			8	650	0.11133			

f) Obtenga las funciones de probabilidad condicionales  $p(x|y,z)$ ,  $p(y|x,z)$  y  $p(z|x,y)$

$x$	$y$	$z$	$p(x,y,z)$	$p_{yz}(y,z)$	$p(x y,z)$	$p_{xz}(x,z)$	$p(y x,z)$	$p_{xy}(x,y)$	$p(z x,y)$
2	3	400	0.055063	0.16583	0.33204	0.11074	0.49721	0.16627	0.33117
2	3	500	0.055217	0.16528	0.33409	0.11132	0.49601	0.16627	0.3321
2	3	650	0.055988	0.167	0.33527	0.1123	0.49856	0.16627	0.33673
2	4	400	0.055681	0.1667	0.33401	0.11074	0.50279	0.1681	0.33124
2	4	500	0.056106	0.16712	0.33573	0.11132	0.50399	0.1681	0.33377
2	4	650	0.056311	0.16808	0.33503	0.1123	0.50144	0.1681	0.33499
4	3	400	0.055497	0.16583	0.33466	0.11098	0.50006	0.16614	0.33404
4	3	500	0.055105	0.16528	0.33341	0.11074	0.49759	0.16614	0.33168
4	3	650	0.055539	0.167	0.33258	0.11144	0.49836	0.16614	0.33429
4	4	400	0.055483	0.1667	0.33283	0.11098	0.49994	0.16703	0.33218
4	4	500	0.055638	0.16712	0.33293	0.11074	0.50241	0.16703	0.33311
4	4	650	0.055904	0.16808	0.33261	0.11144	0.50164	0.16703	0.3347
8	3	400	0.055273	0.16583	0.33331	0.11081	0.4988	0.1657	0.33358
8	3	500	0.054953	0.16528	0.33249	0.11033	0.4981	0.1657	0.33165
8	3	650	0.055469	0.167	0.33216	0.11133	0.49823	0.1657	0.33477
8	4	400	0.055539	0.1667	0.33316	0.11081	0.5012	0.16677	0.33302
8	4	500	0.055372	0.16712	0.33134	0.11033	0.5019	0.16677	0.33202
8	4	650	0.055862	0.16808	0.33236	0.11133	0.50177	0.16677	0.33496

Note que en el caso de estas funciones de probabilidad condicionales, no puede sumarse literalmente todo, la función de probabilidad condicional está acotada específicamente al valor que tomen sus condiciones, es decir, para un solo caso:

$x$	$y$	$z$	$p(x y,z)$	$xp(x y,z)$
2	3	400	0.33204	0.66408
4	3	400	0.33466	1.33862
8	3	400	0.33331	2.66644
Suma=			1	4.66914

g) Obtenga los valores esperados de  $x$  con respecto de  $y$  y  $z$ , de  $y$  con respecto de  $x$  y  $z$ , y de  $z$  con respecto de  $x$  y  $y$ .

$y$	$z$	$E\{x y,z\}$	$x$	$z$	$E\{y x,z\}$	$x$	$y$	$E\{z x,y\}$
3	400	4.66914	2	400	3.50279	2	3	517.393
3	500	4.66179	2	500	3.50399	2	4	517.124
3	650	4.6581	2	650	3.50144	4	3	516.74
4	400	4.66462	4	400	3.49994	4	4	516.987
4	500	4.65389	4	500	3.50241	8	3	516.857
4	650	4.65938	4	650	3.50164	8	4	516.942
			8	400	3.5012			
			8	500	3.5019			
			8	650	3.50177			

h) Determine si las variables  $x$ ,  $y$  y  $z$  son estadísticamente independientes.

$x$	$y$	$z$	$p(x,y,z)$	$p(x)p(y)p(z)$
2	3	400	0.055063	0.055384
2	3	500	0.055217	0.055359
2	3	650	0.055988	0.055806
2	4	400	0.055681	0.055805
2	4	500	0.056106	0.055781
2	4	650	0.056311	0.056231
4	3	400	0.055497	0.055185
4	3	500	0.055105	0.055161
4	3	650	0.055539	0.055606
4	4	400	0.055483	0.055605
4	4	500	0.055638	0.055581
4	4	650	0.055904	0.056029

$x$	$y$	$z$	$p(x,y,z)$	$p(x)p(y)p(z)$
8	3	400	0.055273	0.055069
8	3	500	0.054953	0.055045
8	3	650	0.055469	0.055489
8	4	400	0.055539	0.055488
8	4	500	0.055372	0.055464
8	4	650	0.055862	0.055912

Observe de la tabla anterior que en general  $p(x,y,z)$  es diferente de  $p(x)p(y)p(z)$ , por lo que las variables  $x$ ,  $y$  y  $z$  no son estadísticamente independientes.

#### 4.4.2. Funciones de probabilidad con $n$ variables aleatorias conjuntas continuas

i.  $f(x_1, x_2, \dots, x_n) \geq 0$

ii. 
$$\int_{\forall x_1} \int_{\forall x_2} \dots \int_{\forall x_n} f(x_1, x_2, \dots, x_n) dx_n \dots dx_2 dx_1 = 1 \quad (4.27)$$

Además, se cumple que:

iii. 
$$p(a_1 \leq x_1 \leq b_1, a_2 \leq x_2 \leq b_2, \dots, a_n \leq x_n \leq b_n) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \dots \int_{a_n}^{b_n} f(x_1, x_2, \dots, x_n) dx_n \dots dx_2 dx_1 \quad (4.28)$$

Las funciones marginales están dadas como:

$$f_1(x_1) = \int_{\forall x_2} \int_{\forall x_3} \dots \int_{\forall x_n} f(x_1, x_2, \dots, x_n) dx_n \dots dx_3 dx_2$$

$$f_2(x_2) = \int_{\forall x_1} \int_{\forall x_3} \dots \int_{\forall x_n} f(x_1, x_2, \dots, x_n) dx_n \dots dx_3 dx_1 \quad (4.29)$$

iv.

⋮

$$f_n(x_n) = \int_{\forall x_1} \int_{\forall x_2} \dots \int_{\forall x_{n-1}} f(x_1, x_2, \dots, x_n) dx_n \dots dx_2 dx_1$$

v. La función de probabilidad condicional de  $x_1$  y su valor esperado están dadas por las siguientes expresiones:



$$f_{x_1|x_2, \dots, x_n}(x_1) = \frac{f(x_1, x_2, \dots, x_n)}{\int_{\forall x_1} f(x_1, x_2, \dots, x_n) dx_1} \quad (4.30)$$

$$E\{x_1|x_2, x_3, \dots, x_n\} = \int_{\forall x_1} x_1 f_{x_1|x_2, \dots, x_n}(x_1) dx_1 \quad (4.31)$$

A la gráfica de la esperanza de la expresión 4.27 se le denomina Regresión de  $x_1$  sobre  $(x_2, x_3, \dots, x_n)$ .

Las variables  $(x_1, x_2, \dots, x_n)$  son variables aleatorias independientes, sí y sólo sí, para toda  $(x_1, x_2, \dots, x_n)$  se cumple que:

$$f(x_1, x_2, \dots, x_n) = f_1(x_1) f_2(x_2) \dots f_n(x_n) \quad (4.32)$$

## Ejercicio 4.7

Sea la función de probabilidad continua

$$w = f(x, y, z) = \left\{ \begin{array}{ll} & \left. \begin{array}{l} 0 \leq x \leq 3 \\ \frac{1}{3} xy^2 z^3 \\ 0 \leq y \leq 2 \\ 0 \leq z \leq 1 \\ 0 \quad \text{cualquier otro caso} \end{array} \right\}$$

- a) Compruebe que cumple las características propias de una función de probabilidad continua.

Como se puede apreciar de la función dada anteriormente,  $f(x, y, z) \geq 0$  para todos los valores de  $x, y$  y  $z$  del dominio de definición. Además, se puede apreciar que la integral triple de la función en el intervalo de definición da uno:

$$\begin{aligned} \int_{\forall x} \int_{\forall y} \int_{\forall z} f(x, y, z) dz dy dx &= \frac{1}{3} \int_0^3 \int_0^2 \int_0^1 xy^2 z^3 dz dy dx \\ &= \frac{1}{3} \int_0^3 \int_0^2 xy^2 \left[ \frac{z^4}{4} \right]_0^1 dy dx = \frac{1}{12} \int_0^3 \int_0^2 xy^2 dy dx \\ &= \frac{1}{12} \int_0^3 x \left[ \frac{y^3}{3} \right]_0^2 dx = \frac{2}{9} \int_0^3 x dx = \left[ \frac{x^2}{9} \right]_0^3 = 1 \end{aligned}$$

- b) Calcule la probabilidad de que  $p(1 \leq x \leq 2, y \leq 1, \frac{1}{4} \leq z \leq \frac{3}{4})$

$$p(1 \leq x \leq 2, y \leq 1, \frac{1}{4} \leq z \leq \frac{3}{4}) = \frac{1}{3} \int_1^2 \int_0^1 \int_{\frac{1}{4}}^{\frac{3}{4}} xy^2 z^3 dz dy dx$$

$$\begin{aligned}
&= \frac{1}{3} \int_1^2 \int_0^1 xy^2 \left[ \frac{z^4}{4} \right]_{\frac{1}{4}}^{\frac{3}{4}} dy dx = \frac{1}{12} \int_1^2 \int_0^1 xy^2 \left[ \frac{81}{256} - \frac{1}{256} \right] dy dx \\
&= \frac{1}{12} \left[ \frac{80}{256} \right] \int_1^2 x \left[ \frac{y^3}{3} \right]_0^1 dx = \frac{1}{12} \left[ \frac{5}{16} \right] \frac{1}{3} \int_1^2 x dx \\
&= \frac{5}{576} \left[ \frac{x^2}{2} \right]_1^2 = \frac{5}{1152} [4-1] = \frac{5}{384} = 0.013021
\end{aligned}$$

c) Obtenga su función de probabilidad conjunta acumulada:

$$F(x,y,z) = p(u \leq x, v \leq y, w \leq z) = \frac{1}{3} \int_0^x \int_0^y \int_0^z uv^2w^3 dw dv du$$

$$F(x,y,z) = \frac{1}{3} \int_0^x \int_0^y \left[ \frac{w^4}{4} \right]_0^z uv^2 dv du = \frac{1}{12} \int_0^x \int_0^y uv^2 z^4 dv du$$

$$F(x,y,z) = \frac{1}{12} \int_0^x u \left[ \frac{v^3}{3} \right]_0^y z^4 du = \frac{1}{36} \int_0^x y^3 z^4 u du$$

$$F(x,y,z) = \frac{1}{36} \left[ \frac{u^2}{2} \right]_0^x y^3 z^4 = \frac{1}{72} x^2 y^3 z^4$$

La función de probabilidad acumulada queda definida como:

$$F(x,y,z) = \frac{1}{72} x^2 y^3 z^4 \quad \begin{array}{l} 0 \leq x \leq 3 \\ 0 \leq y \leq 2 \\ 0 \leq z \leq 1 \end{array}$$

Cuando se analizó el caso de una función de probabilidad de una sola variable en el dominio de  $x$ , se dijo que si  $f(x)$  es la función de densidad y  $F(x)$  es la función de probabilidad acumulada, por la regla de Barrow del Cálculo, se puede afirmar que:

$$\text{Si } f(x) \geq 0$$

$$\text{y } \int_{v_x} f(x) dx = 1$$

$$\text{y } \exists F(x) = \int_{-\infty}^x f(t) dt$$

Entonces

$$p(a \leq x \leq b) = \int_a^b f(x) dx = F(b) - F(a)$$

Esta afirmación no se puede generalizar para el caso de varias variables, es decir

$$\text{Si } f(x,y,z) \geq 0$$

$$\text{y } \int_{\forall x} \int_{\forall y} \int_{\forall z} f(x,y,z) dz dy dx = 1$$

$$\text{y } \exists F(x,y,z) = \int_{-\infty}^x \int_{-\infty}^y \int_{-\infty}^z f(u,v,w) dw dv du$$

Entonces

$$p(a \leq x \leq b, c \leq y \leq d, e \leq z \leq f) = \int_a^b \int_c^d \int_e^f f(x,y,z) dz dy dx \neq F(b,d,f) - F(a,c,e)$$

De los dos últimos incisos observe que

$$p\left(1 \leq x \leq 2, y \leq 1, \frac{1}{4} \leq z \leq \frac{3}{4}\right) = \frac{1}{3} \int_1^2 \int_0^1 \int_{\frac{1}{4}}^{\frac{3}{4}} xy^2 z^3 dz dy dx = 0.013021$$

En cambio

$$F\left(2, 1, \frac{3}{4}\right) - F\left(1, 0, \frac{1}{4}\right) = \frac{1}{72}(2)^2(1)^3 \left(\frac{3}{4}\right)^4 - \frac{1}{72}(1)^2(0)^3 \left(\frac{1}{4}\right)^4 = 0.017578$$

d) Determine sus funciones marginales de probabilidad  $g(x)$ ,  $h(y)$  y  $l(z)$ :

$$g(x) = \frac{1}{3} \int_0^2 \int_0^1 xy^2 z^3 dz dy = \frac{x}{12} \int_0^2 y^2 [z^4]_0^1 dy = \frac{1}{12} x \int_0^2 y^2 dy = \frac{1}{36} x [y^3]_0^2 = \frac{2}{9} x$$

$$h(y) = \frac{1}{3} \int_0^3 \int_0^1 xy^2 z^3 dz dx = \frac{y^2}{12} [z^4]_0^1 \int_0^3 x dx = \frac{y^2}{24} [x^2]_0^3 = \frac{3}{8} y^2$$

$$l(z) = \frac{1}{3} \int_0^3 \int_0^2 xy^2 z^3 dy dx = \frac{z^3}{9} \int_0^3 x [y^3]_0^2 dx = \frac{8z^3}{9} \int_0^3 x dx = \frac{4z^3}{9} [x^2]_0^3 = 4z^3$$

e) Calcule las marginales  $u(x,y)$ ,  $v(x,z)$  y  $w(y,z)$ :

$$u(x,y) = \frac{1}{3} \int_0^1 xy^2 z^3 dz = \frac{1}{12} xy^2 [z^4]_0^1 = \frac{1}{12} xy^2$$

$$v(x,z) = \frac{1}{3} \int_0^2 xy^2 z^3 dy = \frac{1}{9} xz^3 [y^3]_0^2 = \frac{8}{9} xz^3$$

$$w(y,z) = \frac{1}{3} \int_0^3 xy^2 z^3 dx = \frac{1}{6} y^2 z^3 [x^2]_0^3 = \frac{3}{2} y^2 z^3$$

- f) Obtenga las funciones de probabilidad condicionales  $q(x|y,z)$ ,  $r(y|x,z)$  y  $s(z|x,y)$ :

$$q(x|y,z) = \frac{f(x,y,z)}{w(y,z)} = \frac{\frac{1}{3}xy^2z^3}{\frac{3}{2}y^2z^3} = \frac{2}{9}x$$

$$r(y|x,z) = \frac{f(x,y,z)}{v(x,z)} = \frac{\frac{1}{3}xy^2z^3}{\frac{8}{9}xz^3} = \frac{3}{8}y^2$$

$$s(z|x,y) = \frac{f(x,y,z)}{u(x,y)} = \frac{\frac{1}{3}xy^2z^3}{\frac{1}{12}xy^2} = 4z^3$$

- g) Obtenga los valores esperados de  $x$  con respecto de  $y$  y  $z$ , de  $y$  con respecto de  $x$  y  $z$ , y de  $z$  con respecto de  $x$  y  $y$ .

$$E\{x|y,z\} = \int_{\forall x} xq(x|y,z)dx = \frac{2}{9} \int_0^3 x^2 dx = \frac{2}{27} [x^3]_0^3 = 2$$

$$E\{y|x,z\} = \int_{\forall y} yr(y|x,z)dy = \frac{3}{8} \int_0^2 y^3 dy = \frac{3}{32} [y^4]_0^2 = \frac{3}{2}$$

$$E\{z|x,y\} = \int_{\forall z} zs(z|x,y)dz = 4 \int_0^1 z^4 dz = \frac{4}{5} [z^5]_0^1 = \frac{4}{5}$$

- h) Determine si las variables  $x$ ,  $y$  y  $z$  son estadísticamente independientes.

Observe que  $f(x,y,z) = g(x)h(y)l(z)$ , es decir

$$f(x,y,z) = \frac{1}{3}xy^2z^3 = \left(\frac{2}{9}x\right)\left(\frac{3}{8}y^2\right)(4z^3)$$

Por lo que las variables  $x$ ,  $y$  y  $z$  son estadísticamente independientes.

#### 4.5 Características de combinaciones lineales de variables aleatorias

Sean  $x_1, x_2, \dots, x_n$

variables aleatorias estadísticamente independientes,  
con funciones de probabilidad:

$$f_1(x_1), f_2(x_2), \dots, f_n(x_n)$$

de medias:

$$\mu_{x_1}, \mu_{x_2}, \dots, \mu_{x_n}$$

y varianzas:

$$\sigma_{x_1}^2, \sigma_{x_2}^2, \dots, \sigma_{x_n}^2,$$

y sean:  $a_1, a_2, \dots, a_n \in \mathfrak{R}$

De acuerdo con la expresión 3.21:

$$\mu_y = a_1 \mu_{x_1} + a_2 \mu_{x_2} + \dots + a_n \mu_{x_n} \quad (4.33)$$

Con base en la expresión 3.36:

$$\sigma_y^2 = a_1^2 \sigma_{x_1}^2 + a_2^2 \sigma_{x_2}^2 + \dots + a_n^2 \sigma_{x_n}^2 \quad (4.34)$$

## Ejercicios propuestos del capítulo 4

- 4.1. En una práctica de baloncesto un jugador se dispone a efectuar tres tiros libres. Se considera que la probabilidad que tiene el jugador de anotar en cada tiro independiente es de 0.6. El número de canastas que el jugador anote en sus dos primeros tiros se representará con una variable aleatoria  $x$  y el número de canastas que anote en sus dos últimos tiros se representará con la variable aleatoria  $y$ .
- Obtener el espacio muestral del experimento aleatorio asignando el valor que le corresponda a las variables aleatorias  $x$  y  $y$  en cada evento.
  - Construir una tabla que represente una función de probabilidad conjunta para las variables aleatorias  $x$  y  $y$ .
- 4.2. Una moneda homogénea se lanza cuatro veces, sea  $x$  la variable aleatoria que denota el número de soles que salen y sea  $y$  la variable aleatoria que denota el número de soles en forma continua que salen.
- Determinar la distribución de probabilidad conjunta de  $(x, y)$ .
  - Determinar las distribuciones marginales de  $x$  y  $y$ .
  - Determinar las distribuciones de probabilidad condicionales  $p(x|y)$  y  $P(y|x)$ .
  - Calcular la covarianza  $\sigma_{xy}^2$  y el coeficiente de correlación  $\rho_{xy}$ .
- 4.3. Un gerente de existencias tiene los registros de demanda para el producto de su compañía durante los últimos 100 días. La variable aleatoria  $x$  representa el número de órdenes recibidas diariamente y la variable aleatoria  $y$  representa el número de unidades por orden. Sus datos se muestran en la siguiente tabla.

$x/y$	1	2	3	4	5	6	7	8	9
1	10	6	3	2	1	1	1	1	1
2	8	5	3	2	1	1	1		
3	8	5	2	1	1				
4	7	4	2	1	1				
5	6	3	1	1					
6	5	3	1	1					

- Obtenga las distribuciones marginales de  $x$  y  $y$ .
  - Obtenga todas las distribuciones condicionales para  $y$  dada  $x$ .
  - ¿Son independientes el número de órdenes recibidas y el tamaño de cada orden?
  - Calcule  $E\{x|y=2\}$
- 4.4. Un lote consta de  $N$  transistores y de estos son defectuosos  $M$  ( $M \leq N$ ). Se seleccionan aleatoriamente dos transistores sin reemplazo a partir de este lote y se determina si son o no defectuosos. La variable aleatoria:

$$X_i = \begin{cases} 1, & \text{Si el } i\text{ésimo transistor no es defectuoso } i=1,2 \\ 0, & \text{Si el } i\text{ésimo transistor es defectuoso} \end{cases}$$

- Determine la función probabilística conjunta para  $x_1$  y  $x_2$ .
  - ¿Cuáles son las funciones probabilísticas marginales para  $x_1$  y  $x_2$ ?
  - ¿Son  $x_1$  y  $x_2$  variables aleatorias independientes?
- 4.5. (Problema 4.1 de la página 123, del libro de Hines, Montgomery, Goldsman y Borror, *Probabilidad y Estadística para Ingeniería*, editorial Patria, cuarta edición, 2013). Un fabricante de refrigeradores sujeta el producto terminado a una inspección final. Le interesan dos categorías de defectos: rayones o fallas en el acabado porcelanizado y defectos mecánicos. El número de cada tipo de defecto es una variable aleatoria. En la siguiente tabla se muestra la función masa de probabilidad conjunta, en donde  $x$  representa la ocurrencia de defectos en el acabado y  $y$  representa la ocurrencia de defectos mecánicos.
- Obtenga las distribuciones marginales de  $x$  y  $y$ .
  - Obtenga la distribución probabilística de defectos mecánicos, dado que no hay defectos en el acabado.

- c) Obtenga la distribución probabilística de defectos en el acabado, dado que no hay defectos mecánicos.

$y \backslash x$	0	1	2	3	4	5
0	11/50	4/50	2/50	1/50	1/50	1/50
1	8/50	3/50	2/50	1/50	1/50	
2	4/50	3/50	2/50	1/50		
3	3/50	1/50				
4	1/50					

- 4.6. La función masa de probabilidad conjunta de dos variables aleatorias discretas  $x$ ,  $y$  está dada por:

$$p_{xy}(x,y) = \begin{cases} cxy & \begin{matrix} x = 1,2,3 \\ y = 1,2,3 \end{matrix} \\ 0 & \text{otro caso} \end{cases}$$

- Determinar el valor de  $c$  para que se verifique que se trata de una función de probabilidad conjunta.
- Obtener las funciones de probabilidad marginales  $g_x(x)$  y  $h_y(y)$ .
- Calcular las funciones de probabilidad condicionales  $p(x|y)$  y  $p(y|x)$ .
- Determinar la media y la varianza de  $x$  y de  $y$ .
- Obtener la covarianza  $\sigma_{xy}^2$ .
- Calcular el coeficiente de correlación  $\rho_{xy}$ .
- Investigar si las variables aleatorias  $x$  y  $y$  son independientes.
- Calcular las siguientes probabilidades:
  - $p(x=1, y=2)$
  - $p(x \leq 2, y \leq 2)$
  - $p(1 \leq x < 3, 1 < y < 3)$
  - $p(x \leq 2, y \geq 2)$
  - $p(x \leq 2 | y=2)$



4.7. Suponga que  $x$  y  $y$  tienen la siguiente distribución de probabilidad conjunta:

$f(x,y)$		$y$		
		1	2	3
$x$	1	0	1/5	2/15
	2	k	1/9	1/4
	3	1/12	0	1/18

- Obtenga el valor de  $k$  para que  $f(x,y)$  sea una función de probabilidad conjunta.
- Obtener las funciones de probabilidad marginales  $g_x(x)$  y  $h_y(y)$ .
- Obtener las funciones de probabilidad condicionales  $p(x|y)$  y  $p(y|x)$ .
- Calcular la media y la varianza de  $x$  y de  $y$ .
- Determinar la covarianza  $\sigma_{xy}^2$ .
- Obtener el coeficiente de correlación de  $x$  con  $y$ .
- Investigar si las variables aleatorias  $x$  y  $y$  son independientes.
- Calcular las siguientes probabilidades:
  - $p(x=1, y=2)$
  - $p(x \leq 2, y \leq 2)$
  - $p(1 \leq x < 3, 1 < y < 3)$
  - $p(x \leq 2, y \geq 2)$
  - $p(x \leq 2 | y=2)$

4.8. Dadas las variables aleatorias  $x$  y  $y$  con densidad de probabilidad conjunta definida como:

$$f_{x,y}(x,y) \begin{cases} \frac{1}{2}(x+3y) & 0 \leq x \leq 1 \\ & 0 \leq y \leq 1 \\ 0 & \text{Cualquier otro caso} \end{cases}$$

- Comprobar que se trata de una función de probabilidad conjunta.
- ¿Tiene sentido hablar de moda de una distribución de probabilidad conjunta? Obténgala.
- Obtener las funciones de probabilidad marginales  $g_x(x)$  y  $h_y(y)$ .
- Determinar las funciones de probabilidad condicionales  $p(x|y)$  y  $p(y|x)$ .
- Calcular la media y la varianza de  $x$  y de  $y$ .

- f) Determinar la covarianza  $\sigma_{x_1x_2}^2$ .
- g) Obtener el coeficiente de correlación de  $x_1$  con  $x_2$ .
- h) Investigar si las variables aleatorias  $x_1$  y  $x_2$  son independientes.
- i) Calcular las siguientes probabilidades:
  - i.  $p(x_1 < 0.3, x_2 < 0.5)$
  - ii.  $p(x_1 > 0.2, x_2 < 0.3)$
  - iii.  $p(0.4 < x_1 < 0.5, 0.2 < x_2 < 0.4)$
  - iv.  $p(0.4 < x_1 < 0.6 \mid x_2 < 0.5)$
  - v. Obtenga el valor de  $k$  tal que
 
$$p(\mu_{x_1} - k\sigma_{x_1} < x_1 < \mu_{x_1} + k\sigma_{x_1}, \mu_{x_2} - k\sigma_{x_2} < x_2 < \mu_{x_2} + k\sigma_{x_2}) = 0.90$$

- 4.9. Una característica de calidad de un producto se comporta como una variable aleatoria sujeta a dos variables aleatorias a su vez. Estas variables presentan intervalos de variación  $1 \leq x_1 \leq 3$ , y  $1 \leq x_2 \leq 4$ . La función de probabilidad conjunta de  $[x_1, x_2]$  es:

$$f(x_1, x_2) = k(5 - x_1 - x_2) \quad \begin{array}{l} 0 \leq x_1 \leq 2 \\ 1 \leq x_2 \leq 3 \end{array}$$

- a) Obtenga el valor de  $k$  para que  $f(x_1, x_2)$  sea una función de probabilidad conjunta.
  - b) Obtener las funciones de probabilidad marginales  $g_{x_1}(x_1)$  y  $h_{x_2}(x_2)$ .
  - c) Obtener las funciones de probabilidad condicionales  $p(x_1|x_2)$  y  $p(x_2|x_1)$ .
  - d) Calcular la media y la varianza de  $x_1$  y de  $x_2$ .
  - e) Determinar la covarianza  $\sigma_{x_1x_2}^2$ .
  - f) Obtener el coeficiente de correlación  $\rho_{x_1x_2}$ .
  - g) Investigar si las variables aleatorias  $x_1$  y  $x_2$  son independientes.
  - h) Calcular las siguientes probabilidades:
    - i.  $p(x_1 = 1, x_2 = 2)$
    - ii.  $p(x_1 \leq 1, x_2 \leq 2)$
    - iii.  $p(1 \leq x_1 < 2, 2 < x_2 < 3)$
    - iv.  $p(x_1 \leq 1, x_2 > 2)$
    - v.  $p(x_1 \leq 1 \mid x_2 \leq 2)$
- 4.10. Se mide en dos puntos de la calle la intensidad del sonido causado por el ruido general de los alrededores. Sean  $x_1$ , y  $x_2$  las variables aleatorias que representan la intensidad del sonido en los dos puntos. Si se sabe que la distribución conjunta de las variables aleatorias  $x_1$  y  $x_2$  está dada por:

$$f_{x_1, x_2} = \begin{cases} 6e^{-(2x_1+3x_2)} & x_1 \geq 0 \text{ y } x_2 \geq 0 \\ 0 & \text{c.o.c} \end{cases}$$

- a) Comprobar que se trata de una función de probabilidad conjunta.
- b) Obtener las funciones de probabilidad marginales  $g_{x_1}(x_1)$  y  $h_{x_2}(x_2)$ .
- c) Determinar las funciones de probabilidad condicionales  $p(x_1|x_2)$  y  $p(x_2|x_1)$ .
- d) Calcular la media y la varianza de  $x_1$  y de  $x_2$ .
- e) Determinar la covarianza  $\sigma_{x_1x_2}^2$ .
- f) Obtener el coeficiente de correlación de  $x$  con  $y$ .
- g) Investigar si las variables aleatorias  $x$  y  $y$  son independientes.
- h) Calcular las siguientes probabilidades:
  - i.  $p(x < 3, y < 0.4)$
  - ii.  $p(x > 0.2, y < 0.3)$
  - iii.  $p(0.4 < x < 0.5, 0.2 < y < 0.4)$
  - iv.  $p(0.4 < x < 0.6 | y < 0.5)$
  - v. Obtenga el valor de  $k$  tal que  $p(\mu_x - k\sigma_x < x < \mu_x + k\sigma_x, \mu_y - k\sigma_y < y < \mu_y + k\sigma_y) = 0.90$ .

4.11. Si  $x^*$  es la variable aleatoria estandarizada, esto es:

$$x^* = \frac{(x - E(x))}{\sigma_x} \quad \text{y} \quad y^* = \frac{(y - E(y))}{\sigma_y}$$

Establezca cómo debería ser el coeficiente de correlación  $\rho_{xy}$ , si  $x = -ay$ .

4.12. Suponga que la variable aleatoria  $x$  está uniformemente distribuida en el intervalo  $[0,1]$ . Asuma que la distribución condicional de  $y$  dado  $x = x$  tiene una distribución binomial con parámetros “ $n$ ” y “ $p = x$ ”, es decir:

$$P(y = y | x = x) = \binom{n}{y} x^y (1-x)^{n-y} \quad \forall y = 0, 1, 2, \dots, n$$

- i) Obtenga el valor esperado de  $y$
- j) Determine la distribución de  $y$
- k) Calcule  $E\{y = y | x = x\}$

## 5. Modelos Probabilísticos de Fenómenos Aleatorios Discretos

En este tema se analizan ciertos fenómenos reales donde se requiere considerar la incertidumbre y cuyos modelos probabilísticos discretos serán deducidos junto con sus principales propiedades, incluyendo ejemplos de aplicación en la práctica profesional de la Ingeniería.

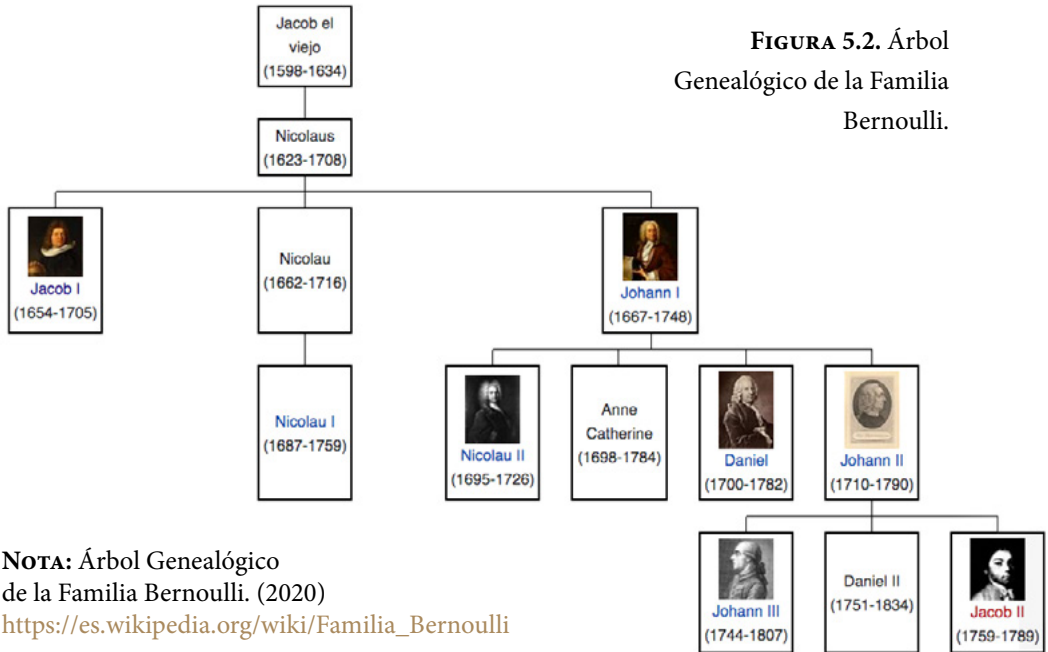
Los modelos probabilísticos que se verán para el caso discreto se resumen en la tabla de la figura 5.1.

FIGURA 5.1. Modelos Probabilísticos Discretos

Modelos Probabilísticos Discretos								
Distribución	Parámetros	Variable	Valores que toma $x$	Función de Probabilidad: $p(x)$	Función de Probabilidad Acumulada: $P(x)=p(t \leq x)$	Media	Varianza	Función Generatriz de Momentos
Bernoulli	$0 < p < 1$ , probabilidad de éxito en cada ensayo	$X =$ éxito en un ensayo	$x = 0$ , fracaso $x = 1$ , éxito	$p(x) = \begin{cases} p^x(1-p)^{(1-x)} & x = 0,1 \\ 0 & \text{en otro caso} \end{cases}$	$P(x) = p(t \leq x) = \begin{cases} 1-p & x = 0 \\ 1 & x = 1 \end{cases}$	$p$	$p(1-p)$	$fgm(t) = pe^t + (1-p)$
Binomial	$n = 1, 2, \dots$ , número de ensayos o tamaño de muestra $0 < p < 1$ , probabilidad constante de éxito en cada ensayo	$x =$ número de éxitos en $n$ ensayos	$x = 0, 1, 2, \dots, n$	$p(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{(n-x)} & x = 0, 1, 2, \dots, n \\ 0 & \text{en otro caso} \end{cases}$	$p(t \leq x) = \sum_{r=0}^{t \wedge x} \binom{n}{r} p^r (1-p)^{(n-r)}$	$np$	$np(1-p)$	$fgm(t) = [pe^t + (1-p)]^n$
Geométrica	$0 < p < 1$ , probabilidad de éxito en cada ensayo	$x =$ número de ensayos hasta tener el primer éxito	$x = 0, 1, 2, \dots$	$p(x) = \begin{cases} p(1-p)^{(x-1)} & x = 0, 1, 2, \dots \\ 0 & \text{en otro caso} \end{cases}$	$p(t \leq x) = \sum_{r=0}^{t \wedge x} p(1-p)^{(t-1)}$	$1/p$	$(1-p)/p^2$	$fgm(t) = pe^t / [1 - (1-p)e^t]$
Pascal (Binomial Negativa)	$0 < p < 1$ , probabilidad de éxito en cada ensayo $r = 1, 2, \dots (r > 0)$	$x =$ número de ensayos hasta tener $r$ éxitos	$x = r, r+1, r+2, \dots$	$p(x) = \begin{cases} \binom{x-1}{r-1} p^r (1-p)^{(x-1)} & x = 0, 1, 2, \dots \\ 0 & \text{en otro caso} \end{cases}$	$p(t \leq x) = \sum_{r=0}^{t \wedge x} \binom{t-1}{r-1} p^r (1-p)^{(t-r)}$	$r/p$	$\frac{r(1-p)}{p^2}$	$fgm(t) = \left[ \frac{pe^t}{1 - (1-p)e^t} \right]^r$
Hiper-geométrica	$N = 1, 2, \dots$ , tamaño del lote o población $n = 1, 2, \dots, N$ , tamaño de muestra $D = 1, 2, \dots, N$ , Número de Defectuosos en el Lote o Población	$x =$ número de éxitos en una muestra de tamaño $n$ , para un lote de tamaño $N$ , donde existen $D$ éxitos	$x = 0, 1, 2, \dots, \min(n, D)$	$p(x) = \begin{cases} \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}} & x = 0, 1, 2, \dots, \min(n, D) \\ 0 & \text{en otro caso} \end{cases}$	$p(t \leq x) = \sum_{r=0}^{t \wedge x} \frac{\binom{D}{r} \binom{N-D}{n-r}}{\binom{N}{n}}$	$n \left( \frac{D}{N} \right)$	$n \left( \frac{D}{N} \right) \left( 1 - \frac{D}{N} \right) \left( \frac{N-n}{N-1} \right)$	
Poisson	$c > 0$ , número de ocurrencias en promedio por unidad	$x =$ número de ocurrencias, éxitos o llegadas por unidad	$x = 0, 1, 2, \dots$	$p(x) = \begin{cases} e^{-c} \frac{c^x}{x!} & x = 0, 1 \\ 0 & \text{en otro caso} \end{cases}$	$p(t \leq x) = \sum_{r=0}^{t \wedge x} e^{-c} \frac{c^r}{r!}$	$c$	$c$	$fgm(t) = e^{c(e^t - 1)}$

## 5.1. Modelo probabilístico de Bernoulli

La familia Bernoulli produjo muchos artistas y científicos notables, en particular un gran número de matemáticos famosos del siglo XVIII, como se muestra en el árbol genealógico que se presenta en la figura 5.2.



La distribución de Bernoulli (o distribución dicotómica), nombrada así por el matemático suizo Jacob Bernoulli (1655-1705), es una distribución de probabilidad discreta, que toma valor 1 para la probabilidad de éxito  $p$  y valor 0 para la probabilidad de fracaso  $q = 1 - p$ . Este modelo corresponde a un experimento en el cual se realiza un solo ensayo, en una población con  $N$  elementos de los cuales existen  $D$  elementos con cierta característica de interés, lo cual lo hace interesante a ojos de quien realiza el experimento y que de obtenerse se le denomina éxito. Por ejemplo, en la baraja americana, con  $n = 52$  cartas, al evento de extraer un As se le considera un éxito; en este caso, la probabilidad de obtener un éxito es  $p = 4/52$ , ya que existen cuatro posibles ases en 52 cartas; asimismo, se aprecia que la probabilidad de no obtener un éxito es  $q = 1 - p = 48/52$ . De la misma forma, al lanzar un dado homogéneo de nueve caras y esperar que la cara que caiga hacia arriba sea un múltiplo de tres, a lo cual se le denominaría éxito, entonces la probabilidad de éxito es  $p = 3/9 = 1/3$ , ya que existen 3 caras cuyo número es múltiplo de 3: el 3, 6 y 9; la probabilidad de obtener un fracaso será  $q = 1 - p = 2/3$ .

En general, se le denominará éxito al evento que tenga la menor probabilidad de los dos posibles resultados, es decir, se considerará que  $p$  siempre es menor que  $q$ . De esta forma, por ejemplo, en un lote con  $N=10000$  artículos, de los cuales existen  $D=450$  defectuosos, la probabilidad de obtener un éxito será la de obtener un artículo defectuoso; en este caso  $p=450/10000=0.045$  y  $q=0.955$ .

Para representar un modelo de Bernoulli, se designa la  $x$  como variable representativa y se dice, como un símil del sistema binario, que  $x=0$  si se obtiene un fracaso y  $x=1$  si se obtiene un éxito. En general, la función de probabilidad de un modelo de Bernoulli, tendrá como función de probabilidad a la siguiente expresión matemática:

$$p(x) = \begin{cases} 1-p & x=0 \\ p & x=1 \end{cases} \quad (5.1)$$

Su función de probabilidad acumulada sería:

$$P(x) = p(t \leq x) = \begin{cases} 1-p & x=0 \\ 1 & x=1 \end{cases} \quad (5.2)$$

La media de la distribución de Bernoulli sería:

$$\mu_x = E\{x\} = (0)(1-p) + (1)(p) = p \quad (5.3)$$

Su varianza:

$$\sigma_x^2 = E\{x^2\} - E^2\{x\} = (0^2)(1-p) + (1^2)(p) - p^2 \quad (5.4)$$

$$\sigma_x^2 = p - p^2 = p(1-p)$$

El coeficiente de asimetría:

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{1-2p}{\sqrt{p(1-p)}} \quad (5.5)$$

El coeficiente de curtosis:

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{1-6p+6p^2}{p(1-p)} \quad (5.6)$$

La función generatriz de momentos:

$$FGM_x(t) E\{e^{tx}\} = e^0(1-p) + e^t p = 1-p + pe^t \quad (5.7)$$

Su función característica:

$$FC_x(t) = E\{e^{itx}\} = e^0(1-p) + e^{it} p = 1-p + pe^{it} \quad (5.8)$$

### Ejercicio 5.1

En un hospital de la zona metropolitana de la Ciudad de México se tienen  $n=520$  pacientes, de los cuales 50 son alérgicos a la penicilina. Se selecciona aleatoriamente por apellido a un paciente de este hospital. Obtenga la función de probabilidad de que sea alérgico a la penicilina, así como sus principales características.

Su función de probabilidad es:

$$p(x) = \left\{ \begin{array}{ll} 0.903846 & x=0 \\ 0.096154 & x=1 \\ 0 & \text{en cualquier caso} \end{array} \right\}$$

Su media y su varianza:

$$\mu_x = 0.096154$$

$$\sigma_x^2 = 0.086908$$

Su coeficiente de asimetría y su coeficiente de curtosis:

$$\gamma_1 = 2.739775$$

$$\gamma_2 = 5.506367$$

Su función generatriz de momentos:

$$FGM_x(t) = 0.90385 + 0.09615e^t$$

Su función característica:

$$FC_x(t) = 0.90385 + 0.09615e^{it}$$



## 5.2. Modelo Probabilístico Hipergeométrico $h(x; n, D, N)$

Este modelo corresponde a un experimento en el cual se realizan  $n$  ensayos de tipo Bernoulli, en una población con  $n$  artículos de los cuales existen  $D$  artículos con cierta característica de interés, que los hace ser interesantes a ojos de quien realiza el experimento y que en cada ensayo de obtenerse se le denomina éxito. Se designa la  $x$  como el número de éxitos obtenidos en  $n$  ensayos.

En este caso la expresión:

$$\binom{D}{x} = \frac{D!}{x!(D-x)!} = \frac{D(D-1)(D-2) \cdots (D-x+1)}{x!} \quad (5.9)$$

Representa el número de posibilidades diferentes que existen de extraer  $x$  elementos exitosos en la muestra de  $D$  que existen en la población. Asimismo,

$$\binom{N-D}{n-x} = \frac{(N-D)!}{(n-x)!(N-D-n+x)!} \quad (5.10)$$

Representa el número de posibilidades diferentes de extraer  $n-x$  elementos fracasados en la muestra de  $N-D$  que existen en la población.

Por el Teorema Multiplicativo Fundamental del Conteo, existen

$$\binom{D}{x} \binom{N-D}{n-x} \quad (5.11)$$

Maneras diferentes de extraer  $x$  elementos exitosos y  $(n-x)$  elementos fracasados en una muestra de tamaño  $n$ , obtenida de una población con  $n$  elementos en total, de los cuales existen  $D$  elementos exitosos. De la misma forma, existen

$$\binom{N}{n} \quad (5.12)$$

formas diferentes de extraer una muestra de tamaño  $n$  de una población de tamaño  $N$ .

Se parte del supuesto que la población de éxitos y fracasos se encuentra mezclada aleatoriamente en un solo contenedor y que cada ensayo se realiza de forma aleatoria y sin reemplazo, por lo que se puede suponer que los elementos posibles son equiprobables, es decir, tienen la misma posibilidad de ser extraídos, con lo cual es factible aplicar la definición clásica de probabilidad de Laplace. De esta forma, la probabilidad de obtener  $x$  elementos exitosos de una muestra de tamaño  $n$ , extraída de una población de tamaño  $N$ , donde existen  $D$  elementos exitosos mezclados con los  $N-D$  elementos fracasados, está dada por la siguiente expresión:

$$h(x;n,D,N) = \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}} \quad x = 0, 1, 2, \dots, \text{mínimo}(D, n) \quad (5.13)$$

La función de probabilidad acumulada hipergeométrica sería:

$$H(d;n,D,N) = p(x \leq d; n, D, N) = \sum_{x=0}^{x=d} \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}} \quad d = 0, 1, 2, \dots, \text{mínimo}(D, n) \quad (5.14)$$

Su media es:

$$\mu_x = E\{x\} = \sum_{x=0}^{x=\text{mínimo}(D,n)} x \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}} = n \left( \frac{D}{N} \right) = np \quad (5.15)$$

Su varianza:

$$\sigma_x^2 = E\{x^2\} - E^2\{x\} = \sum_{x=0}^{x=\text{mínimo}(D,n)} x^2 \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}} - n^2 \left( \frac{D}{N} \right)^2 = n \left( \frac{D}{N} \right) \left( 1 - \frac{D}{N} \right) \left( \frac{N-n}{N-1} \right) = np(1-p) \left( \frac{N-n}{N-1} \right) \quad (5.16)$$

Su coeficiente de asimetría:

$$\gamma_1 = \frac{(N-2D)(N-1)^{1/2} (N-2n)}{[nD(N-D)(N-n)]^{1/2} (N-2)} \quad (5.17)$$

Su coeficiente de curtosis:

$$Y_2 = \left[ \frac{N^2(N-1)}{n(N-2)(N-3)(N-n)} \right] \left[ \frac{N(N+1) - 6N(N-n)}{D(N-D)} + \frac{3n(N-n)(N+6)}{N^2} - 6 \right] \quad (5.18)$$

Su función generadora de momentos:

$$FGM_x(t) = \frac{\binom{N-D}{n} * {}_2F_1(-n, -D; N-D-n+1; e^t)}{\binom{N}{n}} \quad (5.19)$$

En donde:

$${}_2F_1(a, b; c; x) = \sum_{i=0}^{\infty} \frac{a^{(i)} b^{(i)} x^i}{c^{(i)} i!} \quad (5.20)$$

$$c \neq 0, -1, -2$$

$$m_{(i)} = m(m+1) \cdots (m+i+1) \quad m_0 = 1 \quad (5.21)$$

Cabe señalar que, para hacer el cálculo de probabilidades con la distribución hipergeométrica, utilizando Excel, se utiliza el siguiente comando:

$$p(x) = H(x; n, D, N, \text{acumulado}) = \text{DISTR.HIPERGEOM.N}(x, n, D, N, 0 \text{ ó } 1) \quad (5.22)$$

En donde  $x$  es el número de éxitos que se pretende obtener en la muestra,  $n$  es el tamaño de muestra o el número de ensayos en la muestra,  $D$  el número de éxitos en la población o lote,  $N$  el número de unidades en la población,  $\text{acumulado} = \text{verdadero}$  ó  $1$  si se pretende calcular la función de probabilidad acumulada desde cero hasta  $x$ ;  $\text{acumulado} = \text{falso}$  ó  $0$  si se pretende calcular la probabilidad de que el número de éxitos en la muestra sea exactamente  $x$ .

De tal manera que:

$$p(x=5; n=20, D=50, N=1000) = \text{DISTR.HIPERGEOM.N}(5, 20, 50, 1000, 0) = 0.001979$$

$$p(x \leq 5; n=20, D=50, N=1000) = \text{DISTR.HIPERGEOM.N}(5, 20, 50, 1000, 1) = 0.999738$$

$$p(x > 5; n=20, D=50, N=1000) = 1 - \text{DISTR.HIPERGEOM.N}(5, 20, 50, 1000, 1) = 0.000262$$

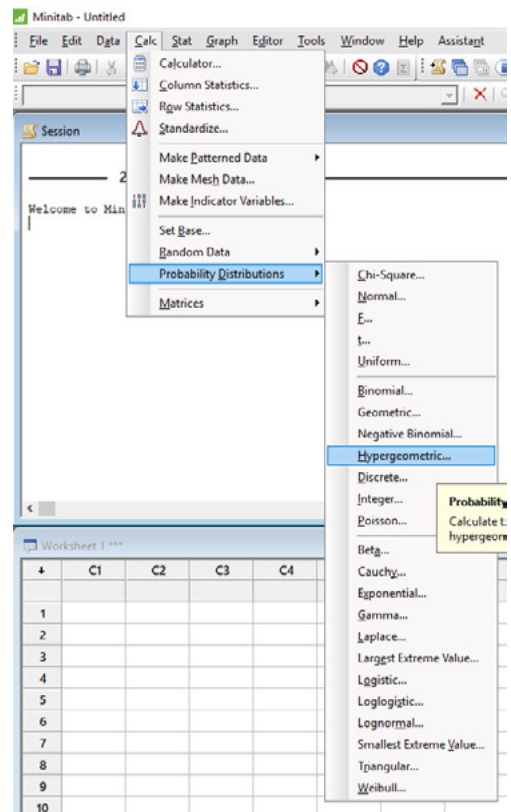
$$p(3 < x \leq 5; n=20, D=50, N=1000) = \text{DISTR.HIPERGEOM.N}(5,20,50,1000,1) - \text{DISTR.HIPERGEOM.N}(3,20,50,1000,1) = 0.014547$$

$$p(3 \leq x \leq 5; n=20, D=50, N=1000) = \text{DISTR.HIPERGEOM.N}(5,20,50,1000,1) - \text{DISTR.HIPERGEOM.N}(2,20,50,1000,1) = 0.073314$$

Para hacer el cálculo de probabilidades con la distribución hipergeométrica, utilizando Minitab, se realizan los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en Hypergeometric, como se muestra en la figura 5.3.

FIGURA 5.3



2. Aparece la pantalla que se muestra en la figura 5.3.a, en la cual, para calcular la probabilidad  $p(x=5, n=20, D=50, N=1000)$  se señala si se trata de una probabilidad en un punto marcando Probability, en Population size (N) se indica el tamaño de la población o lote, en Event count in population (M)

se indica el número de éxitos en la población  $D$ , en Sample size se teclea el tamaño de la muestra  $n$ , en Input constant se indica el valor de  $x$  para el cual se desea calcular la probabilidad y se da un click en Ok obteniéndose lo que se observa en la figura 5.4.b.

Si lo que se pretende calcular es la probabilidad acumulada hasta el valor  $x=5$ , en vez de elegir Probability debe elegirse Cumulative probability y todo lo demás igual, obteniéndose lo que se muestra en las figuras 5.5.a. y 5.5.b.

FIGURAS 5.4.a. y 5.4.b.

Hypergeometric Distribution

Probability  
 Cumulative probability  
 Inverse cumulative probability

Population size (N): 1000  
 Event count in population (M): 50  
 Sample size (n): 20

Input column:   
 Optional storage:   
 Input constant: 5  
 Optional storage:

29/09/2018 08:09:16 p. m.

Welcome to Minitab, press F1 for help.

### Probability Density Function

Hypergeometric with N = 1000, M = 50, and n = 20

x	P( X = x )
5	0.0019787

Para  $p(x \leq 5; n=20, D=50, N=1000)$

FIGURAS 5.5.a. y 5.5.b.

Hypergeometric Distribution

Probability  
 Cumulative probability  
 Inverse cumulative probability

Population size (N): 1000  
 Event count in population (M): 50  
 Sample size (n): 20

Input column:   
 Optional storage:   
 Input constant: 5  
 Optional storage:

29/09/2018 10:17:35 p. m.

Welcome to Minitab, press F1 for help.

### Cumulative Distribution Function

Hypergeometric with N = 1000, M = 50, and n = 20

x	P( X ≤ x )
5	0.999738

En el lenguaje de desarrollo R la sintaxis que se aplica en las distribuciones de probabilidad es una letra inicial seguida del nombre de la distribución. La letra inicial es *d* si se trata de la función de densidad de probabilidad. Si la letra inicial es *p* se trata de la función de probabilidad acumulada. Si la letra inicial es *q* se trata de calcular los cuartiles de la función de probabilidad analizada. Finalmente, si la letra inicial es la *r* se trata de generar números aleatorios con la distribución analizada.

En el caso de la distribución hipergeométrica, para el cálculo de probabilidades en un punto dado, se usa el comando:

$$\text{dhyper}(x, m, n, k, \text{log}=\text{FALSE}) \quad (5.23)$$

En donde  $x$  es el punto donde se quiere calcular la probabilidad;  $m$  es el número de artículos o eventos exitosos, o con cierta característica de interés, por ejemplo, ser defectuoso;  $n$  es el número de artículos complementarios en la población o en el lote que no son exitosos o que no presentan esa característica de interés, por ejemplo, ser no defectuoso;  $k$  es el tamaño de muestra o el número de unidades extraídas de la población o del lote,  $\text{log}$  es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma FALSE.

Por ejemplo,

$$p(x=5; n=20, D=50, N=1000)= \text{dhyper}(5,50,950,20,\text{log}=\text{FALSE})= 0.001979$$

Para el cálculo de la probabilidad acumulada hasta un punto dado, se usa el comando:

$$\text{phyper}(x, m, n, k, \text{lower.tail}=\text{TRUE}, \text{log}=\text{FALSE}) \quad (5.24)$$

En donde  $x$  es el punto hasta donde se quiere calcular la probabilidad acumulada;  $m$  es el número de artículos o eventos exitosos, o con cierta característica de interés, por ejemplo, ser defectuoso;  $n$  es el número de artículos complementarios en la población o en el lote que no son exitosos o que no presentan esa característica de interés, por ejemplo, ser no defectuoso;  $k$  es el número total de artículos en la población o en el lote,  $\text{log}$  es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma FALSE.  $\text{lower.tail}$  es un valor lógico, si se le asigna un valor

verdadero el valor que arroja es la probabilidad acumulada a la izquierda de  $x$ , de lo contrario, si el valor que toma es FALSE, el valor que arroja es la probabilidad a la derecha del valor de  $x$ .

Por ejemplo,

$$p(x \leq 5; n=20, D=50, N=1000) = \text{phyper}(5, 50, 950, 20, \text{lower.tail}=\text{TRUE}, \text{log}=\text{FALSE}) = 0.999738$$

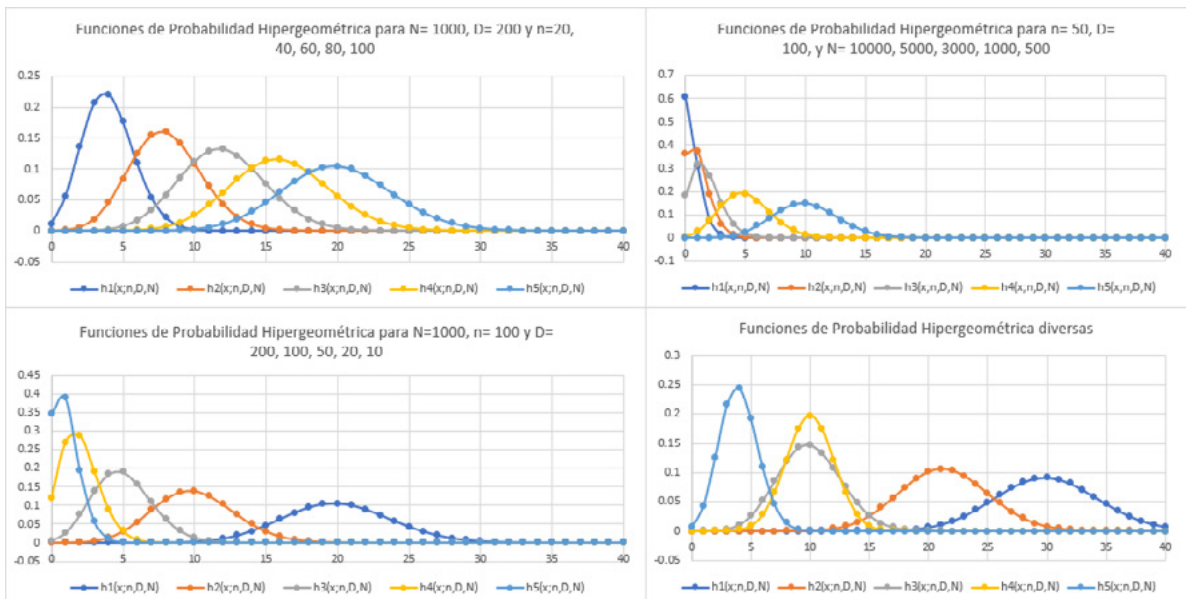
$$p(x > 5; n=20, D=50, N=1000) = \text{phyper}(5, 50, 950, 20, \text{lower.tail}=\text{FALSE}, \text{log}=\text{FALSE}) = 0.000262$$

$$p(3 < x \leq 5; n=20, D=50, N=1000) = \text{phyper}(5, 50, 950, 20, \text{lower.tail}=\text{TRUE}, \text{log}=\text{FALSE}) - \text{phyper}(3, 50, 950, 20, \text{lower.tail}=\text{TRUE}, \text{log}=\text{FALSE}) = 0.014547$$

$$p(3 \leq x \leq 5; n=20, D=50, N=1000) = \text{phyper}(5, 50, 950, 20, \text{lower.tail}=\text{TRUE}, \text{log}=\text{FALSE}) - \text{phyper}(2, 50, 950, 20, \text{lower.tail}=\text{TRUE}, \text{log}=\text{FALSE}) = 0.073314$$

En la figura 5.6 se muestran gráficas de la distribución hipergeométrica para diferentes valores de sus parámetros  $n$ ,  $D$  y  $N$ .

**FIGURA 5.6.** Gráficas de Distribuciones de Probabilidad Hipergeométricas



## Ejercicio 5.2

En los juegos de naipes existen dos tipos de barajas:

1. La baraja americana que consta de 4 figuras diferentes: picas ♠, corazones ♥, diamantes ♦ y tréboles ♣; para cada una de estas figuras existen los números del 1 al 10, el jack J, la reina Q y el rey R, lo que implica que la baraja americana consta de 52 cartas; al uno se le denomina el As, A, y es la carta de mayor peso en el juego de póker.
2. La baraja española que consta de 4 figuras diferentes: espadas 🏆, bastos 🏏, oros 🍀 y copas 🍷; para cada una de estas figuras existen los números del 1 al 7, además de la sota, el caballo y el rey, lo que implica que la baraja española consta de 40 cartas.

Suponga que a un jugador cualquiera, aleatoriamente se le reparten cinco cartas, sin derecho a descartarlas, es decir a cambiarlas por otras; el modelo probabilístico que representa al número de ases que le tocan en cada juego es hipergeométrico, como puede verse a continuación, para el caso de la baraja americana: hay un lote con  $N=52$  cartas, donde existen  $D=4$  ases, se asignan  $n=5$  cartas a un jugador; sea  $x$  = número de ases en la muestra, entonces, los valores que puede tomar  $x$  son 0, 1, 2, 3 ó 4. De esta forma, la probabilidad para cada valor de  $x$  es:

$$p(x=0) = \frac{\binom{4}{0} \binom{52-4}{5-0}}{\binom{52}{5}} = \frac{\binom{4}{0} \binom{48}{5}}{\binom{52}{5}} = \frac{48 \cdot 47 \cdot 46 \cdot 45 \cdot 44 \cdot 43!}{5! \cdot 43!} = \frac{48 \cdot 47 \cdot 46 \cdot 45 \cdot 44 \cdot 43!}{52 \cdot 51 \cdot 50 \cdot 49 \cdot 48 \cdot 47!} = \frac{5! \cdot 43!}{5! \cdot 47!} = 0.658842$$

Nótese que la probabilidad de que a un jugador no le den ases al repartirle cinco cartas es del 65.88%. De la misma forma:

$$p(x=1) = \frac{\binom{4}{1} \binom{52-4}{5-1}}{\binom{52}{5}} = \frac{\binom{4}{1} \binom{48}{4}}{\binom{52}{5}} = 0.299474$$

$$p(x=2) = \frac{\binom{4}{2} \binom{52-4}{5-2}}{\binom{52}{5}} = \frac{\binom{4}{2} \binom{48}{3}}{\binom{52}{5}} = 0.039930$$



$$p(x=3) = \frac{\binom{4}{3}\binom{52-4}{5-3}}{\binom{52}{5}} = \frac{\binom{4}{3}\binom{48}{2}}{\binom{52}{5}} = 0.001736$$

$$p(x=4) = \frac{\binom{4}{4}\binom{52-4}{5-4}}{\binom{52}{5}} = \frac{\binom{4}{4}\binom{48}{1}}{\binom{52}{5}} = 0.000018$$

Como puede apreciarse del último cálculo, la probabilidad de que a un jugador le den un póker de ases (4 ases) al repartirle cinco cartas es de 0.000018, es decir, si dicho jugador lleva a cabo un millón de juegos, en promedio en 18 de ellos le tocará un póker de ases.

x	h(x; n, D, N)	xh(x)	(x-media x) <sup>2</sup>	(x-media x) <sup>3</sup>	(x-media x) <sup>4</sup>
0	0.658842	0.000000	0.097462	-0.037485	0.014417
1	0.299474	0.299474	0.113410	0.069791	0.042948
2	0.039930	0.079860	0.104196	0.168316	0.271895
3	0.001736	0.005208	0.011875	0.031058	0.081229
4	0.000018	0.000074	0.000241	0.000873	0.003155
	Suma=	0.384615	0.327184	0.232552	0.413645

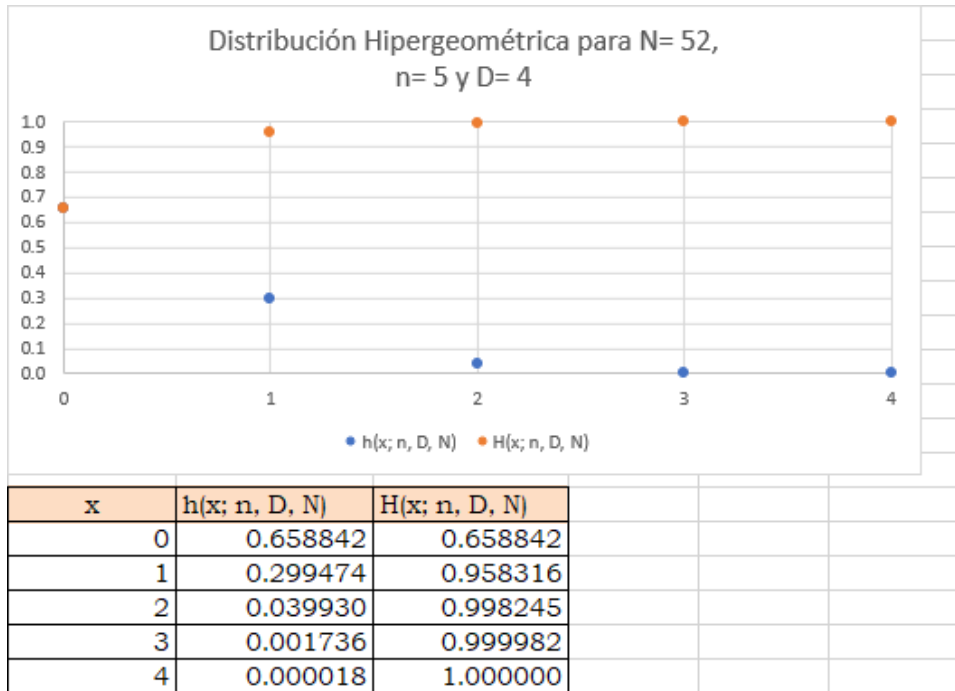
$$\mu_x = n \left( \frac{D}{N} \right) = 5 \left( \frac{4}{52} \right) = 0.384615$$

$$\sigma_x^2 = n \left( \frac{D}{N} \right) \left( 1 - \frac{D}{N} \right) \sqrt{\frac{N-n}{N-1}} = 5 \left( \frac{4}{52} \right) \left( \frac{48}{52} \right) \sqrt{\frac{47}{51}} = 0.327184$$

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{0.232552}{0.327184^{3/2}} = 1.242603$$

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = 0.864056$$

La gráfica del modelo probabilístico hipergeométrico de este ejemplo se muestra a continuación, junto con su función de probabilidad acumulada.

**Ejercicio 5.3**

En un salón de clases se tiene un grupo de  $N= 60$  alumnos, de los cuales  $D= 20$  son hombres. El profesor decide formar un equipo de  $n= 6$  alumnos; de acuerdo a la lista de asistencia, elige aleatoriamente al primero entre los seis primeros y de allí cada 10 alumnos elige a uno hasta completar los seis que requiere.

a) Calcule la probabilidad de que el primer alumno elegido sea hombre.

Para seleccionar al primero se usa un ensayo de Bernoulli, para el cual  $D=20$  y  $N= 60$ , por lo que  $p = D/N = 20/60 = 0.333333$

b) Obtenga la probabilidad de que se hayan elegido a tres hombres en el equipo.

$$p(x=3) = \frac{\binom{20}{3} \binom{60-20}{6-3}}{\binom{60}{6}} = \frac{\frac{20 * 19 * 18}{3!} \frac{40 * 39 * 38}{3!}}{\frac{60 * 59 * 58 * 57 * 56 * 55}{6!}} = 0.224977$$

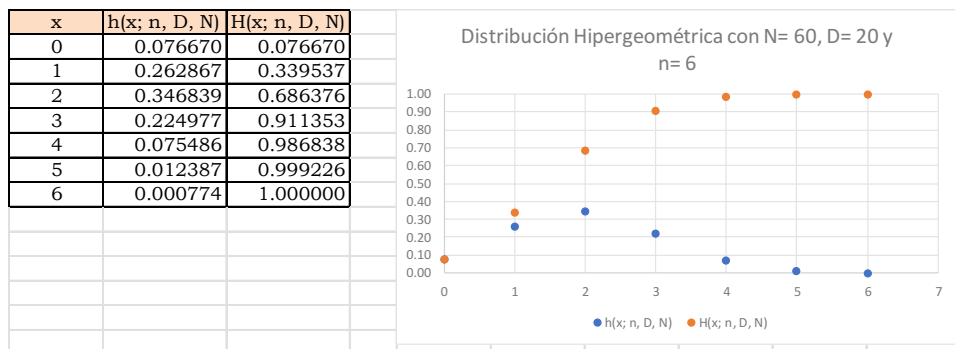
- c) Determine la probabilidad de que existan entre dos y cuatro mujeres en el equipo.

$$p(2 \leq x \leq 4) = \sum_{x=2}^{x=4} \frac{\binom{40}{x} \binom{20}{6-x}}{\binom{60}{6}} = \frac{\binom{40}{2} \binom{20}{4}}{\binom{60}{6}} + \frac{\binom{40}{3} \binom{20}{3}}{\binom{60}{6}} + \frac{\binom{40}{4} \binom{20}{2}}{\binom{60}{6}} = 0.6473$$

- d) Estime la probabilidad de que el equipo esté conformado por hombres.

$$p(x=6) = \frac{\binom{20}{6} \binom{60-20}{6-6}}{\binom{60}{6}} = \frac{20 * 19 * 18 * 17 * 16 * 15}{6!} = \frac{20 * 19 * 18 * 17 * 16 * 15}{60 * 59 * 58 * 57 * 56 * 55} = 0.00077$$

- e) Grafique la función de probabilidad para todos los valores que puede tomar  $x$  y su función de probabilidad acumulada en la misma figura.



- f) Obtenga la media, la mediana y la moda.

$$\mu_x = \sum_{x=0}^{x=6} xh(x) = 2$$

La mediana es el valor de  $x$  para el cual la probabilidad acumulada a la izquierda es 0.5 y la probabilidad acumulada a su derecha también es 0.5. En este caso, nótese que para  $x=1$ ,  $H(x=1)=0.339537$ , y para  $x=2$ ,  $H(x)=0.686376$ , por lo que la mediana cae entre  $x=1$  y  $x=2$ . Se aproximará la mediana por interpolación lineal. Se trata de obtener el valor de  $x$  para el cual  $H(x)=0.5$ .

La ecuación de una recta dados dos de sus puntos está dada por:

$$m_e = \frac{x_2 - x_1}{y_2 - y_1} (0.5 - y_1) + x_1$$

$$m_e = \frac{(2-1)(0.5-0.339537)}{0.686376-0.339537} + 1 = 1.46264$$

La moda es el valor de  $x$  más frecuente, o sea, aquél que proporciona el punto más alto de  $h(x)$ , como se puede apreciar en la gráfica la moda es  $m_o = 2$ .

g) Calcule la varianza, la desviación estándar y el coeficiente de variación.

$$\sigma_x^2 = \sum_{x=0}^{x=6} (x - \mu_x)^2 h(x) = 1.220339$$

$$\sigma_x = \sqrt{\sigma_x^2} = 1.10469$$

$$CV = \frac{\sigma_x}{\mu_x} = 0.55234$$

h) Determine el coeficiente de asimetría y el coeficiente de curtosis.

x	h(x; n, D, N)	H(x; n, D, N)	xh(x)	(x-media) <sup>2</sup> h(x)	(x-media) <sup>3</sup> h(x)	(x-media) <sup>4</sup> h(x)
0	0.076670	0.076670	0	0.306679	-0.61	1.23
1	0.262867	0.339537	0.26287	0.262867	-0.26	0.26
2	0.346839	0.686376	0.69368	0.000000	0.00	0.00
3	0.224977	0.911353	0.67493	0.224977	0.22	0.22
4	0.075486	0.986838	0.30194	0.301942	0.60	1.21
5	0.012387	0.999226	0.06194	0.111486	0.33	1.00
6	0.000774	1.000000	0.00465	0.012387	0.05	0.20
			2	1.220339	0.336645	4.123904

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = 0.24972$$

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = -0.2308$$

### Ejercicio 5.1

Suponga que  $N = 100$  interruptores termomagnéticos especiales son fabricados diariamente en un proceso productivo. Para asegurar la calidad de este proceso, el criterio de aceptación es elegir 10 interruptores y probar su disparo magnético, su disparo térmico y su continuidad; al menos 9 de ellos deben cumplir con las especificaciones. ¿Qué puede usted mencionar acerca de la bondad del plan de muestreo definido?, ¿es estricto o es laxo?

Sea  $x$  el número de interruptores que no cumple las especificaciones de disparo magnético, disparo térmico o continuidad. Como se puede apreciar, se tiene un lote de  $N=100$  interruptores de los cuales existen  $D$  que no cumplen alguna de las especificaciones, pero no se sabe cuánto vale  $D$ . Se toma una muestra aleatoria de  $n=10$  interruptores y debe haber al menos 9 interruptores “buenos” en la muestra, es decir, cuando mucho debe haber un defectuoso en la muestra. La probabilidad de que existan  $x$  defectuosos se comportaría como una distribución hipergeométrica. En este caso se considerará como interruptor defectuoso a aquel que no cumpla alguna de las especificaciones de disparo magnético, disparo térmico o continuidad.

En este caso, la probabilidad de aceptar el lote sería  $\beta = p(x \leq 1)$ .

$$\beta = p(x \leq 1) = h(x \leq 1; n=10, D, N=100) = \sum_{x=0}^{x=1} \frac{\binom{D}{x} \binom{100-D}{10-x}}{\binom{100}{10}}$$

Que también puede ser escrito de la siguiente forma:

$$\beta = \frac{\binom{D}{0} \binom{100-D}{10}}{\binom{100}{10}} + \frac{\binom{D}{1} \binom{100-D}{9}}{\binom{100}{10}}$$

Nótese que la probabilidad de aceptar el lote o rechazarlo depende del valor que tome  $D$ , el cual representa el número de artículos defectuosos en el lote.  $D$  en este caso puede tomar valores desde cero hasta  $N=100$ . Se graficará  $D$  contra  $\beta$ .

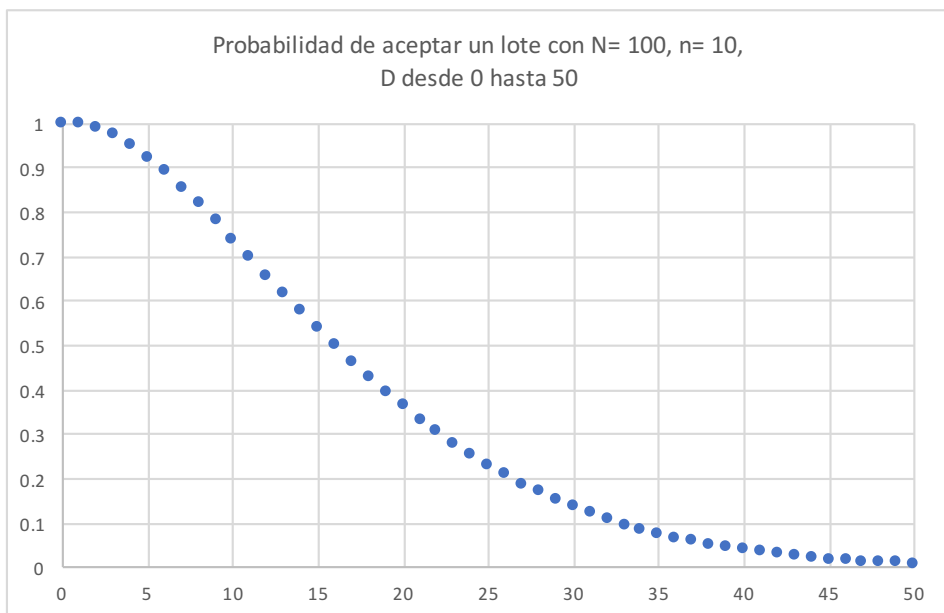
$D$	$\beta=p(x \leq 1)$	25	0.22928
0	1	26	0.20751
1	1	27	0.18732
2	0.99091	28	0.16865
3	0.97421	29	0.15144
4	0.95123	30	0.13562
5	0.92314	31	0.12113
6	0.89099	32	0.1079
7	0.85569	33	0.09584
8	0.81805	34	0.08489
9	0.77877	35	0.07498

**FIGURA 5.7.** Tabla de probabilidades para graficar la Curva Característica de Operación

$D$	$\beta=p(x \leq 1)$	25	0.22928
10	0.73847	36	0.06603
11	0.69767	37	0.05797
12	0.65683	38	0.05074
13	0.61632	39	0.04428
14	0.57647	40	0.03852
15	0.53755	41	0.03339
16	0.49977	42	0.02885
17	0.46331	43	0.02484
18	0.42831	44	0.02131
19	0.39487	45	0.01822
20	0.36305	46	0.01551
21	0.33291	47	0.01316
22	0.30446	48	0.01112
23	0.27772	49	0.00935
24	0.25267	50	0.00783

$\beta$  representa la probabilidad de aceptar un lote con  $D$  defectuosos.

FIGURA 5.8. Curva Característica de Operación



A la curva de la figura anterior se le denomina Curva Característica de Operación, CCO, y su forma establece qué tan estricto es el plan de muestreo aplicado para aceptar el lote. Si la curva es muy estrecha o “picuda” significa que el plan de muestreo es muy estricto. Si la curva es muy amplia implica que el plan de muestreo es muy laxo.

### 5.3 Modelo Probabilístico Binomial $b(x;n,p)$

Con respecto al modelo hipergeométrico, para el cual, en un lote de tamaño  $N$ , se tienen  $D$  artículos con alguna característica de interés; se obtiene una muestra de tamaño  $n$  y se contabiliza  $x$  el número de artículos en la muestra con la característica de interés, se calculará el límite cuando  $N$  tiende a infinito de tal manera que el cociente  $D/N$  tienda a tomar un valor constante  $p$ ; esto sucede si la  $D$  y la  $N$  tienden a crecer en la misma proporción.

$$\lim_{\substack{D \rightarrow p \\ N \rightarrow \infty}} h(x, n, D, N) = \lim_{\substack{D \rightarrow p \\ N \rightarrow \infty}} \frac{\binom{D}{x} \binom{N-D}{n-x}}{\binom{N}{n}}$$

$$\lim_{\substack{D \rightarrow p \\ N \rightarrow \infty}} \frac{\frac{D(D-1)(D-2)\dots(D-x+1)}{x(x-1)(x-2)\dots(1)} \frac{(N-D)(N-D-1)(N-D-2)\dots(N-D-n+x+1)}{(n-x)(n-x-1)(n-x-2)\dots(1)}}{\frac{N(N-1)(N-2)\dots(N-n+1)}{n(n-1)(n-2)\dots(1)}}$$

$$\lim_{\substack{D \rightarrow p \\ N \rightarrow \infty}} \frac{\frac{D \left(\frac{D}{N} - \frac{1}{N}\right) \left(\frac{D}{N} - \frac{2}{N}\right) \dots \left(\frac{D}{N} - \frac{x-1}{N}\right)}{x(x-1)(x-2)\dots(1)} \frac{\left(1 - \frac{D}{N}\right) \left(1 - \frac{D}{N} - \frac{1}{N}\right) \left(1 - \frac{D}{N} - \frac{2}{N}\right) \dots \left(1 - \frac{D}{N} - \frac{n-x-1}{N}\right)}{(n-x)(n-x-1)(n-x-2)\dots(1)}}{1 \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \dots \left(1 - \frac{n-1}{N}\right)}$$

$$\lim_{\substack{D \rightarrow p \\ N \rightarrow \infty}} \frac{n(n-1)(n-2)\dots(1)}{x(x-1)(x-2)\dots(1)(n-x-1)(n-x-2)\dots(1)} \frac{\frac{D \left(\frac{D}{N} - \frac{1}{N}\right) \left(\frac{D}{N} - \frac{2}{N}\right) \dots \left(\frac{D}{N} - \frac{x-1}{N}\right) \left(1 - \frac{D}{N}\right) \left(1 - \frac{D}{N} - \frac{1}{N}\right) \left(1 - \frac{D}{N} - \frac{2}{N}\right) \dots \left(1 - \frac{D}{N} - \frac{n-x-1}{N}\right)}{1 \left(1 - \frac{1}{N}\right) \left(1 - \frac{2}{N}\right) \dots \left(1 - \frac{n-1}{N}\right)}}$$

$$\binom{n}{x} p^x (1-p)^{n-x}$$

Nótese que al calcular este límite doble, aparece una nueva función de probabilidad, la cual se conoce como Función de Probabilidad Binomial, a la que se le denotará como  $b(x;n,p)$ . Esta función se deduce de la función de probabilidad hipergeométrica y aparece cuando el tamaño de la población  $N$  es infinito y la probabilidad de éxito de ensayo a ensayo,  $D/N$ , se mantiene igual a una constante  $p$ .

La función de probabilidad binomial está dada por la siguiente expresión:

$$b(x;n,p) = \binom{n}{x} p^x (1-p)^{n-x} \quad x=0,1,2,\dots,n \quad (5.25)$$

Su función de probabilidad acumulada está dada por la expresión:

$$B(x;n,p) = p(t \leq x) = \sum_{t=0}^{t=x} \binom{n}{t} p^t (1-p)^{n-t} \quad x=0,1,2,\dots,n \quad (5.26)$$

Esta función de probabilidad recibe el nombre de binomial debido a que las combinaciones de  $x$  en  $N$  conforman los coeficientes del binomio de Newton:

$$(a+b)^n = \sum_{x=0}^{x=n} \binom{n}{x} a^x b^{n-x} \quad (5.27)$$

Obsérvese que  $b(x; n, p)$  es no negativa para cualquier valor de  $x$ ; además, de la expresión 5.27 anterior, si  $a = p$  y  $b = 1 - p$  se demuestra que la sumatoria anterior da uno, lo cual es una condición para considerarse una función de probabilidad.

Su media:

$$\mu_x = E\{x\} = \sum_{x=0}^{x=n} x \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=0}^{x=n} x \frac{n!}{x! (n-x)!} p^x (1-p)^{n-x}$$

$$\mu_x = np \sum_{x=1}^{x=n} \frac{(n-1)!}{(x-1)! (n-x)!} p^{x-1} (1-p)^{n-x}$$

$$\mu_x = np \sum_{x=1}^{x=n} \binom{n-1}{x-1} p^{x-1} (1-p)^{n-x} = np \sum_{y=0}^{y=n-1} \binom{n-1}{y} p^y (1-p)^{n-y}$$

$$\mu_x = np (p+1-p)^{n-1}$$

$$\mu_x = np \quad (5.28)$$



Su varianza:

$$\begin{aligned}
 E\{x(x-1)\} &= \sum_{x=0}^{x=n} x(x-1) \binom{n}{x} p^x (1-p)^{n-x} = \sum_{x=0}^{x=n} x(x-1) \frac{n!}{x! (n-x)!} p^x (1-p)^{n-x} \\
 E\{x(x-1)\} &= n(n-1)p^2 \sum_{x=2}^{x=n} \frac{(n-2)!}{(x-2)! (n-x)!} p^{x-2} (1-p)^{n-x} = n(n-1)p^2 \\
 \sigma_x^2 &= E\{x^2\} - E^2\{x\} = E\{x(x-1)\} + E\{x\} - E^2\{x\} = \\
 \sigma_x^2 &= n(n-1)p^2 + np - (np)^2 \\
 \sigma_x^2 &= np(1-p) \tag{5.29}
 \end{aligned}$$

Su función generadora de momentos:

$$\begin{aligned}
 FGM_x(t) &= E\{e^{xt}\} = \sum_{x=0}^{x=n} \binom{n}{x} e^{tx} p^x (1-p)^{n-x} \\
 FGM_x(t) &= \sum_{x=0}^{x=n} \binom{n}{x} (pe^t)^x (1-p)^{n-x} \\
 FMG_x(t) &= (pe^t + 1-p)^n \tag{5.30}
 \end{aligned}$$

Recuerde que con la función generadora de momentos se pueden calcular los momentos de orden  $k$  con respecto al origen, a través de las derivadas de la función generadora de momentos, valuadas en  $t=0$ , como se ilustra a continuación:

$$\begin{aligned}
 \mu'_1 &= \left[ \frac{dFGM_x(t)}{dt} \right]_{t=0} = np \\
 \mu'_2 &= \left[ \frac{d^2FGM_x(t)}{dt^2} \right]_{t=0} = np + n(n-1)p^2 \\
 \mu'_3 &= \left[ \frac{d^3FGM_x(t)}{dt^3} \right]_{t=0} = np + 3n(n-1)p^2 + n(n-1)(n-2)p^3 \\
 \mu'_4 &= \left[ \frac{d^4FGM_x(t)}{dt^4} \right]_{t=0} = np + 7n(n-1)p^2 + 6n(n-1)(n-2)p^3 + n(n-1)(n-2)(n-3)p^4
 \end{aligned}$$

Con los momentos de orden  $k$  con respecto al origen se pueden obtener los momentos de orden  $k$  con respecto a la media, como se ilustra a continuación:

$$\mu_x = \mu'_1 = np$$

$$\sigma_x^2 = \mu_2 = E\{(x - \mu_x)^2\} = \mu'_2 - \mu_x^2 = np(1-p)$$

$$\mu_3 = E\{(x - \mu_x)^3\} = \mu'_3 - 3\mu_x \mu'_2 - \mu_x^3 = np - 3np^2 + 2np^3$$

$$\mu_3 = np(1-p)(1-2p)$$

$$\mu_4 = E\{(x - \mu_x)^4\} = \mu'_4 - 4\mu_x \mu'_3 + 6\mu_x^2 \mu'_2 - 4\mu_x^3 \mu'_1 + \mu_x^4$$

$$\mu_4 = np + n(3n-7)p^2 - 6n(n-2)p^3 + 3n(n-2)p^4$$

Por lo que, el coeficiente de asimetría está dado por:

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{np(1-p)(1-2p)}{np(1-p) \sqrt{np(1-p)}} = \frac{1-2p}{\sqrt{np(1-p)}} \quad (5.31)$$

Su coeficiente de curtosis:

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{np + n(3n-7)p^2 - 6n(n-2)p^3 + 3n(n-2)p^4}{[np(1-p)]^2} - 3 = \frac{1-6p(1-p)}{np(1-p)} \quad (5.32)$$

Cabe señalar que, para hacer el cálculo de probabilidades con la distribución binomial, utilizando Excel, se utiliza el siguiente comando:

$$p(x) = B(x; n, p, \text{acumulado}) = \text{DISTR.BINOM.N}(x, n, p, 0 \text{ ó } 1) \quad (5.33)$$

En donde  $x$  es el número de éxitos que se pretende obtener en la muestra,  $n$  es el tamaño de muestra o el número de ensayos en la muestra,  $p$  la probabilidad de obtener un éxito en un ensayo, acumulado=verdadero o 1 si se pretende calcular la función de probabilidad acumulada desde cero hasta  $x$ ; acumulado=falso o 0 si se pretende calcular la probabilidad de que el número de éxitos en la muestra sea exactamente  $x$ .

De tal manera que:

$$p(x=5; n=20, p=0.05)=\text{DISTR.BINOM.N}(5,20,0.05,0)= 0.002245$$

$$p(x\leq 5; n=20, p=0.05)= \text{DISTR.BINOM.N}(5,20,0.05,1)= 0.999671$$

$$p(x>5; n=20, p=0.05)= 1- \text{DISTR.BINOM.N}(5,20,0.05,1)= 0.000329$$

$$p(3<x\leq 5; n=20, p=0.05)=\text{DISTR.BINOM.N}(5,20,0.05,1)-$$

$$\text{DISTR.BINOM.N}(3,20,0.05,1)= 0.015572$$

$$p(3\leq x\leq 5; n=20, p=0.05)= \text{DISTR.BINOM.N}(5,20,0.05,1)-$$

$$\text{DISTR.BINOM.N}(2,20,0.05,1)= 0.075154$$

Para hacer el cálculo de probabilidades con la distribución binomial, utilizando Minitab, se realiza cada uno de los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en Binomial, como se muestra en la figura 5.2.
2. Aparece la pantalla que se muestra en la figura 5.9, en la cual, para calcular la probabilidad  $p(x=5, n=20, p=0.05)$  se señala si se trata de una probabilidad en un punto marcando Probability, en Number of trials se indica el tamaño de la población o lote 20, en Event probability se indica la probabilidad de éxito en un ensayo 0.05, en Input constant se indica el valor de  $x$  para el cual se desea calcular la probabilidad y se da un click en Ok obteniéndose el resultado.

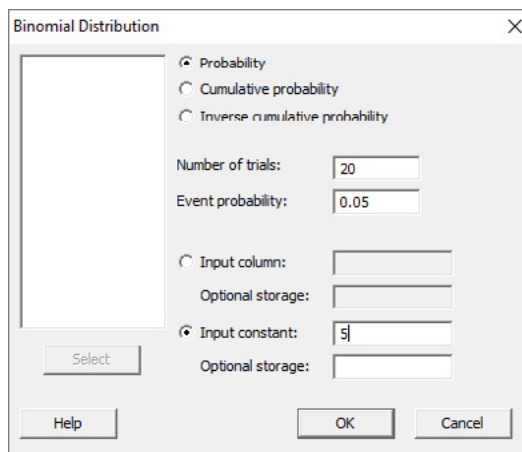


FIGURA 5.9

Si lo que se pretende calcular es la probabilidad acumulada hasta el valor  $x=5$ , en vez de elegir Probability debe elegirse Cumulative probability y todo lo demás igual.

En el lenguaje de desarrollo R, para el caso de la distribución binomial, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo de la probabilidad exactamente en un punto:

$$\text{dbinom}(x, \text{size}, \text{prob}, \text{log} = \text{FALSE}) \quad (5.34)$$

Para el cálculo de la probabilidad acumulada hasta un punto dado:

$$\text{pbinom}(x, \text{size}, \text{prob}, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) \quad (5.35)$$

En donde  $x$  es el punto donde se quiere calcular la probabilidad; size es el tamaño de la muestra o el número de ensayos; prob es la probabilidad de tener éxito en cada ensayo; log es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma FALSE. lower.tail es un valor lógico, si se le asigna un valor verdadero el valor que arroja es la probabilidad acumulada a la izquierda de  $x$ , de lo contrario, si el valor que toma es FALSE, el valor que arroja es la probabilidad a la derecha del valor de  $x$ .

Por ejemplo,

$$p(x=5; n=20, p=0.05) = \text{dbinom}(5, 20, 0.05, \text{log}=\text{FALSE}) = 0.002245$$

$$p(x \leq 5; n=20, p=0.05) = \text{pbinom}(5, 20, 0.05, \text{lower.tail}=\text{TRUE}, \text{log}=\text{FALSE}) \\ = 0.999671$$

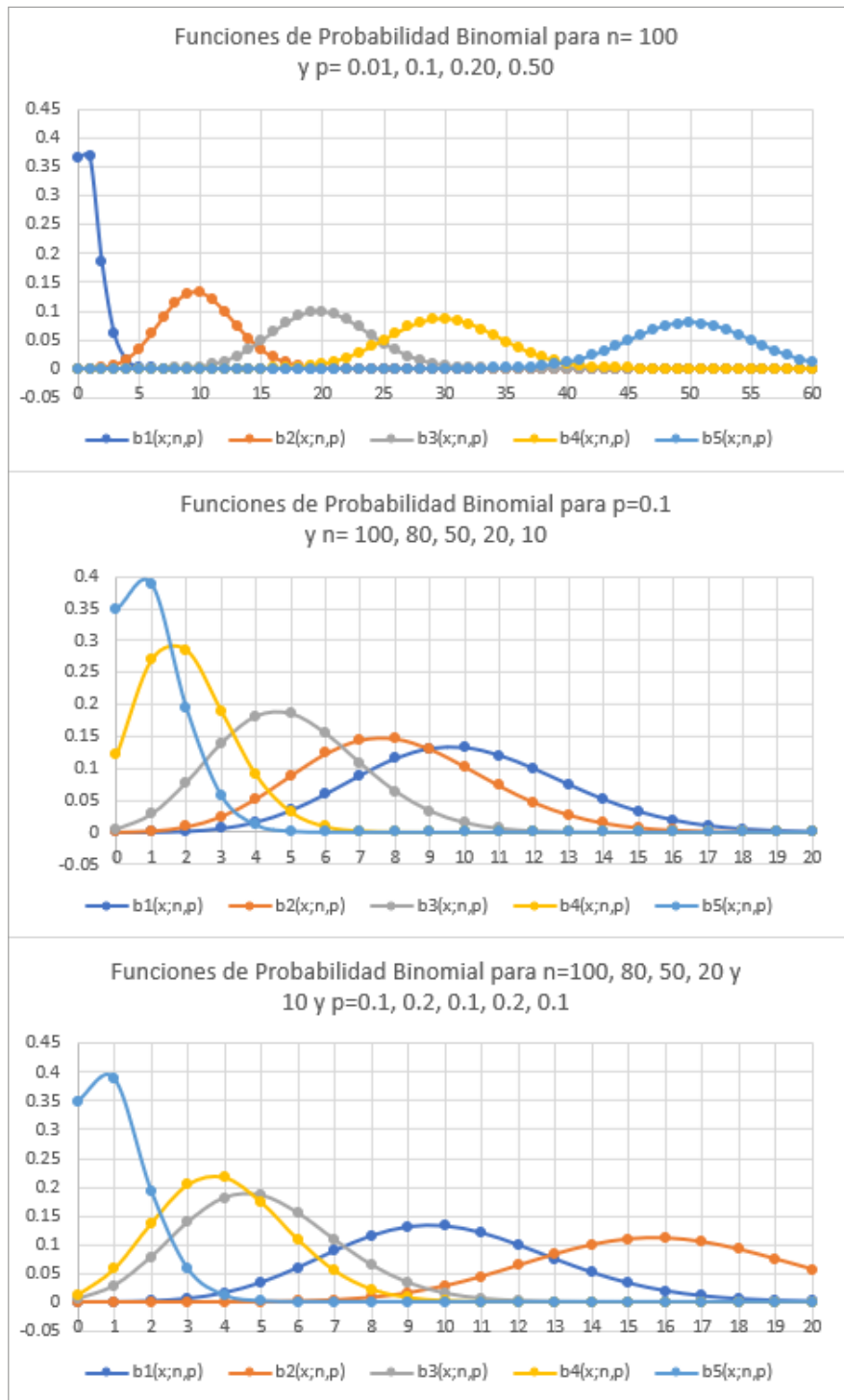
$$p(x > 5; n=20, p=0.05) = \text{pbinom}(5, 20, 0.05, \text{lower.tail}=\text{FALSE}, \text{log}=\text{FALSE}) \\ = 0.000329$$

$$p(3 < x \leq 5; n=20, p=0.05) = \text{pbinom}(5, 20, 0.05, \text{lower.tail}=\text{TRUE}, \text{log}=\text{FALSE}) - \\ \text{pbinom}(3, 20, 0.05, \text{lower.tail}=\text{TRUE}, \text{log}=\text{FALSE}) = 0.015572$$

$$p(3 \leq x \leq 5; n=20, p=0.05) = \text{pbinom}(5, 20, 0.05, \text{lower.tail}=\text{TRUE}, \\ \text{log}=\text{FALSE}) - \text{pbinom}(2, 20, 0.05, \text{lower.tail}=\text{TRUE}, \text{log}=\text{FALSE}) = 0.075154$$

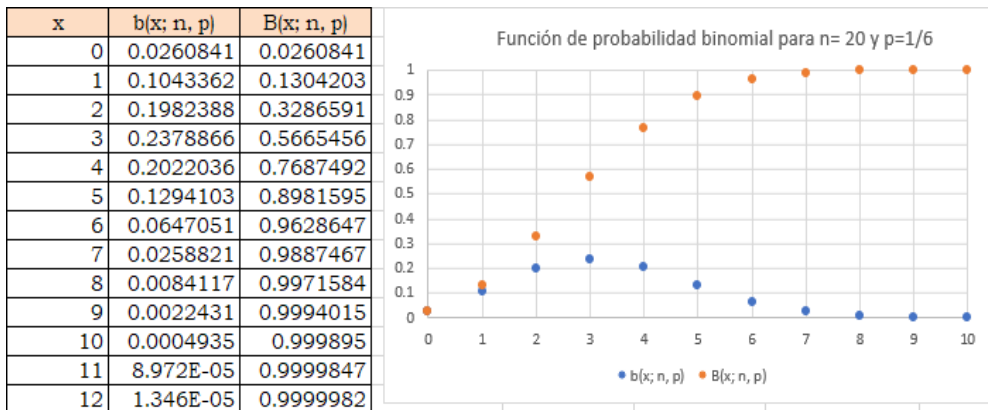
En la figura 5.10 se muestran algunas gráficas de la distribución binomial para diferentes valores de  $n$  y  $p$ .

**FIGURA 5.10.**  
Gráficas de la  
distribución  
binomial para  
diferentes valores  
de  $n$  y  $p$



## Ejercicio 5.5

El ejemplo más común de una distribución de probabilidad binomial, en los juegos de azar, es el lanzamiento simultáneo y repetido de dos dados homogéneos de seis caras planas e iguales. Suponga que se lanzan dichos dados  $n=20$  veces. Se suman los números de las caras superiores de ambos dados y se le llama éxito al evento que caiga como resultado de la suma el número 7. Nótese que al lanzar dos dados homogéneos simultáneamente como un evento de Bernoulli, la probabilidad de que caiga como resultado el número 7 es  $p=6/36=1/6=0.166667$ , correspondiente a los eventos favorables (1,6), (6,1), (2,5), (5,2), (3,4), (4,3), donde la abscisa representa la cara que cae hacia arriba de uno de los dados y la ordenada representa la cara que cae hacia arriba del otro dado. Nótese que se trata de  $n=20$  ensayos de Bernoulli, donde la probabilidad de éxito  $p=0.166667$  se mantiene constante de evento a evento y el número  $N$  de veces que pueden ser tirados simultáneamente los dos dados es infinito, por lo que corresponde exactamente a un modelo binomial. Se muestra a continuación una tabla de resultados de los posibles eventos (sólo se consideró hasta  $x=12$ , pero se pudo haber calculado hasta  $x=20$ ), los parámetros estadísticos que la caracterizan y las gráficas de la función de probabilidad binomial y de su acumulada.



$$\mu_x = np = 20 \left( \frac{1}{6} \right) = 3.333333$$

$$\sigma_x^2 = np(1-p) = 20 \left( \frac{1}{6} \right) \left( \frac{5}{6} \right) = 2.777778$$

$$\gamma_1 = \frac{1-2p}{\sqrt{np(1-p)}} = \frac{1-2/6}{\sqrt{20(1/6)(5/6)}} = 0.4$$

$$\gamma_2 = \frac{1-6p(1-p)}{np(1-p)} = \frac{1-6(1/6)(5/6)}{20(1/6)(5/6)} = 0.06$$

Nótese que una variable aleatoria binomial  $x$  se puede expresar como una suma de  $N$  variables aleatorias tipo Bernoulli:

$$x = x_1 + x_2 + \dots + x_n$$

$$x_i = 0, 1 \quad \forall i = 1, 2, \dots, n$$

Lo que implica que  $x$  puede tomar los valores enteros desde de cero hasta  $n$ .

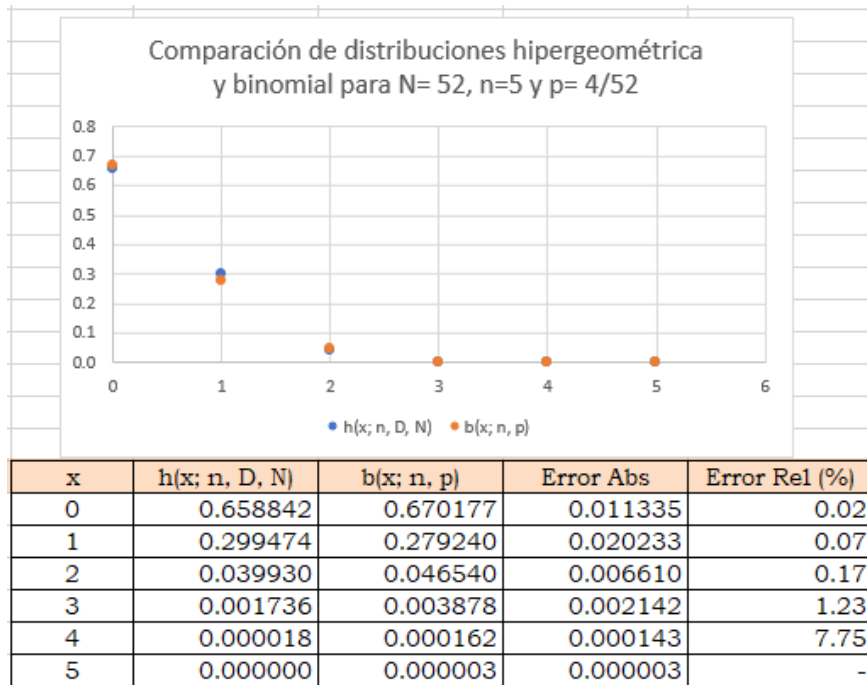
### 5.3.1. Aproximación de la distribución hipergeométrica a través de la binomial

Si  $x$  es una variable aleatoria con distribución hipergeométrica  $h(x; N, n, D)$ , ya se demostró previamente que cuando  $N \rightarrow \infty$  y  $D/N \rightarrow p$  se obtiene una distribución binomial  $b(x; n, p)$ , lo que implica que la distribución binomial puede utilizarse para aproximar probabilidades de una variable aleatoria hipergeométrica.

En general, la aproximación se considera satisfactoria si  $N > 50$  y  $n/N \leq 0.10$ ; mientras más pequeño sea el cociente  $n/N$  la aproximación es mejor.

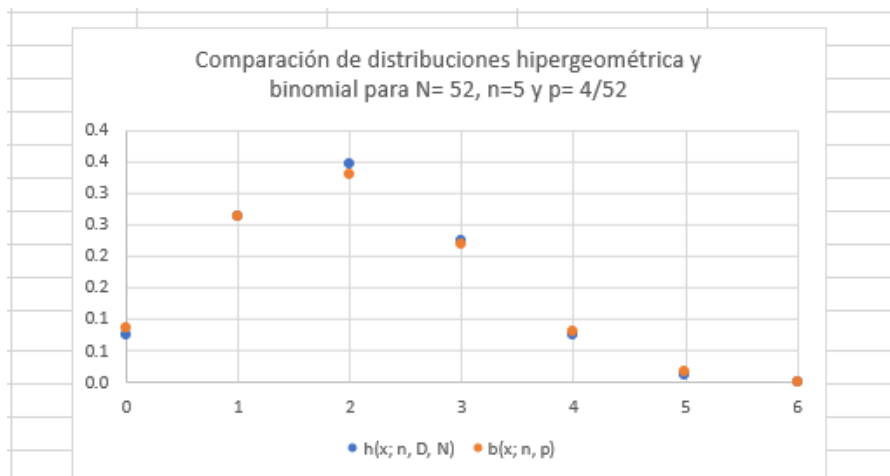
¿Por qué la necesidad de usar la distribución binomial como aproximación de la hipergeométrica?

La razón es histórica, antes del advenimiento de las computadoras, calcular probabilidades con la distribución hipergeométrica resultaba un proceso engorroso por el cálculo de los factoriales y de las combinaciones; era más rápido calcular probabilidades con la distribución binomial, de allí que se buscaban aproximaciones que facilitaran los cálculos. Para el ejemplo 5.2, en donde se hicieron los cálculos con la distribución hipergeométrica, se trazarán ambas gráficas, usando la binomial también y se compararán para ver qué tanto se aproxima la binomial a la hipergeométrica. En este ejemplo se considerará como  $p = D/N = 4/52 = 0.076923$ ; nótese que  $n/N = 4/52 < 0.1$  por lo que el ajuste debe ser adecuado. Las gráficas se muestran en la siguiente figura:



Cabe mencionar que en este ejemplo, la hipergeométrica solo está definida hasta  $x=$  mínimo  $(D, n)= 4$ , en cambio, la binomial está definida hasta  $x= n=5$ , lo cual no corresponde con el modelo real, porque solo se tienen cuatro posibles éxitos. Como se puede apreciar en la figura anterior, los resultados se ajustan adecuadamente, el máximo error absoluto que se comete es de 2 centésimas.

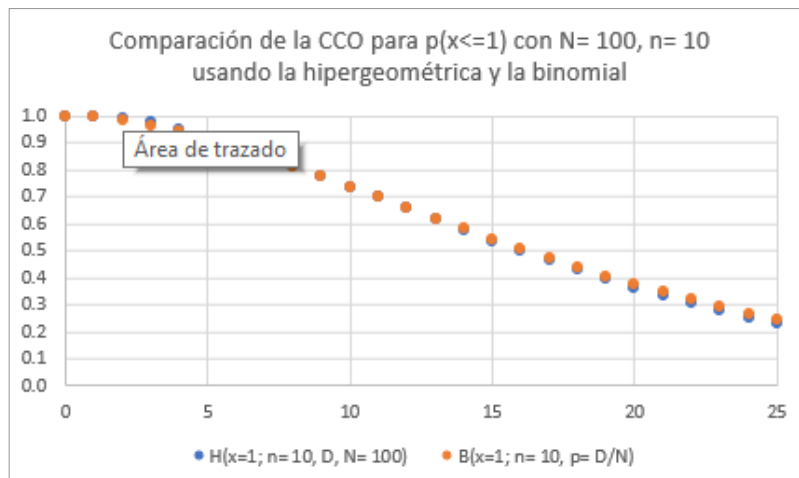
Para el ejercicio 5.3, en donde  $N= 60$ ,  $D= 20$  y  $n= 6$ , la comparación de ambos modelos arroja los siguientes resultados:





x	$h(x; n, D, N)$	$b(x; n, p)$	Error Abs	Error Rel (%)
0	0.076670	0.087791	0.011122	0.15
1	0.262867	0.263374	0.000507	0.00
2	0.346839	0.329218	0.017621	0.05
3	0.224977	0.219479	0.005498	0.02
4	0.075486	0.082305	0.006819	0.09
5	0.012387	0.016461	0.004074	0.33
6	0.000774	0.001372	0.000598	0.77

Para el ejercicio 5.4, en donde  $N = 100$ ,  $D$  se desconoce y  $n = 10$ , antes de querer comparar la hipergeométrica con la binomial, es necesario establecer el mismo eje horizontal, o se grafica usando la  $x$  como número de defectuosos o se grafica usando la  $p$  como porcentaje defectuoso. Recuerde que  $D = pN$  o también que  $p = D/N$ . La comparación de las curvas características de operación con la hipergeométrica y con la binomial arroja los siguientes resultados:



#### 5.4. Modelo Probabilístico de Pascal o Binomial Negativo $bn(x; p, r)$

Suponga que se tiene un ensayo de Bernoulli, para el cual sólo existen dos posibles resultados: éxito o fracaso. La probabilidad de éxito de ensayo a ensayo es  $p$  y la de fracaso  $q = 1 - p$ . Se realiza un experimento conformado por un cierto número de ensayos de Bernoulli, estadísticamente independientes entre sí y que deberán repetirse hasta obtener “ $r$ ” éxitos. En este caso, la variable aleatoria es  $x$ , donde  $x$  representará el número de ensayos que deberán realizarse hasta obtener  $r$  éxitos. Obviamente, que para obtener  $r$  éxitos se deben realizar cuando menos  $r$  ensayos, por lo que  $x = r, r + 1, r + 2, \dots$

Suponga que primero se obtienen  $(x - r)$  fracasos, la probabilidad de que esto suceda estaría dada por  $(1 - p)(1 - p) \dots (1 - p) = (1 - p)^{x - r}$ . Posteriormente, se obtienen  $r$  éxitos, la probabilidad de que esto suceda sería  $pp \dots p = p^r$ . La probabilidad de obtener primero  $(x - r)$  fracasos y luego  $r$  éxitos sería el producto de las probabilidades anteriores, es decir,  $(1 - p)^{x - r} p^r$ .

Ahora bien, los primeros  $r - 1$  éxitos deben aparecer en cualquier orden dentro de los primeros  $x - 1$  ensayos, porque el último ensayo  $x$  debe ser éxito. Esto implica que existen  $C(r - 1, x - 1)$  combinaciones o maneras diferentes de obtener  $r$  éxitos en  $x$  ensayos. De esta forma, la probabilidad de obtener  $r$  éxitos en  $x$  ensayos estaría dada por la siguiente expresión:

$$bn(x; p, r) = \binom{x - 1}{r - 1} (1 - p)^{x - r} p^r \quad x = r, r + 1, r + 2, \dots \quad (5.36)$$

Cabe señalar que el modelo deducido anteriormente, es aplicable cuando el número de éxitos a obtener,  $r$ , en  $x$  intentos es un número natural, pero puede ser extendido al caso en que  $r$  es un número real positivo, para el cual, su función de probabilidad sería:

$$bn(x; p, r) = \frac{\Gamma(r + x)}{x! \Gamma(r)} p^r (1 - p)^x \quad x = 0, 1, 2, \dots \quad (5.37)$$

En dónde, la función

$$\Gamma(r) = \int_0^{\infty} t^{r-1} e^{-t} dt \quad (5.38)$$

Es la función Gamma o Función Factorial Generalizada, la cual recibe este nombre porque si  $r$  es natural  $\Gamma(r)=(r-1)!$

Al modelo con  $r$  número natural se le conoce como Función de Probabilidad de Pascal y al modelo con  $r$  número real se le conoce como Binomial Negativo.

La función de probabilidad acumulada de la expresión 5.36 sería:

$$Bn(x;p,r) = \sum_{t=r}^{t=x} \binom{t-1}{r-1} (1-p)^{t-r} p^r \quad x = r, r+1, r+2, \dots \quad (5.39)$$

La cual también puede ser escrita como:

$$Bn(x;p,r) = \sum_{t=0}^{t=x} \binom{t+r-1}{r-1} (1-p)^t p^r \quad x = r, r+1, r+2, \dots \quad (5.40)$$

Para calcular la función generatriz de momentos considere el binomio de Newton con exponente negativo:

$$\frac{1}{(a-b)^r} = \sum_{k=0}^{k \rightarrow \infty} \binom{r+k}{k} \frac{b^k}{a^{r+k}} \quad (5.41)$$

$$FGM_x(t) = E\{e^{xt}\} = \sum_{x=0}^{x \rightarrow \infty} \binom{x+r-1}{r-1} e^{xt} (1-p)^x$$

$$FGM_x(t) = \sum_{x=0}^{x \rightarrow \infty} \frac{(x+r-1)!}{(r-1)! x!} p^r [(1-p)e^t]^x$$

$$FGM_x(t) = p^r + rp^r [(1-p)e^t] + \frac{r(r+1)}{2!} [(1-p)e^t]^2 + \dots$$

$$FGM_x(t) = \left[ \frac{1}{p} - \frac{(1-p)e^t}{p} \right]^{-r}$$

$$FGM_x(t) = \frac{p^r}{[1 - (1-p)e^t]^r} \quad (5.42)$$

La expresión anterior es válida para cuando lo que se desea contabilizar es el número de fracasos  $x-r$  entre  $r$  éxitos, recuerde que la probabilidad de obtener fracasos es  $q=1-p$ . La función generatriz de momentos también puede ser escrita de la siguiente forma (partiendo de la expresión 5.36):

$$FGM_x(t) = \frac{p^r}{[e^{-t} - (1-p)]^r} \quad (5.43)$$

La función característica sería:

$$FC_x(t) = \frac{p^r}{[e^{-it} - (1-p)]^r} \quad (5.44)$$

Ahora, calculando los momentos con respecto al origen de primero, segundo, tercero y cuarto orden:

$$\mu'_1 = \left[ \frac{dFGM_x(t)}{dt} \right]_{t=0} = \frac{r}{p}$$

$$\mu'_2 = \left[ \frac{d^2FGM_x(t)}{dt^2} \right]_{t=0} = \frac{r(r+1)}{p^2} - \frac{r}{p}$$

$$\mu'_3 = \left[ \frac{d^3FGM_x(t)}{dt^3} \right]_{t=0} = \frac{r(r+1)(r+2)}{p^3} - \frac{3r(r+1)}{p^2} + \frac{r}{p}$$

$$\mu'_4 = \left[ \frac{d^4FGM_x(t)}{dt^4} \right]_{t=0} = \frac{r(r+1)(r+2)(r+3)}{p^4} - \frac{6r(r+1)(r+2)}{p^3} + \frac{7r(r+1)}{p^2} - \frac{r}{p}$$

Se calculan a continuación los momentos con respecto a la media, de primero, segundo, tercero y cuarto orden:

$$\mu_x = \mu'_1 = \frac{r}{p}$$

$$\sigma_x^2 = E\{x^2\} - \mu_x^2 = \frac{r(r+1)}{p^2} - \frac{r}{p} - \frac{r^2}{p^2} = \frac{r(1-p)}{p^2}$$

$$\mu_3 = \frac{r(1-p)(2-p)}{p^3}$$

$$\mu_4 = \frac{3r(r+2)(1-2p)}{p^4} + \frac{r(3r+7-p)}{p^2}$$

Con los momentos anteriores se obtienen los parámetros estadísticos de la distribución binomial negativa:

$$\text{Media} \quad \mu_x = \frac{r}{p} \quad (5.45)$$

$$\text{Varianza} \quad \sigma_x^2 = \frac{r(1-p)}{p^2} \quad (5.46)$$

$$\text{Coeficiente de Asimetría} \quad \gamma_1 = \frac{2-p}{\sqrt{r(1-p)}} \quad (5.47)$$

$$\text{Coeficiente de Curtosis} \quad \gamma_2 = \frac{6 - 6p + p^2}{r(1-p)} \quad (5.48)$$

Cabe señalar que si lo que se contabiliza son el número de fracasos antes de obtener  $r$  éxitos la media es igual a  $r(1-p)/p$ . Todos los demás parámetros estadísticos son iguales.

Para hacer el cálculo de probabilidades con la distribución binomial negativa o de Pascal, utilizando Excel, se utiliza el siguiente comando:

$$p(x)=\text{Bn}(x;r,p,\text{acumulado})=\text{NEGBINOM.DIST}(x-r,r,p,0 \text{ o } 1) \quad (5.49)$$

En donde  $x$  es el número de ensayos a realizar antes de obtener  $r$  éxitos,  $x-r$  el número de fracasos antes de obtener  $r$  éxitos,  $p$  es la probabilidad de obtener un éxito en un ensayo,  $\text{acumulado}=\text{verdadero}$  o  $1$  si se pretende calcular la función de probabilidad acumulada desde cero hasta  $x$ ;  $\text{acumulado}=\text{falso}$  o  $0$  si se pretende calcular la probabilidad de que el número de ensayos para obtener  $r$  éxitos sea exactamente  $x$ .

De tal manera que:

$$p(x=5;r=3, p=0.05)= \text{NEGBINOM.DIST}(2,3,0.05,0)= 0.000677$$

$$p(x\leq 5; r=3, p=0.05)= \text{NEGBINOM.DIST}(2,3,0.05,1)= 0.001158$$

$$p(x>5; r=3, p=0.05)= 1 - \text{NEGBINOM.DIST}(2,3,0.05,1)= 0.998842$$

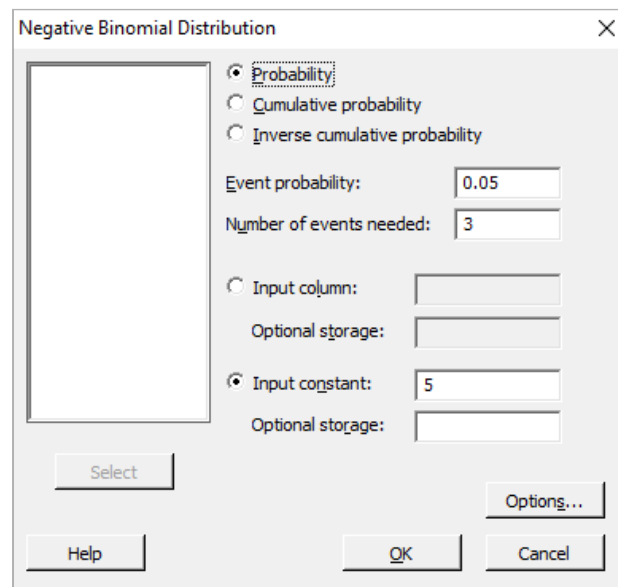
$$p(4 < x \leq 6; r=3, p=0.05) = \text{NEGBINOM.DIST}(3, 3, 0.05, 1) - \text{NEGBINOM.DIST}(1, 3, 0.05, 1) = 0.001749$$

$$p(4 \leq x \leq 6; r=3, p=0.05) = \text{NEGBINOM.DIST}(3, 3, 0.05, 1) - \text{NEGBINOM.DIST}(0, 3, 0.05, 1) = 0.002105$$

Para hacer el cálculo de probabilidades con la distribución binomial negativa o de Pascal, utilizando Minitab, se realizan los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en Binomial negative, como se muestra en la figura 5.2.
2. Aparece la pantalla que se muestra en la figura 5.11, en la cual, para calcular la probabilidad  $p(x=5, r=3, p=0.05)$  se señala si se trata de una probabilidad en un punto marcando Probability, en Event probability se indica la probabilidad de éxito en un ensayo 0.05, en Number of events needed se indica el valor de  $r=3$  y en Input constant el valor de  $x=5$  para el cual se desea calcular la probabilidad y se da un click en Ok obteniéndose el resultado.

FIGURA 5.11



Si lo que se pretende calcular es la probabilidad acumulada hasta el valor  $x=5$ , en vez de elegir Probability debe elegirse Cumulative probability y todo lo demás igual.

En el lenguaje de desarrollo R, para el caso de la distribución binomial negativa o de Pascal, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo de la probabilidad exactamente en un punto:

$$\text{dnbinom}(x, \text{size}, \text{prob}, \text{mu}, \text{log} = \text{FALSE}) \quad (5.50)$$

Para el cálculo de la probabilidad acumulada hasta un punto dado:

$$\text{pnbinom}(x, \text{size}, \text{prob}, \text{mu}, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) \quad (5.51)$$

En donde  $x$  es el número de ensayos fracasados hasta obtener  $\text{size}$  éxitos;  $\text{size}$  es el número de ensayos exitosos a lograr en  $x+r$  ensayos;  $\text{prob}$  es la probabilidad de tener éxito en un ensayo;  $\text{mu}$  es una parametrización alternativa (usada a menudo en Ecología) donde  $\text{prob} = \text{size}/(\text{size}+\text{mu})$ . La varianza es  $\text{mu} + \text{mu}^2/\text{size}$  en esta parametrización;  $\text{log}$  es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma FALSE.  $\text{lower.tail}$  es un valor lógico; si se le asigna un valor verdadero, el valor que arroja es la probabilidad acumulada a la izquierda de  $x$ , de lo contrario, si el valor que toma es Falso, el valor que arroja es la probabilidad a la derecha del valor de  $x$ .

Por ejemplo,

$$p(x=5; r=3, p=0.05) = \text{dnbinom}(2,3,0.05, \text{log}=\text{FALSE}) = 0.000677$$

$$p(x \leq 5; r=3, p=0.05) = \text{pnbinom}(2,3,0.05, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.001158$$

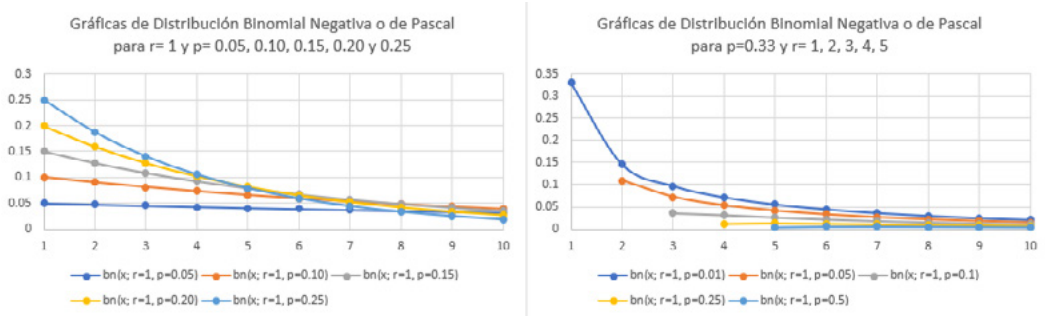
$$p(x > 5; r=3, p=0.05) = \text{pnbinom}(2,3,0.05, \text{lower.tail}=\text{FALSE}, \text{log.p}=\text{FALSE}) = 0.998842$$

$$p(4 < x \leq 6; r=3, p=0.05) = \text{pnbinom}(3,3,0.05, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pnbinom}(1,3,0.05, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.001749$$

$$p(4 \leq x \leq 6; r=3, p=0.05) = \text{pnbinom}(3,3,0.05, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pnbinom}(0,3,0.05, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.002105$$

En la figura 5.12 se muestran algunas gráficas de la distribución binomial negativa o de Pascal para diferentes valores de  $p$  y  $r$ .

**FIGURA 5.12.** Gráficas de la distribución binomial negativa



### Ejercicio 5.6

Según las estadísticas globales del mejor jugador de la NBA de todos los tiempos, Michael Jordan, la probabilidad de que al lanzar un tiro desde cualquier posición, lo acertara en la canasta, era  $p=0.472103$ .

- a) ¿Cuántos tiros tendría que realizar en promedio para obtener tres canastas?, ¿cuál sería su varianza, coeficiente de asimetría y coeficiente de curtosis?

$$\mu_x = \frac{r}{p} = \frac{3}{0.472103} = 6.354545$$

$$\sigma_x^2 = \frac{r(1-p)}{p^2} = \frac{3(1-0.472103)}{0.472103^2} = 7.105537$$

$$\gamma_1 = \frac{2-p}{\sqrt{r(1-p)}} = \frac{2-0.472103}{\sqrt{3(1-0.472103)}} = 1.214112$$

$$\gamma_2 = \frac{6-6p+p^2}{r(1-p)} = \frac{6-6 \cdot 0.472103 + 0.472103^2}{3(1-0.472103)} = 2.140735$$

- b) Calcule la probabilidad de que en un juego cualquiera tenga necesidad de disparar cinco tiros para obtener tres canastas.

$$bn(x=5; p=0.472103, r=3) = \binom{4}{2} (1-0.472103)^2 (0.472103)^3 = 0.175938$$



c) Obtenga la probabilidad de anotar tres canastas cuando más en cinco tiros.

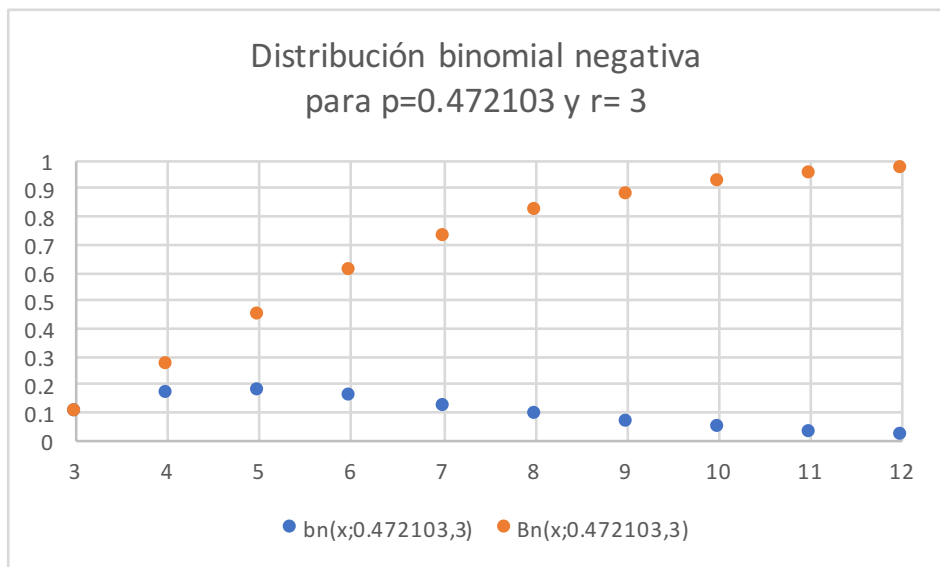
$$Bn(x=; p=0.472103, r=3) = \sum_{t=3}^{t=5} \binom{t-1}{2} (1-0.472103)^{t-3} (0.472103)^3 = 0.447802$$

d) Calcule la probabilidad de que el número de tiros para obtener tres canastas sea mayor de 10.

$$p(x>10) = 1 - Bn(x=10; p=0.472103, r=3) = 0.077201$$

e) Grafique la función de probabilidad del número de tiros que tendría que realizar para obtener tres canastas.

Nótese de la gráfica siguiente que la mediana de esta distribución se encuentra entre cinco y seis y su moda es cinco.



### 5.5. Modelo Probabilístico Geométrico $g(x; p)$

La distribución de Pascal o binomial negativa para  $r=1$  recibe el nombre de Distribución Geométrica. La  $x$  representa el número de ensayos que se tienen que realizar hasta obtener un éxito. Su función de probabilidad es:

$$g(x; p) = bn(x; p, r=1) = (1-p)^{x-1} p \quad x = 1, 2, 3, \dots \quad (5.52)$$

Su función de probabilidad acumulada es:

$$G(x; p) = Bn(x; p, r=1) = \sum_{t=1}^{t=x} (1-p)^{t-1} p \quad x = 1, 2, 3, \dots \quad (5.53)$$

Su media, varianza, coeficiente de asimetría, coeficiente de curtosis y su función generatriz de momentos están dadas por las expresiones:

$$\text{Media} \quad \mu_x = \frac{1}{p} \quad (5.54)$$

$$\text{Varianza} \quad \sigma_x^2 = \frac{1-p}{p^2} \quad (5.55)$$

$$\text{Coeficiente de Asimetría} \quad \gamma_1 = \frac{2-p}{\sqrt{1-p}} \quad (5.56)$$

$$\text{Coeficiente de Curtosis} \quad \gamma_2 = \frac{6-6p+p^2}{1-p} \quad (5.57)$$

$$\text{Función Generatriz de Momentos} \quad FGM_x(t) = \frac{p}{[e^{-t}-1+p]} \quad (5.58)$$

$$\text{Función Característica} \quad FC_x(t) = \frac{p}{[e^{-it}-1+p]} \quad (5.59)$$

Para el cálculo de probabilidades con la distribución geométrica, utilizando Excel se emplea la binomial negativa para  $r=1$ , con el siguiente comando:

$$p(x) = Bn(x; 1, p, \text{acumulado}) = \text{NEGBINOM.DIST}(x-1, 1, p, 0 \text{ ó } 1) \quad (5.60)$$

En donde  $x$  es el número de ensayos a realizar hasta obtener el primer éxito,  $x-1$  el número de fracasos antes de obtener el primer éxito,  $p$  es la probabilidad de obtener un éxito en un ensayo, acumulado = verdadero ó 1 si se pretende calcular la función de probabilidad acumulada desde cero hasta  $x$ ; acumulado = falso ó 0 si se pretende calcular la probabilidad de que el número de ensayos para obtener un éxito sea exactamente  $x$ .

De tal manera que:

$$p(x=5; r=1, p=0.05) = \text{NEGBINOM.DIST}(4, 1, 0.05, 0) = 0.040725$$

$$p(x \leq 5; r=1, p=0.05) = \text{NEGBINOM.DIST}(4, 1, 0.05, 1) = 0.226219$$

$$p(x > 5; r=1, p=0.05) = 1 - \text{NEGBINOM.DIST}(4, 1, 0.05, 1) = 0.773781$$

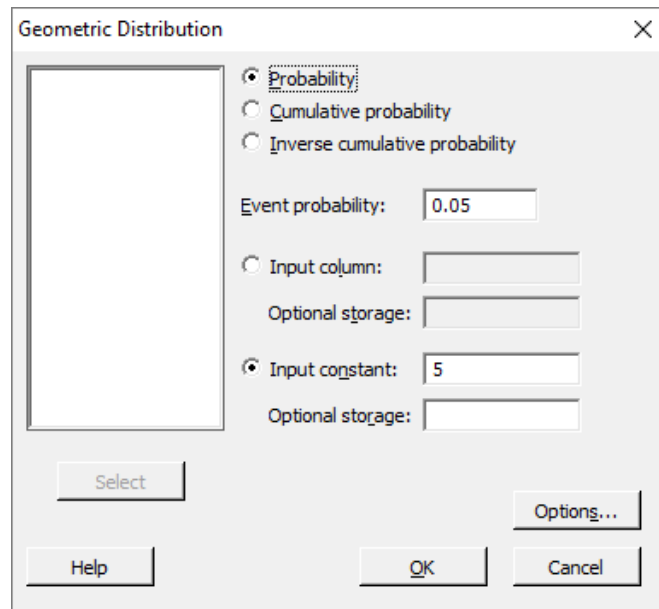
$$p(4 < x \leq 6; r=1, p=0.05) = \text{NEGBINOM.DIST}(5, 1, 0.05, 1) - \text{NEGBINOM.DIST}(3, 1, 0.05, 1) = 0.079414$$

$$p(4 \leq x \leq 6; r=1, p=0.05) = \text{NEGBINOM.DIST}(5, 1, 0.05, 1) - \text{NEGBINOM.DIST}(2, 1, 0.05, 1) = 0.122283$$

Para hacer el cálculo de probabilidades con la distribución geométrica, utilizando Minitab, se realizan los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en Geometric, como se muestra en la figura 5.13.
2. Aparece la pantalla que se muestra en la figura 5.13, en la cual, para calcular la probabilidad  $p(x=5, r=1, p=0.05)$  se señala si se trata de una probabilidad en un punto marcando Probability, en Event probability se indica la probabilidad de éxito en un ensayo 0.05, en Number of events needed se indica el valor de  $r=1$  y en Input constant el valor de  $x=5$  para el cual se desea calcular la probabilidad y se da un click en Ok obteniéndose el resultado.

FIGURA 5.13



Si lo que se pretende calcular es la probabilidad acumulada hasta el valor  $x=5$ , en vez de elegir Probability debe elegirse Cumulative probability y todo lo demás igual.

En el lenguaje de desarrollo R, para el caso de la distribución geométrica, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo de la probabilidad exactamente en un punto:

$$dgeom(x, prob, log = FALSE) \quad (5.61)$$

Para el cálculo de la probabilidad acumulada hasta un punto dado:

$$pgeom(x, prob, lower.tail = TRUE, log.p = FALSE) \quad (5.62)$$

En donde  $x$  es el número de ensayos fracasados hasta obtener un éxito;  $prob$  es la probabilidad de tener éxito en un ensayo;  $log$  es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma FALSE.  $lower.tail$  es un valor lógico; si se le asigna un valor verdadero, el valor que arroja es la probabilidad acumulada a la izquierda de  $x$ , de lo contrario, si el valor que toma es falso, el valor que arroja es la probabilidad a la derecha del valor de  $x$ .

Por ejemplo,

$$p(x=5; r=1, p=0.05) = \text{dgeom}(4, 0.05, \text{log}=\text{FALSE}) = 0.040725$$

$$p(x \leq 5; r=1, p=0.05) = \text{pgeom}(4, 0.05, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.226219$$

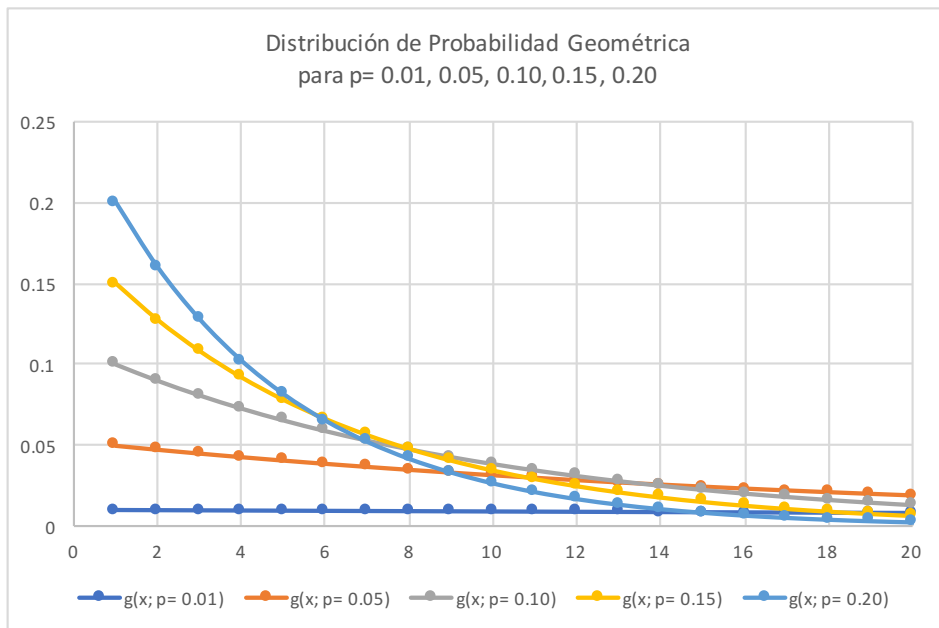
$$p(x > 5; r=1, p=0.05) = \text{pgeom}(4, 0.05, \text{lower.tail}=\text{FALSE}, \text{log.p}=\text{FALSE}) = 0.773781$$

$$p(4 < x \leq 6; r=1, p=0.05) = \text{pgeom}(5, 0.05, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pgeom}(3, 0.05, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.079414$$

$$p(4 \leq x \leq 6; r=1, p=0.05) = \text{pgeom}(5, 0.05, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pgeom}(2, 0.05, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.122283$$

En la figura 5.14 se muestran algunas gráficas de la distribución de probabilidad geométrica para diferentes valores de  $p$ .

**FIGURA 5.14.** Distribución de Probabilidad Geométrica para diferentes valores de  $p$



### 5.5.1. Propiedad de Olvido o de Falta de Memoria de la Distribución Geométrica

Existe una propiedad de al menos dos distribuciones de probabilidad: la distribución geométrica y la distribución exponencial negativa, la cual se conoce como propiedad de olvido o de falta de memoria.

Sea  $T$  una variable aleatoria. Se dice que  $T$  cumple la propiedad de olvido o de falta de memoria, si satisface la siguiente expresión:

$$p(T \geq s + t \mid T > t) = p(T \geq s) \quad (5.63)$$

Como se demostrará a continuación, la distribución geométrica cumple esta propiedad.

Sea  $x$  una variable aleatoria geométrica con parámetro  $p$ , es decir,

$$g(x; p) = (1-p)^{x-1} p$$

Entonces, desarrollando el primer miembro de la expresión 5.63, por definición de probabilidad condicional:

$$p(X \geq x + s \mid X > x) = \frac{p[(X > x + s) \cap (X > x)]}{p(X > x)} = \frac{p(X > x + s)}{p(X > x)}$$

$$p(X \geq x + s \mid X > x) = \frac{p(X \geq x + s - 1)}{p(X > x)} = \frac{(1-p)^{s+t-1}}{(1-p)^t} = (1-p)^{s-1}$$

$$p(X \geq x + s \mid X > x) = p(X > s - 1) = p(X \geq s)$$

#### Ejercicio 5.7

Para destruir una partícula microscópica cancerosa se utiliza un nanorobot armado con un rayo láser. Para destruir dicha partícula basta con acertar con un solo disparo. La probabilidad de acertar con un disparo es 0.33.

c) Obtenga la probabilidad de acertar en el primer disparo.

$$p=0.33$$

d) Existe riesgo de dañar otras partes del cuerpo si se hacen más de cinco disparos y no se acierta, por lo cual, es necesario calcular la probabilidad de destruir la partícula cuando mucho en cinco disparos.

$$p(x \leq 5) = G(5; 0.33) = \sum_{t=1}^{t=5} (1-0.33)^{t-1} (0.33) = 0.864987$$

- e) Calcule la probabilidad de que, para destruir la partícula, el número de disparos este entre tres y cinco.

$$p(3 \leq x \leq 5) = G(5; 0.33) - G(2; 0.33) = \sum_{t=3}^{t=5} (1-0.33)^{t-1} (0.33) = 0.313887$$

- f) Suponga que el nanorobot ya ha disparado tres veces y no ha acertado, ¿cuál es la probabilidad de que acierte después del quinto disparo?

Por la propiedad de olvido o pérdida de memoria:

$$p(X > 3 + 2 \mid X > 3) = p(X > 2)$$

$$p(X > 3 + 2 \mid X > 3) = 1 - (0.67)^2 (0.33) = 0.851863$$

### Ejercicio 5.8

La compañía de seguros GNP se propone realizar presentaciones sucesivas en empresas hasta lograr un contrato anual en alguna de ellas. Para preparar la logística del proceso se requiere una inversión inicial de \$20,000.00. Cada presentación de ventas cuesta alrededor de \$10,000.00.

- a) ¿Cuál es el costo esperado de lograr un contrato?, si la probabilidad de alcanzarlo en cada presentación es  $p=0.1$ .

Sea  $x$  el número de visitas requeridas para lograr un contrato anual. El costo para lograrlo sería  $C(x) = 20000 + 10000x$ .

El costo esperado de lograr un contrato anual sería la esperanza matemática de  $C(x)$ :

$$E\{C(x)\} = E\{20000 + 10000x\} = 20000 + 10000E\{x\} = 20000 + 10000(1/p) = 120000$$

- b) Si la ganancia al lograr un contrato es de \$50,000.00, ¿vale la pena organizar los viajes?

Si se analizan solo los valores esperados, habría pérdida en promedio, pero se revisará estadísticamente la función de probabilidad de la ganancia.

$x$	$g(x; p= 0.1)$	$G(x; p=0.1)$	Ganancia neta
1	0.1	0.1	40000
2	0.09	0.19	30000
3	0.081	0.271	20000
4	0.0729	0.3439	10000
5	0.06561	0.40951	0
6	0.059049	0.468559	-10000
7	0.0531441	0.5217031	-20000
8	0.04782969	0.56953279	-30000
9	0.043046721	0.612579511	-40000
10	0.038742049	0.65132156	-50000

Nótese que si se organizan las visitas hasta un máximo de cuatro, la ganancia se mantiene positiva, luego habría pérdidas. La probabilidad de que se logre un contrato anual cuando mucho en cuatro visitas sería:

$$p(x \leq 4) = G(4; 0.1) = \sum_{t=1}^{t=4} (0.9)^{t-1} (0.1) = 0.3439$$

- c) Si el presupuesto para publicidad es sólo de \$100,000.00, ¿cuál es la probabilidad de que esta suma sea gastada sin lograr un sólo contrato en las visitas?

$x$	$g(x; p= 0.1)$	$G(x; p=0.1)$	Ganancia neta
1	0.1	0.1	40000
2	0.09	0.19	30000
3	0.081	0.271	20000
4	0.0729	0.3439	10000
5	0.06561	0.40951	0
6	0.059049	0.468559	-10000
7	0.0531441	0.5217031	-20000
8	0.04782969	0.56953279	-30000
9	0.043046721	0.612579511	-40000
10	0.038742049	0.65132156	-50000
11	0.034867844	0.686189404	-60000
12	0.03138106	0.717570464	-70000
13	0.028242954	0.745813417	-80000
14	0.025418658	0.771232075	-90000
15	0.022876792	0.794108868	-100000
16	0.020589113	0.814697981	-110000
17	0.018530202	0.833228183	-120000
18	0.016677182	0.849905365	-130000
19	0.015009464	0.864914828	-140000
20	0.013508517	0.878423345	-150000

En 15 visitas el presupuesto se agotaría y no se lograría una sola venta y la probabilidad de que esto ocurriera sería  $p(x > 15) = 1 - 0.794109 = 0.205891$ .



### 5.6. Modelo Probabilístico de Poisson, $p_o(x; c)$

Con respecto al modelo binomial,  $b(x; n, p)$ , para el cual, en un lote de tamaño infinito, se tiene un porcentaje  $p$  de artículos con alguna característica de interés; se obtiene una muestra de tamaño  $n$  y se contabiliza  $x$  el número de artículos en la muestra con la característica de interés, se calculará el límite cuando  $n$  tiende a infinito de tal manera que el producto  $np$  tienda a tomar un valor constante  $c$ ; esto sucede si la  $n$  tiende a crecer a la misma velocidad que la  $p$  tiende a decrecer en la misma proporción.

$$p_o(x; c) = \lim_{\substack{np \rightarrow c \\ n \rightarrow \infty}} b(x; n, p) = \lim_{\substack{np \rightarrow c \\ n \rightarrow \infty}} \binom{n}{x} p^x (1-p)^{n-x}$$

$$p_o(x; c) = \lim_{\substack{np \rightarrow c \\ n \rightarrow \infty}} \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!} p^x (1-p)^{n-x}$$

$$p_o(x; c) = \lim_{\substack{np \rightarrow c \\ n \rightarrow \infty}} \frac{n^x \left[ \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right) \right] p^x}{x!} \left(1 - \frac{c}{n}\right)^{n-x}$$

$$p_o(x; c) = \lim_{\substack{np \rightarrow c \\ n \rightarrow \infty}} \frac{(np)^x \left[ \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right) \right]}{x!} \frac{\left(1 - \frac{c}{n}\right)^n}{\left(1 - \frac{c}{n}\right)^x}$$

Pero  $\lim_{n \rightarrow \infty} \left(1 - \frac{u}{n}\right)^n = e^{-u}$

$$p_o(x; c) = e^{-c} \frac{c^x}{x!} \quad x = 0, 1, 2, \dots \quad (5.64)$$

A la función de probabilidad que se obtiene se le conoce como Función de Probabilidad de Poisson y es un modelo muy utilizado en la práctica profesional de un ingeniero; un gran número de variables aleatorias asociadas a fenómenos físicos son modeladas por medio de la distribución de Poisson.

En la Teoría de Líneas de Espera, el modelo de Poisson permite modelar el número de llegadas a una cola por unidad de tiempo. En Control Estadístico de Procesos, en la fabricación de productos que contienen muchas componentes (como un coche, un equipo de sonido, un refrigerador, una pantalla de televisión, etcétera) la variable relacionada con el número de defectos por unidad

producida puede ser modelada con esta distribución. En la Teoría de Inventarios es sumamente útil modelar la demanda durante el tiempo de entrega de un producto como un modelo de Poisson. Otra aplicación sumamente interesante es la probabilidad de ruina de una compañía aseguradora. Todos estos modelos tienen que ver con lo que se denomina un proceso de Poisson.

El proceso de Poisson, también conocido como ley de los sucesos raros, es un proceso estocástico de tiempo continuo, que consiste en "contar" eventos raros (de ahí el nombre "sucesos raros") que ocurren a lo largo del tiempo. El tiempo entre cada par de eventos consecutivos tiene una distribución exponencial con el parámetro  $\lambda$  (se verá posteriormente), y cada uno de estos tiempos entre llegadas se supone que es independiente de otros tiempos entre llegadas. Es llamado así por el matemático Simeon Denis Poisson (1781-1840).

**FIGURA 5.15.** Simeon Denis Poisson (1781-1840)

**NOTA:** Simeon Denis Poisson. (2020) [https://es.wikipedia.org/wiki/Simeon\\_Denis\\_Poisson](https://es.wikipedia.org/wiki/Simeon_Denis_Poisson)



Formalmente, el proceso de Poisson con intensidad (o tasa)  $\lambda \geq 0$  es un proceso de conteo subyacente en tiempo continuo  $\{x(t), t \geq 0\}$ , donde  $x(t)$  es una colección de variables aleatorias con las siguientes propiedades:

- a) Un evento puede ocurrir aleatoriamente en cualquier momento o en cualquier punto en el espacio, pero  $x(t=0) = 0$ .
- b) Si  $s \leq t$ , entonces  $x(s) \leq x(t)$ .

- c) Las ocurrencias de un evento en un intervalo de tiempo dado (o en un espacio dado), son independientes de las que ocurran en cualquier intervalo de tiempo o de espacio no traslapado. Es decir, para todo  $n > 0$  y  $0 < t_1 < t_2 < \dots < t_n$ , las variables aleatorias  $x(t_1)$ ,  $x(t_2) - x(t_1)$ ,  $\dots$ ,  $x(t_n) - x(t_{n-1})$  son estadísticamente independientes.
- d) Para toda  $h > 0$  y  $t \in R^+$ ,  $x(h)$  y  $x(t+h) - x(t)$  tienen la misma distribución.
- e) La probabilidad de que se presente con exactitud una ocurrencia o éxito en un intervalo con duración  $\Delta t$ , es aproximadamente igual a  $\lambda \Delta t$ . La aproximación es en el sentido de que la probabilidad es  $\lambda \Delta t + o_1(\Delta t)$ , donde la función  $[o_1(\Delta t)/\Delta t] \rightarrow 0$  cuando  $\Delta t \rightarrow 0$ .
- f) La probabilidad de que ocurran exactamente cero ocurrencias o éxitos en el intervalo de tiempo  $[t, t + \Delta t]$  es de aproximadamente  $1 - (\lambda \Delta t)$ . Una vez más, esto es en el sentido de que es igual a  $1 - (\lambda \Delta t) + o_2(\Delta t)$ , donde la función  $[o_2(\Delta t)/\Delta t] \rightarrow 0$  cuando  $\Delta t \rightarrow 0$ .
- g) La probabilidad de que se presenten dos o más ocurrencias o éxitos es igual a una cantidad  $o_3(\Delta t)$ , donde  $o_3(\Delta t)/\Delta t \rightarrow 0$  cuando  $\Delta t \rightarrow 0$ .

Sobre la base de las hipótesis *c* y *d*, por el Teorema de Probabilidad Total:

$$p[x(t+\Delta t) = x] = p[x(t) = x]p[x(\Delta t) = 0] + p[x(t) = x-1]p[x(\Delta t) = 1] + p[x(t) = x-2]p[x(\Delta t) = 2] + \dots \quad (5.65)$$

Entonces, con base en las hipótesis iniciales anteriores, se obtiene, usando la notación  $p_x(t) = p[x(t) = x]$ :

$$p_x(t) = [1 - \lambda \Delta t - o(\Delta t)^2 - \dots] p_x(t) + (\lambda \Delta t) p_{x-1}(t) + o(\Delta t)^2 p_{x-2}(t) + \dots \quad (5.66)$$

Despreciando los términos de alto orden, la ecuación anterior puede ser escrita como:

$$\frac{p_x(t+\Delta t) - p_x(t)}{\Delta t} = \lambda p_x(t) + \lambda p_{x-1}(t) \quad (5.67)$$

En el límite cuando  $\Delta t \rightarrow 0$ , se obtiene la ecuación diferencial para  $p_x(t)$ :

$$\frac{dp_x(t)}{dt} = \lambda p_x(t) + \lambda p_{x-1}(t) \quad (5.68)$$

La ecuación diferencial anterior es aplicable para cualquier valor de  $x \geq 1$ . Para  $x=0$ , el proceso de derivación anterior conduce a lo siguiente:

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t) \quad (5.69)$$

Si el proceso de conteo empieza desde cero, las condiciones iniciales asociadas a las ecuaciones (5.68) y (5.69) son  $p_0(0) = 1.0$  y  $p_x(0) = 0$ .

La solución de la ecuación diferencial (5.68), con la primera de las condiciones iniciales establecidas, para  $x=0$ , por el método de separación de variables es:

$$\frac{dp_0(t)}{dt} = -\lambda p_0(t)$$

$$\frac{dp_0(t)}{p_0(t)} = -\lambda dt$$

$$\ln[p_0(t)] = -\lambda t + k$$

$$p_0(t) = e^k e^{-\lambda t}$$

*Pero*

$$p(t=0) = e^k = 1$$

$$p_0(t) = e^{-\lambda t} \quad (5.70)$$

Para  $x \geq 1$ , las soluciones de la ecuación diferencial (5.68) son:

$$p_1(t) = \lambda t e^{-\lambda t}$$

$$p_2(t) = \frac{(\lambda t)^2}{2!} e^{-\lambda t} \quad (5.71)$$

Para  $x$  en general, la solución es:

$$p_x(t) = \frac{(\lambda t)^x}{x!} e^{-\lambda t} \quad (5.72)$$

que es la función de probabilidad de Poisson, en la cual, el parámetro  $\lambda$  es la tasa media de ocurrencia o de éxito del evento.

La Función de Probabilidad Acumulada de Poisson está dada por la expresión:

$$P_o(x; c) = \sum_{k=0}^{k=x} e^{-c} \frac{c^k}{k!} \quad x=0, 1, 2, \dots \quad (5.73)$$

que también puede ser escrita en términos del tiempo como:

$$P_o(x; c) = \sum_{k=0}^{k=x} e^{-\lambda t} \frac{(\lambda t)^k}{k!} \quad x=0, 1, 2, \dots \quad (5.74)$$

La Función Generatriz de Momentos de la Función de Probabilidad de Poisson está dada por:

$$FGM_x(t) = E\{e^{tx}\} = \sum_{x=0}^{x \rightarrow \infty} e^{tx} e^{-c} \frac{c^x}{x!} e^{-c} \sum_{x=0}^{x \rightarrow \infty} \frac{(ce^t)^x}{x!}$$

Pero

$$\sum_{x=0}^{x \rightarrow \infty} \frac{(u)^x}{x!} = e^u$$

$$FGM_x(t) = e^{c(e^t-1)} \quad (5.75)$$

Su Función Característica está dada por:

$$FC_x(t) = e^{c(e^{it}-1)} \quad (5.76)$$

Los primeros momentos con respecto al origen son:

$$\mu'_1 = \left[ \frac{dFGM_x(t)}{dt} \right]_{t=0} = c$$

$$\mu'_2 = \left[ \frac{d^2FGM_x(t)}{dt^2} \right]_{t=0} = c + c^2$$

$$\mu'_3 = \left[ \frac{d^3 FGM_x(t)}{dt^3} \right]_{t=0} = c + 3c^2 + c^3$$

$$\mu'_4 = \left[ \frac{d^4 FGM_x(t)}{dt^4} \right]_{t=0} = c + 7c^2 + 6c^3 + c^4$$

Se calculan a continuación los momentos con respecto a la media, de primero, segundo, tercero y cuarto orden:

$$\mu_x = \mu'_1 = c$$

$$\sigma_x^2 = E\{x^2\} - \mu_x^2 = (c+c^2) - (c^2) = c$$

$$\mu_3 = c$$

$$\mu_4 = c + 3c^2$$

Con los momentos anteriores se obtienen los parámetros estadísticos de la distribución de Poisson:

$$\text{Media} \quad \mu_x = c = \lambda \quad (5.77)$$

$$\text{Varianza} \quad \sigma_x^2 = c = \lambda \quad (5.78)$$

$$\text{Coeficiente de asimetría} \quad \gamma_1 = \frac{1}{\sqrt{c}} = \lambda^{-1/2} \quad (5.79)$$

$$\text{Coeficiente de Curtosis} \quad \gamma_1 = \frac{1}{c} = \lambda^{-1} \quad (5.80)$$

Para el cálculo de probabilidades con la distribución de Poisson, utilizando Excel se emplea el siguiente comando:

$$p(x) = p_o(x; \text{media, acumulado}) = \text{POISSON.DIST}(x, \lambda, 0 \text{ ó } 1) \quad (5.81)$$

En donde  $x$  es el número de éxitos u ocurrencias que se pretende lograr, el parámetro  $\lambda$  es la tasa media de ocurrencia o de éxito del evento; acumulado= verdadero o 1 si se pretende calcular la función de probabilidad acumulada desde cero hasta  $x$ ; acumulado= falso o 0 si se pretende calcular la probabilidad de que el número de éxitos u ocurrencias sea exactamente  $x$ .

De tal manera que:

$$p(x=3; \lambda=5) = \text{POISSON.DIST}(3,5,0) = 0.140374$$

$$p(x \leq 3; \lambda=5) = \text{POISSON.DIST}(3,5,1) = 0.265026$$

$$p(x > 3; \lambda=5) = 1 - \text{POISSON.DIST}(3,5,1) = 0.734974$$

$$p(4 < x \leq 6; \lambda=5) = \text{POISSON.DIST}(6,5,1) - \text{POISSON.DIST}(4,5,1) = 0.321690$$

$$p(4 \leq x \leq 6; \lambda=5) = \text{POISSON.DIST}(6,5,1) - \text{POISSON.DIST}(3,5,1) = 0.497158$$

Para hacer el cálculo de probabilidades con la distribución de Poisson, utilizando Minitab, se realizan los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en Poisson, como se muestra en la figura 5.2.
2. Aparece la pantalla que se muestra en la figura 5.16, en la cual, para calcular la probabilidad  $p(x=5, r=1, p=0.05)$  se señala si se trata de una probabilidad en un punto marcando Probability, en Event probability se indica la probabilidad de éxito en un ensayo 0.05, en Number of events needed se indica el valor de  $r=1$  y en Input constant el valor de  $x=5$  para el cual se desea calcular la probabilidad y se da un click en Ok obteniéndose el resultado.

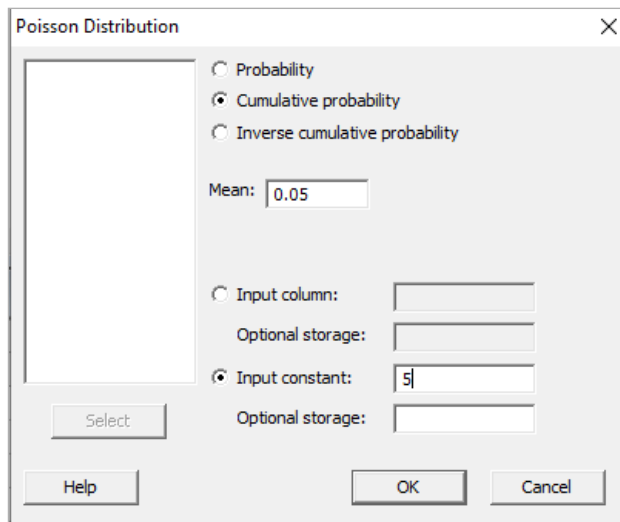


FIGURA 5.16

Si lo que se pretende calcular es la probabilidad acumulada hasta el valor  $x=5$ , en vez de elegir Probability debe elegirse Cumulative probability y todo lo demás igual.

En el lenguaje de desarrollo *R*, para el caso de la distribución de Poisson, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo de la probabilidad exactamente en un punto:

$$\text{dpois}(x, \lambda, \text{log} = \text{FALSE}) \quad (5.82)$$

Para el cálculo de la probabilidad acumulada hasta un punto dado:

$$\text{ppois}(x, \lambda, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) \quad (5.83)$$

En donde  $x$  es el número de éxitos, llegadas u ocurrencias por unidad de tiempo;  $\lambda$  es el número de éxitos, llegadas u ocurrencias en promedio por unidad de tiempo;  $\text{log}$  es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma FALSE.  $\text{lower.tail}$  es un valor lógico; si se le asigna un valor verdadero, el valor que arroja es la probabilidad acumulada a la izquierda de  $x$ , de lo contrario, si el valor que toma es falso, el valor que arroja es la probabilidad a la derecha del valor de  $x$ .

Por ejemplo,

$$p(x=3; \lambda=5) = \text{dpois}(3,5, \text{log}=\text{FALSE}) = 0.140374$$

$$p(x \leq 3; \lambda=5) = \text{ppois}(3,5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.265026$$

$$p(x > 3; \lambda=5) = \text{ppois}(3,5, \text{lower.tail}=\text{FALSE}, \text{log.p}=\text{FALSE}) =$$

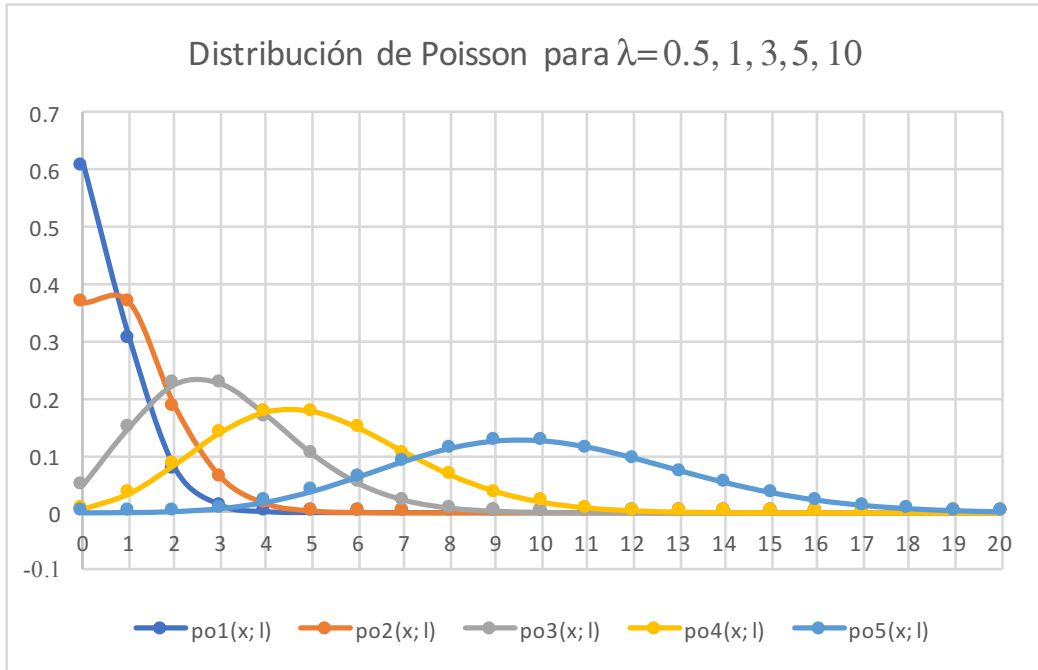
$$p(4 < x \leq 6; \lambda=5) = \text{ppois}(6, 5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{ppois}(4, 5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.321690$$

$$p(4 \leq x < 6; \lambda=5) = \text{ppois}(6,5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{ppois}(3, 5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.497158$$



En la figura 5.17 se muestran algunas gráficas de la distribución de Poisson para diferentes valores de  $\lambda$ .

**FIGURA 5.17.** Gráficas de la distribución de Poisson



### Ejercicio 5.9

Una empresa automotriz ha registrado a lo largo del tiempo, que en el proceso de fabricación de automóviles comete dos defectos en promedio por unidad producida.

- a) Calcule la probabilidad de que en una unidad en particular no aparezcan defectos.

$$p(x=0) = p_o(x=0; \lambda=2) = e^{-2} \frac{2^0}{0!} = 0.135335$$

- b) Calcule la probabilidad de que en una unidad en particular aparezcan menos de tres defectos.

$$p(x<3) = p(x \leq 2) = P_o(x \leq 2; c=2) = \sum_{t=0}^{t=2} e^{-2} \frac{2^t}{t!} = 0.676676$$

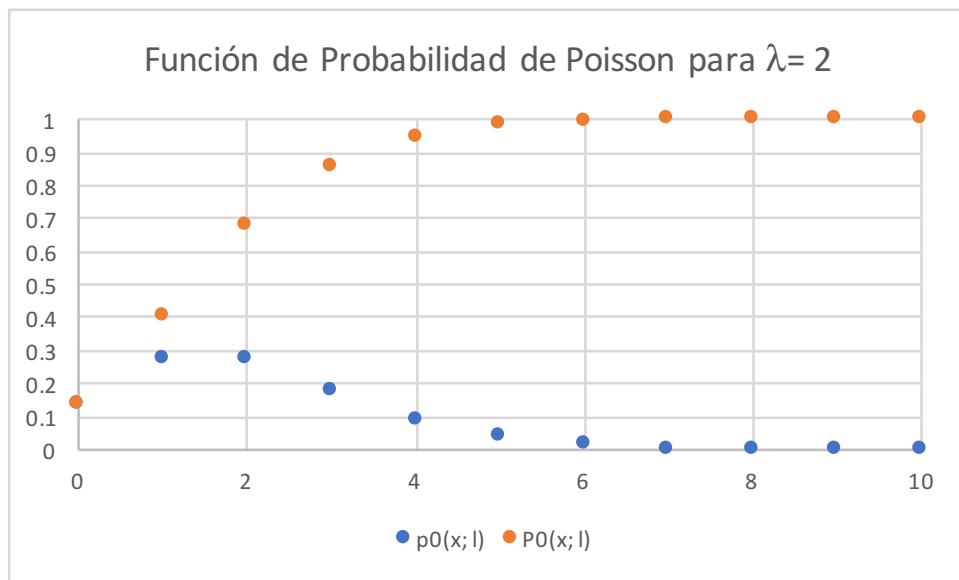
- c) Obtenga la probabilidad de que en una unidad en particular aparezcan entre dos y cinco defectos inclusive.

$$p(2 \leq x \leq 5) = P_o(2 \leq x \leq 5; c=2) = \sum_{t=2}^{t=5} e^{-2} \frac{2^t}{t!} = 0.577431$$

- d) Obtenga la probabilidad de que en una unidad en particular aparezcan más de cinco defectos.

$$p(x > 5) = 1 - p(x \leq 5) = 1 - P_o(x \leq 5; c=2) = 1 - \sum_{t=0}^{t=5} e^{-2} \frac{2^t}{t!} = 0.016564$$

- e) Grafique la función de probabilidad y la función de probabilidad acumulada.



### 5.7 Modelo Probabilístico Multinomial,

$$p_m(x_1, x_2, \dots, x_k; p_1, p_2, \dots, p_k)$$

El modelo multinomial es una generalización del modelo binomial. Suponga que se tiene un espacio de eventos que se particiona en  $k$  subconjuntos mutuamente excluyentes entre sí. Suponga que se eligen  $N$  elementos aleatoriamente y que resultan  $x_1$  elementos de la primera partición,  $x_2$  elementos de la segunda partición y así sucesivamente  $x_k$  elementos de la  $k$ -ésima partición. Suponga que la probabilidad de obtener un elemento de la primera partición es  $p_1$ , un elemento de la segunda partición es  $p_2$  y así sucesivamente, un elemento de la  $k$ -ésima partición es  $p_k$ ; de tal forma que, aplicando el teorema fundamental del conteo, su función de probabilidad está dada por:

$$p(x_1, x_2, \dots, x_k) = \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k} \quad (5.84)$$

para

$$\begin{aligned} x_i &= 0, 1, 2, \dots, n \\ i &= 1, 2, \dots, n \end{aligned}$$

donde

$$\sum_{i=1}^{i=k} x_i = n$$

Debe tenerse en cuenta que las variables aleatorias  $x_1, x_2, \dots, x_k$  no son estadísticamente independientes, porque la suma de todas las  $x_i$  debe ser  $n$ .

Las medias y las varianzas de la distribución multinomial son:

$$\begin{aligned} \mu_i &= E\{x_i\} = np_i \\ \sigma_{x_i}^2 &= np_i(1-p_i) \end{aligned} \quad (5.85)$$

La covarianza entre dos de las variables  $x_i$  y  $x_j$  está dada por:

$$\sigma_{x_i x_j}^2 = \text{cov}(x_i, x_j) = -np_i p_j \quad i \neq j \quad (5.86)$$

La Función Generadora de Momentos de la Distribución Multinomial es:

$$FGM(t) = \left( \sum_{i=1}^{i=k} p_i e^{it} \right)^n \quad (5.87)$$

Su Función Característica está dada por la expresión:

$$FC_x(t) = \left( \sum_{j=1}^{j=k} p_j e^{it_j} \right)^n \quad (5.88)$$

### Ejercicio 5.10

En los juegos de naipes existen dos tipos de barajas:

1. La baraja americana que consta de 4 figuras diferentes: picas ♠, corazones ♥, diamantes ♦ y tréboles ♣; para cada una de estas figuras existen los números del 1 al 10, el jack J, la reina Q y el rey R, lo que implica que la baraja americana consta de 52 cartas; al uno se le denomina el As, A, y es la carta de mayor peso en el juego de póker.
2. La baraja española que consta de 4 figuras diferentes: espadas ♠, bastos ♣, oros 🀄 y copas 🀅; para cada una de estas figuras existen los números del 1 al 7, además de la sota, el caballo y el rey, lo que implica que la baraja española consta de 40 cartas.

Suponga que a un jugador cualquiera, aleatoriamente se le reparten cinco cartas, sin derecho a descartarlas, es decir a cambiarlas por otras; el modelo probabilístico que representa al número de cartas de cada figura en cada juego no es multinomial, ya que el número de cartas totales es finito y al extraer una carta se afecta la probabilidad de obtener otra carta de la misma figura, por lo que la probabilidad de éxito no es constante; sin embargo, si se hiciera un muestreo con reemplazo, es decir, se toma una carta se observa y se reintegra al mazo volviendo a barajar, se tienen probabilidades constantes de éxito de cada figura.

Sea  $x_1$  el número de cartas extraídas de picas,  $x_2$  el número de cartas extraídas de corazones,  $x_3$  el número de cartas extraídas de diamantes y  $x_4$  el número de cartas extraídas de tréboles, para el caso de la baraja americana: hay un lote con  $N= 52$  cartas, donde existen 13 cartas de cada figura, se asignan  $N= 5$  cartas a un jugador. De esta forma, el modelo si se puede considerar como multinomial.

- a) ¿Cuál es la probabilidad de que las cinco cartas sean de corazones?

$$p(x_1=0, x_2=5, x_3=0, x_4=0) = \frac{5!}{0!5!0!0!} \left(\frac{1}{4}\right)^0 \left(\frac{1}{4}\right)^5 \left(\frac{1}{4}\right)^0 \left(\frac{1}{4}\right)^0$$

$$p(x_1=0, x_2=5, x_3=0, x_4=0) = 0.00098$$

- b) ¿Cuál es la probabilidad de obtener tres diamantes y dos tréboles?

$$p(x_1=0, x_2=0, x_3=3, x_4=2) = \frac{5!}{0!0!3!2!} \left(\frac{1}{4}\right)^0 \left(\frac{1}{4}\right)^0 \left(\frac{1}{4}\right)^3 \left(\frac{1}{4}\right)^2$$

$$p(x_1=0, x_2=0, x_3=3, x_4=2) = 0.00488$$

- c) Si el muestreo se hiciera sin reemplazo, ¿Cuál sería la probabilidad de obtener cinco corazones?

$$p(x_1=0, x_2=5, x_3=0, x_4=0) = \frac{13(12)(11)(10)(9)}{52(51)(50)(49)(48)} =$$

$$p(x_1=0, x_2=5, x_3=0, x_4=0) = 0.000495$$

Como se puede apreciar, el resultado obtenido en el inciso (a) es bastante aproximado al resultado real sin reemplazo.

- d) ¿Cuál es la probabilidad de obtener una flor imperial al jugar póker cerrado?

Se le llama Flor a obtener cinco cartas de la misma figura y en orden sucesivo (como una escalera). La Flor Imperial es aquella que está conformada con la misma figura en escalera, pero con las cinco cartas más grandes de dicha figura; en este caso serían A, R, Q, J, 10. De esta forma:

$$p(\text{Flor imperial}) = \frac{20(4)(3)(2)(1)}{52(51)(50)(49)(48)} =$$

$$p(x_1=0, x_2=5, x_3=0, x_4=0) = 1.539 E(-6) = 0.00000154$$

La probabilidad de obtener una flor imperial es de 1.54 en un millón de oportunidades. Este es el juego más alto que existe, muy por arriba del tradicional póker de ases.

## Ejercicios propuestos del capítulo 5

- 5.1. Clasifique las siguientes variables aleatorias como discretas o continuas:
- X: El número de accidentes de automóvil por año en Veracruz.
  - Y: El tiempo invertido para jugar 18 hoyos de golf.
  - M: La cantidad de leche que produce cada año una vaca específica.
  - N: El número de huevos que pone cada mes una gallina.
  - P: El número de permisos de construcción que se autorizan cada mes en determinada ciudad.
  - Q: El peso del grano de maíz producido por acre.
- 5.2. Anote (sin hacer cálculos) el nombre de la distribución teórica que utilizaría para obtener la solución a los siguientes problemas.
- a) Se ha observado que las fallas mecánicas ocurridas en una planta son a razón de una cada dos horas. Si llega a la planta un día cualquiera ¿cuál es la probabilidad de que transcurra al menos una hora antes de que se produzca una falla?  
Distribución \_\_\_\_\_
  - b) Se considera que un competidor de tiro al blanco tiene una probabilidad de acertar de 0.7. Hallar la probabilidad de que su primer blanco lo logre en el cuarto disparo.  
Distribución \_\_\_\_\_
  - c) El promedio de accidentes de tránsito en la Ciudad de México es de 72 por día. Hallar la probabilidad de que en un día se produzcan más de 20 accidentes.  
Distribución \_\_\_\_\_
- 5.3. Se lanza una moneda hasta que se presentan tres águilas en forma sucesiva. ¿Es este un espacio muestral discreto? Explique su respuesta.

- 5.4. En una sucesión de ensayos de Bernoulli, obtenga la probabilidad de obtener un éxito en el  $i$ -ésimo ensayo, dado que se sabe que hay exactamente  $r$  éxitos en la sucesión.
- 5.5. Un lote de siete dispositivos contiene dos defectuosos, un hotel adquiere de forma aleatoria tres. Si  $x$  es el número de aparatos defectuosos adquiridos por el hotel, obtenga la distribución de probabilidad de  $x$ .
- Expresar los resultados de forma gráfica a través de un histograma de probabilidad.
  - Encuentre la distribución acumulada de variable aleatoria  $x$  que representa el número de artículos defectuosos.
  - Utilizando  $p(x)$ , encuentre:
    - $p(x=1)$
    - $p(0 < x \leq 2)$
- 5.6. Determinar la probabilidad de que cuatro de cinco cartas extraídas aleatoriamente y sin reposición, de una baraja normal de 52 cartas, sean ases. ¿Cuál sería la probabilidad si se quiere que salgan los 4 ases primero?
- 5.7. Un estudiante de la Facultad de Ingeniería de la UNAM planea llevar seis materias independientes este semestre. La probabilidad estimada de éxito en cada materia es de 0.9 ¿cuál es la probabilidad de que al menos en cinco de las materias tenga éxito?
- 5.8. Un lote de 30 balatas se somete a procedimiento de prueba de aceptación. El procedimiento consiste en seleccionar seis balatas aleatoriamente, sin reemplazo y probarlas. Si dos o menos balatas fallan, se aceptan las restantes. De otro modo se rechaza el lote. Considérese que el lote contiene siete balatas defectuosas.
- ¿Cuál es la probabilidad exacta de aceptar el lote?
  - ¿Cuál es la probabilidad de aceptar el lote, calculándola mediante la distribución binomial?
- 5.9. Una empresa del sector eléctrico recibe pequeños lotes ( $N=32$ ) de transformadores eléctricos de potencia. Decide rechazar el lote el 95% de las veces si este contiene ocho transformadores defectuosos. Supóngase que decide que la presencia de un dispositivo defectuoso en la muestra es suficiente para rechazar ¿cuál debería ser el tamaño de la muestra que seleccione?

- 5.10. En una fábrica de bicicletas eléctricas, 2% de los motores de bicicletas producidos tiene algún defecto.
- ¿Cuál es la probabilidad de que en un lote de 100 motores se encuentren más de 3 defectuosos?
  - ¿Cuál es la probabilidad de que en un lote de 1000 motores se encuentren entre 25 y 30 defectuosos inclusive?
  - ¿Cuál es la probabilidad de que en un lote de 3 motores exista uno defectuoso?
- 5.11. Un proceso de producción de tarjetas sim opera en promedio, con una fracción de defectos del 2%. Cada dos horas se selecciona del proceso una muestra de tamaño 50. Si la muestra contiene más de tres defectos, el proceso deberá detenerse. Determinar la probabilidad de que el proceso tenga que ser detenido debido al esquema de muestreo.
- 5.12. Sea  $p=1\%$  la probabilidad de que un cierto tipo de led eléctrico falle en una prueba de 24 horas. Calcular la probabilidad de que un tablero consistente de 50 de esos led funcionará 24 horas sin fallas debidas a los focos.
- 5.13. Un fabricante sabe que, en promedio, el 2% de los calentadores eléctricos que produce requerirán reparación en los 90 días siguientes a su venta. Obtener la probabilidad de que entre 1500 de los calentadores al menos 30 requerirán reparación en los primeros 90 días después de su venta. Utilizar una aproximación y explicar por qué se puede aplicar dicha aproximación.
- 5.14. Si hay en promedio 1% de zurdos, calcúlese la probabilidad de tener por lo menos cuatro zurdos entre 200 personas.
- 5.15. La compañía aseguradora ING ha descubierto que el 0.2% de la población de la Ciudad de México recibe heridas, como resultado de cierto tipo de accidentes automovilísticos. La compañía tiene 15000 asegurados cubiertos contra tal accidente.
- ¿Cuál es la probabilidad de que tres o menos reclamaciones se presenten por este tipo de accidente durante el siguiente año?
  - ¿Cinco o más reclamaciones?
- 5.16. Un vendedor de seguros vende pólizas a 5 hombres, todos de la misma edad y con buena salud. De acuerdo con las tablas actuariales, la probabilidad de



- que un hombre de esta edad viva 30 años más es de  $\frac{2}{3}$ . Calcular la probabilidad que a los 30 años vivan: a) los 5 hombres; b) al menos 3; c) solamente 2.
- 5.17. Sean  $x$  y  $y$  dos variables aleatorias discretas con distribución desconocida.
- Para  $x$  suponga que  $n = 15$ ,  $p = 0.25$  y para  $y$  suponga que  $n = 15$  y  $p = 0.5$ , estimar las siguientes probabilidades utilizando para ello el Teorema de Chebyshev:  $p(x = 8)$ ,  $p(x \leq 3)$ ,  $p(y \leq 7)$ ,  $P(y \geq 9)$  y  $p(x + y \geq 12)$ .
  - Calcúlense las probabilidades anteriores mediante el empleo de la distribución binomial.
  - Estímense las probabilidades anteriores mediante el empleo de la distribución normal.
  - Repetir los incisos  $a$ ,  $b$  y  $c$  para  $n = 25$  y comparar los resultados.
- 5.18. En un concurso de tiro un competidor que desea colocarse de líder se dispone a tirar y lo hará hasta que su tiro dé en el blanco. Si se supone que la probabilidad de que dé en el blanco es de 0.7. Calcular la probabilidad de que necesite por lo menos 5 tiros.
- 5.19. Se han capacitado cinco obreros para utilizarlos en manufactura enfocada. La probabilidad de que un obrero capacitado tenga éxito en su actividad es de 0.95. Considerando actividades estadísticamente independientes, ¿cuál es la probabilidad de que el primer obrero que falle sea el quinto?
- 5.20. Un arquero experto da en la diana el 95% de las veces, ¿Cuál es la probabilidad de que falle por primera vez en su decimoquinto disparo?
- 5.21. La probabilidad de que un relevador se detenga en un ensayo es de 0.2. ¿cuántos relevadores deben ser ensayados para que la probabilidad de que por lo menos 3 paren sea más cercana a 0.9?
- 5.22. La probabilidad de que un jugador de dominó acierte en la ficha que tira es 0.75. El experimento habrá de repetirse hasta que ocurran cinco fichas exitosas ¿Cuál es el número esperado de repeticiones necesarias? ¿Cuál es la varianza?
- 5.23. En una gasolinera se ha observado que se presentan un promedio de cuatro automóviles por hora a cargar gasolina, ¿cuál es la probabilidad de que en un día cualquiera se presenten más de 5 automóviles entre 7:00 pm y las 8:00 pm (en una hora)?

- 5.24. Los mecánicos de una agencia automotriz llegan al área de refacciones pidiendo una banda para motor, siguiendo una distribución de Poisson con parámetro  $\lambda = 2$ . Normalmente el área de refacciones tiene un inventario de tres de estas partes de repuesto. Si se solicitan más de tres bandas, los mecánicos deberán ir a conseguir estas piezas a Iztapalapa.
- En cierto día, ¿cuál es la probabilidad de que un mecánico tenga que trasladarse a Iztapalapa?
  - ¿Cuál es la demanda diaria esperada de bandas para motor?
  - ¿Cuántas bandas para motor de repuesto deberán tenerse en el área de refacciones, para dar servicio adecuado el 90% de las veces?
  - ¿Cuál es el número esperado de mecánicos que recibirán servicio adecuado diariamente en el área de refacciones?
  - ¿Cuál es el número esperado de mecánicos que realizarán el viaje a Iztapalapa?
- 5.26. El número de llegadas a la cola de una rosticería sigue una distribución de Poisson con media de 10. Los despachadores de la rosticería pueden manejar hasta 20 llegadas sin tener una sobrecarga. ¿Cuál es la probabilidad de que ocurra una sobrecarga?
- 5.27. La probabilidad de que un dron, después de colocarlo en órbita, funcione de manera adecuada es 0.9. Supóngase que cinco de estos se colocan en órbita y operan de manera independiente:  
¿Cuál es la probabilidad de que, por lo menos, el 80% funcione adecuadamente:
- Para  $n = 10$
  - Para  $n = 20$
  - ¿Son inesperados estos resultados?, ¿por qué?
- 5.28. Suponga que en la Ciudad de México, la probabilidad de que ocurra un sismo de una magnitud mayor a 6 durante cualquier día del año es 0.011. Considerando que existe independencia estadística, ¿cuál es la probabilidad de que el primer sismo ocurra el 28 de enero?
- 5.29. No es el caso, pero suponga que la probabilidad de que un avión sufra un accidente a nivel mundial es 0.00001, ¿considera lógico este valor? Supóngase que 120000 aviones vuelan diariamente a nivel mundial, ¿cuál es la probabilidad de que en un día cualquiera no ocurran accidentes de

avión?, ¿cuál es la probabilidad de que ocurran dos o más accidentes? Fundamente su respuesta.

- 5.30. El número de colonias de bacterias de cierto tipo, en muestras de agua contaminada, tiene una distribución de Poisson con una media de dos colonias por centímetro cúbico.
- Si se toman en forma independiente cuatro muestras de un centímetro cúbico de esta agua, encuentre la probabilidad de que al menos una muestra tenga una o más colonias de bacterias.
  - ¿Cuántas unidades muestrales de un centímetro cúbico deben seleccionarse? para tener una probabilidad de aproximadamente 0.95 de encontrar al menos una colonia de bacterias
- 5.31. Un voceador vende periódicos a razón de “m” por hora en promedio. Si alguien acaba de comprar un periódico, ¿cuál es la probabilidad de que transcurran dos minutos antes de que se venda otro periódico?
- 5.32. El número de clientes que llega a un banco se comporta como una variable aleatoria de Poisson. Si el número promedio es de 120 por hora, ¿cuál es la probabilidad de que en un minuto lleguen por lo menos tres clientes?, ¿puede esperarse que la frecuencia de llegada de los clientes al banco sea constante en un día cualquiera?
- 5.33. El número promedio de llamadas telefónicas que llegan al conmutador de una compañía entre las 10 y las 12 horas del día es a razón de 2 llamadas por minuto. Si el conmutador tiene capacidad para recibir 2 llamadas simultáneamente, ¿cuál es la probabilidad de que una persona que llame entre esas horas reciba la señal de ocupado?
- 5.34. Un libro de Fundamentos de Probabilidad tiene 340 páginas en las cuales pueden existir errores tipográficos. Si en la realidad hay 27 errores distribuidos aleatoriamente entre estas 340 páginas ¿cuál es la probabilidad de que una muestra aleatoria de 40 páginas contenga al menos un error? ¿qué tan grande deberá ser la muestra aleatoria para asegurarse de que al menos cinco errores serán detectados con una probabilidad del 90%?

## 6. Modelos probabilísticos continuos de fenómenos aleatorios

En este tema se analizan ciertos fenómenos reales donde se requiere considerar la incertidumbre y cuyos modelos probabilísticos continuos serán deducidos junto con sus principales propiedades, incluyendo ejemplos de aplicación en la práctica profesional de la Ingeniería.

Las características que son medibles numéricamente con instrumentos son más comunes que las características que no son medibles numéricamente y de éstas, la inmensa mayoría presentan un intervalo de variación continuo, de allí la importancia de conocer estos modelos de probabilidad continuos.

Los modelos probabilísticos que se verán para el caso continuo se resumen en la tabla de la figura 6.1.

FIGURA 6.1. Modelos Probabilísticos Continuos

Modelos Probabilísticos continuos							
Distribución	Parámetros	Valores que toma x	Función Densidad de Probabilidad: f(x)	Función de Distribución o de Probabilidad Acumulada: F(x)=P(t≤x)	Media	Varianza	Función Generatriz de Momentos
Uniforme	a, b,	a ≤ x ≤ b x ∈ ℝ	f(x) = $\begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{en otro caso} \end{cases}$	F(x) = $\begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases}$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$	$\frac{e^{bt} - e^{at}}{(b-a)t}$
	b > a						
Triangular	a, b, c	a ≤ x ≤ c x ∈ ℝ	f(x) = $\begin{cases} \frac{2(x-a)}{(c-a)(b-a)} & \text{si } a \leq x \leq b \\ \frac{2(c-x)}{(c-a)(c-b)} & \text{si } b < x \leq c \\ 0 & \text{en cualquier otro caso} \end{cases}$	F(x) = $\begin{cases} 0 & \text{si } x < a \\ \frac{(x-a)^2}{(c-a)(b-a)} & \text{si } a \leq x \leq b \\ 1 - \frac{(c-x)^2}{(c-a)(c-b)} & \text{si } b < x \leq c \\ 1 & \text{si } x > c \end{cases}$	$\frac{c+b+a}{3}$	$\frac{1}{18} [a^2+b^2+c^2-ab-ac-bc]$	
	c > b > a						
Exponencial	λ > 0	x > 0 x ∈ ℝ <sup>+</sup>	f(x) = $\begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{en otro caso} \end{cases}$	F(x) = $\begin{cases} 0 & x \leq 0 \\ 1 - e^{-\lambda x} & x > 0 \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\left(1 - \frac{t}{\lambda}\right)^{-r}$
Normal	μ ∈ ℝ σ ∈ ℝ	x ∈ ℝ	f(x) = $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$	F(x) = $\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\frac{(t-\mu)^2}{\sigma^2}} dt$	μ	σ	$e^{\left[\mu + \frac{\sigma^2 t^2}{2}\right]}$
LogNormal	μ ∈ ℝ σ ∈ ℝ	x > 0 x ∈ ℝ <sup>+</sup> y = Ln(x)	f(x) = $\frac{1}{x\sigma_y\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(\ln(x)-\mu_y)^2}{\sigma_y^2}}$	F(x) = $\frac{1}{\sigma_y\sqrt{2\pi}} \int_0^x \frac{e^{-\frac{1}{2}\frac{(\ln(t)-\mu_y)^2}{\sigma_y^2}}}{t} dt$	$\mu_x = e^{\left[\mu_y + \frac{\sigma_y^2}{2}\right]}$	$\sigma_x^2 = \mu_x^2 (e^{\sigma_y^2} - 1)$	
Gamma	r > 0 λ > 0	x > 0	f(x) = $\frac{\lambda}{\Gamma(r)} (\lambda x)^{r-1} e^{-\lambda x}$	F(x) = $\frac{\lambda}{\Gamma(r)} \int_0^x (\lambda t)^{r-1} e^{-\lambda t} dt$	$\frac{r}{\lambda}$	$\frac{r}{\lambda^2}$	$\left(1 - \frac{t}{\lambda}\right)^{-1}$
Beta	r > 0 λ > 0	0 ≤ x ≤ 1	f(x) = $\frac{\Gamma(\lambda+r)}{\Gamma(\lambda)\Gamma(r)} x^{\lambda-1} (1-x)^{r-1}$ 0 ≤ x ≤ 1 λ > 0 r > 0	F(x) = $\frac{\Gamma(\lambda+r)}{\Gamma(\lambda)\Gamma(r)} \int_0^x t^{\lambda-1} (1-t)^{r-1} dt$	$\mu_x = \frac{\lambda}{\lambda+r}$	$\sigma_x^2 = \frac{\lambda r}{(\lambda+r)^2 (\lambda+r+1)}$	
Weibull	γ ∈ ℝ δ > 0 β > 0	x ≥ γ	f(x) = $\frac{\beta}{\delta} \left(\frac{x-\gamma}{\delta}\right)^{\beta-1} e^{-\left(\frac{x-\gamma}{\delta}\right)^\beta}$	F(x) = $\frac{\beta}{\delta} \int_\gamma^x \left(\frac{t-\gamma}{\delta}\right)^{\beta-1} e^{-\left(\frac{t-\gamma}{\delta}\right)^\beta} dt$	$\mu_x = \gamma + \delta \Gamma\left(\frac{1}{\beta} + 1\right)$	$\sigma_x^2 = \delta^2 \left\{ \Gamma\left(\frac{2}{\beta} + 1\right) - \left[ \Gamma\left(\frac{1}{\beta} + 1\right) \right]^2 \right\}$	

## 6.1. Distribución uniforme continua

En el ejercicio 3.10 de la página 115 se hizo el análisis estadístico de la Distribución Uniforme Continua, la cual en términos generales se muestra graficada en la figura 3.20; su función de probabilidad es la expresión definida en (3.50), su función de probabilidad acumulada se define en la expresión (3.51), su media en (3.52), su varianza en (3.53), su coeficiente de asimetría en (3.56), su coeficiente de curtosis en (3.57), su función generatriz de momentos en (3.58) y su función característica en (3.59).

Solo se ejemplificará la forma en la que se utiliza esta distribución.

### Ejercicio 6.1

El tiempo de troquelado de una lámina de acero con una máquina de control numérico, en el laboratorio de manufactura avanzada, se comporta como una función de probabilidad uniforme en el intervalo [5, 12].

a) Obtenga sus principales características.

Función de probabilidad:

$$f(x) = \left\{ \begin{array}{ll} \frac{1}{7} & 5 \leq x \leq 12 \\ 0 & \text{Cualquier otro caso} \end{array} \right\}$$

Función de Probabilidad Acumulada:

$$F(x) = \left\{ \begin{array}{ll} 0 & x < 5 \\ \frac{x-5}{7} & 5 \leq x \leq 12 \\ 1 & x > 12 \end{array} \right\}$$

Su media, varianza, coeficiente de asimetría, coeficiente de curtosis, función generatriz de momentos y función característica son:

$$\mu_x = \frac{a+b}{2} = 8.5$$

$$\sigma_x^2 = \frac{(b-a)^2}{12} = \frac{49}{12} = 4.0833$$

$$\gamma_1 = 0$$

$$\gamma_3 = \frac{9}{5} (b-a) - 3 = 9.6$$

$$FGM_x(t) = \frac{e^{12t} - e^{5t}}{7t}; \quad FC_x(t) = \frac{e^{12t} - e^{5t}}{7t}$$

b) Calcule las siguientes probabilidades:

$$p(x \leq 10) = \frac{10-5}{7} = \frac{5}{7} = 0.7143$$

$$p(x > 11) = 1 - p(x \leq 11) = 1 - \frac{11-5}{7} = \frac{1}{7} = 0.1429$$

$$p(6 \leq x \leq 10) = \frac{10-5}{7} - \frac{6-5}{7} = \frac{4}{7} = 0.5714$$

c) Obtenga el valor de  $x$  que hace que se tenga una probabilidad acumulada

$$p(t \leq x) = 0.75$$

$$p(t \leq x) = \frac{x-5}{7} = 0.75$$

$$x = 5 + 7(0.75) = 10.25$$

d) Genere 10 números aleatorios con esta distribución de probabilidad uniforme continua.

$$u = F(t \leq x) = \frac{x-5}{7}$$

Despejando:

$$x = 5 + 7u$$

Se generan 10 números aleatorios entre cero y uno con la siguiente instrucción en Excel:

`u = aleatorio()`

De tal manera que:

11.79	11.49
8.28	10.67
9.21	5.10
5.47	7.87
5.95	10.09

## 6.2. Distribución Triangular

En el ejercicio 3.11 de la página 119 se hizo el análisis estadístico de la Distribución Triangular, la cual en términos generales se muestra graficada en la figura 3.21; su función de probabilidad es la expresión definida en (3.60), su función de probabilidad acumulada se define en la expresión (3.61), su media en (3.62) y su varianza en (3.65). Sólo se ejemplificará la forma en la que se utiliza esta distribución.

### Ejercicio 6.2

(Problema 15.4, página 277, del libro de Raúl Coss Bu, *Análisis y Evaluación de Proyectos de Inversión*, editorial LIMUSA, 1995). La compañía El Encanto desea entrar en un nuevo negocio que promete generar los flujos de efectivo con distribución triangular que se muestran a continuación. Si los flujos de efectivo de un período a otro son independientes, y la TREMA (tasa de rentabilidad mínima atractiva) es  $i=15\%$ , ¿cuál es la probabilidad de que el valor presente esperado sea mayor que cero? Si un proyecto es aceptado cuando  $\text{Prob}(E\{VPN>0\})>0.95$ , ¿qué decisión tomaría la compañía?

Año	Estimación Pesimista	Estimación más Probable	Estimación Optimista
0	-\$300	-\$250	-\$200
1	80	100	120
2	75	100	125
3	80	120	140
4	90	120	150
5	100	130	150

Nótese que el flujo de efectivo en cada período es de tipo triangular y además son simétricas cada una de ellas.

De conformidad con las expresiones (3.62) y (3.65) las medias y las varianzas de cada año, dado que se trata de distribuciones triangulares estarían dadas en la siguiente tabla:

Año	Estimación Pesimista	Estimación más Probable	Estimación Optimista	Media	Varianza
0	-300	-250	-200	-250.00	416.67
1	80	100	120	100.00	66.67
2	75	100	125	100.00	104.17
3	80	120	140	113.33	155.56
4	90	120	150	120.00	150.00
5	100	130	150	126.67	105.56



El valor presente neto de una propuesta de inversión, sin considerar inflación, se calcula de acuerdo con la siguiente expresión:

$$VPN = \sum_{j=0}^{j=n} \frac{x_j}{(1+i)^j} \quad (6.1)$$

Donde  $x_j$  es una variable aleatoria con distribución triangular, que representa el flujo de efectivo neto del período  $j$  y cuya media y varianza son  $\mu_j$  y  $\sigma_j^2$  respectivamente. La  $j$  representa el año y la  $i$  la tasa de rentabilidad mínima atractiva (TREMA).

El Valor Presente Neto Esperado estaría dado por la esperanza matemática del VPN, es decir:

$$E\{VPN\} = E\left\{ \sum_{j=0}^{j=n} \frac{x_j}{(1+i)^j} \right\} = \sum_{j=0}^{j=n} \frac{E\{x_j\}}{(1+i)^j} = \sum_{j=0}^{j=n} \frac{\mu_j}{(1+i)^j} \quad (6.2)$$

De la misma forma, la varianza del VPN estaría dada por la siguiente expresión:

$$Var\{VPN\} = Var\left\{ \sum_{j=0}^{j=n} \frac{x_j}{(1+i)^j} \right\} = \sum_{j=0}^{j=n} \frac{Var\{x_j\}}{(1+i)^{2j}} = \sum_{j=0}^{j=n} \frac{\sigma_j^2}{(1+i)^{2j}} \quad (6.3)$$

En el siguiente cuadro se hacen los cálculos en Excel para estimar el valor presente neto esperado y su varianza.

Año	$(1+i)^j$	Media	Varianza	Media/ $(1+i)^j$	Varianza/ $(1+i)^{2j}$
0	1	-250.00	416.67	-250.00	416.67
1	1.15	100.00	66.67	86.96	50.41
2	1.3225	100.00	104.17	75.61	59.56
3	1.520875	113.33	155.56	74.52	67.25
4	1.749006	120.00	150.00	68.61	49.04
5	2.011357	126.67	105.56	62.98	26.09

Media

VNP= 118.68

Var VPN= 669.01

Como ya se mencionó el flujo de efectivo en cada año tiene distribución triangular y esta es simétrica, por lo que se puede suponer que tiene distribución aproximadamente normal y por el teorema de aditividad de la distribución normal o por el teorema del límite central, se puede suponer que la suma es normal (estrictamente hablando, una distribución triangular es un caso particular de distribución beta como más adelante se verá, y la suma de dos o más distribuciones beta estadísticamente independientes da a su vez una distribución tipo beta, que para este caso particular sería también una triangular).

Calculando la probabilidad de que el valor presente neto esperado sea positivo, suponiendo que la distribución es normal:

$$p(E\{VPN\} > 0) = p(x > 0 \mid x \approx N(\mu_x = 118.68, \sigma_x^2 = 669.01))$$

$$p(E\{VPN\} > 0) = 0.999998$$

Dado que un proyecto es aceptado cuando  $\text{Prob}(E\{VPN > 0\}) > 0.95$ , la decisión debe ser aceptar invertir en el proyecto.

### 6.3. Distribución Exponencial Negativa

La distribución de probabilidad de Poisson que se dedujo en el subtema 5.6, a partir de lo que se conoce como Proceso de Poisson o Ley de los Sucesos Raros que ocurren a lo largo del tiempo, específicamente, la expresión (5.72), establece la probabilidad de que ocurran  $x$  éxitos en un período de tiempo dado por  $\lambda t$ , en donde  $\lambda$  representa el número promedio de éxitos por unidad de tiempo:

$$p_x(x) = \frac{(\lambda t)^x}{x!} e^{-\lambda t}$$

Nótese que si se hace  $x=0$  se obtiene la probabilidad de que en el período de tiempo  $\lambda t$  no se obtenga algún éxito. Su complemento es la probabilidad de que en el período de tiempo  $\lambda t$  se obtenga al menos un éxito o llegada:

$$F(t) = p(u \leq t) = 1 - e^{-\lambda t} \quad (6.4)$$

A la expresión anterior se le conoce como la función de probabilidad acumulada exponencial negativa. La derivada de esta función proporciona la función de densidad exponencial negativa, la cual está dada por la expresión:

$$f(t) = \lambda e^{-\lambda t} \quad t > 0 \quad (6.5)$$

Su función generatriz de momentos sería:

$$FGM_t(s) = \int_0^{\infty} e^{st} \lambda e^{-\lambda t} dt = \lambda \int_0^{\infty} e^{-(\lambda-s)t} dt$$

$$FGM_t(s) = -\frac{\lambda}{\lambda-s} \left[ e^{-(\lambda-s)t} \right]_0^{\infty} = \frac{\lambda}{\lambda-s}$$

$$FGM_t(s) = \frac{\lambda}{\lambda-s} \quad (6.6)$$

Su función característica sería:

$$FC_t(s) = \int_0^{\infty} e^{ist} \lambda e^{-\lambda t} dt = \lambda \int_0^{\infty} e^{-(\lambda-is)t} dt$$

$$FC_t(s) = -\frac{\lambda}{\lambda-is} \left[ e^{-(\lambda-is)t} \right]_0^{\infty} = \frac{\lambda}{\lambda-is} \quad (6.7)$$

$$FC_t(s) = -\frac{\lambda}{\lambda-is}$$

Su media sería:

$$\mu_t = \left[ \frac{d FGM_t(s)}{ds} \right]_{s=0} = \left[ \frac{\lambda}{(\lambda-s)^2} \right]_{s=0} = \frac{1}{\lambda} \quad (6.8)$$

Su varianza sería:

$$\sigma_t^2 = \left[ \frac{d^2 FGM_t(s)}{ds^2} \right]_{s=0} - \mu_t^2$$

$$\sigma_t^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \quad (6.9)$$

Su coeficiente de asimetría:

$$\mu_3 = \left[ \frac{d^3 FGM(s)}{ds^3} \right]_{s=0} - 3\mu_x \left[ \frac{d^2 FGM(s)}{ds^2} \right]_{s=0} + 3\mu_x^3 - \mu_x^3$$

$$\mu_3 = \frac{6}{\lambda^3} - 3 \frac{1}{\lambda} \frac{2}{\lambda^2} + \frac{2}{\lambda^3} = \frac{2}{\lambda^3}$$

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2/\lambda^3}{1/\lambda^3} = 2 \quad (6.10)$$

Su coeficiente de curtosis:

$$\mu_4 = \left[ \frac{d^4 FGM(s)}{ds^4} \right]_{s=0} - 4\mu_x \left[ \frac{d^3 FGM(s)}{ds^3} \right]_{s=0} + 6\mu_x^2 \left[ \frac{d^2 FGM(s)}{ds^2} \right]_{s=0} - 4\mu_x^4 + \mu_x^4$$

$$\mu_4 = \frac{24}{\lambda^4} - 4 \left( \frac{1}{\lambda} \right) \left( \frac{6}{\lambda^3} \right) + 6 \left( \frac{1}{\lambda^2} \right) \left( \frac{2}{\lambda^2} \right) - \frac{3}{\lambda^4} = \frac{9}{\lambda^4}$$

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{9/\lambda^4}{1/\lambda^4} - 3 = 6 \quad (6.11)$$

Para el cálculo de probabilidades con la distribución exponencial negativa, utilizando Excel se emplea el siguiente comando:

$$\text{DISTR.EXP.N}(t, \lambda t, 1) \quad (6.12)$$

En donde  $t$  es el tiempo transcurrido desde el inicio hasta el tiempo  $t$ , el parámetro  $\lambda\tau$  es la tasa media de ocurrencia o de éxito del evento hasta el tiempo  $t$ , acumulado=verdadero o 1 si se pretende calcular la función de probabilidad acumulada desde el inicio hasta  $t$ ; acumulado=falso o 0 si se pretende calcular el valor de la función de densidad de probabilidad en el tiempo  $t$ . De tal manera que:

$$p(t=3; \lambda\tau=2) = \text{DISTR.EXP.N}(3,2,0) = 0.004958$$

Esta no es una probabilidad, es el valor de la función de densidad en  $t= 3$ .

$$p(t \leq 3; \lambda\tau = 2) = \text{DISTR.EXP.N}(3,2,1) = 0.997521$$

$$p(t > 3; \lambda\tau = 2) = 1 - \text{DISTR.EXP.N}(3,2,1) = 0.002479$$

$$p(4 < t \leq 6; \lambda t = 2) = \text{DISTR.EXP.N}(6,2,1) - \text{DISTR.EXP.N}(4,2,1) = 0.000329$$

$$p(4 < t \leq 6; \lambda t = 2) = \text{DISTR.EXP.N}(6,2,1) - \text{DISTR.EXP.N}(4,2,1) = 0.000329$$

Para hacer el cálculo de probabilidades con la distribución exponencial negativa, utilizando Minitab, se realizan los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en Exponential, como se muestra en la figura 6.1.

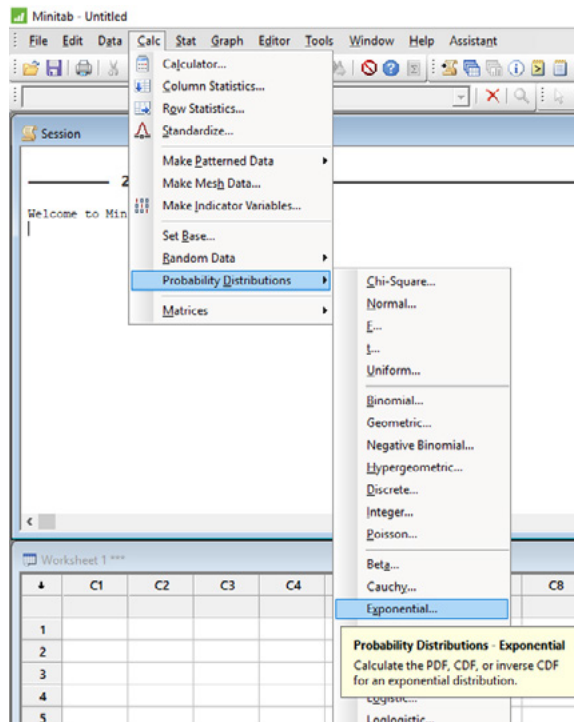
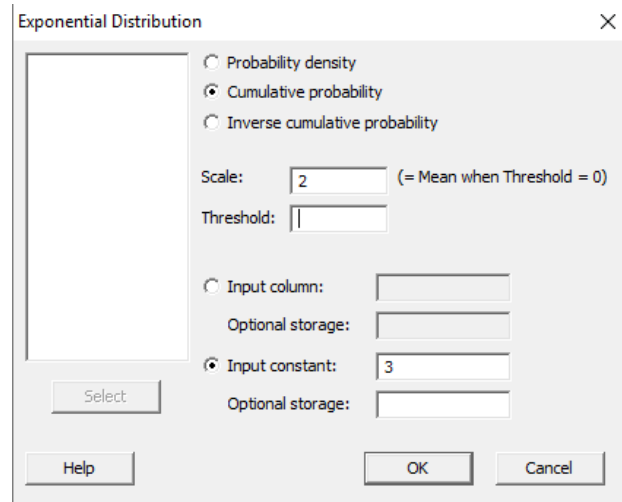


FIGURA 6.1.

2. Aparece la pantalla que se muestra en la figura 6.2, en la cual, para calcular la probabilidad  $p(t \leq 3, \lambda\tau = 2)$  se señala si se trata de una probabilidad en un punto marcando Probability, en Scale se indica el valor de  $\lambda\tau = 2.0$ , en Input constant se indica el valor de  $t = 3$  para el cual se desea calcular la probabilidad y se da un click en Ok obteniéndose el resultado ya mostrado anteriormente.

FIGURA 6.2.



En el lenguaje de desarrollo R, para el caso de la distribución Exponencial Negativa, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo de la función de densidad de probabilidad exactamente en un punto:

$$\text{dexp}(t, \text{rate}=\lambda\tau, \text{log} = \text{FALSE}) \quad (6.13)$$

Para el cálculo de la probabilidad acumulada hasta un punto dado.

$$\text{pexp}(t, \text{rate}=\lambda\tau, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) \quad (6.14)$$

En donde  $t$  es el tiempo entre éxitos, llegadas u ocurrencias;  $\lambda\tau$  es el número de éxitos, llegadas u ocurrencias en promedio en el tiempo  $\lambda\tau$ ;  $\text{log}$  es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma  $\text{FALSE}$ .  $\text{lower.tail}$  es un valor lógico; si se le asigna un valor verdadero, el valor que arroja es la probabilidad acumulada a la izquierda de  $x$ , de lo contrario, si el valor que toma es falso, el valor que arroja es la probabilidad a la derecha del valor de  $t$ .

Por ejemplo,

$$p(t=3; \lambda\tau=2) = \text{dexp}(3, \text{rate}=2, \text{log}=\text{FALSE}) = 0.004958$$

Esta no es una probabilidad, es el valor de la función de densidad en  $t = 3$ .

$$p(t \leq 3; \lambda\tau=2) = \text{pexp}(3, \text{rate}=2, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.997521$$

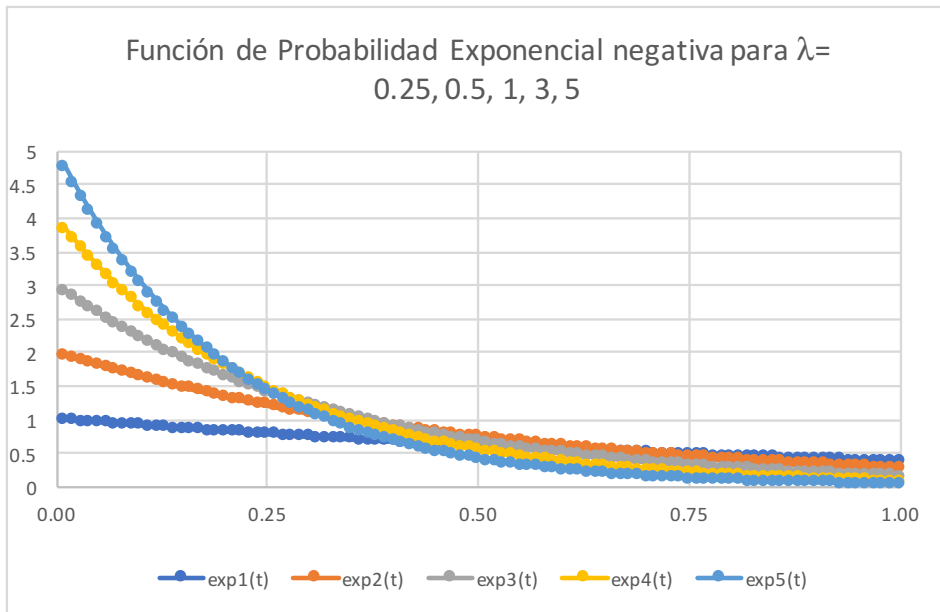
$$p(t > 3; \lambda\tau=2) = \text{pexp}(3, \text{rate}=2, \text{lower.tail}=\text{FALSE}, \text{log.p}=\text{FALSE}) = 0.002479$$

$$p(4 < t \leq 6; \lambda\tau=2) = \text{pexp}(6, \text{rate}=2, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pexp}(4, \text{rate}=2, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.000329$$

$$p(4 <= t <= 6; \lambda\tau=2) = \text{pexp}(6, \text{rate}=2, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pexp}(4, \text{rate}=2, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.000329$$

En la figura 6.3 se muestran algunas gráficas de la distribución exponencial negativa para diferentes valores de  $\lambda\tau$ .

**FIGURA 6.3**



### 6.3.1. Propiedad de Olvido o de Falta de Memoria de la Distribución Exponencial Negativa

La distribución exponencial tiene una muy interesante propiedad, la cual puede ser denominada como “propiedad de olvido o falta de memoria” para variables aleatorias continuas. Recuérdese el concepto de probabilidad condicional:

$$p(T > t+s \mid T > t) = \frac{p(T > t+s)}{p(T > t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s}$$

Lo que implica que:

$$p(T > t+s \mid T > t) = p(T > s) \quad (6.15)$$

Lo anterior implica que:

$$p(T > t+s) = p(T > t) p(T > s) \quad (6.16)$$

Lo que significa que si  $T$  representa el tiempo de vida de un dispositivo con distribución exponencial negativa y este sigue funcionando en el tiempo  $t$ , entonces su tiempo de vida restante tiene la misma distribución exponencial negativa.

#### Ejercicio 6.3

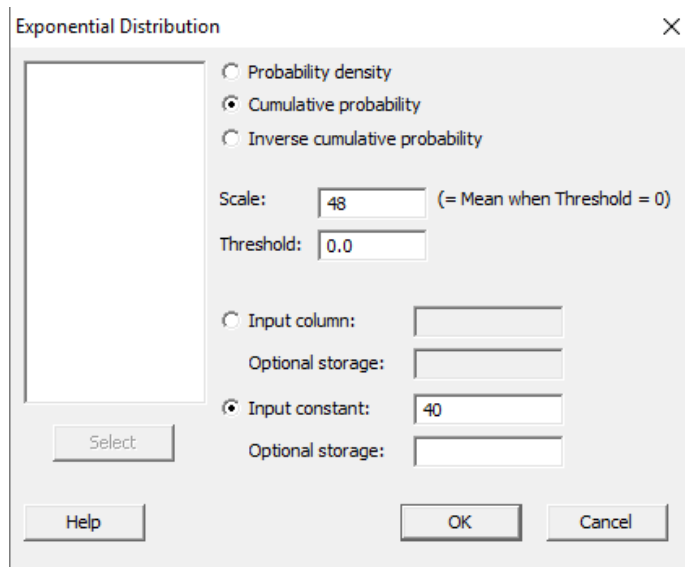
Una compañía que se dedica a reparar instrumentos musicales se tarda en reparar un instrumento en promedio 48 horas. Por estudios previos, la compañía sabe que el tiempo de reparación se comporta como una distribución exponencial negativa. Calcule:

- La probabilidad de que se tarde menos de 40 horas en reparar un instrumento.

Con Minitab, en el menú principal se ingresa a Calc, luego Probability Distributions, posteriormente Exponential y en la pantalla que se obtiene, se teclean los datos mostrados en la figura y se obtiene el siguiente resultado:

$$p(t \leq 40, \lambda\tau = 48) = 0.5654$$





- ii. La probabilidad de que se tarde más de 50 horas en reparar un instrumento.

Con Excel:

$$p(t > 50, \lambda\tau = 1/48) = 1 - \text{DISTR.EXP.N}(50, 1/48, 1) = 0.35287$$

- iii. La probabilidad de que se tarde entre 40 y 60 horas en reparar un instrumento.

Con R:

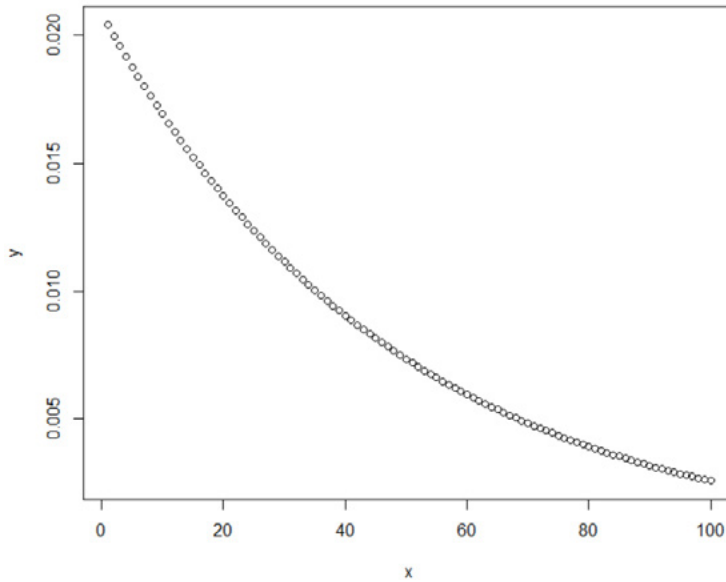
$$p(40 \leq t \leq 60, \lambda\tau = 1/48) = \text{pexp}(60, \text{rate} = 1/48, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) - \text{pexp}(40, \text{rate} = 1/48, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) = 0.148093$$

- iv. Grafique la función de probabilidad.

Con R, se teclean las siguientes líneas:

```
x<-c(1:100)
y=dexp(x,rate=1/48)
plot(x,y)
```

Se obtiene la siguiente gráfica:



- v. Calcule el primer cuartil, el tercer cuartil y el intervalo intercuartil.

Se sabe que:

$$p(t \leq q) = 1 - e^{-\lambda q}$$

$$e^{-\lambda q} = 1 - p(t \leq q)$$

$$e^{\lambda q} = \frac{1}{1 - p(t \leq q)}$$

$$q = \frac{1}{\lambda} \ln \left( \frac{1}{1 - p(t \leq q)} \right)$$

Por lo que los cuartiles están dados por:

$$q_3 = 48 \ln \left( \frac{1}{1 - 0.75} \right) = 66.5421$$

$$q_1 = 48 \ln \left( \frac{1}{1 - 0.25} \right) = 13.8087$$

El intervalo intercuartil sería=  $66.5421 - 13.8087 = 52.7334$ .

vi. Genere diez números aleatorios con dicha distribución.

Con R, utilizando el comando:

```
rexp(10, rate=1/48)
```

Se obtienen los valores 118.45, 1.61, 83.34, 10.49, 112.65, 43.83, 10.65, 4.59, 24.8, 56.11.

vii. Suponga que reparar un producto cuesta  $k$  pesos; pero si lo repara antes de 30 horas, la empresa recibe una bonificación adicional del 15%; si lo entrega entre 30 y 48 horas solo recibe el pago por reparación, y si lo entrega después de 48 horas la empresa recibe una multa de 25% del costo de reparación; obtenga la ganancia promedio de la empresa por cada instrumento reparado.

$$\text{Ganancia\_Promedio} = E\{\text{Ganancia}\} = (1+0.15)kp(t \leq 30) + kp(30 \leq t \leq 48) + (1-0.25)kp(t > 48)$$

$$\text{Ganancia\_Promedio} = 1.15k(0.46474) + k(0.16738) + 0.75k(0.36788)$$

$$\text{Ganancia\_Promedio} = 0.97774k$$

#### Ejercicio 6.4

(Problema 6–15, página 172, William W. Hines et al, *Probabilidad y Estadística para Ingeniería*, Cuarta Edición, Grupo Editorial Patria, 2013). Se están considerando dos procesos de manufactura. El costo por unidad para el proceso I es  $c$  pesos, en tanto que para el proceso II es  $3c$  pesos. Los productos fabricados en cada proceso presentan tiempos de falla con funciones de probabilidad exponenciales con tiempos medios de falla de 25 y 35 horas respectivamente. Si un producto falla antes de 15 horas a un costo de  $z$  pesos, ¿qué proceso recomendaría usted?

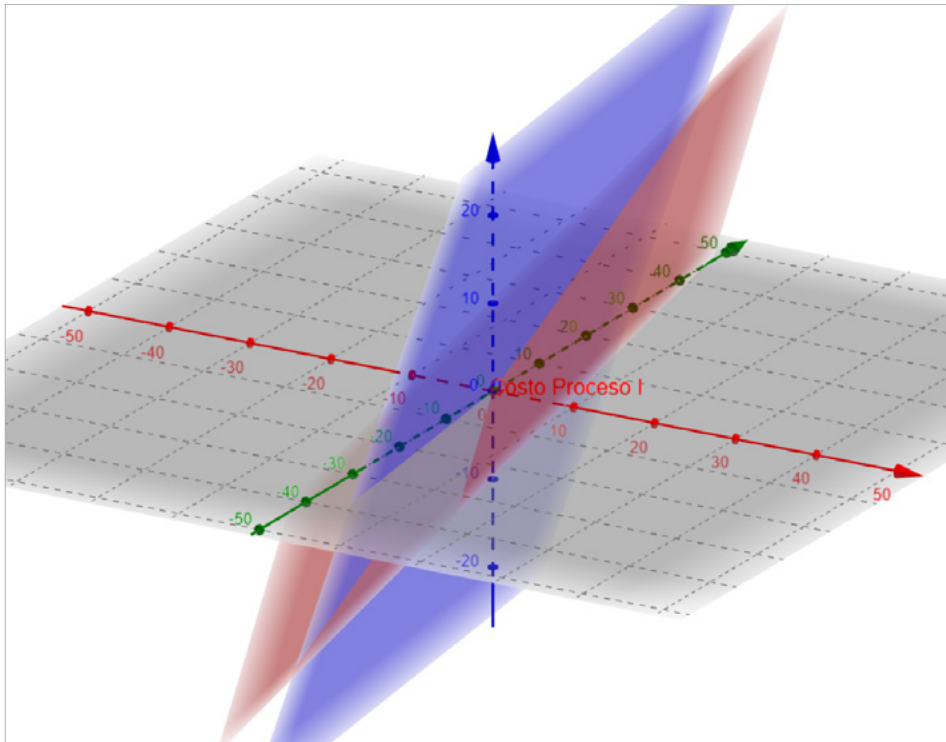
$$\text{Costo por unidad proceso I} = (c+z)p_I(t < 15) + cp(t > 15) = c + zp_I(t < 15)$$

$$\text{Costo por unidad proceso I} = c + 0.451188z$$

$$\text{Costo por unidad proceso II} = (3c+z)p_{II}(t < 15) + zp_{II}(t > 15)$$

$$\text{Costo por unidad proceso II} = 3c + 0.348561z$$

Las ecuaciones obtenidas corresponden a planos y se graficaron en el espacio de tres dimensiones, utilizando para ello Geogebra, lo cual se muestra a continuación:



Como se puede apreciar el costo del proceso I (plano en rojo) está por abajo del costo del proceso II (plano en azul) a partir de la intersección entre ambos planos, es decir, si  $c > 0.051313z$ .

**Ejercicio 6.5.** El tiempo de vida de una lámpara de neón presenta distribución exponencial negativa con una media de 2000 horas de uso. Dicha lámpara ya ha durado 2000 horas de uso. ¿Cuál es la probabilidad de que falle antes de 3000 horas?

Por la propiedad de olvido o pérdida de memoria:

$$p(T \geq 3000 | T > 2000) = p(T \geq 1000)$$

$$p(T < 1000) = 1 - e^{-\frac{1000}{2000}} = 0.393469$$

### 6.3.2. Distribución de probabilidad exponencial negativa desplazada

En la expresión (6.5) la función de probabilidad exponencial negativa está definida a partir de  $t > 0$ . Sin embargo, dicha distribución puede iniciar en cualquier valor  $a$  y su función de probabilidad estaría definida por la siguiente expresión:

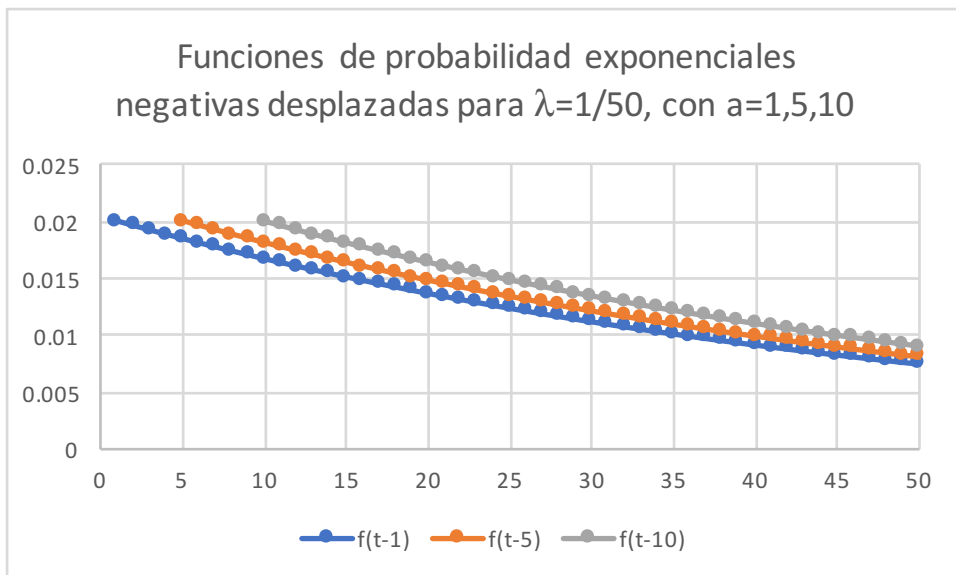
$$f(t) = \lambda e^{-\lambda(t-a)} \quad t > a \quad (6.17)$$

Su función de probabilidad acumulada sería:

$$F(t) = 1 - e^{-\lambda(t-a)} \quad t > a \quad (6.18)$$

En la figura 6.4 se muestran ejemplos de funciones de probabilidad exponenciales negativas desplazadas para una media  $1/\lambda = 10$ .

**FIGURA 6.4**



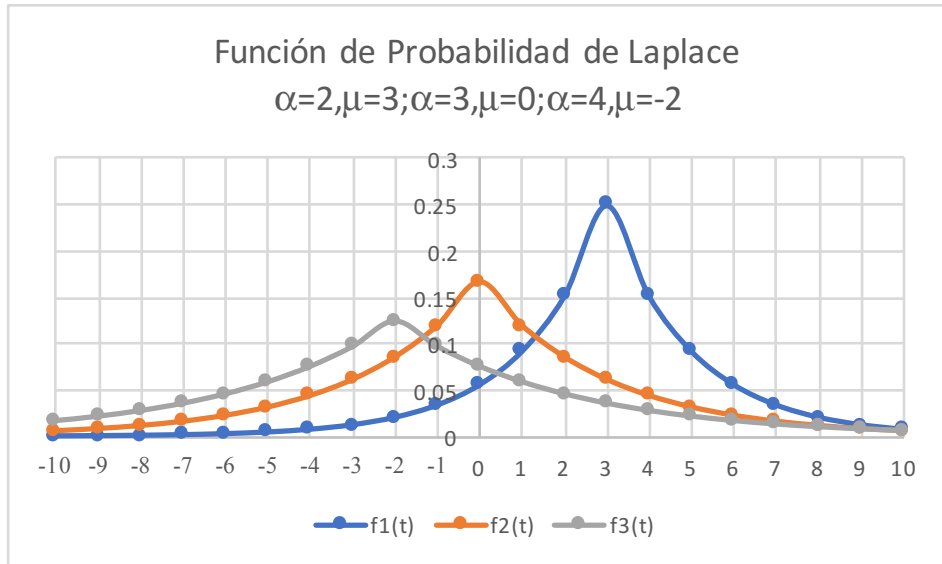
### 6.3.3. Función de Probabilidad de Laplace o Función de Probabilidad Exponencial de doble lado

La función de probabilidad de Laplace, también conocida como función de probabilidad exponencial de doble lado está dada por la expresión:

$$f(x) = \frac{1}{2\alpha} e^{-\frac{|x-\mu|}{\alpha}} \quad x \in \mathfrak{R} \quad (6.19)$$

Donde  $\mu$  es la media y  $\alpha > 0$  es un parámetro de escala que determina la razón de decaimiento de la función; su varianza es  $\sigma_x^2 = 2\alpha^2$ . Algunos ejemplos de funciones de Laplace se muestran en la figura 6.5.

FIGURA 6.5.



Aunque la figura no lo refleja, estas funciones tienen un pico en su media y no son derivables en este punto.

## 6.4. Función de Probabilidad Gamma

La distribución Gamma es una función de probabilidad con dos parámetros: de escala  $\lambda=1/\theta = \beta$  y de forma  $r=k = \alpha$ , definida por la expresión:

$$g_a(t; r, \lambda) = \lambda e^{-\lambda t} \frac{(\lambda t)^{r-1}}{\Gamma(r)} = \frac{1}{\Gamma(k)\theta^k} (\lambda t)^{k-1} e^{-\frac{t}{\theta}} \quad t > 0 \quad (6.20)$$

La cual también se define de la siguiente forma:

$$g_a(t; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} \quad t > 0 \quad (6.21)$$

La parametrización con  $\lambda$  y  $r$  es más común en probabilidad, pero en las aplicaciones es más común usar  $k$  y  $\theta$  en econometría y  $\alpha$  y  $\beta$  en estadística bayesiana.

En donde, a la función:

$$\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx \quad (6.22)$$

Se le conoce como Función Gamma o Función Factorial Generalizada. Esta función cumple tres propiedades muy importantes las cuales se demuestran utilizando la integración por partes y se muestran a continuación:

$$\Gamma(r+1) = \int_0^\infty x^r e^{-x} dx = r\Gamma(r) \quad (6.23)$$

$$\forall n \in \mathbb{N} \quad \Gamma(n+1) = n! \quad (6.24)$$

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty x^{-1/2} e^{-x} dx = \sqrt{\pi} \quad (6.25)$$

Como se puede apreciar en la expresión (6.24) se encuentra la razón de que la Función Gamma también reciba el nombre de Función Factorial Generalizada.

En la figura 2.6 se ilustró la forma que tiene la Función Gamma.

La Función de Probabilidad Acumulada Gamma está representada de la siguiente forma:

$$G_a(t; r, \lambda) = \frac{\lambda}{\Gamma(r)} \int_0^\infty e^{-\lambda t} (\lambda t)^{r-1} dt = \frac{\gamma(r, \lambda t)}{\Gamma(r)} \quad (6.26)$$

En donde a la función:

$$\gamma(r, \lambda t) = \int_0^{\lambda t} v^{r-1} e^{-v} dv \quad (6.27)$$

Se le conoce como Función Gamma Incompleta Inferior.

La Función Generatriz de Momentos de la distribución Gamma está dada por:

$$FGM_t(s) = E\{e^{st}\} = \int_0^{\infty} e^{st} \lambda e^{-\lambda t} \frac{(\lambda t)^{r-1}}{\Gamma(r)} dt = \frac{\lambda^r}{\Gamma(r)} \int_0^{\infty} e^{-(\lambda-s)t} t^{r-1} dt$$

Si se hace:

$$u = (\lambda-s)t \rightarrow t = \frac{u}{(\lambda-s)} \rightarrow du = (\lambda-s) dt$$

$$FGM_t(s) = \frac{\lambda^r}{\Gamma(r)} \int_0^{\infty} e^{-u} \left( \frac{u}{(\lambda-s)} \right)^{r-1} \frac{du}{(\lambda-s)} = \frac{\lambda^r}{\Gamma(r)(\lambda-s)^r} \int_0^{\infty} e^{-u} u^{r-1} du$$

$$FGM_t(s) = \frac{\lambda^r}{\Gamma(r)(\lambda-s)^r} \Gamma(r) = \frac{\lambda^r}{(\lambda-s)^r}$$

$$FGM_t(s) = \left( 1 - \frac{s}{\lambda} \right)^{-r} = (1-\theta s)^{-k} = \left( 1 - \frac{s}{\beta} \right)^{-\alpha} \quad (6.28)$$

Su función característica está dada por la expresión:

$$FC_t(s) = \left( 1 - \frac{is}{\lambda} \right)^{-r} = (1-i\theta s)^{-k} = \left( 1 - \frac{is}{\beta} \right)^{-\alpha} \quad (6.29)$$

La media de la distribución Gamma está dada por:

$$\mu_t = \left[ \frac{dFGM_t(s)}{ds} \right]_{s=0} = \frac{r}{\lambda} \left( 1 - \frac{s}{\lambda} \right)^{-r-1}_{s=0}$$

$$\mu_t = \frac{r}{\lambda} = k\theta = \frac{\alpha}{\beta} \quad (6.30)$$



Su varianza:

$$\sigma_t^2 = \left[ \frac{d^2 FGM_t(s)}{ds^2} \right]_{s=0} - \mu_t^2 = \frac{r(r+1)}{\lambda^2} \left( 1 - \frac{s}{\lambda} \right)_{s=0}^{-r-2} - \frac{r^2}{\lambda^2}$$

$$\sigma_t^2 = \frac{r}{\lambda^2} = k\theta^2 = \frac{\alpha}{\beta^2} \quad (6.31)$$

Su coeficiente de asimetría:

$$\mu_3 = \left[ \frac{d^3 FGM_t(s)}{ds^3} \right]_{s=0} - 3\mu_t \left[ \frac{d^2 FGM_t(s)}{ds^2} \right]_{s=0} + 2\mu_t^3$$

$$\mu_3 = \frac{r(r+1)(r+2)}{\lambda^3} - 3 \frac{r}{\lambda} \frac{r(r+1)}{\lambda^2} + 2 \frac{r^3}{\lambda^3} = \frac{2r}{\lambda^3}$$

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2r/\lambda^3}{(r/\lambda^2)^{3/2}} = \frac{2}{\sqrt{r}} = \frac{2}{\sqrt{k}} = \frac{2}{\sqrt{\alpha}} \quad (6.32)$$

Su coeficiente de curtosis:

$$\mu_4 = \left[ \frac{d^4 FGM_t(s)}{ds^4} \right]_{s=0} - 4\mu_t \left[ \frac{d^3 FGM_t(s)}{ds^3} \right]_{s=0} + 6\mu_t^2 \left[ \frac{d^2 FGM_t(s)}{ds^2} \right]_{s=0} - 3\mu_t^4$$

$$\mu_4 = \frac{r(r+1)(r+2)(r+3)}{\lambda^4} - 4 \frac{r}{\lambda} \frac{r(r+1)(r+2)}{\lambda^3} + 6 \left( \frac{r}{\lambda} \right)^2 \left( \frac{r(r+1)}{\lambda^2} \right) - 3 \left( \frac{r}{\lambda} \right)^4$$

$$\mu_4 = \frac{3r^2 + 6r}{\lambda^4}$$

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{(3r^2 + 6r)/\lambda^4}{r^2/\lambda^4} - 3 = \frac{6}{r} = \frac{6}{k} = \frac{6}{\alpha} \quad (6.33)$$

Para el cálculo de probabilidades con la distribución gamma, al utilizar Excel se emplea el siguiente comando:

$$\text{DISTR.GAMMA.N}(t, \text{alfa}, \text{beta}, 1) \quad (6.34)$$

En donde  $t$  es el tiempo transcurrido desde el inicio hasta el tiempo  $t$ ; el parámetro alfa =  $r = k$  es el parámetro de forma; el parámetro beta =  $\lambda = 1/\theta$  es el parámetro de escala; acumulado = verdadero o 1 si se pretende calcular la función

de probabilidad acumulada desde el inicio hasta  $t$ ; acumulado= falso o 0 si se pretende calcular el valor de la función de densidad de probabilidad en el tiempo  $t$ . De tal manera que:

$$p(t=3; \alpha=2, \beta=0.5) = \text{DISTR.GAMMA.N}(3,2,0.5,0) = 0.029745$$

Esta no es una probabilidad, es el valor de la función de densidad en  $t=3$ .

$$p(t \leq 3; \alpha=2, \beta=0.5) = \text{DISTR.GAMMA.N}(3,2,0.5,1) = 0.982649$$

$$p(t > 3; \alpha=2, \beta=0.5) = 1 - \text{DISTR.GAMMA.N}(3,2,0.5,1) = 0.017351$$

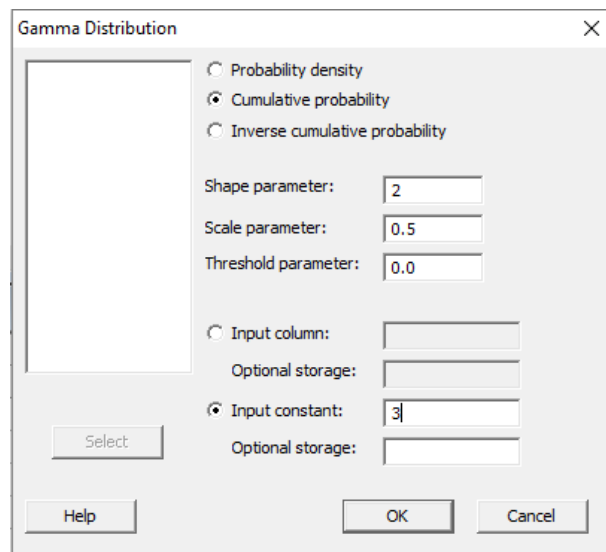
$$p(4 < t \leq 6, \alpha=2, \beta=0.5) = \text{DISTR.GAMMA.N}(6,2,0.5,1) - \text{DISTR.GAMMA.N}(4,2,0.5,1) = 0.002939$$

$$p(4 \leq t < 6, \alpha=2, \beta=0.5) = \text{DISTR.GAMMA.N}(6,2,0.5,1) - \text{DISTR.GAMMA.N}(4,2,0.5,1) = 0.002939$$

Para hacer el cálculo de probabilidades con la distribución Gamma, utilizando Minitab, se realizan los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en Gamma, tal como se ha venido haciendo antes.
2. Aparece la pantalla de la figura 6.6, en la cual se teclean los valores que se muestran en la misma:

**FIGURA 6.6**



Se obtiene lo ya calculado antes:

$$p(t \leq 3; \alpha=2, \beta=0.5) = 0.982649$$

En el lenguaje de desarrollo R, para el caso de la distribución Gamma, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo del valor de la función de densidad de probabilidad exactamente en un punto:

$$\text{dgamma}(t, \text{shape}, \text{scale}, \text{log} = \text{FALSE}) \quad (6.35)$$

Esta no es una probabilidad, es el valor de la función de densidad en  $t$ .

Para el cálculo de la probabilidad acumulada hasta un punto dado.

$$\text{pgamma}(t, \text{shape}, \text{rate}=1, \text{scale}=1/\text{rate}, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) \quad (6.36)$$

En donde  $t$  es el tiempo hasta la falla  $u$  ocurrencia;  $\text{shape}$  el parámetro de forma;  $\text{scale}$  es el parámetro de escala;  $\text{log}$  es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma `FALSE`. `lower.tail` es un valor lógico; si se le asigna un valor verdadero, el valor que arroja es la probabilidad acumulada a la izquierda de  $x$ , de lo contrario, si el valor que toma es falso, el valor que arroja es la probabilidad a la derecha del valor de  $t$ .

Por ejemplo,

$$p(t=3; \alpha=2, \beta=0.5) = \text{dgamma}(3, \text{shape}=2, \text{scale}=0.5, \text{log}=\text{FALSE}) = 0.029745$$

Esta no es una probabilidad, es el valor de la función de densidad en  $t = 3$ .

$$p(t \leq 3; \alpha=2, \beta=0.5) = \text{pgamma}(3, \text{shape}=2, \text{scale}=0.5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.982649$$

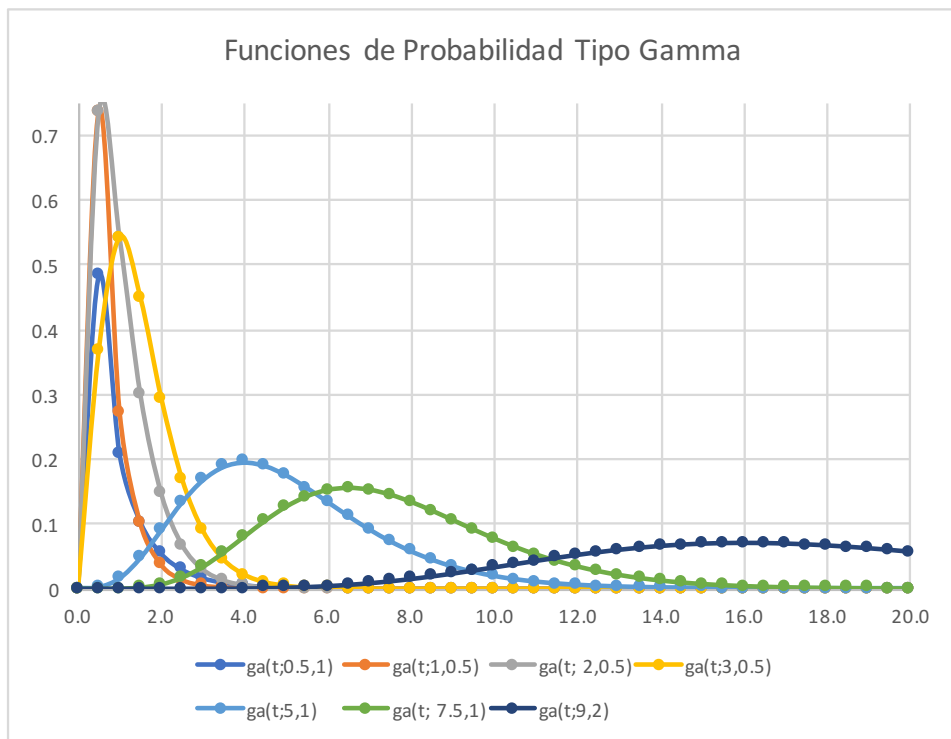
$$p(t > 3; \alpha=2, \beta=0.5) = \text{pgamma}(3, \text{shape}=2, \text{scale}=0.5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.017351$$

$p(4 < t <= 6; \alpha=2, \beta=0.5) = \text{pgamma}(6, \text{shape}=2, \text{scale}=0.5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pgamma}(4, \text{shape}=2, \text{scale}=0.5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.002939$

$p(4 <= t <= 6; \alpha=2, \beta=0.5) = \text{pgamma}(6, \text{shape}=2, \text{scale}=0.5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pgamma}(4, \text{shape}=2, \text{scale}=0.5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.002939$

En la figura 6.7 se muestran diversas funciones de probabilidad tipo Gamma:

FIGURA 6.7



Existen cuatro casos particulares de la distribución Gamma, que por su importancia vale la pena analizar o conocer.

1. Si  $r=1$ , la distribución resultante es una exponencial negativa, ya analizada en el subtema 6.3.

$$\text{gamma}(t; r=1, \lambda) = \lambda e^{-\lambda t} = \text{exp}(t; \lambda) \quad t > 0$$

2. Si  $r = k \in \mathfrak{N}$  es un número natural, la distribución que se origina se denomina Distribución de Erlang, la cual resulta ser una suma de  $k$  variables aleatorias exponenciales negativas estadísticamente independientes, con parámetro  $\lambda$ .

$$gamma(t; k \in \mathfrak{N}, \lambda) = \frac{\lambda^k}{(k-1)!} t^{k-1} e^{-\lambda t} \quad t > 0 \quad (6.37)$$

Esto implica que  $t = t_1 + t_2 + \dots + t_k$  donde cada  $t_i$  tiene distribución exponencial negativa de parámetro  $\lambda$  y es estadísticamente independiente de las otras variables aleatorias que se están sumando.

Dado que se supuso que  $r = k$  es un número natural, la expresión (6.26) puede ser obtenida integrando por partes y obteniendo la expresión:

$$\begin{aligned} G_a(t; k, \lambda) &= \frac{\lambda}{\Gamma(r)} \int_0^\infty e^{-\lambda t} (\lambda t)^{r-1} dt = \\ &= G_a(t; k, \lambda) = 1 - \sum_{j=0}^{j=k-1} e^{-\lambda t} \frac{(\lambda t)^j}{j!} \end{aligned} \quad (6.38)$$

La cual es la suma de los términos de Poisson con media  $\lambda t$ .

3. Si  $\lambda = 1/2$  la distribución que se obtiene se denomina Distribución  $\chi^2$  (se lee Ji Cuadrada) con  $r = k$  grados de libertad. Esta distribución se analizará posteriormente, ya que representa la distribución de probabilidad de la varianza poblacional de una población con distribución normal, la cual también será analizada posteriormente.
4. Si se hace  $y = raíz(t)$ , se obtiene la distribución de Nakagami, cuya función de probabilidad es:

$$f(y) = \frac{2r^r}{\Gamma(r)\alpha^r} y^{2r-1} e^{-\frac{ry^2}{\alpha}} \quad (6.39)$$

Donde  $r \geq 1/2$  y  $\alpha > 0$  son los parámetros de forma y de escala respectivamente. La media de la distribución de Nakagami está dada por la expresión:

$$\mu_y = \sqrt{\frac{r}{\alpha}} \Gamma\left(r + \frac{1}{2}\right) \Gamma\left(\frac{1}{2}\right) \quad (6.40)$$

Su varianza

$$\sigma_y^2 = \alpha \left[ 1 - \frac{1}{r} \Gamma^2\left(r + \frac{1}{2}\right) / \Gamma^2(r) \right] \quad (6.41)$$

## Ejercicio 6.6

En un gimnasio el CrossFit es una de las actividades deportivas más completas, ya que trabaja, de manera variada y no monótona, las diez grandes áreas físicas principales: agilidad, coordinación, equilibrio, flexibilidad, fuerza, potencia, precisión, resistencia cardiorrespiratoria, resistencia muscular y velocidad. Consiste básicamente en una rutina de  $k$  ejercicios diferentes, los cuales se aplican cada uno de ellos durante  $t$  minutos continuos sin parar, y se continua con otro ejercicio diferente durante otros  $t$  minutos de tiempo, y así sucesivamente hasta completar la rutina de  $k$  ejercicios diferentes. Suponga que una rutina de crossfit considera realizar abdominales durante  $t_1$  minutos de tiempo, posteriormente trotar durante  $t_2$  minutos de tiempo, después realizar barra horizontal durante  $t_3$  minutos, y finalmente realizar lagartijas durante  $t_4$  minutos de tiempo. Suponga que el tiempo promedio de cada ejercicio es exponencial con una media de cinco minutos cada uno de ellos.

- a) ¿Cuál es la probabilidad de que una rutina se realice en menos de 18 minutos de tiempo?

Dado que el tiempo total de la rutina es la suma de los tiempos de cada uno de los cuatro ejercicios  $T = t_1 + t_2 + t_3 + t_4$  y cada uno de ellos es exponencial, con parámetro  $\lambda = 1/5$ , entonces, el tiempo total de la rutina tiene distribución tipo Erlang con  $r=4$  y  $\lambda=1/5$ .

$$p(T \leq 18) = \text{pgamma}(18, \text{shape}=4, \text{scale}=5, \text{lower.tail}=\text{TRUE}, \text{log}=\text{FALSE}) = 0.484784$$

- b) ¿Cuál es la probabilidad de que una rutina se realice en más de 18 minutos de tiempo?

$$p(T > 18) = \text{pgamma}(18, \text{shape}=4, \text{scale}=5, \text{lower.tail}=\text{FALSE}, \text{log}=\text{FALSE}) = 0.515216$$

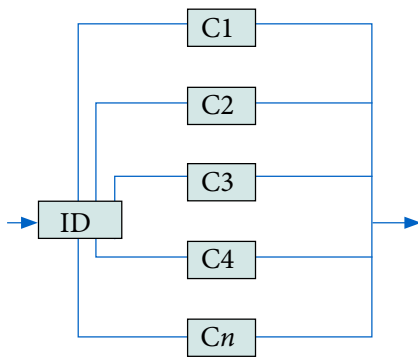
- c) ¿Cuál es la probabilidad de que una rutina se realice entre 18 y 21 minutos de tiempo?

$$p(18 \leq T \leq 21) = \text{pgamma}(21, \text{shape}=4, \text{scale}=5, \text{lower.tail}=\text{TRUE}, \text{log}=\text{FALSE}) - \text{pgamma}(18, \text{shape}=4, \text{scale}=5, \text{lower.tail}=\text{TRUE}, \text{log}=\text{FALSE}) = 0.119813.$$

## Ejercicio 6.7

La confiabilidad es la probabilidad de que un producto y/o servicio se desempeñe adecuadamente, bajo ciertas condiciones dadas durante un cierto período de tiempo. Un sistema en redundancia en espera (stand by system) es un sistema

conformado por varios dispositivos del mismo tipo, ordenados en paralelo, como el que se muestra en la figura 6.8. En un sistema en redundancia en espera, el dispositivo uno se encuentra activo y operando. En el momento que el dispositivo uno falla, el interruptor de decisión ID cambia al componente dos en forma inmediata, el cual permanece operando hasta que falla y nuevamente el interruptor ID cambia al dispositivo tres, y así sucesivamente. En un sistema en redundancia en espera se supone que el interruptor de decisión no falla, por lo que el tiempo total de falla del sistema completo está dado como la suma de los tiempos de falla de cada uno de sus componentes. Si se supone que el tiempo de falla de cada dispositivo es exponencial negativo, con el mismo factor de escala, y cada uno de ellos es estadísticamente independiente de los otros, entonces, el tiempo total de falla de un sistema en redundancia en espera está modelado por medio de una distribución tipo Erlang.



**FIGURA 6.8.** Sistema en Redundancia en Espera

**FUENTE:** FIGURA 17-21, página 655. Probabilidad y Estadística para Ingeniería. Hines et al. Cuarta edición. Editorial Patria, 2013.

Suponga que en la planta de Los Berros Estado de México, la cual bombea ocho metros cúbicos de agua por segundo de la Presa Victoria en Valle de Bravo a la Ciudad de México, se encuentran nueve motores en redundancia en espera, con un tiempo de falla promedio de 5 años cada uno. Calcule la probabilidad de que el sistema colapse en menos de cinco años.

$$P(T=t_1+t_2+t_3+t_4+t_5+t_6 < 5) = \text{pgamma}(5, \text{shape}=9, \text{scale}=5, \text{lower.tail} = \text{TRUE}, \text{log}=\text{FALSE}) = 0.000001125.$$

Puede parecer una probabilidad muy baja, pero, ¿se ha puesto a pensar en lo que ocurriría si se interrumpiera el bombeo repentinamente? Cada metro cúbico de agua pesa una tonelada, es decir, un flujo de ocho metros cúbicos serían ocho toneladas de agua por segundo que bajando por acción de la gravedad arrasarían con lo que se encontrara a su paso.

## 6.5. Función de Probabilidad Tipo Beta

La función de probabilidad tipo Beta está dada por la siguiente expresión:

$$f(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} \quad x \in [0,1] \quad (6.42)$$

Donde, a la función  $B(\alpha, \beta)$  se le conoce como función Beta, la cual se define como:

$$B(\alpha, \beta) = \int_0^1 u^{\alpha-1} (1-u)^{\beta-1} du = \frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha+\beta)} \quad (6.43)$$

Su función de probabilidad acumulada está dada por la siguiente expresión:

$$F(x; \alpha, \beta) = \frac{B(x; \alpha, \beta)}{B(\alpha, \beta)} I_x(\alpha, \beta) \quad (6.44)$$

La media de la distribución beta es:

$$\mu_x = \frac{\alpha}{\alpha+\beta} \quad (6.45)$$

Su varianza:

$$\sigma_x^2 = \frac{\alpha\beta}{(\alpha+\beta)^2 (\alpha+\beta+1)} \quad (6.46)$$

Su coeficiente de asimetría:

$$\gamma_1 = \frac{2(\beta-\alpha) \sqrt{\alpha+\beta+1}}{(\alpha+\beta+2) \sqrt{\alpha\beta}} \quad (6.47)$$

Su coeficiente de curtosis:

$$\gamma_2 = \frac{6[(\alpha-\beta)^2 (\alpha+\beta+1) - \alpha\beta(\alpha+\beta+2)]}{\alpha\beta (\alpha+\beta+2) (\alpha+\beta+3)} \quad (6.48)$$

Su función generatriz de momentos:

$$FGM_x(t) = 1 + \sum_{k=1}^{\infty} \left[ \prod_{r=0}^{k-1} \frac{\alpha+r}{\alpha+\beta+r} \right] \frac{t^k}{k!} \quad (6.49)$$



Su función característica:

$$FC_x(t) = 1 + \sum_{k=1}^{\infty} \left[ \prod_{r=0}^{k-1} \frac{\alpha+r}{\alpha+\beta+r} \right] \frac{(it)^k}{k!} \quad (6.50)$$

Para el cálculo de probabilidades con la distribución beta, utilizando Excel se emplea el siguiente comando:

$$\text{DISTR.BETA.N}(x, \text{alfa}, \text{beta}, \text{acumulado}, [a], [b]) \quad (6.51)$$

En donde  $x$  es un valor entre  $a$  y  $b$ ; los parámetros de forma alfa y beta;  $a$  es el límite inferior de definición de  $x$ ,  $b$  es el límite superior de definición de  $x$ ; acumulado= verdadero ó 1 si se pretende calcular la función de probabilidad acumulada hasta  $x$ ; acumulado= falso ó 0 si se pretende calcular el valor de la función de densidad en  $x$ . De tal manera que:

$$p(x=2; \alpha=8, \beta=10, a=1, b=3) = \text{DISTR.BETA.N}(2, 8, 10, 0, 1, 3) = 1.48376$$

Ésta no es una probabilidad es el valor de la función de densidad en  $x=2$ .

$$p(x \leq 2; \alpha=8, \beta=10, a=1, b=3) = \text{DISTR.BETA.N}(2, 8, 10, 1, 1, 3) = 0.68547$$

$$p(x > 2; \alpha=8, \beta=10, a=1, b=3) = 1 - \text{DISTR.BETA.N}(2, 8, 10, 1, 1, 3) = 0.31453$$

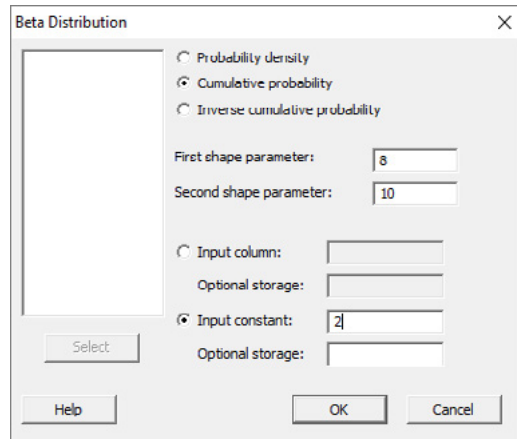
$$p(1 < x \leq 2.5; \alpha=8, \beta=10, a=1, b=3) = \text{DISTR.BETA.N}(2.5, 8, 10, 1, 1, 3) - \text{DISTR.BETA.N}(1, 8, 10, 1, 1, 3) = 0.9969$$

$$p(1 \leq x < 2.5; \alpha=8, \beta=10, a=1, b=3) = \text{DISTR.BETA.N}(2.5, 8, 10, 1, 1, 3) - \text{DISTR.BETA.N}(1, 8, 10, 1, 1, 3) = 0.9969$$

Para hacer el cálculo de probabilidades con la distribución Beta, utilizando Minitab, se realizan los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en Beta, tal como se ha venido haciendo antes.
2. Aparece la pantalla de la figura 6.9, en la cual se teclean los valores que se muestran en la misma:

FIGURA 6.9



Se obtiene lo ya calculado antes:

$$p(x \leq 2; \alpha=8, \beta=10, a=1, b=3) = 0.68547$$

En el lenguaje de desarrollo R, para el caso de la distribución Beta, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo del valor de la función de densidad de probabilidad exactamente en un punto:

$$\text{dbeta}(x, \text{shape1}, \text{shape2}, \text{ncp}=0, \text{log} = \text{FALSE}) \quad (6.52)$$

Esta no es una probabilidad, es el valor de la función de densidad en  $x$ .

Para el cálculo de la probabilidad acumulada hasta un punto dado.

$$\text{pbeta}(x, \text{shape1}, \text{shape2}, \text{ncp}=0, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) \quad (6.53)$$

En donde  $x$  es un valor entre los límites de definición del intervalo de  $x$ ;  $\text{shape1}$  el parámetro de forma 1;  $\text{shape2}$  el parámetro de forma 2;  $\text{ncp}=0$  usa el algoritmo para una distribución no centrada, el cual no es el mismo que se usa si se omite  $\text{ncp}$ ;  $\text{log}$  es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma FALSE.  $\text{lower.tail}$  es un valor lógico; si se le asigna un valor verdadero, el valor que arroja es la probabilidad acumulada a la izquierda de  $x$ , de lo contrario, si el valor que toma es falso, el valor que arroja es la probabilidad a la derecha del valor de  $x$ .

Por ejemplo,

$$p(x=2; \alpha=8, \beta=10) = \text{dbeta}(2, 8, 10, \text{ncp}=0, \text{log}=\text{FALSE}) = 1.48376$$

Esta no es una probabilidad, es el valor de la función de densidad en  $x=2$ .

$$p(x \leq 2; \alpha=8, \beta=10) = \text{pbeta}(2, 8, 10, \text{ncp}=0, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.68547$$

$$p(x > 2; \alpha=8, \beta=10) = 1 - \text{pbeta}(2, 8, 10, \text{ncp}=0, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.31453$$

$$p(1 < x \leq 2.5; \alpha=8, \beta=10) = \text{pbeta}(2.5, 8, 10, \text{ncp}=0, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pbeta}(1, 8, 10, \text{ncp}=0, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.9969$$

$$p(1 \leq x \leq 2.5; \alpha=8, \beta=10, a=1, b=3) = \text{pbeta}(2.5, 8, 10, \text{ncp}=0, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pbeta}(1, 8, 10, \text{ncp}=0, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.9969$$

Cuando el rango finito de la variable aleatoria tipo beta,  $x$ , necesita ser extendido más allá de los límites  $[0, 1]$ , por ejemplo, al intervalo  $[a, b]$ , la función de probabilidad tipo beta puede ser definida mediante la siguiente expresión:

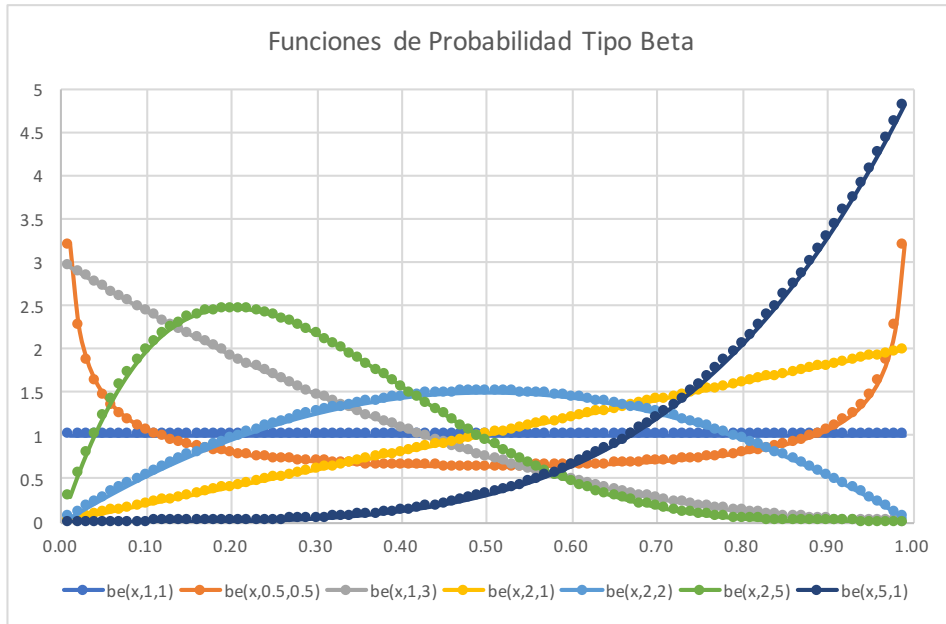
$$f(x) = \frac{1}{B(\alpha, \beta) (b-a)^{\alpha+\beta+1}} (x-a)^{\alpha-1} (b-x)^{\beta-1} \quad x \in [a, b] \quad (6.54)$$

En la figura 6.10 se muestran diversas funciones de probabilidad tipo Beta.

De la figura 6.10 se puede resumir la forma que presenta la distribución tipo Beta, dependiendo de los valores que tomen sus parámetros de forma:

Parámetros	Forma
$\alpha=1, \beta=1$	Uniforme
$\alpha=1, \beta > 2$ (ó $\beta=1, \alpha > 0$ )	Estrictamente convexa
$\alpha=1, 1 < \beta < 2$ (ó $\beta=1, 1 < \alpha < 2$ )	Estrictamente cóncava
$\alpha=1, \beta=2$ (ó $\alpha=2, \beta=1$ )	Línea recta
$\alpha=\beta$	Simétrica alrededor de 0.5

FIGURA 6.10.



Existen cinco casos particulares de la distribución Beta que por su importancia vale la pena analizar o conocer:

1. Si  $\alpha=\beta=1$  aparece la función de probabilidad Uniforme Continua, ya analizada en el ejercicio 3.9 y en el subtema 6.1 y cuya gráfica se muestra en la figura 6.10.
2. Si  $x$  es una variable aleatoria  $\text{beta}(\alpha, \beta)$ , y se hace la transformación  $y=\beta x$  para  $\alpha=1$  y  $\beta \rightarrow \infty$  se obtiene la distribución exponencial negativa, ya analizada en el subtema 6.3.
3. Si  $\alpha=\beta=2$  aparece la función de probabilidad Triangular, ya analizada en el ejercicio 3.10 y en el subtema 6.2.
4. Si  $\alpha=\beta=1/2$  aparece la distribución de probabilidad conocida como Función de Probabilidad Angseno, cuya expresión matemática está dada por:

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}} \quad x \in [0,1] \quad (6.55)$$

Su función de probabilidad acumulada está dada por la expresión:

$$F(x) = \frac{2}{\pi} \operatorname{arcseno}(x) \quad x \in [0,1] \quad (6.56)$$

La media de esta distribución es  $1/2$  y su varianza  $1/6$ .

5. Si  $\beta=1$  se obtiene una forma de la distribución de la función de potencia, cuya función de probabilidad está dada por la expresión:

$$f(x) = \alpha x^{\alpha-1} \quad x \in [0,1] \quad (6.57)$$

Su función de probabilidad acumulada sería:

$$F(x) = x^\alpha \quad x \in [0,1] \quad (6.58)$$

Su media y su varianza:

$$\mu_x = \frac{\alpha}{\alpha+1} \quad (6.59)$$

$$\sigma_x^2 = \frac{\alpha}{(\alpha+1)^2(\alpha+2)} \quad (6.60)$$

### Ejercicio 6.8

Suponga que la fracción de alumnos que acredita la asignatura de Probabilidad en la Facultad de Ingeniería de la UNAM presenta una distribución beta con una media de 0.45 y una desviación estándar de 0.05.

- a) Obtenga los parámetros de forma  $\alpha$  y  $\beta$  de dicha distribución.

$$\mu_x = \frac{\alpha}{\alpha+\beta} = 0.45$$

$$\sigma_x = \sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}} = 0.25 \Rightarrow \sqrt{\frac{0.45\beta}{(\alpha+\beta)(\alpha+\beta+1)}} = 0.25$$

$$\alpha = \frac{0.45}{0.55} \beta$$

$$0.45\beta = (0.25)^2 (\alpha+\beta)(\alpha+\beta+1)$$

$$\alpha = 1.332 \quad \beta = 1.628$$

b) Determine su función de probabilidad y grafíquela

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} \quad x \in [0,1]$$

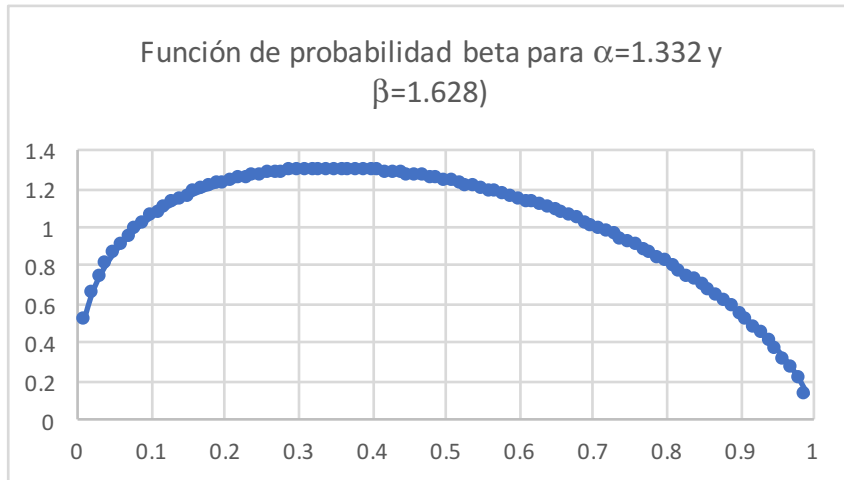
$$\Gamma(\alpha+\beta) = \Gamma(1.332+1.628) = 1.928135379$$

$$\Gamma(\alpha) = \Gamma(1.332) = 0.8931376$$

$$\Gamma(\beta) = \Gamma(1.628) = 0.896973933$$

$$\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} = 2.406796262$$

$$f(x) = 2.406796x^{0.332} (1-x)^{0.628}$$



c) Por petición de la administración de la División de Ciencias Básicas de la Facultad de Ingeniería se solicita elevar la fracción de aprobados a más de 0.70, ¿cuál es la probabilidad de que esto ocurra?

$$p(x>0.7; \alpha=1.332, \beta=1.628) = 1 - \text{DISTR.BETA.N}(0.7, 1.332, 1.628) = 0.194301$$

d) ¿Cuál es la probabilidad de que aprueben entre 0.6 y 0.8 de los alumnos?

$$p(0.6 < x < 0.8; \alpha=1.332, \beta=1.628) = \text{DISTR.BETA.N}(0.8, 1.332, 1.628) - \text{DISTR.BETA.N}(0.6, 1.332, 1.628) = 0.199053$$

e) Obtenga el intervalo interdecil de esta distribución.

Utilizando R:

$q_3 = \text{qbeta}(0.90, 1.332, 1.628, \text{ncp}=0, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.803646$

$q_1 = \text{qbeta}(0.10, 1.332, 1.628, \text{ncp}=0, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.117664$

Intervalo Interdecil =  $0.803646 - 0.117664 = 0.685982$

f) Genere 10 números aleatorios con esta distribución.

Utilizando r con el comando:

`rbeta(10, 1.332, 1.628, ncp=0)`

Se generan los siguientes números:

0.1325152    0.5001743

0.1776874    0.8621432

0.1342789    0.7647708

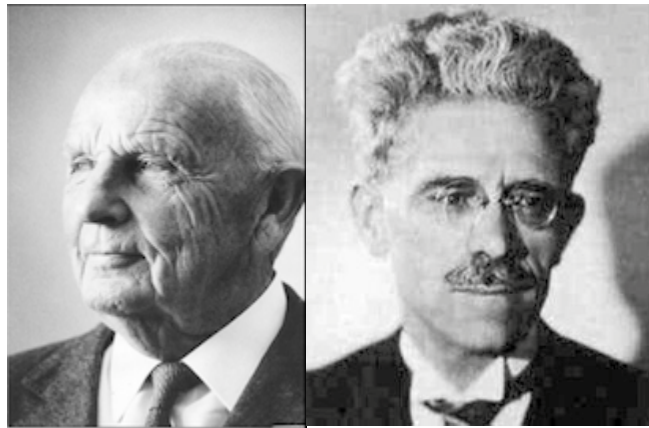
0.5229006    0.5439200

0.6674203    0.6455420

## 6.6. Función de Probabilidad de Weibull

La función de probabilidad tipo Weibull tiene un papel muy preponderante como modelo de confiabilidad para establecer el tiempo de falla en componentes y sistemas eléctricos y mecánicos.

Recibe su nombre de Waloddi Weibull, 1887-1979 ([https://es.wikipedia.org/wiki/Waloddi\\_Weibull](https://es.wikipedia.org/wiki/Waloddi_Weibull)) que la describió detalladamente en 1951, aunque fue descubierta inicialmente en 1927 por Maurice Fréchet, 1878-1973 ([https://es.wikipedia.org/wiki/Maurice\\_Fr%C3%A9chet](https://es.wikipedia.org/wiki/Maurice_Fr%C3%A9chet)).



**FIGURA 6.11.** Waloddi Weibull y Maurice Fréchet

**NOTAS** (2017 y 2018) [https://es.wikipedia.org/wiki/Waloddi\\_Weibull](https://es.wikipedia.org/wiki/Waloddi_Weibull)  
[https://es.wikipedia.org/wiki/Maurice\\_Fr%C3%A9chet](https://es.wikipedia.org/wiki/Maurice_Fr%C3%A9chet)

Su función de probabilidad está dada por la expresión:

$$f(x) = \frac{\beta}{\delta} \left( \frac{x-\gamma}{\delta} \right)^{\beta-1} e^{-\left(\frac{x-\gamma}{\delta}\right)^\beta} \quad x \geq \gamma \quad (6.61)$$

La cual presenta tres parámetros:  $\gamma \in \mathfrak{R}$  parámetro de localización, a partir del cual inicia su intervalo de definición,  $\delta > 0$  parámetro de escala y  $\beta > 0$  parámetro de forma.

Cabe señalar que una parte considerable de las fuentes de referencia usa expresiones aparentemente diferentes de la señalada en (6.61). Para mostrar que se trata de la misma expresión, solo se hará un cambio de variable  $t = x - \gamma$ , con lo cual se expresan las siguientes definiciones:

$$f(t) = \frac{k}{\lambda} \left( \frac{t}{\lambda} \right)^{k-1} e^{-\left(\frac{t}{\lambda}\right)^k} \quad t \geq 0 \quad (6.62)$$



$$f(t) = \frac{r}{\alpha} \left(\frac{t}{\alpha}\right)^{r-1} e^{-\left(\frac{t}{\alpha}\right)^r} \quad t \geq 0 \quad (6.63)$$

$$f(t) = \frac{\alpha}{\beta^\alpha} t^{\alpha-1} e^{-\left(\frac{t}{\beta}\right)^\alpha} \quad t \geq 0 \quad (6.64)$$

Como se puede apreciar en las cuatro expresiones anteriores, se describe la misma función de probabilidad, solo se tiene que considerar el cambio de nomenclatura en el nombre de los factores de cada una de ellas, siguiendo el orden en que se presentan las expresiones (6.61), (6.62), (6.63), (6.64):

$\beta = k = r = \alpha$       parámetro de forma  
 $\delta = \lambda = \alpha = \beta$     parámetro de escala  
 $\gamma$                       parámetro de localización

Con relación a la expresión (6.61) su función de probabilidad acumulada está dada por la expresión:

$$F(x) = 1 - e^{-\left(\frac{x-\gamma}{\delta}\right)^\beta} \quad x \geq \gamma \quad (6.65)$$

Su media:

$$\mu_x = \gamma + \delta \Gamma\left(1 + \frac{1}{\beta}\right) \quad (6.66)$$

Su varianza:

$$\sigma_x^2 = \delta^2 \left\{ \Gamma\left(1 + \frac{2}{\beta}\right) - \left[ \Gamma\left(1 + \frac{1}{\beta}\right) \right]^2 \right\} \quad (6.67)$$

Con relación a la expresión (6.61) su Función Generatriz de Momentos es:

$$FGM_x(s) = e^{\gamma s} \sum_{n=0}^{\infty} \frac{s^n \delta^n}{n!} \Gamma\left(1 + \frac{n}{\beta}\right) \quad \beta \geq 1 \quad (6.68)$$

Su función característica:

$$FC_x(s) = e^{\gamma s} \sum_{n=0}^{\infty} \frac{(is)^n \delta^n}{n!} \Gamma\left(1 + \frac{n}{\beta}\right) \quad \beta \geq 1 \quad (6.69)$$

No se citan los coeficientes de asimetría y de curtosis porque sus expresiones no son simples, pero el lector interesado puede consultar [https://es.wikipedia.org/wiki/Distribuci%C3%B3n\\_de\\_Weibull](https://es.wikipedia.org/wiki/Distribuci%C3%B3n_de_Weibull)

Para el cálculo de probabilidades con la distribución Weibull, utilizando Excel, este software define a la función de probabilidad Weibull como se establece en la expresión (6.64),  $\alpha$  corresponde con el parámetro de forma  $\beta$  definido en la expresión (6.61), y  $\beta$  corresponde con el parámetro de escala  $\delta$  de la expresión (6.61) a través del siguiente comando:

$$\text{DISTR.WEIBULL}(t, \text{alfa}, \text{beta}, \text{acumulado}) \quad (6.70)$$

En donde  $t > 0$ ; acumulado= verdadero ó 1 si se pretende calcular la función de probabilidad acumulada hasta  $t$ ; acumulado= falso ó 0 si se pretende calcular el valor de la función de densidad en  $t$ . De tal manera que:

$$p(t=2; \alpha=3, \beta=5, \gamma=0) = \text{DISTR.WEIBULL}(2, 3, 5, 0) = 0.090048$$

Esta no es una probabilidad, es el valor de la función de densidad en  $t=2$ .

$$p(t \leq 2; \alpha=3, \beta=5, \gamma=0) = \text{DISTR.WEIBULL}(2, 3, 5, 1) = 0.061995$$

Observe que:

$$p(t < t_0; \alpha, \beta, \gamma \neq 0) = \text{DISTR.WEIBULL}(t_0 - \gamma, \alpha, \beta, 1)$$

Por ejemplo:

$$p(t \leq 2; \alpha=3, \beta=5, \gamma=1) = \text{DISTR.WEIBULL}(1, 3, 5, 1) = 0.007968$$

$$p(t > 2; \alpha=3, \beta=5, \gamma=1) = 1 - \text{DISTR.WEIBULL}(1, 3, 5, 1) = 0.992032$$

$$p(3 < t \leq 5; \alpha=3, \beta=4, \gamma=1) = \text{DISTR.WEIBULL}(4, 2, 4, 1) - \text{DISTR.WEIBULL}(2, 2, 4, 1) = 0.410921$$

$$p(3 \leq t < 5; \alpha=2, \beta=4, \gamma=1) = \text{DISTR.WEIBULL}(4, 2, 4, 1) - \text{DISTR.WEIBULL}(2, 2, 4, 1) = 0.410921$$

Para hacer el cálculo de probabilidades con la distribución Weibull, utilizando Minitab, la función de probabilidad que se utiliza es la misma que para Excel, expresión (6.64), por lo que se realizan los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en Weibull, tal como se ha venido haciendo antes.
2. Aparece la pantalla de la figura 6.12, en la cual se teclean los valores que se muestran en la misma:

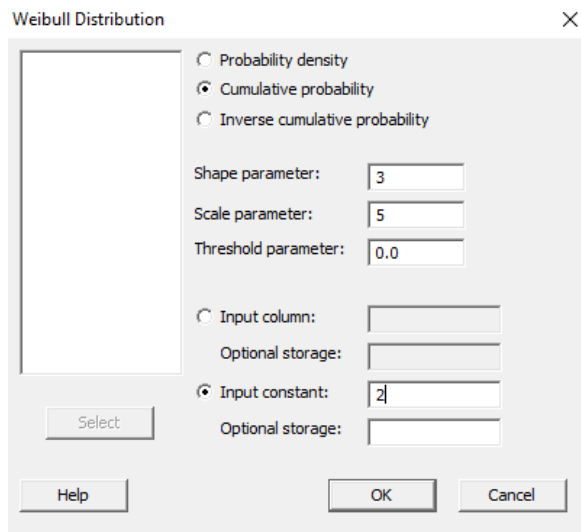


FIGURA 6.12

Se obtiene lo ya calculado antes:

$$p(t \leq 2; \alpha=3, \beta=5, \gamma=0) = 0.061995$$

En el lenguaje de desarrollo R, para el caso de la distribución Weibull, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo del valor de la función de densidad de probabilidad exactamente en un punto:

$$\text{dweibull}(t, \text{shape}, \text{scale}=1, \text{log} = \text{FALSE}) \quad (6.71)$$

Esta no es una probabilidad, es el valor de la función de densidad en  $t$ .

Para el cálculo de la probabilidad acumulada hasta un punto dado.

$$\text{pweibull}(t, \text{shape}, \text{scale}=1, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) \quad (6.72)$$

En donde  $t$  es un valor entre los límites de definición del intervalo de  $t$ ;  $\text{shape}$  el parámetro de forma;  $\text{scale}$  el parámetro de escala;  $\text{log}$  es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma `FALSE`.  $\text{lower.tail}$  es un valor lógico; si se le asigna un valor verdadero, el valor que arroja es la probabilidad acumulada a la izquierda de  $t$ , de lo contrario, si el valor que toma es falso, el valor que arroja es la probabilidad a la derecha del valor de  $t$ .

Por ejemplo,

$$p(t=2; \alpha=3, \beta=5, \gamma=0) = \text{dweibull}(2, 3, \text{scale}=5, \text{log}=\text{FALSE}) = 0.090048$$

Esta no es una probabilidad, es el valor de la función de densidad en  $t = 2$ .

$$p(t \leq 2; \alpha=3, \beta=5, \gamma=0) = \text{pweibull}(2, 3, \text{scale}=5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.061995$$

Observe que:

$$p(t < t_0; \alpha, \beta, \gamma \neq 0) = \text{pweibull}(t_0 - \gamma, \alpha, \text{scale}=\beta, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE})$$

Por ejemplo:

$$p(t \leq 2; \alpha=3, \beta=5, \gamma=1) = \text{pweibull}(1, 3, \text{scale}=5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.007968$$

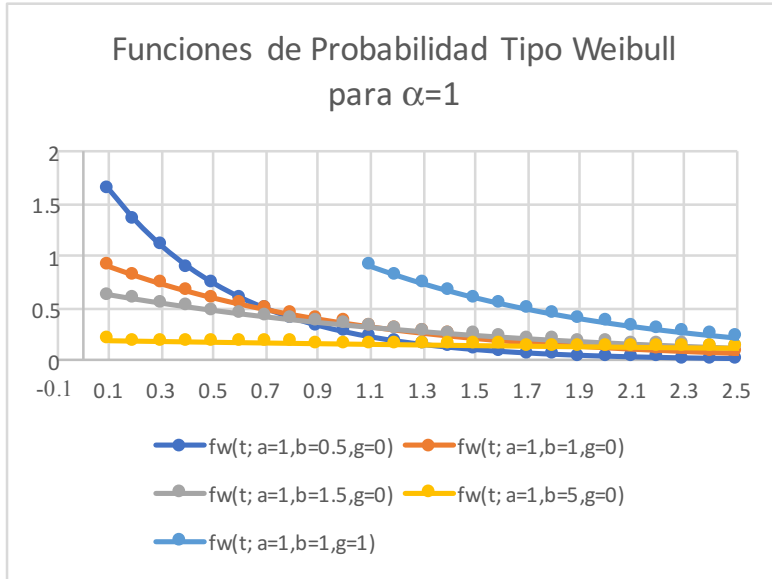
$$p(t > 2; \alpha=3, \beta=5, \gamma=1) = \text{pweibull}(1, 3, \text{scale}=5, \text{lower.tail}=\text{FALSE}, \text{log.p}=\text{FALSE}) = 0.992032$$

$$p(3 < t \leq 5; \alpha=2, \beta=4, \gamma=1) = \text{pweibull}(4, 2, \text{scale}=4, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pweibull}(2, 2, \text{scale}=4, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.410921$$

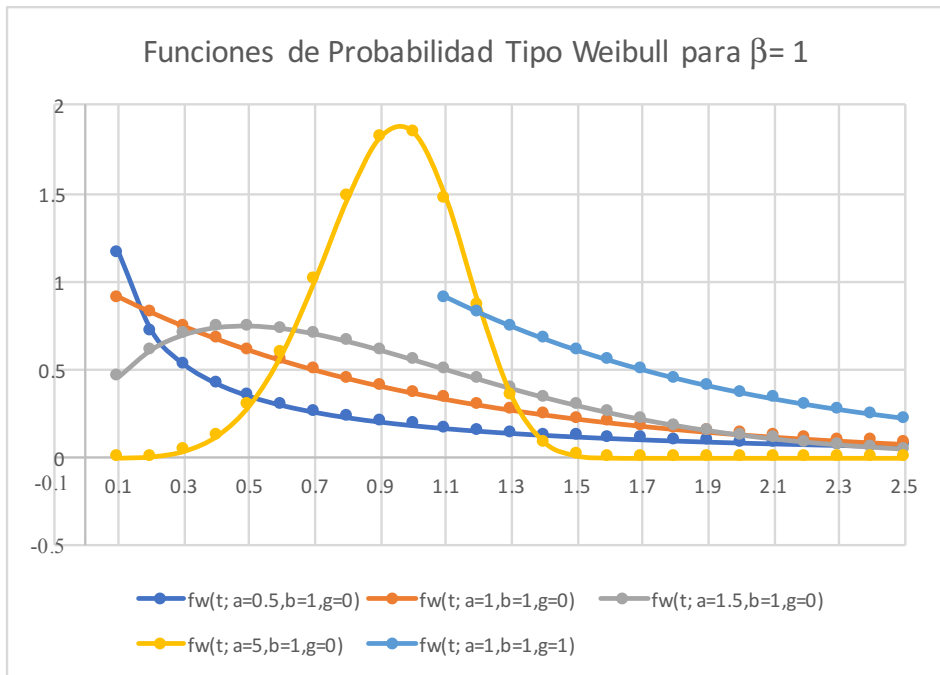
$$p(3 \leq t < 5; \alpha=2, \beta=4, \gamma=1) = \text{pweibull}(4, 2, \text{scale}=4, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pweibull}(2, 2, \text{scale}=4, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.410921$$

En la figura 6.13.a y 6.13.b se muestran diversas funciones de probabilidad tipo Weibull.

**FIGURA 6.13.a.**



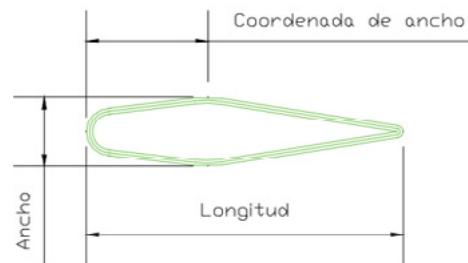
**FIGURA 6.13.b**



## Ejercicio 6.9

**Ejercicio 6.9.** En el artículo “Aplicaciones de la Distribución Weibull en Ingeniería de Confiabilidad”, en las Memorias del Coloquio Mexicano de Economía Matemática y Econometría, los autores Raquel Salazar Moreno, Abraham Rojano Aguilar, Esther Figueroa Hernández y Francisco Pérez Soto (<http://www.eumed.net/libros-gratis/2013/1255/aplicaciones-distribucion-weibull-ingenieria-confiabilidad.pdf>), logran ajustar un modelo Weibull al largo y ancho de una semilla de avena, utilizando el método de máxima verosimilitud para estimar el valor de los parámetros (en el libro de Fundamentos de Estadística se describirá este método). En la figura 6.14 se muestra la geometría aproximada de un grano de avena.

FIGURA 6.14



**NOTAS** Geometría de un grano de avena (sf. 2013), página 156 de las Memorias del Coloquio Mexicano de Economía Matemática y Econometría <http://www.eumed.net/libros-gratis/2013/1255/aplicaciones-distribucion-weibull-ingenieria-confiabilidad.pdf>

Las funciones de probabilidad obtenidas en el artículo antes mencionado fueron las siguientes:

$$f(l) = \frac{5.5474}{8.4388} \left( \frac{l-3.2666}{8.4388} \right)^{4.5474} e^{-\left( \frac{l-1.3339}{1.3368} \right)^{3.9894}}$$

$$f(a) = \frac{3.9894}{1.3368} \left( \frac{a-1.3339}{1.3368} \right)^{2.9894} e^{-\left( \frac{a-1.3339}{1.3368} \right)^{3.9894}}$$

a) Obtenga las medias del largo y del ancho del grano de avena.

$$\mu_l = \gamma_l + \delta_l \Gamma \left( 1 + \frac{1}{\beta_l} \right) = 3.2666 + 8.4388 * \Gamma \left( 1 + \frac{1}{5.5474} \right) = 11.061106$$

$$\mu_a = \gamma_a + \delta_a \Gamma \left( 1 + \frac{1}{\beta_a} \right) = 1.3339 + 1.3368 * \Gamma \left( 1 + \frac{1}{3.9894} \right) = 2.545396$$

b) Calcule las varianzas del largo y del ancho del grano de avena.

$$\sigma_l^2 = \delta_l^2 \left\{ \Gamma\left(1 + \frac{2}{\beta_l}\right) - \left[ \Gamma\left(1 + \frac{1}{\beta_l}\right) \right]^2 \right\} = (8.4388)^2 * \left\{ \Gamma\left(1 + \frac{2}{8.4388}\right) - \left[ \Gamma\left(1 + \frac{1}{8.4388}\right) \right]^2 \right\} = 1.264529$$

$$\sigma_a^2 = \delta_a^2 \left\{ \Gamma\left(1 + \frac{2}{\beta_a}\right) - \left[ \Gamma\left(1 + \frac{1}{\beta_a}\right) \right]^2 \right\} = (1.3368)^2 * \left\{ \Gamma\left(1 + \frac{2}{1.3368}\right) - \left[ \Gamma\left(1 + \frac{1}{1.3368}\right) \right]^2 \right\} = 0.86108$$

c) Obtenga las funciones de probabilidad acumulada del largo y del ancho del grano de avena.

$$F(l) = 1 - e^{-\left(\frac{l-\gamma_l}{\delta_l}\right)^{\beta_l}} = 1 - e^{-\left(\frac{l-3.2666}{8.4388}\right)^{5.5474}}$$

$$F(a) = 1 - e^{-\left(\frac{a-\gamma_a}{\delta_a}\right)^{\beta_a}} = 1 - e^{-\left(\frac{a-1.3339}{1.3368}\right)^{3.9894}}$$

d) Determine la función de probabilidad conjunta del largo y del ancho del grano de avena.

Dado que ambas variables se suponen estadísticamente independientes:

$$f(l,a) = f(l) * f(a) = 1.961778 \left(\frac{l-3.2666}{8.4388}\right)^{4.5474} * \left(\frac{a-1.3339}{1.3368}\right)^{2.9894} e^{-\left[\left(\frac{l-1.3339}{1.3368}\right)^{3.9894} + \left(\frac{a-1.3339}{1.3368}\right)^{3.9894}\right]}$$

e) Determine la fracción de granos que tiene un largo menor a 5 milímetros y un ancho menor a 1.5 milímetros.

$$p(l < 5) = F(l=5) = 1 - e^{-\left(\frac{5-3.2666}{8.4388}\right)^{5.5474}} = 0.000153739$$

$$p(a < 1.5) = F(a=1.5) = 1 - e^{-\left(\frac{1.5-1.3339}{1.3368}\right)^{3.9894}} = 0.000243647$$

$$p(l < 5, a < 1.5) = p(l < 5) * p(a < 1.5) = 3.7458E - 08$$

f) Determine la fracción de granos que tiene un largo de entre 8 y 12 milímetros y un ancho de entre 2 y 3 milímetros.

$$p(8 \leq l \leq 12) = F(l=12) - F(l=8) = e^{-\left(\frac{8-3.2666}{8.4388}\right)^{5.5474}} - e^{-\left(\frac{12-3.2666}{8.4388}\right)^{5.5474}} = 0.662057027$$

$$p(1.5 \leq a < 3) = F(a=3) - F(a=1.5) = e^{-\left(\frac{1.5-1.3339}{1.3368}\right)^{3.9894}} - e^{-\left(\frac{3-1.3339}{1.3368}\right)^{3.9894}} = 0.939787939$$

$$p(8 \leq l \leq 12, 1.5 \leq a < 3) = p(8 \leq l \leq 12) * p(1.5 \leq a < 3) = 0.622193209$$

### 6.6.1. Función de Probabilidad de Rayleigh

La función de probabilidad de Rayleigh es un caso particular de la distribución de Weibull. Si se hace  $\beta = 2$ ,  $\delta^2 = 2\alpha^2$  y  $\gamma = 0$ , se obtiene la siguiente expresión:

$$f(x) = \frac{x}{\alpha^2} e^{-\frac{x^2}{2\alpha^2}} \quad x > 0 \quad (6.73)$$

Su función de probabilidad acumulada es:

$$F(x) = 1 - e^{-\frac{x^2}{2\alpha^2}} \quad x > 0 \quad (6.74)$$

Su media y su varianza están dadas por:

$$\mu_x = \alpha \sqrt{\pi/2} \quad (6.75)$$

$$\sigma_x^2 = \frac{(4-\pi)\alpha^2}{2} \quad (6.76)$$

Suponga que las componentes  $x$  y  $y$  de un vector en dos dimensiones  $r=(x,y)$ , presentan distribución normal estándar (más adelante se verá). Entonces la magnitud de este vector  $|r| = \text{raíz}(x^2+y^2)$  presenta distribución de Rayleigh, como se demostrará posteriormente. Lo mismo es aplicable para el caso de un número complejo. La distribución de Rayleigh tiene aplicaciones en los modelos de variación de la velocidad del viento, en los modelos de variación de magnitudes de ciertas señales eléctricas, en el ruido de imágenes de ultrasonido y en otras aplicaciones de telecomunicaciones.



## 6.6.2. Función de Probabilidad de Rice

Este modelo es un caso general de la distribución de Rayleigh y su función de probabilidad está dada por la expresión:

$$f(x) = \frac{x}{\alpha^2} e^{-\frac{x^2+c^2}{2\alpha^2}} I_0\left(\frac{cx}{\alpha^2}\right) \quad x > 0 \quad (6.77)$$

Donde  $\alpha > 0$  es un parámetro de escala,  $c > 0$  es un parámetro de localización, e  $I_0(x)$  es la función de Bessel modificada de primer tipo con orden cero.

La función de Bessel modificada de primer tipo con orden  $\alpha$  está definida por la siguiente expresión:

$$I_\alpha(x) = \sum_{n=0}^{\infty} \frac{1}{n! \Gamma(n+\alpha+1)} \left(\frac{x}{2}\right)^{2n+\alpha} \quad (6.78)$$

Su función de probabilidad acumulada está dada por:

$$F(x) = 1 - \int_{\frac{x}{\alpha}}^{\infty} w e^{-\left(\frac{w^2}{2} - \frac{c^2}{2\alpha^2}\right)} I_0\left(\frac{cw}{\alpha}\right) dw \quad (6.79)$$

## 6.7. Función de Probabilidad de Pareto

Su función de probabilidad está dada por la expresión:

$$f(x) = \left( \frac{r\alpha^r}{x^{r+1}} \right) \quad x > \alpha \quad (6.80)$$

Donde  $\alpha > 0$  es un parámetro de escala y  $r > 0$  es un parámetro de forma.

Su función de probabilidad acumulada está dada por la expresión:

$$F(x) = \left[ 1 - \left( \frac{\alpha}{x} \right)^r \right] \quad x > \alpha \quad (6.81)$$

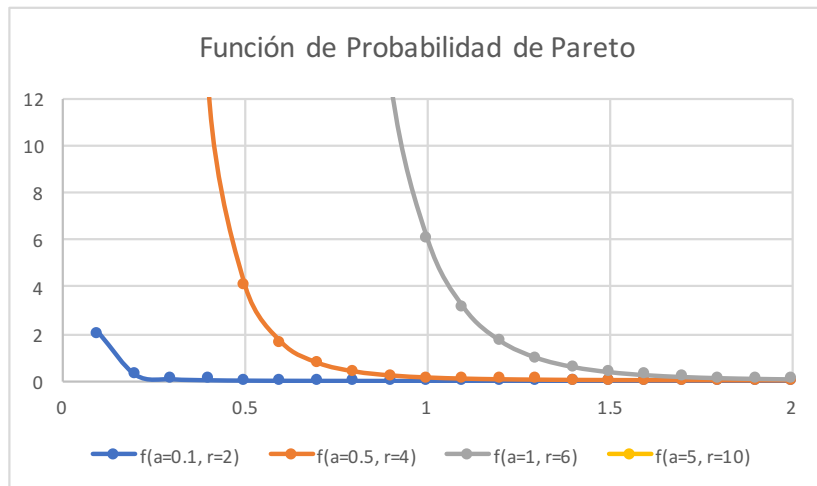
Su media y desviación estándar son:

$$\mu_x = \frac{r\alpha}{r-1} \quad r > 1 \quad (6.82)$$

$$\sigma_x^2 = \frac{r\alpha^2}{[(r-1)(r-2)]} \quad \alpha > 2 \quad (6.83)$$

En la figura 6.15 se muestran algunas funciones de probabilidad de Pareto, la cual se caracteriza por una muy larga cola derecha para pequeños valores de  $r$ .

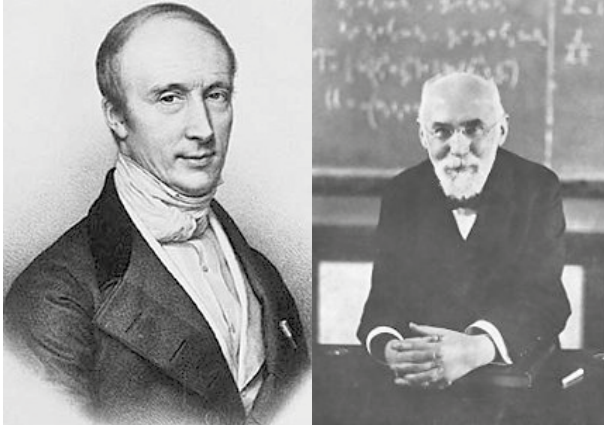
FIGURA 6.15. Función de Probabilidad de Pareto



La función de probabilidad de Pareto se aproxima a la función Delta de Dirac para grandes valores de  $r$ .

## 6.8. Función de Probabilidad de Cauchy–Lorentz

Llamada así en honor de Augustin Cauchy (1789-1857) y Hendrik Antoon Lorentz (1853-1928); su importancia en la física es dada por ser la solución de la ecuación diferencial que describe la resonancia forzada. En espectroscopia describe la forma de las líneas espectrales que son ampliadas por diversos mecanismos, en particular, el mecanismo de ensanchamiento por colisión.



**FIGURA 6.16.**  
Cauchy y Lorentz

**NOTAS** (2019 y 2020)

[https://es.wikipedia.org/wiki/Augustin\\_Louis\\_Cauchy](https://es.wikipedia.org/wiki/Augustin_Louis_Cauchy)

[https://es.wikipedia.org/wiki/Hendrik\\_Antoon\\_Lorentz](https://es.wikipedia.org/wiki/Hendrik_Antoon_Lorentz)

En la hidrología, se utiliza la distribución de Cauchy para analizar variables aleatorias como valores máximos de la precipitación y la descarga de ríos y además, para describir épocas de sequía.

Su función de probabilidad está dada por la siguiente expresión:

$$f(x) = \frac{\alpha / \pi}{(x-c)^2 + \alpha^2} \quad x \in \mathfrak{R} \quad (6.84)$$

Donde  $\alpha > 0$  es un parámetro de escala y  $c \in \mathfrak{R}$  es un parámetro de localización.

Su función de probabilidad acumulada está dada por:

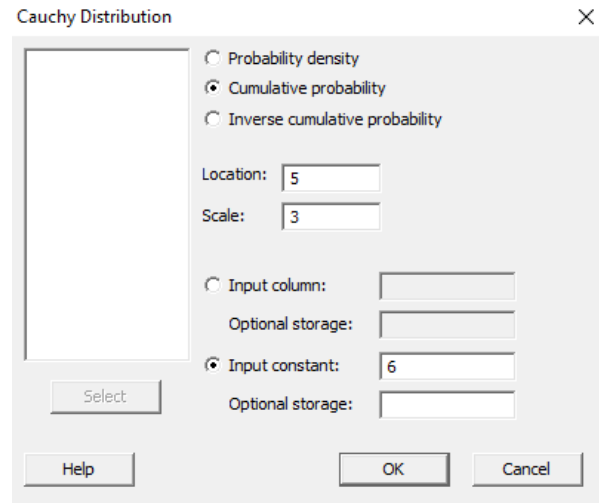
$$F(x) = \frac{1}{2} + \frac{1}{\pi} \operatorname{ang} \tan \left( \frac{x-c}{\alpha} \right) \quad (6.85)$$

En general, la distribución de Cauchy-Lorenz no cuenta con media, ni varianza, ni función generadora de momentos.

Excel no incluye un comando para calcular probabilidades con la función de probabilidad de Cauchy-Lorenz, pero las expresiones (6.84) y (6.85) pueden ser usadas directamente. Para hacer el cálculo de probabilidades con la distribución de Cauchy Lorenz, utilizando Minitab, se realizan los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en Cauchy, tal como se ha venido haciendo antes.
2. Aparece la pantalla de la figura 6.17, en la cual se teclean los valores que se muestran en la misma:

FIGURA 6.17



Se obtiene:

$$p(x \leq 6; c=5, \alpha=3) = 0.602416$$

En el lenguaje de desarrollo R, para el caso de la distribución de Cauchy-Lorenz, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo del valor de la función de densidad de probabilidad exactamente en un punto:

$$dcauchy(x, location = 0, scale = 1, log = FALSE) \quad (6.86)$$

Esta no es una probabilidad, es el valor de la función de densidad en  $x$ .

Para el cálculo de la probabilidad acumulada hasta un punto dado.

$$\text{pcauchy}(x, \text{location} = c, \text{scale} = \alpha, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) \quad (6.87)$$

En donde  $x$  es un valor entre los límites de definición del intervalo de  $x$ ; location el parámetro de localización; scale el parámetro de escala; log es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma FALSE. lower.tail es un valor lógico; si se le asigna un valor verdadero, el valor que arroja es la probabilidad acumulada a la izquierda de  $t$ , de lo contrario, si el valor que toma es falso, el valor que arroja es la probabilidad a la derecha del valor de  $t$ .

Por ejemplo,

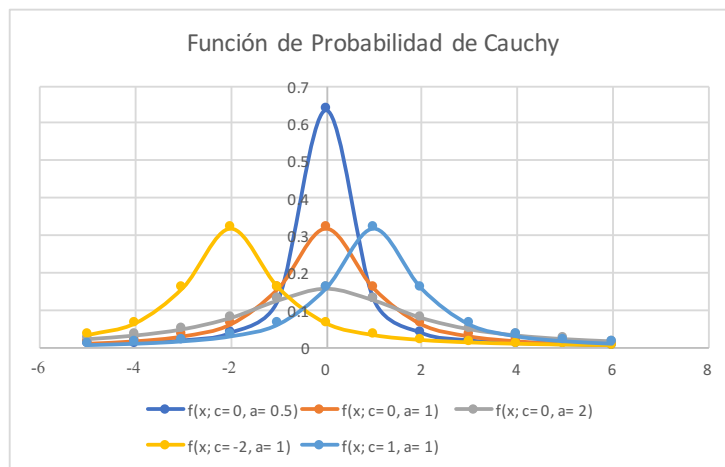
$$p(x=6; c=5, \alpha=3) = \text{dcauchy}(6, \text{location}=5, \text{scale}=3, \text{log}=\text{FALSE}) = 0.095493$$

Esta no es una probabilidad, es el valor de la función de densidad en  $x=6$ .

$$p(x \leq 6; c=5, \alpha=3) = \text{pcauchy}(6, \text{location}=5, \text{scale}=3, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.602416$$

En la figura 6.18 se muestran diversas funciones de probabilidad tipo Cauchy.

**FIGURA 6.18.**



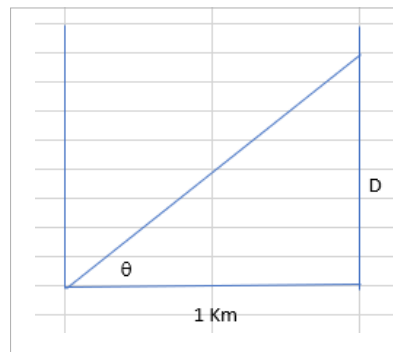
Es posible generar números aleatorios con función de probabilidad tipo Cauchy, al despejar la variable aleatoria  $x$  de su función de probabilidad acumulada, dada en la expresión (6.80):

$$x = c + \alpha \tan \left[ \pi \left( u - \frac{1}{2} \right) \right] \quad (6.88)$$

Donde  $c$  es el parámetro de localización,  $\alpha$  es el parámetro de escala y  $u$  es un número aleatorio uniforme en el intervalo  $[0,1]$ .

### Ejercicio 6.10

(tomado de la página [http://carleos.epv.uniovi.es/~carleos/docencia/teloydisren/descriptiva+probabilidad/l\\_edyp/lib\\_edyp\\_html/node56.html](http://carleos.epv.uniovi.es/~carleos/docencia/teloydisren/descriptiva+probabilidad/l_edyp/lib_edyp_html/node56.html)). Un ejercicio de orientación para una persona ciega consiste en hacerla andar en línea recta entre dos paredes paralelas que distan un kilómetro entre sí. El grado de desorientación  $D$  es la distancia entre el lugar más cercano desde el punto de partida a la segunda pared y el punto en el que la persona ciega alcanzó la segunda pared. Suponiendo que el ángulo  $\theta$  que forma la primera pared y la dirección escogida por esa persona sigue una distribución uniforme en  $[-\pi/2$  y  $\pi/2]$ , obtenga la función de probabilidad del grado de desorientación.



Para calcular el grado de desorientación véase que  $D = \tan(\theta)$  y que  $f(\theta) = 1/\pi$  para  $\theta \in (-\pi/2, \pi/2)$ . A partir de aquí se calcula la función de distribución de  $D$ .

$$F(d) = p(D \leq d) = p(\tan(\theta) \leq d) = p(\theta \leq \arctan(d)) = \arctan(d) + \pi/2$$

Con lo cual la función de densidad es:

$$f_D(d) = \frac{1/\pi}{1+d^2}$$

Se trata de una distribución Tipo Cauchy de parámetros  $\alpha = 1$  y  $c = 0$ .

## Ejercicio 6.11

Demostrar que la función de probabilidad de Cauchy-Lorenz puede ser generada a partir del cociente de dos variables aleatorias  $u$  y  $v$  estadísticamente independientes con distribución uniforme en el intervalo de  $[-1,1]$  y  $u^2+v^2 < 1$ .

Para deducir la función de probabilidad que presenta el cociente de dos variables aleatorias estadísticamente independientes con distribución uniforme en el intervalo  $[-1, 1]$  se aplica el procedimiento establecido en el subtema 4.3 de la página 162 de este libro:

Sean las variables  $u, v$  continuas, estadísticamente independientes, con distribución uniforme ambas en el intervalo  $[-1, 1]$ , es decir

$$f_u(u) = \frac{1}{2} \quad u \in [-1,1]$$

$$g_v(v) = \frac{1}{2} \quad v \in [-1,1]$$

Por lo que su función de probabilidad conjunta está dada por la expresión:

$$h(u,v) = f_u(u) g_v(v) = \frac{1}{4} \quad \begin{array}{l} u \in [-1,1] \\ v \in [-1,1] \end{array}$$

Sea la variable aleatoria  $s = u^2 + v^2$ , y una segunda variable aleatoria  $t = u/v$ .

De estas dos expresiones se despejan  $u$  y  $v$ , y se derivan parcialmente con respecto a  $u$  y a  $v$  respectivamente:

$$u^2 = \frac{st^2}{t^2+1} \quad v^2 = \frac{s}{t^2+1}$$

*Derivando:*

$$2u \frac{\partial u}{\partial s} = \frac{t^2}{t^2+1} \rightarrow \frac{\partial u}{\partial s} = \frac{t^2}{2u(t^2+1)}$$

$$2u \frac{\partial u}{\partial t} = \frac{2ts}{(t^2+1)^2} \rightarrow \frac{\partial u}{\partial t} = \frac{ts}{u(t^2+1)^2}$$

$$2v \frac{\partial v}{\partial s} = \frac{1}{t^2+1} \rightarrow \frac{\partial v}{\partial s} = \frac{1}{2v(t^2+1)}$$

$$2v \frac{\partial v}{\partial t} = \frac{-2st}{(t^2+1)^2} \rightarrow \frac{\partial v}{\partial t} = \frac{-st}{v(t^2+1)^2}$$

Por lo que el Jacobiano de la transformación sería:

$$J \begin{pmatrix} u & v \\ s & t \end{pmatrix} = \begin{vmatrix} \frac{t^2}{2u(t^2+1)} & \frac{1}{2v(t^2+1)} \\ \frac{st}{u(t^2+1)^2} & \frac{-st}{v(t^2+1)^2} \end{vmatrix} = \frac{-st^3}{2uv(t^2+1)^3} - \frac{-st}{2uv(t^2+1)^3} = \frac{-st}{2uv(t^2+1)^2}$$

Pero:

$$u^2v^2 = \left( \frac{st^2}{t^2+1} \right) \left( \frac{s}{t^2+1} \right) = \frac{s^2 t^2}{(t^2+1)^2} \rightarrow uv = \frac{st}{t^2+1}$$

$$J \begin{pmatrix} u & v \\ s & t \end{pmatrix} = -\frac{1}{2(t^2+1)}$$

$$h(s,t) = h(u,v) \left| J \begin{pmatrix} u & v \\ s & t \end{pmatrix} \right| = \frac{1}{8(t^2+1)}$$

Nótese que  $s = u^2 + v^2 < 1$  implica que  $0 < s < 1$ . Por lo que la función  $h(s,t)$  está compuesta de dos funciones estadísticamente independientes:

$$h(s,t) = \left( \frac{\pi}{8} \right) \left( \frac{1/\pi}{1+t^2} \right)$$

Nótese que la primera de ellas, para  $s = u^2 + v^2$ , corresponde con una variable aleatoria uniforme y la segunda para  $t = u/v$ , corresponde con una función de probabilidad de Cauchy.

### Ejercicio 6.12

Obtenga el primero, segundo y el tercer cuartil de la función de probabilidad de Cauchy.

Utilizando la expresión (6.83):

$$\begin{aligned} q_1 &= c + \alpha \tan(\pi(0.25 - 0.5)) = c - \alpha \\ \text{Mediana} = q_2 &= c + \alpha \tan(\pi(0.5 - 0.5)) = c \\ q_3 &= c + \alpha \tan(\pi(0.75 - 0.5)) = c + \alpha \end{aligned}$$

### Ejercicio 6.13

Calcule la moda de la función de probabilidad de Cauchy.

Derivando, igualando a cero y despejando el valor de  $x$ , se obtiene el valor de la moda, ya que representa el máximo valor que alcanza la función de densidad de probabilidad. En este caso la moda es igual a la mediana:  $c$ .



## 6.9. Función de Probabilidad Logística

$$f(x) = \frac{e^{-\frac{x-\mu}{\alpha}}}{\alpha \left(1 + e^{-\frac{x-\mu}{\alpha}}\right)^2} \quad x \in \mathfrak{R} \quad (6.89)$$

Donde  $m$  es un parámetro de localización y  $\alpha$  es un parámetro de escala.

Su función de probabilidad acumulada está dada por la expresión:

$$F(x) = \frac{1}{1 + e^{-\frac{x-\mu}{\alpha}}} \quad x \in \mathfrak{R} \quad (6.90)$$

Su media y su varianza están dadas a continuación:

$$\mu_x = \mu \quad (6.91)$$

$$\sigma_x^2 = \frac{(\pi\alpha)^2}{3} \quad (6.92)$$

## Ejercicios propuestos del capítulo 6

- 6.1. Sea la variable aleatoria “X” con distribución uniforme en el intervalo (0,1), comparar  $P(|x-\mu_x| < k\sigma_x)$  con los valores dados por la desigualdad de Tchebychev para “k” igual a 1.25 y 2.0.
- 6.2. Sea  $x$  una variable aleatoria con distribución uniforme sobre el intervalo  $(a,b)$ .
- ¿Cuál es la probabilidad de que  $x$  tome un valor que se encuentre a una desviación estándar de la media?
  - ¿Puede tomar  $x$  un valor que se encuentre a dos desviaciones estándar de la media?
- 6.3. En una compañía telefónica se reciben en promedio 2 llamadas por minuto; en un periodo de 5 minutos, nos interesa conocer la función de distribución hasta la próxima llamada. ¿Cuál es la probabilidad de que la siguiente llamada ocurra dentro de los próximos 5 minutos? Si el operador abandona el conmutador durante un minuto, ¿Cuál es la probabilidad de que no se reciban llamadas durante la ausencia del operador?
- 6.4. Una persona llega a la estación del metro en forma aleatoria entre las 10 y las 11 de la mañana. Si el metro pasa por la estación precisamente cada 10 minutos a partir de las 6 de la mañana, ¿Cuál es la probabilidad de que el usuario tenga que esperar más de 5 minutos?
- 6.5. Un fabricante de aparatos de televisión a color ofrece una garantía de un año en la que se compromete a reemplazar sin costo para el cliente el cinescopio si éste falla. El fabricante estima que el tiempo transcurrido hasta la falla,  $t$  es una variable aleatoria exponencial con la siguiente función de probabilidad:

$$f(t) = \frac{1}{9} e^{-\frac{t}{9}} \quad t > 0$$

- a) ¿A qué porcentaje de los aparatos tendrá que proporcionarles servicio el fabricante?
- b) Si el beneficio por venta es de \$50,000.00 y reemplazar un cinescopio cuesta \$50,000.00, ¿Cuál es el beneficio esperado del negocio?
- 6.6. Suponiendo que el conmutador de una oficina de asesorías recibe un promedio de 0.6 llamadas por minuto, calcular la probabilidad de que en un intervalo de 4 minutos haya al menos tres llamadas.
- 6.7. La vida o duración de un sistema electrónico está dada como la suma de las vidas de sus componentes. El tiempo medio entre fallas es de 4 horas para cada componente. Si el sistema tiene cuatro componentes entonces su tiempo de falla está dado por  $T = t_1 + t_2 + t_3 + t_4$ . ¿Cuál es la probabilidad de que el sistema operé al menos durante 24 horas?
- 6.8. Sea  $x$  una variable aleatoria con distribución beta y parámetros  $\alpha=3$  y  $\beta=1$ .
- a) Graficar la función de densidad de probabilidad.
- b) Obtener la media, la varianza y el coeficiente de asimetría.
- c) ¿Cuál es la probabilidad de que  $x$  tome un valor que se encuentre dentro de una desviación estándar a partir de la media? ¿A dos desviaciones estándar de la media?
- d) Determinar los cuartiles de la distribución.
- 6.9. Para un artículo se necesitan cuatro etapas. Si el tiempo total necesario para armar un artículo, en horas, es una variable aleatoria con distribución gama y parámetro de escala  $\theta=2$ , ¿Cuál es la probabilidad de armar un artículo en menos de 15 horas?
- 6.10. Si los parámetros de la distribución beta son enteros, puede demostrarse que la función de distribución acumulativa beta se encuentra relacionada con la distribución binomial en la siguiente forma:

$$P(X < p) = I_p(\alpha, \beta) = \sum_{y=\alpha}^n \frac{n!}{(n-y)! y!} p^y (1-p)^{n-y}$$

En donde  $n = \alpha + \beta - 1$  y  $0 < p < 1$ . Si  $x$  es una variable aleatoria con una distribución beta con parámetros  $\alpha=2$  y  $\beta=3$ , emplear la relación anterior para obtener:  $p(x < 0.1)$ ,  $p(x < 0.25)$  y  $p(x < 0.5)$

- 6.11. Un sistema está formado por dos componentes independientes  $A$  y  $B$ . El sistema permanecerá operando mientras uno o ambos componentes funcionen. Si el tiempo de vida de la componente  $A$  es una variable aleatoria de Weibull con  $\alpha = 1/2$ ,  $\theta = 10$ , y si el tiempo de vida  $B$  es también una variable de Weibull con  $\alpha = 2$  y  $\theta = 12$ . ¿Cuál es la probabilidad de que el sistema trabaje más de 20 horas?

FIGURA 7.1. Modelos de Distribuciones Muestrales

Modelos Probabilísticos de Distribuciones Muestrales						
Distribución	Parámetros	Valores que toma $x$	Función Densidad de Probabilidad: $f(x)$	Función de Distribución o de Probabilidad Acumulada: $F(x)=p(t \leq x)$	Media	Varianza
Ji Cuadrada ( $\chi^2$ )	$k \in \mathfrak{N}$	$u = \chi^2 = z_1^2 + z_2^2 + \dots + z_k^2$ $u = \chi^2 > 0$	$f(u) = \frac{1}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} u^{\left(\frac{k}{2}-1\right)} e^{-\frac{u}{2}}$	$F(u) = \frac{1}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} \int_0^u v^{\left(\frac{k}{2}-1\right)} e^{-\frac{v}{2}}$	$k$	$2k$
t de Student	$k \in \mathfrak{N}$	$t = \frac{z}{\sqrt{\frac{\chi^2}{k}}}$ $t \in \mathfrak{R}$	$f(t) = \frac{\Gamma\left[\frac{k+1}{2}\right]}{\sqrt{\pi k} \Gamma\left[\frac{k}{2}\right]} \frac{1}{\left[\frac{t^2}{k} + 1\right]^{\left(\frac{k+1}{2}\right)}}$	$F(t) = \frac{\Gamma\left[\frac{k+1}{2}\right]}{\sqrt{\pi k} \Gamma\left[\frac{k}{2}\right]} \int_{-\infty}^t \frac{1}{\left[\frac{s^2}{k} + 1\right]^{\left(\frac{k+1}{2}\right)}} \frac{1}{\left[\frac{s^2}{k} + 1\right]^{\left(\frac{k+1}{2}\right)}} ds$	0	$\frac{k}{(k-2)}$
F de Fisher	$u, v \in \mathfrak{N}$ $u > 0$ $v > 0$	$f = \frac{w}{\frac{y}{v}} > 0$ $w, y$ son $\chi^2$	$h(f) = \frac{\Gamma\left(\frac{u+v}{2}\right) \left(\frac{u}{v}\right)^{\left(\frac{u}{2}\right)} f^{\left(\frac{u}{2}-1\right)} \Gamma\left(\frac{u}{2}\right) \Gamma\left(\frac{v}{2}\right) \left[\frac{u}{v} f + 1\right]^{\frac{(u+v)}{2}}}{\Gamma\left(\frac{u}{2}\right) \Gamma\left(\frac{v}{2}\right) \left[\frac{u}{v} f + 1\right]^{\frac{(u+v)}{2}}}$	$H(f) = \frac{\Gamma\left(\frac{u+v}{2}\right) \left(\frac{u}{v}\right)^{\left(\frac{u}{2}\right)} \int_0^f \frac{t^{\left(\frac{u}{2}-1\right)} \Gamma\left(\frac{u}{2}\right) \Gamma\left(\frac{v}{2}\right) \left[\frac{u}{v} t + 1\right]^{\frac{(u+v)}{2}} dt}{\Gamma\left(\frac{u}{2}\right) \Gamma\left(\frac{v}{2}\right) \left[\frac{u}{v} t + 1\right]^{\frac{(u+v)}{2}}}$	$\mu_f = \frac{v}{(v-2)}$  $v > 2$	$\sigma_f^2 = \frac{2v^2(u+v-2)}{u(v-2)^2(v-4)}$  $v > 4$

## 7. La Función de Probabilidad Normal y las Funciones de Probabilidad Muestrales

**FIGURA 7.2.** Billete de 10 marcos alemanes, con la figura de Gauss y en el fondo la distribución normal



Figura tomada de Deutsche Bundesbank, Frankfurt am Main, Germany – [http://www.bundesbank.de/Redaktion/DE/Standardartikel/Kerngeschaefsfelder/Bargeld/dm\\_banknoten.html#doc18118bodyText2](http://www.bundesbank.de/Redaktion/DE/Standardartikel/Kerngeschaefsfelder/Bargeld/dm_banknoten.html#doc18118bodyText2), Dominio público, <https://commons.wikimedia.org/w/index.php?curid=3813487>

A partir del siglo XVII, diferentes científicos se abocaron a determinar la forma analítica de una curva que se ajustara a la distribución de un conjunto de datos de errores de medición cometidos. Alrededor de 1632, Galileo Galilei (1564-1642), astrónomo, filósofo, ingeniero, matemático y físico italiano (magnífico representante del Renacimiento), [https://es.wikipedia.org/wiki/Galileo\\_Galilei](https://es.wikipedia.org/wiki/Galileo_Galilei), fue el primero que argumentó por escrito que debía hacerse un tratamiento sistemático y científico a la medición de los errores, resumiendo en cinco puntos su planteamiento:

1. Existe un y solo un valor numérico que representa la distancia real de una estrella al centro de la tierra.
2. Las observaciones que se realicen respecto a esta distancia contienen errores ocasionados por diversas fuentes: el observador, los instrumentos de medición utilizados y las condiciones observacionales, entre otras.
3. Las mediciones obtenidas a partir de las observaciones realizadas se distribuyen simétricamente alrededor del valor real, lo que quiere decir que los errores cometidos se distribuyen alrededor del cero.
4. Los errores menores se presentan de manera más frecuente que los errores más grandes.
5. La distancia a obtener es función de las observaciones angulares, de modo que, un ajuste pequeño en las observaciones puede resultar en una variación grande en la distancia calculada.

En 1744, Pierre Simon Laplace (1749-1827), [https://es.wikipedia.org/wiki/Pierre-Simon\\_Laplace](https://es.wikipedia.org/wiki/Pierre-Simon_Laplace), astrónomo, físico y matemático francés presentó una primera curva, obtenida bajo el supuesto que esta debe ser simétrica y no decreciente para valores positivos de  $x$ . Al ser dicha función no decreciente, el cociente entre ella y su derivada debía ser una constante negativa, lo cual puede ser escrito matemáticamente como:

$$\frac{d\phi(x)}{dx} = -m\phi(x) \quad (7.1)$$

La cual representa una ecuación diferencial de primer orden, que por el método de separación de variables y su integración se resuelve como:

$$\phi(x) = \frac{m}{2} e^{-|x|} \quad (7.2)$$

El problema de esta función es de que no es derivable en el origen.

Para 1777, Daniel Bernoulli (1700–1782), [https://es.wikipedia.org/wiki/Daniel\\_Bernoulli](https://es.wikipedia.org/wiki/Daniel_Bernoulli), matemático, estadístico, físico y médico suizo, de la portentosa familia Bernoulli, bajo el mismo supuesto de errores acotados, supuso sobre su curva propuesta, que fuera derivable en el valor donde alcanza su máximo.

Hizo entonces su aparición el genial Karl Friedrich Gauss (1777-1855), [https://es.wikipedia.org/wiki/Carl\\_Friedrich\\_Gauss](https://es.wikipedia.org/wiki/Carl_Friedrich_Gauss), matemático, astrónomo,

geobotánico y físico alemán, conocido como el príncipe de las matemáticas, quien se basó en el Método de los Mínimos Cuadrados (se analizará en el libro de Fundamentos de Estadística) para obtener la curva que mejor se ajustara a los datos que se tenían. Este criterio lo sustentó en los tres puntos siguientes:

1. Los errores menores son más probables que los mayores.
2. La probabilidad que se cometa en la medición un error de magnitud real  $\varepsilon$  o  $-\varepsilon$  resulta ser la misma.
3. En presencia de varias observaciones sobre la misma cantidad, el valor más probable de esa cantidad es el promedio de las observaciones.

Por el método de los mínimos cuadrados Gauss llegó a la siguiente ecuación diferencial:

$$\frac{\phi'(x)}{\phi(x)} = -h^2x \quad (7.3)$$

Sobre la base de estas afirmaciones, Gauss concluyó que la densidad de probabilidad para el error (esto es, la curva del error) estaba dada por:

$$\phi(x) = \frac{h}{\sqrt{\pi}} e^{-h^2x^2} \quad (7.4)$$

Por lo anterior, a esta función de probabilidad se le conoce como función de probabilidad de Gauss o gaussiana. Durante los siglos XVIII y XIX se realizaron varios intentos de establecer a esta distribución como la ley probabilística base de todas las variables continuas; debido a esto, se llegó a utilizar el término de Normal, lo cual se debe también a que en muchos aspectos resulta ser la piedra angular de la estadística.

Se dice que una variable aleatoria  $x$  presenta Distribución Normal con media  $\mu_x (-\infty < \mu_x < \infty)$  y desviación estándar  $\sigma_x > 0$ , si tiene función de densidad de probabilidad dada por la siguiente expresión:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad x \in \mathfrak{R} \quad (7.5)$$



Esta función de densidad de probabilidad, resulta ser tan fundamental en el análisis de la incertidumbre y en la teoría de errores, que vale la pena hacer un análisis de la misma.

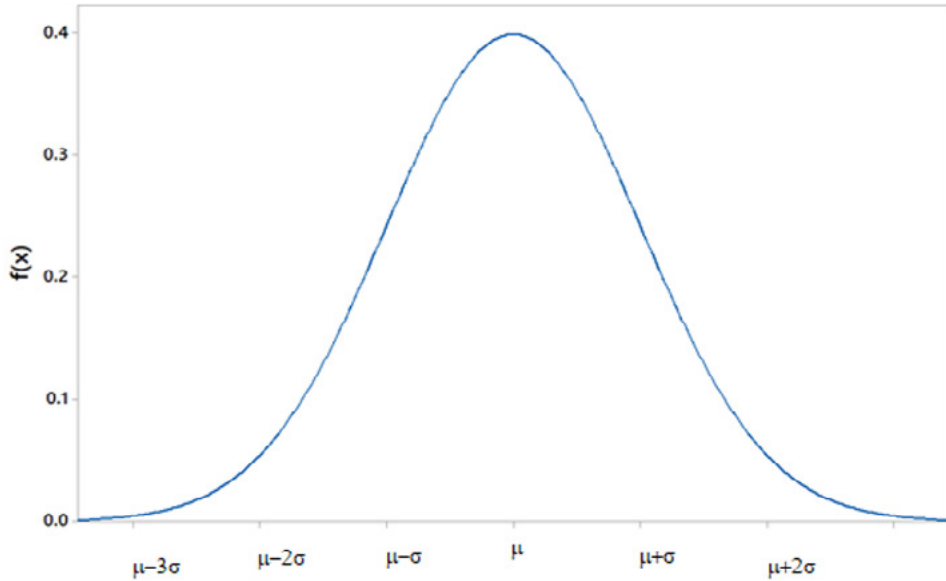
1.  $f(x) > 0 \quad \forall x \in \mathfrak{R}$ , es positiva para cualquier valor de  $x$  en todos los reales.
2. El eje de las  $x$  es una asíntota inferior de la curva, es decir:

$$\lim_{x \rightarrow -\infty} \left[ \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \right] = \lim_{x \rightarrow \infty} \left[ \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \right] = 0$$

3. La integral de menos infinito a infinito de esta función es uno.

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$$

4. Es simétrica con respecto a la recta  $x = \mu$ .
5. En el punto  $x = \mu$  se presenta el máximo absoluto de esta función, donde tiene su media, igual a su mediana y a su moda.
6. Esta función decae tan drásticamente desde su máxima altura  $\left( \mu, \frac{1}{\sigma\sqrt{2\pi}} \right)$  hacia ambos lados del punto máximo y se aproxima tan rápido al eje de las  $x$  que basta con graficarla de  $\mu - 3\sigma$  a  $\mu + 3\sigma$ .
7. Esta función presenta dos puntos de inflexión (punto donde la curva cambia de concavidad) en  $x = \mu - \sigma$  y  $x = \mu + \sigma$ .
8. El área bajo la curva entre los puntos de inflexión  $\mu - \sigma$  y  $\mu + \sigma$  es aproximadamente igual a 68.27%; entre los puntos  $\mu - 2\sigma$  y  $\mu + 2\sigma$  es de 95.45% y entre  $\mu - 3\sigma$  y  $\mu + 3\sigma$  es de aproximadamente 99.73%.

FIGURA 7.3. Distribución Normal  $x \sim N(\mu, \sigma)$ 

Su Función de Probabilidad Acumulada sería:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt \quad x \in \mathfrak{R} \quad (7.6)$$

La cual desafortunadamente no tiene una expresión analítica, ya que la integral sólo puede ser resuelta numéricamente. Dado que esta función de probabilidad es muy común, generalmente se hace un cambio de variable:

$$z = \frac{x - \mu}{\sigma} \quad (7.7)$$

A este cambio de variable se le conoce como estandarización de la variable aleatoria  $x$  y a la función resultante se le conoce como función de probabilidad normal estándar:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \quad z \in \mathfrak{R} \quad (7.8)$$

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{1}{2}t^2} dt \quad (7.9)$$

La función de densidad de probabilidad normal estándar establecida en la expresión 7.8 presenta las siguientes propiedades:

1.  $f(z) > 0 \quad \forall z \in \mathfrak{R}$ .
2. El eje de las  $z$  es una asíntota inferior de la curva.
3. La integral de menos infinito a infinito de esta función es uno.

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2}{2}} dz = 1$$

4. Es simétrica con respecto a la recta  $z=0$ .
5. En el origen se presenta el máximo absoluto de esta función, donde tiene su media, igual a su mediana y a su moda  $\left(0, \frac{1}{\sqrt{2\pi}}\right)$
6. Esta función decae tan drásticamente desde su máxima altura  $\left(0, \frac{1}{\sqrt{2\pi}}\right)$  hacia ambos lados del punto máximo y se aproxima tan rápido al eje de las  $z$  que basta con graficarla de  $-3$  a  $+3$ .
7. Esta función presenta dos puntos de inflexión (punto donde la curva cambia de concavidad) en  $z=-1$  y  $z=+1$ .
8. El área bajo la curva entre los puntos de inflexión  $-1$  y  $+1$  es aproximadamente igual a 68.27%; entre los puntos  $-2$  y  $+2$  es de 95.45% y entre  $-3$  y  $+3$  es de aproximadamente 99.73%.

La Función Generatriz de Momentos de la Función de Probabilidad Normal se obtiene de la siguiente forma:

$$FGM_x(t) = E\{e^{xt}\} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{xt} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}} dx$$

$$FGM_x(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{xt - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Pero:

$$xt - \frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2 = -\frac{1}{2} \left[ \frac{x - (\mu+t)}{\sigma} \right]^2 + \mu t + \frac{1}{2} \sigma^2 t^2$$

$$FGM_x(t) = \frac{e^{\mu t + \frac{1}{2} \sigma^2 t^2}}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} \left[ \frac{x - (\mu+t)}{\sigma} \right]^2} dx$$

Pero:

$$\frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2} \left[ \frac{x - (\mu+t)}{\sigma} \right]^2} dx = 1$$

$$FGM_x(t) = e^{\mu t + \frac{1}{2} \sigma^2 t^2} \quad (7.10)$$

Su función característica:

$$FC_x(t) = e^{i\mu t - \frac{1}{2} \sigma^2 t^2} \quad (7.11)$$

Su media y su varianza son:

$$\mu_x = \left[ \frac{dFGM_x(t)}{dt} \right]_{t=0} = \left[ (\mu + \sigma^2 t) e^{\mu + \frac{1}{2} \sigma^2 t} \right]_{t=0} = \mu_x \quad (7.12)$$

$$\sigma_x^2 = \left[ \frac{d^2 FGM_x(t)}{dt^2} \right]_{t=0} - \mu^2$$

$$\sigma_x^2 = \left[ (\mu + \sigma^2 t)^2 e^{\mu + \frac{1}{2} \sigma^2 t} + \sigma^2 e^{\mu + \frac{1}{2} \sigma^2 t} \right]_{t=0} - \mu^2 = \sigma^2 \quad (7.13)$$

Su coeficiente de asimetría y su coeficiente de curtosis son cero, tal y como se demuestra al obtener la tercera derivada y la cuarta derivada de la Función Generatriz de Momentos y aplicar la definición de estos conceptos.

Para el cálculo de probabilidades con la distribución Normal, utilizando Excel, se utiliza el siguiente comando:

$$\text{DISTR.NORM.N}(x, \mu_x, \sigma_x, \text{Acumulado}) \quad (7.14)$$

En donde  $x \in \mathfrak{R}$ ;  $\mu_x$  representa la media de  $x$ ,  $\sigma_x$  la desviación estándar de  $x$ , acumulado=verdadero o 1 si se pretende calcular la función de probabilidad acumulada hasta  $x$ ; acumulado=falso o 0 si se pretende calcular el valor de la función de densidad en  $x$ . De tal manera que:

$$p(x=2; \mu=3, \sigma=5) = \text{DISTR.NORM.N}(2, 3, 5, 0) = 0.078209$$

Esta no es una probabilidad, es el valor de la función de densidad en  $x=2$ .

$$p(x \leq 2; \mu=3, \sigma=5) = \text{DISTR.NORM.N}(2, 3, 5, 1) = 0.42074$$

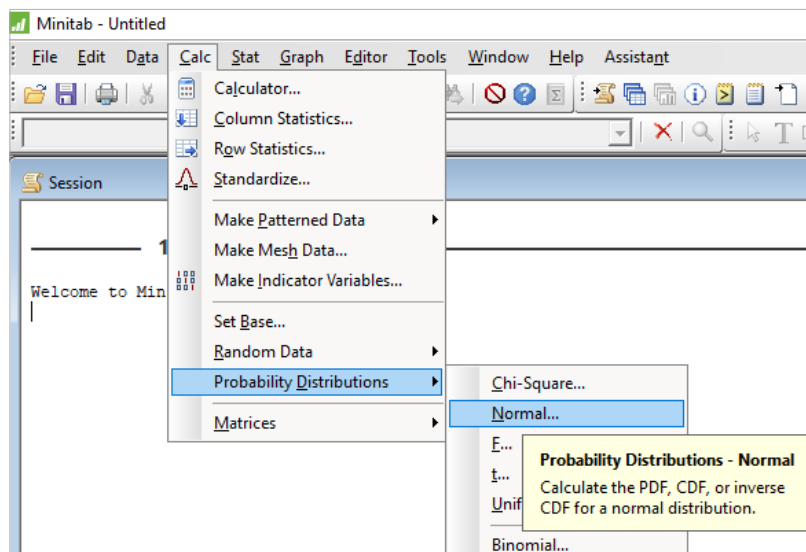
$$p(x > 2; \mu=3, \sigma=5) = 1 - \text{DISTR.NORMAL}(2, 3, 5, 1) = 0.938005$$

$$p(2 \leq x \leq 4; \mu=3, \sigma=5) = \text{DISTR.NORM.N}(4, 3, 5, 1) - \text{DISTR.NORM.N}(2, 3, 5, 1) = 0.15852.$$

Para hacer el cálculo de probabilidades con la distribución Normal, utilizando Minitab, la función de probabilidad que se utiliza es la misma que para Excel, expresión 7.6, por lo que se realizan los siguientes pasos:

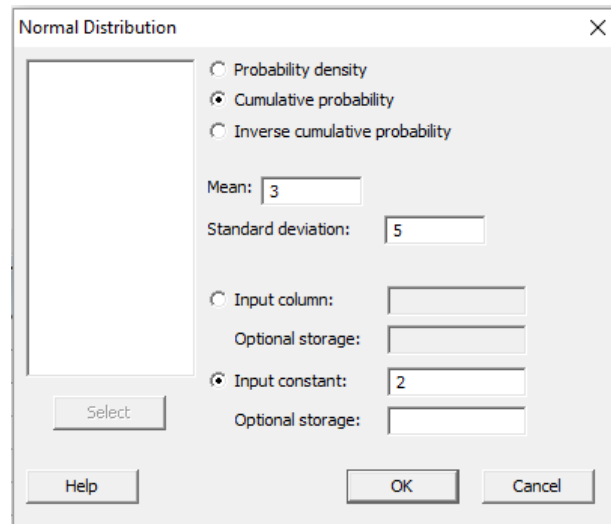
1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en Normal, tal como se muestra en la figura 7.4.

**FIGURA 7.4**



2. Aparece la pantalla de la figura 7.5, en la cual se teclean los valores que se muestran en la misma:

FIGURA 7.5



Se obtiene lo ya calculado antes:

$$p(x \leq 2; \mu=3, \sigma=5) = 0.42074$$

En el lenguaje de desarrollo R, para el caso de la distribución Normal, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo del valor de la función de densidad de probabilidad exactamente en un punto:

$$\text{dnorm}(x, \text{mean}=\text{, sd}=\text{, log} = \text{FALSE}) \quad (7.15)$$

Esta no es una probabilidad, es el valor de la función de densidad en  $x$ .

Para el cálculo de la probabilidad acumulada hasta un punto dado.

$$\text{pnorm}(x, \text{mean}=\text{, sd}=\text{, lower.tail} = \text{TRUE, log.p} = \text{FALSE}) \quad (7.16)$$

En donde  $x$  es un valor real; mean es la media o parámetro de localización; sd es la desviación estándar o el parámetro de escala; log es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad

calculada, por omisión se toma FALSE. lower.tail es un valor lógico; si se le asigna un valor verdadero, el valor que arroja es la probabilidad acumulada a la izquierda de  $x$ , de lo contrario, si el valor que toma es falso, el valor que arroja es la probabilidad a la derecha del valor de  $x$ .

Por ejemplo,

$$p(x=2; \mu_x=3, \sigma_x=5) = \text{dnorm}(x=2; \text{mean}=3 \text{ sd}=5 \text{ log}=\text{FALSE}) = 0.078209$$

Esta no es una probabilidad, es el valor de la función de densidad en  $x = 2$ .

$$p(x \leq 2; \mu_x=3, \sigma_x=5) = \text{pnorm}(2, \text{mean}=3, \text{sd}=5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.42074$$

$$p(x > 2; \mu_x=3, \sigma_x=5) = \text{pnorm}(2, \text{mean}=3, \text{sd}=5, \text{lower.tail}=\text{FALSE}, \text{log.p}=\text{FALSE}) = 0.57926$$

$$p(3 \leq x \leq 4; \mu_x=3, \sigma_x=5) = \text{pnorm}(4, \text{mean}=3, \text{sd}=5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pnorm}(3, \text{mean}=3, \text{sd}=5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.07926$$

En la figura 7.6.a y 7.6.b se muestran diversas funciones de probabilidad tipo Normal.

**FIGURA 7.6.a**

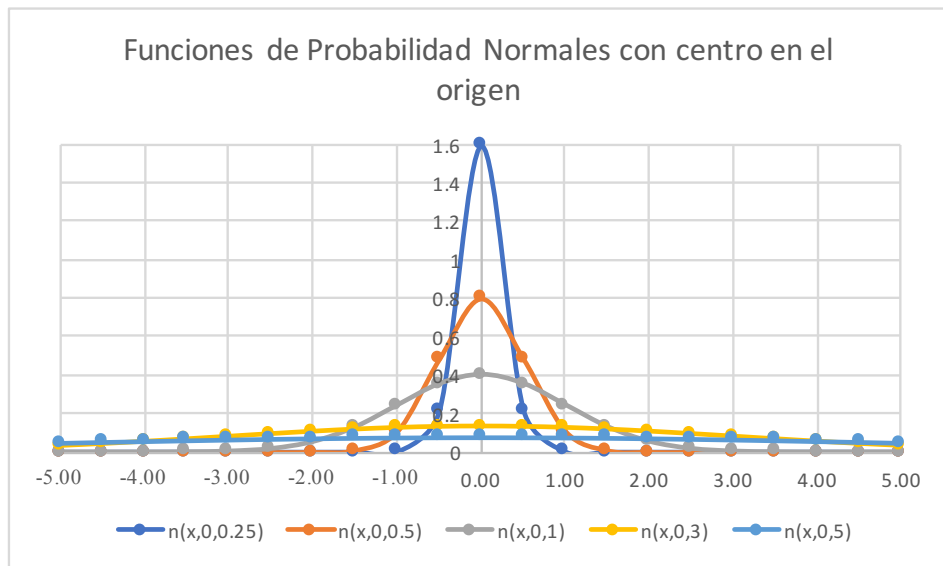
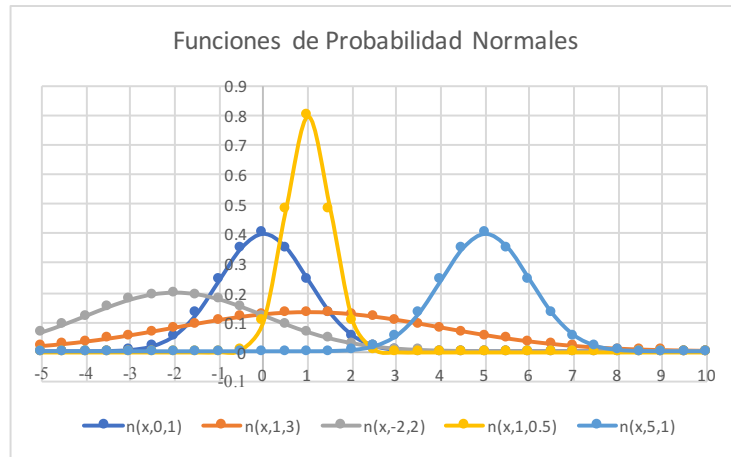


FIGURA 7.6.b



Nótese de las figuras anteriores que para una desviación estándar fija la forma de la curva es la misma, esto implica que la forma puntiaguda o aplanada de la curva normal depende del valor de la desviación estándar, para una desviación estándar pequeña la curva es picuda y para una grande la curva es aplanada. El valor de referencia de la desviación estándar es el de la curva normal estándar, cuyo valor es uno. Para una media fija la posición de la curva es la misma. La función de probabilidad normal estándar cuya expresión analítica es la ecuación 7.8, se usa como referencia para las demás, nótese que tiene como media cero y desviación estándar uno.

## Ejercicio 7.1

La distancia  $x$  a la cual un competidor puede efectuar un tiro, en una competencia de lanzamiento de jabalina, se considera que es una variable aleatoria con distribución normal de media  $m_x=50$  m y desviación estándar  $s_x=5$  m. Si el competidor en cuestión efectúa un tiro, calcular la probabilidad de que su lanzamiento:

- a) Haya sido menor de 40 m.

Usando Excel:

$$p(x < 40) = \text{DISTR.NORM.N}(40, 50, 5, 1) = 0.02275 = 2.275\% = 22750 \text{ ppm.}$$

- b) Haya sido mayor de 55 m.

Usando R:

$$p(x > 55) = \text{pnorm}(55, \text{mean}=50, \text{sd}=5, \text{lower.tail} = \text{FALSE}, \text{log.p} = \text{FALSE}) = 0.158655$$

$$p(x > 55) = 1 - 0.841345 = 0.158655 = 15.8655\% = 158655 \text{ ppm.}$$



c) Haya caído entre los 50 y los 60 m.

Usando R:

$$p(50 \leq x < 60) = \text{pnorm}(60, \text{mean}=50, \text{sd}=5, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) \\ - \text{pnorm}(50, \text{mean}=50, \text{sd}=5, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) = 0.47725$$

d) Obtenga el valor de  $k$  para el cual se cumple que:  $p(50-5k \leq x \leq 50+5k) = 0.90$ .

Nótese que 50 es la media y que 5 es la desviación estándar, por lo que se trata de una curva, cuya área entre  $50-5k$  y  $50+5k$  es 0.90 y está centrada, por lo que a la izquierda de  $50-5k$  o a la derecha de  $50+k$  se tiene un área de 0.05 en ambas colas, como se muestra en la figura 7.6 siguiente, trazada usando Minitab.

Como se puede apreciar en la figura 7.7,  $p(x < 50-5k) = 0.05$ , por lo que usando R se obtiene que:

$$q = \text{qnorm}(0.05, \text{mean}=50, \text{sd}=5, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) = 41.77573$$

$$\text{Por lo que } 50 - 5k = 41.77573 \rightarrow k = 1.644854$$

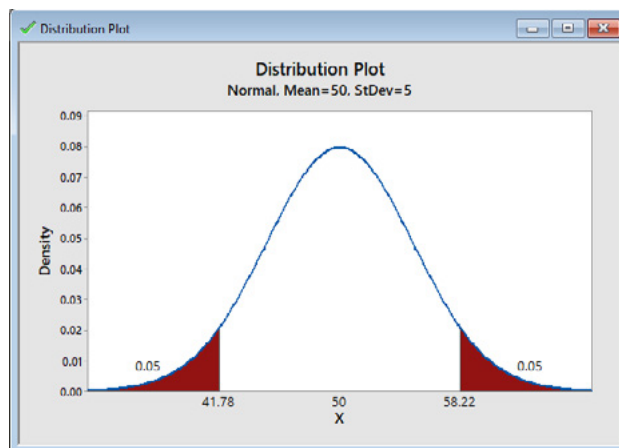


FIGURA 7.7

### Ejercicio 7.2

Una pieza torneada presenta como una de sus características principales el diámetro. La especificación establece que la pieza debe tener un diámetro medio de diseño de 3" (' representa pies y " representa pulgadas), con una variación permitida de  $2.8'' \leq \phi \leq 3.2''$ . Al hacer un análisis estadístico se determinó que el diámetro de la pieza presenta distribución normal con media  $\mu_\phi = 2.9''$  y desviación estándar  $\sigma_\phi = 0.0667''$ . Si la pieza presenta un diámetro menor de 2.8" se

considera inservible y debe ser desechada a un costo de \$12.00; si la pieza presenta un diámetro por arriba de 3.2" puede ser retrabajada a un costo de \$5.00.

- a) Suponga que si el diámetro de la pieza cae entre los límites de especificación, aunque no sea 3.0, no se genera ningún costo. Calcule la pérdida promedio que se tiene en un lote de  $N= 10,000$  piezas.

Costo promedio de pérdida por cada pieza=  $12 \cdot p(\phi < 2.8) + 0 \cdot p(2.8 \leq \phi \leq 3.2) + 5 \cdot p(x > 3.2)$

$p(\phi < 2.8) = \text{pnorm}(2.8, \text{mean}=2.9, \text{sd}=0.0667, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) = 0.066904$

$p(\phi < 3.2) = \text{pnorm}(3.2, \text{mean}=2.9, \text{sd}=0.0667, \text{lower.tail} = \text{FALSE}, \text{log.p} = \text{FALSE}) = 3.433801\text{E}-6$

Costo promedio de pérdida por cada pieza= 0.802865

Costo promedio de pérdida de un lote con 10,000 piezas= \$8,028.65

- b) Suponga que la función de pérdida por cada pieza está dada por la siguiente expresión:

$$c(\phi) = \left\{ \begin{array}{lll} 12 & \text{si} & x < 2.8 \\ 300(x-2)^2 & \text{si} & 2.8 \leq x \leq 3 \\ 125(x-3)^2 & \text{si} & 3 < x \leq 3.2 \\ 5 & \text{si} & x > 3.2 \end{array} \right\}$$

Calcule el costo de la pérdida promedio para un lote de  $N= 10,000$  piezas.

La pérdida promedio por cada pieza estaría dada por la esperanza matemática de la función de pérdida:

$$E\{c(\phi)\} = 12p(x < 2.8) + \frac{300}{0.06667\sqrt{2\pi}} \int_{2.8}^{3.0} (x-2)^2 e^{-\frac{1}{2}\left(\frac{x-3}{0.667}\right)^2} dx$$

$$+ \frac{125}{0.06667\sqrt{2\pi}} \int_3^{3.2} (x-3)^2 e^{-\frac{1}{2}\left(\frac{x-3}{0.667}\right)^2} dx + 5p(x > 3.2)$$

En donde las integrales que aparecen son resueltas por integración numérica utilizando Excel, por la Regla Trapezoidal:

Costo promedio por pieza=  $E\{c(f)\} = \$1.720454$

Costo promedio de pérdida por un lote de 10,000 piezas= \$17,204.54

### Ejercicio 7.3

Una aplicación requiere que un dispositivo electrónico dure 500 horas. Se funde el dispositivo y se va por un repuesto, pero no hay del tipo que siempre se ha usado, solo se tienen dos dispositivos electrónicos parecidos. El primero de ellos tiene una duración con distribución normal de media 480 horas y desviación estándar de 50 horas. El segundo dispositivo tiene una duración con distribución normal de media 470 horas y desviación estándar de 60 horas.

a) Determine cuál de los dos dispositivos pondría como reemplazo.

En este ejercicio se debe elegir el dispositivo más confiable de los dos. La confiabilidad de un dispositivo es la probabilidad de que dicho dispositivo se desempeñe adecuadamente, bajo ciertas condiciones dadas durante un cierto período de tiempo.

En este caso se debe elegir aquel dispositivo que presente la mayor probabilidad  $p(t > 500)$ , por lo que:

$P(t_1 > 500) = \text{pnorm}(500, \text{mean}=480, \text{sd}=50, \text{lower.tail}=\text{FALSE}, \text{log.p}=\text{FALSE}) = 0.344578$

$P(t_2 > 500) = \text{pnorm}(500, \text{mean}=470, \text{sd}=60, \text{lower.tail}=\text{FALSE}, \text{log.p}=\text{FALSE}) = 0.308538$

Como se puede apreciar el dispositivo más confiable es el primero.

b) Suponga que el segundo dispositivo puede regular su desviación estándar, ¿para qué valor límite sería mejor el segundo dispositivo?

En este caso  $p(t_2 > 500) = p(z > z_0) = 0.344578$

$z_0 = \text{qnorm}(0.344578, \text{mean}=0, \text{sd}=1) = 0.4$

Pero  $z_0 = (500 - 470) / \text{sd} \rightarrow \text{sd} > (500 - 470) / 0.4 = 75$

El segundo dispositivo con media de 470 sería mejor que el primero si su desviación estándar fuera mayor de 75.

## 7.1. Distribución Lognormal

La distribución lognormal es el modelo matemático que representa a la distribución de una variable aleatoria cuyo logaritmo se comporta como una variable aleatoria normal. En el ámbito de la Ingeniería la función de probabilidad lognormal ha sido utilizada para describir el tiempo de falla en Confiabilidad, pero presenta muchas aplicaciones en otros campos de Física, Biología y Ciencias Sociales.

A continuación, se deducirá la función de probabilidad lognormal utilizando para ello la expresión (3.70) de la página 134 de este libro.

*Suponga*

$$x \approx N(\mu, \sigma)$$

*donde*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

Sea  $y = e^x$ , lo que implica que  $x = \text{Ln}(y)$  y que  $dx/dy = 1/y$

$$f_y(y) = f_x(x) \left| \frac{dx}{dy} \right| = \frac{1}{y\sigma\sqrt{2\pi}} e^{-\left(\frac{\text{Ln}(y)-\mu}{\sigma}\right)^2} \quad y > 0 \quad (7.17)$$

La media de la distribución lognormal está dada por la expresión:

$$\mu_y = e^{\mu + \frac{\sigma^2}{2}} \quad (7.18)$$

Su varianza está dada por la expresión:

$$\sigma_y^2 = (e^{\sigma^2} - 1) e^{2\mu + \sigma^2} \quad (7.19)$$

Para el cálculo de probabilidades con la distribución lognormal, utilizando Excel, se usa el siguiente comando:

$$\text{DISTR.LOGNORM}(y, \mu_y, \sigma_y, \text{Acumulado}) \quad (7.20)$$

En donde  $y \in \mathcal{R}^+$ ;  $\mu_y$  representa la media de  $y$ ,  $\sigma_y$  la desviación estándar, acumulado= verdadero o 1 si se pretende calcular la función de probabilidad acumulada hasta  $y$ ; acumulado= falso o 0 si se pretende calcular el valor de la función de densidad en  $y$ . De tal manera que:

$$p(y=2; \mu=3, \sigma=5) = \text{DISTR.LOGNORM}(2, 3, 5, 0) = 0.035866$$

Esta no es una probabilidad, es el valor de la función de densidad en  $y=2$ .

$$p(y \leq 2; \mu=3, \sigma=5) = \text{DISTR.NORM}(2, 3, 5, 1) = 0.322266$$

$$p(y > 2; \mu=3, \sigma=5) = 1 - \text{DISTR.LOGNORM}(2, 3, 5, 1) = 0.677733$$

$$p(2 \leq y \leq 4; \mu=3, \sigma=5) = \text{DISTR.LOGNORM}(4, 3, 5, 1) - \text{DISTR.LOGNORM}(2, 3, 5, 1) = 0.051179$$

Para hacer el cálculo de probabilidades con la distribución lognormal, utilizando Minitab, se realizan los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en Lognormal, tal como se muestra en la figura 7.4.
2. Aparece la pantalla de la figura 7.8, en la cual se teclean los valores que se muestran en la misma:

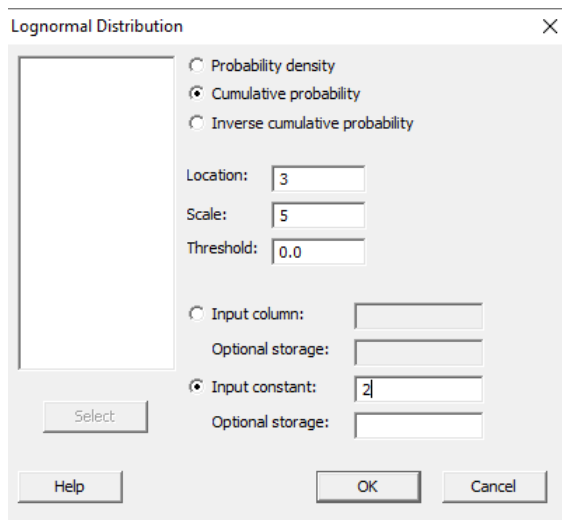


FIGURA 7.8

Se obtiene lo ya calculado antes:

$$p(y \leq 2; \mu=3, \sigma=5) = 0.322266$$

En el lenguaje de desarrollo R, para el caso de la distribución lognormal, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo del valor de la función de densidad de probabilidad exactamente en un punto:

$$\text{dlnorm}(y, \text{meanlog}=\mu, \text{sdlog}=\sigma, \text{log} = \text{FALSE}) \quad (7.21)$$

Esta no es una probabilidad, es el valor de la función de densidad en  $y$ .

Para el cálculo de la probabilidad acumulada hasta un punto dado.

$$\text{plnorm}(y, \text{meanlog}=\mu, \text{sdlog}=\sigma, \text{lower.tail} = \text{TRUE}, \text{log.p} = \text{FALSE}) \quad (7.22)$$

En donde  $y$  es un valor real positivo;  $\text{meanlog}$  es la media de  $y$  o parámetro de localización;  $\text{sdlog}$  es la desviación estándar o el parámetro de escala;  $\text{log}$  es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma `FALSE`. `lower.tail` es un valor lógico; si se le asigna un valor verdadero, el valor que arroja es la probabilidad acumulada a la izquierda de  $y$ , de lo contrario, si el valor que toma es falso, el valor que arroja es la probabilidad a la derecha del valor de  $y$ .

Por ejemplo,

$$p(y=2; \mu=3, \sigma=5) = \text{dlnorm}(2; \text{meanlog}=3, \text{sdlog}=5, \text{log}=\text{FALSE}) = 0.035866$$

Esta no es una probabilidad, es el valor de la función de densidad en  $x=2$ .

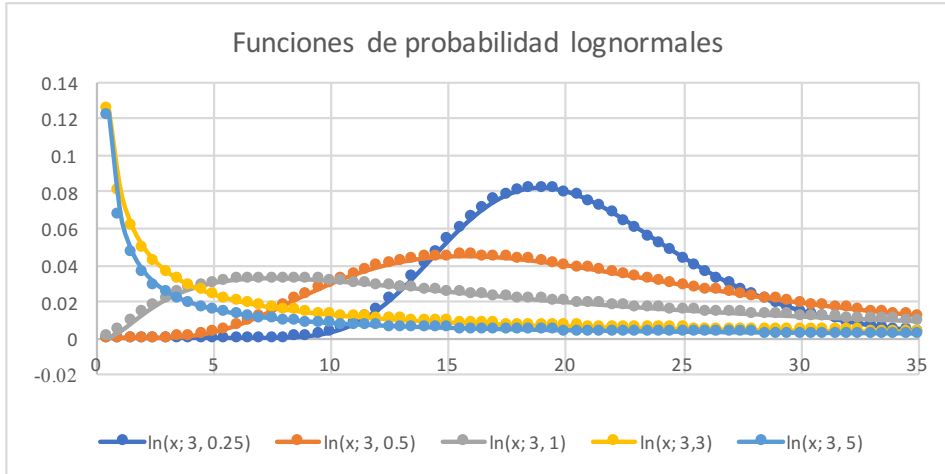
$$p(y \leq 2; \mu=3, \sigma=5) = \text{plnorm}(2, \text{meanlog}=3, \text{sdlog}=5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.322266$$

$$p(y > 2; \mu=3, \sigma=5) = \text{plnorm}(2, \text{mean}=3, \text{sd}=5, \text{lower.tail}=\text{FALSE}, \text{log.p}=\text{FALSE}) = 0.677734$$

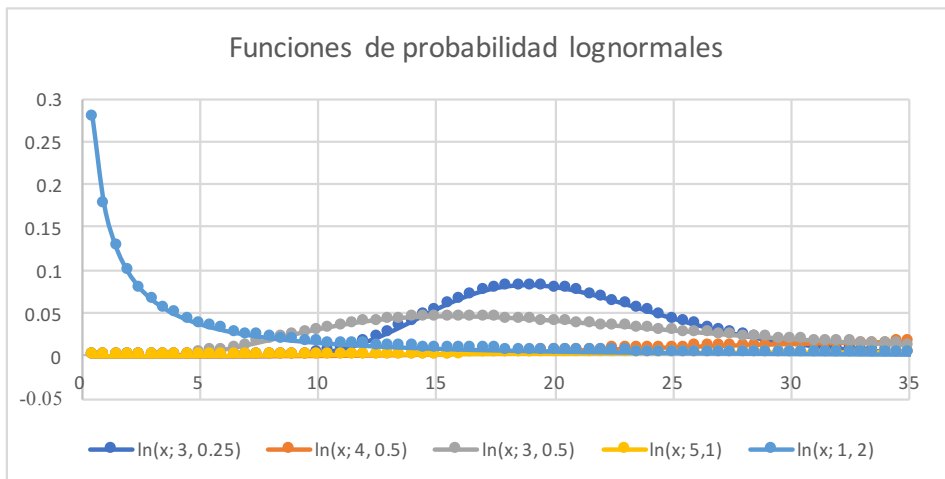
$$p(2 \leq x \leq 4; \mu=3, \sigma=5) = \text{plnorm}(4, \text{meanlog}=3, \text{sdlog}=5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{plnorm}(2, \text{meanlog}=3, \text{sdlog}=5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.051179$$

En la figura 7.9.a y 7.9.b se muestran diversas funciones de probabilidad tipo lognormal.

**FIGURA 7.9.a**



**FIGURA 7.9.b**



La distribución lognormal se ha descrito como el modelo de distribución de datos de vida útil más comúnmente utilizado para muchas aplicaciones de alta tecnología. La distribución se basa en el modelo de crecimiento multiplicativo, lo que significa que en cualquier instante de tiempo, el proceso sufre un incremento aleatorio de degradación que es proporcional a su estado actual. El efecto

multiplicador de todos estos incrementos independientes aleatorios se acumula para generar la falla. Por lo tanto, la distribución suele utilizarse para modelar partes o componentes que fallan principalmente debido a esfuerzo o fatiga.

## Ejercicio 7.4

En el área de confiabilidad de una empresa se registraron las tasas de tiempos de falla ( $\lambda$ ) en la apertura de válvulas de seguridad que fallaron trabajando en condiciones normales de funcionamiento, arrojando los siguientes valores: 0.000002, 0.000008, 0.000091, 0.0000016.

- a) Obtenga la media muestral y la desviación estándar muestral del logaritmo de estas medidas.

$\lambda$	$\text{Ln}(\lambda)$
2E-06	-13.1223634
0.000008	-9.43348392
9.1E-05	-9.30465105
1.6E-06	-13.3455069
Media=	-11.3015013
Desv Est=	2.233860451

$$\hat{\mu}_\lambda \approx \bar{\lambda} = -11.302$$

$$\hat{\sigma}_\lambda \approx S_{n-1} = 2.23386$$

- b) Determine su función de probabilidad suponiendo que el comportamiento probabilístico del logaritmo de estas lecturas es normal, lo que implica que su función de probabilidad es lognormal.

Dado que no se dispone de datos sobre la media poblacional y la desviación estándar poblacional se utilizarán como aproximaciones (estimadores) las de la muestra obtenidas en el inciso (a).

$$f_\lambda(\lambda) = \frac{1}{2.23386\sqrt{2\pi\lambda}} e^{-\left(\frac{\text{Ln}(\lambda) + 11.302}{2.3386}\right)^2} \quad \lambda > 0$$

- c) Calcule la probabilidad de que  $\lambda > 0.0001$

$$p(\lambda > 0.0001) = \text{plnorm}(0.0001, \text{meanlog}=11.302, \text{sdlog}=2.23386, \text{lower.tail}=\text{FALSE}, \text{log.p}=\text{FALSE}) = 0.174548 = 17.4548\% = 174548 \text{ ppm}$$



## 7.2 Propiedad Reproductiva o Aditiva de la Distribución Normal

Suponga que se tienen dos variables aleatorias  $x_1 \sim N(\mu_{x1}, \sigma_{x1})$  y  $x_2 \sim N(\mu_{x2}, \sigma_{x2})$  estadísticamente independientes con distribución normal, y que se define la variable  $y = x_1 + x_2$ , ¿qué distribución de probabilidad presenta la variable aleatoria  $y$ ?

De acuerdo con lo señalado en la expresión (3.19), la media de una suma de dos variables aleatorias es igual a la suma de las medias de cada una de ellas. Asimismo, de conformidad con la expresión (3.32), la varianza de una suma de dos variables aleatorias estadísticamente independientes es igual a la suma de las varianzas de cada una de ellas. Sin embargo, persiste la duda, ¿qué distribución de probabilidad presenta la suma de dos variables aleatorias normales estadísticamente independientes?

Para demostrar que la suma de dos variables aleatorias normales es a su vez normal, se usará el concepto de función generatriz de momentos de una distribución normal, la cual se define en la expresión (7.10). Las funciones generatrices de momentos de las variables  $x_1$  y  $x_2$  son:

$$FGM_{x_1}(t) = e^{\mu_1 t + \frac{1}{2} \sigma_1^2 t^2}$$

$$FGM_{x_2}(t) = e^{\mu_2 t + \frac{1}{2} \sigma_2^2 t^2}$$

Dado que las variables aleatorias  $x_1$  y  $x_2$  son estadísticamente independientes, la función generatriz de momentos conjunta se obtiene como el producto de las funciones generatrices de momentos de cada una de ellas:

$$FGM_{x_1, x_2}(x_1 + x_2) = FGM_{x_1}(x_1) FGM_{x_2}(x_2)$$

$$FGM_{x_1, x_2}(x_1 + x_2) = \left( e^{\mu_1 t + \frac{1}{2} \sigma_1^2 t^2} \right) \left( e^{\mu_2 t + \frac{1}{2} \sigma_2^2 t^2} \right)$$

$$FGM_{x_1 + x_2}(x_1 + x_2) = e^{(\mu_1 + \mu_2)t + \frac{1}{2}(\sigma_1^2 + \sigma_2^2)t^2}$$

$$FGM_y(y) = e^{\mu_y t + \frac{1}{2} \sigma_y^2 t^2}$$

Como se puede apreciar en la última expresión la función generatriz de momentos obtenida corresponde con la de una variable aleatoria normal, por la propiedad de unicidad, que establece que la función generatriz de momentos de existir es única, por lo que queda demostrado que la suma de dos variables aleatorias normales estadísticamente independientes arroja a su vez una variable aleatoria normal de media la suma de las medias y de varianza la suma de las varianzas de las variables aleatorias que se están sumando.

De la misma forma se demuestra que la diferencia de dos variables aleatorias normales estadísticamente independientes es a su vez una variable aleatoria normal, cuya media se obtiene como la diferencia entre las medias de las variables que se restan y de varianza la suma de las variancias de dichas variables; basta con sustituir en la suma  $x_1+(-x_2)$ .

Este resultado se puede generalizar para el caso de una combinación lineal de  $n$  variables aleatorias normales estadísticamente independientes, cuya función de probabilidad también es normal.

#### Teorema de Aditividad o de Reproducibilidad de la distribución normal

Sean

$$\begin{aligned}x_1 &\approx N(\mu_1, \sigma_1) \\x_2 &\approx N(\mu_2, \sigma_2) \\&\vdots \\x_n &\approx N(\mu_n, \sigma_n)\end{aligned}$$

$n$  variables aleatorias normales, estadísticamente independientes y sea  
 $y = a_1x_1 + a_2x_2 + \dots + a_nx_n$

Entonces

$$y \approx N(\mu_y, \sigma_y) \tag{7.23}$$

Donde

$$\begin{aligned}\mu_y &= a_1 \mu_1 + a_2 \mu_2 + \dots + a_n \mu_n \\ \sigma_y^2 &= a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \dots + a_n^2 \sigma_n^2\end{aligned}$$

Demostración:

Como las variables son estadísticamente independientes, su función generatriz de momentos conjunta se obtiene multiplicando las funciones generatrices de momentos de cada una de las variables aleatorias definidas:

$$FGM_y(y) = FGM_{x_1}(x_1) FGM_{x_2}(x_2) \cdots FGM_{x_n}(x_n)$$

$$FGM_y(y) = \left( e^{a_1 \mu_1 t + \frac{1}{2} a_1^2 \sigma_1^2 t^2} \right) \left( e^{a_2 \mu_2 t + \frac{1}{2} a_2^2 \sigma_2^2 t^2} \right) \cdots \left( e^{a_n \mu_n t + \frac{1}{2} a_n^2 \sigma_n^2 t^2} \right)$$

$$FGM_y(y) = e^{(a_1 \mu_1 + a_2 \mu_2 + \cdots + a_n \mu_n) t + \frac{1}{2} (a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \cdots + a_n^2 \sigma_n^2) t^2}$$

$$FGM_y(y) = e^{\mu_y t + \frac{1}{2} \sigma_y^2 t^2}$$

Como se puede apreciar en la última expresión la función generatriz de momentos obtenida corresponde con la de una variable aleatoria normal, por lo que queda demostrado que la combinación lineal de  $n$  variables aleatorias normales estadísticamente independientes arroja a su vez una variable aleatoria normal.

### Ejercicio 7.5

Sean  $x_1 \sim N(3,3)$ ,  $x_2 \sim N(4,2)$  y  $x_3 \sim N(5,1)$  variables aleatorias normales estadísticamente independientes y sea  $y = 2x_1 + 3x_2 + 4x_3$ , calcule:

a) La media y la varianza de  $y$ .

$$\begin{aligned} \mu_y &= 2 * (3) + 3 * (4) + 4 * (5) = 38 \\ s_y^2 &= 2^2 * (3)^2 + 3^2 * (2)^2 + 4^2 * (1)^2 = 88 \quad \rightarrow \quad s_y = 9.3808 \end{aligned}$$

b)  $p(y <= 38) = 0.5$

Dado que se pide que sea menor o igual a su media y la normal es simétrica alrededor de su media.

c)  $p(y > 50) = 1 - \text{pnorm}(50, \text{mean}=88, \text{sd}= 9.3808, \text{lower.tail}=\text{FALSE}, \text{log.p}=\text{FALSE}) = 1 - 0.999975 = 0.000025$

### Corolario de normalidad de la media muestral

La media muestral de una muestra de  $n$  unidades muestrales,  $\bar{x}$ , obtenida aleatoriamente de una población normal de media  $\mu_x$  con desviación estándar  $\sigma_x$  es normal, de media la media poblacional y de desviación estándar la desviación estándar poblacional entre la raíz de  $n$ , es decir

$$\bar{x} \sim N\left(\mu_x, \frac{\sigma_x}{\sqrt{n}}\right)$$

Demostración:

Dado que  $x_i$ ,  $i = 1, 2, \dots, n$ , es una unidad muestral de la población  $x \sim N(\mu_x, \sigma_x)$  se cumple que  $E\{x_i\} = \mu_x$  y  $\text{var}\{x_i\} = \sigma_x^2$ , por lo que:

$$\mu_{\bar{x}} = E\{\bar{x}\} = E\left\{\frac{1}{n}(x_1 + x_2 + \dots + x_n)\right\} = \frac{1}{n}[n\mu_x] = \mu_x$$

$$\sigma_{\bar{x}}^2 = \text{var}\{\bar{x}\} = \text{var}\left\{\frac{1}{n}(x_1 + x_2 + \dots + x_n)\right\} = \frac{1}{n^2}[n\sigma_x^2] = \frac{\sigma_x^2}{n}$$

$$\sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}$$

$$\bar{x} \sim N\left(\mu_x, \frac{\sigma_x}{\sqrt{n}}\right) \quad (7.24)$$

La expresión 7.24 se obtuvo del teorema de aditividad o reproducibilidad de la distribución normal.

A las funciones de probabilidad de parámetros muestrales como la media muestral,  $\bar{x}$ , la varianza muestral,  $S_{n-1}^2$ , la fracción defectuosa de una muestra,  $\hat{p}$ , etcétera, se les denomina funciones de probabilidad muestrales.

#### Ejercicio 7.6

Se usa una carretilla de construcción para transportar tabiques rojos recocidos, cuyas dimensiones estandarizadas son  $6 \times 12 \times 24$  cm. Suponga que el peso de un tabique con estas dimensiones presenta una distribución normal con media de 2.25 kg con una desviación estándar de 100 gramos. Si la carga máxima que resiste la carretilla es de 200 kg:

- a) ¿Cuál sería el peso medio de cien ladrillos y cuál sería su desviación estándar?

$$\text{Peso medio de 100 ladrillos} = 100(2.25) = 225 \text{ kg}$$

$$\text{Varianza del peso de 100 ladrillos} = (100)^2 * (0.1)^2 = 100 \text{ kg}^2$$

$$\text{Desviación estándar del peso de 100 ladrillos} = 10 \text{ kg}$$

- b) ¿Cuál sería el número máximo de ladrillos que podría transportar la carretilla, con una probabilidad del 99.73% de que resista la carga?

$$\text{Peso medio de } n \text{ ladrillos} = 2.25 * n$$

$$\text{Varianza del peso de } n \text{ ladrillos} = n^2 * (0.1)^2 = 0.01 * n^2$$

$$\text{Desviación estándar del peso de } n \text{ ladrillos} = 0.1 * n$$

$$p(y \leq 200) = p\left(z \leq \frac{200 - 2.25n}{0.1n}\right) = 0.9973$$

$$z = 3$$

$$3 \leq \frac{200 - 2.25n}{0.1n}$$

Si se despeja  $n$  en la última expresión:

$$n \leq \frac{200}{2.25} = 88.89$$

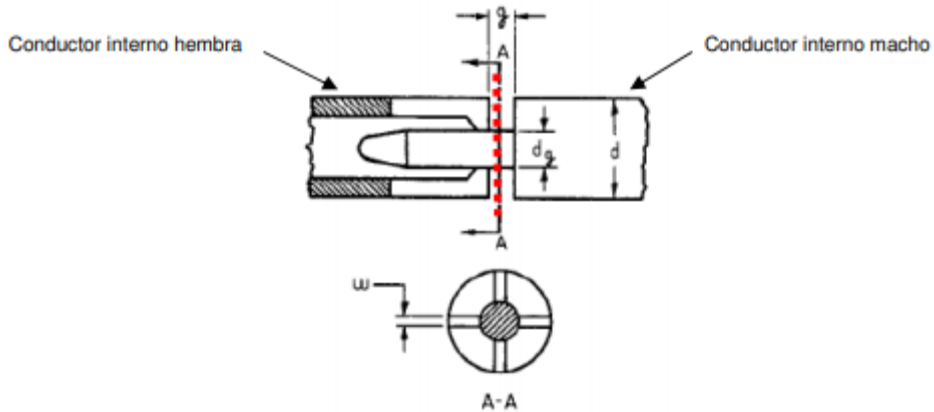
$$n \leq 88$$

### Ejercicio 7.7

Los conectores de audio analógicos conocidos como audiojack de señales analógicas se utilizan para conectar micrófonos, auriculares y otros sistemas a dispositivos electrónicos como computadoras, tabletas o celulares como el que se muestra en la figura siguiente. Se le denomina claro a la diferencia entre el diámetro interno y el diámetro externo. Suponga que un conector macho TRRS presenta un diámetro externo de 3.5 mm de media con una desviación estándar de 0.03 mm. Se supone además que el diámetro interno del conector hembra, a

donde va insertado, presenta una distribución normal con media de 3.52 mm y una desviación estándar de 0.04 mm. Calcule la probabilidad de que el conector macho no entre en el conector hembra.

FIGURA 7.10. Conectores de audio



Se puede apreciar que el claro está definido como  $x_1 - x_2$ ; dado que se trata de dos distribuciones normales, la diferencia también es normal, de media  $\mu_y = \mu_{x_1} - \mu_{x_2} = 3.52 - 3.50 = 0.02$  y de varianza  $\sigma_y^2 = \sigma_{x_1}^2 + \sigma_{x_2}^2 = (0.04)^2 + (0.03)^2 = 0.0025$ , lo que implica que su desviación estándar sea  $\sigma_y = 0.05$ . Si no existe claro el conector macho no podría entrar en el conector hembra, por lo que la probabilidad de que esto ocurra está dada por:

$$p(y \leq 0) = p\left(z \leq \frac{0 - 0.02}{0.05}\right) = 0.344578$$

### 7.3. Teoremas del Límite Central

El Teorema del Límite Central no es un único teorema, sino que consiste en un conjunto de teoremas diferentes o complementarios acerca del comportamiento probabilístico de una suma de variables aleatorias, los cuales afirman que la distribución de probabilidad de dicha suma se aproxima a una distribución normal. El término Central se debe a Polyá (1920), dada su importancia capital en la teoría de probabilidades, ya que, en muchos campos de aplicación, se observan en todo momento distribuciones normales, o casi normales.

La primera versión del Teorema del Límite Central fue dada por De Moivre en su libro “The Doctrine of Chances” de 1733, para el caso especial  $p=1/2$ , el cual generalizó Laplace para el caso  $p$  arbitrario. Esta primera versión corresponde con la aproximación de la distribución binomial a través de la distribución normal. Para ello, se utiliza el siguiente Lema conocido como la Fórmula de Stirling (James Stirling, 1692-1770). En matemáticas, un lema es una proposición demostrada (teorema), utilizada para establecer un teorema menor o una premisa auxiliar que forma parte de un teorema más general.

Lema: La fórmula de Stirling parte de un límite:

$$\lim_{n \rightarrow \infty} \left[ \frac{n!}{\sqrt{nn^n} e^{-n}} \right] = \sqrt{2\pi} \quad (7.25)$$

Lo que conlleva a suponer que para  $n$  lo suficientemente grande, se puede suponer la siguiente aproximación:

$$n! \approx \sqrt{2\pi n} \left( \frac{n}{e} \right)^n \quad (7.26)$$

El lector interesado en consultar la demostración de la Fórmula de Stirling puede consultar la siguiente referencia en su página 196 <https://www.cimat.mx/~jortega/MaterialDidactico/EPyE10/Cap7LaV2.pdf>

#### Teorema de De Moivre-Laplace

Sea  $x$  una variable aleatoria con distribución binomial, con parámetro  $p$ , de media  $np$  y varianza  $np(1-p)$ ; su función de probabilidad está dada por

$$p_x(x) = \binom{n}{x} p^x (1-p)^{n-x} \quad x = 0, 1, 2, \dots, n$$

La variable aleatoria  $x$  es el resultado de una suma de variables aleatorias con función de probabilidad de Bernoulli:

$$x = x_1 + x_2 + \dots + x_n$$

donde

$$p(x_i) = \begin{cases} 1-p & x_i = 0 \\ p & x_i = 1 \end{cases}$$

$$\forall i = 1, 2, \dots, n$$

Sea  $x$  una variable aleatoria normal con media  $np$  y varianza  $np(1-p)$ , con función de probabilidad dada por la expresión:

$$\phi(x) = \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{1}{2} \left( \frac{x-np}{\sqrt{np(1-p)}} \right)^2}$$

Con función de probabilidad acumulada

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2} \left( \frac{t-np}{\sqrt{np(1-p)}} \right)^2} dt$$

Entonces:

$$p(a \leq x \leq b) = \lim_{n \rightarrow \infty} \left[ \sum_{x=a}^{x=b} \binom{n}{x} p^x (1-p)^{n-x} \right] = \Phi(b) - \Phi(a) \quad (7.27)$$

La expresión anterior implica también que

$$\lim_{n \rightarrow \infty} \binom{n}{x} p^x (1-p)^{n-x} = \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{1}{2} \left( \frac{x-np}{\sqrt{np(1-p)}} \right)^2} \quad (7.28)$$

Este teorema establece que una variable aleatoria binomial, con media  $np$  y varianza  $np(1-p)$  tiende a comportarse como una variable aleatoria normal, con la misma media y varianza, en la medida en que el tamaño de muestra tienda a infinito.

La demostración de este teorema la puede consultar en la dirección <https://www.cimat.mx/~jortega/MaterialDidactico/EPyE10/Cap7LaV2.pdf> página 199.



Experimentalmente se puede comprobar que se puede considerar que existe una adecuada aproximación de la binomial a través de la normal, si se cumple que para  $p < 0.5$ ,  $np > 5$ . Para explicar la forma de operar esto se mencionan algunos ejemplos. Suponga que  $p = 0.1$ , para que la binomial se comporte como una normal, debe cumplirse que  $n > 50$ . Si  $p = 0.01$ ,  $n > 500$ ; si  $p = 0.001$ ,  $n > 5000$ .

En la figura 5.10 de la página 211 de este libro se muestran diversos ejemplos de distribuciones de probabilidad binomiales y claramente se pueden apreciar aquellas que tienen una forma parecida a una normal.

Al aplicar esta afirmación debe tenerse en cuenta que la variable aleatoria binomial es discreta y la normal es continua, por lo que puede inducir a errores. Para poder usar esta aproximación debe discretizarse la variable aleatoria normal  $x$ , la forma de hacerlo es tomando como entero el intervalo entre  $x-0.5$  y  $x+0.5$ , como se expresa a continuación, bajo la premisa de que  $n$  debe ser un número lo suficientemente grande como para considerar válida la aproximación:

$$\begin{aligned}
 p(x < a) &\approx \Phi(a-0.5) \\
 p(x \leq a) &\approx \Phi(a+0.5) \\
 p(x = a) &\approx \Phi(a+0.5) - \Phi(a-0.5) \\
 p(x > b) &= 1 - p(x \leq b) \approx 1 - \Phi(b+0.5) \\
 p(x \geq b) &= 1 - p(x < b) \approx 1 - \Phi(b-0.5) \\
 p(a \leq x \leq b) &\approx \Phi(b+0.5) - \Phi(a-0.5) \\
 p(a < x < b) &\approx \Phi(b+0.5) - \Phi(a+0.5) \\
 p(a \leq x < b) &\approx \Phi(b-0.5) - \Phi(a-0.5) \\
 p(a < x \leq b) &\approx \Phi(b-0.5) - \Phi(a+0.5)
 \end{aligned} \tag{7.29}$$

En la anterior exposición se partió de la suposición de que  $x$  es binomial, pero, ¿que sucedería si  $x$  no fuera binomial sino hipergeométrica? También se puede aproximar una variable hipergeométrica a través de una normal. La diferencia sería que en el caso de la hipergeométrica su media es la misma  $np$ ; sin embargo, su varianza debe multiplicarse por un factor de corrección, como se muestra:

$$\sigma_x^2 = np(1-p) \left( \frac{N-n}{N-1} \right)$$

En la figura 5.6 de la página 196 de este libro se muestran diversos ejemplos de distribuciones de probabilidad hipergeométricas y claramente se pueden apreciar aquellas que tienen una forma parecida a una normal.

## Ejercicio 7.8

Suponga que se lanza repetidamente un dado homogéneo de seis caras 6000 veces.

- a) Calcule la probabilidad de que el número de veces que cae el número tres en la cara superior del dado está comprendido entre 900 y 1050.

La probabilidad de que caiga el tres en la cara superior del dado al efectuar un lanzamiento es  $p=1/6=0.166667$ , suponiendo que el dado es homogéneo y no está cargado. Cada lanzamiento del dado es estadísticamente independiente de los demás lanzamientos. El dado se lanza  $n=6000$ . El número de veces que un dado puede ser lanzado es infinito, por lo cual, la función de probabilidad de este evento es binomial.

Utilizando R:

$$p(900 \leq x \leq 1050) = \text{pbinom}(1050, 6000, 0.166667, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pbinom}(899, 6000, 0.166667, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.958999$$

- b) Estime la probabilidad de que el número de veces que cae el número tres en la cara superior del dado está comprendido entre 900 y 1050, utilizando para ello a la distribución normal y obtenga el error absoluto y el error relativo que se cometen por esta aproximación.

$$\text{La media sería } np = 6000(0.166667) = 1000$$

La varianza sería  $np(1-p) = 1000(1-0.166667) = 833.33$ , por lo que su desviación estándar sería 28.8675.

Utilizando R:

$$p(900 \leq x \leq 1050) = \text{pnorm}(1050.5, \text{mean}=1000, \text{sd}=28.8675, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pnorm}(899.5, \text{mean}=1000, \text{sd}=28.8675, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.959637$$

$$\text{Error absoluto} = |\text{valor real} - \text{valor aproximado}| = 0.000638$$

$$\text{Error relativo} = \text{Error absoluto} * 100\% / \text{valor real} = 0.07\%$$

## Ejercicio 7.9

En un lote de  $N=600$  prendas íntimas para mujer, de color rojo, existen  $D=60$  que están manchadas con un tinte indeleble, por lo cual se consideran defectuosas. Suponga que se obtiene una muestra de tamaño  $n=50$  y sea  $x$  la variable que representa el número de prendas defectuosas en la muestra.

- a) Calcule la probabilidad de que  $x$  se encuentre entre cinco y ocho inclusive, considerando el modelo real de la hipergeométrica.

Utilizando Excel:

$$p(5 \leq x \leq 8) = \text{DIST.HIPERGEOM.N}(8, 50, 60, 600, 1) - \text{DIST.HIPERGEOM.N}(4, 50, 60, 600, 1) = 0.526171$$

- b) Determine la probabilidad de que  $x$  se encuentre entre cinco y ocho inclusive, considerando el modelo aproximado de la binomial y estime el error absoluto y el error relativo que se comete con esta aproximación.

Utilizando R:

$$p(5 \leq x \leq 8) = \text{pbinom}(8, 50, 0.1, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{pbinom}(4, 50, 0.1, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.510934$$

$$\text{Error absoluto} = |\text{valor real} - \text{valor aproximado}| = 0.015237$$

$$\text{Error relativo} = \text{Error absoluto} * 100\% / \text{valor real} = 2.90\%$$

- c) Determine la probabilidad de que  $x$  se encuentre entre cinco y ocho inclusive, considerando el modelo aproximado de la Poisson y estime el error absoluto y el error relativo que se comete con esta aproximación.

Utilizando R:

$$p(5 \leq x \leq 8) = \text{ppois}(8, 5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - \text{ppois}(4, 5, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.491413$$

$$\text{Error absoluto} = |\text{valor real} - \text{valor aproximado}| = 0.034758$$

$$\text{Error relativo} = \text{Error absoluto} * 100\% / \text{valor real} = 6.61\%$$

- d) Determine la probabilidad de que  $x$  se encuentre entre cinco y ocho inclusive, considerando el modelo aproximado de la Normal y estime el error absoluto y el error relativo que se comete con esta aproximación.

Utilizando R:

```
p(5≤x≤8)=
pnorm(8.5,mean=5,sd=2.01414,lower.tail=TRUE,log.p=FALSE)-
pnorm(4.5,mean=5,sd=2.01414,lower.tail=TRUE,log.p=FALSE)= 0.556897
```

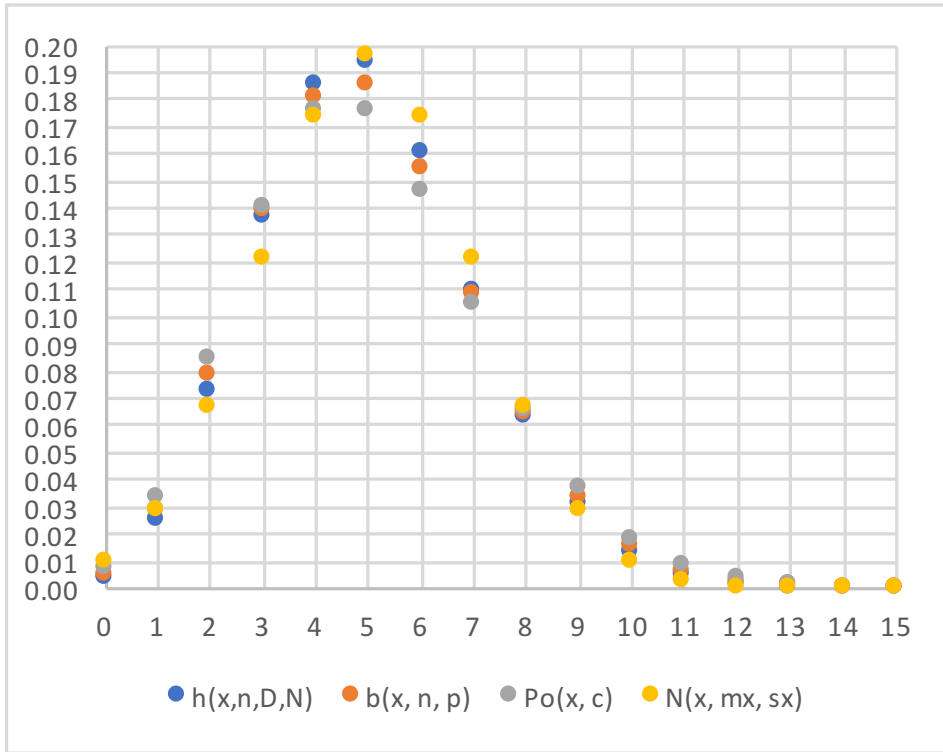
Error absoluto=|valor real-valor aproximado|= 0.030726

Error relativo= Error absoluto \* 100%/valor real= 5.84%

- e) Grafique la función de probabilidad de  $x$ , suponiendo que tiene distribución hipergeométrica, binomial, de Poisson o Normal.

En la figura 7.11 se muestran las gráficas de cada una de las distribuciones de probabilidad graficadas en Excel para este ejemplo. La primera corresponde con el modelo real que es hipergeométrico con parámetros  $N=600$ ,  $D=60$  y  $n=50$ . La segunda corresponde con un modelo aproximado a una binomial de parámetros  $n=50$  y  $p=D/N=60/600=0.1$ . La tercera corresponde con un modelo de Poisson con  $c=np=50(0.1)=5$ . La cuarta es un modelo normal con media  $\mu_x=np=5$  y varianza  $\sigma_x^2=np(1-p)(N-n)/(N-1)=5(0.9)(550)/599=2.01414$ .

FIGURA 7.11. Comparación de funciones de probabilidad discretas



Teorema del límite central para una suma de variables aleatorias con la misma distribución de probabilidad.

Teorema: Sean  $n$  variables aleatorias estadísticamente independientes,  $x_1, x_2, \dots, x_n$ , con medias  $\mu_1, \mu_2, \dots, \mu_n$  y con desviaciones estándar  $\sigma_1, \sigma_2, \dots, \sigma_n$ , con la misma distribución de probabilidad, sin importar de que distribución de probabilidad se trate. Sea  $y = x_1 + x_2 + \dots + x_n$ , entonces,

$$\text{Lim}_{n \rightarrow \infty} [y] \approx N(\mu_y, \sigma_y) \quad (7.30)$$

donde

$$\begin{aligned} \mu_y &= \mu_1 + \mu_2 + \dots + \mu_n \\ \sigma_y^2 &= \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2 \end{aligned}$$

Teorema del límite central para una combinación lineal de variables aleatorias con cualquier distribución de probabilidad

Teorema: Sean  $n$  variables aleatorias estadísticamente independientes,  $x_1, x_2, \dots, x_n$ , con medias  $\mu_1, \mu_2, \dots, \mu_n$  y con desviaciones estándar  $\sigma_1, \sigma_2, \dots, \sigma_n$ , con cualquier distribución de probabilidad.

Sea  $y = a_1 x_1 + a_2 x_2 + \dots + a_n x_n$ , entonces,

$$\lim_{n \rightarrow \infty} [y] \approx N(\mu_y, \sigma_y) \quad (7.31)$$

donde

$$\mu_y = a_1 \mu_1 + a_2 \mu_2 + \dots + a_n \mu_n$$

$$\sigma_y^2 = a_1^2 \sigma_1^2 + a_2^2 \sigma_2^2 + \dots + a_n^2 \sigma_n^2$$

Para una demostración formal de los diversos enunciados del teorema del límite central consultar el artículo *Teorema del Límite Central* de Jimena Blaio-tta y Pablo Delieutraz de la Universidad de Buenos Aires, 2004, en la siguiente dirección electrónica:

[https://www.u-cursos.cl/ingenieria/2009/2/MA3401/1/material\\_docente/bajar?id\\_material=260765](https://www.u-cursos.cl/ingenieria/2009/2/MA3401/1/material_docente/bajar?id_material=260765)

#### 7.4. Distribución de probabilidad normal de la media aritmética (promedio) de una muestra

Dado que la media aritmética de una muestra se calcula como una suma de variables aleatorias idénticamente distribuidas, se puede afirmar que la media de la media muestral es su media poblacional como a continuación se muestra:

$$\begin{aligned}\bar{x} &= \frac{1}{n} [x_1 + x_2 + \dots + x_n] \\ \mu_{\bar{x}} &= E\{\bar{x}\} = E\left\{\frac{1}{n} [x_1 + x_2 + \dots + x_n]\right\} \\ \mu_{\bar{x}} &= \frac{1}{n} [E\{x_1\} + E\{x_2\} + \dots + E\{x_n\}] \\ \mu_{\bar{x}} &= \frac{1}{n} [n\mu_x] = \mu_x\end{aligned}\tag{7.32}$$

Lo mismo se puede deducir para el caso de la varianza de la media aritmética de una muestra:

$$\begin{aligned}\sigma_{\bar{x}}^2 &= \text{var}(\bar{x}) = \text{var}\left\{\frac{1}{n} [x_1 + x_2 + \dots + x_n]\right\} \\ \sigma_{\bar{x}}^2 &= \frac{1}{n^2} [\text{var}(x_1) + \text{var}(x_2) + \dots + \text{var}(x_n)] \\ \sigma_{\bar{x}}^2 &= \frac{1}{n^2} [n\sigma_x^2] = \frac{\sigma_x^2}{n} \\ \sigma_{\bar{x}} &= \frac{\sigma_x}{\sqrt{n}}\end{aligned}\tag{7.33}$$

Por el teorema de aditividad de la distribución normal, se puede afirmar que si una población presenta distribución normal, entonces la media aritmética de una muestra de dicha población también presenta distribución normal. De la misma forma, por el Teorema del Límite central, se puede afirmar que si se extrae una muestra de tamaño grande de una población con cualquier distribución, la media muestral tenderá a tener distribución normal en la medida en que  $n$  tienda a infinito.

Esto se puede resumir de la siguiente forma:

Sea

$$\bar{x} = \frac{1}{n} [x_1 + x_2 + \dots + x_n]$$

Entonces

$$\bar{x} \sim N\left(\mu_x, \frac{\sigma_x}{\sqrt{n}}\right) \quad (7.34)$$

Bajo uno de los dos supuestos siguientes:  $x$  es normal o  $n$  tiende a ser muy grande (tiende a infinito).

En la práctica profesional de la Ingeniería, el pensar en una muestra de tamaño infinito es ilusorio; en realidad, los lotes presentan tamaño finito y las muestras con mayor razón; surge entonces la pregunta, ¿de qué tamaño debe ser la muestra para suponer que la media muestral de una población con distribución diferente a la normal es normal? La respuesta desde luego depende de la distribución real que presenten los datos. Algunas reglas empíricas o “de dedo”, establecen experimentalmente lo siguiente:

- i. Si la distribución de probabilidad de una población es parecida a una normal, es decir, si la gráfica de la función de probabilidad de una población presenta una forma parecida al perfil de una campana casi simétrica basta tomar  $n \geq 4$ .
- ii. Si la distribución de probabilidad de una población es parecida a una distribución uniforme, basta tomar  $n \geq 12$ .
- iii. Si la distribución de probabilidad de una población presenta la mayor parte de sus medidas en sus extremos (cóncava hacia arriba y no simétrica), por ejemplo, la distribución exponencial negativa, a lo cual se le podría denominar comportamiento antinormal,  $n \geq 100$ .

#### Ejercicio 7.10

Un autobús de segunda clase viaja entre la Ciudad de México y la Ciudad de Cosamaloapan Veracruz, pero se detiene en promedio 15 minutos con una desviación estándar de tres minutos en las terminales de seis ciudades intermedias en su ruta. La media y la desviación estándar de los tiempos de viaje entre cada ciudad intermedia se muestran en la siguiente tabla.



Pares de Ciudades	Tiempo promedio de viaje (min)	Desv Est (min)
México-Puebla	135	15
Puebla-Orizaba	130	12
Orizaba-Córdoba	35	3
Córdoba-Tinajas	40	5
Tinajas-Tierra Blanca	29	6
Tierra Blanca-Ciudad Alemán	50	4
Ciudad Alemán-Tres Valles	15	2
Tres Valles-Cosamaloapan	58	7

- a) ¿Qué distribución de probabilidad presenta el tiempo total de viaje? Fundamente su respuesta. ¿Cómo se comprobaría?

$$T = t_1 + t_2 + t_3 + t_4 + t_5 + t_6 + t_7 + t_8 + 6t_9$$

En donde los  $t_i$ ,  $i=1,2,3,4,5,6,7,8$ , representan a los tiempos de viaje entre ciudades intermedias y  $t_9$  representa el tiempo de espera en cada ciudad intermedia.

El tiempo total de viaje es la suma de los ocho tiempos de viaje entre las ciudades intermedias más seis tiempos de espera en dichas ciudades; esto implica que se trata de una suma de 14 tiempos de viaje y de espera. Si se supone que el tiempo de viaje entre las ciudades intermedias y el tiempo de espera en cada ciudad intermedia se comporta como una función de probabilidad parecida a una normal, o es más, parecida a una distribución uniforme, dado que  $n > 12$ , se puede suponer que la distribución de probabilidad del tiempo total de viaje es una normal.

Para comprobarlo se tendrían que obtener muestras periódicas de tiempos de viaje entre ciudades y determinar si con estas muestras se puede determinar que cada tiempo de viaje presenta un comportamiento probabilístico parecido a una normal o a una uniforme.

- b) Obtenga la media y la desviación estándar del tiempo total de viaje.

$$\mu_T = 135 + 130 + 35 + 40 + 29 + 50 + 6(15) = 509 \text{ min}$$

$$\sigma_T^2 = 15^2 + 12^2 + 3^2 + 5^2 + 6^2 + 4^2 + 2^2 + 7^2 + 6^2(3^2) = 832 \text{ min}^2$$

$$\sigma_T = 28.8444 \text{ min}$$

- c) Calcule la probabilidad de que el autobús complete su jornada en un máximo de ocho horas (480 min).

$$p(x \leq 480) = \\ \text{pnorm}(480, \text{mean}=509, \text{sd}=28.8444, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = \\ 0.157354 = 15.734\%$$

- d) Si la línea de autobuses afirma que el tiempo total de viaje es menor a 7 horas, ¿cuáles serían sus conclusiones sobre esta afirmación?

$$p(x \leq 420) = \\ \text{pnorm}(420, \text{mean}=509, \text{sd}=28.8444, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = \\ 0.001016$$

Rechazaría esa afirmación, ya que la probabilidad de que llegue antes de siete horas de viaje es menor al 1%.

- e) ¿Entre qué tiempo mínimo y qué tiempo máximo, ambos centrados alrededor de la media, se tendría un 95% de probabilidad de llegar?

$$p(\mu_T - k\sigma_T \leq T \leq \mu_T + k\sigma_T) = 0.95$$

Estandarizando:

$$p(-k \leq z \leq k) = 0.95 \quad \rightarrow \quad 2p(0 \leq z \leq k) = 0.95 \quad \rightarrow \quad p(z \leq k) = 0.975$$

$$T_{\max} = \text{qnorm}(0.975, \text{mean}=509, \text{sd}=28.8444, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 565.534 \text{ min}$$

$$T_{\min} = \text{qnorm}(0.025, \text{mean}=509, \text{sd}=28.8444, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 452.466 \text{ min}$$

El tiempo de viaje estaría entre 7.54 y 9.42 horas al 95% de probabilidad.

## 7.5. Distribución de probabilidad de la varianza de una muestra

### Distribución de Probabilidad Ji Cuadrada

Sean  $z_1, z_2, \dots, z_k$ , variables aleatorias normales estándar, estadísticamente independientes, es decir,

$$z_i \approx N(0,1) \quad i = 1, 2, \dots, k$$

Sea

$$\chi^2 = z_1^2 + z_2^2 + \dots + z_k^2 \quad (7.35)$$

Cabe señalar que en la expresión anterior  $\chi^2$  (ji cuadrada) representa a una variable, no es que esté elevada al cuadrado, es solo notación, se indica así solo para recordar que se trata de una suma de cuadrados.

La función de probabilidad de esta variable aleatoria ji cuadrada es:

$$f(\chi^2) = \frac{1}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} (\chi^2)^{\frac{k}{2}-1} e^{-\frac{\chi^2}{2}} \quad \chi^2 > 0 \quad (7.36)$$

Dónde  $\Gamma(k/2)$  está definida en la expresión (6.22) como la Función Gamma o Función Factorial Generalizada, la cual cumple las propiedades establecidas en las expresiones (6.23), (6.24) y (6.25).

Para deducir la expresión (7.36) primero se deducirá la función de probabilidad que presenta una variable aleatoria normal estándar al cuadrado y luego se deducirá la función de probabilidad de una suma de  $k$  variables aleatorias normales estándar elevadas al cuadrado.

Sea  $z \sim N(0,1)$ , lo que implica que

$$f_z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

y sea  $u = z^2$  entonces  $z = \pm \text{raíz}(u)$

$$\left| \frac{dz}{du} \right| = \left| \pm \frac{1}{2\sqrt{u}} \right| = \frac{1}{2\sqrt{u}}$$

La función de probabilidad de  $u$  estaría dada por:

$$\begin{aligned} F_u(u) &= p(U \leq u) = p(z^2 \leq u) \\ &= p(|z| \leq \sqrt{u}) = F_u(\sqrt{u}) - F_u(-\sqrt{u}) \\ f_u(u) &= 2f_z(z) \left| \frac{dz}{du} \right| = \frac{1}{\sqrt{2\pi}} e^{-\frac{u}{2}} \frac{1}{\sqrt{u}} \end{aligned}$$

Pero

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

Entonces

$$f_u(u) = \frac{1}{2^{\frac{1}{2}} \Gamma\left(\frac{1}{2}\right)} e^{-u} u^{-\frac{1}{2}} \quad (7.37)$$

A la cual se le denomina función densidad de probabilidad ji cuadrada con un grado de libertad.

Si se compara la expresión anterior (7.37) con la expresión (6.21) de la página 268, correspondiente a una función de probabilidad tipo gamma, se puede apreciar claramente que la expresión (7.37) obtenida anteriormente corresponde con una gamma con  $\alpha=1/2$  y  $\beta=1/2$ .

Una propiedad básica de la función de probabilidad gamma es que cumple la propiedad de aditividad, es decir, la suma de  $k$  variables aleatorias tipo gamma con el mismo parámetro  $\alpha$  da a su vez una variable tipo gamma con parámetro  $k\alpha$ , esto implica que la variable

$$v = u_1 + u_2 + \dots + u_k$$

posee distribución gamma con parámetros  $\alpha = k/2$  y  $\beta=1/2$ , es decir, tiene la distribución que se presenta en la expresión (7.36).

La función de probabilidad acumulada ji cuadrada está dada por la expresión:

$$F(\chi^2) = \frac{1}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} \int_0^{\chi^2} u^{-\frac{k}{2}-1} u^{-\frac{u}{2}} du = \frac{\gamma\left(\frac{k}{2}, \frac{\chi^2}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} \quad (7.38)$$

Donde  $\gamma(a,b)$  representa a la función gamma incompleta inferior.

Su función generadora de momentos se expresa como:

$$FGM_{\chi^2}(t) = (1-2t)^{-\frac{k}{2}} \quad t < 1/2 \quad (7.39)$$

Su función característica es:

$$FC_{\chi^2}(t) = (1-2it)^{-\frac{k}{2}} \quad t < 1/2 \quad (7.40)$$

La media y la varianza de la función de probabilidad ji cuadrada son:

$$\mu_{\chi^2} = k \quad (7.41)$$

$$\sigma_{\chi^2}^2 = 2k \quad (7.42)$$

Su coeficiente de asimetría y su coeficiente de curtosis son:

$$\gamma_{1\chi^2} = \sqrt{\frac{8}{k}} \quad (7.43)$$

$$\gamma_{2\chi^2} = \frac{12}{k} \quad (7.44)$$

Para hacer el cálculo de probabilidades con la distribución ji cuadrada, utilizando Excel, se utiliza el siguiente comando:

$$p(\chi^2, k) = \text{DISTR.CHICUAD}(\chi^2, k, \text{acumulado}) \quad (7.45)$$

En donde  $\chi^2$  es el valor de la variable ji cuadrada para el cual se pretende evaluar su probabilidad,  $k$  es el número de grados de libertad o sea el número de variables aleatorias normales estándar estadísticamente independientes que se están sumando, acumulado= verdadero o 1 si se pretende calcular la función

de probabilidad acumulada desde cero hasta  $\chi^2$ ; acumulado= falso o 0 si se pretende calcular el valor de la función de densidad para un valor  $\chi^2$ .

Cabe señalar que Excel también maneja otro comando que proporciona la probabilidad de que ji cuadrada sea mayor que  $\chi^2$ , es decir proporciona la probabilidad a la derecha del valor  $\chi^2$ .

$$p(\chi^2)=\text{DISTR.CHICUAD.CD}(\chi^2, k, \text{acumulado}) \quad (7.46)$$

De tal manera que:

$$p(\chi^2=5, k=10)= \text{DISTR.CHICUAD}(5, 10, 0)= 0.066801$$

Esta no es una probabilidad, es el valor de la función de densidad en el punto  $x=5$  para 10 grados de libertad.

$$p(\chi^2 \leq 5, k=10)= \text{DISTR.CHICUAD}(5, 10, 1)= 0.108822$$

o también

$$p(\chi^2 > 5, k=10)= 1-\text{DISTR.CHICUAD.CD}(5, 10)= 0.108822$$

$$p(\chi^2 > 5; k= 10)= \text{DISTR.CHICUAD.CD}(5, 10)= 0.891178$$

$$p(4 \leq \chi^2 \leq 6; k= 10)= \text{DISTR.CHICUAD}(6, 10) - \text{DISTR.CHICUAD}(4, 10)= 0.132084$$

o también

$$p(4 \leq \chi^2 \leq 6, k=10)= \text{DISTR.CHICUAD.CD}(4, 10) - \text{DISTR.CHICUAD.CD}(6, 10)= 0.132084$$

También se pueden obtener valores de la ji cuadrada para probabilidades dadas, a través del siguiente comando:

$$\chi^2_{p_0, k} = \text{INV.CHICUAD}(p_0, k) \quad (7.47)$$

El cual proporciona el valor de ji cuadrada para el cual la probabilidad a la izquierda de él es  $p_0$ , que también puede ser estimado a través de la siguiente expresión:

$$\chi^2_{p_0, k} = \text{INV.CHICUAD.CD}(1-p_0, k) \quad (7.48)$$

Como se puede apreciar a continuación:

Obtener el valor de  $\chi^2$  para el cual  $p(\chi^2 \leq \chi_0^2) = 0.90$  con  $k = 10$  grados de libertad:

$$\chi^2_{0.9,10} = \text{INV.CHICUAD}(0.9, 10) = 15.987$$

O también

$$\chi^2_{0.9,10} = \text{INV.CHICUAD.CD}(0.1, 10) = 15.987$$

$$P(\chi^2_{0.05,10} \leq \chi^2 \leq \chi^2_{0.95,10}) = 0.90$$

En este último caso se calculan dos valores:

$$\chi^2_{0.9,10} = \text{INV.CHICUAD}(0.95, 10) = 18.307$$

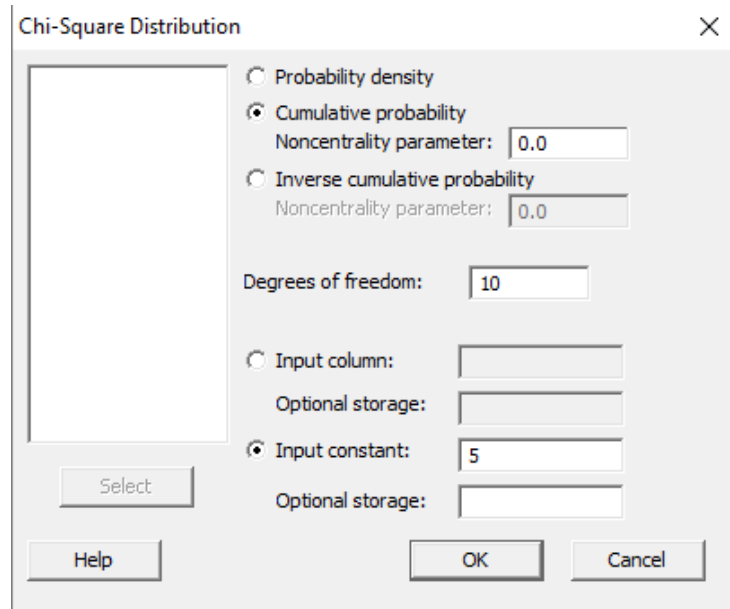
$$\chi^2_{0.05,10} = \text{INV.CHICUAD}(0.05, 10) = 3.94$$

Para hacer el cálculo de probabilidades con la distribución ji cuadrada, utilizando Minitab, se realizan los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en Chi-Square, como se muestra en la figura 7.4. Cabe señalar que la letra griega  $\chi$  en inglés se pone como chi, pero en español es ji y corresponde con la letra c.
2. Aparece la pantalla que se muestra en la figura 7.12, en la cual, para calcular la probabilidad  $p(\chi^2 \leq 5, k = 10)$  se señala Cumulative probability, en Degrees of freedom se indica 10 y en input constant se indica el valor de ji cuadrada para el cual se desea calcular la probabilidad y se da un click en Ok obteniéndose el resultado ya obtenido previamente:

$$p(\chi^2 \leq 5, k = 10) = 0.108822$$

FIGURA 7.12.



En el lenguaje de desarrollo R, para el caso de la distribución ji cuadrada, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo de la función de densidad exactamente en un punto:

$$dchisq(x, df, ncp=0, log = FALSE) \quad (7.49)$$

Para el cálculo de la probabilidad acumulada hasta un punto dado:

$$pchisq(x, df, ncp=0, lower.tail = TRUE, log.p = FALSE) \quad (7.50)$$

Para el cálculo de los percentiles, deciles o cuartiles se usa la expresión:

$$qchisq(q, df, ncp= 0, lower.tail= TRUE, log.p= FALSE) \quad (7.51)$$

Para generar  $n$  números aleatorios con distribución ji cuadrada con  $k$  grados de libertad se usa la expresión:

$$rchisq(n, df, ncp=0) \quad (7.52)$$

En donde  $x$  es el valor que toma la variable aleatoria ji cuadrada,  $df$  es el número de grados de libertad,  $ncp$  es el parámetro de no centralidad que para el caso



que se analiza debe ser cero porque tal como fue definida como una suma de cuadrados no puede tomar valores negativos, `log` es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma `FALSE`. `lower.tail` es un valor lógico; si se le asigna un valor verdadero, el valor que arroja es la probabilidad acumulada a la izquierda de  $x$ , de lo contrario, si el valor que toma es Falso, el valor que arroja es la probabilidad a la derecha del valor de  $x$ .

Por ejemplo,

$$f(x=5, k=10)=dchisq(5,10, ncp=0, log=FALSE)= 0.066801$$

Esta no es una probabilidad, es el valor de la función de densidad en el punto  $x=5$  para 10 grados de libertad.

$$p(\chi^2 \leq 5, k=10)= pchisq(5, 10, ncp=0, lower.tail=TRUE, log.p=FALSE)= 0.108822$$

$$p(\chi^2 > 5; k= 10)= pchisq(5, 10, ncp=0, lower.tail=FALSE, log.p=FALSE)= 0.891178$$

$$p(4 \leq \chi^2 \leq 6, k=10)= pchisq(6, 10, ncp=0, lower.tail=TRUE, log.p=FALSE)-pchisq(4, 10, ncp=0, lower.tail=TRUE, log.p=FALSE)= 0.132084$$

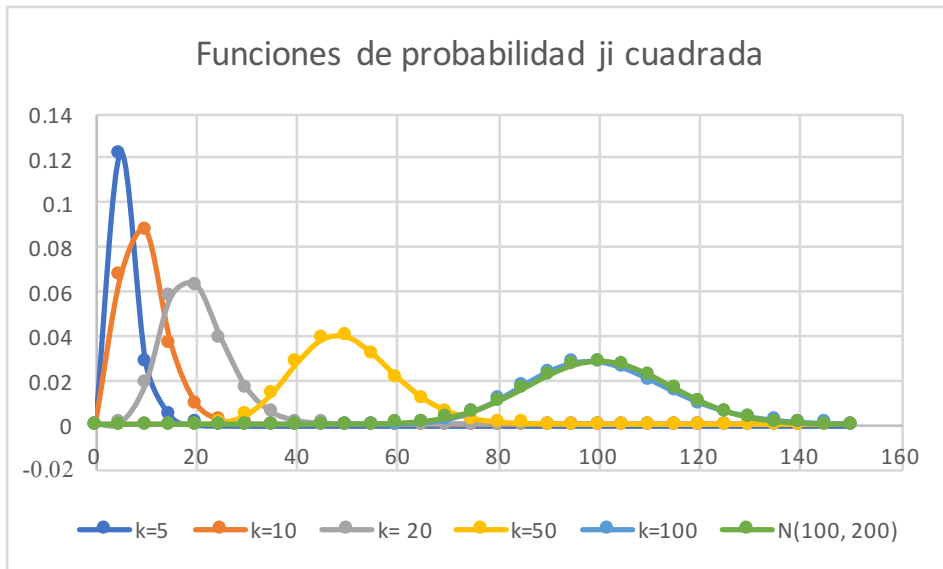
$$\chi^2_{0.9,10}= qchisq(0.9, 10, ncp=0, lower.tail=TRUE, log.p=FALSE)= 15.987$$

Generar cinco números aleatorios con distribución ji cuadrada con 10 grados de libertad:

$$rchisq(5,10,ncp= 0)= 4.677 \quad 9.179 \quad 15.376 \quad 7.076 \quad 13.697$$

En la figura 7.11 se muestran algunas gráficas de la distribución Ji Cuadrada para los siguientes grados de libertad  $k=5, 10, 20, 50, 100$ . Nótese que en la medida que el número de grados de libertad crece la gráfica de la ji cuadrada se aproxima más a una normal, lo que refleja el cumplimiento del Teorema del Límite Central. La última gráfica que se muestra en color verde corresponde con una normal de media  $k=100$  y varianza  $k=200$ , nótese como esta es muy parecida a una ji cuadrada con 100 grados de libertad.

FIGURA 7.13. Gráficas de la distribución ji cuadrada



### 7.5.1. Teorema de aditividad de la distribución ji cuadrada

Sean

$$\chi_{1, k_1}^2, \chi_{2, k_2}^2, \dots, \chi_{p, k_p}^2$$

$p$  variables aleatorias con distribución ji cuadrada, estadísticamente independientes con  $k_1, k_2, \dots, k_p$  grados de libertad respectivamente. Entonces, la suma

$$y = \chi_{1, k_1}^2 + \chi_{2, k_2}^2 + \dots + \chi_{p, k_p}^2 \quad (7.53)$$

También presenta distribución ji cuadrada con  $k_y = k_1 + k_2 + \dots + k_p$  grados de libertad.

Prueba: Por definición, cada una de las variables aleatorias ji cuadrada es una suma de  $k_i$  variables aleatorias normales estándar elevadas al cuadrado, es decir

$$\chi_{i, k_j}^2 = \sum_{j=1}^{j=k} z_{ij}^2 \quad i = 1, 2, \dots, p$$

Por lo tanto

$$y = \sum_{i=1}^{i=p} \sum_{j=1}^{j=k_j} z_{ij}^2$$

Dado que todas las variables aleatorias  $z_{ij}$  son estadísticamente independientes porque las  $\chi_i^2$  son estadísticamente independientes, la variable aleatoria  $y$  es exactamente la suma de los cuadrados de  $k_y = k_1 + k_2 + \dots + k_p$  variables aleatorias normales estándar estadísticamente independientes, que por la definición de ji cuadrada también se comporta como una variable aleatoria ji cuadrada con  $k_y$  grados de libertad.

La importancia del comportamiento probabilístico de una variable aleatoria ji cuadrada se debe a que como se verá a continuación el parámetro estadístico denominado varianza de una muestra se comporta como una ji cuadrada.

La varianza de una muestra se define de la siguiente forma:

$$S_{n-1}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

En donde se supondrá que cada elemento de la muestra  $x_i$  pertenece a una población normal con media  $\mu_x$  y varianza  $\sigma_x^2$  o en su defecto el tamaño de muestra se considerará muy grande con el objeto de aplicar el teorema de aditividad de la distribución normal o el teorema del límite central. Si en la expresión anterior se multiplica de ambos lados por  $(n-1)$ , por aproximación heurística se sustituye la media muestral por la media poblacional, y se divide entre la varianza poblacional, se obtiene que:

$$\frac{(n-1)S_{n-1}^2}{\sigma_x^2} = \sum_{i=1}^n \left( \frac{x_i - \mu_x}{\sigma_x} \right)^2 = \sum_{i=1}^n z_i^2$$

Como se puede apreciar, se trata de una suma de cuadrados de variables aleatorias normales estándar, por lo que se puede afirmar que el estadístico correspondiente al primer miembro de la expresión anterior es una variable aleatoria ji cuadrada; aunque no todos los sumandos son estadísticamente independientes, por la propia definición de varianza muestral.

Como se mostrará a continuación los grados de libertad de esta variable aleatoria son  $n-1$ .

$$\begin{aligned} \sum_{i=1}^{i=n} (x_i - \mu_x)^2 &= \sum_{i=1}^n (x_i - \bar{x} + \bar{x} - \mu_x)^2 \\ &= \sum_{i=1}^n (x_i - \bar{x})^2 + \sum_{i=1}^n (\bar{x} - \mu_x)^2 + 2(\bar{x} - \mu_x) \sum_{i=1}^n (x_i - \bar{x}) \\ &= \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \mu_x)^2 \end{aligned}$$

Por lo anterior y a partir de la definición de varianza muestral:

$$\frac{\sum_{i=1}^{i=n} (x_i - \mu_x)^2}{\sigma_x^2} = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{\sigma_x^2} + \frac{(\bar{x} - \mu_x)^2}{\frac{\sigma_x^2}{n}}$$

Por el teorema de aditividad de la distribución normal o por el teorema del límite central la media muestral presenta distribución normal de media  $\mu_x$  y varianza  $\sigma_x^2/n$ , lo que implica que el segundo término del segundo miembro de la expresión anterior presenta distribución ji cuadrada con un grado de libertad y dado que el primer miembro presenta distribución ji cuadrada con  $n$  grados de libertad, por el teorema de aditividad de la distribución ji cuadrada, el primer término del segundo miembro debe presentar distribución ji cuadrada con  $n-1$  grados de libertad, lo que se representa como:

$$\frac{(n-1)S_{n-1}^2}{\sigma_x^2} \approx \chi_{n-1}^2 \tag{7.54}$$

Textualmente esto significa que si se multiplica la varianza muestral por  $(n-1)$  y se divide entre la varianza poblacional, la variable que resulta es una ji cuadrada con  $n-1$  grados de libertad, resultado teórico muy importante para las aplicaciones que se verán posteriormente.

## 7.6. Distribución de probabilidad de la media de una muestra para una población con varianza desconocida

### Distribución de probabilidad t de Student

Sean  $z \sim N(0,1)$ , una variable aleatoria normal estándar, y  $v \sim \chi^2 k$  una variable aleatoria ji cuadrada con  $k$  grados de libertad, ambas estadísticamente independientes. Sea la variable aleatoria

$$t = \frac{Z}{\sqrt{\frac{V}{k}}} \quad (7.55)$$

A la variable aleatoria “t” se le denomina t de Student, por el pseudónimo con que firmaba su autor William Sealy Gosset (1876-1937) ya que trabajaba en las destilerías Guinness en Dublín, la cual le prohibía a sus empleados publicar artículos independientemente de la información que contuviesen, debido a que otro investigador de Guinness había publicado anteriormente un artículo que contenía secretos industriales de la destilería.



**FIGURA 7.14.** William Sealy Gosset (1876–1937)

**NOTA(2020)**

[https://es.wikipedia.org/wiki/William\\_Sealy\\_Gosset](https://es.wikipedia.org/wiki/William_Sealy_Gosset)

La función de densidad de probabilidad de la variable aleatoria t de Student o de Gosset está dada por la expresión:

$$f(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi k} \Gamma\left(\frac{k}{2}\right)} \frac{1}{\left[\frac{t^2}{k} + 1\right]^{\frac{(k+1)}{2}}} \quad t \in \mathfrak{R} \quad (7.56)$$

donde como ya se dijo a la expresión  $\Gamma(k/2)$  se le conoce como Función Gamma o Función Factorial Generalizada.

### Demostración

Dado que  $z$  y  $v$  son estadísticamente independientes, su función de densidad conjunta es:

$$f(z, v) = g_z(z) h_v(v) = \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}} \frac{v^{\frac{k}{2}-1} e^{-\frac{v}{2}}}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)}$$

Por otra parte, del procedimiento establecido en el subtema 4.3, página 162, se definen las variables:

$$t = \frac{z}{\sqrt{\frac{v}{k}}}$$

$$s = v$$

Obsérvese que  $z \in \mathfrak{R}$ ,  $v \in \mathfrak{R}^+$  por ser una suma de cuadrados, lo que implica que  $t \in \mathfrak{R}$ ,  $s \in \mathfrak{R}^+$

De las expresiones anteriores se despejan  $z$  y  $v$ :

$$z = t \sqrt{\frac{s}{k}}$$

$$v = s$$

y se sustituyen en la función de probabilidad conjunta:

$$f(t, s) = \frac{e^{-\frac{t^2 s}{2k}} s^{\frac{k}{2}-1} e^{-\frac{s}{2}}}{\sqrt{2\pi} 2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)}$$

Se obtiene el Jacobiano de la Transformación:

$$J \begin{pmatrix} z \\ t \end{pmatrix} \begin{pmatrix} v \\ s \end{pmatrix} = \begin{vmatrix} \frac{\partial z}{\partial t} & \frac{\partial z}{\partial s} \\ \frac{\partial v}{\partial t} & \frac{\partial v}{\partial s} \end{vmatrix} = \begin{vmatrix} \sqrt{\frac{s}{k}} & \frac{t}{2\sqrt{ks}} \\ 0 & 1 \end{vmatrix} = \sqrt{\frac{s}{k}}$$

Por lo cual la función de probabilidad conjunta de  $t$  y  $s$  está definida por la expresión:

$$f(t,s) = \frac{e^{-\frac{t^2}{2k}}}{\sqrt{2\pi}} \frac{s^{\frac{k}{2}-1}}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} e^{-\frac{s}{2}} \left| J \begin{pmatrix} z \\ t \end{pmatrix} \begin{pmatrix} v \\ s \end{pmatrix} \right| = \frac{1}{\sqrt{2\pi k}} \frac{s^{\frac{k}{2}-1}}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} e^{-\frac{s}{2}} e^{-\frac{t^2}{2k}}$$

De donde integrando con respecto a  $s$ , desde cero a infinito:

$$f(t) = \int_0^{\infty} f(t,s) ds = \frac{1}{\sqrt{2\pi k}} \frac{s^{\frac{k}{2}-1}}{2^{\frac{k}{2}} \Gamma\left(\frac{k}{2}\right)} \int_0^{\infty} e^{-\frac{s}{2}} e^{-\frac{t^2}{2k}} ds$$

Aplicando las propiedades de la función Gamma se obtiene la expresión (7.56), lo que queda demostrado.

La función de probabilidad acumulada  $t$  de Student está dada por la expresión:

$$F(t) = \frac{\Gamma\left(\frac{k+1}{2}\right)}{\sqrt{\pi k} \Gamma\left(\frac{k}{2}\right)} \int_{-\infty}^t \frac{du}{\left[\frac{u^2}{k} + 1\right]^{\frac{k+1}{2}}} = B_{(t+\sqrt{t^2+k})/2\sqrt{t^2+k}}\left(\frac{k}{2}, \frac{k}{2}\right) \quad (7.57)$$

Donde  $B_y(a,b)$  es la función beta incompleta como se vio anteriormente.

La media y la varianza de la función de probabilidad  $t$  de Student son:

$$\mu_t = 0 \quad k > 1 \quad (7.58)$$

$$\sigma_t^2 = \frac{k}{k-2} \quad k > 2 \quad (7.59)$$

Su coeficiente de asimetría y su coeficiente de curtosis son:

$$\gamma_{1t} = 0 \quad k > 3 \quad (7.60)$$

$$\gamma_{2t} = \frac{6}{k-4} \quad k > 4 \quad (7.61)$$

Para hacer el cálculo de probabilidades con la distribución t de Student, utilizando Excel, se utiliza el siguiente comando:

$$p(t, k) = \text{DISTR.T.N}(t, k, \text{acumulado}) \quad (7.62)$$

En donde  $t$  es el valor de la variable t de Student para el cual se pretende evaluar su probabilidad,  $k$  es el número de grados de libertad, acumulado= verdadero o 1 si se pretende calcular la función de probabilidad acumulada desde cero hasta  $t$ ; acumulado= falso o 0 si se pretende calcular el valor de la función de densidad para un valor  $t$ .

Cabe señalar que Excel también maneja otro comando que proporciona la probabilidad de que la variable sea mayor que  $t$ , es decir proporciona la probabilidad a la derecha del valor  $t$ .

$$p(t) = \text{DISTR.T.CD}(t, k) \quad (7.63)$$

De tal manera que:

$$p(t=5, k=10) = \text{DISTR.T.N}(5, 10, 0) = 0.000396$$

Esta no es una probabilidad, es el valor de la función de densidad en el punto  $t=5$  para 10 grados de libertad.

$$p(t \leq 5, k=10) = \text{DISTR.T.N}(5, 10, 1) = 0.999731$$

o también

$$p(t \leq 5, k=10) = 1 - \text{DISTR.T.CD}(5, 10) = 0.999731$$

$$p(t > 5, k=10) = \text{DISTR.T.CD}(5, 10) = 0.000269$$

$$p(4 \leq t \leq 6, k=10) = \text{DISTR.T.N}(6, 10, 1) - \text{DISTR.T.N}(4, 10, 1) = 0.001193$$



o también

$$p(4 \leq t \leq 6, k=10) = \text{DISTR.T.CD}(4, 10) - \text{DISTR.T.CD}(6, 10) = 0.001193$$

También se pueden obtener valores de la *t* de Student para probabilidades dadas, a través del siguiente comando:

$$t_{p0,k} = \text{INV.T}(p0, k) \quad (7.64)$$

El cual proporciona el valor de *t* para el cual la probabilidad a la izquierda es  $p_0$ , como se puede apreciar a continuación:

Obtener el valor de  $t_0$  para el cual  $p(t \leq t_0) = 0.90$  con  $k = 10$  grados de libertad:

$$t_{0.9,10} = \text{INV.T}(0.9, 10) = 1.372$$

$$p(t_{0.05,10} \leq t \leq t_{0.95,10}) = 0.90$$

En este último caso se calculan dos valores:

$$t_{0.95,10} = \text{INV.T}(0.95, 10) = 1.812$$

$$t_{0.05,10} = \text{INV.T}(0.05, 10) = -1.812$$

El segundo siempre será el mismo que el primero pero con signo negativo, esto se debe a la simetría de la *t* de Student con respecto al origen.

Para hacer el cálculo de probabilidades con la distribución *t* de Student, utilizando Minitab, se realizan los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en *t*, como se muestra en la figura 7.4.
2. Aparece la pantalla que se muestra en la figura 7.15, en la cual, para calcular la probabilidad  $p(t \leq 5, k=10)$  se señala Cumulative probability Noncentrality parameter 0.0, en Degrees of freedom se indica 10 y en input constant se indica el valor de *t* para el cual se desea calcular la probabilidad y se da un click en Ok obteniéndose el resultado ya obtenido previamente:

$$p(t \leq 5, k=10) = 0.999731$$

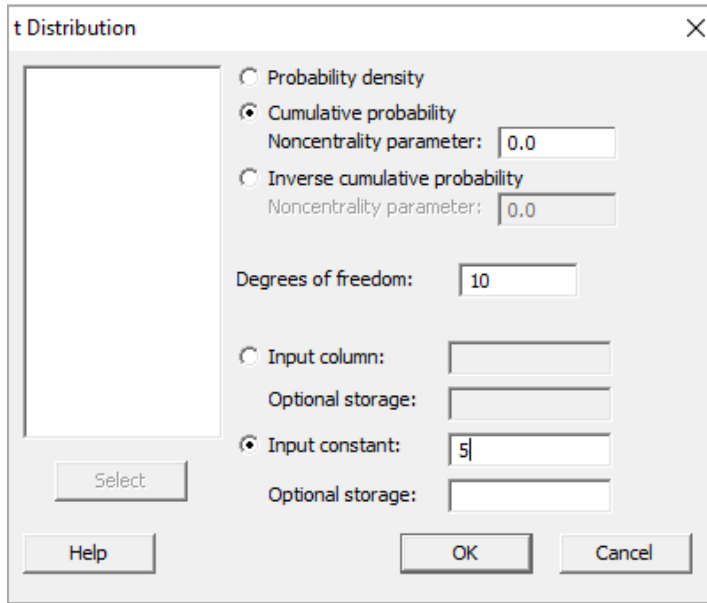


FIGURA 7.15.

En el lenguaje de desarrollo R, para el caso de la distribución *t* de Student, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo de la función de densidad exactamente en un punto:

$$dt(x, df, ncp, log = FALSE) \quad (7.65)$$

Para el cálculo de la probabilidad acumulada hasta un punto dado:

$$pt(x, df, ncp, lower.tail = TRUE, log.p = FALSE) \quad (7.66)$$

Para el cálculo de los percentiles, deciles o cuartiles se usa la expresión:

$$qt(q, df, ncp, lower.tail = TRUE, log.p = FALSE) \quad (7.67)$$

Para generar  $n$  números aleatorios con distribución *t* de Student con  $df$  grados de libertad se usa la expresión:

$$rt(n, df, ncp) \quad (7.68)$$

En donde  $x$  es el valor que toma la variable aleatoria *t* de Student,  $df$  es el número de grados de libertad,  $ncp$  es el parámetro de no centralidad (ya que puede no

estar centrada con respecto al origen), `log` es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma `FALSE`. `lower.tail` es un valor lógico; si se le asigna un valor verdadero, el valor que arroja es la probabilidad acumulada a la izquierda de  $x$ , de lo contrario, si el valor que toma es Falso, el valor que arroja es la probabilidad a la derecha del valor de  $x$ .

Por ejemplo,

$$f(t=5, k=10)=dt(5,10, 0, log=FALSE)= 0.000396$$

Esta no es una probabilidad, es el valor de la función de densidad en el punto  $t=5$  para 10 grados de libertad.

$$p(t \leq 5, k=10)= pt(5, 10, 0, lower.tail=TRUE, log.p=FALSE)= 0.999731$$

$$p(t > 5; k= 10)= pt(5, 10, 0, lower.tail=FALSE, log.p=FALSE)= 0.000269$$

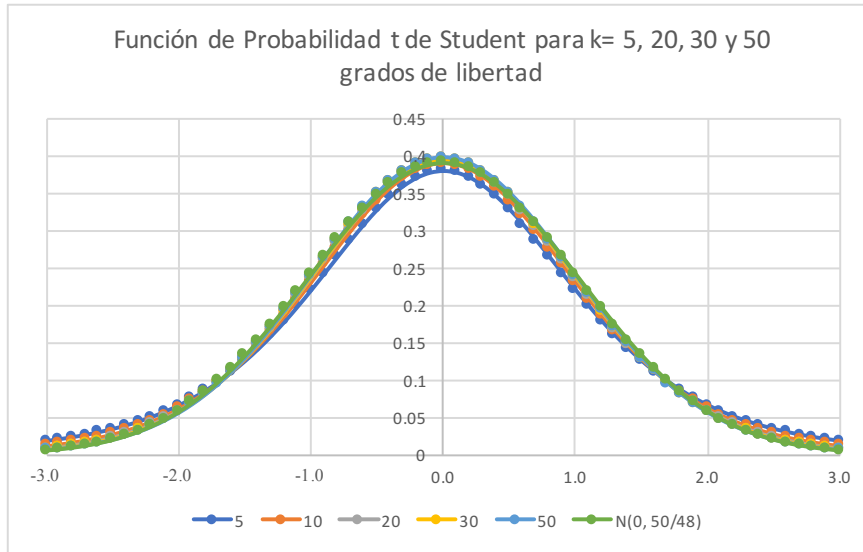
$$p(4 \leq t \leq 6, k=10)= pt(6, 10, 0, lower.tail=TRUE, log.p=FALSE) - pt(4, 10, 0, lower.tail=TRUE, log.p=FALSE)= 0.001193$$

$$t_{0.9,10}= qt(0.9, 10, 0, lower.tail=TRUE, log.p=FALSE)= 1.372$$

Generar cinco números aleatorios con distribución  $t$  de Student con 10 grados de libertad:

$$rt(5,10, 0)= \quad 0.238 \quad 0.244 \quad 2.219 \quad 0.584 \quad 0.883$$

En la figura 7.16 se muestran algunas gráficas de la distribución  $t$  de Student para los siguientes grados de libertad  $k= 5, 10, 20, 30, 50$ . Nótese que en la medida que el número de grados de libertad crece la gráfica de la  $t$  de Student se aproxima más a una normal, lo que refleja el cumplimiento del Teorema del Límite Central. La última gráfica que se muestra en color naranja corresponde con una normal de media cero y varianza  $50/48$ , nótese como ésta es muy parecida a una  $t$  de Student con 50 grados de libertad; de hecho se puede considerar que para  $k > 30$  el ajuste con la normal como aproximación de la  $t$  de Student es bastante adecuado.

**FIGURA 7.16.** Gráficas de la distribución t de Student

La importancia del comportamiento probabilístico de una variable aleatoria t de Student se debe a que, como se verá a continuación, el parámetro estadístico denominado media estandarizada de una muestra se comporta como una t de Student.

Por el teorema de aditividad de la distribución normal o por el teorema del límite central la media muestral presenta distribución normal de media  $\mu_x$  y varianza  $\sigma_x^2/n$ , lo que implica que la variable:

$$z = \frac{\bar{x} - \mu_x}{\frac{\sigma_x}{\sqrt{n}}} \approx N(0,1)$$

por otra parte, la variable:

$$\frac{(n-1)S_{n-1}^2}{\sigma_x^2} \approx \chi_{n-1}^2$$

Por definición de t de Student

$$t = \frac{z}{\sqrt{\frac{v}{k}}} = \frac{\frac{\bar{x} - \mu_x}{\frac{\sigma_x}{\sqrt{n}}}}{\sqrt{\frac{(n-1)S_{n-1}^2}{(n-1)\sigma_x^2}}} = \frac{\bar{x} - \mu_x}{\frac{S_{n-1}}{\sqrt{n}}}$$

Esto implica que la variable

$$\frac{\bar{x} - \mu_x}{\frac{S_{n-1}}{\sqrt{n}}} \approx t_{n-1} \quad (7.69)$$

Presenta distribución t de Student con  $(n-1)$  grados de libertad, y para valores de  $n > 30$  su función de probabilidad es aproximadamente normal, con media cero y varianza  $n/(n-2)$ .

## 7.7. Distribución de probabilidad del cociente entre varianzas de dos poblaciones

Distribución de probabilidad F de Fisher-Snedecor

Sean  $u \chi_{k_1}^2$  y  $v \chi_{k_2}^2$  dos variables aleatorias ji cuadrada con  $k_1$  y  $k_2$  grados de libertad respectivamente, ambas estadísticamente independientes. Sea la variable aleatoria

$$F = \frac{\frac{u}{k_1}}{\frac{v}{k_2}} \quad (7.70)$$

A la variable aleatoria "F" se le denomina F de Fisher-Snedecor, por los dos principales científicos que se dedicaron a analizarla: Sir Ronald Aylmer Fisher (1890-1962), ya citado en la página 11 de este libro, y George Waddel Snedecor (1881-1974).

La función de densidad de probabilidad de la variable aleatoria F de Fisher-Snedecor está dada por la expresión:

$$f(F) = \frac{\Gamma\left(\frac{k_1 + k_2}{2}\right) \left(\frac{k_1}{k_2}\right)^{\frac{k_1}{2}}}{\Gamma\left(\frac{k_1}{2}\right) \Gamma\left(\frac{k_2}{2}\right)} \frac{F^{\frac{k_1}{2}-1}}{\left[\frac{k_1}{k_2} F + 1\right]^{\frac{(k_1+k_2)}{2}}} \quad (7.71)$$

La cual sigue la distribución F con  $k_1$  grados de libertad en el numerador y  $k_2$  grados de libertad en el denominador, lo que suele abreviarse como  $F_{k_1, k_2}$ .

Demostración

Dado que  $u$  y  $v$  son estadísticamente independientes, su función de densidad conjunta es:

$$f(u, v) = g_u(u) h_v(v) = \frac{u^{\frac{k_1}{2}-1} e^{-\frac{u}{2}}}{2^{\frac{k_1}{2}} \Gamma\left(\frac{k_1}{2}\right)} \frac{v^{\frac{k_2}{2}-1} e^{-\frac{v}{2}}}{2^{\frac{k_2}{2}} \Gamma\left(\frac{k_2}{2}\right)}$$

Por otra parte, del procedimiento establecido en el subtema 4.3, página 162, se definen las variables:

$$F = \frac{u/k_1}{v/k_2}$$

$$G = v$$

Obsérvese que  $u$  y  $v \in \mathfrak{R}^+$  por ser sumas de cuadrados, lo que implica que:  $F > 0$  y  $G > 0$ .

De las expresiones anteriores se despejan  $u$  y  $v$ :

$$u = \frac{k_1}{k_2} FG$$

$$v = G$$

y se sustituyen en la función de probabilidad conjunta:

$$f(F, G) = \frac{\left(\frac{k_1}{k_2} FG\right)^{\frac{k_1}{2}-1} e^{-\frac{k_1}{2k_2} FG}}{2^{\frac{k_1}{2}} \Gamma\left(\frac{k_1}{2}\right)} \frac{G^{\frac{k_2}{2}-1} e^{-\frac{G}{2}}}{2^{\frac{k_2}{2}} \Gamma\left(\frac{k_2}{2}\right)}$$

Se obtiene el Jacobiano de la Transformación:

$$J \begin{pmatrix} u & v \\ F & G \end{pmatrix} = \begin{vmatrix} \frac{\partial u}{\partial F} & \frac{\partial u}{\partial G} \\ \frac{\partial v}{\partial F} & \frac{\partial v}{\partial G} \end{vmatrix} = \begin{vmatrix} \frac{k_1}{k_2} G & \frac{k_1}{k_2} F \\ 0 & 1 \end{vmatrix} = \frac{k_1}{k_2} G$$

Por lo cual la función de probabilidad conjunta de  $t$  y  $s$  está definida por la expresión:

$$f(F, G) = \frac{\frac{k_1}{k_2} \left(\frac{k_1}{k_2} FG\right)^{\frac{k_1}{2}-1} G^{\frac{k_2}{2}} e^{-\frac{G}{2} \left(\frac{k_1}{k_2} F + 1\right)}}{2^{\frac{k_1}{2}} \Gamma\left(\frac{k_1}{2}\right) 2^{\frac{k_2}{2}} \Gamma\left(\frac{k_2}{2}\right)}$$

De donde al integrar con respecto a  $G$ , desde cero a infinito:

$$f(F) = \int f(F,G)dG = \int_0^\infty \frac{\frac{k_1}{k_2} \left( \frac{k_1}{k_2} FG \right)^{\frac{k_1}{2}-1} G^{\frac{k_2}{2}} e^{-\frac{G}{2} \left( \frac{k_1}{k_2} F+1 \right)} dG}{2^{\frac{k_1}{2}} \Gamma\left(\frac{k_1}{2}\right) 2^{\frac{k_2}{2}} \Gamma\left(\frac{k_2}{2}\right)}$$

Al aplicar las propiedades de la función Gamma se obtiene la expresión (7.71), lo que queda demostrado.

La función de probabilidad acumulada  $F$  de Fisher Snedecor está dada por la expresión:

$$F(f) = \frac{\Gamma\left(\frac{k_1+k_2}{2}\right) \left(\frac{k_1}{k_2}\right)^{\frac{k_1}{2}}}{\Gamma\left(\frac{k_1}{2}\right) \Gamma\left(\frac{k_2}{2}\right)} \int_0^f \frac{u^{\frac{k_1}{2}-1}}{\left[\frac{k_1}{k_2}u+1\right]^{\frac{(k_1+k_2)}{2}}} du \quad (7.72)$$

La media y la varianza de la función de probabilidad  $F$  de Fisher Snedecor están dadas por las expresiones:

$$\mu_F = \frac{k_2}{k_2-2} \quad k_2 > 2 \quad (7.73)$$

$$\sigma_F^2 = \frac{2k_2^2 (k_1+k_2-2)}{k_1(k_2-2)^2 (k_2-4)} \quad k_2 > 4 \quad (7.74)$$

Para hacer el cálculo de probabilidades con la distribución  $F$  de Fisher Snedecor, utilizando Excel, se utiliza el siguiente comando:

$$p(f, k_1, k_2) = \text{DISTR.FN}(f, k_1, k_2, \text{acumulado}) \quad (7.75)$$

En donde  $f$  es el valor de la variable  $F$  de Fisher Snedecor para el cual se pretende evaluar su probabilidad,  $k_1$  es el número de grados de libertad en el numerador,  $k_2$  es el número de grados de libertad en el denominador, acumulado= verdadero o 1 si se pretende calcular la función de probabilidad acumulada desde cero hasta  $f$ ; acumulado= falso o 0 si se pretende calcular el valor de la función de densidad para un valor  $f$ .



Cabe señalar que Excel también maneja otro comando que proporciona la probabilidad de que la variable sea mayor que  $f$ , es decir proporciona la probabilidad a la derecha del valor  $t$ .

$$p(t)=\text{DISTR.FCD}(f, k_1, k_2) \quad (7.76)$$

De tal manera que:

$$p(f=5, k_1=8, k_2=12)= \text{DISTR.FN}(5, 8, 12, 0)= 0.00533$$

Esta no es una probabilidad, es el valor de la función de densidad en el punto  $f=5$  para 8 grados de libertad en el numerador y 12 grados de libertad en el denominador.

$$p(f \leq 5, k_1=8, k_2=12)= \text{DISTR.FN}(5, 8, 12, 1)= 0.993425$$

o también

$$p(f \leq 5, k_1=8, k_2=12)= 1-\text{DISTR.FCD}(5, 8, 12, 1)= 0.993425$$

$$p(f > 5, k_1=8, k_2=12)= \text{DISTR.FCD}(5, 8, 12)= 0.006575$$

$$p(4 \leq f \leq 6, k_1=8, k_2=12)= \text{DISTR.FN}(6, 8, 12, 1) - \text{DISTR.FN}(4, 8, 12, 1)= 0.01258$$

$$p(4 \leq f \leq 6, k_1=8, k_2=12)= \text{DISTR.FCD}(4, 8, 12) - \text{DISTR.FCD}(6, 8, 12)= 0.012576$$

Se pueden obtener valores de la  $F$  de Fisher Snedecor para probabilidades dadas, a través del siguiente comando:

$$F_{p_0, k_1, k_2} = \text{INV.F}(p_0, k_1, k_2) \quad (7.77)$$

El cual proporciona el valor de  $F$  para el cual la probabilidad a la izquierda de él es  $p_0$ , como se puede apreciar a continuación:

Obtener el valor de  $f_0$  para el cual  $p(f \leq f_0)=0.90$  con  $k_1=8$  y  $k_2=12$  grados de libertad:

$$f_{0.9,8,12}= \text{INV.F}(0.9, 8, 12)= 2.24457$$

$$P(f_{0.05,8,12} \leq f \leq f_{0.95,8,12})= 0.90$$

En este último caso se calculan dos valores:

$$f_{0.95,8,12} = \text{INV.F}(0.95, 8, 12) = 2.84857$$

$$f_{0.05,8,12} = \text{INV.F}(0.05, 10) = 0.30451$$

Para hacer el cálculo de probabilidades con la distribución F de Fisher Snedecor, utilizando Minitab, se realizan los siguientes pasos:

1. En el menú superior de Minitab dar click en Calc, después en Probability Distributions y posteriormente en F, como se muestra en la figura 7.4.
2. Aparece la pantalla que se muestra en la figura 7.17, en la cual, para calcular la probabilidad  $p(f \leq 5, k_1 = 8, k_2 = 10)$  se señala Cumulative probability, Noncentrality parameter 0.0, en Numerator degrees of freedom se indica 8, Denominator degrees of freedom se indica 12 y en Input constant se indica el valor de f para el cual se desea calcular la probabilidad y se da un click en Ok obteniéndose el resultado ya estimado previamente:

$$p(f \leq 5, k_1 = 8, k_2 = 12) = 0.993425$$

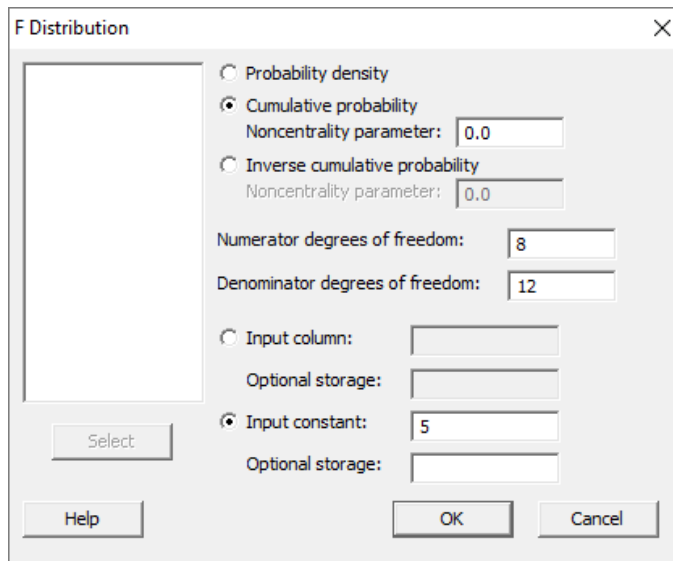


FIGURA 7.17.

En el lenguaje de desarrollo R, para el caso de la distribución F de Fisher Snedecor, para el cálculo de probabilidades se usa uno de los siguientes comandos:

Para el cálculo de la función de densidad exactamente en un punto:

$$df(x, df_1, df_2, ncp, \log = \text{FALSE}) \quad (7.78)$$

Para el cálculo de la probabilidad acumulada hasta un punto dado:

$$pf(x, df_1, df_2, ncp, \text{lower.tail} = \text{TRUE}, \log.p = \text{FALSE}) \quad (7.79)$$

Para el cálculo de los percentiles, deciles o cuartiles se usa la expresión:

$$qf(q, df_1, df_2, ncp, \text{lower.tail} = \text{TRUE}, \log.p = \text{FALSE}) \quad (7.80)$$

Para generar  $n$  números aleatorios con distribución F de Fisher con  $df_1$  grados de libertad en el numerador y  $df_2$  grados de libertad en el denominador se usa la expresión:

$$rf(n, df_1, df_2, ncp) \quad (7.81)$$

En donde  $x$  es el valor que toma la variable aleatoria F de Fisher Snedecor,  $df_1$  es el número de grados de libertad en el numerador,  $df_2$  es el número de grados de libertad en el denominador,  $ncp$  es el parámetro de no centralidad (ya que puede no estar centrada con respecto al origen),  $\log$  es un valor lógico, si se le asigna un valor verdadero el valor que arroja es el logaritmo de la probabilidad calculada, por omisión se toma FALSE.  $\text{lower.tail}$  es un valor lógico; si se le asigna un valor verdadero, el valor que arroja es la probabilidad acumulada a la izquierda de  $x$ , de lo contrario, si el valor que toma es Falso, el valor que arroja es la probabilidad a la derecha del valor de  $x$ .

Por ejemplo,

$$f(x=5, k_1=8, k_2=12) = df(5, 8, 10, 0, \log=\text{FALSE}) = 0.005330$$

Esta no es una probabilidad, es el valor de la función de densidad en el punto  $x=5$  para 8 grados de libertad en el numerador y 12 en el denominador.

$$p(x \leq 5, k_1=8, k_2=12) = pf(5, 8, 12, 0, \text{lower.tail}=\text{TRUE}, \log.p=\text{FALSE}) = 0.993425$$

$$p(x > 5; k_1=8, k_2=12) = pf(5, 10, 0, \text{lower.tail}=\text{FALSE}, \log.p=\text{FALSE}) = 0.006575$$

$$p(4 \leq f \leq 6, k_1 = 8, k_2 = 12) = pf(6, 8, 12, 0, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) - pt(4, 8, 12, 0, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 0.012576$$

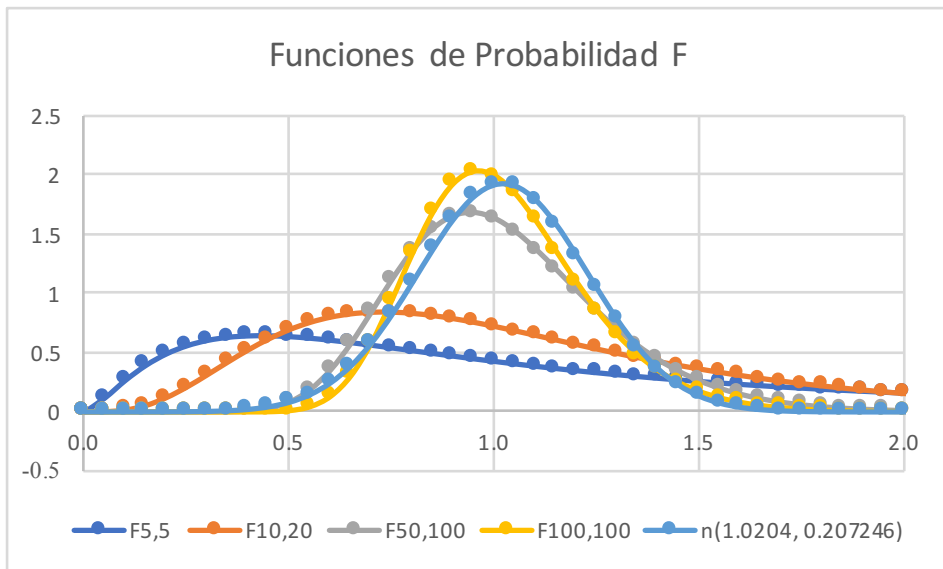
$$f_{0.9,8,12} = qf(0.9, 8, 10, 0, \text{lower.tail}=\text{TRUE}, \text{log.p}=\text{FALSE}) = 2.24457$$

Generar cinco números aleatorios con distribución F de Fisher Snedecor con 8 grados de libertad en el numerador y 12 grados de libertad en el denominador:

$$rf(5, 8, 12, 0) = 1.06988 \quad 0.77028 \quad 3.17687 \quad 0.96042 \quad 2.379$$

En la figura 7.18 se muestran algunas gráficas de la distribución F de Fisher Snedecor. Nótese que en la medida que el número de grados de libertad crece, la gráfica de la F de Fisher se aproxima más a una normal, lo que refleja el cumplimiento del teorema del límite central. La última gráfica que se muestra en color azul claro corresponde con una normal de media  $k_2/(k_2-2) = 100/98 = 1.0204$  y desviación estándar raíz  $(2k_2^2(k_1+k_2-2) / k_1(k_2-2)^2(k_2-4)) = 0.207246$ , nótese como esta es parecida a una F de Fisher con 100 grados de libertad en el numerador y 100 en el denominador.

**FIGURA 7.18.** Gráficas de la distribución F de Fisher



La importancia del comportamiento probabilístico de una variable aleatoria F de Fisher se debe a que, como se verá a continuación, el parámetro estadístico denominado cociente entre varianzas se comporta como una F de Fisher Snedecor.

En el subtema anterior se demostró que la variable:

$$\frac{(n-1)S_{n-1}^2}{\sigma_x^2} \sim \chi^2_{n-1}$$

Presenta función de probabilidad ji cuadrada con  $n-1$  grados de libertad.

Por definición de F de Fisher Snedecor

$$F_{k_1-1, k_2-1} = \frac{\frac{(k_1-1)S_1^2}{\sigma_1^2} / (k_1-1)}{\frac{(k_2-1)S_2^2}{\sigma_2^2} / (k_2-1)} = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \quad (7.82)$$

Esto implica que la variable de la expresión 7.82 presenta distribución F de Fisher Snedecor, con  $k_1-1$  grados de libertad en el numerador y  $k_2-1$  grados de libertad en el denominador.

## Ejercicios propuestos del capítulo 7

- 7.1. Sea  $x \sim N(50, 10)$ . Determinar las siguientes probabilidades:
- $p(X < 40)$
  - $p(X < 65)$
  - $p(X > 55)$
  - $p(X > 35)$
  - $p(40 < x < 45)$
  - $p(38 < X < 62)$
  - Calcular el valor de  $k$  tal que  $p(x < \mu_x - k\sigma_x) = 0.95$
  - Calcular el valor de  $k$  tal que  $p(\mu_x - k\sigma_x < x < \mu_x + k\sigma_x) = 0.95$
- 7.2. La cantidad de agua que pasa por una estación de aforo de un canal tiene distribución normal de media  $10 \text{ m}^3/\text{s}$  y desviación estándar  $2 \text{ m}^3/\text{s}$ . Si la capacidad máxima del canal es de  $14 \text{ m}^3/\text{s}$  calcular la probabilidad de que el caudal se desborde del canal.
- 7.3. Sea  $X \sim N(\mu, \sigma)$ . Determinar la media y la varianza de  $x$  si los percentiles al 40% y al 80% son  $p_{0.4} = 50$  y  $p_{0.8} = 100$ .
- 7.4. La distancia  $x$  a la cual un competidor puede efectuar un tiro, en una competencia de lanzamiento de Jabalina, se considera que es una variable aleatoria normal con parámetros  $\mu = 50 \text{ m}$  y  $\sigma^2 = 25 \text{ m}^2$ . Si el competidor en cuestión efectúa un tiro, calcular la probabilidad de que su lanzamiento:
- Haya sido de más de 55 m.
  - Haya sido entre 50 y 60 m.
- 7.5. La longitud de las flechas de ciertas máquinas está distribuida normalmente. Si 13.47% de las flechas miden más de 174.4 centímetros y 8.08% miden menos de 164.4 centímetros, ¿Cuál es la media y la desviación estándar de las longitudes de las flechas?

- 7.6. El tiempo programado de vuelo de un avión de México a Acapulco es de 50 minutos, pero por situaciones meteorológicas impredecibles la duración programada puede variar normalmente llegando a ser de más de 60 minutos en el 5% de los casos. Calcular la probabilidad de que algún vuelo sea:
- Menor de 45 minutos.
  - De 48 a 53 minutos.
- 7.7. Un estudio mostró que la vida media de los focos de una cierta marca puede modelarse con una variable aleatoria con distribución normal, de media igual a 624 días y desviación estándar igual a 90 días. Si la fábrica desea garantizar sus focos por 30 meses, ¿Cuál es el porcentaje de focos que tendrían que ser rechazados por el fabricante al no cumplir con la garantía?
- 7.8. Cierta tipo particular de transistor deberá utilizarse en un circuito durante 50 horas, se dispone de dos marcas de transistores, una que tiene una distribución de duración exponencial con  $\mu_1=45$ , y la otra con una distribución de duración normal con  $\mu_2=48$  y  $\sigma_2^2=9$ , ¿Qué transistor deberá preferirse? Justificar su respuesta.
- 7.9. Una fábrica de pistones cuyos diámetros se encuentran adecuadamente clasificados por una distribución normal con un diámetro promedio de 5 cm y una desviación estándar igual a 0.001 cm establece que para que cada pistón sirva, su diámetro debe encontrarse entre 4.998 y 5.002 cm. Si el diámetro del pistón es menor que 4.998 se desecha; si es mayor que 5.002 el pistón puede reprocesarse.
- ¿Qué porcentaje de pistones servirá?
  - ¿Qué porcentaje será reprocesado?
  - ¿Qué porcentaje será desechado?
- 7.10. El diámetro exterior de una flecha “D” es una variable aleatoria distribuida normalmente con una media de 4 pulgadas y desviación estándar de 0.1 pulgadas. Se especifica que el diámetro “D” debe ser de  $4 \pm 0.05$  pulgadas para obtener una ganancia de \$1.00. Si el diámetro real se diferencia del valor especificado por más de 0.05 pulgadas, pero menos de 0.08 pulgadas, la pérdida del fabricante es de \$0.50. Si el diámetro real se diferencia del diámetro especificado en más de 0.08 pulgadas, la pérdida es de \$1.00.

- a) Calcular las probabilidades de que la variable se encuentre entre los intervalos mencionados.
  - b) Calcular el valor esperado de la ganancia.
- 7.11. La demanda mensual de cierto producto A tiene una distribución normal con una media de 500 unidades y desviación estándar igual a 80 unidades. La demanda mensual de cierto producto B tiene una distribución normal con una media de 600 unidades y desviación estándar igual a 100 unidades. Un comerciante que vende estos productos tiene en su almacén 280 unidades de A y 650 de B al comienzo de un mes, ¿Cuál es la probabilidad de que, en el mes, se vendan todas las unidades de ambos productos? Puede suponerse independencia entre ambos eventos.
- 7.12. Un fabricante de aviones desea obtener remaches para montar los propulsores de sus aviones. El esfuerzo a la tensión mínimo necesario de cada remache es de 25,000 lb. Se pide a tres fabricantes de remaches (A, B y C) que proporcionen toda la información pertinente con respecto a los remaches que producen. Los tres fabricantes aseguran que la resistencia a la tensión de sus remaches se encuentra distribuida, de manera aproximadamente normal, con un valor medio de 28 000, 30 000 y 29 000 lb, respectivamente.
- a) ¿Tiene el fabricante la suficiente información para hacer una selección? ¿Por qué?
  - b) Supóngase que las desviaciones estándar para A, B y C son 1 000, 1 800, 1 200 respectivamente. ¿Cuál es la probabilidad de que un remache producido ya sea por A, B o C no reúna los requisitos mínimos?
  - c) Si usted fuera el fabricante de aviones, ¿podría elegir entre A, B y C, con base en su respuesta al inciso b)? ¿por qué?
- 7.13. Una flecha cuyo diámetro es  $N(1, 0.03)$  debe insertarse en un collarín cuyo diámetro es  $N(1.04, 0.04)$ . Determinar la probabilidad de que no entre la flecha en el collarín.
- 7.14. En un proceso de producción 10% de sus artículos son defectuosos. Diariamente se selecciona una muestra aleatoria de 200 artículos y se cuenta el número de defectuosos  $x$ .
- a) Identifique, fundamentando porqué, de qué tipo de distribución se trata.
  - b) ¿Se podría aproximar a la distribución normal? ¿Qué condiciones son necesarias?



- c) Utilizando la aproximación a la normal calcule  $p(x \leq 20)$ ,  $p(x=15)$ ,  $p(15 \leq x \leq 25)$
- 7.15. Un ensamble para la flecha de una máquina está constituido por cinco secciones. La longitud de cada sección se comporta como una variable aleatoria normal, independiente de las demás secciones. Las secciones extremas tienen una longitud promedio de 1.001 cm cada una. Las tres secciones de en medio tienen una longitud promedio de 1.999 cm cada una. La desviación estándar en la longitud de cada sección es la misma y tiene valor de 0.004 cm.
- ¿Cuál es la longitud promedio de dicho ensamble?
  - ¿Cuál es la desviación estándar de este ensamble?
  - ¿Qué porcentaje de los ensambles tiene una longitud mayor de 8.002 cm?
- 7.16. Los pesos de 600 ejes producidos en un torno tienen distribución normal de media 53 kg. y desviación estándar de 2.5 kg. Estos se empaquetan en cajas de 10, las que soportan hasta 540 Kg de peso. Si se envían 35 de estas cajas, calcular cuántas cajas cabe esperar se rompan por exceso de peso.
- 7.17. Cierta proceso de manufactura produce pernos que deben tener un diámetro desde 1.2 hasta 1.25 pulgadas. Las especificaciones de la máquina establecen que el diámetro medio es de 1.21 con una desviación estándar de 0.02. ¿Qué porcentaje de los pernos está fuera de especificaciones?
- 7.18. En los siguientes incisos grafique la función de probabilidad ji cuadrada, obtenga los valores que se le solicitan e indíquelos en la gráfica, empleando tablas, Excel, Minitab y R.
- $X_{0.90, 5}^2$
  - $X_{0.95, 10}^2$
  - $X_{0.025, 20}^2$
  - $X_{\alpha}^2$  tal que  $P\{X_{12}^2 \leq X_{\alpha}^2, 12\} = 0.975$
  - Genere 10 números aleatorios con distribución ji cuadrada con ocho grados de libertad.
- 7.19. En los siguientes incisos grafique la función de probabilidad t de Student, obtenga los valores que se le solicitan e indíquelos en la gráfica, empleando tablas, Excel, Minitab y R.
- $t_{0.90, 5}$

- b)  $t_{0.95, 10}$
- c)  $t_{0.025, 20}$
- d)  $t_{\alpha}$ , tal que  $p\{t_{10} \leq t_{\alpha, 10}\} = 0.975$
- e) Genere 10 números aleatorios con distribución  $t_8$ .

7.20. En los siguientes incisos grafique la función de probabilidad F de Fisher Snedecor, obtenga los valores que se le solicitan e indíquelos en la gráfica, empleando tablas, Excel, Minitab y R.

- a)  $F_{0.90, 5, 5}$
- b)  $F_{0.95, 5, 10}$
- c)  $F_{0.025, 10, 20}$
- d)  $F_{\alpha}$ , tal que  $p\{F_{10, 20} \leq F_{\alpha, 10, 20}\} = 0.975$
- e) Genere 10 números aleatorios con distribución  $F_{8, 12}$ .

## Referencias Bibliográficas

1. William W. Hines, Douglas C. Montgomery, David M. Goldsman, Connie M. Borror. *Probabilidad y Estadística para Ingeniería*. Editorial Patria. Cuarta Edición. México, 2013.
2. Maria Dolores Ugarte, Ana F. Militino, Alan T. Arnholt. *Probability and Statistics with R*. Chapman and Hall/CRC, 2015.
3. John J. Shynk. *Probability, Random Variables, and Random Processes. Theory and Signal Processing Applications*. John Wiley & Sons Inc. 2013.
4. Alfredo H. S. Ang and Wilson H. Tang. *Probability Concepts in Engineering Planning and Design*. John Wiley & Sons, Inc, 1975.
5. Bernard Ostle, *Estadística Aplicada*. Editorial LIMUSA, 1983.
6. Jack R. Benjamin & C. Allin Cornell. *Probability, Statistics and Decision for Civil Engineers*. McGraw–Hill Book Company, 1970.
7. William Feller. *An Introduction to Probability Theory and Its Applications*, Volume I, 3rd Edition, John Wiley & Sons, 1968.

## Referencias Mesográficas

1. <https://www.nationalgalleries.org/sites/default/files/styles/postcard/public/externals/41629.jpg?itok=6ukIZbgA>
2. [https://upload.wikimedia.org/wikipedia/commons/0/05/Disguised\\_Vishnu\\_and\\_a\\_Brahman%2C\\_Folio\\_from\\_a\\_Razmnama%2C.jpg](https://upload.wikimedia.org/wikipedia/commons/0/05/Disguised_Vishnu_and_a_Brahman%2C_Folio_from_a_Razmnama%2C.jpg)
3. [https://es.wikipedia.org/wiki/Pierre\\_de\\_Fermat](https://es.wikipedia.org/wiki/Pierre_de_Fermat)
4. [https://es.wikipedia.org/wiki/Blaise\\_Pascal](https://es.wikipedia.org/wiki/Blaise_Pascal)
5. [https://es.wikipedia.org/wiki/Familia\\_Bernoulli](https://es.wikipedia.org/wiki/Familia_Bernoulli)
6. [https://es.wikipedia.org/wiki/Jakob\\_Bernoulli](https://es.wikipedia.org/wiki/Jakob_Bernoulli)
7. [https://es.wikipedia.org/wiki/Abraham\\_de\\_Moivre](https://es.wikipedia.org/wiki/Abraham_de_Moivre)
8. [https://es.wikipedia.org/wiki/Carl\\_Friedrich\\_Gauss](https://es.wikipedia.org/wiki/Carl_Friedrich_Gauss)
9. [https://es.wikipedia.org/wiki/Pierre-Simon\\_Laplace](https://es.wikipedia.org/wiki/Pierre-Simon_Laplace)
10. [https://es.wikipedia.org/wiki/Andr%C3%A9\\_Kolmog%C3%B3rov](https://es.wikipedia.org/wiki/Andr%C3%A9_Kolmog%C3%B3rov)
11. [https://es.wikipedia.org/wiki/Ronald\\_Fisher](https://es.wikipedia.org/wiki/Ronald_Fisher)
12. [https://es.wikipedia.org/wiki/Historia\\_de\\_la\\_estad%C3%ADstica](https://es.wikipedia.org/wiki/Historia_de_la_estad%C3%ADstica)
13. [http://www.estadisticaparatodos.es/historia/histo\\_proba.html](http://www.estadisticaparatodos.es/historia/histo_proba.html)
14. [https://www.sis.net/documentos/Digitalizados/22939\\_Repensando%20la%20investigacion.pdf](https://www.sis.net/documentos/Digitalizados/22939_Repensando%20la%20investigacion.pdf)
15. [https://es.wikipedia.org/wiki/Georg\\_Cantor](https://es.wikipedia.org/wiki/Georg_Cantor)
16. [https://es.wikipedia.org/wiki/Funci%C3%B3n\\_gamma](https://es.wikipedia.org/wiki/Funci%C3%B3n_gamma)
17. [https://es.wikipedia.org/wiki/Ley\\_de\\_los\\_grandes\\_n%C3%BAmeros](https://es.wikipedia.org/wiki/Ley_de_los_grandes_n%C3%BAmeros)
18. <http://ilcolibri.altervista.org/webtraining/wp-content/uploads/2015/12/Estadisticas-dado.png>
19. <file:///D:/Sem%202015-2/ProbyEst2015-2/Escuelas%20de%20Probabilidad.pdf>



*Fundamentos de Probabilidad  
y Aplicaciones, con R, Minitab y Excel*

se publicó digitalmente en el repositorio de la  
Facultad de Ingeniería el 8 de mayo de 2023.

Segunda edición electrónica de un ejemplar  
(10 MB) en formato PDF.

El cuidado de la edición y diseño estuvieron a  
cargo de la Unidad de Apoyo Editorial de la  
Facultad de Ingeniería. La familia tipográfica  
utilizada fueron Minion Pro y Chivo con sus  
respectivas variantes.