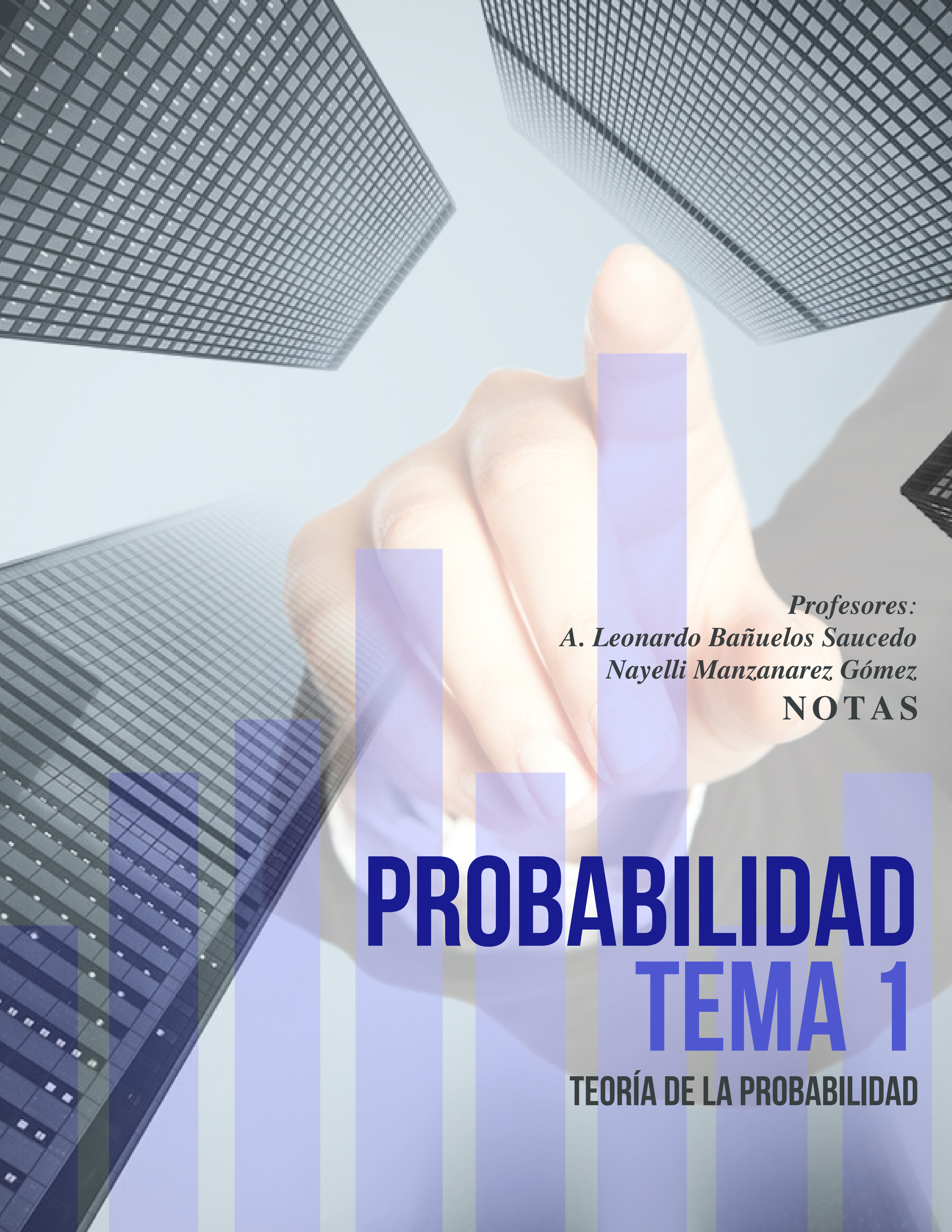


Profesores:
A. Leonardo Bañuelos Saucedo
Nayelli Manzanarez Gómez

NOTAS

PROBABILIDAD

TEMA 1

TEORÍA DE LA PROBABILIDAD

TEMA I
TEORÍA DE LA PROBABILIDAD

INTRODUCCIÓN

De manera sencilla, puede decirse que la probabilidad es la rama de las matemáticas que estudia los fenómenos con incertidumbre, pero ¿por qué estudiar Probabilidad en Ingeniería? Hasta este momento, en prácticamente todas las asignaturas cursadas en la Facultad de Ingeniería, se han estudiado modelos determinísticos, esto es, los problemas estudiados tienen siempre solución única o un conjunto de respuestas perfectamente delimitado. En Cinemática, por ejemplo, el tiempo que tarda un objeto en llegar al suelo en caída libre depende de la altura y de la aceleración de la gravedad; en Termodinámica, la presión manométrica en el fondo de un recipiente depende de la altura de la columna de agua y del peso específico del fluido; en Cálculo, el valor de una integral $\int_a^b f(x)\,dx$ depende de la función y de los límites de integración; pero en todos los casos anteriores, se ha supuesto que es posible definir con precisión los parámetros que permitirán utilizar la expresión que proporcionará el resultado, y en cada caso el resultado siempre es único. La gran mayoría de los modelos determinísticos son idealizaciones o simplificaciones de la realidad, mismas que la naturaleza del problema permite realizar; esto no siempre es posible y se debe recurrir a modelos probabilísticos o aleatorios, en el caso de la caída libre, se desprecia la resistencia del aire sobre el cuerpo que cae, así como la variación de la aceleración de la gravedad; en el caso de la presión manométrica, se considera que la densidad es constante (cuerpo homogéneo), etc. Cuando en un experimento no se pueden realizar las simplificaciones necesarias para conocer con precisión el resultado, se debe recurrir a los modelos aleatorios, por ejemplo:

- El numero de personas que llegan a una fila bancaria en una hora.
- La cantidad de lluvia en un lugar determinado, etc.

En el lanzamiento de un moneda al aire, donde se pretende observar la cara que queda hacia arriba (volado), se sabe que existen dos resultados: águila o sol; sin embargo, no se pueden hacer más consideraciones que permitan predecir el resultado del lanzamiento. En este tipo de experimentos se deben utilizar modelos aleatorios. Así, en muchos problemas de ingeniería, es imposible predecir con exactitud las variables que intervienen en un problema específico, como por ejemplo, el número de llamadas telefónicas que debe enlazar una central en cierta hora del día, el número de personas que llegan a una fila bancaria para ser atendidas, la interferencia electromagnética en las comunicaciones, la cantidad de lluvia en un lugar determinado en un mes, el comportamiento de la bolsa de valores, etc. Y para poder estudiar éstos y otros problemas es necesario estudiar y comprender la Probabilidad.

DEFINICIONES

Un experimento es cualquier procedimiento capaz de generar resultados observables. Los experimentos se pueden clasificar en determinísticos y aleatorios.

$$\textit{Experimento} \begin{cases} \textit{Determinístico} \\ \textit{Aleatorio} \end{cases}$$

Un experimento determinístico es aquel que al repetirse bajo las mismas condiciones controlables presenta siempre el mismo resultado.

Un experimento aleatorio es aquel que al repetirse bajo las mismas condiciones aparentes puede presentar diferentes resultados.

La clasificación de los experimentos en determinísticos y aleatorios depende de la información y el conocimiento que se tenga del experimento, y dependiendo del experimento que se desee estudiar se utilizan modelos matemáticos de tipo determinístico o de tipo aleatorio (probabilístico).

Los modelos matemáticos de tipo determinístico permiten predecir los resultados de un experimento a partir de ciertas condiciones, ejemplos de modelos determinísticos son:

$s = -\frac{1}{2} g t^2 + v_o t + s_0$ Tiro parabólico

$F = k \delta$ Ley de Hooke

$V = R I$ Ley de Ohm

Los modelos matemáticos de tipo probabilístico no permiten predecir el resultado de un experimento, únicamente indican la frecuencia con la que cabe esperar determinado resultado al repetir el experimento un número muy grande de veces, o bien, la certidumbre que se tiene con respecto a la obtención de algún resultado en una sola ejecución del experimento, ejemplos de modelos aleatorios son:

El resultado de un volado.
El resultado de lanzar un dado.
El comportamiento en el precio de una acción.
La velocidad del aire.

Dependiendo de si el resultado que se puede observar toma valores de un conjunto continuo o discreto, el experimento aleatorio se puede clasificar en experimento continuo o experimento discreto [1].

Ejemplos de experimentos aleatorios discretos son el volado y el lanzamiento de un dado, mientras que el precio de una acción y la velocidad del aire son continuos.

Los conceptos básicos de la Probabilidad tienen su fundamento en la Teoría de Conjuntos, en donde el concepto básico es la pertenencia; en Probabilidad, el concepto básico es la ocurrencia.

El conjunto de todos los resultados posibles recibe el nombre de espacio muestral, se denota por S, o bien por Ω. En algunos casos, el analista puede definir más de un espacio muestral para una experimento, en función de lo que desee observar; sin embargo, siempre existirá una definición de espacio muestral que contenga todas las definiciones parciales y es recomendable definir el espacio muestral en términos de los que se observa en el experimento. A cualquier subconjunto del espacio muestral se le llama evento, y se acostumbra denotar mediante las primeras letras mayúsculas del alfabeto.

Ejemplo 1.1 Espacio Muestral

Definir el espacio muestral para el experimento: lanzar un dado.

Resolución
Una primera definición podría ser:

---

[1] Experimento continuo. Si el resultado se toma de un conjunto continuo de valores, por ejemplo. Velocidad del aire, precio de una acción.

Experimento discreto. Si el resultado se toma de un conjunto discreto de valores.

$S_1 = \{ \; par \; , \; impar \; \}$
Una definición completa sería:
$S_2 = \{ \; 1, \; 2, \; 3, \; 4, \; 5, \; 6 \; \}$
que es lo que se observa al lanzar el dado.

Los espacios muestrales con un número grande o infinito de puntos muestrales se describen mejor mediante un enunciado o regla. Por ejemplo: si los posibles resultados de un experimento son el conjunto de ciudades en el mundo con una población de más de un millón de habitantes, entonces el espacio muestral puede describirse por comprensión de la siguiente forma:

$$S = \{ \; x \mid x \textit{ es una ciudad con una población de más de un millón} \; \}$$

Existen algunos eventos que por su importancia tienen un nombre en particular, en ocasiones se les llama eventos notables. El evento seguro es aquel que con ocurrirá con certeza, es el espacio muestral S. El evento imposible o vacío es aquel que nunca ocurrirá, se denota por $\varnothing$ (conjunto vacío). El evento simple o elemental es un solo elemento del espacio muestral, se suele denotar E_i . Los eventos excluyentes o mutuamente excluyentes (conjuntos disjuntos) son aquellos cuya intersección es el evento imposible (conjunto vacío). Los eventos colectivamente exhaustivos son aquellos eventos cuya unión forma todo el espacio muestral.

Las definiciones razonables de los espacios muestrales utilizan eventos simples, que permiten contestar a varias preguntas de probabilidad relacionadas con el mismo experimento.

Ejemplo 1.2 Espacio Muestral y Eventos

Del experimento "Lanzamiento de dos monedas", obtener el espacio muestra y un evento que no sea simple.

Resolución

Espacio muestral $S = \{ \; (sol, sol), \; (aguila, aguila), \; (sol, aguila), \; (aguila, sol) \; \}$
Un posible evento A seria: $A = \{ \; (sol, sol), \; (aguila, aguila) \; \}$

La probabilidad estudia los fenómenos en los cuales no se tiene la certeza del resultado que ocurrirá en particular; y puesto que pueden ocurrir varios resultados, es conveniente conocer algunas técnicas que permitan la cuantificación rápida de distintas opciones. Así, como antecedente de la probabilidad, es necesario conocer algunos temas del álgebra básica y de la teoría de conjuntos, los cuales servirán para realizar cálculos probabilísticos de una manera más simple. La teoría de conjuntos se desarrolla en el apéndice A de este tema, y la forma de cuantificar los elementos de cada evento requieren de técnicas conteo.

TÉCNICAS DE CONTEO

Cuando resulta difícil o tedioso el contar el número de elementos de un conjunto finito por medio de la numeración directa, deben emplearse algunas técnicas especiales, las cuales reciben el nombre de técnicas de conteo. En algunos experimentos es útil listar los elementos del espacio muestral para obtener mayor información mediante un diagrama de árbol. Para visualizar esto considérese el problema de adquirir un equipo de medición de entre 3 marcas distintas (A, B y C); y para cada marca se puede adquirir un equipo barato o caro (Ba o Ca); y además cada equipo se puede clasificar en fácil de operar o difícil de operar(F o D). Algunas de las opciones que se pueden generar son: equipo A, barato y difícil de operar; equipo B, caro y fácil de operar; equipo C, barato y fácil de operar; etc. Todos las posibles resultados se pueden escribir en forma gráfica utilizando un diagrama de árbol [2] como el que se muestra en la figura 1.

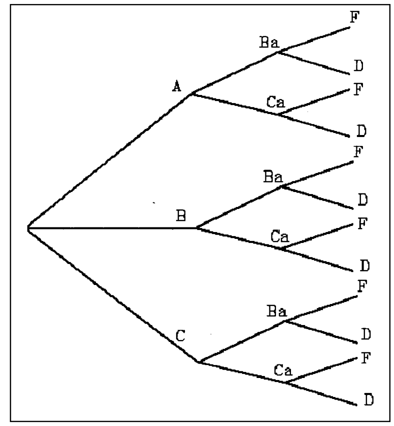


Figura 1. Diagrama de árbol

Debe observarse que en total existen 12 posibilidades, cada posibilidad corresponde a una rama del diagrama de árbol, este resultado se obtiene al multiplicar las posibilidades que existen en cada nodo, es decir, $3 \times 2 \times 2 = 12$. Lo que lleva a la regla de multiplicación de opciones, la cual se enuncia a continuación.

Teorema 1.1 (Multiplicación de opciones o Regla de la multiplicación)
Si los conjuntos $A_1, A_2, \ldots, A_k$ contienen los elementos $n_1, n_2, \ldots, n_k$, respectivamente, entonces existen $n_1 \cdot n_2 \cdot \ldots \cdot n_k$ maneras de elegir primero un elemento de A_1, luego un elemento de $A_2, \ldots$ y finalmente un elemento de A_k.

Ejemplo 1.3
En una caja existen 7 plumas, 3 lápices y 4 plumones todos diferentes entre sí. Determinar el número de formas en las que se pueden seleccionar una pluma, un lápiz y un plumón.

Resolución
Utilizando la regla de la multiplicación se tiene:

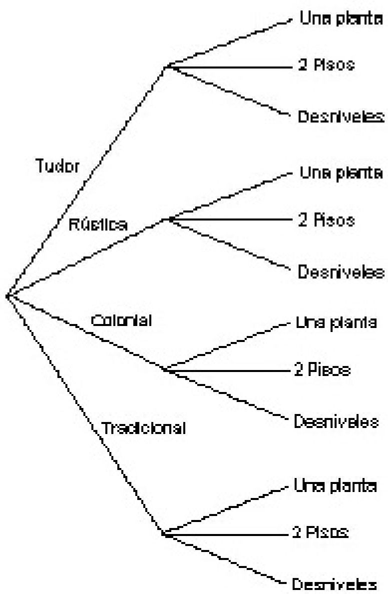
$$7 \times 3 \times 4 = 84 \text{ formas diferentes.}$$

Ejemplo 1.4

Supongamos que un ingeniero ofrece a los futuros compradores la elección del estilo de la fachada entre

[2] Representación gráfica construida con líneas ramificadas unidas por nodos que permite ilustrar un proceso a pasos.

tudor, rústica, colonial y tradicional; en una planta, dos niveles y desniveles. ¿En cuántas formas diferentes un comprador puede ordenar una de estas casas?



Resolución

Todos los posibles resultados se pueden escribir en forma gráfica utilizando un diagrama de árbol.

$(4)(3) = 12$ formas diferentes

Ejemplo 1.5

Miguel va a armar una computadora por sí mismo. Tiene la opción de comprar los procesadores de dos marcas, un disco duro de cuatro marcas, la memoria de tres marcas y un conjunto de accesorios en cinco tiendas locales. ¿ De cuántas formas diferentes puede Miguel comprar las partes?

Resolución

$n_1 = 2,\ n_2 = 4,\ n_3 = 3,\ n_4 = 5$

$2 \times 4 \times 3 \times 5 = 120$ formas diferentes de comprar las partes.

Ejemplo 1.6
De cuantas formas diferentes se puede contestar un examen de 10 preguntas de opción múltiple, con cinco incisos por pregunta.

Resolución
Nuevamente, utilizando la regla de la multiplicación se tiene:

$$5 \times 5 \times 5 \times 5 \times 5 \times 5 \times 5 \times 5 \times 5 \times 5 = 9\,765\,625$$

Ejemplo 1.7
Puede comprarse un medicamento para la cura del asma ya sea líquido, en tabletas o en cápsulas, a 5 diferentes fabricantes, y todas las presentaciones en concentración regular o alta. ¿En cuántas formas diferentes puede un médico recetar la medicina a un paciente que sufre de este padecimiento?

Resolución
Generalizando la regla de la multiplicación se tienen:
$(5)(3)(2) = 30$ formas diferentes.

En el ejemplo 1.6 se observa que el orden en el que se eligen las respuestas es importante, en general, si r objetos se seleccionan de un conjunto de n objetos distintos, cualquier arreglo ordenado de ellos recibe el nombre de permutación [3]. Una permutación difiere de otra si los elementos que la contienen cambian de orden.

Para determinar una fórmula que proporcione el número total de permutaciones de r objetos seleccionados de un conjunto de n objetos distintos, considérese que se desea obtener el número de placas de identificación que se pueden generar, si las placas tienen 4 dígitos y no se permite repetición. Entonces para el primer dígito se puede seleccionar una de las diez posibilidades que son $\{0,1,2,3,4,5,6,7,8,9\}$, para el segundo dígito se tienen nueve posibilidades, para el tercer dígito se tienen ocho posibilidades y para el cuarto dígito se tienen siete posibilidades, lo que genera un total de

$$10 \times 9 \times 8 \times 7 = 5040 \text{ posibilidades.}$$

En general, la primera elección se realiza de entre los n objetos, la segunda elección se realiza de entre los $n - 1$ objetos restantes y así sucesivamente hasta la r-ésima elección en la cual se selecciona de entre los $n - (r - 1) = n - r + 1$ objetos restantes. Utilizando la regla de la multiplicación y denotando por ${}_nP_r$ al número de permutaciones de r objetos seleccionados de un conjunto de n objetos distintos es $n(n-1)(n-2)\ .\ .\ .\ (n-r+1)$. Si además se introduce la notación factorial la fórmula de la permutación se simplifica.

Definición 2.1
El factorial de un número entero no negativo se denota mediante el símbolo !, y corresponde al producto de los números naturales menores o igual que él.
Entonces el factorial del número n es:

$$n! = n(n-1)(n-2)\ .\ .\ .\ 1$$

Y además $0! = 1$.

Utilizando la notación factorial el número de permutaciones de r elementos tomados de un conjunto de n objetos distintos está dado por el siguiente teorema:

---

[3] Algunos autores les llaman ordenaciones, siendo para ellos las permutaciones un caso particular de las ordenaciones en el cual $n = r$.

Teorema 2.2

El número de permutaciones de r objetos tomados de un conjunto de n objetos distintos es

$$ {}_nP_r = P_r^n = \frac{n!}{(n-r)!} $$

Como puede observarse en el teorema 2.2, existen dos formas de denotar a las permutaciones, sólo con subíndices, o con subíndice y superíndice.

Ejemplo 1.8

El director de un canal de televisión, debe elegir 7 de entre 20 programas para la transmisión en un cierto día, ¿de cuántas formas posibles puede armar la barra de programación?

Resolución

Claramente debe utilizar permutaciones de 7 elementos tomados de un grupo de 20, por lo que se tiene:

$$ {}_{20}P_7 = \frac{20!}{(20-7)!} = 390\,700\,800 \quad \text{formas distintas} $$

Ejemplo 1.9

a) ¿Cuántas permutaciones diferentes pueden hacerse con las letras de la palabra columna?

b) ¿Cuántas de estas permutaciones empiezan con la letra m?

Resolución

a) Puesto que no hay letras iguales:

$$ P_n^n = P_7^7 = 7! = 5040 \text{ permutaciones} $$

b) Puesto que las palabras deben empezar con la letra m, sólo se permutan las otras 6 letras, por lo que:

$$ P_6^6 = 6! = 720 \text{ permutaciones} $$

Ejemplo 1.10

En una competencia de maratón, hay 40 participantes. Determinar de cuántas formas diferentes se pueden repartir el primero, el segundo y el tercer lugar.

Resolución

Puesto que es una selección de tres personas de 40, importando el orden, se tienen:

$$ P_3^{40} = \frac{40!}{(40-3)!} = 59\,280 \quad \text{formas posibles.} $$

Un caso especial de las permutaciones es aquel en el cual se permite la repetición de los objetos que la forman, dando origen a las permutaciones con repetición. Por ejemplo, si se desea conocer el número de "combinaciones" distintas que puede tener un candado de cuatro dígitos cuyas numeraciones van del cero al seis se debe utilizar la regla de la multiplicación, teniéndose $7 \times 7 \times 7 \times 7 = 2401$ formas distintas.

Teorema 1.3

Si se seleccionan r objetos de un conjunto de n elementos, con repetición, entonces se tienen

$$_{n}PR_{r} = n^{r}$$

formas distintas de efectuar la selección.

Ejemplo 1.11

¿Cuántas placas distintas de automóviles pueden generarse, si cada placa posee tres dígitos y tres letras? Considérese un alfabeto de 26 letras.

Resolución

Para los números se tienen $_{10}PR_{3} = 10^{3} = 1000$ formas distintas.

Para las letras se tienen $_{26}PR_{3} = 26^{3} = 17\,576$ formas distintas.

Finalmente, utilizando la regla de la multiplicación se tienen:

$_{10}PR_{3} \times {}_{26}PR_{3} = 1000 \times 17\,576 = 17\,576\,000$ placas distintas.

En ocasiones se deben encontrar arreglos de objetos, de los cuales algunos son indistinguibles, por ejemplo al sacar las monedas de una caja que contiene cinco monedas de un peso, cuatro de cinco pesos y tres de diez pesos. En este caso es indistinto seleccionar cualquiera de las monedas de la misma denominación, lo que da lugar a las permutaciones con grupos de elementos iguales.

Teorema 1.4

Si se tiene un conjunto de n elementos, con grupos de elementos iguales, de los cuales n_1 son del tipo 1, n_2 son del tipo 2, etc. de tal forma que

$$n = n_1 + n_2 + \dots + n_k \quad .$$

Entonces el número de permutaciones distinguibles de los n objetos es

$$_{n}P_{n_1, n_2, \dots, n_k} = P^{n}_{n_1, n_2, \dots, n_k} = \frac{n!}{n_1!\; n_2!\; \dots\; n_k!}$$

Ejemplo 1.12

Determinar el número de permutaciones distinguibles de las letras de la palabra cocodrilo.

Resolución

En la palabra cocodrilo se observan:

2 letras c
3 letras o
1 letra d
1 letra r
1 letra i

1 letra 1

Por lo que se tienen

$$_{9}P_{2,3,1,1,1,1} = \frac{9!}{2!\ 3!\ 1!\ 1!\ 1!\ 1!} = 30\,240 \text{ permutaciones distinguibles.}$$

Un caso muy especial de las permutaciones, se tiene cuando se desea generar el arreglo alrededor de un círculo, dando origen a las permutaciones circulares (PC). Debe observarse que existe diferencia entre las permutaciones en línea y las circulares, puesto que en las circulares se considera el mismo arreglo si sus elementos tienen el mismo precedente y consecuente en el sentido de las manecillas del reloj.

Teorema 1.5

Si n elementos se acomodan de manera que formen un círculo, entonces existen

$$PC_n = (n-1)!$$

formas distintas de ordenar los elementos.

Ejemplo 1.13

En un partido de baloncesto los 10 jugadores se colocan alrededor del círculo central. ¿De cuántas formas se pueden acomodar?

Resolución

Utilizando las permutaciones circulares se tiene

$PC_{10} = (10-1)! = 362\,880$ formas distintas.

Si al seleccionar un grupo de r objetos de un conjunto de n elementos, no interesa el orden en el que se seleccionan, sino simplemente los objetos seleccionados, se generan entonces las combinaciones. Una combinación de r elementos tomados de un conjunto S que contiene n elementos, es un subconjunto de S que tiene r elementos distintos. Si se denota por $_{n}C_{r}$, o por C_r^n, al número de combinaciones que se pueden realizar al tomar r elementos de un conjunto de n, se tiene:

Teorema 1.6 Combinaciones

El número de formas en las cuales r objetos pueden seleccionarse, sin importar el orden de un conjunto de n objetos distintos es

$$_{n}C_{r} = C_r^n = \binom{n}{r} = \frac{n!}{r!\ (n-r)!}$$

Las combinaciones tienen varias formas de denotarse, como puede verse en el teorema 1.6. Por otro lado debe observase que la fórmula del teorema de combinaciones se puede obtener a partir de la fórmula de las permutaciones, si C_r^n denota el número de combinaciones (sin importar el orden), de la regla de la multiplicación, para obtener las distintas posibilidades importando el orden se tiene que multiplicar por las permutaciones de los r elementos tomados de r en r, y esto es igual al factorial de r, entonces:

$$C_r^n \; r! = P_r^n$$

de donde

$$C_r^n = \frac{P_r^n}{r!}$$

y finalmente

$$C_r^n = \frac{n!}{r! \; (n-r)!}$$

Ejemplo 1.14
De un grupo de 10 alumnos se piensa otorgar becas a 4 de ellos. ¿De cuántas maneras se puede hacer la selección?

Resolución
Puesto que para realizar la selección, no importa el orden, se utilizan combinaciones, de donde se tienen

$C_4^{10} = \dfrac{10!}{4! \; (10-4)!} = 210$ formas diferentes

---

Ejemplo 1.15
Una caja de cartón con 12 baterías para radio contiene una que está defectuosa. ¿En cuántas formas diferentes un inspector puede elegir tres de las baterías y obtener
a) la defectuosa;
b) ninguna defectuosa?

Resolución

a) Existen ${}_{12}C_3 = \begin{pmatrix} 12 \\ 3 \end{pmatrix} = 220$ formas distintas de seleccionar 3 baterías, y

${}_{11}C_3 = \begin{pmatrix} 11 \\ 3 \end{pmatrix} = 165$ formas de seleccionar las no defectuosas.

$\therefore$ Existen $220 - 165 = 55$ formas de seleccionar la defectuosa.

b) 165 formas.

Además de las combinaciones, se tienen las combinaciones con repetición, en las cuales se permite repetir los objetos en una misma combinación.

Teorema 1.7 Combinaciones con repetición
El número de formas en las cuales r objetos pueden seleccionarse, sin importar el orden, de un conjunto de n objetos distintos considerando que se pueden repetir es

$$CR_r^n = C_r^{n+r-1}$$

Para la obtención de la fórmula debe observarse que el hecho de que se puedan repetir los elementos es equivalente a que la combinación se efectúe considerando que todos los elementos se repiten r veces cuando se seleccionan los primeros $r-1$ elementos, puesto que cuando se selecciona el r-ésimo ya no debe considerarse la repetición. Por lo que la selección se hace considerando $n+r-1$ elementos, dando lugar a la fórmula $CR_r^n = C_r^{n+r-1}$.

Ejemplo 1.16
Determinar el número de combinaciones con repetición que se pueden formar con cuatro números tomando de tres en tres.

Resolución
Se tienen

$$CR_3^4 = C_3^{4+3-1} = C_3^6 = \frac{6!}{3!\ (6-3)!} = 20$$ formas posibles.

Es común denotar las combinaciones C_r^n utilizando la notación $\binom{n}{r}$, y a los números que se obtienen de las combinaciones al variar r desde cero hasta n se les llama números combinatorios. Es común leer la expresión $\binom{n}{r}$ como "las combinaciones de n sobre r", o "n combinación r".

Los números combinatorios tienen las siguientes propiedades:

- Los números combinatorios para $r = 0$ y para $r = n$ son iguales a la unidad.
  Demostración
  De la fórmula de combinaciones se tiene:
  $$\binom{n}{0} = \frac{n!}{0!\ (n-0)!} = 1$$
  $$\binom{n}{n} = \frac{n!}{n!\ (n-n)!} = 1$$
- Los números combinatorios que se tienen para valores de $r = 1, 2, \ldots, n-1$ son simétricos entre sí.
  $$\binom{n}{r} = \binom{n}{n-r}$$

Demostración
$$\binom{n}{n-r} = \frac{n!}{(n-r)!\ (n-n+r)!} = \frac{n!}{r!\ (n-r)!} = \binom{n}{r}$$

- La suma de los números combinatorios con numerador igual a n y denominador consecutivo, $r = k-1$ y $r = k$ respectivamente, es igual al número combinatorio de numerador $n = n+1$ y denominador $r = k$.
  $$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$$
  Demostración

$$\binom{n}{k-1} + \binom{n}{k} = \frac{n!}{(k-1)!\,(n-k+1)!} + \frac{n!}{k!\,(n-k)!}$$
$$= \frac{n!\,k + n!\,(n-k+1)}{k!\,(n-k+1)!} = \frac{n!\,(k+n-k+1)}{k!\,(n-k+1)!}$$
$$= \frac{(n+1)!}{k!\,(n+1-k)!} = \binom{n+1}{k}$$

Pascal, matemático francés (1623-1662), estudió las propiedades de los números combinatorios y construyó un triángulo con los números combinatorios conocido como triángulo de Pascal[4].

$$\begin{array}{ccccccc}
 & & & \binom{0}{0} & & & \\
 & & \binom{1}{0} & & \binom{1}{1} & & \\
 & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & \\
\binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} \\
 & & & \vdots & & &
\end{array}$$

sustituyendo los números combinatorios se tiene:

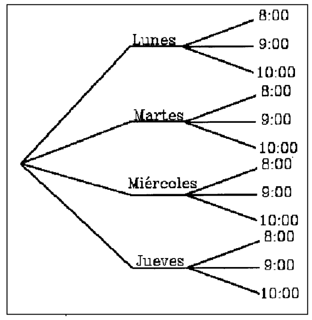
$$\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1 \\
 & & & & \vdots & & & &
\end{array}$$

En donde se puede observar con claridad las tres propiedades de los números combinatorios.

Ejemplo 1.17
Un estudiante debe efectuar una práctica de campo, y puede realizarla de lunes a jueves en cualquiera de los horarios siguientes: 8:00, 9:00 o 10:00 A.M. Dibujar el diagrama de árbol que muestre las distintas formas en las que puede realizar la práctica.

Resolución
El diagrama de árbol se muestra a continuación

[4] Pascal introdujo la notación; sin embargo, otros matemáticos previos a él como Tartaglia y Omar Khayyám ya habían estudiado sus propiedades.



Ejemplo 1.18
En un lote de 50 componentes electrónicos, se tienen 3 defectuosos.
a) ¿De cuántas formas posibles se pueden tomar 5 de ellos y encontrar un defectuoso?
b) ¿De cuántas formas posibles se pueden tomar 10 de ellos y encontrar dos defectuosos?

Resolución
a) Debe observarse que se realiza una selección de cuatro de los 47 componentes no defectuosos y uno de los tres defectuosos, sin importar el orden, por lo que se tienen:

$$C_4^{47}\ C_1^3 = \binom{47}{4}\binom{3}{1} = 535\,095 \quad \text{formas posibles.}$$

b) Puesto que tampoco importa el orden, se tienen

$$C_8^{47}\ C_2^3 = \binom{47}{8}\binom{3}{2} = 943\,372\,485 \quad \text{formas posibles.}$$

Ejemplo 1.19

Un testigo de un accidente de tránsito en el que el causante huyó, le indica al policía que el número de la matrícula del automóvil tenía las letras ALB seguidas por tres dígitos, el primero de los cuales era un cinco. Si el testigo no puede recodar los otros dos dígitos pero está seguro de que los tres eran diferentes, encuentre el número máximo de registros de automóvil que debe verificar la policía.

Resolución
Como el primer número tiene que ser cinco el número de formas queda en función de los otros dos, donde el segundo tiene $n_1 = 9$ posibilidades y el tercero $n_2 = 8$ posibilidades, entonces por la regla de la multiplicación se deben verificar $(n_1)(n_2) = (9)(8) = 72$ registros.

Ejemplo 1.20
En un estudio de economía de combustibles, se prueban tres carros de carreras con cinco diferentes marcas de gasolina, en siete sitios de prueba en distintas regiones del país. Si se utilizan dos pilotos en el estudio y las pruebas se realizan una vez bajo cada conjunto de condiciones, ¿cuántas se necesitarán?

Resolución
De la regla de la multiplicación se tienen:
$(3)(5)(7)(2) = 210$ formas diferentes.

Después de repasar brevemente algunas de las técnicas para contar, debemos retomar el propósito del capítulo, el cual es que el alumno conozca el significado de la probabilidad y pueda desarrollar varias relaciones entre probabilidades utilizando la definición de probabilidad y los teoremas fundamentales. Históricamente los primeros usos de la probabilidad fueron en los juegos de azar, y se utilizó lo que en la actualidad se llama "enfoque clásico". El primer dado data de principios del tercer milenio a.C. Los griegos ya reconocían la importancia del azar, y tenían una Diosa de la suerte, llamada Tique. El primer autor reconocido en probabilidad es el matemático, médico, físico y astrólogo italiano Gerolamo Cardano (1501-1576). Cardano escribió sobre los juegos de azar en su libro, "Liber de ludo aleae" (Manual sobre juegos de azar) , escrito en la década de 1560 pero publicado póstumamente en 1663. Cardano fue consultado para el célebre problema del duque de Toscana (1560), pero no encontró una solución satisfactoria. El problema fue resuelto por Galileo Galilei 50 años después. Gerolamo predijo el día de su muerte, y para no fallar como lo había hecho en otras predicciones, se quito la vida ese día.

Abraham de Moivre escribió en el siglo XVII el libro "La doctrina de las probabilidades" (The Doctrine of Chances) que plantea teoremas de suma importancia en el desarrollo de la probabilidad. Como anécdota curiosa, Abraham también predijo el día de su muerte.

Posteriormente se utilizaron enfoques de frecuencia y personales, y se estableció el desarrollo axiomático de la probabilidad. Los enfoques o escuelas de la probabilidad se estudian a continuación.

ENFOQUES DE LA PROBABILIDAD

La probabilidad tiene tres enfoques, interpretaciones o escuelas.

1. Interpretación Clásica

La interpretación clásica de la probabilidad dice que si el espacio muestral consta de $\mathrm{n}(S)$ eventos simples, todos igualmente factibles, y si el evento A tiene $\mathrm{n}(A)$ eventos simples, entonces la probabilidad del evento A, denotada por $P(A)$, se calcula como el cociente $\dfrac{\mathrm{n}(A)}{\mathrm{n}(S)}$.

$$P(A) = \frac{\mathrm{n}(A)}{\mathrm{n}(S)}$$

Los inconvenientes de la interpretación clásica son:

a. Es necesario que todos los resultados del experimento tengan la misma probabilidad.
b. En ocasiones es muy difícil o imposible determinar la cardinalidad de A y de S.

2. Interpretación Frecuentista

La interpretación frecuentista de la probabilidad dice que la probabilidad de un evento A es la frecuencia relativa con la que ha ocurrido dicho evento en un número muy grande de experimentaciones.

$$P(A) = \lim_{N \to \infty} \frac{\mathrm{fr}(A)}{N}$$

donde $\mathrm{fr}(A)$ es la frecuencia o número de veces que se observa el evento A en N repeticiones del experimento.

La interpretación frecuentista también se conoce como de Von Mises, en honor a Richard Von Mises, matemático y físico austrohúngaro.
El principal inconveniente del enfoque frecuentista es que no es aplicable cuando el experimento no se puede repetir un gran número de veces.
Mientras más grande sea el número de repeticiones del experimento, se obtendrá una mejor aproximación de la probabilidad del evento A.

3. Interpretación Subjetiva
Es una medida entre 0 y 1 (inclusive), que se asocia de manera subjetiva.
Se requiere de una persona experimentada para determinar la probabilidad.

El inconveniente de este enfoque es que suele cambiar el valor asignado cuando se le pregunta a otro experto.

A la interpretación clásica también se le conoce como a priori y de Laplace, a la interpretación frecuentista se le llama también de frecuencia relativa, objetiva, empírica, a posteriori y estadística; y a la interpretación subjetiva también se le llama personal.

La definición de la función de probabilidad consiste en tres axiomas.

Definición Axiomática

Los axiomas son verdades tan evidentes que se admiten sin demostración. La probabilidad está basada en los axiomas de Probabilidad desarrollados en 1933 por el matemático Ruso Kolmogórov.

Para un espacio muestral S, la probabilidad de cada evento $A \subset S$ es un número que se asigna al evento, se denota $P(A)$, y tiene las siguientes propiedades:

1. $P(A)$ es un número no negativo; esto es, $P(A) \geq 0$

2. La probabilidad del espacio muestral S es la unidad, esto es:
$$P(S) = 1$$
3. Si A y B son eventos que se excluyen mutuamente en S, entonces la probabilidad de la unión de los eventos es igual a la suma de sus probabilidades; es decir:

$$P(A \cup B) = P(A) + P(B)$$

Debe observarse que los axiomas de la probabilidad no relacionan la probabilidad con los experimentos en sí. Los axiomas sólo sientan las bases que debe cumplir una función de probabilidad. La relación debe establecerse a través de alguno de los enfoques de la probabilidad.

Un punto muy importante que debe resaltarse es el hecho de que si un evento tiene probabilidad cero, en el axioma uno, no significa que sea un evento imposible. Para tener el evento vacío (o imposible) una probabilidad de cero es condición necesaria, pero no suficiente.

Ejemplo 1.21

Si se extrae una carta de un paquete bien barajado de 52 cartas. Determinar la probabilidad de obtener:

a) Un rey rojo.
b) Un 3, 4, 5 ó 6.
c) Una carta negra.
d) Un as rojo o una reina negra.

Resolución

a) Sea R el evento en el cual se obtiene un rey rojo.

$$P(R) = \frac{n(R)}{n(S)} = \frac{2}{52} = \frac{1}{26}$$

b) Sea A el evento en el cual se obtiene un 3, 4, 5 ó 6.

$$P(A) = \frac{n(A)}{n(S)} = \frac{16}{52} = \frac{4}{13}$$

c) Sea N el evento en el cual se obtiene una carta negra.

$$P(N) = \frac{n(N)}{n(S)} = \frac{26}{52} = \frac{1}{2}$$

d) Sea B el evento en el cual se obtiene un as rojo o una reina negra.

$$P(B) = \frac{n(B)}{n(S)} = \frac{4}{52} = \frac{1}{13}$$

Ejemplo 1.22

Una familia tiene 3 hijos. Determinar todas las posibles permutaciones, con respecto al sexo de los hijos. Bajo suposiciones adecuadas ¿cuál es la probabilidad de que, exactamente dos de los hijos tengan el mismo sexo? ¿Cuál es la probabilidad de tener un varón y dos mujeres? ¿Cuál es la probabilidad de tener 3 hijos del mismo sexo?

Resolución

Denotando H para hombre y M para mujer, las posibles permutaciones son:

HHH, HHM, HMH, MHH, MHM, MMH, HMM, MMM

Considerando que cada permutación tiene la misma probabilidad, y si A es el evento en el cual se tienen exactamente dos hijos del mismo sexo:

$$P(A) = \frac{6}{8} = \frac{3}{4}$$

Sea B el evento en el cual se tienen un varón y dos mujeres

$$P(B) = \frac{3}{8}$$

Sea C el evento en el cual se tienen tres hijos del mismo sexo.

$$P(C) = \frac{2}{8} = \frac{1}{4}$$

Para resolver los ejemplos anteriores se utilizó el enfoque clásico de la probabilidad, en el cual se supone que cada carta tiene la misma probabilidad de ocurrencia, o que cada punto posible de la combinación de hijos e hijas tiene la misma probabilidad; cuando se dispone de información histórica puede utilizarse el enfoque frecuentista, como lo ilustra el siguiente ejemplo.

Ejemplo 1.23

Se toma la medida en centímetros de 200 jugadores de básquetbol, y se construye la siguiente distribución de frecuencias.

Estatura *cm*	Frecuencia
171 - 180	10
181 - 190	12
191 - 200	25
201 - 210	132
211 - 220	21

Si se selecciona un jugador de este grupo al azar, calcular la probabilidad para cada uno de los siguientes eventos

a) A: Mida entre 191 y 200 [cm]

b) B: Mida entre 201 y 210 [cm]

c) C: Mida entre 191 y 210 [cm]

d) D: Mida a lo sumo 200 [cm]

Resolución

Con $N = 200$ observaciones

a) $$P(A) = \frac{\text{fr}(A)}{N} = \frac{25}{200} = 0.125$$

b) $$P(B) = \frac{\text{fr}(B)}{N} = \frac{132}{200} = 0.66$$

c) $$P(C) = P(A \cup B) = \frac{\text{fr}(A \cup B)}{N} = \frac{25 + 132}{200} = 0.785$$

d) $$P(D) = \frac{\text{fr}(D)}{N} = \frac{10 + 12 + 25}{200} = \frac{47}{200} = 0.235$$

La probabilidad[5] es una función que tiene como dominio el conjunto de todos los eventos de un espacio

[5]La definición formal para un curso intermedio de probabilidad diría que un espacio de probabilidad es una terna (Ω, F, P) en donde Ω es un conjunto arbitrario, F es una σ-álgebra de subconjuntos de Ω y $P : F \rightarrow [0, 1]$ es una medida de probabilidad.

muestral (conjunto potencia de S), y como recorrido el intervalo de números reales $[0,1]$. Se denota por

$$P : 2^S \rightarrow [0,1]$$

Teorema 1.9 Teoremas Elementales de la Probabilidad
1) $P(\varnothing) = 0$
2) Si A es un evento cualquiera, entonces $\overline{A}$ también lo es, y su probabilidad es
$P(\overline{A}) = 1 - P(A)$
3) Sean A y B dos eventos cualesquiera
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Demostración de 1

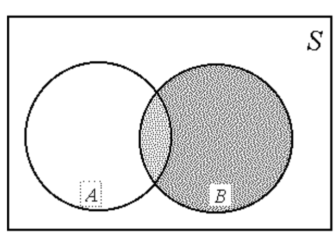
Cualquier evento A puede escribirse como
$A = A \cup \varnothing$
Por lo que
$P(A) = P(A \cup \varnothing)$
Puesto que A y $\varnothing$ son excluyentes, $A \cap \varnothing = \varnothing$, y por el axioma 3
$P(A) = P(A \cup \varnothing) = P(A) + P(\varnothing)$
$P(\varnothing) = P(A) - P(A) = 0$

Demostración de 2

El espacio muestral puede escribirse como
$S = \overline{A} \cup A$
de donde
$P(S) = P(\overline{A} \cup A)$
por el axioma 3
$P(S) = P(\overline{A}) + P(A)$
y del axioma 2
$1 = P(\overline{A}) + P(A)$
Finalmente
$P(\overline{A}) = 1 - P(A)$

Demostración de 3

Se parte del diagrama de Venn, y se descompone la unión



$A \cup B = A \cup (B \cap \overline{A})$

$B = (A \cap B) \cup (B \cap \overline{A})$

Por lo tanto

$P(A \cup B) = P(A) + P(B \cap \overline{A})$

$P(B) = P(A \cap B) + P(B \cap \overline{A})$

Restando la segunda ecuación de la primera se tiene:

$P(A \cup B) - P(B) = P(A) - P(A \cap B)$

de donde se obtiene la expresión buscada.

■

El teorema tres puede generalizarse para tres o más eventos. El siguiente teorema muestra la probabilidad de la unión de tres eventos.

Teorema 1.10

Sean A , B y C tres eventos cualesquiera, entonces

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

Teorema 1.11

Si $A \subset B$, entonces $P(A) \leq P(B)$.

Ejemplo 1.24

Un sistema electrónico consta de dos aparatos A y B, y para poder operar requiere que ambos aparatos funcionen. La probabilidad de que el aparato A funcione en un momento dado es 0.9 y la probabilidad de que el aparato B funcione es 0.85. La probabilidad de que ambos aparatos fallen simultáneamente es de 0.05. Calcular la probabilidad de que en un momento dado el sistema funcione.

Resolución

Sean los eventos

$A \triangleq$ el aparato A funciona.

$B \triangleq$ el aparato B funciona.

Datos: $P(A) = 0.9$

$P(B) = 0.85$

$P(\overline{A} \cap \overline{B}) = 0.05$

Análisis

$P(A \cap B) = P(A) + P(B) - P(A \cup B)$ . . . (a)

De las leyes de De Morgan[6] se sabe que $\overline{(A \cup B)} = \overline{A} \cap \overline{B}$, por lo que

$P[\overline{(A \cup B)}] = P(\overline{A} \cap \overline{B}) = 0.05$

$P(A \cup B) = 1 - P[\overline{(A \cup B)}] = 1 - 0.05$

$P(A \cup B) = 0.95$

sustituyendo en (a)

$P(A \cap B) = 0.9 + 0.85 - 0.95 = 0.8$

La probabilidad de que el sistema funcione es de 0.8.

Ejemplo 1.25

En un laboratorio de computadoras se tienen 15 computadoras de las cuales cinco están obsoletas, si una persona toma al azar tres de ellas, ¿cuál es la probabilidad de que por lo menos una de las tres esté obsoleta?

Resolución

Sean los eventos:

$A \triangleq$ Al menos una de las tres computadoras está obsoleta.

$D_1 \triangleq$ Exactamente una computadora está obsoleta (y dos no).

$D_2 \triangleq$ Exactamente dos computadoras están obsoletas (y una no).

$D_3 \triangleq$ Exactamente las tres computadoras están obsoletas.

Análisis:

Es claro que la probabilidad buscada es:

$P(A) = P(D_1 \cup D_2 \cup D_3)$

Y puesto que son eventos mutuamente excluyentes:

$P(A) = P(D_1) + P(D_2) + P(D_3)$

Para calcular cada una de las probabilidades se utiliza el enfoque clásico, considerando que la selección de cualquiera de las computadoras tiene la misma probabilidad. Para realizar el conteo se utilizan combinaciones y la regla de la multiplicación.

$$P(D_1) = \frac{\binom{5}{1}\binom{10}{2}}{\binom{15}{3}} = \frac{5(45)}{455} \approx 0.4945$$

$$P(D_2) = \frac{\binom{5}{2}\binom{10}{1}}{\binom{15}{3}} \approx 0.2197$$

---

[6] Si no se recuerdan las leyes de De Morgan, debe repasarse el apéndice A de este capítulo.

$$P(D_3) = \frac{\binom{5}{3}\binom{10}{0}}{\binom{15}{3}} \approx 0.02197$$

Finalmente

$$P(A) \approx 0.4945 + 0.2197 + 0.02197 \approx 0.7363$$

La probabilidad de que por lo menos una esté obsoleta es 0.7363.

Ejemplo 1.26

Una bolsa con premios contiene 500 sobres, 75 de los cuales contienen \$1000 en efectivo, 150 contienen \$250 y 275 contienen \$100.

a) ¿Cuál es el espacio muestral para las diferentes cantidades de dinero?

b) Asignar probabilidades a los elementos del espacio muestral y después obtener la probabilidad de que el primer sobre que se seleccione contenga menos de \$1000

Resolución

a) El espacio muestral para las diferentes cantidades de dinero es:

$$S = \{\ \$100\ ,\ \$250\ ,\ \$1000\ \}$$

b) Las probabilidades para la primera extracción son, respectivamente: $\frac{275}{500}$, $\frac{150}{500}$ y $\frac{75}{500}$

Sea el evento A, el primer sobre seleccionado contiene menos de \$1000, entonces, considerando los sobres de \$100 y de \$250, se tiene:

$$P(A) = \frac{275 + 150}{500} = \frac{425}{500} = 0.85$$

Ejemplo 1.27

Si se seleccionan al azar tres libros de un estante que contiene cinco libros de cálculo, tres libros de termodinámica y un diccionario, determinar las siguientes probabilidades:

a) Que se tome el diccionario.

b) Que se escojan dos libros de cálculo y un libro de termodinámica.

Resolución

a) $A \triangleq$ Se toma el diccionario

$$P(A) = \frac{\binom{1}{1}\binom{8}{2}}{\binom{9}{3}} = \frac{1}{3}$$

b) $B \triangleq$ Se toman dos libros de cálculo y un libro de termodinámica.

$$P(B) = \frac{\binom{5}{2}\binom{3}{1}\binom{1}{0}}{\binom{9}{3}} = \frac{30}{84} \approx 0.3571$$

Ejemplo 1.28

La probabilidad de que una industria transnacional se ubique en México es de 0.7; de que se localice en Argentina 0.4, y de que se encuentre en al menos uno de los dos países es de 0.8. Determinar la probabilidad de que la industria se localice en:
a) ambas ciudades,
b) ninguna de las ciudades.

Resolución
Sean los eventos:
$M \triangleq$ La industria se localiza en México.
$A \triangleq$ La industria se localiza en Argentina.
a) Que se localice en ambas ciudades significa que se localiza en México y en Argentina, por lo que se tiene una intersección.

$$P(M \cap A) = P(M) + P(A) - P(M \cup A)$$
$$P(M \cap A) = 0.7 + 0.4 - 0.8 = 0.3$$

b) De forma similar, pero utilizando leyes de De Morgan y el teorema del complemento se tiene:

$$P(\bar{M} \cap \bar{A}) = 1 - P(M \cup A) = 1 - 0.8 = 0.2$$

Ejemplo 1.29
Un circuito electrónico se selecciona al azar de un lote de 1000 circuitos. Los defectos de manufactura se clasifican en tres diferentes tipos, denominados A, B y C. Los defectos tipo A ocurren el $2\ \%$ de las veces, los del tipo B el $1\ \%$, y los del tipo C el $1.5\ \%$. Además, se sabe que el $0.5\ \%$ tienen los defectos tipo A y B; el $0.6\ \%$, los defectos A y C, y el $0.4\ \%$ presenta los defectos B y C, en tanto que el $0.2\ \%$ tiene los tres defectos. ¿Cuál es la probabilidad de que el circuito flexible seleccionado tenga al menos uno de los tres tipos de defectos?

Resolución
Sean los eventos
$A \triangleq$ El circuito tiene el defecto tipo A.
$B \triangleq$ El circuito tiene el defecto tipo B.
$C \triangleq$ El circuito tiene el defecto tipo C.
Del enunciado se tienen las probabilidades:

$$P(A) = 0.02 \quad , \quad P(B) = 0.01 \quad , \quad P(C) = 0.015$$
$$P(A \cap B) = 0.005, \quad P(A \cap C) = 0.006$$
$$P(B \cap C) = 0.004, \quad P(A \cap B \cap C) = 0.002$$

Por lo que, para que tenga al menos uno de los defectos se busca una unión.

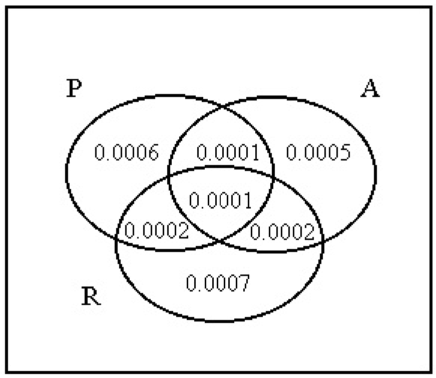
$$P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$$
$$P(A \cup B \cup C) = 0.02 + 0.01 + 0.015 - 0.005 - 0.006 - 0.004 + 0.002$$
$$P(A \cup B \cup C) = 0.032$$

Ejemplo 1.30

Supóngase que, en el mantenimiento de un enorme archivo médico que sirve para cobrar el seguro médico, la probabilidad de un error en el procesamiento de registros es 0.0010, la probabilidad de error en el archivado es 0.0009, la probabilidad de un error en la recuperación de datos es 0.0012, la probabilidad de un error en el procesamiento de registros y en el archivado es 0.0002, la probabilidad de un error en el procesamiento de registro, así como en recuperación es de 0.0003, la probabilidad de un error en archivado, así como en recuperación es de 0.0003 y la probabilidad de un error en el procesado de registros, en el archivado y la recuperación es de 0.0001. ¿Cuál es la probabilidad de incurrir al menos en uno de estos errores?

Resolución

Sea P el procesamiento, A el archivado y R la recuperación, del enunciado puede obtenerse el siguiente diagrama de Venn, en el que se indican las probabilidades para cada una de las intersecciones.



Incurrir en al menos uno de los errores es la unión de los tipos de errores, por lo que:

$P(P \cup A \cup R) = 0.0024$

PROBABILIDAD CONDICIONAL E INDEPENDENCIA ESTADÍSTICA

En diversos estudios se ha establecido cierta "relación" entre el fumar y el enfisema pulmonar, por ejemplo, dos médicos británicos reportaron que de 135 hombres afectados con enfisema 122 (90%) eran fumadores. En términos simples, la inferencia que puede obtenerse es: si una persona tiene enfisema pulmonar, es más probable que sea un fumador a que no lo sea; pero, ¿podrá hacerse la afirmación inversa? ¿Si una persona es fumadora, sus posibilidades de enfermarse serán mucho mayores que las de una persona que no lo es? Para contestar preguntas así, es necesario definir el concepto de probabilidad condicional, el cual es fundamental en el estudio de la probabilidad.

Definición 1.2

La probabilidad condicional de un evento A, dado que ya ocurrió el evento B, es:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$

siempre que $P(B) > 0$.

Justificación

Considerando eventos igualmente probables,

$$P(A \cap B) = \frac{n(A \cap B)}{n(S)} \quad , \qquad P(B) = \frac{n(B)}{n(S)}$$

sustituyendo en $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$

$$P(A \mid B) = \frac{\dfrac{n(A \cap B)}{n(S)}}{\dfrac{n(B)}{n(S)}} = \frac{n(A \cap B)}{n(B)}$$

La anterior justificación no es una demostración, la definición de la probabilidad condicional es eso, una definición; es decir no es un teorema ni un axioma. La función $P(A \mid B)$ cumple con las propiedades de una función de probabilidad.

Del planteamiento inicial, si consideramos que los datos reportados por los médicos provienen de una población de N personas, en donde nos interesan dos eventos:

$A \triangleq$ La persona es fumadora.
$B \triangleq$ La persona está enferma de enfisema pulmonar.

y las cardinalidades de los eventos son: $n(B) = 135$ y $n(B \cap A) = 122$

La probabilidad de seleccionar aleatoriamente a una persona fumadora, del conjunto de personas con enfisema es:

$$\frac{n(B \cap A)}{n(B)} = \frac{122}{135} = 0.9$$

En palabras se dice: "La probabilidad del evento A (la persona es fumadora) dado que el evento B (la persona está enferma de enfisema pulmonar) ha ocurrido, es 0.9".

En símbolos se escribe:

$$P(A \mid B) = 0.9$$

Por otro lado, la probabilidad de que una persona seleccionada aleatoriamente de la población sea fumadora y esté enferma de enfisema es:

$$P(A \cap B) = \frac{n(B \cap A)}{N}$$

de manera similar la probabilidad de que una persona seleccionada aleatoriamente de la población tenga enfisema es:

$$P(B) = \frac{n(B)}{N}$$

De donde

$$P(A \mid B) = \frac{n(B \cap A)}{n(B)} = \frac{\frac{n(B \cap A)}{N}}{\frac{n(B)}{N}} = \frac{P(A \cap B)}{P(B)}$$

que coincide con la definición dada anteriormente.

En los problemas que involucran probabilidades condicionales, dependiendo de la forma en la que se proporciona la información, puede convenir realizar una tabla con las probabilidades conocidas, hacer un diagrama de árbol o simplemente utilizar las definiciones y teoremas correspondientes. En el siguiente ejemplo se acomodan los elementos en una tabla, de forma tal que las probabilidades correspondientes pueden obtenerse directamente del cociente "casos a favor" entre "casos totales".

El siguiente ejemplo proporciona los datos en una tabla.

Ejemplo 1.31

En un seminario especial de la Facultad de Ingeniería hay 100 alumnos. Algunos de estos alumnos estudian (E) mientras otros no lo hacen ($\overline{E}$). Por otro lado y por la forma de evaluar del profesor, algunos alumnos acreditan el curso (A) mientras que otros no lo acreditan ($\overline{A}$). La tabla siguiente muestra el número de alumnos en cada categoría. Si se selecciona a un alumno al azar y se descubre que ha acreditado, ¿cuál es la probabilidad de que haya estudiado?

	E	$\overline{E}$	Total
A	60	10	70
$\overline{A}$	10	20	30
Total	70	30	100

Resolución

Utilizando las letras proporcionadas para la definición de eventos, se tiene:

$$P(E \mid A) = \frac{P(A \cap E)}{P(A)} = \frac{\frac{60}{100}}{\frac{70}{100}} = \frac{6}{7}$$

El siguiente ejemplo ilustra el uso directo de la definición de probabilidad condicional.

Ejemplo 1.32

La probabilidad de que un automóvil al que se le llena el tanque de gasolina necesite también un cambio de aceite es de 0.25; la de que requiera un nuevo filtro de aceite, 0.40 y de que le haga falta tanto cambio de aceite como de filtro, 0.14.

a) Si debe cambiarse el aceite ¿cuál es la probabilidad de que necesite un filtro nuevo?

b) Si necesita un filtro nuevo, ¿cuál es la probabilidad de que requiera que se le cambie el aceite?

Resolución

Sean los eventos

$A \triangleq$ Requiere un cambio de aceite.

$F \triangleq$ Requiere un cambio de filtro.

Del enunciado

$$P(A) = 0.25$$
$$P(F) = 0.40$$
$$P(A \cap F) = 0.14$$

a) $$P(F \mid A) = \frac{P(F \cap A)}{P(A)} = \frac{0.14}{0.25} = 0.56$$

b) $$P(A \mid F) = \frac{P(A \cap F)}{P(F)} = \frac{0.14}{0.40} = 0.35$$

La consecuencia más importante de la definición de probabilidad condicional se obtiene escribiéndola de la siguiente manera:

$$P(A \cap B) = P(B \mid A)\ P(A)$$

o bien

$$P(A \cap B) = P(A \mid B)\ P(B)$$

que se conoce como el teorema de multiplicación.

El siguiente ejemplo ilustra el uso de los diagramas de árbol para el cálculo de probabilidades utilizando el teorema de la multiplicación.

Ejemplo 1.33

Súpongase que en la compañía de seguros "El cóndor", el 65% de sus clientes son mayores de 35 años, y estos clientes tienen una probabilidad de accidentarse de 0.06, mientras que los de 35 años o menos, tienen una probabilidad de accidentarse de 0.15. Determinar la probabilidad que una persona se accidente y sea de 35 años o menos.

Resolución

Se definen los eventos:

$A \triangleq$ Mayor de 35 años.

$B \triangleq$ De 35 años o menos.

$C \triangleq$ Se accidenta.

Del enunciado las probabilidades son:

$P(A) = 0.65$

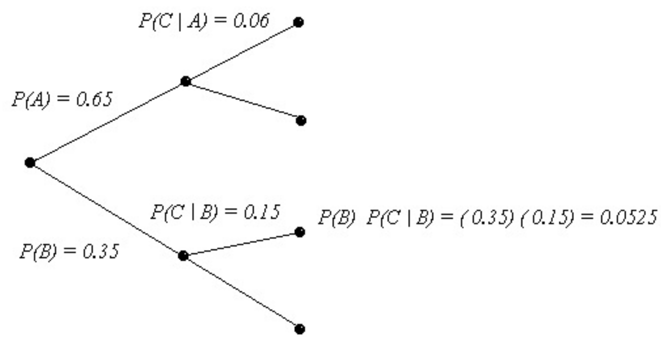
$P(B) = 0.35$ (Puesto que es el complemento de A)

$P(C \mid A) = 0.06$

$P(C \mid B) = 0.15$

y se busca la probabilidad de la intersección $P(C \cap B)$

Las probabilidades se pueden colocar en un diagrama de árbol de la siguiente manera.



La probabilidad buscada se obtiene de la multiplicación de las probabilidad es de las ramas de interés, esto es:

$$P(C\cap B) = P(B)\ P(C \mid B) = (0.35)\ (0.15) = 0.0525$$

Definición 1.3 Independencia

Dos eventos A y B son estadísticamente independientes, si y sólo si

$$P(A\cap B) = P(A)\ P(B)$$

En otro caso los eventos son dependientes.

Y de la definición de probabilidad condicional, la condición de independencia puede escribirse como:

Dos eventos son independientes si

$$P(A \mid B) = P(A)$$

y

$$P(B \mid A) = P(B)$$

En términos simples, la independencia de eventos significa que la ocurrencia o la no ocurrencia de un evento no tiene influencia en la ocurrencia o no ocurrencia de otro evento.

En un diagrama de Venn, los eventos independientes deben tener intersección. Si en un diagrama de Venn dos eventos no tienen intersección, entonces son invariablemente dependientes.

El siguiente ejemplo muestra el uso de la condición de independencia para el cálculo de probabilidades.

Ejemplo 1.34

Supóngase que A y B son dos eventos independientes asociados con un experimento. Si la probabilidad de que A o B ocurran es igual a 0.6, mientras que la probabilidad de que ocurra A es igual a 0.4, determinar la probabilidad de que B ocurra.

Resolución

Datos: $P(A \cup B) = 0.6$, $P(A) = 0.4$

Y puesto que A y B son independientes:

$P(A \cap B) = P(A)\ P(B)$

y del teorema de la unión de eventos

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

de donde

$P(A \cup B) = P(A) + P(B) - P(A)\ P(B)$

sustituyendo

$0.6 = 0.4 + P(B) - 0.4\ P(B)$

Despejando

$$P(B) = \frac{0.2}{0.6} = \frac{1}{3}$$

PROBABILIDAD TOTAL Y TEOREMA DE BAYES

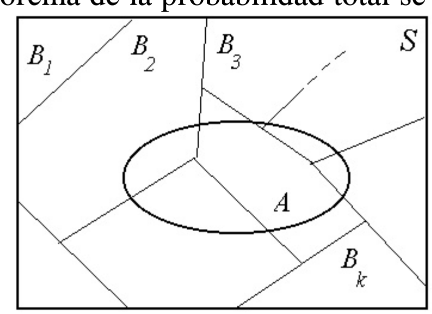
El teorema o regla de la multiplicación es útil para determinar la probabilidad de un evento que depende de otros, y esta idea se generaliza a través del teorema de la probabilidad total.

Teorema 1.12 Teorema de la Probabilidad Total

Sean los eventos $B_1, B_2, \ldots, B_k$ una partición exhaustiva y mutuamente excluyente del espacio muestral S, de tal forma que $P(B_i) \neq 0$ para $i = 1, 2, \ldots, k$, entonces para cualquier evento A en S

$$P(A) = \sum_{i=1}^{k} P(B_i \cap A) = \sum_{i=1}^{k} P(B_i)\ P(A \mid B_i)$$

En un diagrama de Venn, el teorema de la probabilidad total se ilustra de la siguiente forma.



Ejemplo 1.35

Una caja contiene 9000 componentes de los cuales 5% son defectuosos, una segunda caja contiene 7000 componentes de los cuales 40% son defectuosos. Existen otras dos cajas con 6000 componentes cada una y 10% de defectuosos. si se selecciona aleatoriamente una caja, y de ella una componente, ¿cuál es la probabilidad de que la componente seleccionada sea defectuosa?

Resolución
Sean los eventos:

$C_i \triangleq$ La caja seleccionada es la i-ésima, $i = 1, 2, 3, 4$.

$D \triangleq$ La componente seleccionada es defectuosa.

Del enunciado se tienen los datos:

$$P(C_i) = \frac{1}{4} \quad ; \quad i = 1, 2, 3, 4.$$

$$P(D \mid C_1) = 0.05 \qquad P(D \mid C_2) = 0.4$$

$$P(D \mid C_3) = 0.1 \qquad P(D \mid C_4) = 0.1$$

Por lo que utilizando el teorema de la probabilidad total

$$P(D) = \sum_{i=1}^{4} P(D \cap C_i)$$

$$P(D) = P(D \cap C_1) + P(D \cap C_2) + P(D \cap C_3) + P(D \cap C_4)$$

Y del teorema de la multiplicación

$$P(D \cap C_i) = P(C_i)\ P(D \mid C_i)$$

finalmente

$$P(D) = \frac{1}{4}(0.05 + 0.4 + 0.1 + 0.1) = 0.1625$$

Un teorema de gran importancia en la toma de decisiones y en la estadística es el Teorema de Bayes, formulado por el matemático inglés Thomas Bayes y que determina la probabilidad de un evento a través de los efectos observados.

Teorema 1.13 Teorema de Bayes

Sean los eventos $B_1, B_2, \ldots, B_k$ una partición exhaustiva y mutuamente excluyente del espacio muestral S, de tal forma que $P(B_i) \neq 0$ para $i = 1, 2, \ldots, k$, entonces para cualquier evento A en S tal que $P(A) \neq 0$ se tiene

$$P(B_r \mid A) = \frac{P(B_r \cap A)}{\sum_{i=1}^{k} P(B_i \cap A)} = \frac{P(B_r)\ P(A \mid B_r)}{\sum_{i=1}^{k} P(B_i)\ P(A \mid B_i)}$$

para $r = 1, 2, 3, \ldots, k$.

Este teorema es un consecuencia de la definición de probabilidad condicional y del teorema de probabilidad total.

Ejemplo 1.36
Del ejemplo anterior de las cajas con componentes.
¿Cuál es la probabilidad de que la caja seleccionada haya sido la segunda, si la componente seleccionada fue defectuosa?

Resolución

$$P(C_2 \mid D) = \frac{P(C_2 \cap D)}{P(D)} = \frac{P(C_2)P(D \mid C_2)}{\sum_{i=1}^{4} P(C_i)\ P(D \mid C_i)}$$

Y utilizando el resultado del ejemplo anterior

$$P(C_2 \mid D) = \frac{0.1}{0.1625} = 0.6154$$

Ejemplo 1.37

La policía planea reforzar el respeto a los límites de velocidad mediante la utilización de sistemas de radar en 4 diferentes sitios dentro de la ciudad. Los sistemas de radar en cada sitio L_1, L_2, L_3 y L_4 se ponen a funcionar, respectivamente, el 40%, 30%, 20% y 30% del tiempo, y si una persona que conduce a gran velocidad rumbo a su trabajo tiene, respectivamente, las probabilidades de 0.2, 0.1, 0.5 y 0.2 de pasar por alguno de estos sitios.

a) ¿cuál es la probabilidad de que le levanten una multa?

b) Si la persona recibe una infracción por conducir a gran velocidad rumbo a su trabajo, ¿cuál es la probabilidad de que haya pasado el radar que se localiza en el sitio L_2?

Resolución

Sea L_i el evento en el cual el conductor pasa por el radar i, $i = 1,2,3,4$; e I el evento en el cual le levantan una infracción(multa).

$P(L_1) = 0.2$ $\quad P(L_2) = 0.1$ $\quad P(L_3) = 0.5$ $\quad P(L_4) = 0.2$

$P(I \mid L_1) = 0.4$ $\quad P(I \mid L_2) = 0.3$ $\quad P(I \mid L_3) = 0.2$ $\quad P(I \mid L_4) = 0.3$

a) $$P(I) = \sum_{i=1}^{4} P(I \cap L_i)$$

$$P(I) = \sum_{i=1}^{4} P(L_i)\,P(I \mid L_i)$$

$$P(I) = P(L_1)\,P(I \mid L_1) + P(L_2)\,P(I \mid L_2) + \\ + P(L_3)\,P(I \mid L_3) + P(L_4)\,P(I \mid L_4)$$

Sustituyendo:

$$P(I) = (0.2)(0.4) + (0.1)(0.3) + (0.5)(0.2) + (0.2)(0.3)$$

$$P(I) = 0.27$$

b) $$P(L_2 \mid I) = \frac{P(L_2)\,P(I \mid L_2)}{P(I)}$$

$$P(L_2 \mid I) = \frac{0.1\,(0.3)}{0.27} = \frac{1}{9}$$

Ejemplo 1.38

Un inversionista está pensando en comprar un número muy grande de acciones de una compañía. La cotización de las acciones en la bolsa, durante los seis meses anteriores, es de gran interés para el inversionista. Con base en esta información, se observa que la cotización se relaciona con el producto nacional bruto. Si el PNB aumenta, la probabilidad de que el valor de las acciones aumente es de 0.8. Si el PNB es el mismo, la probabilidad de que las acciones aumenten su valor es de 0.2. Si el PNB disminuye, la probabilidad es de sólo 0.1. Si para los siguientes seis meses se asignan las probabilidades 0.4, 0.3 y 0.3 a los eventos; el PNB aumenta, es el mismo y disminuye, respectivamente, determinar la probabilidad de que las acciones aumenten su valor en los próximos seis meses.

Resolución

Sean los eventos:

$A \triangleq$ El PNB aumenta.

$M \triangleq$ El PNB es el mismo.

$D \triangleq$ El PNB disminuye.

$V \triangleq$ Las acciones aumentan su valor.

Del enunciado se tienen los datos:

$P(V|A) = 0.8 \quad , \quad P(A) = 0.4$

$P(V|M) = 0.2 \quad , \quad P(M) = 0.3$

$P(V|D) = 0.1 \quad , \quad P(D) = 0.3$

Empleando el teorema de probabilidad total, se tiene que:

$$\begin{aligned} P(V) &= P(V \cap A) + P(V \cap M) + P(V \cap D) \\ &= P(A)\,P(V \mid A) + P(M)\,P(V \mid M) + P(D)\,P(V \mid D) \\ &= (0.4)(0.8) + (0.3)(0.2) + (0.3)(0.1) \\ &= 0.41 \end{aligned}$$

Ejemplo 1.39

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, 0.4 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.

Resolución

Sean los eventos:

$A \triangleq$ La formación es de tipo I.

$B \triangleq$ La formación es de tipo II.

$C \triangleq$ La formación es de tipo III.

$E \triangleq$ Se encuentra petróleo.

Datos:

$P(A) = 0.35 \quad , \quad P(E|A) = 0.4$

$P(B) = 0.4$, $P(E \mid B) = 0.2$
$P(C) = 0.25$, $P(E \mid C) = 0.3$

$$P(B \mid \overline{E}) = \frac{P(B)\ P(\overline{E} \mid B)}{P(A)\ P(\overline{E} \mid A) + P(B)\ P(\overline{E} \mid B) + P(C)\ P(\overline{E} \mid C)}$$
$$= \frac{0.4\,(0.8)}{0.35\,(0.6) + 0.4\,(0.8) + 0.25\,(0.7)} = \frac{64}{141} \approx 0.4539$$

Ejemplo 1.40

Supóngase que la probabilidad de que los Patriotas de Nueva Inglaterra ganen el campeonato de la Conferencia Americana es de 0.25, y la probabilidad de que lo obtengan los Acereros de Pitsburgh es 0.20. Además, la probabilidad de que el campeón de la Conferencia Americana gane el Super Tazón es 0.45, 0.55 o 0.35, dependiendo de si los Patriotas, los Acereros o algún otro equipo gana el campeonato.

a) ¿Cuál es la probabilidad de que un equipo de la conferencia americana gane el Super Tazón?
b) Si un equipo de la Conferencia Americana gana el Super Tazón, ¿cuál es la probabilidad de que los Patriotas de Nueva Inglaterra ganen el titulo de su conferencia?

Resolución
Sean los eventos

$CA \triangleq$ Un equipo de la Conferencia Americana gana el Super Tazón.
$PN \triangleq$ Los Patriotas de Nueva Inglaterra ganan el campeonato de la Conferencia Americana.
$AP \triangleq$ Los Acereros de Pitsburgh ganan el campeonato de la Conferencia Americana.
$O \triangleq$ El campeonato de la Conferencia Americana no lo ganan ni los Patriotas ni los Acereros.

Datos:
$P(PN) = 0.25$ $P(CA \mid PN) = 0.45$
$P(AP) = 0.2$ $P(CA \mid AP) = 0.55$
$P(O) = 0.55$ $P(CA \mid O) = 0.35$

a) $P(CA) = P(PN)\ P(CA \mid PN) + P(AP)\ P(CA \mid AP) + P(O)\ P(CA \mid O)$
$= 0.25\,(0.45) + 0.2\,(0.55) + 0.55\,(0.35)$
$= 0.415$

b) $$P(PN \mid CA) = \frac{P(PN)\ P(CA \mid PN)}{P(CA)}$$
$$= \frac{0.25\,(0.45)}{0.415} = 0.271$$

Ejemplo 1.41
En cierta ciudad, durante el mes de mayo, la probabilidad de que un día lluvioso sea seguido por otro también lluvioso es de 0.8 y la probabilidad de que un día soleado anteceda a un día lluvioso es de 0.6. Suponiendo que cada día se clasifica como lluvioso o soleado y que el tiempo de cualquier día depende solamente del tiempo del día anterior, determinar la probabilidad de que en la ciudad citada un día lluvioso de mayo anteceda a dos días lluviosos más, luego, a un día soleado y finalmente a otro día lluvioso.

Resolución

Eventos:
$S_i \triangleq$ El día i es soleado.
$L_i \triangleq$ El día i es lluvioso.
$P(L_2\, L_3\, S_4\, L_5 \mid L_1) = (0.8)\,(0.8)\,(1 - 0.8)\,(0.6) = 0.0768$

AUTOEXAMEN TEMA I

1.- Una compañía planea construir cinco almacenes adicionales en sitios nuevos. Se consideran 10 sitios. El número total de elecciones que pueden considerarse es: $\square$

A) 50 B) 120 C) 252 D) $10!$ E) 30240

2.- El pedido de una computadora personal digital puede especificar uno de cinco tamaños de memoria, cualquier tipo de monitor de tres posibles, cualquier tamaño de disco duro de entre cuatro posibles y puede incluir o no una tableta para lápiz electrónico. El número de sistemas distintos que pueden ordenarse es: $\square$

A) 14 B) 60 C) 16 D) 120
E) Ninguno de los anteriores.

3.- Un contratista desea construir 9 casas, cada una con diseño diferente. El número de formas diferentes en las que puede ubicar estas casas si 6 terrenos están de un lado de la calle y 3 están en el lado opuesto es: $\square$

A) $(6)\,(3) = 18$ B) $(6!)\,(3!)\,(2) = 8640$ C) $\binom{9}{6}\binom{9}{3} = 7056$

D) $9! = 362880$ E) $P_6^9\ \ P_3^9 = 30481920$

4.- Un club de tenis está formado por cinco hombres y cuatro mujeres. Si un equipo de dobles mixtos se forma con un hombre y una mujer, el número de partidos de dobles mixtos que se pueden realizar con los miembros del club es:
A) 380 B) 40 C) 190 D) 120 E) 240

5.- Un examen de 20 preguntas, tiene 5 opciones por pregunta. El número de formas diferentes para contestar todo el examen es: $\square$

A) 100 B) 32×10^5 C) 15504 D) 5^{20}
E) Ninguna de las anteriores.

6.- De un grupo de 15 graduados en Ingeniería se seleccionan 10 de manera aleatoria. Sea P la probabilidad de que 4 de los mejores 5 graduados estén incluidos en la selección de 10, entonces la afirmación correcta es: $\square$

A) $0 \le P \le \frac{1}{5}$ B) $\frac{1}{5} < P \le \frac{2}{5}$ C) $\frac{2}{5} < P \le \frac{3}{5}$

D) $\frac{3}{5} < P \le \frac{4}{5}$ E) $\frac{4}{5} < P \le 1$

7.- Si $P(A) = 0.7$, $P(B) = 0.5$ y $P(\overline{A \cup B}) = 0.1$, entonces $P(A \mid B)$ es:

A) $\frac{3}{7}$ B) $\frac{3}{5}$ C) $\frac{5}{7}$ D) $\frac{7}{9}$ E)1

8.- Suponer que $P(A \mid B) = 0.8$, $P(A) = 0.5$ y $P(B) = 0.2$. Entonces $P(B \mid A)$.
A) 0.16 B) 0.1 C) 0.125 D) 0.32
E) Ninguna de las anteriores.

9.- La probabilidad del espacio muestral, $P(S)$, debe cumplir con: $\square$

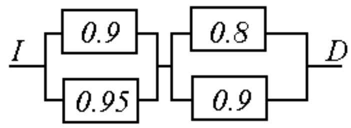
A) $P(S) < 1$ B) $P(S) = 0$ C) $P(S) = 1$
D) $0 < P(S) < 1$ E) $0 \le P(S) < 1$

10.- Un lote de 50 tornillos tiene 30 que son más gruesos que la dimensión requerida. Suponer que del lote se seleccionan 3 tornillos al azar, sin reemplazo. La probabilidad de que los tres tornillos sean más gruesos es: $\square$

A) 0.207 B) 0.06 C) 0.18 D) 0.02
E) Ninguna de las anteriores.

11.- Dados los eventos A y B, para los cuales $P(A) = 0.4$, $P(B) = 0.3$ y $P(A|B) = 0.1$, entonces $P(B|A)$ es: $\square$
A) 0.3 B) 0.03 C) 0.075
D) 0.9 E) Ninguna de los anteriores.

12.- El siguiente circuito trabaja si y sólo si, existe una trayectoria en funcionamiento de izquierda a derecha. El dibujo indica la probabilidad de que cada dispositivo funcione. Supóngase que la probabilidad de que un dispositivo funcione no depende del funcionamiento de los demás. Determinar la probabilidad de que el circuito funcione. $\square$



A) 0.6156 B) 0.3844 C) 0.975 D) 0.9751 E) 0.9999

13.- En una cierta gasolinería, se sabe que el 70% de los clientes utilizan Magna y que el 30% utilizan Premium. De los clientes que utilizan gasolina Magna, el 40% llena el tanque, mientras que el 55% de los que utilizan Premium llenan el tanque. La probabilidad de que un cliente cualquiera llene el tanque de gasolina es: $\square$

A) 0.43 B) 0.445 C) 0.95 D) 0.7272
E) Ninguna de las anteriores.

14.- Supóngase que la compañía de seguros "El cóndor" clasifica a las personas en alto y bajo riesgo. Los registros históricos indican que la probabilidad de que una persona clasificada en alto riesgo tenga un accidente es 0.3, mientras que para una persona de bajo riesgo es 0.08. Si el 70% de la población está clasificada como de bajo riesgo, entonces la probabilidad de que una persona que no tuvo accidente el año pasado esté clasificada como de bajo riesgo es: $\square$

A) 0.21 B) 0.246 C) 0.7 D) 0.754
E) Ninguna de las anteriores.

15.- Si los eventos A y B son mutuamente excluyentes entonces se puede afirmar sobre ellos que: $\square$

A) Son colectivamente exhaustivos.
B) Su unión es el vacío.
C) Son dependientes.
D) $P(A \mid B) = P(A)$
E) Ninguna de las anteriores.

Apéndice A CONJUNTOS

TEORÍA ELEMENTAL: CONJUNTOS Y SUBCONJUNTOS

CONCEPTO

La teoría de conjuntos se encuentra en los fundamentos de la matemática, y de la necesidad de entender algunos conceptos básicos que se utilizan en diversas asignaturas de ingeniería.

El concepto de conjunto es fundamental en la matemática; sin embargo, no existe una definición de conjunto[7]. Intuitivamente, se dice que un conjunto es un lista, colección o clase de objetos bien definidos. Los objetos se llaman elementos o miembros del conjunto.

Ejemplos de conjuntos
- Los alumnos de la Facultad de Ingeniería.
- Los Ingenieros Industriales.
- Los números pares.
- Los números 1, 3, 5, 7, 9, 11, . . .

NOTACIÓN

Los conjuntos se denotan por letras mayúsculas, generalmente las primeras del abecedario.

$$A,\ B,\ C,\ \dots$$

y los elementos de los conjuntos se representan por la misma letra del conjunto pero minúscula.

$$a,\ b,\ c,\ \dots$$

Un conjunto se define a través de los elementos que lo forman (pertenencia), si el elemento a pertenece al conjunto A entonces se escribe:

$$a \in A$$

Cuando un elemento, por ejemplo y, no pertenece al conjunto A se escribe:

$$y \notin A$$

Al escribir la asignación o igualdad de un conjunto se utilizan llaves, como se muestra a continuación

$$A = \{ 1 , 2 , 3 , 4 \}$$

o bien

$$B = \{ \ x \mid x \text{ es un número par } \}$$ [8]

Obsérvese la forma en la que se definieron los dos últimos conjuntos. Cuando se define un conjunto escribiendo los elementos que los forman y separándolos con comas, se utiliza la forma tabular, y se dice que el conjunto se define por extensión. Mientras que cuando el conjunto se define mediante una regla, se utiliza la forma constructiva y se dice que el conjunto se define por comprensión.

Ejercicio

---

[7] Otros conceptos fundamentales de las matemáticas no tienen definición formal, como por ejemplo: punto, recta, plano, etc.

[8] Recuérdese que la línea vertical " | " se lee "tal que".

Definir 5 conjuntos por comprensión y 5 por extensión.

CARDINALIDAD

La cardinalidad de un conjunto es el número de elementos que pertenecen a un conjunto. La cardinalidad se denota generalmente por n, aunque también se llega a utilizar el símbolo $\#$.

Ejemplo A.1

Sea el conjunto
$C = \{ x \mid x \text{ es un día de la semana} \}$
Obtener la cardinalidad de C.

Resolución

Es claro que el conjunto es
$C = \{ Lunes, Martes, Miércoles, Jueves, Viernes, Sábado, Domingo \}$
Por lo que
$n(C) = 7$

No todos los conjuntos tienen cardinalidad, esto se debe a que existen conjunto finitos e infinitos. Intuitivamente un conjunto es finito si consta de un cierto número de elementos distintos, es decir, que se pueda terminar de contar a los elementos del conjunto, en caso contrario se dice que el conjunto es infinito. Además existen conjuntos numerables y no numerables. Intuitivamente, un conjunto numerable es aquel que se puede contar, (o bien, poner en relación 1 a 1 con el conjunto de los números naturales) y un conjunto no numerable es aquel que no se puede contar.

Ejemplo A.2

Clasificar los siguientes conjuntos.
a) $A = \{ 1, 2, 3, 4 \}$
b) $B = \{ x \mid x \text{ es un número par} \}$
c) $C = \{ x \mid x \text{ es un número real entre } 0 \text{ y } 1 \}$

Resolución

a) Finito.
b) Infinito numerable.
c) Infinito no numerable.

IGUALDAD DE CONJUNTOS

Dos conjuntos A y B son iguales si ambos tiene los mismos elementos; es decir, si cada elemento que pertenece a A también pertenece a B y cada elemento que pertenece a B también pertenece a A.

La igualdad se denota con el símbolo $=$

$$A = B$$

Debe observarse que la igualdad de conjuntos está basada en los elementos que lo componen, no en la forma en la que se describe, así los siguientes pares de conjuntos son iguales.

$A = \{ x \mid x \text{ es un día de la semana} \}$
$B = \{ Lunes, Martes, Miércoles, Jueves, Viernes, Sábado, Domingo \}$

$C = \{ x \mid x^2 - 1 = 0 \}$
$D = \{ -1, 1 \}$

$E = \{ 1, 2, 3 \}$
$F = \{ 3, 2, 1, 3 \}$

En efecto los conjuntos E y F son iguales, debe entenderse que los elementos 1,2,3 son abstracciones, y no importa el orden en el que aparecen ni que se repitan. Más aún, el 3 es una abstracción, representa "el número 3", y aunque aparezca repetido es el mismo número 3 como abstracción y no es "otro" número 3, por lo que los elementos de los conjuntos son los mismos y su cardinalidad es también la misma $n(E) = n(F) = 3$.

Ejemplo A.3

Determinar cuáles de los siguientes conjuntos son iguales.
a) $A = \{ 1, 3 \}$
b) $B = \{ x \mid x \text{ es un número de un dado} \}$
c) $C = \{ x \mid x^2 - 4x + 3 = 0 \}$
d) $D = \{ x \mid x \text{ es el número de caras cuando se lanzan seis monedas} \}$

Resolución

$A = \{ 1, 3 \}$
$B = \{ 1, 2, 3, 4, 5, 6 \}$
$C = \{ 1, 3 \}$ puesto que son las raíces
$D = \{ 0, 1, 2, 3, 4, 5, 6 \}$
Por lo que los únicos eventos iguales son A y C, $A = C$.

SUBCONJUNTOS

Si todo elemento que pertenece al conjunto A, también es elemento del conjunto B entonces A es un subconjunto de B. Lo cual se denota $A \subset B$ (y se lee A es subconjunto de B).

Por ejemplo, si los conjuntos son:

$A = \{1, 2, 3, 4\}$
$B = \{5, 4, 3, 2, 1\}$

se observa con claridad que $A \subset B$.

Aunque es poco común, para el ejemplo anterior, también puede decirse que B es un superconjunto de A, lo cual se denota $B \supset A$.

La igualdad de conjuntos puede también definirse a partir de la definición de subconjuntos. Dos conjuntos A y B son iguales si y sólo si $A \subset B$ y $B \subset A$.

Todo conjunto es subconjunto de sí mismo; es decir, para el conjunto A se puede escribir $A \subset A$.

Puesto que todo conjunto es subconjunto de sí mismo, cuando A es subconjunto de B, pero A y B no son iguales, entonces se dice que A es un subconjunto propio de B.

Algunos textos utilizan el símbolo $\subseteq$ para indicar que A es un subconjunto de B, $A \subseteq B$; dejando el símbolo $\subset$ para indicar que A es subconjunto propio de B, $A \subset B$. En este curso se utilizará solamente el símbolo $\subset$, que no distingue entre subconjunto y subconjunto propio.

Para indicar que A no es un subconjunto de B se cancela el símbolo de subconjunto, teniéndose $A \not\subset B$.

CONJUNTO UNIVERSAL

Todos los conjuntos que se utilizan en un problema o en una investigación se consideran subconjuntos de un conjunto fijo llamado conjunto universal, el cual se denota generalmente por U.

Ejemplo A.4
Escribir los elementos de cada uno de los siguientes conjuntos universo.

a) El conjunto de los enteros entre 1 y 50 divisibles entre 8.
b) El conjunto $S = \{ x \mid x^2 + 4x - 5 = 0 \}$.
c) El conjunto de resultados cuando una moneda se lanza al aire hasta que resultan una águila o tres soles.
d) El conjunto $S = \{ x \mid x \text{ es un continente} \}$.
e) El conjunto $S = \{ x \mid 2x - 4 \geq 0 \;\; y \;\; x < 1 \}$.

Resolución

a) $S = \{8, 16, 24, 32, 40, 48\}$.
b) Las soluciones de $x^2 + 4x - 5 = 0$ son -5 y 1, por lo tanto $S = \{-5, 1\}$.
c) Denotando A para águila y O para sol.
$S = \{ A, OA, OOA, OOO \}$.
d) S = {América, África, Europa , Asia, Antártida y Oceanía}.

e) Resolviendo la inecuación (desigualdad) se tiene que $x \geq 2$ pero de la condición $x < 1$, se concluye que $S = \varnothing$.

Ejemplo A.5

Describir por comprensión el conjunto universo S consistente de todos los puntos en el primer cuadrante, dentro de un círculo de radio 3 con centro en el origen.

Resolución

$$S = \{ (x, y) \mid x^2 + y^2 < 9 ; x > 0 , y > 0 \}$$

Ejemplo A.6

Un experimento consiste en lanzar primero un dado y después lanzar una moneda, siempre y cuando el número en el dado sea par. Si el resultado del dado en non, la moneda se lanza dos veces. Al utilizar la notación 4H, por ejemplo, se indica el elemento donde el número resultante en el dado es un 4 y la moneda cae cara; 3HT para señalar el elemento de que el dado cae 3 y en el la moneda se observan una cara y una cruz. Describir el conjunto universo completo.

Resolución

S = {1HH, 1HT, 1TH, 1TT, 2H, 2T, 3HH, 3HT, 3TH, 3TT, 4H, 4T, 5HH, 5HT, 5TH, 5TT, 6H, 6T}

Obsérvese que el conjunto universo está formado por 18 elementos.

Ejemplo A.7

Un experimento consiste en lanzar un par de dados, uno verde y uno rojo y registrar los números que resultan. Si x es el resultado del dado verde y y el dado rojo, describir el conjunto universo S.

a) Por extensión (lista de los elementos (x, y)).

b) Por comprensión (utilizando el método de la regla).

Resolución

a)

$$S = \left\{ \begin{array}{l} (1,1) , (1,2) , \ldots , (1,6) \\ (2,1) , (2,2) , \ldots , (2,6) \\ \quad\quad\quad . \\ \quad\quad\quad . \\ \quad\quad\quad . \\ (6,1) , (6,2) , \ldots , (6,6) \end{array} \right\}$$

b) $S = \{ (x,y) \mid x = 1,2 \ldots ,6 \ , \ y = 1,2, \ldots 6 \}$

CONJUNTO VACÍO

Intuitivamente, el conjunto vacío es aquel que carece de elementos.

Por ejemplo, si A es el conjunto de personas mayores de 300 años con vida, entonces A es vacío.

En matemáticas, otro ejemplo sería

$$B = \{ x \mid x = x + 1 \}$$

y es vacío puesto que no existe un número que sea igual a sí mismo más uno.

El conjunto vacío se denota por: $\varnothing$, o bien por { }.

El conjunto vacío $\varnothing$ se considera subconjunto de cualquier conjunto.

Debe observarse la diferencia entre los conjuntos $\varnothing$ y $\{\varnothing\}$. El primero es el conjunto vacío, mientras que el segundo es el conjunto que contiene al conjunto vacío. De hecho, es un conjunto de conjuntos, concepto que no se desarrolla en el curso propedéutico. La diferencia es difícil de entender y deberá considerarse como una definición.

COMPLEMENTO

El complemento de un conjunto A, que se denota mediante $\overline{A}$, A^c o A' es el conjunto de todos los elementos que no pertenecen a A.

Ejemplo A.8
Si el conjunto universo es $U = \{1, 2, 3, 4, 5, 6, 7, 8\}$ y el conjunto $A = \{2, 4, 6\}$, entonces el complemento de A es:
$\overline{A} = \{1, 3, 5, 7, 8\}$

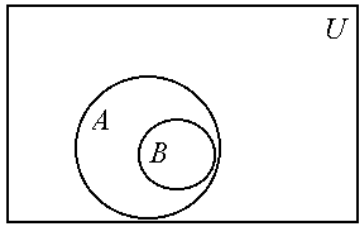
CONJUNTOS DISJUNTOS O EXCLUYENTES

Si dos conjuntos A y B no tienen ningún elemento en común, es decir, ningún elemento de A está en B y viceversa, entonces se dice que A y B son disjuntos.

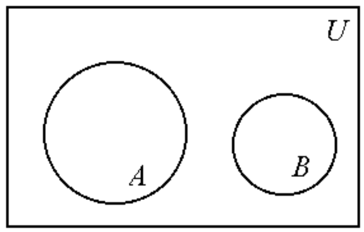
DIAGRAMAS DE VENN-EULER

Una manera sencilla de ilustrar los conjuntos y subconjuntos, así como las relaciones entre ellos es mediante la utilización de los diagramas de Venn-Euler. Los diagramas de Venn-Euler, o simplemente diagramas de Venn, son las representación de los conjuntos mediante un área plana, por lo general el área se delimita por una circunferencia.

En la siguiente figura se ilustra el conjunto universo, U; y los conjuntos A y B, donde B es subconjunto de A.



En la siguiente figura se ilustra el conjunto universo, U; y los conjuntos A y B, donde A y B son conjuntos disjuntos.



CONJUNTO POTENCIA

La familia de todos los subconjuntos de un conjunto A se llama conjunto potencia, y se denota generalmente por 2^A.

Ejemplo A.9
Si $A = \{1, 2\}$, entonces el conjunto potencia de A es
$2^A = \{\varnothing, \{1\}, \{2\}, \{1,2\}\}$

La cardinalidad del conjunto potencia de A está dada por $n(2^A) = 2^{n(A)}$

El conjunto potencia es un conjunto de conjuntos, es decir, sus elementos son a su vez conjuntos.

Obsérvese que los conjuntos que forman al conjunto potencia de un conjunto A, incluyen al conjunto vacío (puesto que el vacío es subconjunto de cualquier conjunto) y al mismo conjunto A (puesto que todo conjunto es subconjunto de sí mismo).

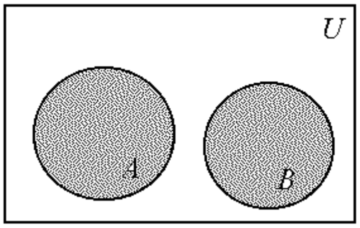
OPERACIONES FUNDAMENTALES CON CONJUNTOS

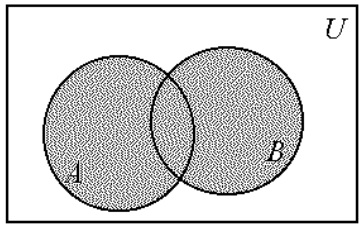
UNIÓN

La unión de dos conjuntos A y B es el conjunto de todos los elementos que pertenecen a A o a B. Se denota por $A \cup B$.

La unión es una operación conmutativa, es decir: $A \cup B = B \cup A$

En un diagrama de Venn-Euler, dados dos conjuntos A y B, la unión de ambos se puede indicar mediante un rayado, coloreado, punteado, etc.





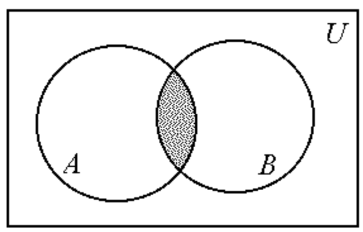
INTERSECCIÓN

La intersección de dos conjuntos A y B es el conjunto de los elementos que pertenecen tanto a A como a B. Se denota por $A \cap B$.

Dos conjuntos disjuntos no tienen intersección.

La intersección es una operación conmutativa, es decir: $A \cap B = B \cap A$

En un diagrama de Venn-Euler, la intersección es el área común de los dos conjuntos.

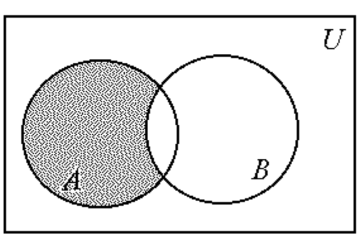


DIFERENCIA

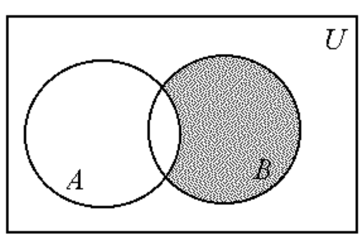
La diferencia de dos conjuntos A y B es el conjunto de los elementos que pertenecen a A, pero no a B. Se denota por $A - B$. La diferencia de dos conjuntos no es una operación conmutativa, es decir, en general $A - B \neq B - A$.

En ocasiones la diferencia de dos conjuntos se denota por: A / B, o bien $A \sim B$; pero es preferible utilizar la notación tradicional $A - B$. Ejemplo con diagramas de Venn-Euler.

$A - B$



$B - A$



La diferencia entre dos conjuntos no es una operación que se acostumbre utilizar en probabilidad.

Ejemplo A.10

Si el conjunto universal se define como
$U = \{1, 2, 3, 4, 5, 6, 7\}$ y se definen los subconjuntos
$A = \{1, 2, 3\}$
$B = \{2, 4, 6\}$
$C = \{1, 3, 5, 7\}$
Determinar:

a) $\overline{A}$
b) $\overline{B}$
c) La relación entre $\overline{C}$ y B
d) $A \cup B$
e) $A \cup C$
f) $B \cup C$
g) $A \cap B$
h) $A \cap C$
i) $B \cap C$
j) $U - A$
k) $A - B$
l) $\overline{A} - \overline{B}$

Resolución

a) $\overline{A} = \{4, 5, 6, 7\}$
b) $\overline{B} = \{1, 3, 5, 7\}$
c) Puesto que $\overline{C} = \{2, 4, 6\}$, entonces $\overline{C} = B$
d) $A \cup B = \{1, 2, 3, 4, 6\}$
e) $A \cup C = \{1, 2, 3, 5, 7\}$
f) $B \cup C = U$, es decir, el conjunto universo
g) $A \cap B = \{2\}$
h) $A \cap C = \{1, 3\}$
i) $B \cap C = \varnothing$
j) $U - A = \overline{A} = \{4, 5, 6, 7\}$
En general, para cualquier conjunto A, $U - A = \overline{A}$.
k) $A - B = \{1, 3\}$
l) $\overline{A} - \overline{B} = \{4, 6\}$

LEYES DE CONJUNTOS

Las operaciones de conjuntos pueden extenderse al involucrar cualquier número finito de conjuntos, en particular al utilizar la unión e intersección para tres conjuntos se tienen las siguientes leyes:

Leyes de identidad

$A \cup \varnothing = A$ $\qquad A \cap U = A$

$A \cup U = U$ $\qquad A \cap \varnothing = \varnothing$

Leyes asociativas

$$A \cup (B \cup C) = (A \cup B) \cup C$$
$$A \cap (B \cap C) = (A \cap B) \cap C$$

Leyes distributivas

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

Leyes de De Morgan

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$
$$\overline{A \cap B} = \overline{A} \cup \overline{B}$$

PRODUCTO CARTESIANO

Si A y B son dos conjuntos, entonces el conjunto de todos los pares ordenados (a, b) tales que $a \in A$ y $b \in B$ se denomina conjunto del producto cartesiano de A y B, o simplemente producto cartesiano. El producto cartesiano se denota por $A \times B$, y por comprensión se define como:

$$A \times B = \{ (a, b) \mid a \in A \ , \ b \in B \}$$

El producto cartesiano no es conmutativo.

La cardinalidad del producto cartesiano es $n(A \times B) = \left(n(A) \right) \left(n(B) \right)$

Ejemplo A.11
Sean los conjuntos $A = \{ 1 , 2 , 3 \}$ y $B = \{ 3 , 4 \}$
Obtener $A \times B$
Resolución

$$A \times B = \left\{ \begin{array}{c} (1,3) \ , \ (1,4) \ , \\ (2,3) \ , \ (2,4) \ , \\ (3,3) \ , \ (3,4) \end{array} \right\}$$

Ejemplo A.12
Un experimento consiste en lanzar un par de dados, uno verde y uno rojo y registrar los números que resultan. Sea x es el resultado del dado verde y y el dado rojo, listar los elementos que corresponden a los siguientes conjuntos.

a) A, en el que la suma sea mayor que 8.
b) B, de que ocurra un dos en cualquiera de los dados.
c) C, de que se obtiene un número mayor de 4 en el dado verde.
d) $A \cap C$.
e) $A \cap B$.
f) $B \cap C$.

Respuestas

a) A = {(3,6), (4,5), (4,6), (5,4), (5,5), (5,6), (6,3), (6,4), (6,5), (6,6)}

b) B = {(1,2), (2,2), (3,2), (4,2), (5,2), (6,2), (2,1), (2,3), (2,4), (2,5), (2,6)}

c) C = {(5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)}

d) $A \cap C$ = {(5,4), (5,5), (5,6), (6,3), (6,4), (6,5)}

e) $A \cap B = \varnothing$

f) $B \cap C$ = {(5,2), (6,2)}

Apéndice B TEOREMA DEL BINOMIO

Una suma de dos elementos $a + b$ se denomina binomio. Es común utilizar n binomio elevado a una potencia natural, $(a + b)^n$; y el teorema del binomio proporciona una fórmula que permite desarrollar la expresión anterior en una suma.

Teorema 1.8 (Del Binomio)

El desarrollo del binomio $(a + b)^n$ está dado por la expresión

$$(a + b)^n = \sum_{r=0}^{n} \binom{n}{r} a^r \, b^{n-r}$$

Y los números combinatorios $\binom{n}{r}$ reciben el nombre de coeficientes binomiales.

Ejemplo B.1

Demostrar que $\sum_{i=0}^{n} \binom{n}{i} = 2^n$.

Resolución

Del Binomio de Newton se sabe que

$$(a + b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$$

por lo que $$(1 + 1)^n = \sum_{i=0}^{n} \binom{n}{i} 1^i 1^{n-i}$$

es decir: $$2^n = \sum_{i=0}^{n} \binom{n}{i}$$

BIBLIOGRAFÍA

Hines, William W. y Montgomery, Douglas C. - Probabilidad y Estadística para ingeniería, cuarta edición.- CECSA.- México, 2005.

Milton, Susan J. Y Arnold, Jesse C.- Probabilidad y Estadística para con aplicaciones para ingeniería y ciencias computacionales, cuarta edición.- McGraw-Hill.- México, 2004.

Devore, Jay L.- Probabilidad y Estadística para ingeniería y ciencias, séptima edición.- Cengage Learning.- México, 2008.

Mendenhall, William III. et al.- Introducción a la Probabilidad y Estadística.- Décimo cuarta edición.- Cengage Learning.- México 2015.

Wackerly Dennis D.- Mendenhall, William, et al.- Estadística Matemática con Aplicaciones, sexta edición.- Editorial Thomson.- México, 2002.

Walpole, Ronald E., et al.- Probability and Statistics for Engineers and Scientists.- Pearson.- USA, 2007.

Montgomery, Douglas C. y Runger, George C.-Probabilidad y Estadística aplicadas a la Ingeniería, segunda edición.- Limusa-Wiley.- México, 2002.

Scheaffer, Richard L. y McClave, James T.- Probabilidad y Estadística para Ingeniería.- Grupo Editorial Iberoamérica.- México, 1993.

Canavos, George C.- Probabilidad y Estadística Aplicaciones y Métodos.- McGraw-Hill.- México, 1988.

Meyer, Paul L.- Probabilidad y Aplicaciones Estadísticas.- Addison Wesley Iberoamericana.- México, 1992.

Spiegel, Murray R. et al.- Probabilidad y Estadística, cuarta edición.- Mc Graw-Hill.-México 2013.

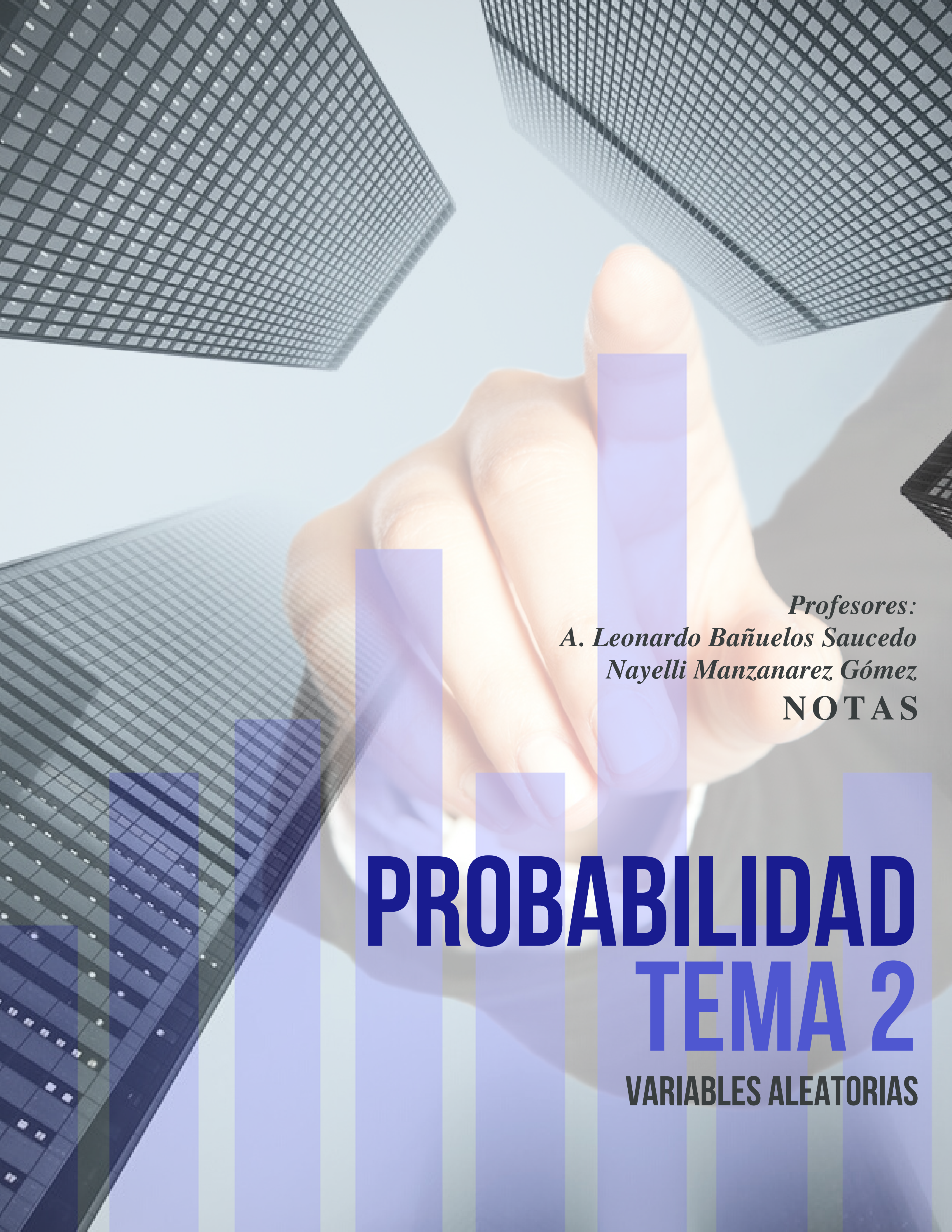
Borras García, Hugo E., et al.- Apuntes de Probabilidad y Estadística.-Facultad de Ingeniería.- México, 1985.

Rosenkrantz, Walter A.- Introduction to Probability and Statistics for Scientists and Engineers.- McGraw-Hill.- EE.UU., 1997.

Ziemer, Rodger E.- Elements of Engineering Probability & Statistics.- Prentice Hall.- USA 1997.

Lipschutz, Seymour.- Teoría de Conjuntos y Temas Afines.- McGraw-Hill.- México, 1991.

Aguilar Juárez, Patricia.- Apuntes de clase, FI, UNAM.



Profesores:
A. Leonardo Bañuelos Saucedo
Nayelli Manzanarez Gómez

NOTAS

PROBABILIDAD
TEMA 2
VARIABLES ALEATORIAS

TEMA II
VARIABLES ALEATORIAS

INTRODUCCIÓN

El objetivo de este tema es conocer y utilizar el concepto de variable aleatoria para resolver problemas probabilísticos.

¿Qué es la variable aleatoria? En el tema anterior, se describieron los resultados de un experimento (eventos) en palabras; claramente esto dificultaba el análisis de algunos problemas. Es mucho más fácil describir y manipular cuando se utilizan números.

El propósito de la variable aleatoria es "mapear" cada evento de un espacio muestral en un punto de un eje real, esta idea puede resultar abstracta, es mucho más fácil pensar en la idea de que suele ser más fácil describir, manipular y analizar resultados cuando se utilizan números.

La "variable aleatoria " es la regla de correspondencia para realizar el mapeo de un conjunto (espacio muestral) a otro conjunto (números reales) por lo que la variable aleatoria es en sí una función.

LA VARIABLE ALEATORIA

Los fenómenos que más interesan a los ingenieros son aquellos que pueden ser identificados por números, los cuales reciben el nombre de eventos numéricos. Por ejemplo, un médico está interesado en el evento de que diez de diez pacientes se curen de cierta enfermedad.

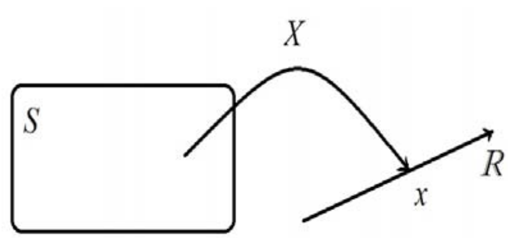


Fig. 2.1. Variable Aleatoria

Definición 2.1

Una variable aleatoria (v.a.) es una función, cuyos valores son números reales, definida en un espacio muestral. De una manera simple puede denotarse por $\boldsymbol{X} : \boldsymbol{S} \rightarrow \mathbb{R}$

En otras palabras una variable aleatoria es una función que asigna números reales a cada posible resultado de un experimento aleatorio; esto es, es una función cuyo dominio de definición es el espacio muestral de un experimento (o con mayor precisión, el conjunto potencia del espacio muestral) y cuyo rango toma del conjunto de los números reales.

Usualmente, se denotan a las variables aleatorias (vv.aa.) utilizando las últimas letras mayúsculas del alfabeto.

Los valores de la imagen de dicha función, se conocen como valores de la variable aleatoria (v.a.) y se denotan con la misma letra que la función, pero con minúsculas.

S))))))))))))))))))))))))))))))))))Q

Ejemplo 2.1

Considerando las familias que tiene dos hijos, si se desea conocer el sexo de los hijos entonces el espacio muestral es:

$S = \{ (F,F)\ ,\ (F,M)\ ,\ (M,F)\ ,\ (M,M) \}$

Utilizando F para el sexo femenino y M para el masculino.
Asignar valores a la variable aleatoria.

Resolución

Podrían realizarse varias asignaciones para cada uno de los elementos del espacio muestral, por ejemplo la variable aleatoria X, asigna al primer elemento el número 1, al segundo elemento el número 2 y así sucesivamente.

$X:\ 1,\ 2,\ 3,\ 4$

O de forma similar pero iniciando con cero para el primer elemento

$Y:\ 0,\ 1,\ 2,\ 3$

O contando el número de hombres en cada evento del espacio muestral

$W:\ 0,\ 1,\ 1,\ 2.$

etc.

S))))))))))))))))))))))))))))))))))Q

La definición (en su construcción y significado) de una variable aleatoria (v.a.) es arbitraria. Una forma adecuada de construir (definir) una v.a. es hacerlo de tal manera que dicha variable responda a la pregunta de interés.

En el ejemplo anterior, si nos interesa conocer el número de hijos varones en una familia, la definición más adecuada de la v.a. es la de W que asocia números reales de la siguiente manera:

$$(F,F) \to 0 \quad ; \quad (F,M),(M,F) \to 1 \quad ; \quad (M,M) \to 2$$

esto es, a cada punto del espacio muestral se le asocia el número de hijos varones

S))))))))))))))))))))))))))))))))))Q

Ejemplo 2.2

Considérese el lanzamiento de dos dados.

a) Definir el espacio muestral del experimento.

b) Definir una variable aleatoria para la suma de los dados.

c) Calcular la probabilidad de que la suma sea par.

Resolución

a) $$S = \left\{ \begin{array}{cccccc} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{array} \right\}$$

b) La definición es la siguiente:
Sea X la variable aleatoria que representa la suma de los resultados en el lanzamiento de dos dados.

Los posibles valores que toma la suma de los dados (posibles valores x de X) son:

2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

Esto implica que el rango[1] de la variable aleatoria X es:

$$R_X = \{ 2 , 3 ,4 , 5 , 6 , 7 , 8 , 9 , 10 , 11 , 12 \}$$

c) Sea A el evento en el cual la suma de los resultados es par, es evidente que:

$$A = \{ X=2 , X=4 , X=6 , X=8 , X=10 , X=12 \}$$

Por lo que

$$P(A) = P(X=2) + P(X=4) + P(X=6) + P(X=8) + + P(X=10) + P(X=12)$$
$$P(A) = \frac{1}{36} + \frac{3}{36} + \frac{5}{36} + \frac{5}{36} + \frac{3}{36} + \frac{1}{36}$$
$$P(A) = \frac{1}{2}$$

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.3

En una ciudad se observa el tiempo que transcurre de un sismo a otro, el cual se representa mediante la v.a. T. Obtener el rango de T.

[1]Rango en el sentido de rango de una función, como el estudiado en los cursos de Cálculo.

Resolución

Considerando que de un sismo a otro debe de transcurrir algún tiempo, y que no se sabe cuánto tardará en ocurrir el nuevo sismo, se tiene:

$R_T = \{ t \mid t > 0 \ ; \ t \in \mathbb{R} \}$

S))))))))))))))))))))))))))))))))))))))Q

Las variables aleatorias pueden clasificarse en discretas, continuas o mixtas. Se estudiarán las características de las discretas y de las continuas, dejando las mixtas como una combinación de los casos anteriores.

$$\textit{Variables aleatorias} \begin{cases} \textit{Discretas} \\ \textit{Continuas} \end{cases}$$

Una variable aleatoria discreta toma valores de un conjunto numerable, mientras que una variable aleatoria continua toma valores de un conjunto continuo.

VARIABLES ALEATORIAS DISCRETAS

Definición 2.2

Una variable aleatoria se dice discreta si solamente puede tomar valores de un conjunto numerable de valores.

Una vez definida una variable aleatoria discreta, se debe definir la forma en la que se asignarán las probabilidades a cada valor que puede tomar la variable aleatoria.

Definición 2.3

Sea X una v.a. discreta, se define su función de probabilidad [2] $f_X(x)$ como:

$$f_X(x) = P(X = x)$$

donde $f: R_X \rightarrow [0, 1]$

Para cualquier función de probabilidad de una v.a. discreta debe cumplirse lo siguiente:

[2] También llamada: función masa de probabilidad o distribución de probabilidad. El nombre de función masa proviene de las similitudes de la función de probabilidad con el estudio de las masas en la mecánica.
Es muy común la notación $p_X(x)$, donde se resalta el hecho de que la función proporciona probabilidad. En estas notas se utiliza la notación $f_X(x)$ para hacer énfasis en que es una función.

Definición 2.4 Propiedades de la función de Probabilidad

1) La función de probabilidad siempre está entre cero y uno, inclusive.
$0 \le f_X(x) \le 1 \quad , \quad \forall\ x$

2) La suma de las probabilidades de la función de probabilidad es la probabilidad del espacio muestral.
$$\sum_{\forall x} f_X(x) = 1$$

3) La probabilidad de que la variable aleatoria este en un intervalo, es la suma de los valores de la función de probabilidad para dicho intervalo
$$P(a \le X \le b) = \sum_{x=a}^{b} f_X(x)$$

Debe observarse la analogía de estas propiedades con los axiomas de la probabilidad. Para determinar si una función es una función de probabilidad, se deben cumplir las propiedades anteriores, en particular deben probarse (1) y (2).

S)))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.4

Considérese el lanzamiento de una moneda. Se desea observar el número de lanzamientos hasta que "caiga" por primera vez un sol. Obtener una expresión para la función de probabilidad y verificar que cumple con las primeras dos propiedades.

Resolución

Sea X la v.a. que representa el número de lanzamientos necesarios para observar por primera vez un sol.

El rango de la v.a. es $R_X = \{1, 2, 3, 4, \ldots\}$

$$f_X(1) = P(X=1) = \frac{1}{2}$$

$$f_X(2) = P(X=2) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

$$f_X(3) = P(X=3) = \frac{1}{8}$$

.
.
.

En general

$$f_X(x) = \begin{cases} \left(\dfrac{1}{2}\right)^x & ; \quad x = 1, 2, 3 \ldots \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

))

Verificando la propiedad (1)

$$0 \le \left(\frac{1}{2}\right)^{x} \le 1 \quad , \quad \forall\, x$$

Verificando (2)

Debe cumplirse que $\sum_{\forall x} f_X(x) = 1$

es decir $\sum_{x=1}^{\infty} \left(\frac{1}{2}\right)^{x} = 1$

De la serie geométrica se sabe que: $\sum_{n=0}^{\infty} r^{n}$ converge a $\frac{1}{1-r}$ cuando $|r| < 1$ y diverge cuando $|r| \ge 1$.

Relacionando la función de probabilidad con la serie geométrica se tiene que $r = \frac{1}{2}$.

Sólo falta restar el primer término cuando $k = 0$, de donde

$$\sum_{x=1}^{\infty} \left(\frac{1}{2}\right)^{x} = \left[\sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^{k}\right] - 1 = \left[\frac{1}{1-\frac{1}{2}}\right] - 1 = 1$$

Por lo que sí se satisface (2).

S))))))))))))))))))))))))))))))))))Q

Debe observarse con mucho cuidado el hecho de que una vez definida la v.a., debe poderse obtener el rango de dicha variable, es decir el conjunto de valores que la variable aleatoria puede tomar; sin embargo, el dominio de la función de probabilidad puede extenderse a todos los reales, para facilitar la notación en análisis posteriores.

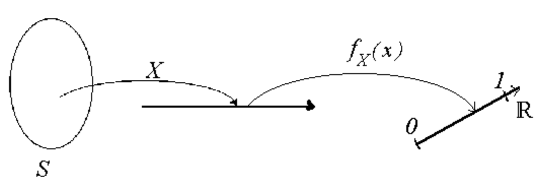


Fig. 2.2. Variable Aleatoria y Función de Probabilidad

S))))))))))))))))))))))))))))))))))Q

Ejemplo 2.5

Considérese la v.a. X cuyos posibles valores son 0, 1, 2, 3 y 4; y que tiene la siguiente función de probabilidad.

x	0	1	2	3	4
$f_X(x)$	$\frac{1}{6}$	$\frac{1+k}{6}$	$\frac{1+2k}{6}$	$\frac{1+3k}{6}$	$\frac{1+4k}{6}$

a) Determinar el valor de la constante k para que $f_X(x)$ sea una función de probabilidad.

b) Calcular $P(X < 3)$.

Resolución

a) Para que $f_X(x)$ sea una función de probabilidad, deben cumplirse las propiedades 1 y 2.

Si $k > 0$ entonces se cumple que $f_X(x) > 0$

Y por otro lado:

$$\sum_{x=0}^{4} f_X(x) = \frac{1}{6} + \frac{1+k}{6} + \frac{1+2k}{6} + \frac{1+3k}{6} + \frac{1+4k}{6} = 1$$

De donde $k = \frac{1}{10}$

b) $P(X < 3) = P(X=0) + P(X=1) + (X=2)$

$$= \frac{1}{6} + \frac{1+\frac{1}{10}}{6} + \frac{1+\frac{2}{10}}{6} = \frac{11}{20}$$

S))))))))))))))))))))))))))))))))))Q

Los problemas que generalmente resultan más interesantes son aquellos en los que se debe obtener la función de probabilidad para una variable aleatoria discreta a partir de un problema en particular.

S))))))))))))))))))))))))))))))))))Q

Ejemplo 2.6

Al examinar muestras de cierto mineral se encontró que podrían tener hasta dos tipos de impurezas. Estudios realizados en las muestras concluyeron que el 20% de las muestras no revelaban ninguna impureza, el 40% tenían la impureza del tipo A y el 50% la impureza del tipo B (naturalmente, algunas tenían ambas impurezas). Si se selecciona una muestra al azar, obtener la distribución de probabilidad para Y, esto es, el número de impurezas encontradas en la muestra.

Resolución

Sea A el evento en el cual se tiene la impureza A, y B el evento en el cual se tiene la impureza B, entonces, la siguiente tabla muestra las probabilidades de las intersecciones de los eventos:

	A	$\bar{A}$	Total
B	0.1	0.4	0.5
$\bar{B}$	0.3	0.2	0.5
Total	0.4	0.6	1

De donde

$P(Y = 0) = P(\bar{A} \cap \bar{B}) = 0.2$

$P(Y = 1) = P(\bar{A} \cap B) + P(A \cap \bar{B}) = 0.4 + 0.3 = 0.7$

$P(Y = 2) = P(A \cap B) = 0.1$

Por lo que la función de probabilidad es:

y	0	1	2
$f_Y(y)$	0.2	0.7	0.1

S))))))))))))))))))))))))))))))))))))Q

Una función de probabilidad generalmente se representa de manera gráfica utilizando líneas verticales que representan la probabilidad, como se muestra en la siguiente figura.

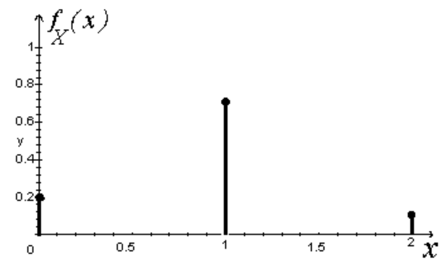


Fig. 2.3. Gráfica de una función de Probabilidad

Sin embargo, en este texto se preferirá la representación puntual, para hacer énfasis en que se trata de una función muy semejante a las estudiadas en los cursos de cálculo, pero que solamente toma valores puntuales. La representación que se utilizará es entonces:

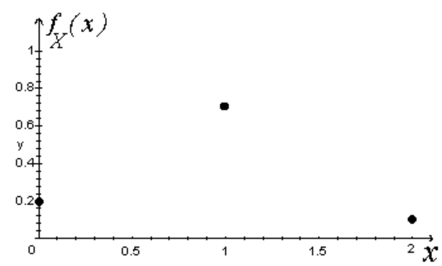
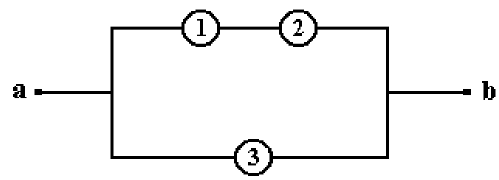


Fig. 2.4. Gráfica de una función de Probabilidad

))))))))))))))))))))))))))))))))))Q

Ejemplo 2.7

En la siguiente figura se muestra un sistema con tres interruptores. La probabilidad de que cada interruptor funcione en forma correcta es de 0.9 y operan de forma independiente. Si un interruptor trabaja en forma correcta, la corriente puede pasar por él. Determinar la distribución de probabilidad Y, el número de trayectorias cerradas (trayectorias que funcionan) de a a b cuando los tres interruptores se encienden.



Resolución

El rango de la v.a. es $R_Y = \{ 0 \ , \ 1 \ , \ 2 \}$

Sea S_i el evento en el cual el interruptor i funciona, para $i = 1 \ , \ 2 \ , \ 3$.

$$
\begin{aligned}
P(Y = 0) &= P\left(\bar{S}_1 S_2 \bar{S}_3 \cup S_1 \bar{S}_2 \bar{S}_3 \cup \bar{S}_1 \bar{S}_2 \bar{S}_3 \right) \\
&= (0.1)(0.9)(0.1) + (0.9)(0.1)(0.1) + \\
&\quad + (0.1)(0.1)(0.1) \\
&= 0.019
\end{aligned}
$$

$$
\begin{aligned}
P(Y = 1) &= P\left(S_1 S_2 \bar{S}_3 \cup S_1 \bar{S}_2 S_3 \cup \bar{S}_1 S_2 S_3 \cup \bar{S}_1 \bar{S}_2 S_3 \right) \\
&= 3(0.9)(0.9)(0.1) + (0.1)(0.1)(0.9) \\
&= 0.252
\end{aligned}
$$

$$
\begin{aligned}
P(Y = 2) &= P\left(S_1 S_2 S_3 \right) = (0.9)(0.9)(0.9) \\
&= 0.729
\end{aligned}
$$

Finalmente la distribución es:

y	0	1	2
$f_Y(y)$	0.019	0.252	0.729

S))))))))))))))))))))))))))))))))))Q

Un ejemplo más elaborado de la obtención de una función de probabilidad es el mostrado en el siguiente ejemplo.

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.8

De la gente que llega a un banco de sangre como donador, uno de cada tres tiene sangre tipo O^+, y uno de 15 sangre tipo O^-. Para las siguientes tres personas que llegan al banco, sea X el número que tiene sangre O^+ y Y las que tienen O^-.
Si se supone que las proporciones de los tipos de sangre se mantienen constantes (independencia), determinar las distribuciones de probabilidad para X y Y.

Resolución
Se tiene que:

$R_X = \{ 0 , 1 , 2 , 3 \}$, Probabilidad de que una persona tenga O^+ , $\frac{1}{3}$.

$R_Y = \{ 0 , 1 , 2 , 3 \}$, Probabilidad de que una persona tenga O^- , $\frac{1}{15}$

$$P(X = 0) = \left(\frac{2}{3} \right)^3$$

$$P(X = 1) = \binom{3}{1} \left(\frac{1}{3} \right) \left(\frac{2}{3} \right)^2$$

En general:

$$P(X = x) = \binom{3}{x} \left(\frac{1}{3} \right)^x \left(\frac{2}{3} \right)^{3-x} \quad , \quad x = 0 , 1 , 2 , 3 .$$

$$P(Y = y) = \binom{3}{y} \left(\frac{1}{15} \right)^y \left(\frac{14}{15} \right)^{3-y} \quad , \quad y = 0 , 1 , 2 , 3 .$$

En forma tabular, se tiene:

x	0	1	2	3
$f_X(x)$	$\frac{8}{27}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{1}{27}$

y	0	1	2	3
$f_Y(y)$	$\frac{2744}{3375}$	$\frac{588}{3375}$	$\frac{42}{3375}$	$\frac{1}{3375}$

S))))))))))))))))))))))))))))))))))))Q

VARIABLES ALEATORIAS CONTINUAS

Definición 2.5

Una v.a. X se dice continua si puede tomar valores de un conjunto infinito no numerable de valores.

Las variables aleatorias continuas también tienen una función que proporciona información sobre la probabilidad, dicha función se define a continuación.

Definición 2.6 Propiedades de la Función de Densidad

Sea X una variable aleatoria continua, se define su función de densidad $f_X(x)$, como una función con las siguientes propiedades:

1) La función de densidad no es negativa.
$f_X(x) \geq 0 \quad \forall\, x$

2) La probabilidad del espacio muestral es la unidad.
$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$

3) La probabilidad de que la variable aleatoria tome valores en un intervalo, es la integral de la función de densidad en dicho intervalo.
$P(a < X < b) = \int_a^b f_X(x)\, dx$

Recordando las propiedades de la integral definida, se tienen los siguientes resultados:

a) $P(X = x) = 0$

b) $P(a \leq X \leq b) = P(a < X < b)$

En términos simples, las propiedades de la función de densidad se basan en los axiomas de la probabilidad y expresan las mismas ideas: la función de densidad no debe ser negativa (1); y que la probabilidad del espacio muestral debe ser igual a la unidad (2). De los resultados debe destacarse el hecho de que la probabilidad de que una variable aleatoria continua tome exactamente un valor es cero (a), lo cual debe interpretarse como: la probabilidad de que una variable aleatoria continua tome exactamente un valor es *prácticamente* cero.

Una función de densidad, se representa gráficamente de igual forma que una función continua en el cálculo, como puede observarse en la figura 5.

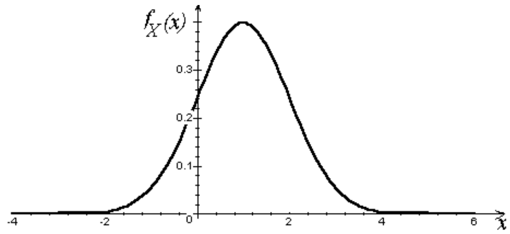


Fig. 2.5. Gráfica de una función de densidad

Debe observarse de la propiedad (3), que para obtener la probabilidad de que una variable aleatoria esté dentro de cierto intervalo, se integra sobre ese intervalo; recordando la interpretación geométrica de la integral, se puede decir que la probabilidad coincide con el "área" bajo la curva $f_X(x)$.

El siguiente ejemplo muestra el uso de la propiedad 2 de las funciones de densidad, la propiedad 3 para el cálculo de probabilidades y adicionalmente, ilustra la forma en la que se pueden calcular probabilidades condicionales, utilizando una variación de la definición de probabilidad condicional vista en el capítulo anterior,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.9

El tiempo requerido por los estudiantes para presentar un examen de una hora es una variable aleatoria con una función de densidad dada por

$$f_Y(y) = \begin{cases} cy^2 + y & ; \quad 0 \le y \le 1 \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

a) Determinar el valor de c.

b) Trazar la gráfica de $f_Y(y)$.

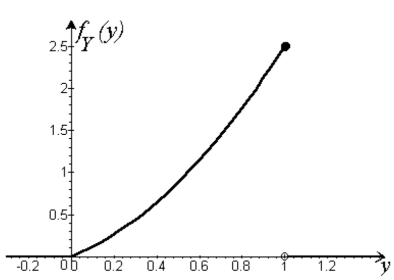
c) Calcular la probabilidad de que un estudiante termine en menos de media hora.

d) Dado que cierto estudiante necesita al menos 15 minutos para presentar el examen, obtener la probabilidad de que necesite al menos 30 minutos para terminarlo.

Resolución

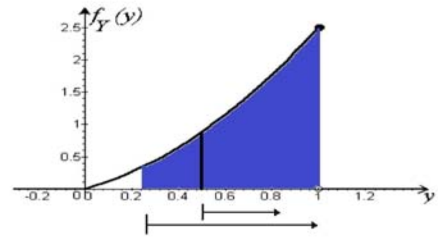
a) $\int_0^1 (cy^2 + y)\, dy = 1$, de donde $\frac{c}{3} + \frac{1}{2} = 1 \quad \Rightarrow c = \frac{3}{2}$

b)



c) $P(0 \le Y \le 0.5) = \int_0^{\frac{1}{2}} \left(\frac{3}{2} y^2 + y \right) dy = \frac{3}{16}$

d) $$P\left(Y \geq \frac{1}{2} \,\middle|\, Y \geq \frac{1}{4} \right) = \frac{P\left(Y \geq \frac{1}{2} \right)}{P\left(Y \geq \frac{1}{4} \right)}$$



$$= \frac{\int_{\frac{1}{2}}^{1} \left(\frac{3}{2} y^2 + y \right) dy}{\int_{\frac{1}{4}}^{1} \left(\frac{3}{2} y^2 + y \right) dy} = \frac{104}{123} = 0.8455$$

S)))))))))))))))))))))))))))))))))))Q

Ejemplo 2.10

La longitud X de una pieza mecánica se ajusta a 60 milímetros, pero en realidad es una variable aleatoria con la siguiente función de densidad.

$$f_X(x) = \begin{cases} \frac{1}{2} & ; \quad 59 \leq x \leq 61 \\ 0 & : \quad en \ otro \ caso \end{cases}$$

a) Calcular la probabilidad de que la longitud sea mayor que 60 [mm]

b) Si se seleccionan aleatoriamente dos piezas, en forma independiente, calcular la probabilidad de que ambas sean de más de 60 [mm].

Resolución

a) $$P(X > 60) = \int_{60}^{61} \frac{1}{2} \, dx = \frac{1}{2}$$

b) Sea X_i la variable aleatoria que representa la longitud de la pieza i

$$f_{X_i}(x_i) = \begin{cases} \frac{1}{2} & 59 \leq x_i \leq 61 \\ 0 & en \ otro \ caso \end{cases} \qquad \text{para} \qquad i = 1, 2$$

por lo que:

$$P(X_1 > 60 \cap X_2 > 60) = P(X_1 > 60)\, P(X_2 > 60)$$

$$= \left(\frac{1}{2} \right) \left(\frac{1}{2} \right)$$

Recuérdese que la selección se realiza de manera independiente

$$= \frac{1}{4}$$

S)))))))))))))))))))))))))))))))))))Q

FUNCIÓN DE DISTRIBUCIÓN

La función de distribución, también llamada función de distribución acumulativa, muestra el comportamiento acumulado de una variable aleatoria. Cuando se obtiene, ayuda para calcular probabilidades sin necesidad de volver a realizar sumas o integrales.

Definición 2.7

Si X es una variable aleatoria, entonces su función de distribución $F_X(x)$, se define como una función que asocia a cada valor real, la probabilidad de que la variable aleatoria asuma valores menores o iguales que él.

$$F_X(x) = P(X \leq x)$$

$F_X : \mathbb{R} \rightarrow [0, 1]$

La función de distribución se obtiene de la siguiente manera:

$$F_X(x) = \begin{cases} \displaystyle\sum_{i=-\infty}^{x} f_X(i) & ; \ X \textit{ discreta} \\ \displaystyle\int_{-\infty}^{x} f_X(t)\, dt & ; \ X \textit{ continua} \end{cases}$$

Sin importar si la variable aleatoria es discreta o continua, la función de distribución tiene las siguientes propiedades, que se deducen directamente de la definición.

Propiedades de la Función de Distribución

1) $0 \leq F_X(x) \leq 1 \quad , \quad -\infty < x < \infty$

2) Para el mayor valor en el rango de la variable aleatoria X, $F_X(x) = 1$; es decir:

$$\lim_{x \to \infty} F_X(x) = 1$$

Para un valor menor al primer valor en el rango de la variable aleatoria X, $F_X(x) = 0$; es decir:

$$\lim_{x \to -\infty} F_X(x) = 0$$

3) La función $F_X(x)$ es no decreciente; es decir:

$$\text{Si } a \leq b \text{ entonces } F_X(a) \leq F_X(b)$$

4) La probabilidad de que una variable aleatoria esté en el intervalo $(a, b]$ está dada por:

$$P(a < X \leq b) = F_X(b) - F_X(a)$$

Para generalizar la propiedad (4) al intervalo cerrado $[a, b]$, deben utilizarse los casos discreto y continuo.

$$P(a \le X \le b) = \begin{cases} F_X(b) - F_X(a) + f_X(a) & ; \; X \text{ discreta} \\ F_X(b) - F_X(a) & ; \; X \text{ continua} \end{cases}$$

La representación gráfica de la función de distribución para una v.a discreta es una función escalonada con discontinuidades de salto, mientras que para una variable aleatoria continua es una función continua no decreciente.

Puesto que para una variable aleatoria continua, la probabilidad de que la variable esté dentro de un intervalo coincide con el área bajo la curva para ese intervalo, la función de distribución permite el cálculo de la probabilidad simplemente como una resta de áreas (propiedad 4), como se ilustra en la siguiente figura.

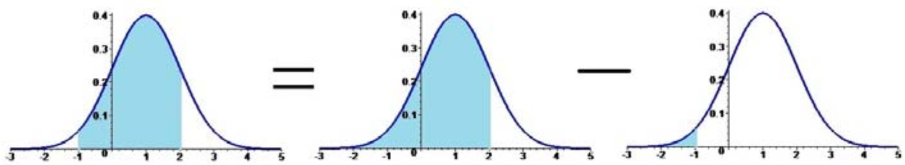


Fig. 2.6. Uso de la función de distribución para el cálculo de probabilidades. Variable aleatoria continua

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.11

Sea X una variable aleatoria con función de probabilidad $f_X(x)$

x	-5	-1	1	1.5	3
$f_X(x)$	0.2	0.01	0.3	0.29	0.2

a) Construir la función de distribución de X en forma tabular.

b) Trazar su gráfica.

c) $P(X < 1)$

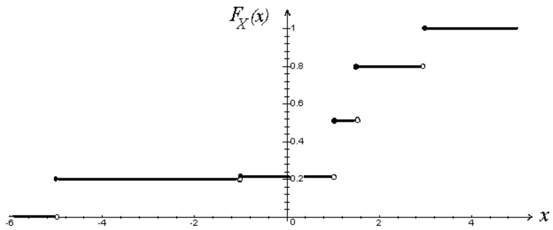
d) $P(-1 < X \le 1)$

Resolución

a) La forma tabular de la función de distribución se obtiene directamente de la función de probabilidad, sumando las casillas a la izquierda y la del valor que se desea.

x	-5	-1	1	1.5	3
$F_X(x)$	0.2	0.21	0.51	0.8	1

b)



c) $P(X < 1) = F_X(-1) = 0.21$

d) $P(-1 < X \leq 1) = F_X(1) - F_X(-1) = 0.3$

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.12
El tiempo requerido por los estudiantes para presentar un examen de una hora es una variable aleatoria con una función de densidad dada por

$$f_Y(y) = \begin{cases} \frac{3}{2}y^2 + y & ; \quad 0 \leq y \leq 1 \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

a) Obtener $F_Y(y)$.

b) Trazar la gráfica $F_Y(y)$.

c) Utilizar $F_Y(y)$ del inciso (a) para encontrar $F_Y(-1)$, $F_Y(0)$ y $F_Y(1)$.

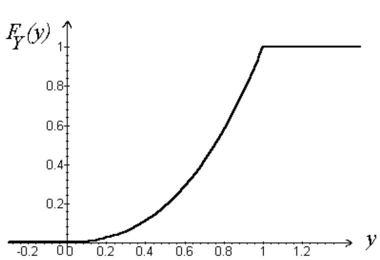
Resolución

a) $$F_Y(y) = \int_0^y \left(\frac{3}{2}t^2 + t \right) dt = \frac{y^3}{2} + \frac{y^2}{2}, \; 0 \leq y \leq 1$$

Finalmente

$$F_Y(y) = \begin{cases} 0 & ; \quad y < 0 \\ \frac{y^3}{2} + \frac{y^2}{2} & ; \quad 0 \leq y \leq 1 \\ 1 & ; \quad y > 1 \end{cases}$$

b)



))

c) $F_Y(-1) = 0$

$F_Y(0) = 0$

$F_Y(1) = \frac{1}{2} + \frac{1}{2} = 1$

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.13

Una gasolinería tiene dos bombas que pueden bombear cada una hasta 10 000 litros de gasolina por mes. La cantidad total de gasolina bombeada en un mes es una variable aleatoria Y (expresada en diez mil litros), con una función de densidad de probabilidad dada por

$$f_Y(y) = \begin{cases} y & ; \quad 0 < y < 1 \\ 2 - y & ; \quad 1 \le y < 2 \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

a) Trazar la gráfica de $f_Y(y)$.

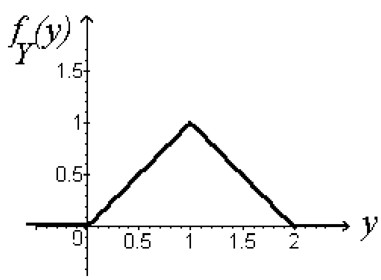
b) Obtener $F_Y(y)$ y construir su gráfica.

c) Calcular la probabilidad de que la gasolinera bombee entre 8 000 y 12 000 litros en un mes.

d) Si se sabe que la gasolinera ha bombeado más de 10 000 litros en un mes en particular, obtener la probabilidad de que haya bombeado más de 15 000 litros durante el mes.

Resolución

a)



b) Para $0 < y < 1$, $F_Y(y) = \int_0^y t\, dt = \frac{y^2}{2}$

Para $1 \le y < 2$

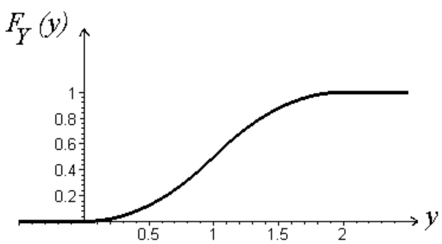
En este intervalo debe considerarse lo que se acumula de $y = 1$ en adelante, así como todo lo acumulado en el intervalo anterior.

$$F_Y(y) = \int_0^1 t\, dt + \int_1^y (2 - t)\, dt = 2y - \frac{y^2}{2} - 1$$

Finalmente

$$F_Y(y) = \begin{cases} 0 & ; \quad y \le 0 \\ \dfrac{y^2}{2} & ; \quad 0 < y < 1 \\ 2y - \dfrac{y^2}{2} - 1 & ; \quad 1 \le y < 2 \\ 1 & ; \quad y \ge 2 \end{cases}$$

Obsérvese que la función de distribución es no decreciente, por lo que inicia teniendo el valor cero y termina teniendo el valor de uno.



c) $P(\, 0.8 \le Y \le 1.2 \,) = F_Y(1.2) - F_Y(0.8)$

$$= (\, 2.4 - 0.72 - 1 \,) - 0.32 = 0.36$$

d) $P(\, Y > 1.5 \mid Y > 1 \,) = \dfrac{P(\, Y > 1.5 \,)}{P(\, Y > 1 \,)}$

$$= \frac{1 - (\, 3 - 1.125 - 1 \,)}{0.5} = 0.25$$

S))))))))))))))))))))))))))))))))))))Q

VALOR ESPERADO DE UNA VARIABLE ALEATORIA

El concepto de valor esperado es sin duda uno de los más importantes en el estudio de las distribuciones de probabilidad. Tiene sus orígenes en los juegos de azar, debido a que los apostadores querían saber cuál era su expectativa en un juego después de participar en él muchas veces.

Por ejemplo, considérese una pequeña rifa que depende del resultado de un dado. El boleto para participar en la rifa cuesta un peso, y el apostador recibe cinco pesos si el resultado es 6 (tiene una ganancia de 4 pesos), y en caso contrario pierde el valor del boleto con el que participó en la rifa. Después de un gran número de juegos, ¿cual será su pérdida o ganancia? Para contestar se debe considerar a la distribución de probabilidad como la frecuencia relativa a largo plazo de los resultados que se obtendrán, de manera que, si X es la variable aleatoria que representa el resultado del dado, la probabilidad de ganar es $P(\,X = 6\,) = \dfrac{1}{6}$, mientras que la probabilidad de perder es $P(\,X \ne 6\,) = \dfrac{5}{6}$, de donde la ganancia esperada promedio por rifa es:

$$4\left(\frac{1}{6}\right) + (-1)\left(\frac{5}{6}\right) = -\frac{1}{6}$$

Es decir, después de jugar un gran número de veces, el apostador pierde un sexto de peso en promedio por rifa.

Debe observarse que la operación que se realizó es un promedio ponderado de la ganancia de cada rifa. De hecho de esta forma se define el valor esperado de una variable aleatoria discreta. Para el caso de una variable aleatoria continua la definición es similar.

Definición 2.8

Sea X una variable aleatoria con distribución de probabilidad $f_X(x)$. El valor esperado de X, $\mathrm{E}(X)$ es:

$$\mathrm{E}(X) = \begin{cases} \sum_{\forall x} x\, f_X(x) & ;\ X \textit{ discreta} \\ \int_{-\infty}^{\infty} x\, f_X(x)\, dx & ;\ X \textit{ continua} \end{cases}$$

El valor esperado es muy utilizado en la teoría de decisiones. Es decir, en situaciones de incertidumbre se toma una decisión basada en la expectativa de repetir un gran número de veces una situación similar y determinar el valor esperado como resultado de la repetición.

Es muy común denotar al valor esperado de una variable aleatoria X mediante μ_X, y en ocasiones, simplemente como μ.

S))))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.14

Un automovilista desea asegurar su automóvil por 50 000 U.M.[3]. La compañía aseguradora estima que una pérdida total puede ocurrir con una probabilidad de 0.002, un 50% de pérdida con una probabilidad de 0.01 y un 25% de pérdida con una probabilidad de 0.1. Ignorando todas las otras pérdidas parciales, ¿qué prima deberá cobrar anualmente la compañía aseguradora para tener una utilidad promedio por automóvil de 500 U.M.?

Resolución

Sea c la prima y X la variable aleatoria que representa la utilidad, entonces

x	$c - 50000$	$c - 25000$	$c - 12500$	c
$f_X(x)$	0.002	0.01	0.1	0.888

Para que la utilidad promedio sea de 500, se debe satisfacer $\mathrm{E}(X) = 500$, de donde

$$\mathrm{E}(X) = (c - 50000)(0.002) + (c - 25000)(0.01) + \\ + (c - 12500)(0.1) + (c)(0.888)$$

[3] Unidades Monetarias

$\mathrm{E}(X) = 500$

Por lo que

$c = 2100$ U.M.

La prima que deberá cobrar es de 2100 U.M.

S))))))))))))))))))))))))))))))))))Q

Ejemplo 2.15

La función de densidad de la variable aleatoria continua X, el número total de horas, en unidades de 100 horas, que una familia utiliza una aspiradora durante un año está dado por

$$f_X(x) = \begin{cases} x & ; \quad 0 < x < 1 \\ 2 - x & ; \quad 1 \leq x < 2 \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

Obtener el número promedio de horas por año que la familia utiliza la aspiradora.

Resolución

$$\mathrm{E}(X) = \int_0^1 x^2\, dx + \int_1^2 x(2 - x)\, dx = 1$$

La aspiradora se usa en promedio 100 horas al año.

S))))))))))))))))))))))))))))))))))Q

PROPIEDADES DEL VALOR ESPERADO

1) El valor esperado de una constante c, es la misma constante.

$\mathrm{E}(c) = c$

2) El valor esperado de una variable aleatoria por una constante, es la constante por el valor esperado de la variable aleatoria

$\mathrm{E}(cX) = c\,\mathrm{E}(X)$

3) El valor esperado de la cantidad $aX + b$ donde a y b son constantes, es el producto de a por el valor esperado de X más b.

$\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b$

4) Si $g(x)$ es una función de X, entonces:

$$\mathrm{E}\big(g(X)\big) = \begin{cases} \sum\limits_{\forall x} g(x)\, f_X(x) & ; \quad X \textit{ discreta} \\ \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx & ; \quad X \textit{ continua} \end{cases}$$

5) El valor esperado de una suma de funciones es igual a la suma de los valores esperados.

Si $g_1(X)$ y $g_2(X)$ son funciones de X, entonces

$$E\left(g_1(X) + g_2(X)\right) = E\left(g_1(X)\right) + E\left(g_2(X)\right)$$

Las demostraciones son inmediatas a partir de la definición del valor esperado y el uso de las propiedades de la suma y de la integral.

El valor esperado de una variable aleatoria discreta puede ser interpretado como el centro de masa de una distribución de masas $f_X(x_i)$ colocadas en los puntos x_i del eje x.

El valor esperado de una variable aleatoria continua puede ser interpretado como la abscisa del centroide de la figura formada por la función $f_X(x)$ junto con el eje x (y los extremos de $f_X(x)$ si los hubiera).

MOMENTOS DE UNA VARIABLE ALEATORIA

Sea X una variable aleatoria, se define su r-ésimo momento con respecto al origen como:

$$\mu_r^{/} = E(X^r)$$

y el r-ésimo momento con respecto a la media[4] como

$$\mu_r = E\left[(X - \mu_X)^r\right]$$

El primer momento con respecto al origen es el valor esperado

$$\mu_1^{/} = E(X) = \mu_X$$

Por las características del operador valor esperado, es posible obtener momentos con respecto a la media a partir de momentos con respecto al origen, como en el siguiente ejemplo

$$\mu_2 = E\left[(X - \mu_X)^2\right] = E(X^2 - 2X\mu_X + \mu_X^2)$$

$$\mu_2 = E(X^2) - 2\mu_X E(X) + \mu_X^2 = E(X^2) - \mu_X^2$$

Es decir, el segundo momento con respecto a la media se puede obtener a través de dos momentos con respecto al origen, y de similar forma el tercer momento con respecto a la media se puede obtener con tres momentos con respecto al origen, y así sucesivamente.

FUNCIÓN GENERADORA DE MOMENTOS

Debido a que los momentos permitirán calcular algunas de las características numéricas de las variables aleatorias, es conveniente estudiar una forma alterna para obtener los momentos. La función generadora de momentos permite calcular con facilidad algunos momentos, que utilizando la definición resultan mucho más complicados.

---

[4] o valor esperado.

Definición 2.9

Sea X una variable aleatoria se define su función generadora de momentos como

$$M_X(\theta) = \mathrm{E}\left(e^{\theta X}\right)$$

y si el valor esperado existe, genera todos los momentos con respecto al origen.

Al calcular el valor esperado $\mathrm{E}\left(e^{\theta X}\right)$, θ se considera como constante, de manera que $M_X(\theta)$ es finalmente sólo función de θ y no contiene X. La variable θ de la función generadora es una variable muda, puesto que para calcular el momento debe valuarse la variable. Es muy común que se denote a la función generadora como $m_X(t)$; sin embargo, aquí se preferirá la mayúscula M, debido a que la minúscula se utiliza en el curso de estadística para definir los momentos muestrales, y se prefiere θ sobre t debido a que t se utiliza como variable aleatoria en una función de densidad muy importante que se estudiará en detalle en el tema 6.

La razón por la cual la función $M_X(\theta)$ genera los momentos con respecto al origen y la forma en que lo hace puede observarse a continuación

La serie de Maclaurin es:

$$f(x) = \frac{f(0)\ x^0}{0!} + \frac{f'(0)\ x}{1!} + \frac{f''(0)\ x^2}{2!} + \ . \ . \ .$$

Expresando $e^{\theta X}$ en serie de Maclaurin se tiene:

Si

$$\begin{aligned} f(X) &= e^{\theta X} & \Rightarrow \quad f(0) &= 1 \\ f'(X) &= \theta\ e^{\theta X} & \Rightarrow \quad f'(0) &= \theta \\ f''(X) &= \theta^2\ e^{\theta X} & \Rightarrow \quad f''(0) &= \theta^2 \\ f'''(X) &= \theta^3\ e^{\theta X} & \Rightarrow \quad f'''(0) &= \theta^3 \end{aligned}$$

sustituyendo

$$e^{\theta X} = 1 + \theta\ X + \frac{\theta^2 X^2}{2!} + \frac{\theta^3 X^3}{3!} + \ . \ . \ .$$

por lo que sustituyendo en la definición de la función generadora de momentos

$$M_X(\theta) = \mathrm{E}\left(e^{\theta X}\right) = \mathrm{E}\left(1 + \theta X + \frac{\theta^2 X^2}{2!} + \frac{\theta^3 X^3}{3!} + \ . \ . \ .\right)$$

de donde

$$M_X(\theta) = \mathrm{E}(1) + \mathrm{E}\left(\theta X\right) + \mathrm{E}\left(\frac{\theta^2 X^2}{2!}\right) + \mathrm{E}\left(\frac{\theta^3 X^3}{3!}\right) + \ . \ . \ .$$

$$M_X(\theta) = 1 + \theta\ \mathrm{E}\left(X\right) + \frac{\theta^2}{2}\ \mathrm{E}\left(X^2\right) + \frac{\theta^3}{3!}\ \mathrm{E}\left(X^3\right) + \ . \ . \ .$$

Para obtener el n-ésimo momento con respecto al origen se deriva n veces con respecto de θ y se valúa en $\theta = 0$.

$$\boxed{\left.\frac{d^n M_X(\theta)}{d\theta^n}\right|_{\theta=0} = \mathrm{E}\left(X^n\right) = \mu_n'}$$

y para ello, la función $M_X(\theta)$ debe ser n veces derivable en cero.

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.16

Considérese la variable aleatoria X con función de probabilidad

x	-4	-2	0	1	10	20
$f_X(x)$	0.01	0.3	0.25	0.04	0.24	0.16

Construir la función generadora de momentos y utilizarla para calcular la media de la variable aleatoria.

Resolución

$M_X(\theta) = \mathrm{E}\left(e^{\theta X}\right)$

$$\mathrm{E}\left(e^{\theta X}\right) = \sum_{\forall x} e^{\theta x} f_X(x) = e^{-4\theta}(0.01) + e^{-2\theta}(0.3) + e^{0(\theta)}(0.25) + $$
$$+ e^{\theta}(0.04) + e^{10\theta}(0.24) + e^{20\theta}(0.16)$$

Por lo que

$$M_X'(\theta) = -4\, e^{-4\theta}(0.01) - 2\, e^{-2\theta}(0.3) + e^{\theta}(0.04)$$
$$+ 10\, e^{10\theta}(0.24) + 20\, e^{20\theta}(0.16)$$

Valuando en cero: $\left. M_X'(\theta)\right|_{\theta=0} = 5$

Por lo que $\mathrm{E}(X) = \mu_X = 5$

S))))))))))))))))))))))))))))))))))))Q

Para una variable aleatoria discreta, cuya función de probabilidad está definida por una tabla, resulta más laborioso utilizar la función generadora para obtener su media o valor esperado, que haber utilizado simplemente la definición. En general, la función generadora facilita el análisis cuando la obtención por medio de la definición es laboriosa, y lo dificulta cuando la obtención del valor esperado es muy fácil por medio de la definición.

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.17

La función de densidad de probabilidad de una variable aleatoria está determinada por:

$$f_X(x) = \begin{cases} \dfrac{x}{16} e^{-\frac{x}{4}} & ; \quad x > 0 \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

a) Determinar la función generadora de momentos de X.

b) Utilizar la función generadora de momentos para obtener la media de X.

Resolución

a) La función generadora de momentos está definida por

$$M_X(\theta) = \mathrm{E}(e^{\theta X})$$

por lo que:

$$M_X(\theta) = \int_0^\infty \frac{x}{16} e^{\left(\theta x - \frac{x}{4}\right)} dx$$

$$M_X(\theta) = \int_0^\infty \frac{x}{16} e^{-\left(\frac{1}{4} - \theta\right)x} dx$$

integrando por partes se obtiene

$$M_X(\theta) = \lim_{b \to \infty} \left[\frac{x e^{\left(\theta - \frac{1}{4}\right)x}}{16\left(\theta - \frac{1}{4}\right)} \right]_0^b - \int_0^\infty \frac{e^{\left(\theta - \frac{1}{4}\right)x}}{16\left(\theta - \frac{1}{4}\right)} dx$$

$$M_X(\theta) = \lim_{b \to \infty} \left[\frac{x e^{\left(\theta - \frac{1}{4}\right)x}}{16\left(\theta - \frac{1}{4}\right)} - \frac{e^{\left(\theta - \frac{1}{4}\right)x}}{16\left(\theta - \frac{1}{4}\right)^2} \right]_0^b$$

Y recordando que la función generadora debe ser continua y n veces derivable en el origen, entonces

$$\lim_{x \to \infty} e^{\left(\theta - \frac{1}{4}\right)x} = \lim_{x \to \infty} e^{-\left(\frac{1}{4} - \theta\right)x} = 0$$

de donde

$$M_X(\theta) = \frac{1}{(4\theta - 1)^2}$$

b) Para obtener la media, o valor esperado, se sabe que:

$$\mu_X = \frac{d}{d\theta} M_X(\theta) \Big|_{\theta = 0}$$

por lo que

$$\frac{d}{d\theta}\, M_X(\theta) = \frac{-8}{(4\theta - 1)^3} \,, \qquad \frac{d}{d\theta}\, M_X(\theta)\bigg|_{\theta = 0} = 8$$

$$\mu_X = 8$$

S))))))))))))))))))))))))))))))))))Q

CARACTERÍSTICAS NUMÉRICAS DE LA VARIABLE ALEATORIA

Si bien, el comportamiento probabilístico de las variables aleatorias queda completamente especificado mediante las funciones de probabilidad o de densidad, según sea el caso, en ocasiones es conveniente trabajar con algunas características numéricas o parámetros que describen el comportamiento de la variable aleatoria. Un parámetro es una característica numérica que describe de forma parcial o completa a una distribución de probabilidad.

Las características numéricas se clasifican en tres:

- Medidas de tendencia central.
- Medidas de dispersión.
- Parámetros de forma.

MEDIDAS DE TENDENCIA CENTRAL

Las medidas de tendencia central son la media, la moda y la mediana.

Media

La media o valor esperado, que también recibe el nombre de esperanza matemática, se estudió antes. En muchas aplicaciones se considera como el valor más representativo de una variable aleatoria. Se denota por μ.

Mediana

Es aquel valor para el cual la probabilidad de que la variable aleatoria tome valores menores o iguales a dicho valor es 0.5. Se denota por $\tilde{\mu}$ o por $\tilde{x}$.

Matemáticamente, la mediana es el valor $\tilde{x}$ tal que $P(X \leq \tilde{x}) = 0.5$.

Moda

Es aquel valor para el cual la función de probabilidad o función de densidad, según sea el caso, toma su valor máximo. Se denota por x_{mo}.

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.18

Considérese la variable aleatoria X con función de probabilidad

x	-4	-2	0	1	5	10
$f_X(x)$	0.01	0.3	0.25	0.04	0.24	0.16

Obtener la moda de la variable aleatoria.

Resolución

Al observar las probabilidades de la tabla se observa que:

$$x_{mo} = -2$$

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.19

Sea la función de densidad

$$f_X(x) = \begin{cases} 2\,e^{-2x} & ; \quad x > 0 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

Obtener la media, la mediana y la moda.

Resolución

La media está dada por $\mu_X = \int_0^{\infty} x\,(2\,e^{-2x})\,dx$

Integrando por partes se tiene que: $\mu_X = \dfrac{1}{2}$

La mediana está dada por $P(X \le \tilde{x}) = 0.5$

de donde $\int_0^{\tilde{x}} 2\,e^{-2x}\,dx = 0.5$

$$\left[-e^{-2x}\right]_0^{\tilde{x}} = 0.5$$

$$1 - e^{-2\tilde{x}} = 0.5$$

$$\tilde{x} = -\frac{\ln 0.5}{2} \approx 0.34657$$

La moda es el valor en el cual se encuentra el máximo, por lo que derivando e igualando a cero para obtener el máximo:

$$f_X'(x) = -4\,e^{-2x} = 0$$

No existe solución.

Por otro lado, no se puede considerar ningún extremo del intervalo, puesto que en ambos casos es abierto $(0, \infty)$.

Se concluye que la función no tiene moda.
S))))))))))))))))))))))))))))))))))Q

Como se pudo observar en el ejemplo anterior, no es necesario que existan todas las medidas de tendencia central, de hecho la media y la moda pueden no existir. La media no existe cuando la integral para las vv.aa. continuas no converge; la moda no existe en casos como el anterior, o bien, cuando para cualquier valor de la variable la función toma el mismo valor[5].

MEDIDAS DE DISPERSIÓN

Las medidas de dispersión indican la lejanía de los valores que puede tomar la variable aleatoria. Las principales medidas de dispersión son el rango, la desviación media, la variancia, la desviación estándar y el coeficiente de variación.

Rango

Es la medida de dispersión más simple. El rango se define como la diferencia entre el mayor valor que puede asumir la variable y el menor valor que puede asumir.

$$Rango = Mayor\ valor - menor\ valor$$

Desviación media

La desviación media de una variable aleatoria es el valor esperado de la diferencia en valor absoluto entre los valores de X y su media. Se denota DM_X

Matemáticamente:

$$DM_X = \mathrm{E}\left(|X - \mu_X| \right) = \begin{cases} \displaystyle\sum_{\forall x} |x - \mu_X| \ f_X(x) & ; \quad X \ discreta \\ \displaystyle\int_{-\infty}^{\infty} |x - \mu_X| \ f_X(x)\ dx & ; \quad X \ continua \end{cases}$$

En otras palabras, puede decirse que la desviación media es el promedio de los valores absolutos de las dispersiones alrededor de la media.

Debido a la dificultad matemática que implica trabajar con el valor absoluto, recuérdese que el valor absoluto está definido mediante dos reglas de correspondencia, se utiliza una función cuadrática para eliminar el problema que se genera por las diferencias positivas y negativas, dando lugar a la variancia.

[5] Algunos autores dicen que en este caso todos los valores son modas.

Varianza

La varianza o variancia se denota por $Var(X)$ o bien por σ_X^2 y se define como el promedio del cuadrado de la diferencia de la variable aleatoria y su media.

Matemáticamente:

$$\operatorname{Var}(X) = \begin{cases} \displaystyle\sum_{\forall x} (x-\mu_X)^2 \, f_X(x) & ; \quad X \textit{ discreta} \\ \displaystyle\int_{-\infty}^{\infty} (x-\mu_X)^2 \, f_X(x)\, dx & ; \quad X \textit{ continua} \end{cases}$$

De la definición de la variancia, y utilizando las propiedades del valor esperado, se puede escribir la variancia como:

$$\operatorname{Var}(X) = \operatorname{E}\left[(X-\mu_X)^2\right] = \operatorname{E}(X^2) - \mu_X^2$$

Debe recordarse que en este momento se está considerando que $f_X(x)$ es una caracterización exacta de la distribución de frecuencias de una población, es decir, $f_X(x)$ proporciona toda la información de la población, por lo que se puede escribir:

$$\operatorname{E}(X) = \mu_X$$

$$\operatorname{Var}(X) = \sigma_X^2$$

en el curso de Estadística se estudiarán los casos en los cuales no se cuenta con toda la información de la población.

La variancia es el segundo momento con respecto a la media.

CARACTERÍSTICAS DE LA VARIANCIA

1) La variancia de una constante es cero.
   $\operatorname{Var}(c) = 0$
2) La variancia de una constante por la variable aleatoria es el cuadrado de la constante por la variancia de la variable aleatoria
   $\operatorname{Var}(cX) = c^2 \operatorname{Var}(X)$

Debido a que la variancia queda en unidades cuadradas, se obtiene la raíz cuadrada para regresar a las unidades originales, lo que da lugar a la desviación estándar.

Desviación estándar

La desviación estándar se define como la raíz cuadrada positiva de la variancia. Se denota por σ_X. Esto es:

$$\sigma_X = \sqrt{\mathrm{Var}(X)} = \sqrt{\sigma_X^2}$$

Coeficiente de Variación

El coeficiente de variación mide la dispersión relativa a la media de una distribución de probabilidad, y se utiliza para comparar la dispersión de dos distribuciones de probabilidad.
Se define como la desviación estándar entre la media. Se denota CV_X.

$$CV_X = \frac{\sigma_X}{\mu_X}$$

El coeficiente de variación es sensible a la posición de los datos, es decir, a la media de los datos, y para que se pueda interpretar de forma adecuada, todos los datos deben ser positivos de tal forma que la media también sea positiva. Es muy común escribir el coeficiente de variación como porcentaje, cuando se hace esto, un coeficiente de variación menor al 15% indica datos homogéneos, mientras que un coeficiente de variación mayor o igual al 15% indica datos heterogéneos.

PARÁMETROS DE FORMA

Los parámetros de forma son el coeficiente de sesgo y el coeficiente de curtosis.

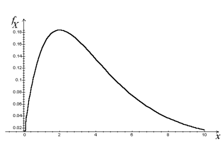
Coeficiente de sesgo

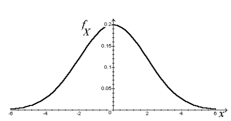
El coeficiente de sesgo mide el grado de simetría de una distribución de probabilidad. Se define como el tercer momento con respecto a la media estandarizado. Se denota por $\alpha_3(X)$.

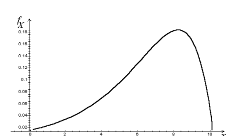
$$\alpha_3(X) = \frac{\mu_3}{\sigma_X^3}$$

El valor del coeficiente de sesgo se compara contra cero, que indica una distribución simétrica con respecto a la media.

$$\alpha_3(X) \begin{cases} > 0 \;; & \textit{sesgo positivo} \\ = 0 \;; & \textit{distribución simétrica} \\ < 0 \;; & \textit{sesgo negativo} \end{cases}$$


Sesgo positivo


Simétrica


Sesgo negativo

El sesgo positivo también se llama sesgo hacia la derecha, y el sesgo negativo hacia la izquierda.

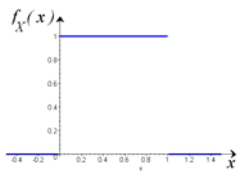
Coeficiente de curtosis

El coeficiente de curtosis mide el grado de aplanamiento, o bien, indica que tan puntiaguda es la distribución. Se define como el cuarto momento con respecto a la media estandarizado. Se denota por $\alpha_4(X)$.

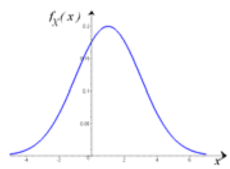
$$\alpha_4(X) = \frac{\mu_4}{\sigma_X^4}$$

El valor del coeficiente de curtosis se compara con el número tres.

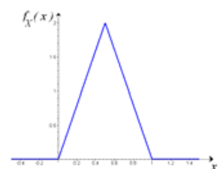
$$\alpha_4(X) \begin{cases} < 3 & \textit{platicúrtica} \\ = 3 & \textit{mesocúrtica} \\ > 3 & \textit{leptocúrtica} \end{cases}$$



Platicúrtica



Mesocúrtica



Leptocúrtica

S))))))))))))))))))))))))))))))))))))Q
Ejemplo 2.20
Considérese una variable aleatoria continua con la función de densidad dada por

$$f_X(x) = \begin{cases} x & ; \quad 0 < x < 1 \\ 2 - x & ; \quad 1 \le x < 2 \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

Determinar todas las características numéricas de la variable aleatoria X.

Resolución
Medidas de tendencia central:

Media, $\mu_X = \mathrm{E}(X) = \int_0^1 x^2\, dx + \int_1^2 x(2 - x)\, dx = 1$

Moda, $x_{mo} = 1$

Mediana,
Para encontrar la mediana deben de utilizarse en orden las reglas de correspondencia de la función de densidad, hasta obtener la regla con la cual se acumule la probabilidad de 0.5
Para la primera regla de correspondencia, con $0 \le x \le 1$, se tiene:

$$0.5 = P(X \le \tilde{x}) = \int_0^{\tilde{x}} x\, dx = \frac{\tilde{x}}{2}$$

Esto es $0.5 = \dfrac{\tilde{x}}{2}$, de donde $\tilde{x} = 1$, y puesto que el valor obtenido sí se encuentra dentro de intervalo $0 \le x \le 1$, entonces la mediana es $\tilde{x} = 1$.

Medidas de dispersión:

Rango, $R_X = 2 - 0 = 2$

Desviación media

$$DM_X = \int_0^1 \left| x - \mu_X \right| x \, dx + \int_1^2 \left| x - \mu_X \right| (2 - x) \, dx$$

por simetría

$$DM_X = 2 \int_0^1 (1 - x) \, x \, dx = \frac{1}{3}$$

Variancia

$$\sigma_X^2 = \int_0^1 (x - 1)^2 x \, dx + \int_1^2 (x - 1)^2 (2 - x) \, dx$$

$$\sigma_X^2 = \frac{1}{6}$$

Desviación estándar

$$\sigma_X = \sqrt{\sigma_X^2} = \frac{\sqrt{6}}{6}$$

Coeficiente de variación

$$CV_X = \frac{\sigma_X}{\mu_X} = \frac{\frac{\sqrt{6}}{6}}{1} = \frac{\sqrt{6}}{6}$$

Parámetros de forma:

Sesgo, $\alpha_3 = \dfrac{\mu_3}{\sigma_X^3}$

pero

$$\mu_3 = \mathrm{E}\left[(X - \mu_X)^3 \right] = \int_0^1 (x - 1)^3 x \, dx + \int_1^2 (x - 1)^3 (2 - x) \, dx$$

$$\mu_3 = 0$$

por lo que $\alpha_3 = 0$ La distribución es simétrica.

Curtosis $\alpha_4 = \dfrac{\mu_4}{\sigma_X^4}$

pero

$$\mu_4 = \mathrm{E}\left[(X - \mu_X)^4 \right] = \int_0^1 (x - 1)^4 x \, dx + \int_1^2 (x - 1)^4 (2 - x) \, dx$$

$$\mu_4 = \frac{1}{15}$$

por lo que

$$\alpha_4 = \frac{\frac{1}{15}}{\left(\frac{1}{6}\right)^2} = \frac{12}{5} = 2.4$$

La distribución es platicúrtica.

S))))))))))))))))))))))))))))))))))Q

Ejemplo 2.21 Rendimiento Esperado

Un empresario debe decidir entre invertir en el proyecto A o el B. Según los estudios realizados por sus analistas cada proyecto puede generar distintos rendimientos dependiendo de situaciones relacionadas con la competencia y la política nacional. En resumen se tiene la siguiente información:

Probabilidad	$R_A\ (\%)$	$R_B\ (\%)$
0.1	-25	-40
0.2	5	0
0.4	15	16
0.2	30	40
0.1	45	66

a) Determinar el rendimiento esperado de los activos A y B.

b) Determinar la desviación estándar y coeficiente de variación de los rendimientos de los proyectos A y B.

c) Jerarquizar los proyectos con base en el rendimiento esperado, el riesgo del proyecto y el riesgo por rendimiento prometido.

Resolución

a) Sean R_A y R_B los rendimientos de los proyectos A y B.

$$E\left(R_A\right) = \sum_{\forall\, r_A} r_A\, f\left(r_A\right) = -25\,(0.1) + 5\,(0.2) + $$
$$+ 15\,(0.4) + 30\,(0.2) + 45\,(0.1)$$

$$E\left(R_A\right) = 15$$

$$E\left(R_B\right) = \sum_{\forall\, r_B} r_B\, f\left(r_B\right) = -40\,(0.1) + 0\,(0.2) + $$
$$+ 16\,(0.4) + 40\,(0.2) + 66\,(0.1)$$

$$E\left(R_B\right) = 17$$

b) $$\text{Var}\left(R_A\right) = E\left[\ \left(R_a - E\left(R_A\right)\right)^2\ \right] = E\left(R_A^2\right) - \left[\ E\left(R_A\right)\right]^2$$

$$\text{Var}\left(R_A\right) = -(25)^2(0.1) + (5)^2(0.2) + (15)^2(0.4) + $$
$$+ (30)^2(0.2) + (45)^2(0.1) - (15)^2$$

$$\text{Var}\left(R_A\right) = 315 \qquad \Rightarrow \qquad \sigma_{R_A} = 17.748$$

$$\text{Var}\left(R_B\right) = 729 \qquad \Rightarrow \qquad \sigma_{R_B} = 27$$

$$CV_{R_A} = \frac{17.748}{15} = 1.183$$

$$CV_{R_B} = \frac{27}{17} = 1.588$$

c) Con base en el rendimiento esperado, el mejor proyecto es el B .
Con base en el riesgo del proyecto, el mejor proyecto es el A .
Con base en el riesgo por rendimiento prometido, el mejor proyecto es el A .

S))))))))))))))))))))))))))))))))))))Q

TÓPICOS ESPECIALES: TEOREMA DE CHEBYSHEV

El teorema de Chebyshev o la desigualdad de Chebyshev proporciona una cota para la probabilidad de que una variable aleatoria se aleje un cierto número de desviaciones estándar de la media. La cota no siempre genera un intervalo reducido; sin embargo, se requiere muy poca información de la variable aleatoria para generar el intervalo, de hecho, basta con conocer la media y la desviación estándar.

Teorema de Chebyshev

La probabilidad de que cualquier variable aleatoria X tome un valor dentro de k desviaciones estándar de la media es por lo menos $1 - \frac{1}{k^2}$; es decir,

$$P(\mu_X - k\sigma < X < \mu_X + k\sigma) \geq 1 - \frac{1}{k^2}$$

De hecho, la desigualdad de Chebyshev se puede escribir también como:

$$P(\,|X - \mu_X| < k\sigma\,) \geq 1 - \frac{1}{k^2}$$

o bien, utilizando el complemento

$$P(\,|\,X - \mu_X\,| \geq k\sigma\,) \leq \frac{1}{k^2}$$

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.22

Una variable aleatoria X tiene una media $\mu = 12$, con una variancia $\sigma^2 = 9$, y una distribución de probabilidades conocida. Utilizando el teorema de Chebyshev, obtener:

a) $P(\,6 < X < 18\,)$.

b) $P(\,3 < X < 21\,)$.

Resolución

a) $P(6 < X < 18) = P[\,12 - 2(3) < X < 12 + 2(3)\,] \geq \dfrac{3}{4}$

b) $P(3 < X < 21) = P[\,12 - 3(3) < X < 12 + 3(3)\,] \geq \dfrac{8}{9}$

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 2.23

Tener a mano un suministro adecuado de refacciones es función importante del almacén de una gran empresa electrónica. Se estudió la demanda mensual de tarjetas para impresoras de microcomputadoras durante algunos meses y se vio que el promedio o media es 28 y la desviación estándar es 4. ¿Cuántas tarjetas de impresora deben tener a la mano al principio de cada mes para asegurar que la demanda será mayor que la oferta cuando mucho con una probabilidad de 0.10?

Resolución

Sea X la variable aleatoria que representa la demanda.

Utilizando el teorema de Chebyshev con $k = \sqrt{10}$ se tiene :

$$P(\,X - \mu \geq k\sigma\,) \leq P(\,|\,X - \mu\,| \geq k\sigma\,) \leq \frac{1}{k^2} = \frac{1}{(\sqrt{10}\,)^2} = 0.1$$

es decir :

$P(\,X - 28 \geq \sqrt{10}\,(4)\,) \leq 0.1$

$P(\,X \geq 40.65\,) \leq 0.1$

Por lo que deben existir 41 tarjetas de impresora en el inventario.

S))))))))))))))))))))))))))))))))))))Q

AUTOEXAMEN TEMA II

1.- Sea X una variable aleatoria con función de densidad

$$f_X(x) = \begin{cases} cx \ , & 0 \le x \le 10 \\ 0 \ , & \textit{en otro caso} \end{cases}$$

donde $c \in \mathbb{R}$. La probabilidad de que X esté entre 1 y 2 es: $\square$

A) $\frac{1}{100}$ B) $\frac{3}{100}$ C) $\frac{5}{100}$ D) $\frac{7}{100}$ E) $\frac{9}{100}$

2.- Los siguientes enunciados se relacionan con una distribución de probabilidad discreta, EXCEPTO: $\square$

A) La suma de todas la probabilidades es uno.
B) Se puede presentar como una fórmula.
C) A cada valor de la variable se le asigna una probabilidad.
D) Se puede representar en forma gráfica.
E) Su función de distribución es continua.

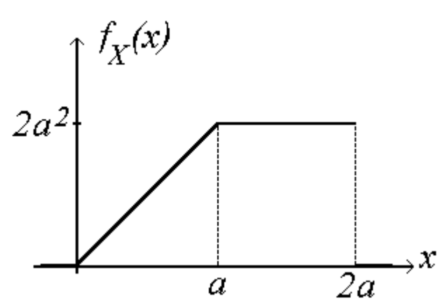
3.- Sea Y una variable aleatoria discreta con función de distribución

$$F_Y(y) = \begin{cases} 0 & ; \quad y < -1 \\ 0.2 & ; \quad -1 \le y < 0 \\ 0.5 & ; \quad 0 \le y < 3 \\ 1 & ; \quad y \ge 3 \end{cases}$$

Entonces $P(0 \le Y \le 2)$ es: $\square$

A) 1 B) 0 C) 0.2 D) 0.3 E) 0.5

4.- Si la variable aleatoria X tiene la función de densidad que se muestra en la figura, entonces el valor de a es: $\square$



A) 2 B) 1 C) $\frac{1}{2}$ D) $\frac{1}{\sqrt{2}}$ E) $\frac{1}{\sqrt[3]{3}}$

5.- Dada la variable aleatoria Y con función de densidad

$$f_Y(y) = \begin{cases} a\,e^{-3y}, & y > 0 \\ 0, & \textit{en otro caso} \end{cases}$$

entonces el valor de $a \in \mathbb{R}$ es: $\square$

A) $\frac{1}{3}$ B) 1 C) 2 D) 3

E) Ninguna de las anteriores.

6.- Una compañía de seguros debe determinar la cuota anual a cobrarse por un seguro de cincuenta mil pesos para hombres cuya edad se encuentra entre los 30 y 35 años. Con base en las tablas actuariales el número de fallecimientos al año, para este grupo, es de 5 por cada mil. Si X es la v.a. que representa la ganancia de la compañía de seguros, entonces el monto de la cuota anual para que la compañía no pierda, a pesar de tener un número grande de tales seguros es: $\square$

A) 100 B) 250 C) 5000 D) 50,000

E) Ninguno de los anteriores.

7.- Dada la función de densidad

$$f_X(x) = \begin{cases} x & ; \; 0 \le x \le 1 \\ 2 - x & ; \; 1 < x \le 2 \\ 0 & ; \; \textit{en otro caso} \end{cases}$$

la función de distribución acumulativa es: $\square$

A) $$F_X(x) = \begin{cases} 0 & ; \; x < 0 \\ \frac{1}{2}x^2 & ; \; 0 \le x \le 1 \\ 2x - \frac{1}{2}x^2 - \frac{3}{2} & ; \; 1 < x \le 2 \\ 1 & ; \; x > 2 \end{cases}$$

B) $F_X(x) = \begin{cases} \frac{1}{2}x^2 & , \ 0 \leq x \leq 1 \\ 2x - \frac{1}{2}x^2 - \frac{3}{2} & , \ 1 < x \leq 2 \\ 0 & , \ \textit{en otro caso} \end{cases}$

C) $F_X(x) = \begin{cases} 0 & , \ x < 0 \\ \frac{1}{2}x^2 & , \ 0 \leq x \leq 1 \\ 2x - \frac{1}{2}x^2 - 1 & , \ 1 < x \leq 2 \\ 1 & , \ x > 2 \end{cases}$

D) $F_X(x) = \begin{cases} \frac{1}{2}x^2 & , \ 0 \leq x \leq 1 \\ 2x - \frac{1}{2}x^2 - \frac{3}{2} & , \ 1 < x \leq 2 \\ 1 & , \ \textit{en otro caso} \end{cases}$

E) Ninguna de las anteriores.

8.- El tiempo que toma reparar una computadora personal es una variable aleatoria con función de densidad

$$f_X(x) = \begin{cases} \frac{1}{2} & , \ 0 < x < 2 \\ 0 & , \ \textit{en otro caso} \end{cases}$$

El costo de la reparación depende del tiempo, y es igual a $40 + 30\sqrt{x}$. El costo esperado al reparar una computadora personal es: $\boxed{}$

A) 68.28 B) 70 C) 1 D) 16.56 E) Ninguna de las anteriores.

9.- Dada la función de probabilidad

X	0	1	2	3
$f_X(x)$	0.1	0.2	0.4	0.3

Su función generadora de momentos es:

1) $0.1 + 0.2\,e^{t} + 0.4\,e^{2t} + 0.3\,e^{3t}$

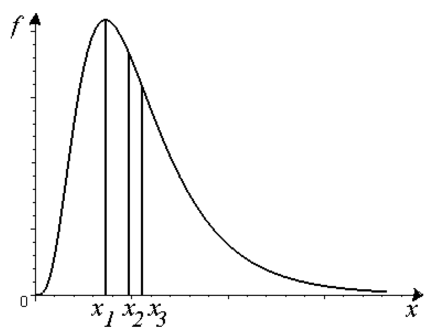
2) $e^{0.2t} + 2\,e^{0.4t} + 3\,e^{0.3t}$

3) $0.1\,t + 0.2\,t + 2\,(0.4)\,t + 3\,(0.3)\,t$

4) $e^{0.1t} + e^{0.2t} + e^{0.4t} + e^{0.3t}$

5) $e^{t}\left(\dfrac{0.1}{0!} + \dfrac{0.2}{1!} + \dfrac{0.4}{2!} + \dfrac{0.3}{3!}\right)$

10.- Si la variable aleatoria X tiene una distribución asimétrica a la derecha entonces x_1, x_2 y x_3 corresponden



en ese orden, a: $\boxed{}$

A) La media, la moda y la mediana.
B) La moda, la media y la mediana.
C) La mediana, la media y la moda.
D) La moda, la mediana y la media.
E) Ninguna de las anteriores.

BIBLIOGRAFÍA

Hines, William W. y Montgomery, Douglas C. - Probabilidad y Estadística para ingeniería, cuarta edición..- CECSA.- México, 2005.

Milton, Susan J. Y Arnold, Jesse C.- Probabilidad y Estadística para con aplicaciones para ingeniería y ciencias computacionales, cuarta edición.- McGraw-Hill.- México, 2004.

Devore, Jay L.- Probabilidad y Estadística para ingeniería y ciencias, séptima edición.- Cengage Learning.- México, 2008.

Mendenhall, William III. et al.- Introducción a la Probabilidad y Estadística.- Décimo cuarta edición.- Cengage Learning.- México 2015.

Wackerly Dennis D.- Mendenhall, William, *et al.*- Estadística Matemática con Aplicaciones, sexta edición.- Editorial Thomson.- México, 2002.

Walpole, Ronald E., *et al.*- Probability and Statistics for Engineers and Scientists.- Pearson.- USA, 2007.

Montgomery, Douglas C. y Runger, George C.-Probabilidad y Estadística aplicadas a la Ingeniería, segunda edición.- Limusa-Wiley.- México, 2002.

Scheaffer, Richard L. y McClave, James T.- Probabilidad y Estadística para Ingeniería.- Grupo Editorial Iberoamérica.- México, 1993.

Canavos, George C.- Probabilidad y Estadística Aplicaciones y Métodos.- McGraw-Hill.- México, 1988.

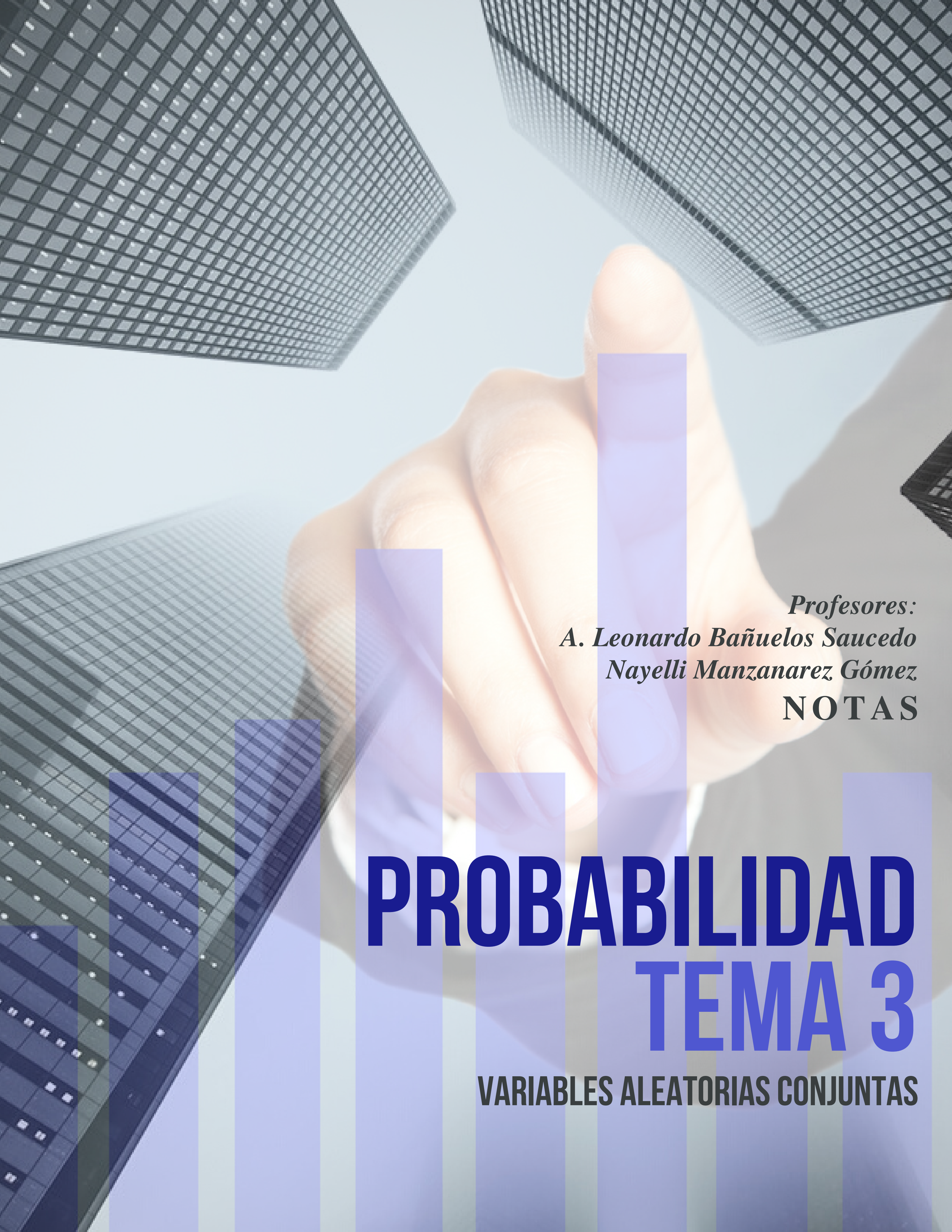
Meyer, Paul L.- Probabilidad y Aplicaciones Estadísticas.- Addison Wesley Iberoamericana.- México, 1992.

Spiegel, Murray R. et al.- Probabilidad y Estadística, cuarta edición.- Mc Graw-Hill.-México 2013.

Borras García, Hugo E., *et al.*- Apuntes de Probabilidad y Estadística.-Facultad de Ingeniería.- México, 1985.

Rosenkrantz, Walter A.- Introduction to Probability and Statistics for Scientists and Engineers.- McGraw-Hill.- EE.UU., 1997.

Ziemer, Rodger E.- Elements of Engineering Probability & Statistics.- Prentice Hall.- USA 1997.



Profesores:
A. Leonardo Bañuelos Saucedo
Nayelli Manzanarez Gómez

NOTAS

PROBABILIDAD
TEMA 3
VARIABLES ALEATORIAS CONJUNTAS

TEMA III
VARIABLES ALEATORIAS CONJUNTAS

INTRODUCCIÓN

En el capítulo anterior se estudiaron variables aleatorias unidimensionales; es decir, se estudió cada variable de manera independiente; sin embargo, en muchas ocasiones es necesario estudiar dos o más características (variables) de un experimento. En cualquier sistema físico, económico o social, que se estudie a través de modelos aleatorios, pueden existir varias variables, y debido a la interrelación entre ellas, deben estudiarse modelos que describan el comportamiento probabilístico conjunto de dichas variables.

Definición 3.1

Si $X_1, X_2, X_3, \ldots, X_n$ son variables aleatorias definidas sobre un mismo espacio muestral, dichas variables aleatorias reciben el nombre de *VARIABLES ALEATORIAS CONJUNTAS.*

Así como se dijo en el tema anterior, que una variable aleatoria responde a una pregunta de interés en un fenómeno, las variables aleatorias conjuntas se puede decir que representan varias características (o responden a varias preguntas) de interés en un mismo fenómeno. Y al igual que en el tema anterior, existen variables aleatorias conjuntas discretas y continuas.

VARIABLES ALEATORIAS CONJUNTAS DISCRETAS

Definición 3.2

Si $X_1, X_2, X_3, \ldots, X_n$ son variables aleatorias conjuntas discretas, se define su función de probabilidad conjunta como:

$$f_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n) = P\left(X_1 = x_1 \cap X_2 = x_2 \cap \ldots \cap X_n = x_n\right)$$
$$= P\left(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\right)$$

Es bastante común utilizar p_{XY} para denotar la función de probabilidad conjunta; sin embargo, aquí se prefiere la notación $f_{XY}(x,y)$, puesto que también es una función y facilita la generalización de diversas expresiones.

En particular si $n = 2$

$$f_{XY}(x,y) = P\left(X = x \cap Y = y\right) = P\left(X = x, Y = y\right)$$

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 3.1

Considérese el lanzamiento de dos dados.
El espacio muestral es

$$S = \left\{ \begin{matrix} (1,1), & (1,2), & (1,3), & (1,4), & (1,5), & (1,6) \\ (2,1), & (2,2), & (2,3), & (2,4), & (2,5), & (2,6) \\ (3,1), & (3,2), & (3,3), & (3,4), & (3,5), & (3,6) \\ (4,1), & (4,2), & (4,3), & (4,4), & (4,5), & (4,6) \\ (5,1), & (5,2), & (5,3), & (5,4), & (5,5), & (5,6) \\ (6,1), & (6,2), & (6,3), & (6,4), & (6,5), & (6,6) \end{matrix} \right\}$$

Si interesa el resultado del primer dado y la suma de los puntos en los dos lanzamientos, entonces Sea

X_1 : El resultado del primer lanzamiento

X_2 : La suma de los resultados de los dos lanzamientos.

$R_{X_1} = \{ 1 , 2 , \ldots , 6 \}$

$R_{X_2} = \{ 2 , 3 , 4 , \ldots , 12 \}$

La función de probabilidad conjunta para X_1 y para X_2 se construye de la siguiente forma:

$$f_{X_1 X_2}(1 , 2) = P\left(X_1 = 1 , X_2 = 2\right) = \frac{1}{36}$$

$$f_{X_1 X_2}(1 , 3) = P\left(X_1 = 1 , X_2 = 3\right) = \frac{1}{36}$$

.

.

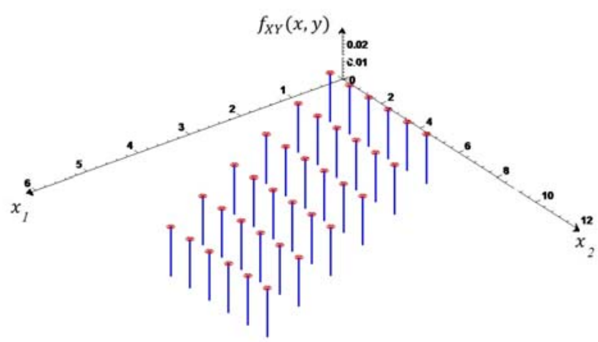
.

Anotando las probabilidades en una tabla se tiene:

$f_{X_1X_2}(x_1\ ,\ x_2)$		x_1					
		1	2	3	4	5	6
x_2	2	$\frac{1}{36}$	0	0	0	0	0
	3	$\frac{1}{36}$	$\frac{1}{36}$	0	0	0	0
	4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	0	0	0
	5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	0	0
	6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	0
	7	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
	8	0	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
	9	0	0	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
	10	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
	11	0	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$
	12	0	0	0	0	0	$\frac{1}{36}$

GRÁFICA DE UNA FUNCIÓN DE PROBABILIDAD CONJUNTA

Al igual que para las variables aleatorias unidimensionales, lo más común es dibujar rectas verticales cuya altura representa la probabilidad, también se pueden dibujar los puntos; sin embargo, como puede observarse, la lectura de una gráfica de una función de probabilidad conjunta se complica y sólo puede hacerse para el caso de dos variables.



Gráfica 3.1. Gráfica del ejemplo 3.1

CARACTERÍSTICAS DE LA FUNCIÓN DE PROBABILIDAD CONJUNTA

Si X , Y son dos variables aleatorias conjuntas discretas. Su función de probabilidad conjunta tiene las siguientes características, basadas en los axiomas de la probabilidad

1) La función siempre está entre cero y uno.

$$0 \leq f_{XY}(x, y) \leq 1 \quad \forall \quad (x, y)$$

2) La suma de la función para todos los valores de las variables es la unidad.

$$\sum_{\forall R_X} \sum_{\forall R_Y} f_{XY}(x, y) = 1$$

3) La probabilidad de que las variables se encuentren en un intervalo, es la suma de la función de probabilidad par dicho intervalo.

$$P\left(x_0 \leq X \leq x_1, \ y_0 \leq Y \leq y_1\right) = \sum_{x = x_0}^{x_1} \sum_{y = y_0}^{y_1} f_{XY}(x, y)$$

FUNCIONES DE PROBABILIDAD: MARGINAL Y CONDICIONAL

Partiendo de un función de probabilidad conjunta de dos variables aleatorias, puede interesarnos el comportamiento aleatorio de una sola de las variables, ya sea porque la otra variable no nos interese en absoluto o porque ya tomó algún valor. Esto da origen a las funciones de probabilidad marginales y las funciones de probabilidad condicionales, que puede generalizarse para funciones de n variables aleatorias.

Si X y Y son dos variables aleatorias conjuntas discretas, entonces se pueden hacer las siguientes definiciones.

La función de probabilidad marginal de X (función de probabilidad de X al margen de Y) como

$$f_X(x) = \sum_{\forall R_Y} f_{XY}(x, y)$$

La función de probabilidad marginal de Y (función de probabilidad de Y al margen de X) como

$$f_Y(y) = \sum_{\forall R_X} f_{XY}(x, y)$$

La función de probabilidad condicional de X dado que $Y = y_0$ es:

$$f_{X|y_0}\left(x \mid Y = y_0\right) = \begin{cases} \dfrac{f_{XY}(x, y_0)}{f_Y(y_0)} & ; \quad f_Y(y_0) > 0 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

La función de probabilidad condicional de Y dado que $X = x_0$ es:

$$f_{Y\,|\,x_0}\left(y \mid X = x_0\right) = \begin{cases} \dfrac{f_{XY}\left(x_0\,,\,y\right)}{f_X\left(x_0\right)} & ; \quad f_X\left(x_0\right) > 0 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 3.2

Para la función de probabilidad conjunta discreta estudiada anteriormente, del lanzamiento de dos dados, obtener las marginales.

Resolución

Una forma de escribir las marginales es agregando una columna y un renglón a la tabla de la función de probabilidad.

$f_{X_1 X_2}\left(x_1\,,\,x_2\right)$		x_1						
		1	2	3	4	5	6	$f_{X_2}\left(x_2\right)$
x_2	2	$\frac{1}{36}$	0	0	0	0	0	$\frac{1}{36}$
	3	$\frac{1}{36}$	$\frac{1}{36}$	0	0	0	0	$\frac{2}{36}$
	4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	0	0	0	$\frac{3}{36}$
	5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	0	0	$\frac{4}{36}$
	6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	0	$\frac{5}{36}$
	7	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{6}{36}$
	8	0	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{5}{36}$
	9	0	0	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{4}{36}$
	10	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{3}{36}$
	11	0	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{2}{36}$
	12	0	0	0	0	0	$\frac{1}{36}$	$\frac{1}{36}$
	$f_{X_1}\left(x_1\right)$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	1

O bien, escribiendo las tablas por separado

x_1	1	2	3	4	5	6
$f_{X_1}(x_1)$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$

x_2	2	3	4	5	6	7	8	9	10	11	12
$f_{X_2}(x_2)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

S))))))))))))))))))))))))))))))))))Q

Ejemplo 3.3

Cierto artículo se fabrica en dos líneas de producción diferentes. La capacidad en cualquier día dado para cada línea es de dos artículos. Supóngase que el número de artículos producidos por la línea uno en un día cualquiera es una variable aleatoria X y que el número de artículos producidos por la línea dos está dado por Y. Con base en datos estadísticos se obtuvo la siguiente tabla que corresponde a la función de probabilidad conjunta de X y Y.

$f_{XY}(x, y)$		x		
		0	1	2
y	0	0.10	0.20	0.20
	1	0.04	0.08	0.08
	2	0.06	0.12	0.12

a) Calcular la probabilidad de que en un día dado el número de artículos producidos en la línea uno sea mayor que el número de artículos producidos en la línea dos .

b) Obtener las funciones de probabilidad marginales de X y Y

c) Determinar la función de probabilidad condicional de X dado $Y=1$

d) Calcular $P\left(X < 2 \mid Y = 1\right)$

Resolución

a) $P(X>Y) = P(X=1, Y=0) + P(X=2, Y=0) + P(X=2, Y=1)$

$$= 0.2 + 0.2 + 0.08 = 0.48$$

b)

$f_{XY}(x\,,\,y)$		$x=0$	$x=1$	$x=2$	$f_Y(y)$
y	0	0.10	0.20	0.20	0.5
	1	0.04	0.08	0.08	0.2
	2	0.06	0.12	0.12	0.3
	$f_X(x)$	0.20	0.40	0.40	1.0

c) $f_{X\,|\,Y\,=\,1}\left(x\,|\,Y\,=\,1\right)$

$$P\left(X\,=\,0\,|\,Y\,=\,1\right)\,=\,\frac{f_{XY}(X\,=\,0\,,\,Y\,=\,1)}{f_Y(1)}\,=\,\frac{0.04}{0.2}\,=\,0.2$$

$$P\left(X\,=\,1\,|\,Y\,=\,1\right)\,=\,\frac{f_{XY}(X\,=\,1\,,\,Y\,=\,1)}{f_Y(1)}\,=\,\frac{0.08}{0.2}\,=\,0.4$$

$$P\left(X\,=\,2\,|\,Y\,=\,1\right)\,=\,\frac{f_{XY}(X\,=\,2\,,\,Y\,=\,1)}{f_Y(1)}\,=\,\frac{0.08}{0.2}\,=\,0.4$$

Construyendo la tabla $f_{X\,|\,Y\,=\,1}\left(x\,|\,Y\,=\,1\right)$

x	0	1	2
$f_{X\,\mid\,Y\,=\,1}\left(x\,\mid\,Y\,=\,1\right)$	0.2	0.4	0.4

d) De la tabla:

$$P\left(X<2\,|\,Y\,=\,1\right)\,=\,0.2\,+\,0.4\,=\,0.6$$

S))))))))))))))))))))))))))))))))))))Q

VARIABLES ALEATORIAS CONJUNTAS CONTINUAS

Cuando nos interesa el comportamiento probabilístico de dos o mas variables aleatorias continuas, es necesario estudiar la función de densidad conjunta.

Definición 3.3

Si $X_1, X_2, X_3, \ldots, X_n$ son variables aleatorias conjuntas continuas, entonces su función de densidad conjunta se define como una función con las siguientes propiedades

1) La función no es negativa

$$f_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n) \geq 0 \qquad \forall \ x_1, x_2, \ldots, x_n$$

2) La integral múltiple sobre todo el rango de las variables aleatorias es la unidad.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n = 1$$

3) La probabilidad de que las variables se encuentren en una región es igual a la integral de la función de densidad integrada sobre dicha región

$$P\left((x_1, x_2, \ldots, x_n) \in \mathbb{R}\right) = \iint \cdots \int_{\mathbb{R}} f_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n$$

Al igual que en los casos anteriores, las primeras dos propiedades nos sirven para identificar a una función de densidad conjunta, mientras que la tercera propiedad nos sirve para calcular probabilidades.

S))))))))))))))))))))))))))))))))))Q

Ejemplo 3.4

Sean X, Y dos variables aleatorias conjuntas con función de densidad

$$f_{XY}(x, y) = \begin{cases} k(x + y) & ; \ 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

a) Calcular el valor de k para el cual $f_{XY}(x,y)$ es función de densidad

b) $P(0.5 \leq X \leq 1,\ 0.5 \leq Y \leq 1)$

c) $P(X + Y \leq 1)$

Resolución

a) Para el caso de dos variables aleatorias la propiedad 2 puede reescribirse como

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)\, dx\, dy = 1$$

Por lo que

$$1 = \int_0^1 \int_0^1 k(x + y)\, dx\, dy$$

))

$$1 = k \int_0^1 \left[\frac{x^2}{2} + yx \right]_0^1 dy = k \int_0^1 \left(\frac{1}{2} + y \right) dy$$

$$1 = k \left[\frac{1}{2} y + \frac{y^2}{2} \right]_0^1$$

$$= k \left(\frac{1}{2} + \frac{1}{2} \right) = k\,(1) \quad \Rightarrow \quad k = 1$$

b) $P\,(0.5 \le X \le 1\,,\ 0.5 \le Y \le 1) = \displaystyle\int_{0.5}^1 \int_{0.5}^1 (x + y)\, dx\, dy$

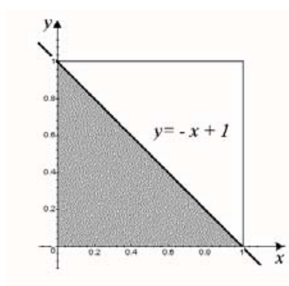
$$= \int_{0.5}^1 \left[\frac{x^2}{2} + xy \right]_{0.5}^1 dy$$

$$= \int_{0.5}^1 \left[\left(\frac{1}{2} + y \right) - \left(\frac{1}{8} + \frac{y}{2} \right) \right] dy$$

$$= \int_{0.5}^1 \left(\frac{y}{2} + \frac{3}{8} \right) dy = \left[\frac{y^2}{4} + \frac{3}{8} y \right]_{0.5}^1$$

$$= \frac{4}{16} + \frac{6}{16} - \frac{4}{16} = \frac{6}{16} = 0.375$$

c) Para obtener $P\,(X + Y \le 1)$, se dibuja la región de integración



$$P\,(X + Y \le 1) = \int_0^1 \int_0^{-x+1} (x + y)\, dy\, dx = \int_0^1 \left[xy + \frac{y^2}{2} \right]_{y=0}^{y=-x+1} dx$$

$$= \int_0^1 \left[\frac{(-x+1)^2}{2} + x(-x+1) \right] dx = \int_0^1 \left(\frac{-x^2+1}{2} \right) dx$$

$$= \left[-\frac{x^3}{6} + \frac{x}{2} \right]_0^1 = -\frac{1}{6} + \frac{1}{2} = \frac{1}{3} \approx 0.333$$

S))))))))))))))))))))))))))))))))))Q

FUNCIONES DE DENSIDAD: MARGINAL Y CONDICIONAL

Al igual que para las variables aleatorias conjuntas discretas, para las continuas también se definen las funciones de densidad marginal y condicional.

Si X y Y son dos variables aleatorias conjuntas continuas, entonces se define:

La función de densidad marginal de X (función de probabilidad de X al margen de Y) como

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x\,,\,y)\;dy$$

La función de densidad marginal de Y (función de probabilidad de Y al margen de X) como

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x\,,\,y)\;dx$$

La función de densidad condicional de X dado que $Y = y_0$ es:

$$f_{X\,|\,y_0}\left(x \,\middle|\, Y = y_0\right) = \begin{cases} \dfrac{f_{XY}(x\,,\,y_0)}{f_Y(y_0)} & ; \quad f_Y(y_0) > 0 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

La función de densidad condicional de Y dado que $X = x_0$ es:

$$f_{Y\,|\,x_0}\left(y \,\middle|\, X = x_0\right) = \begin{cases} \dfrac{f_{XY}(x_0\,,\,y)}{f_X(x_0)} & ; \quad f_X(x_0) > 0 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

Obsérvese que la distribución condicional, se obtiene con la misma expresión tanto para variables aleatorias discretas como continuas.

S))))))))))))))))))))))))))))))))))Q

Ejemplo 5.5

Sean X_1 y X_2 las proporciones de dos sustancias distintas que se encuentran en una muestra de una mezcla de reactivos que se usa como insecticida. Supóngase que X_1 y X_2 tienen una densidad de probabilidad conjunta representada por

$$f_{X_1 X_2}(x_1, x_2) = \begin{cases} 2 \ ; & 0 \le x_1 \ , \ 0 \le x_2 \le 1 \ , \ x_1 + x_2 \le 1 \\ 0 \ ; & \textit{en otro caso} \end{cases}$$

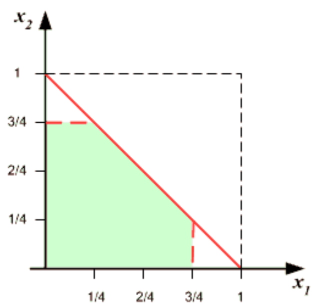
a) Calcular $P\left(X_1 \le \frac{3}{4} \, , X_2 \le \frac{3}{4} \right)$.

b) Determinar $P\left(X_1 \le \frac{1}{2} \, , X_2 \le \frac{1}{2} \right)$.

c) Calcular $P\left(X_1 \le \frac{1}{2} \;\middle|\; X_2 \le \frac{1}{2} \right)$.

Resolución

a) La región es

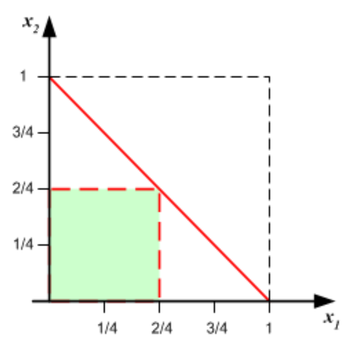


$$P\left(X_1 \le \frac{3}{4}, X_2 \le \frac{3}{4} \right) = \int_0^{\frac{1}{4}} \int_0^{\frac{3}{4}} 2 \ dx_2 dx_1 + \int_{\frac{1}{4}}^{\frac{3}{4}} \int_0^{1-x_1} 2 \, dx_2 \, dx_1$$

$$P\left(X_1 \le \frac{3}{4}, X_2 \le \frac{3}{4} \right) = 2\left(\frac{1}{4} \right)\left(\frac{3}{4} \right) + \int_{\frac{1}{4}}^{\frac{3}{4}} 2(1 - x_1) \, dx_1$$

$$P\left(X_1 \le \frac{3}{4} \, , X_2 \le \frac{3}{4} \right) = \frac{3}{8} + \frac{1}{2} = \frac{7}{8}$$

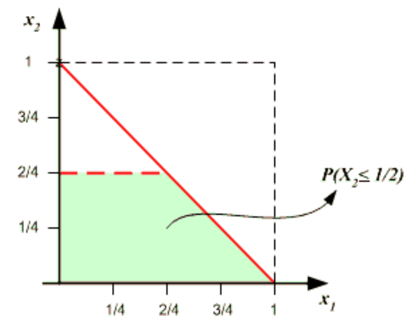
b) La región es:



$$P\left(X_1 \le \frac{1}{2} \, , X_2 \le \frac{1}{2} \right) = 2\left(\frac{1}{2} \right)\left(\frac{1}{2} \right) = \frac{1}{2}$$

c) $$P\left(X_1 \le \frac{1}{2} \;\middle|\; X_2 \le \frac{1}{2} \right) = \frac{P\left(X_1 \le \frac{1}{2} \,,\, X_2 \le \frac{1}{2} \right)}{P\left(X_2 \le \frac{1}{2} \right)}$$

La región de integración para $P\left(X_2 \le \frac{1}{2} \right)$ es:



$$P\left(X_1 \le \frac{1}{2} \;\middle|\; X_2 \le \frac{1}{2} \right) = \frac{\frac{1}{2}}{\int_0^{\frac{1}{2}} \int_0^{1 - x_2} 2 \; dx_1 \, dx_2} = \frac{2}{3}$$

S))))))))))))))))))))))))))))))))))Q

Ejemplo 3.6

Sea la función de densidad conjunta

$$f_{XY}(x,y) = \begin{cases} x^3 + \dfrac{x y^2}{a} & ; \quad 0 \le x \le 1 \;,\; 0 \le y \le 2 \\ 0 & ; \quad en \;\; otro \;\; caso \end{cases}$$

a) Obtener el valor de $a \in \mathbb{R}$.

b) Calcular $P(X \le Y)$.

c) Obtener las funciones marginales $f_X(x)$ y $f_Y(y)$.

d) Obtener las funciones de densidad condicional $f_{X|y_0}(x \mid y_0)$ y $f_{Y|x_0}(y \mid x_0)$.

Resolución

a) Para que la función sea una función de densidad conjunta debe de cumplirse que $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{XY}(x,y) \; dx \; dy = 1$, por lo que:

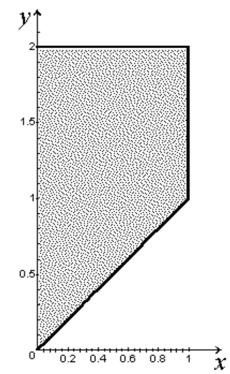
$$\int_0^2 \int_0^1 \left(x^3 + \frac{x y^2}{a} \right) \; dx \; dy = 1$$

de donde

$$\int_0^2 \frac{1}{4}\,\frac{a+2y^2}{a}\,dy = \frac{1}{6}\,\frac{3a+8}{a} = 1$$

$$a = \frac{8}{3}$$

b) La región de integración es:



Por lo que

$$P(X \le Y) = \int_1^2\int_0^1 \left(x^3 + \frac{3}{8}xy^2 \right) dx\,dy + \int_0^1\int_0^y \left(x^3 + \frac{3}{8}xy^2 \right) dx\,dy$$

o bien utilizando el complemento

$$P(X \le Y) = 1 - P(X > Y) = 1 - \int_0^1\int_y^1 \left(x^3 + \frac{3}{8}xy^2 \right) dx\,dy$$

$$P(X \le Y) = 1 - \int_0^1 \left(\frac{1}{4} + \frac{3}{16}y^2 - \frac{7}{16}y^4 \right) dy$$

$$P(X \le Y) = 1 - \frac{9}{40} = \frac{31}{40} = 0.775$$

c) Las marginales se obtienen de

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\,dy$$

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\,dx$$

Para la marginal de x

$$f_X(x) = \int_0^2 \left(x^3 + \frac{3}{8}xy^2 \right) dy = 2x^3 + x \quad ; \quad 0 \le x \le 1$$

$$f_X(x) = \begin{cases} 2x^3 + x \ , & 0 \le x \le 1 \\ 0 \ , & en\ otro\ caso \end{cases}$$

Para la marginal de y

$$f_Y(y) = \int_0^1 \left(x^3 + \frac{3}{8} x y^2 \right) dx = \frac{1}{4} + \frac{3}{16} y^2 \quad ; \quad 0 \le y \le 2$$

$$f_Y(y) = \begin{cases} \frac{1}{4} + \frac{3}{16} y^2 \ , & 0 \le y \le 2 \\ 0 \ , & en\ otro\ caso \end{cases}$$

d) Para las funciones de densidad condicional

$$f_{X|y_0}\left(x \mid Y = y_0\right) = \frac{f_{XY}\left(x , y_0\right)}{f_Y\left(y_0\right)}$$

$$= \frac{x^3 + \frac{3}{8} x y_0^2}{\frac{1}{4} + \frac{3}{16} y_0^2} = \frac{2x(8x^2 + 3y_0^2)}{4 + 3y_0^2}$$

$$f_{X|y_0}\left(x \mid Y = y_0\right) = \begin{cases} \dfrac{2x(8x^2 + 3y_0^2)}{4 + 3y_0^2} & ; \quad 0 \le x \le 1 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

y siempre que $0 \le y_0 \le 2$

Por otro lado

$$f_{Y|x_0}\left(y \mid X = x_0\right) = \frac{f_{XY}\left(x_0 , y\right)}{f_X\left(x_0\right)}$$

$$= \frac{x_0^3 + \frac{3}{8} x_0 y^2}{2x_0^3 + x_0} = \frac{8x_0^2 + 3y^2}{8(2x_0^2 + 1)}$$

$$f_{Y|x_0}\left(y \mid X = x_0\right) = \begin{cases} \dfrac{8x_0^2 + 3y^2}{8(2x_0^2+1)} & ; \quad 0 \le y \le 2 \\ 0 & ; \quad en \;\; otro \;\; caso \end{cases}$$

y siempre que $0 \le x_0 \le 1$

S)))))))))))))))))))))))))))))))))))Q

FUNCIÓN DE DISTRIBUCIÓN CONJUNTA

Proporciona el comportamiento probabilístico acumulado conjunto de una serie de variables aleatorias.

Definición 3.4

Si X y Y son dos variables aleatorias conjuntas se define $F_{XY}(x, y)$ como

$F_{XY}(x, y) = P\;(X \le x,\; Y \le y)$.

Si X y Y son dos variables aleatorias conjuntas discretas o continuas, entonces:

$F_{XY}(x, y) = \displaystyle\sum_{u=-\infty}^{x} \sum_{v=-\infty}^{y} f_{XY}(u, v)$ para vv.aa. discretas

$F_{XY}(x, y) = \displaystyle\int_{u=-\infty}^{u=x} \int_{v=-\infty}^{v=y} f_{XY}(u, v)\; dv\; du$ para vv.aa. continuas

Y $F_{XY}(x, y)$ tiene las siguientes propiedades:

Propiedades de la Función de distribución Conjunta

1) F_{XY} es una función no decreciente

2) $F_{XY}(-\infty, y) = 0$

$F_{XY}(x, -\infty) = 0$

$F_{XY}(\infty, \infty) = 1$

3) $F_{XY}(\infty, y) = F_Y(y)$

$F_{XY}(x, \infty) = F_X(x)$

4) $P\left(x_1 < X \le x_2,\, y_1 < Y \le y_2\right) = F_{XY}\left(x_2, y_2\right) - F_{XY}\left(x_1, y_2\right)$
$\qquad - F_{XY}\left(x_2, y_1\right) + F_{XY}\left(x_1, y_1\right)$

5) Para vv.aa. continuas, si F_{XY} tiene derivadas parciales de orden superior a dos.

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial y\; \partial x}$$

))

S))))))))))))))))))))))))))))))))))Q

Ejemplo 3.7

Utilizando la función de probabilidad conjunta del ejemplo 3.3, obtener su función de distribución conjunta.

Resolución

De la definición de función de distribución conjunta se tiene:

$F_{XY}(x, y)$		$x=0$	$x=1$	$x=2$
y	0	0.1	0.3	0.5
	1	0.14	0.42	0.7
	2	0.2	0.6	1

S))))))))))))))))))))))))))))))))))Q

Obsérvese que para construir la tabla de una función de distribución conjunta discreta sólo se requiere sumar todos los valores por arriba y hacia la izquierda (incluyendo valores diagonales) de la posición que se desea obtener de la tabla de la función de la probabilidad conjunta.

INDEPENDENCIA DE VARIABLES ALEATORIAS CONJUNTAS

Después de estudiar las funciones de probabilidad y de densidad marginales, se observa que el conocimiento (o la información) sobre el valor de una de las variables aleatorias puede modificar el comportamiento probabilístico de la otra variable. Esto sólo sucede si las variables conjuntas son dependientes.

Definición 3.5

Si X y Y son variables aleatorias conjuntas, se dice que son independientes si y sólo si

$$f_{XY}(x, y) = f_X(x)\ f_Y(y)$$

para todo x y y.

En general, para n variables aleatorias conjuntas $X_1, X_2, \ldots X_n$, con función de probabilidad o de densidad conjunta $f_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n)$ y marginales $f_{X_i}(x_i)$ para $i = 1, 2, \ldots, n$, son independientes si y sólo si

$$f_{X_1 X_2 \ldots X_n}(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1)\ f_{X_2}(x_2)\ \ldots\ f_{X_n}(x_n)$$

S))))))))))))))))))))))))))))))))))Q

Ejemplo 3.8

Determinar si las vv.aa. conjuntas X y Y con función de densidad conjunta dada por:

$$f_{XY}(x,y) = \begin{cases} \frac{8}{3}x(x+y) & ; \quad 0 \le x \le 1 \ , \ 0 \le y \le x \\ 0 & ; \quad en \ otro \ caso \end{cases}$$

son independientes o no. Justificar su respuesta.

Resolución
Para determinar si las vv. aa. son independientes o no, deben obtenerse las marginales primero.

Para la marginal de x

$$f_X(x) = \int_0^x \frac{8}{3}x(x+y)\,dy$$

$$f_X(x) = \frac{8}{3}\left(x^3 + \frac{x^3}{2} \right) = 4x^3 \quad ; \quad 0 \le x \le 1$$

De donde

$$f_X(x) = \begin{cases} 4x^3 & ; \quad 0 \le x \le 1 \\ 0 & ; \quad en \ otro \ caso \end{cases}$$

Para la marginal de y

$$f_Y(y) = \int_y^1 \frac{8}{3}x(x+y)\,dx$$

$$f_Y(y) = \frac{8}{3}\left[\frac{x^3}{3} + \frac{x^2 y}{2} \right]_y^1 = \frac{8}{9} + \frac{4}{3}y - \frac{20y^3}{9} \quad ; \quad 0 \le y \le 1$$

De donde

$$f_Y(y) = \begin{cases} \frac{8}{9} + \frac{4}{3}y - \frac{20y^3}{9} & ; \quad 0 \le y \le 1 \\ 0 & ; \quad en \ otro \ caso \end{cases}$$

Y puesto que

$$f_X(x)\,f_Y(y) = \frac{32}{9}x^3 + \frac{16}{3}x^3 y - \frac{80}{9}x^3 y^3 \ne f_{XY}(x,y)$$

se concluye que las variables X y Y no son independientes (son dependientes).

S))))))))))))))))))))))))))))))))))))Q

CARACTERÍSTICAS NUMÉRICAS DE LAS VARIABLES ALEATORIAS CONJUNTAS

Las características numéricas de las variables aleatorias conjuntas son valores que describen el comportamiento probabilístico conjunto.

MEDIDAS DE TENDENCIA CENTRAL: VALOR ESPERADO Y CURVA DE REGRESIÓN

Cuando se estudian variables aleatorias conjuntas, el valor esperado sigue siendo de gran importancia, pero como es de suponerse, se pueden generar infinidad de valores esperados a partir de cualquier función de las variables aleatorias.

Definición 3.6

Si X, Y son variables aleatorias conjuntas con función de probabilidad o de densidad conjunta $f_{XY}(x, y)$ y si $g(x, y)$ es una función de dichas variables aleatorias. Entonces el valor esperado de $g(x, y)$ se define como:

$$\mathrm{E}(g(X, Y)) = \begin{cases} \displaystyle\sum_{\forall R_x} \sum_{\forall R_y} g(x, y)\, f_{XY}(x, y) & \textit{caso discreto} \\ \displaystyle\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y)\, f_{XY}(x, y)\, dx\, dy & \textit{caso continuo} \end{cases}$$

PROPIEDADES DEL VALOR ESPERADO

a) $\mathrm{E}(c) = c$

b) $\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$

c) Si X, Y son variables aleatorias independientes entonces

$$\mathrm{E}\left(g_1(X)\, g_2(Y)\right) = \mathrm{E}\left(g_1(X)\right) \mathrm{E}\left(g_2(Y)\right)$$

S))))))))))))))))))))))))))))))))))Q

Ejemplo 3.9

En un proceso de extracción de cierto mineral se han identificado varios tipos de impurezas. En una muestra específica de este mineral, X denota la proporción de impurezas en la muestra y Y la proporción de la impureza tipo I entre todas las impurezas encontradas. Supóngase que se puede elaborar un modelo de la distribución conjunta de X y Y mediante la función de densidad de probabilidad siguiente:

$$f_{XY}(x, y) = \begin{cases} 2 - 2x & ; \quad 0 \le x \le 1 \quad , \quad 0 \le y \le 1 \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

Encontrar el valor esperado de la proporción de impurezas tipo I.

Resolución

Puesto que X es la proporción de impurezas en la muestra y Y la proporción de impurezas de tipo I en relación al total de las impurezas muestrales, tenemos que XY es la proporción de impurezas tipo I en la muestra entera. Es decir, se debe obtener $E(XY)$

$$E(XY) = \int_0^1 \int_0^1 2xy(1-x)\,dy\,dx = 2\int_0^1 x(1-x)\left(\frac{1}{2}\right)dx$$

$$= \int_0^1 \left(x - x^2\right)dx = \left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^1 = \frac{1}{2} - \frac{1}{3} = \frac{1}{6}$$

La muestra tendrá $\frac{1}{6}$ de impureza tipo I.

))))))))))))))))))))))))))))))))))))Q

Es importante observar que si X, Y son 2 variables aleatorias conjuntas, aún cuando se imponga la condición $X = x_0$ a la variable aleatoria X, la otra variable aleatoria conjunta Y sigue siendo una variable aleatoria, y como tal posee un comportamiento probabilístico propio, descrito por $f_{Y|x_0}\left(Y \mid X = x_0\right)$, y por lo tanto también se pueden determinar sus características numéricas.

Esto mismo ocurre con la condición $\left(X \mid Y = y_0\right)$.

Así, si X y Y son dos variables aleatorias conjuntas y se conoce con certeza el valor que tomará X, el mejor pronóstico que se puede realizar del valor que tomará Y es el valor esperado, calculado desde luego a partir de la distribución condicional correspondiente. Por lo tanto, si se sabe que $X = x_0$, el valor pronosticado para Y que se denota por $\hat{y}_0$ es

$$\hat{y}_0 = \mathrm{E}\left[\,Y \mid X = x_0\,\right] = \mu_{Y|x_0}$$

Definición 3.7

Si X, Y son 2 variables aleatorias conjuntas, se define el valor esperado condicional como:

$$\hat{y}_0 = \mu_{Y|x_0} = \mathrm{E}\left[\,Y \mid X = x_0\,\right] = \begin{cases} \displaystyle\int_{-\infty}^{\infty} y\, f_{Y|x_0}\left(y \mid x_0\right) dy & si\ \ Y\ es\ continua \\ \displaystyle\sum_{\forall R_Y} y\, f_{Y|x_0}\left(y \mid x_0\right) & si\ \ Y\ es\ discreta \end{cases}$$

$$\hat{x}_0 = \mu_{X|y_0} = \mathrm{E}\left[\,X \mid Y = y_0\,\right] = \begin{cases} \displaystyle\int_{-\infty}^{\infty} x\, f_{X|y_0}\left(x \mid y_0\right) dx & si\ \ X\ es\ continua \\ \displaystyle\sum_{\forall R_X} x\, f_{X|y_0}\left(x \mid y_0\right) & si\ \ X\ es\ discreta \end{cases}$$

))

La ***curva de regresión*** de Y dado x, es el lugar geométrico de los valores esperados de las distribuciones condicionales de Y. Su ecuación es: $\hat{y} = \mu_{Y|x}$

La ***curva de regresión*** de X dado y, es el lugar geométrico de los valores esperados de las distribuciones condicionales de X. Su ecuación es: $\hat{x} = \mu_{X|y}$

El concepto de curva de regresión tiene su principal utilidad en los pronósticos, puesto que para cada x proporciona el valor esperado (o pronóstico) de Y, o viceversa. En un curso de Estadística, puede estudiarse un caso particular de la curva de regresión, llamado regresión por mínimos cuadrados.

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 3.10
Sea la función de densidad conjunta

$$f_{XY}(x,y) = \begin{cases} x^3 + \dfrac{3}{8}xy^2 & ; \quad 0 \le x \le 1 \ , \ 0 \le y \le 2 \\ 0 & ; \quad en \ otro \ caso \end{cases}$$

Obtener las esperanzas condicionales para cualquier valor (curvas de regresión).

Resolución
De un ejemplo anterior se sabe que para la función de densidad conjunta, las condicionales son:

$$f_{X|y_0}\left(X \mid Y = y_0\right) = \begin{cases} \dfrac{2x(8x^2 + 3y_0^2)}{4 + 3y_0^2} & ; \quad 0 \le x \le 1 \\ 0 & ; \quad en \ otro \ caso \end{cases}$$

siempre que $0 \le y_0 \le 2$

$$f_{Y|x_0}\left(Y \mid X = x_0\right) = \begin{cases} \dfrac{8x_0^2 + 3y^2}{8(2x_0^2 + 1)} & ; \quad 0 \le y \le 2 \\ 0 & ; \quad en \ otro \ caso \end{cases}$$

siempre que $0 \le x_0 \le 1$

Y las esperanzas condicionales están dadas por:

$$\hat{y}_0 = \mu_{Y \mid x_0} = \mathrm{E}\left[\, Y \mid X = x_0 \,\right] = \int_{-\infty}^{\infty} y \, f_{Y \mid x_0}\left(y \mid x_0\right) \, dy$$

$$\hat{x}_0 = \mu_{X \mid y_0} = \mathrm{E}\left[\, X \mid Y = y_0 \,\right] = \int_{-\infty}^{\infty} x \, f_{x \mid y_0}\left(X \mid y_0\right) \, dx$$

Por lo que

$$\hat{y}_0 = \mu_{Y \mid x_0} = \int_0^2 \frac{1}{8} \, \frac{y\,(8\,x_0^2 + 3\,y^2)}{2\,x_0^2 + 1} \; dy$$

$$\hat{y}_0 = \mu_{Y \mid x_0} = \frac{1}{2} \, \frac{3 + 4\,x_0^2}{2\,x_0^2 + 1} \quad \text{siempre que} \quad 0 \le x_0 \le 1$$

$$\hat{x}_0 = \mu_{X \mid y_0} = \int_0^1 2 \, \frac{x^2\,(8\,x^2 + 3\,y_0^2)}{4 + 3\,y_0^2} \; dx$$

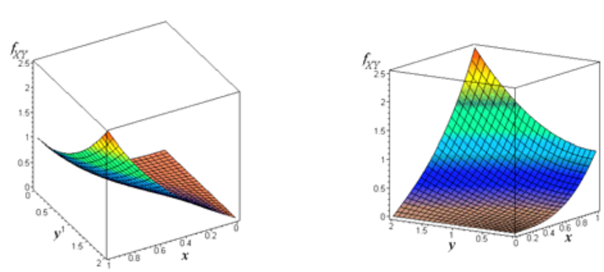
$$\hat{x}_0 = \mu_{X \mid y_0} = \frac{2}{5} \, \frac{8 + 5\,y_0^2}{4 + 3\,y_0^2} \quad \text{siempre que} \quad 0 \le y_0 \le 2$$

S)))))))))))))))))))))))))))))))))))Q

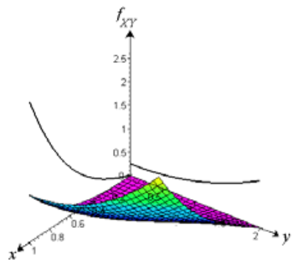
Las funciones marginales y las curvas de regresión pueden interpretarse con mayor facilidad a través de sus gráficas.

Función de densidad conjunta $f_{XY}(x, y) = \begin{cases} x^3 + \dfrac{3}{8} x y^2 & ; \quad 0 \le x \le 1 \;,\; 0 \le y \le 2 \\ 0 & ; \quad en \;\; otro \;\; caso \end{cases}$

Su gráfica desde dos perspectivas es:



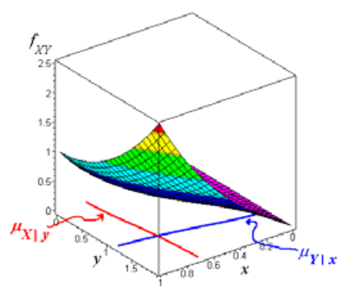
Las funciones condicionales dibujadas junto a la función de densidad son:



Debe observarse que las funciones condicionales se dibujan en el plano (y cada una en su plano); sin embargo, aquí se dibujan junto con la función de densidad conjunta para que se observe que tienen la misma tendencia. Es decir, para cada valor y_0

la condicional $f_{X\,|\,y_0}(x\,|\,y_0)$ acumula la probabilidad de la función conjunta $f_{XY}(x\,,y)$ para el valor constante y_0 y luego la refiere a la unidad.

Las esperanzas condicionales son:



Que igualmente deben dibujarse en el plano. La función $\hat{y} = \mathrm{E}(Y\,|\,x)$ debe dibujarse en el plano xy, mientras que $\hat{x} = \mathrm{E}(X\,|\,y)$ debe dibujarse en el plano yx.

Por otro lado, un teorema de gran aplicación cuando se están estudiando variables aleatorias independientes es el siguiente.

Teorema 3.1

Sean $X_1\,,\,X_2\,,\,.\,.\,.\,.\,,\,X_n$,variables aleatorias independientes cada una con función generadora de momentos $M_{X_1}(\theta)\,,\,M_{X_2}(\theta)\,,\,.\,.\,.\,.\,,\,M_{X_n}(\theta)$, respectivamente.

Si $Y = X_1 + X_2 + \,.\,.\,.\, + X_n$, entonces:

$$M_Y(\theta) = M_{X_1}(\theta)\;M_{X_2}(\theta)\;.\,.\,.\;M_{X_n}(\theta)$$

Demostración

Utilizando la definición de la función generadora de momentos se tiene:

$$\begin{aligned} M_Y(\theta) &= E\left(e^{\theta Y}\right) \\ &= E\left(e^{\theta\,(X_1 + X_2 + \,.\,.\,.\, + X_n)}\right) \\ &= E\left(e^{\theta X_1}\, e^{\theta X_2}\, .\,.\,.\, e^{\theta X_n}\right) \end{aligned}$$

Por independencia

$$\begin{aligned} M_Y(\theta) &= E\left(e^{\theta X_1}\right) E\left(e^{\theta X_2}\right) .\,.\,.\, E\left(e^{\theta X_n}\right) \\ &= M_{X_1}(\theta)\, M_{X_2}(\theta)\, .\,.\,.\, M_{X_n}(\theta) \end{aligned}$$

□

■

El teorema anterior indica que si sumamos n variables aleatorias independientes, entonces la función generadora de momentos de la suma es la multiplicación de las n funciones generadoras de momentos. Este teorema es de utilidad cuando se definen variables aleatorias a partir de la suma de otras variables aleatorias, como se estudiará en los temas 4 y 5 y en el curso de Estadística.

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 3.11

Sean X y Y dos variables aleatorias conjuntas independientes, con funciones de densidad y generadores de momentos:

$$f_X(x) = \begin{cases} 2\,e^{-2x} & ; \quad x > 0 \\ 0 & ; \quad en\ otro\ caso \end{cases} \qquad , \qquad M_X(\theta) = \frac{2}{2 - \theta}$$

$$f_Y(y) = \begin{cases} e^{-y} & ; \quad y > 0 \\ 0 & ; \quad en\ otro\ caso \end{cases} \qquad , \qquad M_Y(\theta) = \frac{1}{1 - \theta}$$

Obtener la función generadora de momentos de la variable aleatoria T definida mediante:

$$T = X + Y$$

Resolución

Por independencia

$$\begin{aligned} M_T(\theta) &= M_X(\theta)\ M_Y(\theta) \\ &= \left(\frac{2}{2 - \theta}\right)\left(\frac{1}{1 - \theta}\right) \end{aligned}$$

S))))))))))))))))))))))))))))))))))))Q

MEDIDAS DE DISPERSIÓN DE VARIABLES ALEATORIAS CONJUNTAS: COVARIANCIA Y CORRELACIÓN

COVARIANCIA

La covariancia es una medida de dispersión que indica, en promedio, qué tanto se alejan conjuntamente los valores de sus medias respectivas, *i.e.* qué tanto varían en conjunto (qué tanto covarían).

Definición 3.8

Si X y Y son 2 variables aleatorias conjuntas, entonces se define la covariancia $Cov\left(X, Y\right)$ como:

$$Cov\left(X, Y\right) = \sigma_{XY} = \mathrm{E}\left[\left(X - \mu_X\right)\left(Y - \mu_Y\right)\right]$$

Aplicando las propiedades del valor esperado, se tiene que la covariancia se puede calcular mediante:

$$Cov\left(X, Y\right) = \mathrm{E}\left[\left(X - \mu_X\right)\left(Y - \mu_Y\right)\right] = \mathrm{E}\left(XY\right) - \mu_X \mu_Y$$

Teorema 3.2

Si X y Y son dos variables aleatorias conjuntas independientes, entonces:

$$Cov\left(X, Y\right) = 0$$

El valor esperado de la covariancia depende de la escala en la que están X y Y. Para referir la relación entre X y Y a una escala fija se debe utilizar el coeficiente de correlación.

Definición 3.9

El coeficiente de correlación, denotado por ρ es:

$$\rho = \frac{Cov\left(X, Y\right)}{\sqrt{\mathrm{Var}\left(X\right)}\sqrt{\mathrm{Var}\left(Y\right)}} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

El coeficiente de correlación es una cantidad adimensional que mide la asociación lineal entre las dos variables aleatorias, el valor de ρ está contenido en el intervalo $\left[-1, 1\right]$.

El coeficiente de correlación tiene las siguientes propiedades.

1) Si X, Y son variables aleatorias independientes, entonces $\rho = 0$

2) Para cualesquiera variables aleatorias conjuntas $-1 \leq \rho \leq 1$

3) $\rho = \pm 1$, si y sólo si $Y = mX + b$ para m y b constantes.

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 3.12

Dada la función de densidad conjunta

$$f_{XY}(x\,,\,y) = \begin{cases} 4\,xye^{-\left(x^2+y^2\right)} & ; \quad 0 \le x < \infty \,,\, 0 \le y < \infty \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

Obtener:

a) Las densidades marginales de X y Y

b) Las funciones de densidad condicional de X y Y

c) La expresión de las esperanzas condicionales de X y Y

d) El coeficiente de correlación.

Resolución

a) $$f_X(x) = \int_0^{\infty} 4\,xye^{-\left(x^2+y^2\right)}\,dy = -2\,xe^{-\left(x^2+y^2\right)}\Big]_{y=0}^{y=\infty}$$

$$= 2\,xe^{-x^2}$$

$$f_X(x) = \begin{cases} 2\,xe^{-x^2} & ; \quad 0 \le x < \infty \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

De manera análoga, por simetría

$$f_Y(y) = \begin{cases} 2\,ye^{-y^2} & : \quad 0 \le y < \infty \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

b) $$f_{X\,|\,y_0}\left(x \mid y_0\right) = \frac{f_{XY}(x\,,\,y_0)}{f_Y(y_0)} = \frac{4\,xy_0e^{-\left(x^2+y^2\right)}}{2\,y_0e^{-y_0^2}}$$

$$= 2\,xe^{-x^2}$$

$$f_{X\,|\,y_0}\left(x \mid y_0\right) = \begin{cases} 2\,xe^{-x^2} & ; \quad x \ge 0 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

de forma análoga, por simetría

$$f_{Y\,|\,x_0}\left(y \mid x_0\right) = \begin{cases} 2\,y e^{-y^2} & y \geq 0 \\ 0 & en\ otro\ caso \end{cases}$$

c) $\hat{x}_0 = \mu_{X\,|\,y_0} = \mathrm{E}\left[\, X \mid Y = y_0 \,\right] = \displaystyle\int_0^{\infty} x \left(2\,x e^{-x^2}\right) dx$

$$= \int_0^{\infty} 2\,x^2 e^{-x^2}\, dx$$

Sea $u = x\,,$ $du = dx$

$dv = 2\,x e^{-x^2}\, dx\,,$ $v = -\,e^{-x^2}$

entonces:

$$\hat{x}_0 = -\,x e^{-x^2}\Big]_0^{\infty} + \int_0^{\infty} e^{-x^2}\, dx = 0 + \frac{\sqrt{\pi}}{2}$$

$\hat{x}_0 = \dfrac{\sqrt{\pi}}{2}$, y d e forma análoga $\hat{y}_0 = \dfrac{\sqrt{\pi}}{2}$

d) Puesto que $f_{XY}(x\,,\,y) = f_X(x)\ f_Y(y)$, entonces las variables aleatorias son independientes y $\mathrm{Cov}\,(\,X,\ Y\,) = 0$, por lo que $\rho = 0$.

S))))))))))))))))))))))))))))))))))Q

Definición 3.10

El coeficiente de determinación lineal, denotado por ρ^2 es:

$$\rho^2 = \frac{\left[\,Cov\,(\,X,\ Y\,)\right]^2}{\mathrm{Var}\,(X)\ \ \mathrm{Var}\,(Y)} = \frac{\sigma_{XY}^2}{\sigma_X^2\,\sigma_Y^2}$$

El coeficiente de determinación, proporciona el grado de explicación de una variable (generalmente y) respecto a la otra (generalmente x), es decir, qué tanto se explica y, conociendo x.

Teorema 3.3

Sean $X_1, X_2, \ldots, X_n$ variables aleatorias con parámetros

$$E(X_i) = \mu_i \qquad i = 1, 2, \ldots, n$$

$$\text{Var}(X_i) = \sigma_i^2 \qquad i = 1, 2, \ldots, n$$

y $a_1, a_2, \ldots, a_n$ son constantes. Entonces la variable aleatoria Y con la combinación lineal siguiente

$$Y = a_1 X_1 + a_2 X_2 + \ldots + a_n X_n = \sum_{i=1}^{n} a_i X_i$$

tiene como parámetros

$$E(Y) = \sum_{i=1}^{n} a_i \mu_i$$

$$\text{Var}(Y) = \sum_{i=1}^{n} a_i^2 \text{ Var}(Y_i) + 2 \sum \sum_{i<j} a_i a_j \text{Cov}(X_i, X_j)$$

La obtención de la variancia de una combinación lineal se facilita si se utiliza la forma matricial

$$\text{Var}(Y) = \left[a_1, a_2, \ldots, a_n\right] \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \ldots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \ldots & \sigma_{2n} \\ . & & & & \\ . & & & & \\ . & & & & \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \ldots & \sigma_n^2 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_n \end{bmatrix}$$

donde la matriz

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \ldots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \ldots & \sigma_{2n} \\ . & & & & \\ . & & & & \\ . & & & & \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \ldots & \sigma_n^2 \end{bmatrix}$$

recibe el nombre de ***matriz de varianza-covarianza***, y tiene diversas aplicaciones en ingeniería.

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 3.13

Supóngase que Y_1, Y_2 y Y_3 son variables aleatorias con

$E(Y_1) = 2$ $\quad E(Y_2) = -1$ $\quad E(Y_3) = 4$

$\mathrm{Var}(Y_1) = 4$ $\quad \mathrm{Var}(Y_2) = 6$ $\quad \mathrm{Var}(Y_3) = 8$

$\mathrm{Cov}(Y_1, Y_2) = 1$ $\quad \mathrm{Cov}(Y_1, Y_3) = -1$ $\quad \mathrm{Cov}(Y_2, Y_3) = 0$

Obtener $\mathrm{E}(3Y_1 + 4Y_2 - 6Y_3)$ y $\mathrm{Var}(3Y_1 + 4Y_2 - 6Y_3)$

Resolución

$$E(3Y_1 + 4Y_2 - 6Y_3) = 3\,\mathrm{E}(Y_1) + 4\,\mathrm{E}(Y_2) - 6\,\mathrm{E}(Y_3)$$

$$E(3Y_1 + 4Y_2 - 6Y_3) = 3(2) + 4(-1) - 6(4) = -22$$

$$\begin{aligned}\mathrm{Var}(3Y_1 + 4Y_2 - 6Y_3) &= 9\,\mathrm{Var}(Y_1) + 16\,\mathrm{Var}(Y_2) + 36\,\mathrm{Var}(Y_3) + \\ &+ 2(3)(4)\,\mathrm{Cov}(Y_1, Y_2) + \\ &+ 2(3)(-6)\,\mathrm{Cov}(Y_1, Y_3) + \\ &+ 2(4)(-6)\,\mathrm{Cov}(Y_2, Y_3)\end{aligned}$$

$$\begin{aligned}\mathrm{Var}(3Y_1 + 4Y_2 - 6Y_3) &= 9(4) + 16(6) + 36(8) + 2(3)(4)(1) \\ &+ 2(3)(-6)(-1) + 2(4)(-6)(0)\end{aligned}$$

$$\mathrm{Var}(3Y_1 + 4Y_2 - 6Y_3) = 480$$

S))))))))))))))))))))))))))))))))))))Q

AUTOEXAMEN TEMA III

1.- El valor de la constante c, para que la función

$$f_{XY}(x,y) = \begin{cases} c \;; & si \quad 0 \le x \le 0.25 \;,\; 0 \le y \le 2000 \\ 0 \;; & en \; otro \; caso \end{cases}$$

sea una función de densidad conjunta es: ☐

A) $\dfrac{1}{2000}$ B) $\dfrac{1}{500}$ C) 1 D) 4

E) Ninguna de las anteriores.

2.- Sean X y Y dos variables aleatorias conjuntas con función de densidad

$$f_{XY}(x,y) = \begin{cases} 6x^2y \;; & 0 \le x \le 1 \;,\; 0 \le y \le 1 \\ 0 \;; & en \; otro \; caso \end{cases}$$

Entonces $P(X \le 0.5 \;,\; Y \le 0.5)$ es: ☐

A) 0.03125 B) 0.75 C) 0.25 D) 0.6562

E) Ninguna de las anteriores.

3.- La ganancia de un almacén de la central de abastos depende del tiempo de llegada del proveedor (que llega entre 1 y 4 a.m.) y la calidad del producto (que convencionalmente varía de 1 a 2). Llamemos X al tiempo de llegada y Y a la calidad.

Ambas variables tienen una función de densidad conjunta

$$f_{XY}(x,y) = \begin{cases} \dfrac{3x - y}{18} \;; & 1 < x < 4 \;,\; 1 < y < 2 \\ 0 \;; & en \; otro \; caso \end{cases}$$

Entonces la función de densidad de probabilidad del tiempo de llegada es: ☐

A) $\dfrac{5}{4} - \dfrac{1}{6}y$ B) $\dfrac{1}{6}x - \dfrac{1}{12}$ C) $\dfrac{1}{2}x - \dfrac{5}{12}$ D) $\dfrac{1}{4} - \dfrac{1}{18}y$ E) 1

4.-- Si X, Y son variables aleatorias conjuntas con función de densidad

$$f_{XY}(x,y) = \begin{cases} 8xy & 0 < x < y \;,\; 0 < y < 1 \\ 0 & en \; otro \; caso \end{cases}$$

entonces la densidad condicional de X dado Y cuando $0 < x < y$, $0 < y < 1$ es: ☐

A) $2x$ B) $\dfrac{2x}{y^2}$ C) $2y$ D) $\dfrac{2x}{1 - y^2}$

E) Ninguna de las anteriores

5.- Sean X y Y dos variables aleatorias conjuntas con función de probabilidad

$f_{XY}(x,y)$		y	
		1	2
x	0	0.1	0.3
	1	0.4	0.2

Entonces, el valor esperado de XY es: $\boxed{}$

A) 0.6 B) 0.8 C) 0.9 D) 1.5
E) Ninguno de los anteriores.

6.- Dos líneas de ensamble producen cierto número de piezas en un día según la distribución de probabilidad conjunta siguiente

$f_{XY}(x,y)$		x		
		0	1	2
	0	0.2	0.1	0.05
	1	0.1	0.15	0.1
y	2	0.05	0.12	0.13

Entonces el número esperado de piezas que producen las dos líneas en un día cualquiera es: $\boxed{}$

A) 2 B) 0.93 C) 1.86 D) 1.88
E) Ninguna de las anteriores.

7.- Si dos variables aleatorias tienen la densidad conjunta

$$f_{XY}(x,y) = \begin{cases} 4xy & 0<x<1 \ , \ 0<y<1 \\ 0 & \textit{en otro caso} \end{cases}$$

entonces su covariancia, $\text{Cov}(X,Y)$, es: $\boxed{}$

A) $\frac{2}{3}$ B) $\frac{4}{9}$ C) 0 D) 1

E) Ninguna de la anteriores.

8.- Si X y Y son variables aleatorias conjuntas con $\mu_x = 10$, $\mu_y = 15$, $\sigma_x^2 = 3$, $\sigma_y^2 = 4$, $\text{Cov}(X,Y) = 2$ entonces $\text{Var}(2X - Y)$ es: $\boxed{}$

A) 2 B) 16 C) 0 D) 24 E) 8

9.- Las variables aleatorias X y Y representan los rendimientos de las acciones de dos empresas. Sus parámetros son: $\mu_X = 10$, $\sigma_X = 15$, $\mu_Y = 15$, $\sigma_Y = 20$, y $\rho_{XY} = -0.1$. La media y la desviación estándar de la combinación lineal $R = 0.4X + 0.6Y$ son respectivamente: $\boxed{}$

A) 12.86, 18 B) 13, 18 C)13, 13.41
D) 13, 12.86 E) Ninguna de las anteriores.

10.- Si las variables aleatorias conjuntas X y Y tienen la función de probabilidad conjunta

F_{XY}		x	
		0	1
y	1	0.1	0.3
	2	0.4	0.2

Entonces el coeficiente de correlación entre X y Y es: $\boxed{}$

A) -0.408 B) 0.7 C) -0.1
D) -1.66 E) Ninguna de las anteriores.

S)))))))))))))))))))))))))))))))))))Q

BIBLIOGRAFÍA

Hines, William W. y Montgomery, Douglas C. - Probabilidad y Estadística para ingeniería, cuarta edición..- CECSA.- México, 2005.

Milton, Susan J. Y Arnold, Jesse C.- Probabilidad y Estadística para con aplicaciones para ingeniería y ciencias computacionales, cuarta edición.- McGraw-Hill.- México, 2004.

Devore, Jay L.- Probabilidad y Estadística para ingeniería y ciencias, séptima edición.- Cengage Learning.- México, 2008.

Mendenhall, William III. et al.- Introducción a la Probabilidad y Estadística.- Décimo cuarta edición.- Cengage Learning.- México 2015.

Wackerly Dennis D.- Mendenhall, William, *et al*.- Estadística Matemática con Aplicaciones, sexta edición.- Editorial Thomson.- México, 2002.

Walpole, Ronald E., *et al*.- Probability and Statistics for Engineers and Scientists.- Pearson.- USA, 2007.

Montgomery, Douglas C. y Runger, George C.-Probabilidad y Estadística aplicadas a la Ingeniería, segunda edición.- Limusa-Wiley.- México, 2002.

Scheaffer, Richard L. y McClave, James T.- Probabilidad y Estadística para Ingeniería.- Grupo Editorial Iberoamérica.- México, 1993.

Canavos, George C.- Probabilidad y Estadística Aplicaciones y Métodos.- McGraw-Hill.- México, 1988.

Meyer, Paul L.- Probabilidad y Aplicaciones Estadísticas.- Addison Wesley Iberoamericana.- México, 1992.

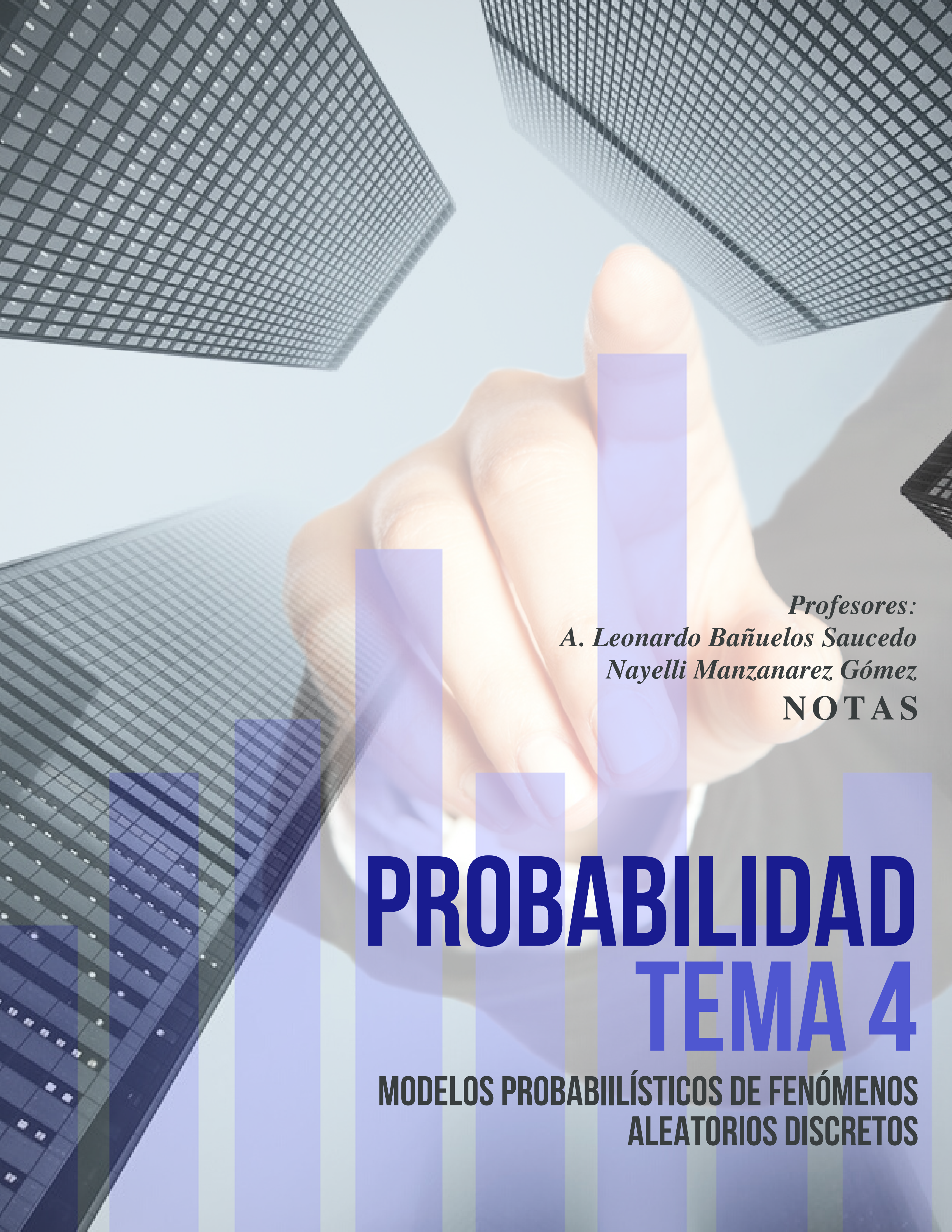
Spiegel, Murray R. et al.- Probabilidad y Estadística, cuarta edición.- Mc Graw-Hill.-México 2013.

Borras García, Hugo E., *et al*.- Apuntes de Probabilidad y Estadística.-Facultad de Ingeniería.- México, 1985.

Rosenkrantz, Walter A.- Introduction to Probability and Statistics for Scientists and Engineers.- McGraw-Hill.- EE.UU., 1997.

Ziemer, Rodger E.- Elements of Engineering Probability & Statistics.- Prentice Hall.- USA 1997.

Lipschutz, Seymour.- Teoría de Conjuntos y Temas Afines.- McGraw-Hill.- México, 1991.



Profesores:

A. Leonardo Bañuelos Saucedo

Nayelli Manzanarez Gómez

NOTAS

PROBABILIDAD

TEMA 4

MODELOS PROBABIILÍSTICOS DE FENÓMENOS ALEATORIOS DISCRETOS

TEMA IV
MODELOS PROBABILÍSTICOS DE FENÓMENOS ALEATORIOS DISCRETOS

INTRODUCCIÓN

Las variables aleatorias sirven para modelar problemas con incertidumbre en los cuales la pregunta que se desea contestar está claramente delimitada. En muchas ocasiones los problemas de interés, coinciden en la forma en la que se define la variable aleatoria, y con ello su función de probabilidad o de densidad y estas coincidencias permiten la formulación de modelos. Los modelos o distribuciones de las variables aleatorias pueden clasificarse en continuos y discretos. En este tema se estudiarán los modelos discretos de una variable aleatoria.

DISTRIBUCIONES DISCRETAS

Una gran cantidad de fenómenos pueden modelarse utilizando distribuciones discretas o continuas. Para el caso de las distribuciones discretas considérense los siguientes ejemplos.

- En el lanzamiento de una moneda, ¿cuál es la probabilidad de obtener una águila?
- En el lanzamiento de n monedas, ¿cuál es la probabilidad de obtener x águilas?
- En un metro cuadrado de tela, ¿cuál es la probabilidad de observar x defectos? etc.

Las distribuciones discretas que se estudiaran en este tema son:
Uniforme
Bernoulli
Binomial
Geométrica
de Pascal (Binomial negativa)
Hipergeométrica
de Poisson

DISTRIBUCIÓN DISCRETA UNIFORME

La distribución discreta uniforme es una de las más simples de todas las distribuciones discretas de probabilidad. Es aquella en la cual la v.a. asume cada uno de sus valores con la misma probabilidad.

Definición 4.1

Sea X una variable aleatoria. que toma los valores $x_1, x_2, \ldots, x_k$, con igual probabilidad

$$f_X(x;k) = \frac{1}{k} \quad , \quad x = x_1, x_2, \ldots, x_k$$

entonces la distribución es discreta uniforme con parámetro k.

Se denota $X \sim \textit{uniforme}(k)$.

En la notación $f_X(x;k)$, la variable aleatoria es X, mientras que k es un parámetro del cual depende la distribución[1]. Puede utilizarse la notación indicando el parámetro, o simplemente $f_X(x)$. En las definiciones aquí mostradas se escribirá la función utilizando el parámetro, pero en la solución de problemas se escribirá simplemente $f_X(x)$.

Teorema 4.1

Si X es una variable aleatoria con distribución discreta uniforme, entonces:

$$\mu_X = \text{E}(X) = \frac{\sum_{i=1}^{k} x_i}{k}$$

$$\sigma_X^2 = \text{Var}(X) = \frac{\sum_{i=1}^{k} (x_i - \mu_X)^2}{k}$$

Demostración

De la definición de valor esperado

$$\mu_X = \text{E}(X) = \sum_{i=1}^{k} x_i \; f_X(x_i, k) = \sum_{i=1}^{k} x_i \left(\frac{1}{k} \right)$$

Y de la definición de la variancia

$$\sigma_X^2 = \text{Var}(X) = \sum_{i=1}^{k} (x_i - \mu_X)^2 \; f_X(x_i; k)$$

$$\sigma_X^2 = \sum_{i=1}^{k} (x_i - \mu_X)^2 \left(\frac{1}{k} \right)$$

■

Debe observase la similitud entre las fórmulas de la media y la variancia de una v.a. discreta y las fórmulas utilizadas en la estadística descriptiva para la media y la variancia de un conjunto de datos, o bien, las fórmulas tradicionales para la media y la varianza de un conjunto de datos.

S))))))))))))))))))))))))))))))))))))))Q

Ejemplo 4.1

Se selecciona a un alumno de un grupo de 10 para supervisar cierto trabajo, escogiendo aleatoriamente el nombre de una caja que contiene 10 papeles numerados del 1 al 10.

a) Obtener la fórmula para la distribución de probabilidad de X que representa el número del papel que se saca. ¿Cuál es la probabilidad de que el número que se saque sea menor a 4?

b) Determinar la media y la variancia de X.

Resolución

a) $X \sim uniforme(10)$

[1] En realidad k no es un parámetro para esta distribución. La distribución discreta uniforme no tiene parámetros, pero en este momento es de utilidad introducir la notación $f_X(x; k)$.

$$f_X(x) = \begin{cases} \dfrac{1}{10} & , \quad x = 1,2,3,\ldots,10 \\ 0 & , \quad en\ otro\ caso \end{cases}$$

$$P(X < 4) = P(X \leq 3) = \frac{3}{10}$$

b) $$E(X) = \sum_{x=1}^{10} x\left(\frac{1}{10}\right)$$

$$= \frac{1}{10}\sum_{i=1}^{10} x_i = \frac{1}{10}\left[\ 1+2+3+4+5+6+7+8+9+10\ \right]$$

$$= \frac{1}{10}(55) = 5.5$$

$$Var(X) = \sum_{x=1}^{10}(x_i - \mu)^2\left(\frac{1}{10}\right) = 8.25$$

S)))))))))))))))))))))))))))))))))))))Q

DISTRIBUCIÓN DE BERNOULLI

Es el caso más sencillo para modelar un experimento. Tiene únicamente dos resultados posibles; satisfactorio o insatisfactorio, bueno o malo, alto o bajo, etc. que en general, pueden denominarse éxito (e) o fracaso (f), con probabilidades p y $q = 1 - p$, respectivamente. Recibe el nombre de distribución de Bernoulli en honor de Jakob Bernoulli (1654-1705), destacado matemático suizo perteneciente a un familia distinguida por su aportaciones a la física y matemática.

Definición 4.2

Sea X la variable aleatoria que representa el número de éxitos que se obtienen al realizar un ensayo de Bernoulli, entonces X tiene una distribución de Bernoulli con parámetro p.

$$f_X(x;p) = \begin{cases} 1-p & ; \quad x = 0 \\ p & ; \quad x = 1 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

Se denota $X \sim Bernoulli(p)$

La distribución de Bernoulli también puede escribirse como:

$$f_X(x) = \begin{cases} p^x (1-p)^{1-x} & ; \quad x = 0,1 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

o bien

$$f_X(x) = \begin{cases} p^x \, q^{1-x} & ; \quad x = 0, 1 \\ 0 & ; \quad en \ otro \ caso \end{cases}$$

renombrando $q = 1 - p$.

En forma tabular la distribución de Bernoulli se escribe

x	0	1
$f_X(x)$	q	p

Teorema 4.2

Si X es una variable aleatoria con distribución de Bernoulli, entonces:

$$\mu_X = \mathrm{E}(X) = p$$

$$\sigma_X^2 = \mathrm{Var}(X) = p\,q$$

$$M_X(\theta) = q + e^{\theta} p$$

La demostración de la media y la variancia es inmediata a partir de la definición.

$$\mu_X = \mathrm{E}(X) = \sum_{x=0}^{1} x \, f_X(x) = (0)(1-p) + (1)\,p = p$$

$$\sigma_X^2 = \mathrm{Var}(X) = \mathrm{E}(X^2) - \mu_X^2 = \sum_{x=0}^{1} x^2 \, f_X(x) - (p)^2$$

$$= (0)^2 (1-p) + (1)^2 p - (p)^2 = p(1-p) = p\,q$$

La función generadora también es inmediata a partir de la definición

$$M_X(\theta) = \mathrm{E}\left(e^{\theta X}\right) = e^{\theta(0)} q + e^{\theta(1)} p = q + e^{\theta} p$$

□

La limitante de la distribución Bernoulli es que sólo se tienen dos resultados y el experimento sólo se realiza una vez, por lo que sirve para el lanzamiento de una moneda (que también se puede modelar con distribución uniforme si la moneda es justa), para el lanzamiento de un dado pero cuando interesa la observación de un número en particular contra la no ocurrencia de dicho número, etc.

La principal utilidad de la distribución de Bernoulli se obtiene cuando se generaliza la definición de la variable aleatoria de Bernoulli. Para ello, debe considerarse el proceso de Bernoulli.

PROCESO DE BERNOULLI

Si se repite un ensayo de Bernoulli, se tiene entonces un proceso de Bernoulli.

El proceso de Bernoulli tiene las siguientes propiedades.

1) El experimento consiste de n ensayos de Bernoulli.
2) Los resultados de cada ensayo pueden clasificarse como éxito o fracaso.
3) La probabilidad de éxito p, permanece constante para todos los ensayos.
4) Los ensayos son independientes.

A partir del proceso de Bernoulli se pueden definir otras variables aleatorias, como la binomial, la geométrica y la de Pascal.

DISTRIBUCIÓN BINOMIAL

Empleando el proceso de Bernoulli, es decir, realizando n ensayos, nos puede interesar contabilizar el número total de éxitos que se observan en los n ensayos. Cuando esto nos interesa se define la variable aleatoria Binomial. Puesto que cada ensayo que se realiza es un experimento de Bernoulli, la variable aleatoria que contabiliza los éxitos es la suma de n variables aleatorias independientes de Bernoulli.

Definición 4.3

Sea X la variable aleatoria que representa el número de éxitos que se observan al realizar un proceso de Bernoulli, entonces X recibe el nombre de variable aleatoria binomial, con distribución

$$f_X(x;n,p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & ; \quad x = 0, 1, \ldots, n \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

Se denota: $X \sim Binomial(n,p)$ [2]

La expresión de la distribución binomial es inmediata al observar que se desean x éxitos en n ensayos, por lo que se tiene $n - x$ fracasos; puesto que el orden en el que se obtienen los éxitos y los fracasos no es importante, deben de considerarse todas las posibilidades, por lo que

$$P(X=x) = ppp \ldots qqq + qpp \ldots qqp + \ldots$$
$$P(X=x) = p^x q^{n-x} + p^x q^{n-x} + \ldots$$

[2] Algunos autores denotan a la variable aleatoria binomial como $X \sim b(n,p)$ o $X \sim bin(n,p)$

$$P(X=x) = \binom{n}{x} p^x q^{n-x}$$

El nombre de distribución *binomial*, proviene del hecho de que los términos de probabilidad que se obtienen de la expresión

$$P(X=x) = \binom{n}{x} p^x q^{n-x}$$

corresponden a los términos del desarrollo del binomio $(q + p)^n$.

Por ejemplo, para $n = 5$ se tiene

$$(q+p)^5 = q^5 + 5q^4p + 10q^3p^2 + 10q^2p + 5qp^4 + p^5$$

mientras que al desarrollar la distribución binomial se tiene:

x	0	1	2	3	4	5
$f_X(x)$	q^5	$5q^4p$	$10q^3p^2$	$10q^2p^3$	$5qp^4$	p^5

y claramente se observan los mismos términos

Teorema 4.3

Sea X una variable aleatoria con distribución binomial y parámetros n y p, entonces:

$\mu_X = np$

$\sigma_X^2 = npq$

$M_X(\theta) = (q + e^\theta p)^n$

Demostración

Sea X una variable aleatoria binomial, con parámetros n y p, entonces X es la suma de n variables independientes de Bernoulli con parámetro p:

$$X \sim Binomial(n, p)$$

$$X = \sum_{i=1}^{n} X_i$$

$$X_i \sim Bernoulli(p)$$

De donde, para el valor esperado

$$\mathrm{E}[X] = \mathrm{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{E}[X_i]$$

$$= \sum_{i=1}^{n} p = np$$

Para la varianza

$$\mathrm{Var}[X] = \mathrm{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathrm{Var}[X_i]$$

$$= \sum_{i=1}^{n} pq = npq$$

Y para la función generadora, utilizando el teorema para la suma de variables aleatorias independientes

$$M_{X_i}(\theta) = q + e^{\theta} p$$

$$M_X(\theta) = \prod_{i=1}^{n} (q + e^{\theta} p) = (q + e^{\theta} p)^n$$

□

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 4.3
Supóngase que los motores de un aeroplano a escala operan en forma independiente y tienen una probabilidad de funcionamiento correcto de 0.6. Suponiendo que un aeroplano a escala puede realizar un vuelo adecuado en tanto se mantengan funcionando al menos la mitad de sus motores, determinar qué aeroplano, uno de 4 motores o uno de 2, tiene mayor probabilidad de terminar su vuelo adecuadamente.

Resolución

Sea X la variable aleatoria que representa el número de motores que funcionen en el aeroplano de 2 motores, y sea Y la variable aleatoria que representa el número de motores que funcionen en el aeroplano de 4 motores.

$X \sim Binomial\,(n=2\ ,\ p=0.6)\ ,\ \ Y \sim Binomial\,(n=4\ ,\ p=0.6)$

Entonces, para que los aeroplanos tengan un vuelo adecuado se tiene:

$$P(X \geq 1) = \sum_{x=1}^{2} \binom{2}{x} (0.6)^x\ (0.4)^{2-x} = 0.84$$

$$P(Y \geq 2) = \sum_{x=2}^{4} \binom{4}{x} (0.6)^x\ (0.4)^{4-x} = 0.8208$$

Por lo que el aeroplano con 2 motores tiene una probabilidad ligeramente mayor de terminar su vuelo en forma adecuada.

S))))))))))))))))))))))))))))))))))))Q

DISTRIBUCIÓN GEOMÉTRICA

Definición 4.4

Sea X la variable aleatoria que representa el número de ensayos de Bernoulli que se requieren para observar por primera vez un éxito, entonces X tiene una distribución geométrica con parámetro p.

$$f_X(x) = \begin{cases} q^{x-1}p & ; \ x = 1, 2, \ldots \\ 0 & ; \ en \ otro \ caso \end{cases}$$

Se denota por $X \sim Geométrica(p)$.

La expresión de la distribución geométrica se deduce con facilidad al plantear las probabilidades de una serie de fracasos consecutivos y finalmente un éxito.

$$P(X=1) = P(éxito) = P(e) = p$$
$$P(X=2) = P(fracaso\ éxito) = P(fe) = qp$$
$$P(X=3) = P(ffe) = qqp = q^2 p$$
$$P(X=4) = P(fffe) = qqqp = q^3 p$$

.
.
.

$$P(X=x) = q^{x-1}p$$

La distribución geométrica en efecto es una función de probabilidad. Para probar que la suma de todos los posibles valores de x es 1 se utiliza la *serie geométrica*:

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1-r} \ , \quad |r| < 1$$

por lo que

$$\sum_{k=0}^{\infty} p q^k = \sum_{x=1}^{\infty} p\, q^{x-1}$$

$$\sum_{k=0}^{\infty} p q^k = p \sum_{k=0}^{\infty} q^k$$

$$= p \ \frac{1}{1-q} = p \frac{1}{p} = 1$$

Teorema 4.4

Sea X una variable aleatoria con distribución geométrica con parámetro p entonces

$$\mu_X = \frac{1}{p} \qquad , \qquad \sigma_X^2 = \frac{q}{p^2}$$

y función generadora de momentos $M_X(\theta) = \dfrac{p\,e^{\theta}}{1 - q\,e^{\theta}}$

Demostración de la media

$$\mu_X = \mathrm{E}[X] = \sum_{x=1}^{\infty} x\,p\,q^{x-1} = p\,\frac{d}{dq}\sum_{x=1}^{\infty} q^x$$

$$\mu_X = p\,\frac{d}{dq}\left(\frac{q}{1-q}\right) = \frac{1}{p}$$

■

S))))))))))))))))))))))))))))))))))Q

Ejemplo 4.3

Se supone que el 30% de los aspirantes para cierto trabajo industrial tiene un entrenamiento avanzado en programación computacional. Los aspirantes son entrevistados, uno tras otro, y son seleccionados al azar del conjunto de aspirantes. Determinar la probabilidad de que se encuentre el primer aspirante con un entrenamiento avanzado en programación en la quinta entrevista.

Resolución

Sea X la variable aleatoria que representa el número de entrevistas para obtener al primer aspirante con entrenamiento avanzado.

$X \sim Geométrica\,(p = 0.3)$

$P(X = 5) = (0.7)^{5-1}(0.3) = 0.072$

S))))))))))))))))))))))))))))))))))Q

Al calcular probabilidades con la distribución geométrica, puede ser de gran utilidad su función de distribución acumulada,

$$F_X(x) = 1 - q^{[\![x]\!]} \quad , \quad x \ge 1$$

en donde q es la probabilidad de fracaso y $[\![x]\!]$ es la función parte entera.

DISTRIBUCIÓN DE PASCAL (Distribución Binomial Negativa)

La distribución de Pascal es una generalización de la distribución geométrica, así mismo la distribución Binomial Negativa es una generalización de la de Pascal, por lo que algunos autores utilizan simplemente el nombre de distribución Binomial Negativa. La distribución de Pascal, es nombrada así en honor del matemático francés Blaise Pascal (1623-1662).

La distribución de Pascal modela el número de ensayos de Bernoulli que se requieren para observar el r-ésimo éxito, esto es, es una suma de r variables aleatorias con distribución Geométrica.

Definición 4.5

Sea X una variable aleatoria que representa el número de ensayos de Bernoulli que se requieren para observar el r-ésimo éxito, si en cada uno de los ensayos se tiene una probabilidad de éxito p, entonces X tiene una distribución de Pascal con parámetros r y p.

$$f_X(x) = \begin{cases} \binom{x-1}{r-1} p^r q^{x-r} & ; \quad x = r, r+1, \ldots \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

Se denota por $X \sim Pascal(r, p)$. [3]

Teorema 4.5

Sea X una variable aleatoria con distribución de Pascal con parámetros r y p, entonces

$$\mu_X = \frac{r}{p} \qquad \sigma_X^2 = \frac{r\,q}{p^2}$$

y función generadora de momentos $M_X(\theta) = \left(\dfrac{p\,e^\theta}{1 - q\,e^\theta} \right)^r$

La media y la variancia se desprenden directamente de la suma de r variables aleatorias independientes con distribución geométrica.

Sea

$$X = X_1 + X_2 + X_3 + \ldots + X_r$$

La suma de r variables aleatorias independientes con distribución geométrica, todas con media $\mu_{X_i} = \dfrac{1}{p}$

[3] También se denota $X \sim BinNeg(r, p)$, y en algunos textos $X \sim b^*(r, p)$.

y varianza $\sigma^2_{X_i} = \dfrac{q}{p^2}$ entonces:

$$E(X) = E(X_1 + X_2 + X_3 + ... + X_r)$$
$$= E(X_1) + E(X_2) + E(X_3) + ... + E(X_r)$$
$$= \frac{1}{p} + \frac{1}{p} + \frac{1}{p} + ... + \frac{1}{p}$$
$$= \frac{r}{p}$$

Y para la variaza se utiliza un procedimiento similar.

El nombre de Distribución *binomial negativa* [4] proviene del hecho de que los valores de la distribución $f_X(x) = \binom{x-1}{r-1} p^r q^{x-r}$, para $x = r, r+1, ...$ son los términos sucesivos de la expansión binomial $\left(\frac{1}{p} - \frac{q}{p}\right)^{-r}$, aunque en realidad el modelo binomial negativo es una generalización del modelo de Pascal, en donde r puede tomar valores reales positivos. Algunos autores definen a la variable binomial negativa como el número de fallas para observar el r-ésimo éxito, lo que modifica el rango de la variable aleatoria y su media, pero no su variancia.

S)))))))))))))))))))))))))))))))))))))Q

Ejemplo 4.4

Los alumnos de cierta Facultad, son cuestionado con respecto a su calificación en el examen diagnóstico. La Facultad debe encontrar a tres alumnos que hayan acreditado el examen. Se sabe por experiencia que el 40% de los alumnos acreditan dicho examen. Encontrar la probabilidad de que se tenga que preguntar a 10 empleados para encontrar tres que hayan acreditado el examen.

Resolución

Sea X la variable aleatoria que representa el número de alumnos a los que se les debe preguntar para encontrar al tercero que haya acreditado el examen, entonces:

$X \sim Pascal(r = 3, p = 0.4)$

$$P(X = 10) = \binom{9}{2}(0.4)^3(0.6)^7 = 0.06$$

---

4 El desarrollo del binomio de Newton para exponente negativo está dado por:

$$(a+b)^{-n} = a^{-n} - n a^{-n-1} b + \frac{n(n+1)}{2!} a^{-n-2} b^2 - \frac{n(n+1)(n+2)}{3!} a^{-n-3} b^3 + \ldots$$

S))))))))))))))))))))))))))))))))))Q

Ejemplo 4.5

Un gran lote de llantas contiene 10% de llantas defectuosas, de ese lote se deben elegir cuatro que se colocarán en un auto.

a) Obtener la probabilidad de que seis llantas deban seleccionarse del lote para obtener cuatro en buen estado.

b) Calcular el valor esperado y la variancia del número de selecciones que deben efectuarse para obtener cuatro llantas sin defectos.

Resolución

a) Sea Y el número de llantas que deben ser seleccionadas hasta encontrar cuatro en buen estado.

$$Y \sim Pascal\ (r = 4 \,,\, p = 0.9)$$

$$P (Y = 6) = \binom{6 - 1}{4 - 1} (0.9)^4 \ (0.1)^2 = 0.06561$$

b) $$\mathrm{E} [Y] = \frac{r}{p} = \frac{4}{0.9} = 4.444$$

$$\mathrm{Var} [Y] = \frac{r (1 - p)}{p^2} = \frac{4 (0.1)}{(0.9)^2} = 0.4938$$

S))))))))))))))))))))))))))))))))))Q

DISTRIBUCIÓN HIPERGEOMÉTRICA

Cuando se desea calcular la probabilidad de obtener x éxitos en n ensayos, pero la probabilidad en cada ensayo no se mantiene constante, sino que se ve modificada en función de las extracciones anteriores, la distribución binomial no puede utilizarse.

En estos casos se utiliza la distribución hipergeométrica.

En la práctica la diferencia entre la distribución binomial y la distribución hipergeométrica radica en el hecho de que para el caso de la binomial se considera que la población es *infinita* (o finita con reemplazo), por lo que la probabilidad de éxito se mantiene constante en cada ensayo; mientras que para la hipergeométrica se considera que la población es *finita* de tamaño N (y sin reemplazo) y en ella existen r elementos que satisfacen la característica de interés (éxitos), por lo que en cada ensayo la probabilidad de éxito se modifica.

Definición 4.6

Si la variable aleatoria X representa el número de éxitos en n ensayos extraídos de una población de tamaño N, con r elementos que tienen la característica de interés, entonces X es una variable aleatoria hipergeométrica con parámetros N, n y r; y con función de probabilidad:

$$f_X(x) = \begin{cases} \dfrac{\dbinom{r}{x}\dbinom{N-r}{n-x}}{\dbinom{N}{n}} & ; \quad x = \max(0, n+k-N), \ldots, \min(k, n) \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

Se denota $X \sim \textit{Hipergeométrica}(r, n, N)$ [5]

La deducción de la fórmula para la distribución hipergeométrica es inmediata al considerar número de casos a favor entre número de casos totales aplicando las técnicas de conteo.

Debe observarse que la distribución hipergeométrica converge a la binomial cuando $N \to \infty$. En la práctica, la distribución binomial representa una buena aproximación a la hipergeométrica siempre que $n < 0.1\ N$, esto es, la muestra es menos de la décima parte de la población

Teorema 4.6

Sea X una variable aleatoria con distribución hipergeométrica con parámetros N, n y r, entonces:

$$\mu_X = E[X] = n\left(\frac{r}{N}\right)$$

$$\sigma_X^2 = \mathrm{Var}[X] = n\left(\frac{r}{N}\right)\left(1 - \frac{r}{N}\right)\left(\frac{N-n}{N-1}\right)$$

El cociente $\dfrac{r}{N}$ debe interpretarse como la proporción de elementos que cumplen la característica de interés, o bien, como la probabilidad de observar un éxito en la primera extracción. Entonces si se sustituye $p = \dfrac{r}{N}$ el valor esperado y la variancia de una distribución hipergeométrica se pueden escribir

$$E[X] = n\,p$$

$$\mathrm{Var}[X] = n\,p\,q\left(\frac{N-n}{N-1}\right)$$

El valor esperado coincide con el de la distribución binomial y la variancia sólo difiere en el factor $\dfrac{N-n}{N-1}$,

[5] o $X \sim h(N, n, r)$.

el cual recibe el nombre de *factor de corrección por población finita* y se denota $FCPF$.

$$FCPF = \frac{N-n}{N-1}$$

Si $N \to \infty$ entonces $FCPF \to 1$ y los parámetros de la distribución hipergeométrica coinciden con los de la binomial.

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 4.6
¿Cuál es la probabilidad de que una auditora de Hacienda detecte solamente dos declaraciones de impuestos con deducciones ilegales, si se seleccionan aleatoriamente seis de 18 declaraciones, ocho de las cuales contienen deducciones ilegales?

Resolución
Sea X la variable aleatoria que representa al número de declaraciones ilegales detectadas, entonces:

$X \sim$ Hipergeométrica $(r = 8 \,, n = 6 \,, N = 18)$

$$P(X = 2) = \frac{\binom{8}{2}\binom{10}{4}}{\binom{18}{6}} = 0.3167$$

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 4.7
Un cargamento de 120 alarmas contra robo contiene cinco defectuosas. Si tres de ellas son seleccionadas aleatoriamente y embarcadas para un cliente, obtener la probabilidad de que al cliente le toque una defectuosa, utilizando
a) la fórmula de la distribución hipergeométrica;
b) la fórmula de la distribución binomial como una aproximación.

Resolución

a) Sea X la variable aleatoria que representa el número de alarmas contra robo defectuosos de una muestra de tres, si en un lote de 120 sólo cinco están defectuosos

$X \sim$ *Hipergeométrica* $(r = 5, n = 3, N = 120)$

$$P(X = 1) = \frac{\binom{5}{1}\binom{115}{2}}{\binom{120}{3}} = 0.1167$$

b) Sea Y la variable aleatoria que representa el número de alarmas defectuosas.

$$Y \sim Binomial\left(n = 3, p = \frac{5}{120} \right)$$

$$P(Y = 1) = \binom{3}{1}\left(\frac{5}{120}\right)^{1}\left(\frac{115}{120}\right)^{2}$$

$$P(Y = 1) = 0.1148$$

S))))))))))))))))))))))))))))))))))Q

PROCESO DE POISSON

Cuando nos interesa contar el número de eventos o sucesos en un intervalo de tiempo, surge el proceso de Poisson, desarrollado por el francés SiméonDenis Poisson (1781-1840), y que es la base para la distribución de Poisson

Un proceso de Poisson[6] , es un proceso con tasa o intensidad λ, si satisface las siguientes propiedades:

1) La probabilidad de que el número de ocurrencias en el instante cero sea cero es igual a uno, esto es, el proceso inicia con cero ocurrencias.
$P\left(N(0) = 0\right) = 1$, esto es, $N(0) = 0$.
2) Para toda s entre 0 y t , $0 < s < t$, $N(t) - N(s)$ tiene distribución de Poisson con parámetro $\lambda(t - s)$, esto implica que el parámetro λ es escalable para distintos intervalos.
3) Independencia. El número de ocurrencias en cualquier intervalo de tiempo es independiente del número de ocurrencias en cualquier otro intervalo de tiempo.

DISTRIBUCIÓN DE POISSON

La distribución de Poisson es una de las distribuciones discretas que tienen más aplicación. Sirve cuando se desea calcular la probabilidad de ocurrencias de un evento en un intervalo continuo determinado. En particular, se puede modelar el número de llegadas por unidad de tiempo.

Deducción de la función de Probabilidad de Poisson a partir de la Distribución Binomial

Considérese la distribución binomial

$$f_X(x; n, p) = \frac{n!}{(n-x)!\ x!}\ p^x\ (1-p)^{n-x}$$

si para $n \geq 1$ se cumple la relación $p = \dfrac{\lambda}{n}$ de modo que $n \to \infty,\ p \to 0,\ np = \lambda$

entonces

$$f_X(x; \lambda) = \frac{n!}{(n-x)!\ x!} \left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x}$$

[6] Aquí se considerará el caso de homogéneo.

$$= \frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^n \frac{n!}{(n-x)!\; n^x \left(1-\frac{\lambda}{n}\right)^x}$$

$$= \frac{\lambda^x}{x!}\left(1-\frac{\lambda}{n}\right)^n \frac{n(n-1)(n-2)\ldots(n-x+1)}{\left[n\left(1-\frac{\lambda}{n}\right)\right]^x}$$

Cuando $n \to \infty$, $\left(1-\frac{\lambda}{n}\right)^n \to e^{-\lambda}$, mientras que el tercer factor tiende a uno, por lo que:

$$f_X(x;\lambda) = \frac{\lambda^x}{x!}\, e^{-\lambda}$$

Para modificar la escala se sustituye λ por λt,

$$f_X(x;\lambda t) = \frac{(\lambda t)^x}{x!}\, e^{-\lambda t}$$

que es la distribución de Poisson.

La escala t no es indispensable para la distribución de Poisson, puesto que puede considerarse que se estudia sobre 1 unidad (de tiempo) con lo que $t = 1$.

Definición 4.7

Sea X una variable aleatoria con distribución de Poisson, con parámetro λ entonces

$$f_X(x) = \begin{cases} \frac{\lambda^x}{x!}\, e^{-\lambda} & ; \quad x = 0, 1, \ldots \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

Se denota $X \sim Poisson(\lambda)$.

Comúnmente, las unidades de la intensidad λ se consideran $\left[\frac{ocurrencias}{tiempo}\right]$, pero en general son $\left[\frac{eventos}{intervalo\ continuo}\right]$.

En ocasiones se escribe la distribución de Poisson mediante el parámetro de escala θ, que es el recíproco de λ, esto es $\theta = \frac{1}{\lambda}$.

Teorema 4.7

Sea X una variable aleatoria con distribución de Poisson y parámetro λ, entonces:

$$\mu_X = \mathrm{E}[X] = \lambda \qquad \sigma_X^2 = \mathrm{Var}[X] = \lambda$$

Demostración del valor esperado

La demostración del valor esperado se simplifica si se utiliza la fórmula de recurrencia para la distribución de Poisson,

$$R_P(x) = \frac{f_X(x+1;\lambda)}{f_X(x;\lambda)} = \frac{\lambda}{x+1}$$

de donde

$$f_X(x+1;\lambda) = \frac{\lambda}{x+1} f_X(x;\lambda), \quad x = 0, 1, \ldots$$

o bien

$$(x+1)\, f_X(x+1;\lambda) = \lambda f_X(x;\lambda)$$

por lo que

$$\sum_{x=0}^{\infty} (x+1)\, f_X(x+1;\lambda) = \sum_{x=0}^{\infty} \lambda f_X(x;\lambda)$$

$$= \lambda \sum_{x=0}^{\infty} f_X(x;\lambda)$$

$$= \lambda$$

y para el lado izquierdo, si $y = x + 1$

$$\sum_{x=0}^{\infty} (x+1)\, f_X(x+1;\lambda) = \sum_{y=1}^{\infty} y\, f_X(y;\lambda)$$

y reescribiendo a y como x, sin cometer error, puesto que para la suma funciona como variable muda,

$$\sum_{x=0}^{\infty} (x+1)\, f_X(x+1;\lambda) = \sum_{x=1}^{\infty} x\, f_X(x;\lambda)$$

$$= \mathrm{E}(X)$$

Por lo que $\mathrm{E}[X] = \lambda$

S))))))))))))))))))))))))))))))))))))Q

Ejemplo 4.8

En promedio, en una cierta intersección ocurren tres accidentes viales por mes. Determinar las probabilidades de que en un determinado mes en esta intersección ocurran

a) exactamente cinco accidentes;

b) menos de tres accidentes.

Resolución

Sea X la variable aleatoria que representa el número de accidentes por mes.

$X \sim Poisson(\lambda = 3)$

a) $$P(X = 5) = \frac{e^{-3}(3)^5}{5!} = 0.10081$$

b) $$P(X < 3) = P(X \le 2) = \sum_{x=0}^{2} \frac{e^{-3}(3)^x}{x!} = 0.4232$$

S))))))))))))))))))))))))))))))))))Q

Ejemplo 4.9

Si X tiene una distribución de Poisson con parámetro de escala θ, con $\theta = \frac{1}{\lambda}$, y si $P(X = 0) = 0.2$, calcular $P(X > 2)$.

Nota: La distribución de Poisson en términos de θ, se escribe:

$$P(X = x) = \frac{e^{-\frac{1}{\theta}}\left(\frac{1}{\theta}\right)^x}{x!} \; ; \; x = 0, 1, \ldots$$

Resolución

De $P(X = 0) = 0.2$

$$= \frac{e^{-\frac{1}{\beta}}\left(\frac{1}{\beta}\right)^0}{0!}$$

$$0.2 = e^{-\frac{1}{\beta}}$$

$$\ln 0.2 = -\frac{1}{\beta} \quad \Rightarrow \quad \beta = -\frac{1}{\ln 0.2}$$

$$\beta = 0.6213$$

$$\therefore \; P(X > 2) = 1 - P(X \le 2)$$

$$= 1 - \big(P(X = 0) + P(X = 1) + P(X = 2)\big)$$

$$= 1 - (0.2 + 0.3219 + 0.259)$$

$$= 0.2191$$

S))))))))))))))))))))))))))))))))))Q

APROXIMACIÓN DE LA DISTRIBUCIÓN DE POISSON A LA BINOMIAL

Puesto que la distribución de Poisson puede obtenerse como un caso límite de la distribución binomial, no debe de sorprender que sirva para aproximar probabilidades binomiales. Cuando n es grande y p pequeña, las probabilidades binomiales se pueden aproximar mediante la distribución de Poisson utilizando $\lambda = np$

Una regla aceptable para emplear la aproximación es que si $n \geq 20$ entonces $p \leq 0.05$; o si $n \geq 100$ la aproximación es muy buena para $np \leq 10$. En general, mientras más grande sea n y más pequeña sea p, tanto mejor será la aproximación.

S)))))))))))))))))))))))))))))))))))Q

Ejemplo 4.10

Si el 0.8% de los fusibles depositados en un lote están defectuosos, utilizar la aproximación de Poisson para determinar la probabilidad de que cuatro fusibles estén defectuosos en una muestra aleatoria de 400.

Resolución

$n = 400 \qquad p = 0.008 \qquad$ de donde $\qquad \lambda = np = 3.2$

Sea X la variable aleatoria que representa el número de defectuosos entre los 400, $X \sim Binomial\,(\,400\;,\;0.008\,)$ y Y la variable aleatoria que representa el número de defectuosos, $Y \sim Poisson\,(\,3.2\,)$

$$P(\,X = 4\,) \approx P(\,Y = 4\,) = \frac{(\,3.2\,)^4\,e^{-3.2}}{4!} = 0.178$$

S)))))))))))))))))))))))))))))))))))Q

Después de estudiar las variables aleatorias discretas, deben de estudiarse aquellos modelos que sirven para los experimentos cuyo resultado tome valores de un conjunto continuo, lo que da lugar a las distribuciones continuas.

Tópico Especial: Distribución Binomial Negativa

Si la variable aleatoria X representa el número de ensayos necesarios para observar el $r-ésimo$ éxito, entonces X es una variable aleatoria con distribución de Pascal, $X \sim Pascal(r\,,p)$; pero al modelar sobre el mismo experimento, se puede definir a la variable aleatoria Y, que representa el número de fracasos hasta que se observen los primeros r éxitos, teniéndose la relación

$$X = Y + r$$

y Y es una variable aleatoria que algunos autores llaman binomial negativa. Se puede denotar por

$$Y \sim BinNeg(r\,,p)$$

aunque muchos autores utilizan la notación b^* o bn.

La función de probabilidad se deduce a partir de la distribución de Pascal, teniéndose:

$$f_Y(y) = P(Y = y) = P(X - r = y) = P(X = y + r)$$

$$= \binom{y + r - 1}{r - 1} p^r (1 - p)^y$$

Utilizando la propiedad de simetría del coeficiente binomial, la distribución binomial negativa puede escribirse como

$$f_Y(y) = \begin{cases} \binom{y+r-1}{y} p^r (1-p)^y & ; \quad y = 0, 1, 2, \ldots \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

donde r es un entero positivo.

La distribución de pascal, o la binomial negativa, pueden generalizarse para casos en los que r no es un entero, para ello sólo es necesario generalizar la definición del coeficiente binomial a través de la función gamma.

Definición 4.11
Si n y k son números reales, entonces

$$\binom{n}{k} = \frac{\Gamma(n+1)}{\Gamma(n-k+1)\,\Gamma(k+1)}$$

donde Γ es la función gamma.

Utilizando esta extensión de la definición del coeficiente binomial, la distribución binomial negativa puede escribirse como

$$f_Y(y) = \frac{\Gamma(y+r)}{\Gamma(r)\,y!} p^r (1-p)^y \; ; \qquad y = 0, 1, 2, \ldots$$

$$r > 0 \quad , \quad 0 < p < 1$$

O bien, en forma más compacta

$$f_Y(y) = \binom{-r}{y} p^r (p-1)^y$$

$$= \binom{-r}{y} p^r (-q)^y$$

en donde el coeficiente binomial pude calcularse como:

$$\binom{-r}{y} = \frac{(-r)(-r-1)\,(-r-2)\,.\,.\,.\,(-r-y+1)}{y\,!}$$

para r y y enteros.

Y de la última expresión para $f_Y(y)$ puede observarse la razón de que se le denomine distribución binomial negativa. Para hacerlo evidente considérese la probabilidad del espacio muestral, y a partir de ahí, se escriben las probabilidades de la distribución, esto es:

$$1 = p^r p^{-r} = p^r (1-q)^{-r}$$

Y utilizando el binomio de Newton (que ahora puede generalizarse también)

$$1 = p^r \sum_{y=0}^{\infty} \binom{-r}{y} (-q)^y$$

y los términos

$$p^r \binom{-r}{y} (-q)^y$$

son las probabilidades de la variable aleatoria Y, el número de fracasos para observar los primeros r éxitos.

Finalmente

$$p^r \sum_{y=0}^{\infty} \binom{-r}{y} (-q)^y = \sum_{y=0}^{\infty} \binom{-r}{y} p^r (-q)^y = 1$$

Y del inicio

$$1 = p^r (1-q)^{-r} = \left(\frac{1}{p} - \frac{q}{p} \right)^{-r}$$

De donde

$$\sum_{y=0}^{\infty} \binom{-r}{y} p^r (-q)^y = \left(\frac{1}{p} - \frac{q}{p} \right)^{-r}$$

y se observa que el desarrollo del binomio negativo, proporciona los términos de probabilidad de la distribución, de ahí el nombre de binomial negativa.

S))))))))))))))))))))))))))))))))))Q

Tópico Especial: Deducción de la función de Probabilidad de Poisson a partir del proceso de Poisson

Sea $f_X(x\,;t)$ la probabilidad de tener x ocurrencias en el intervalo de tiempo t, entonces, si se cumplen las características del proceso de Poisson

1. La probabilidad de una sola ocurrencia en un intervalo muy pequeño dt es $\lambda\, dt$.
2. El intervalo dt es tan pequeño que la probabilidad de tener más de una ocurrencia es despreciable.
3. En un intervalo, los eventos son independientes.

Para observar x ocurrencias en el intervalo $t + dt$, se tienen dos diferentes (excluyentes) probabilidades.

1. Existen x ocurrencias en t, con probabilidad $f_X(x\,;t)$ y ninguna en dt, con probabilidad $1 - \lambda\, dt$.
2. Existen $x-1$ ocurrencias en t, con probabilidad $f_X(x-1\,;t)$ y una ocurrencia en dt con probabilidad $\lambda\, dt$.

De la suposición de independencia, la probabilidad de observar x ocurrencias en $t + dt$ es:

$$f_X(x\,;t+dt) = f_X(x\,;t)(1-\lambda\, dt) + f_X(x-1\,;t)(\lambda\, dt)$$

manipulando algebraicamente

$$f_X(x\,;t+dt) = f_X(x\,;t) - \lambda f_X(x\,;t)\, dt + \lambda f_X(x-1\,;t)\, dt$$

$$f_X(x\,;t+dt) - f_X(x\,;t) = -\lambda f_X(x\,;t)\, dt + \lambda f_X(x-1\,;t)\, dt$$

$$\frac{f_X(x\,;t+dt) - f_X(x\,;t)}{dt} = \lambda \left[f_X(x-1\,;t) - f_X(x\,;t) \right]$$

si se toma el límite cuando $dt \to 0$, entonces el lado izquierdo de la igualdad se convierte en la derivada de $f_X(x\,;t)$, es decir, $\frac{d}{dt} f_X(x\,;t)$

$$\frac{d}{dt} f_X(x\,;t) = \lambda \left[f_X(x-1\,;t) - f_X(x\,;t) \right] \qquad \text{. . . (a)}$$

que es una ecuación diferencial lineal con respecto a t y una ecuación en diferencias finitas de primer orden con respecto a x.

En particular, si $x = 0$, entonces (a) se convierte en

$$\frac{d}{dt} f_X(0\,;t) = \lambda \left[f_X(-1\,;t) - f_X(0\,;t) \right]$$

y puesto que $f_X(-1\,;t) = 0$ (la probabilidad de tener -1 ocurrencias es cero)

$$\frac{d}{dt} f_X(0\,;t) = -\lambda\, f_X(0\,;t)$$

Resolviendo por separación de variables la ecuación diferencial se tiene:

$$\frac{d\, f_X(0\,;t)}{f_X(0\,;t)} = -\lambda\, dt$$

$$\ln f_X(0\,;t) = -\lambda t + c_1$$

$$f_X(0\,;t) = e^{-\lambda t + c_1}$$

$$f_X(0\,;t) = c\; e^{-\lambda t}$$

Y puesto que la probabilidad de tener cero ocurrencias en un intervalo de longitud $t = 0$, debe ser uno, entonces $c = 1$, por lo que:

$$f_X(0\,;t) = e^{-\lambda t}$$

Si $x = 1$, entonces (a) se convierte en

$$\frac{d}{dt} f_X(1\,;t) = \lambda \left[f_X(0\,;t) - f_X(1\,;t) \right]$$

y puesto que $f_X(0\,;t) = e^{-\lambda t}$ entonces

$$\frac{d}{dt} f_X(1\,;t) + \lambda f_X(1\,;t) = \lambda e^{-\lambda t}$$

Se tiene una ecuación diferencial no homogénea. Para resolverla se debe considerar la condición inicial $f_X(1\,;0) = 0$, puesto que la probabilidad de tener una ocurrencia en un intervalo de longitud cero, $t = 0$, es cero.

Simplificando la notación, si $y = f_X(1\,;t)$ entonces:

$$y' + \lambda\, y = \lambda\; e^{-\lambda t}$$

aplicando la transformada de Laplace[7]

$$s\;Y - 0\; + \lambda\;Y\; = \lambda\;\frac{1}{s + \lambda}$$

$$Y = \frac{\lambda}{(s + \lambda)^2}$$

antitransformando

$$y\; = \lambda\;e^{-\lambda t}\;t$$

$$= \lambda\;t\;e^{-\lambda t}$$

En la notación original

$$f_X(1\,;t) = \lambda\,t\,e^{-\lambda t}$$

Procediendo de manera similar, para $x\; = \;2$, de (a) se obtiene:

$$f_X(2\,;t)\; = \frac{(\lambda\,t)^2\;e^{-\lambda t}}{2}$$

Al continuar el proceso se deduce que

$$f_X(x\,;t)\; = \begin{cases} \dfrac{(\lambda\,t)^x\;e^{-\lambda t}}{x!} & ; \quad x\; = \;0\,,1\,,\ldots \\ 0 & ; \quad en\;\;otro\;\;caso \end{cases}$$

expresión que recibe el nombre de distribución de Poisson.

S))))))))))))))))))))))))))))))))))))Q

Tópico adicional

Demostración del valor esperado y la varianza de la variable aleatoria binomial utilizando la función generadora de momentos

Primero obtendremos la función generadora

$$M_X(\theta)\; = \;\mathrm{E}\,(e^{\theta X})\; = \sum_{\forall x}\;e^{\theta x}\,f_X(x)$$

sustituyendo la función de probabilidad

---

[7] Algunas fórmulas básicas son:

$$\mathcal{L}\,\{y'\}\; = \;s\,Y - y(0), \qquad \mathcal{L}\,\{e^{at}\}\; = \;\frac{1}{s - a}$$

$$\mathcal{L}\,\{e^{at}f(t)\}\; = \;F(s - a),\;\; \mathcal{L}\,\{t\}\; = \;\frac{1}{s^2}$$

$$M_X(\theta) = \sum_{x=0}^{n} e^{\theta x} \binom{n}{x} p^x (1-p)^{n-x}$$

reacomodando términos

$$M_X(\theta) = \sum_{x=0}^{n} \binom{n}{x} \left(e^{\theta} p \right)^x (1-p)^{n-x}$$

lo que debe traernos a la mente el binomio de Newton

$$(a+b)^n = \sum_{x=0}^{n} \binom{n}{x} a^x b^{n-x}$$

que al hacer las asignaciones

$$e^{\theta} p = a$$
$$1 - p = b$$

nos da la simplificación

$$M_X(\theta) = \sum_{x=0}^{n} \binom{n}{x} \left(e^{\theta} p\right)^x (1-p)^{n-x} = \left(e^{\theta} p + (1-p)\right)^n$$

y renombrando al complemento de p como q,

$$q = 1 - p$$

se obtiene la función generadora de momentos de la variable aleatoria binomial

$$M_X(\theta) = \left(e^{\theta} p + q\right)^n$$

Puede verse con facilidad que la función generadora valuada en 0 es 1, puesto que $p + q = 1$

$$M_X(0) = \left(e^{0} p + q\right)^n = (p+q)^n = 1$$

Para obtener la media derivamos a la función generadora

$$\frac{d M_X(\theta)}{d\theta} = n\left(e^{\theta} p + q\right)^{n-1} e^{\theta} p$$

acomodando términos

$$\frac{d M_X(\theta)}{d\theta} = n p e^{\theta} \left(e^{\theta} p + q\right)^{n-1}$$

valuando en $\theta = 0$

$$\left.\frac{d M_X(\theta)}{d\theta}\right|_{\theta=0} = n p$$

Por lo que el valor esperado de la v.a. es

$$\mathrm{E}(X) = n p$$

Para obtener la variancia necesitamos el segundo momento con respecto al origen, que encontramos con la segunda derivada de la función generadora

$$\mathrm{E}(X^2) = \frac{d^2 M_X(\theta)}{d\theta^2} = \left(npe^{\theta}\right)\left((n-1)(e^{\theta}p+q)^{n-2}e^{\theta}p\right) + \left(e^{\theta}p+q\right)^{n-1}\left(npe^{\theta}\right)$$

Valuando en el origen

$$\left.\frac{d^2 M_X(\theta)}{d\theta^2}\right|_{\theta=0} = n\,p\,(n-1)\,p + n\,p$$

$$\left.\frac{d^2 M_X(\theta)}{d\theta^2}\right|_{\theta=0} = n\,p^2\,(n-1) + n\,p$$

$$\left.\frac{d^2 M_X(\theta)}{d\theta^2}\right|_{\theta=0} = n\,p - n\,p^2 + n^2\,p^2$$

Y puesto que la varianza de X es igual al segundo momento con respecto al origen menos la media al cuadrado

$$\mathrm{Var}[X] = \mathrm{E}[X^2] - \mu_X^2$$

sustituyendo $\mathrm{Var}[X] = \left(n\,p - n\,p^2 + n^2\,p^2\right) - (np)^2$

simplificando $\mathrm{Var}[X] = n\,p - n\,p^2$

factorizando $\mathrm{Var}[X] = n\,p\,(1-p)$

y finalmente la varianza de la v.a. binomial es:

$$\mathrm{Var}[X] = n\,p\,q$$

Tópico adicional

Demostración de la varianza de la variable aleatoria geométrica

Si X es una variable aleatoria con distribución geométrica, entonces:

$$X \sim \textit{Geométrica}(p)$$

$$f_X(x) = \begin{cases} q^{x-1}p & ; \; x = 1, 2, \ldots, n \\ 0 & ; \; \textit{en otro caso} \end{cases}$$

La media ya se demostró y es igual a $\mu_X = \dfrac{1}{p}$.

Para la varianza

$$\mathrm{Var}[X] = \mathrm{E}[X^2] - \mu_X^2$$

Y se utilizará $\mathrm{E}[X(X-1)] = \mathrm{E}[X^2] - \mu_X$, de donde $\mathrm{E}[X^2] = \mathrm{E}[X(X-1)] + \mu_X$

Por lo que

$$\text{Var}[X] = \text{E}[X(X-1)] + \mu_X - \mu_X^2 \qquad \text{... (a)}$$

Y ahora se debe encontrar $E[X(X-1)]$

$$\text{E}[X(X-1)] = \sum_{x=1}^{\infty} x(x-1)\,p\,q^{(x-1)}$$

factorizando a p y una q se tiene

$$\text{E}[X(X-1)] = p\,q \sum_{x=1}^{\infty} x(x-1)\,q^{(x-2)}$$

que claramente nos recuerda la estructura de una segunda derivada, esto es

$$\begin{aligned}\text{E}[X(X-1)] &= p\,q \sum_{x=1}^{\infty} \frac{d^2}{dq^2}\,q^x \\ &= p\,q\,\frac{d^2}{dq^2} \sum_{x=1}^{\infty} q^x \\ &= p\,q\,\frac{d^2}{dq^2}\,\frac{q}{1-q} \qquad \text{...(b)}\end{aligned}$$

Ahora se requiere la segunda derivada $\dfrac{d^2}{dq^2}\,\dfrac{q}{1-q}$

$$\frac{d^2}{dq^2}\,\frac{q}{1-q} = \frac{d}{dq}\left(\frac{d}{dq}\,\frac{q}{1-q}\right) = \frac{d}{dq}\,\frac{1}{(1-q)^2} = \frac{2}{(1-q)^3} \qquad \text{...(c)}$$

Sustituyendo (c) en (b)

$$\text{E}[X(X-1)] = p\,q\,\frac{2}{(1-q)^3} = \frac{2q}{p^2} \qquad \text{...(d)}$$

y sustituyendo (d) en (a), junto con el valor de la media $\mu_X = \dfrac{1}{p}$,

$$\begin{aligned}\text{Var}[X] &= \text{E}[X(X-1)] + \mu_X - \mu_X^2 \\ &= \frac{2q}{p^2} + \frac{1}{p} - \left(\frac{1}{p}\right)^2 = \frac{2q + p - 1}{p^2} \\ &= \frac{q}{p^2}\end{aligned}$$

))

Demostración de la función generadora de momentos de la variable aleatoria geométrica

Sea

$$X \sim \mathit{Geométrica}\,(p)$$

con función de probabilidad

$$f_X(x) = \begin{cases} q^{x-1}p & ; \; x = 1, 2, \ldots \\ 0 & ; \; \textit{en otro caso} \end{cases}$$

entonces su función generadora de momentos está dada por:

$$M_X(\theta) = \mathrm{E}\,[\,e^{\theta X}\,]$$

$$\begin{aligned} M_X(\theta) &= \sum_{x=1}^{\infty} e^{\theta x}\, q^{x-1}\, p \\ &= \sum_{x=1}^{\infty} e^{\theta(x-1)}\, e^{\theta}\, q^{x-1}\, p \\ &= e^{\theta}\, p \sum_{x=1}^{\infty} e^{\theta(x-1)}\, q^{x-1} \\ &= e^{\theta}\, p \sum_{y=0}^{\infty} e^{\theta y}\, q^{y} \\ &= e^{\theta}\, p \sum_{y=0}^{\infty} (e^{\theta}\, q)^{y} \\ &= e^{\theta}\, p\, \frac{1}{1 - q\, e^{\theta}} \\ &= \frac{p\, e^{\theta}}{1 - q\, e^{\theta}} \end{aligned}$$

AUTOEXAMEN TEMA IV

1.- En promedio, un jugador de beisbol conecta un "hit" en uno de tres intentos. Asumiendo que los eventos son independientes, la probabilidad de que conecte exactamente 3 "hits" en 6 intentos es: ☐

A) $\frac{160}{3^6}$ B) $\frac{160}{3^5}$ C) $\frac{1}{2}$ D) $\frac{80}{3^6}$ E) $\frac{40}{3^6}$

2.- En cierta ciudad, la probabilidad de que ocurra una tormenta en cualquier día durante la primavera es 0.03. Suponiendo independencia, la probabilidad de que la primera tormenta ocurra el día 20 de la primavera es: ☐

A) 0.0168 B) 0.03 C) 0.5438 D) 0.5606
E) Ninguna de las anteriores.

3.- Un gerente de personal está entrevistando a empleados potenciales con el fin de cubrir dos vacantes. La probabilidad de que el entrevistado tenga las cualidades necesarias y acepte un ofrecimiento es 0.8. La probabilidad de que el gerente deba entrevistar cuando mucho a 3 personas es: ☐

A) 0.64 B) 0.36 C) 0.0256 D) 0.896
E) Ninguna de las anteriores.

4.- Con el propósito de decidir si se aceptan los lotes de mercancía que envía una fábrica, se lleva a cabo un procedimiento que consiste en seleccionar diez artículos al azar de cada lote y determinar el número de defectuosos. Un lote se rechaza siempre que se encuentren dos o más artículos defectuosos entre los diez seleccionados. Supóngase que el número de artículos en cada lote es grande y que cada lote contiene un 5% de defectuosos, entonces la probabilidad de rechazar el lote es: ☐

A) 0.0746 B) 0.914 C) 0.086 D) 0.599 E) 0.401

5.- En un crucero poco transitado, sólo el 20% de los automóviles hacen alto total antes de cruzar. De los siguientes 40 automóviles que transitan por el crucero, el número esperado de ellos que hacen alto total es: ☐

A) Prácticamente cero B) 6.4 C) 8 D) 32 E) Ninguna de las anteriores.

6.- Un excursionista desea seleccionar el tipo de linterna de mano que debe llevar en su siguiente campamento. Puede llevar una linterna de una sola batería de 6 V u otra que utiliza dos baterías tamaño D de 1.5 V. En este momento tiene dos baterías de 6 V y cuatro de 1.5 V y no puede comprar más baterías. Si la probabilidad de que cualquier batería funcione es 0.6 y funcionan de manera independiente; entonces, las probabilidades de que las linternas de 6 V y de 3 V funcionen son, respectivamente: ☐

A) 0.16, 0.179 B) 0.48, 0.346 C) 0.84, 0.821 D) 0.36, 0.13
E) Ninguna de las anteriores.

7.- La probabilidad de que un alumno de cierta universidad tenga teléfono celular es 0.2. Entonces, la probabilidad de que el décimo alumno entrevistado aleatoriamente de esta universidad sea el quinto con teléfono celular es: $\square$

A) 0.0001 B) 0.0032 C) 0.026 D) 0.0132 E) Ninguna de las anteriores.

8.- En cierto día, se puso en práctica la segunda etapa del plan de contingencia ambiental, en la Ciudad de México. Si en la zona que le corresponde a un supervisor hay 50 empresas, de las cuales 15 no siguieron los lineamientos de la contingencia, entonces la probabilidad de que 3 empresas de las 10 que verifica no sigan los lineamientos es: $\square$

A) 0.2979 B) 0.6 C) 0.09 D) 0.2668 E) Ninguna de las anteriores.

9.- Si los aviones que llegan a un aeropuerto siguen un proceso de Poisson con una tasa media de 8 aviones por hora. La probabilidad de que lleguen exactamente 10 aviones en las siguientes 2 horas es lo más cercana a: $\square$

A) 0 B) 0.1126 C) 0.099 D) 0.0341 E) Ninguna de las anteriores.

10.- En un sistema de cómputo de tiempo compartido, el número de peticiones de telepuerto es 0.2 por milisegundo, en promedio, y sigue una distribución de Poisson. Entonces la probabilidad de que no lleguen peticiones durante los tres milisegundos siguientes es: $\square$

A) 0.5488 B) 0.8187 C) 0.1822 D) 0.2729 E) Ninguna de las anteriores

BIBLIOGRAFÍA

Hines, William W. y Montgomery, Douglas C. - Probabilidad y Estadística para ingeniería, cuarta edición..- CECSA.- México, 2005.

Milton, Susan J. Y Arnold, Jesse C.- Probabilidad y Estadística para con aplicaciones para ingeniería y ciencias computacionales, cuarta edición.- McGraw-Hill.- México, 2004.

Devore, Jay L.- Probabilidad y Estadística para ingeniería y ciencias, séptima edición.- Cengage Learning.- México, 2008.

Mendenhall, William III. et al.- Introducción a la Probabilidad y Estadística.- Décimo cuarta edición.- Cengage Learning.- México 2015.

Wackerly Dennis D.- Mendenhall, William, *et al*.- Estadística Matemática con Aplicaciones, sexta edición.- Editorial Thomson.- México, 2002.

Walpole, Ronald E., *et al*.- Probability and Statistics for Engineers and Scientists.- Pearson.- USA, 2007.

Montgomery, Douglas C. y Runger, George C.-Probabilidad y Estadística aplicadas a la Ingeniería, segunda edición.- Limusa-Wiley.- México, 2002.

Scheaffer, Richard L. y McClave, James T.- Probabilidad y Estadística para Ingeniería.- Grupo Editorial Iberoamérica.- México, 1993.

Canavos, George C.- Probabilidad y Estadística Aplicaciones y Métodos.- McGraw-Hill.- México, 1988.

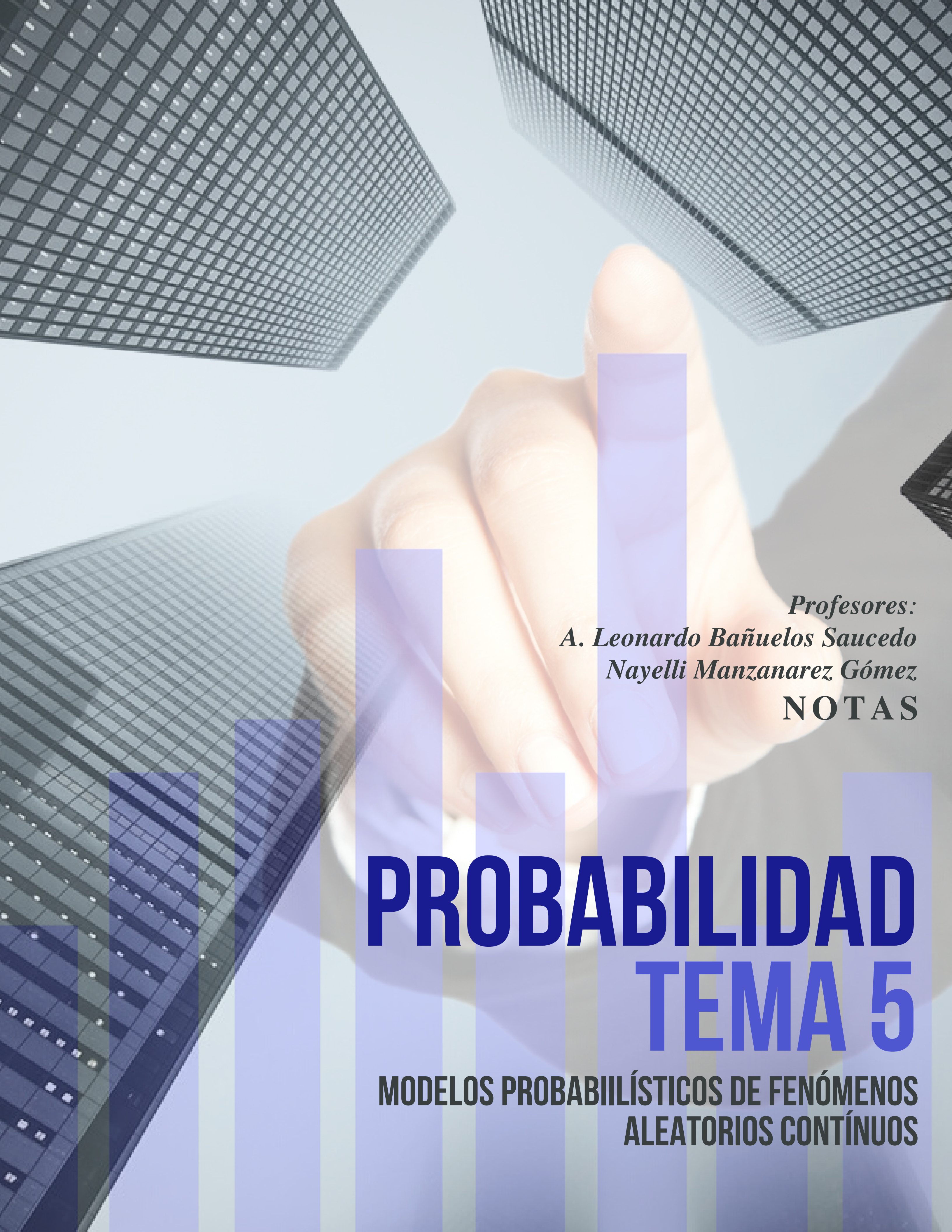
Meyer, Paul L.- Probabilidad y Aplicaciones Estadísticas.- Addison Wesley Iberoamericana.- México, 1992.

Spiegel, Murray R. et al.- Probabilidad y Estadística, cuarta edición.- Mc Graw-Hill.-México 2013.

Borras García, Hugo E., *et al*.- Apuntes de Probabilidad y Estadística.-Facultad de Ingeniería.- México, 1985.

Rosenkrantz, Walter A.- Introduction to Probability and Statistics for Scientists and Engineers.- McGraw-Hill.- EE.UU., 1997.

Ziemer, Rodger E.- Elements of Engineering Probability & Statistics.- Prentice Hall.- USA 1997.



Profesores:
A. Leonardo Bañuelos Saucedo
Nayelli Manzanarez Gómez

NOTAS

PROBABILIDAD
TEMA 5
MODELOS PROBABIILÍSTICOS DE FENÓMENOS ALEATORIOS CONTÍNUOS

TEMA V
MODELOS PROBABILÍSTICOS DE FENÓMENOS ALEATORIOS CONTINUOS

INTRODUCCIÓN

Después de estudiar las variables aleatorias discretas, deben de estudiarse aquellos modelos que sirven para los experimentos cuyo resultado tome valores de un conjunto continuo, lo que da lugar a las distribuciones continuas. En este tema se describen los modelos continuos más comunes y utilizados en ingeniería y en la toma de decisiones a través de la inferencia estadística.

DISTRIBUCIONES CONTINUAS

Las distribuciones continuas que se estudiarán en este capítulo, son:

Uniforme continua.
Exponencial.
Erlang.
Normal.
Gamma.
Ji cuadrada.
T de Student.
F de Fisher.
Weibull.
Lognormal.

DISTRIBUCIÓN CONTINUA UNIFORME

La distribución continua uniforme es las más simple de todas las distribuciones continuas de probabilidad, como su nombre lo indica, la probabilidad para la variable aleatoria es uniforme, o bien, constante.

Definición 5.1

Sea X una variable aleatoria que se distribuye

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & ; \quad a \le x \le b \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

entonces X tiene una distribución continua uniforme con parámetros a y b.

Se denota $X \sim Uniforme\,(a, b)$.

En la notación de la distribución uniforme continua, debe observarse que se tienen dos parámetros, por lo que no se puede confundir con la distribución discreta uniforme, puesto que la discreta se denota sólo con un elemento.

Teorema 5.1

Si X es una variable aleatoria con distribución continua uniforme, entonces:

$$\mu_X = \mathrm{E}[X] = \frac{a+b}{2}$$

$$\sigma_X^2 = \mathrm{Var}[X] = \frac{(b-a)^2}{12}$$

Las demostraciones son inmediatas al aplicar la definición de valor esperado y variancia.

Para la variable aleatoria continua uniforme, la función de distribución acumulada puede obtenerse integrando directamente, de donde:

$$F_X(x) = \begin{cases} 0 & ; \quad x < a \\ \dfrac{x-a}{b-a} & ; \quad a \le x \le b \\ 1 & ; \quad x > b \end{cases}$$

que se utiliza en el siguiente ejemplo.

Ejemplo 5.1

El tiempo de un viaje (ida y vuelta) de los camiones que transportan el concreto hacia una obra de construcción en una carretera, está distribuido uniformemente en un intervalo de 50 a 70 minutos. ¿Cuál es la probabilidad de que la duración del viaje sea mayor a 65 minutos si se sabe que la duración del viaje es mayor a 55 minutos?

Resolución

Sea X la variable aleatoria que representa el tiempo del viaje, entonces

$$X \sim \mathit{Uniforme}(50, 70), \quad f_X(x) = \begin{cases} \dfrac{1}{20} & ; \quad 50 \le x \le 70 \\ 0 & ; \quad \textit{en otro caso} \end{cases}$$

$$P(X > 65 \mid X > 55) = \frac{P(X > 65)}{P(X > 55)} = \frac{1 - \dfrac{65-50}{20}}{1 - \dfrac{55-50}{20}} = \frac{1}{3}$$

DISTRIBUCIÓN EXPONENCIAL

Dentro de un proceso de Poisson considérese la variable aleatoria T, que representa el tiempo que transcurre entre dos ocurrencias sucesivas de un evento, o bien, el tiempo que transcurre para el siguiente evento, esta definición de variable recibe el nombre de distribución exponencial.

Para obtener la función de densidad de una variable aleatoria con distribución exponencial, se obtendrá primero la función de distribución acumulativa, y puesto que la variable aleatoria T es continua la función de densidad se obtiene mediante derivación , con la relación

$$f_T(t) = \frac{d}{dt} F_T(t)$$

Si X es una variable aleatoria con distribución de Poisson, entonces

$$f_X(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

y agregando a la distribución la escala para el tiempo, se tiene:

$$f_X(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}$$

por lo que:

$$F_T(t) = P(T \le t)$$

$$= 1 - P(T > t)$$ Si la ocurrencia se da después de t, entonces antes de t el número de ocurrencias es igual a cero.

$$= 1 - P(X = 0)$$ y calculando la probabilidad de $P(X=0)$ mediante la distribución de Poisson.

$$= 1 - \frac{(\lambda t)^0 \; e^{-\lambda t}}{0!} = 1 - e^{-\lambda t}$$

$$F_T(t) = 1 - e^{-\lambda t}$$

Y derivando para obtener la función de densidad se tiene:

$$\frac{d}{dt} F_T(t) = \lambda e^{-\lambda t}$$

Definición 5.1

Sea T la variable aleatoria que representa el intervalo (generalmente tiempo), que transcurre entre dos ocurrencias sucesivas de un evento, entonces T tiene una distribución exponencial con parámetro λ y función de densidad

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & ; \quad t > 0 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

Se denota por $T \sim exponencial(\lambda)$ [1].

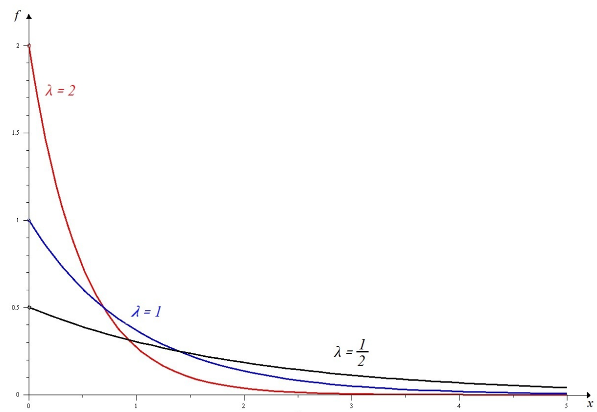


Fig. 5.1. Distribución exponencial

En ocasiones algunos autores utilizan el parámetro β para definir a la función de densidad exponencial, donde $\beta = \frac{1}{\lambda}$, por lo que la función de densidad queda:

$$f_T(t) = \begin{cases} \frac{1}{\beta} e^{-\frac{1}{\beta} t} & ; \quad t > 0 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

[1] También se denota $T \sim \exp(\lambda)$.

Teorema 5.2

Sea T una variable aleatoria con distribución exponencial y parámetro λ, entonces:

$$\mu_T = \mathrm{E}[\,T\,] = \frac{1}{\lambda}$$

$$\sigma_T^2 = \mathrm{Var}[\,T\,] = \frac{1}{\lambda^2}$$

y la función generadora de momentos es:

$$M_T(\,\theta\,) = \frac{\lambda}{\lambda - \theta}$$

Para obtener probabilidades con la distribución exponencial, se puede utilizar la función de densidad e integrarla, pero es más sencillo utilizar la función de distribución acumulada. El ejemplo 5.2 muestra los dos caminos.

Ejemplo 5.2

El tiempo en horas, que tarda un gerente en entrevistar a un aspirante para un trabajo, tiene una distribución exponencial con $\lambda = 2$. Los aspirantes están programados en intervalos de $\frac{1}{4}$ de hora, empezando a las 8:00 a.m., y los aspirantes llegan exactamente a tiempo. Cuando el aspirante con cita a las 8:15 a.m. llega a la oficina del gerente, ¿cuál es la probabilidad de que tenga que esperar para poder ver al gerente?

Resolución

Sea T la variable aleatoria que representa el tiempo que tarda una entrevista, entonces:

$T \sim \mathit{Exponencial}\,(\,\lambda = 2\,)$

$$P\left(\,T > \frac{1}{4}\,\right) = \int_{\frac{1}{4}}^{\infty} 2e^{-2t}\,dt = e^{-\frac{1}{2}}$$

O bien, utilizando la función de distribución acumulada, $F_T(\,t\,) = 1 - e^{-\lambda t}$, se tiene:

$$P\left(\,T > \frac{1}{4}\,\right) = 1 - P\left(\,T \le \frac{1}{4}\,\right) = 1 - F_T\left(\frac{1}{4}\right) = e^{-\frac{1}{2}} = 0.6065$$

Los problemas de probabilidad, pueden involucrar más de una distribución para resolverlos. El siguiente ejemplo muestra el caso en el cual se utiliza la variable aleatoria exponencial y luego se calcula un valor esperado de una variable aleatoria discreta.

Ejemplo 5.3

Un fabricante de monitores proporciona una garantía por un año (8760 horas). Los monitores se utilizan en terminales de aeropuerto para programas de vuelo y están encendidos en uso continuo. La vida media de los monitores es de 20 000 horas, y siguen una distribución exponencial. El costo de fabricación, venta y entrega es de \$3 000 y el monitor se vende en el mercado en \$4 000. Cuesta \$1 500 reemplazar el monitor cuando falla, incluyendo refacciones y mano de obra. El fabricante no tiene obligación de sustituir el monitor si ya ha habido una primera sustitución. ¿Cuál es la utilidad esperada del fabricante?

Resolución

Sean U la utilidad , y T el tiempo de vida de un monitor.

$$T \sim Exponencial\left(\lambda = \frac{1}{20000} \right)$$

$$U(t) = \begin{cases} 1000 & ; \quad t \geq 8760 \\ -500 & ; \quad t < 8760 \end{cases}$$

Por lo tanto

$$\mathrm{E}(U) = 1000 \; P(T \geq 8760) - 500 \, P(T < 8760)$$

$$P(T \geq 8760) = \int_{8760}^{\infty} \frac{1}{20000} \, e^{-\frac{1}{20000}t} \, dt = e^{-876/2000} \approx 0.64533$$

$$P(T < 8760) = 1 - P(T \geq 8760) \approx 0.35467$$

Finalmente:

$$\mathrm{E}(U) = 1000\,(0.64533) - 500\,(0.35467) = 467.99$$

La utilidad será de \$467.99 por monitor.

La distribución exponencial posee una característica muy importante, la cual se conoce como *falta de memoria* o *amnesia*. Es la única variable aleatoria continua con recorrido en los reales positivos que tiene esta propiedad.

La propiedad de falta de memoria dice que:

$$\boxed{P(T > t + s \mid T > t) = P(T > s)}$$

esto se comprueba con facilidad

$$P(T > t + s \mid T > t) = \frac{P(T > t + s)}{P(T > t)}$$

$$= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = e^{-\lambda s}$$

La interpretación contradice la intuición, pero básicamente la propiedad de amnesia indica que el tiempo para la siguiente ocurrencia "reinicia" cada vez que nos percatamos de que ha transcurrido el tiempo t sin ocurrencias.

DISTRIBUCIÓN ERLANG

En el capítulo anterior se estudiaron las distribuciones Geométrica y de Pascal (Binomial negativa), siendo la distribución de Pascal un caso general de la distribución Geométrica, de forma similar, se estudiará la distribución Erlang, que es un caso general de la distribución exponencial.

La distribución exponencial modela el tiempo entre ocurrencias, o dicho de otra forma, el tiempo para que ocurra la siguiente (o primer) ocurrencia, de forma que la primera generalización natural es una variable aleatoria que modele el tiempo hasta la r-ésima ocurencia. El matemático danés Agner Krarup Erlang (1878-1929) desarrolló el modelo que lleva su nombre al estudiar las llamadas telefónicas en centrales de conmutación.

Definición 5.2

Si X es una v.a. con distribución Erlang con parámetros λ y r, entonces X es la suma de r variables aleatorias independientes con distribución exponencial, cada una con parámetro λ, y tiene la función de densidad

$$f_X(x) = \lambda e^{-\lambda x} \frac{(\lambda x)^{r-1}}{(r-1)!} \quad \text{, para} \quad x > 0$$

Con $r = 1, 2, 3, \ldots$

Y se denota $X \sim Erlang(\lambda, r)$

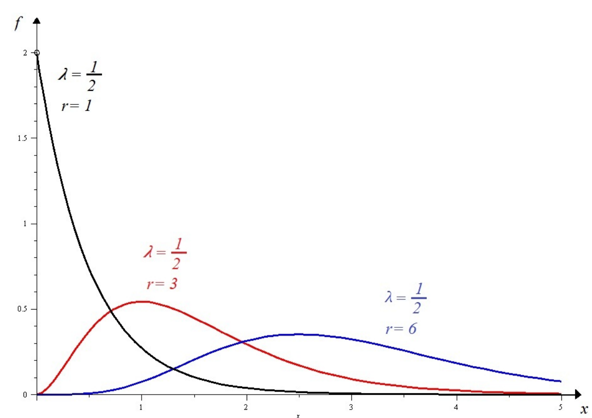


Fig. 5.2. Distribución Erlang

Los parámetros λ y r reciben los nombres de parámetro de escala y parámetro de forma, respectivamente. El máximo de la gráfica se obtiene en $x = \frac{r-1}{\lambda}$, por lo que la moda es $\boxed{x_{mo} = \frac{r-1}{\lambda}}$

La media y la variancia de la distribución Erlang se deducen inmediatamente a partir de su definición.

Sean T_1 , T_2 , T_3 , ... , T_r , r variables aleatorias independientes con distribución exponencial y parámetro λ, entonces:

$$X = T_1 + T_2 + T_3 + \ldots + T_r$$

$$\begin{aligned} \mathrm{E}[X] &= \mathrm{E}[\,T_1 + T_2 + T_3 + \ldots + T_r\,] \\ &= \mathrm{E}[\,T_1\,] + \mathrm{E}[\,T_2\,] + \mathrm{E}[\,T_3\,] + \ldots + \mathrm{E}[\,T_r\,] \\ &= \frac{1}{\lambda} + \frac{1}{\lambda} + \frac{1}{\lambda} + \ldots + \frac{1}{\lambda} \\ &= \frac{r}{\lambda} \end{aligned}$$

Y para la variancia

$$\begin{aligned} \mathrm{Var}[X] &= \mathrm{Var}[\,T_1 + T_2 + T_3 + \ldots + T_r\,] \\ &= \mathrm{Var}[\,T_1\,] + \mathrm{Var}[\,T_2\,] + \mathrm{Var}[\,T_3\,] + \ldots + \mathrm{Var}[\,T_r\,] \\ &= \frac{1}{\lambda^2} + \frac{1}{\lambda^2} + \frac{1}{\lambda^2} + \ldots + \frac{1}{\lambda^2} \\ &= \frac{r}{\lambda^2} \end{aligned}$$

Teorema 5.3

Si X es una v.a. con distribución Erlang con parámetros λ y r, entonces su media y su variancia son:

$$\mu_X = \mathrm{E}[X] = \frac{r}{\lambda}$$

$$\sigma_X^2 = \mathrm{Var}[X] = \frac{r}{\lambda^2}$$

Y su función generadora de momentos esta dada por: $M_X(\theta) = \left(\frac{\lambda}{\lambda - \theta}\right)^r$

Puede verificarse que la distribución Erlang es un caso general de la exponencial, puesto que si a partir de la distribución Erlang se sustituye $r = 1$ (la primera ocurrencia) se tiene:

$$f_X(x) = \lambda e^{-\lambda x} \frac{(\lambda x)^{1-1}}{(1-1)!} \qquad , x > 0$$

$$f_X(x) = \lambda e^{-\lambda x} (1) \qquad , x > 0$$

$$f_X(x) = \lambda e^{-\lambda x} \qquad , x > 0$$

que es la distribución exponencial.

Para calcular probabilidades de una variable aleatoria con distribución Erlang, se debe integrar la función de densidad, o bien, contar con la función de distribución acumulada

$$F_X(x) = \int_0^x \lambda e^{-\lambda \tau} \frac{(\lambda \tau)^{r-1}}{(r-1)!} \, d\tau \qquad , x > 0$$

$$F_X(x) = 1 - \int_x^\infty \lambda e^{-\lambda \tau} \frac{(\lambda \tau)^{r-1}}{(r-1)!} \, d\tau \qquad , x > 0$$

La última integral se realiza por partes y se obtiene:

$$F_X(x) = 1 - \sum_{k=0}^{r-1} \frac{e^{-\lambda x} (\lambda x)^k}{k!} \qquad , x > 0$$

y debe observarse que la suma es precisamente la función de distribución acumulada de Poisson con parámetro λx, lo que en la práctica permite calcular probabilidades de una manera bastante sencilla o inclusive utilizar tablas de la función de distribución de Poisson.

El término $\sum_{k=0}^{r-1} \frac{e^{-\lambda x} (\lambda x)^k}{k!}$, es igual a $P(X > x)$, esto es

$$P(X > x) = \sum_{k=0}^{r-1} \frac{e^{-\lambda x} (\lambda x)^k}{k!}$$

tiene la siguiente interpretación: la probabilidad de que el tiempo hasta la r-ésima ocurrencia exceda cierto valor x es igual a la probabilidad de que el número de ocurrencias de Poisson en el tiempo x sea menor o igual que $r - 1$.

Teorema 5.4

Si X es una v.a. con distribución Erlang con parámetros λ y r, entonces su función de distribución es:

$$F_X(x) = \begin{cases} 1 - \sum_{k=0}^{r-1} \frac{e^{-\lambda x} (\lambda x)^k}{k!} & , \quad x > 0 \\ 0 & , \quad en\ otro\ caso \end{cases}$$

Ejemplo 5.4

La vida útil de un sistema electrónico esta dada por la suma de las vidas de sus tres componentes, esto es: $T = T_1 + T_2 + T_3$. Los tiempos de vida de los componentes son independientes y cada uno sigue una distribución exponencial con media de 4 horas. Determinar la probabilidad de que el sistema funcione por lo menos 16 horas.

Resolución

Puesto que se tiene una suma de variables aleatorias independientes con distribución exponencial, el tiempo de vida del sistema sigue una distribución Erlang

$$T = T_1 + T_2 + T_3 \sim Erlang\left(\lambda = \frac{1}{4},\ r = 3\right)$$

Entonces:

$$P(T > 16) = 1 - F_T(16)$$

Y del teorema 5.4

$$P(T > 16) = \sum_{k=0}^{3-1} \frac{e^{-\left(\frac{1}{4}\right)(16)}\left(\left(\frac{1}{4}\right)(16)\right)^k}{k!} = \sum_{k=0}^{2} \frac{e^{-4}(4)^k}{k!} = 13\,e^{-4}$$

$$P(T > 16) \approx 0.2381$$

DISTRIBUCIÓN GAMMA

Así como la distribución Binomial negativa es una generalización de la distribución de Pascal en la cual el parámetro r puede tomar cualquier valor real positivo, la distribución Erlang también se generaliza al caso en el que el parámetro r sea un real mayor que cero, y no solamente entero positivo. Cuando esto ocurre, se tiene la distribución Gamma, que se define utilizando los parámetros λ y r (escala y forma), al igual que la distribución Erlang. Las principales aplicaciones de la distribución Gamma, surgen de la Inferencia estadística en las pruebas de bondad de ajuste, cuando se tienen datos de una muestra y se desea modelarlos mediante una función de densidad, en estos casos, la densidad Gamma es muy útil debido a la versatilidad de sus parámetros.

Definición 5.3

Si X es una v.a. con distribución Gamma con parámetros $\lambda > 0$ y $r > 0$, entonces su función de densidad es:

$$f_X(x) = \begin{cases} \dfrac{\lambda}{\Gamma(r)} \, (\lambda x)^{r-1} \, e^{-\lambda x} \, , & x > 0 \\ 0 & , \; en \; otro \; caso \end{cases}$$

$X \sim Gamma(\lambda, r)$.

Donde $\Gamma(r)$ es la función Gamma valuada en r. La función $\Gamma(n)$ es una generalización del concepto de factorial y esta definida como:

$$\Gamma(n) = \int_0^\infty x^{n-1} \, e^{-x} \, dx$$

Tiene la propiedad de recurrencia

$$\Gamma(n) = (n-1) \, \Gamma(n-1)$$

y si n es un entero positivo, se puede calcular con el factorial

$$\Gamma(n) = (n-1)!$$

y además, $\Gamma(1) = 1$ y $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

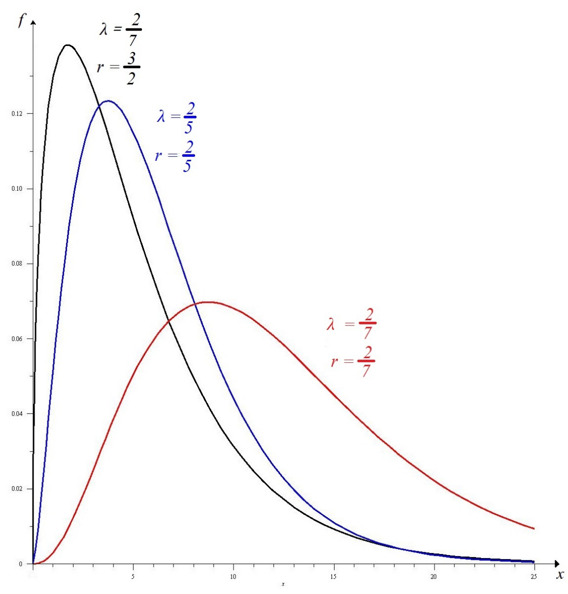


Fig. 5.3. Distribución Gamma

Teorema 5.5

Si X es una v.a. con distribución Gamma con parámetros $\lambda > 0$ y $r > 0$, entonces:

$$\mu_X = \mathrm{E}[X] = \frac{r}{\lambda}$$

$$\sigma_X^2 = \mathrm{Var}[X] = \frac{r}{\lambda^2}$$

$$M_X(\theta) = \left(\frac{\lambda}{\lambda - \theta}\right)^r = \left(1 - \frac{\theta}{\lambda}\right)^{-r}$$

Y puesto que la distribución Erlang es un caso particular de la Gamma, las expresiones para el cálculo de la media, la varianza y la función generadora de la distribución Gamma sirven para la distribución Erlang.

La función generadora de momentos se puede obtener multiplicando las funciones generadoras de momentos de cada una de las variables aleatorias exponenciales que se suman para construir la variable aleatoria Erlang.

También es común encontrar a la distribución Gamma escrita en términos de los parámetros α y β, que se relacionan con los parámetros r y λ de la siguiente forma.

Si $r = \alpha$ y $\lambda = \frac{1}{\beta}$, entonces la función de densidad Gamma puede escribirse como:

$$f_X(x) = \begin{cases} \dfrac{1}{\beta^\alpha\, \Gamma(\alpha)}\, x^{\alpha - 1}\, e^{-\frac{x}{\beta}} & , \quad x > 0 \\ 0 & , \quad \textit{en otro caso} \end{cases}$$

y sus parámetros serían

$$\mathrm{E}[X] = \alpha\beta$$

$$\mathrm{Var}[X] = \alpha\beta^2$$

Puesto que la integral de la función de densidad Gamma es bastante complicada, la gran mayoría de los ejercicios son resueltos con el caso particular de la distribución Erlang, y en ocasiones utilizando tablas de la distribución Gamma incompleta. Estas tablas se basan en un caso particular de la distribución Gamma (llamada distribución Gamma estándar), para el cual $\beta = \frac{1}{\lambda} = 1$, de tal forma que el único parámetro restante es r (o bien α), y la función de distribución de la Gamma estándar se llama función Gamma incompleta, y mediante un cambio de variable permite obtener las probabilidades de una variable aleatoria con distribución Gamma.

Teorema 5.6

Si X es una v.a. con distribución Gamma con parámetros $\lambda > 0$ y $r > 0$, entonces:

$$P(X \le x) = F_X(x\,;\,\lambda, r) = F_{Gi}(\lambda x\,;r)$$

donde $F_{Gi}(\lambda x\,;r)$ es la función Gamma incompleta.

Ejemplo 5.5

Sea X una v.a. con distribución Gamma con parámetros $r = 5.5$ y $\lambda = 1.5$, obtener $P(X \leq 3)$.

Resolución

Utilizando el teorema 5.6 se tiene:

$$P(X \leq 3) = F_X(3\ ;\ 1.5,\ \ 5.5) = F_{Gi}(4.5\ ; 5.5)$$

Y de tablas

Función Gamma incompleta

	r				
λ x	4	4.5	5	5.5	6
3.0	0.3528	0.2601	0.1847	[illegible]	0.0839
3.5	0.4634	0.3629	0.2746	[illegible]	0.1424
4.0	0.5665	0.4659	0.3712	[illegible]	0.2149
4.5	[illegible]	[illegible]	[illegible]	0.3781	0.2971
[illegible]	0.7350	0.6495	0.5595	0.4096	0.3840
5.5	0.7983	0.7243	0.6425	0.5567	0.4711
6.0	0.8488	0.7867	0.7149	0.6364	0.5543
6.5	0.8882	0.8374	0.7763	0.7067	0.6310

$$P(X \leq 3) = 0.3781$$

Ejemplo 5.6

Supóngase que el tiempo en el que un estudiante resuelve un ejercicio en cierto examen sigue una distribución Gamma con media 5.5 minutos y variancia de 5.5.

a) Determinar los parámetros λ y r de la distribución.

b) Obtener la probabilidad de que el estudiante resuelva el ejercicio en más de 11 minutos.

Resolución

a) Sea X el tiempo que tarda un estudiante en resolver el ejercicio, $X \sim Gamma(\lambda, r)$

$$\mathrm{E}[X] = \frac{r}{\lambda}$$

$$\mathrm{Var}[X] = \frac{r}{\lambda^2}$$

De donde

$$\frac{r}{\lambda} = 5.5$$

$$\frac{r}{\lambda^2} = 5.5$$

Por lo que

$$\lambda = 1$$

$$r = 5.5$$

b) $P(X > 11) = 1 - F_X(11\ ;\ 1,\ \ 5.5) = F_{Gi}(11\ ; 5.5)$

$= 1 - .9756$

$= 0.0244$

DISTRIBUCIÓN NORMAL

La *distribución normal* es una de las más utilizadas en la práctica. Muchos problemas reales tienen un comportamiento que se puede aproximar al de la distribución normal. Fue descubierta por Abraham de Moivre (1667-1754) en 1733 como una forma límite de la distribución binomial, después la estudió Laplace, aproximadamente en el año de 1775 y en ocasiones se le conoce como distribución gausiana debido a que Gauss la citó en un artículo en 1809.

Durante los siglos XVIII y XIX se hicieron esfuerzos para establecer el modelo normal como la ley básica que rige las variables aleatorias continuas: de ahí el nombre "normal". Estos esfuerzos fracasaron.

Definición 5.4

Si una variable aleatoria X tiene una función de densidad dada por:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma_X} e^{-\frac{1}{2}\frac{(x-\mu_X)^2}{\sigma_X^2}} \quad ; \quad -\infty < x < \infty$$

Donde μ_X y σ_X son constantes tales que $-\infty < \mu_X < \infty$, $\sigma_X > 0$, entonces X tiene una distribución normal con parámetros, media μ_X y varianza σ_X^2, lo cual se denota por:

$$X \sim N(\mu_X, \sigma_X^2)$$

Propiedades de la función de densidad normal

1) $f_X(x) > 0$

2) $\int_{-\infty}^{\infty} f_X(x)\ dx = 1$

3) $\lim_{x \to \infty} f_X(x) = 0, \quad \lim_{x \to -\infty} f_X(x) = 0$

4) $f_X(\mu + x) = f_X(\mu - x)$ Es simétrica

5) El máximo ocurre en $x = \mu$

6) Puntos de inflexión en $x = \mu \pm \sigma$

7) Curtosis igual a 3

Puesto que la distribución normal es simétrica coinciden la media, la mediana y la moda en el mismo valor.

Por otro lado, la curtosis de la distribución normal es 3, y es por eso que la curtosis de cualquier variable se compara contra dicho valor.

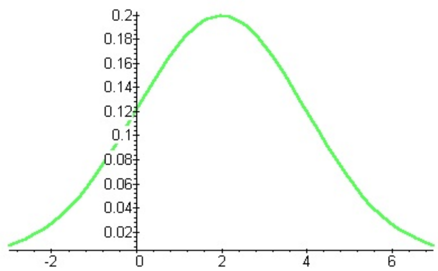


Fig. 5.4. Distribución normal

La gráfica anterior muestra una curva normal con parámetros media 2 y variancia 4.

Para obtener la probabilidad de que la variable aleatoria X se encuentre entre a y b, es necesario resolver la integral:

$$\int_a^b \frac{1}{\sqrt{2\pi}\,\sigma_X} e^{-\frac{1}{2}\frac{(x-\mu_X)^2}{\sigma_X^2}} \, dx$$

que equivale a obtener el área bajo la curva normal; sin embargo, la integral requiere de una función especial, llamada función de error, que no se estudia en cursos básicos de matemáticas, por lo que es más común evaluarla utilizando métodos numéricos o tablas de la *distribución normal estándar*.

Una variable normal estandarizada es aquella que tiene media 0 y variancia 1.

Teorema 5.7 Aditividad de la Distribución Normal

Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes, y todos con distribución normal con media μ_i y variancia σ_i^2, entonces la variable aleatoria Y definida como

$$Y = X_1 + X_2 + \ldots + X_n = \sum_{i=1}^{n} X_i$$

tiene distribución normal con media $\sum_{i=1}^{n} \mu_i$ y variancia $\sum_{i=1}^{n} \sigma_i^2$.

DISTRIBUCIÓN NORMAL ESTÁNDAR

La distribución normal estándar es un caso particular de la distribución normal, la cual tiene como parámetros 0 y 1, es decir, tiene una media de cero y una variancia de 1. Se denota mediante $Z \sim N(0,1)$. El procedimiento para obtener una variable aleatoria con distribución normal estándar a partir de una variable aleatoria con distribución normal y parámetros cualesquiera es mediante un corrimiento y un escalamiento, lo que lleva a la expresión:

$$Z = \frac{X - \mu_X}{\sigma_X}$$

donde $X \sim N(\mu_X, \sigma_X^2)$.

La figura siguiente muestra una curva normal estándar.

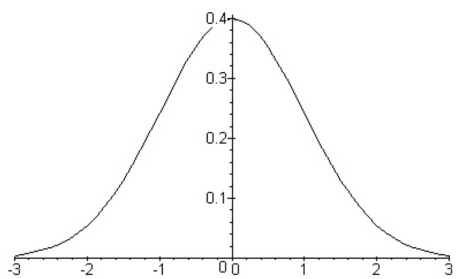


Fig. 5.5 Distribución normal estándar

La tabla del apéndice 1 proporciona las probabilidades que se obtienen con la función de distribución acumulativa normal estándar (áreas bajo la curva normal estándar de cola izquierda). Por ejemplo, la probabilidad de que X sea menor que 1.66, $P(X < 1.66)$, es equivalente a encontrar el área bajo la curva normal de menos infinito a 1.66. Esta probabilidad se obtiene directamente de la tabla. La primera columna proporciona enteros y un decimal mientras que el primer renglón proporciona el segundo decimal, así, para encontrar el 1.66 se localiza en la primera columna el 1.6 y sobre ese renglón se cruza con la columna cuyo encabezado es 0.06 obteniéndose el valor de 0.9515.

Al obtener cualquier probabilidad que no proporcione directamente la tabla, deberá recordarse que la distribución normal estándar es simétrica con respecto al origen, y que el área total bajo la curva es la unidad.

Ejemplo 5.7

Sea Z una variable aleatoria con media 0 y desviación estándar 1, obtener:

a) $P(Z \leq -1.84)$

b) $P(-0.86 \leq Z \leq 1.97)$

Resolución

Utilizando tablas de distribución normal estándar

a) $P\ (Z \le -1.84) = 0.0329$
Que se obtiene directamente de la tabla.

b) $P\ (-0.86 \le Z \le 1.97) = 0.7807$
Que se obtiene al realizar una diferencia de los valores encontrados en la tabla, para 1.97 y -0.86.

Ejemplo 5.8

Los pesos de un número grande de perros de lana miniatura están distribuidos aproximadamente en forma normal con una media de 8 kilogramos y una desviación estándar de 0.9 kilogramos. Si se registran las mediciones y se cierran a décimas de kilogramo, encontrar la fracción de estos perros de lana con pesos

a) arriba de 9.5 kilogramos;
b) cuando mucho 8.6 kilogramos;
c) entre 7.3 y 9.1 kilogramos inclusive.

Resolución

a) $X \triangleq$ es la variable aleatoria que representa el peso de los perros de lana miniatura

$X \sim N(8\ ,\ 0.9^2)$

$$z = \frac{(9.5 - 8)}{0.9} \approx 1.67$$ (Debe redondearse a dos decimales)

La fracción de perros de lana miniatura con peso por encima de 9.5 kilogramos es:

$$P\ (X > 9.5) \approx P\ (Z > 1.67) = 1 - 0.9525 = 0.0475$$

b) $$z = \frac{(8.6 - 8)}{0.9} \approx 0.67$$ (Obsérvese nuevamente el redondeo)

La fracción de perros de lana miniatura que pesan cuando mucho 8.6 kilogramos es:

$$P\ (X < 8.6) \approx P\ (Z < 0.67) = 0.7486$$

c) $$z_1 = \frac{(7.3 - 8)}{0.9} = -0.78 \quad , \quad z_2 = \frac{(9.1 - 8)}{0.9} = 1.22$$

La fracción de perros de lana miniatura que pesan entre 7.3 y 9.1 kilogramos es:

$$P\ (7.3 < X < 9.1) \approx P\ (-0.78 < Z < 1.22)$$
$$P\ (7.3 < X < 9.1) \approx 0.8888 - 0.2177 = 0.6711$$

Teorema 5.8

Sea X una variable aleatoria normal, $X \sim N(\mu_X, \sigma_X^2)$, entonces la función generadora de momentos está dada por:

$$M_X(\theta) = e^{\mu_X \theta + \frac{1}{2}\sigma_X^2 \theta^2}$$

Demostración

Si X es una variable aleatoria con distribución normal y parámetros μ_X y σ_X^2 entonces

$X \sim N\left(\mu_X, \sigma_X^2\right)$

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2} \qquad ; \qquad -\infty < x < \infty$$

Y la función generadora de momentos es

$$M_X(\theta) = E\left[e^{\theta X}\right]$$

$$= \int_{-\infty}^{\infty} \frac{e^{\theta x}}{\sqrt{2\pi}\,\sigma_X} e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2} dx$$

$$= \frac{1}{\sqrt{2\pi}\,\sigma_X} \int_{-\infty} e^{\theta x} e^{-\frac{1}{2}\left(\frac{x-\mu_X}{\sigma_X}\right)^2} dx$$

Realizando el cambio de variable $z = \dfrac{x-\mu_X}{\sigma_X}$

$$M_X(\theta) = \frac{1}{\sqrt{2\pi}} e^{\mu_X \theta} \int_{-\infty}^{\infty} e^{\theta \sigma_X z} e^{-\frac{z^2}{2}} dz$$

$$= \frac{e^{\mu_X \theta}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{z^2 - 2\theta\sigma_X z}{2}} dz$$

$$= \frac{e^{\mu_X \sigma}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(z-\theta\sigma_X)^2}{2} + \frac{\theta^2 \sigma_X^2}{2}} dz$$

$$= e^{\mu_X \theta + \frac{\sigma_X^2 \theta^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(z-\theta\sigma_X)^2}{2}} dz$$

y se observa que la integral junto con el coeficiente $\frac{1}{\sqrt{2\pi}}$, es la integral de una función de densidad normal con media $\theta\sigma_X$ y variancia uno, por lo que su valor es 1.

Finalmente la función generadora es:

$$\boxed{\mu_X(\theta) = e^{\mu_X\theta + \frac{\sigma_X^2\theta^2}{2}}}$$

□

APROXIMACIÓN DE LA DISTRIBUCIÓN NORMAL A LA BINOMIAL

Las probabilidades asociadas con una distribución binomial, pueden obtenerse fácilmente cuando n es pequeña, utilizando la expresión

$$X \sim Binomial\,(n, p)$$

$$P(X = x) = \binom{n}{x} p^x q^{n-x}$$

pero si n es grande, el cálculo se complica. Una forma de resolver estos problemas es mediante la aproximación de Poisson, vista en el tema anterior, pero tiene el inconveniente de que n debe ser grande y p pequeña. Ahora se estudiará otra forma de aproximar probabilidades binomiales utilizando la distribución normal.

Teorema 5.9

Si X es una v.a. con distribución binomial con parámetros $\mu_X = np$ y $\sigma_X^2 = npq$, entonces la forma límite de la distribución de

$$Z = \frac{X - np}{\sqrt{npq}}$$

cuando $n \to \infty$, es la distribución normal estándar $Z \sim N(0, 1)$.

En general, la aproximación es adecuada cuando

$$np > 5 \qquad \text{y} \qquad p \leq \frac{1}{2}$$

o bien

$$nq > 5 \qquad \text{y} \qquad p > \frac{1}{2}$$

Una forma de ilustrar la aproximación, es mediante el histograma para la distribución binomial.

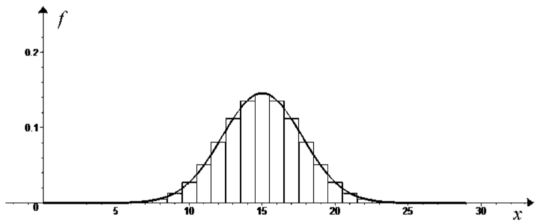


Fig. 5.6. Aproximación de la distribución normal a la binomial

Finalmente, para realizar la aproximación, es necesario realizar un ajuste por continuidad, es decir, la variable aleatoria binomial es discreta, mientras que la variable aleatoria normal es continua, por lo que debe realizarse un pequeño ajuste para mejorar la aproximación.

Si para una variable aleatoria X con distribución binomial se desea calcular la probabilidad de que este entre x_1 y x_2, el ajuste por continuidad está dado por:

$$P(x_1 \le X \le x_2) \approx P\left(\frac{x_1 - d - \mu}{\sigma} \le Z \le \frac{x_2 + d - \mu}{\sigma} \right)$$

donde d es una diferencial de media unidad, *i.e.* representa media unidad de las estudiadas.

Ejemplo 5.9

Si 20% de los residentes de cierta ciudad prefieren un teléfono blanco que cualquier otro color disponible, determinar la probabilidad de que entre los siguientes 1000 teléfonos que se instalen en esta ciudad

a) entre 170 y 185 inclusive sean blancos; y

b) al menos 210 pero no más de 225 sean blancos.

Resolución

$\mu = np = 1000(0.2) = 200$

$\sigma = \sqrt{npq} = \sqrt{200(0.8)} = 12.469$

Utilizando aproximación normal, y con la variable aleatoria X que represente al número de teléfonos blancos que se instalan.

a) $P(170 \le X \le 185) \approx P\left(\frac{170 - 0.5 - 200}{12.649} \le Z \le \frac{185 + 0.5 - 200}{12.649} \right)$

$P(-2.41 < Z \le -1.15) = 0.1171$

b) $P(210 \le X \le 225) \approx P\left(\frac{210 - 0.5 - 200}{12.649} \le Z \le \frac{225 + 0.5 - 200}{12.649} \right)$

$P(0.75 < Z \le 2.02) = 0.2049$

DISTRIBUCIÓN Ji CUADRADA

Por las características especiales que presenta la distribución normal, se han estudiado distribuciones que puedan generarse a partir de ella, tal es el caso de la distribución Ji cuadrada, también llamada chi-cuadrada.

Definición 5.5

Sean $Z_1, Z_2, \ldots, Z_\nu$; ν variables aleatorias independientes con distribución normal estándar, entonces:

$$X^2 = Z_1^2 + Z_2^2 + \ldots + Z_\nu^2$$

es una variable aleatoria con función de densidad

$$f_{X^2}(x) = \begin{cases} \dfrac{1}{2^{\frac{\nu}{2}}\,\Gamma\left(\frac{\nu}{2}\right)}\, x^{\left(\frac{\nu}{2}-1\right)}\, e^{-\frac{x}{2}} & x > 0 \\ 0 & \textit{en otro caso} \end{cases}$$

que recibe el nombre de Ji cuadrada con ν grados de libertad. Se denota mediante el símbolo $\chi^2_{(\nu)}$.

Debe observarse que la distribución Ji cuadrada es un caso particular de la distribución Gamma, en donde $\lambda = \frac{1}{2}$ y $r = \frac{\nu}{2}$.

Por otro lado, también debe observarse la notación. La distribución se llama ji-cuadrada, ji es la letra griega que se utiliza para denotarla, y se escribe al cuadrado para recordar que proviene de una suma de cuadrados. La ji minúscula es χ, mientras que la mayúscula es X. La variable aleatoria ji cuadrada será entonces X^2, y un valor de la variable aleatoria será χ^2, mientras que para denotar el tipo de distribución, se utiliza la ji minúscula pero agregando en un subíndice los grados de libertad entre paréntesis, para no confundirse con un valor de la variable aleatoria. En general los grados de libertad se denotan con la letra griega ν (ni). Esto es $X^2 \sim \chi^2_{(\nu)}$, que se lee: la variable aleatoria ji-cuadrada tiene distribución ji cuadrada con ni grados de libertad. Mientras que la expresión $X^2 \geq \chi^2$ se lee: la variable aleatoria ji cuadrada es mayor o igual que cierto valor ji cuadrada. En resumen:

X^2 Variable aleatoria ji cuadrada.

χ^2 Valor que toma la variable aleatoria ji cuadrada.

$\chi^2_{(\nu)}$ Distribución ji cuadrada con ν (ni) grados de libertad.

La distribución Ji cuadrada, al igual que la distribución normal, presenta la característica de aditividad.

Teorema 5.10 Aditividad de la distribución Ji cuadrada

Sean $X_1^2, X_2^2, \ldots, X_n^2$ variables aleatorias independientes con distribución Ji cuadrada y $\nu_1, \nu_2, \ldots, \nu_n$ grados de libertad, respectivamente. Entonces la variable

$Y = X_1^2 + X_2^2 + \ldots + X_n^2 = \sum_{i=1}^{n} X_i^2$ tiene una distribución Ji cuadrada con $\nu = \sum_{i=1}^{n} \nu_i$ grados de libertad.

La demostración del teorema anterior resulta como consecuencia directa de la definición de la distribución Ji cuadrada, que es una suma de variables aleatorias con distribución normal estándar al cuadrado; por lo que una suma de Ji cuadradas sigue siendo una suma de normales estándar al cuadrado lo que resulta en una Ji cuadrada.

Teorema 5.11

Si X^2 es una variable aleatoria que tiene una distribución Ji cuadrada con ν grados de libertad, $X^2 \sim \chi^2_{(\nu)}$, entonces

$$E[X^2] = \nu \quad , \quad Var[X^2] = 2\nu$$

y la función generadora de momentos de X^2 es:

$$M_{X^2}(\theta) = (1 - 2\theta)^{-\frac{\nu}{2}}$$

La distribución Ji cuadrada presenta un sesgo positivo, según se muestra en la siguiente figura.

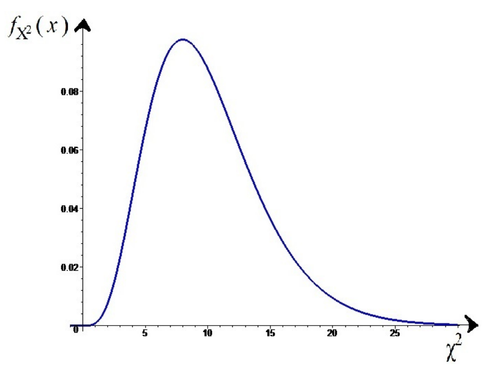


Fig. 5.7 Distribución Ji cuadrada

Para calcular probabilidades con distribución Ji cuadrada se común utilizar tablas. Las tablas de la distribución Ji-cuadrada generalmente proporcionan el valor de χ^2 en función del área de "cola derecha" α, y el valor de Ji-cuadrada se denota χ^2_α cuando $P\left(X^2 \geq \chi^2\right) = \alpha$, como se observa en la siguiente gráfica.

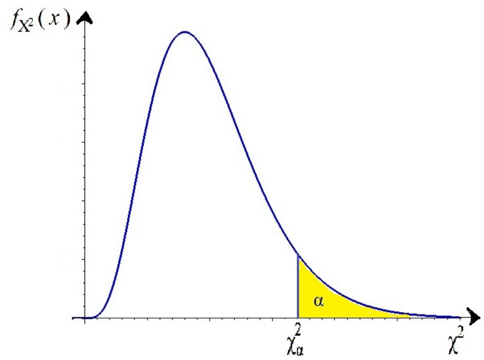


Fig. 5.8 Probabilidad de cola derecha

Con lo que se agrega la notación:

χ^2_α Valor de ji-cuadrada para el cual la probabilidad de cola derecha es α.

Ejemplo 5.10
Utilizar las tablas de la distribución Ji-cuadrada para obtener las siguientes probabilidades

a) Si $X^2 \sim \chi^2_{(15)}$, obtener $P\left(X^2 \geq 15.73\right)$

b) Si $X^2 \sim \chi^2_{(20)}$, **Resolución**

a) De tablas de la distribución i-cuadrada, y observando que las tablas proporcionan probabilidades de "cola derecha" o de lado derecho, entonces, se busca en el renglón el número de grados de libertad y posteriormente el valor de $\chi^2 = 15.73$, es decir:

Distribución Ji cuadrada

	α					
ν	**0.9**	**0.5**	**0.4**	**0.3**	**0.2**	**0.1**
12	6.304	11.340	12.584	14.011	15.812	18.549
13	7.042	12.340	13.636	15.119	16.985	19.812
14	7.790	13.339	14.685	16.222	18.151	21.064
15	8.547	14.339	$\boxed{15.733}$	17.322	19.311	22.307
16	9.312	15.339	16.780	18.418	20.465	23.542
17	10.085	16.338	17.824	19.511	21.615	24.769
18	10.865	17.338	18.868	20.601	22.760	25.989
19	11.651	18.338	19.910	21.689	23.900	27.204
20	12.443	19.337	20.951	22.775	25.037	28.412
21	13.240	20.337	21.992	23.858	26.171	29.615
22	14.041	21.337	23.031	24.939	27.301	30.813

por lo que la probabilidad es: $P\left(X^2 \geq 15.73\right) = 0.4$

b) De forma similar, pero utilizando el complemento

$$P\left(X^2 < 12.44\right) = 1 - P\left(X^2 \geq 12.44\right) = 1 - 0.9 = 0.1$$

DISTRIBUCIÓN *t* de STUDENT

El matemático inglés William S. Gosset (1876-1937) trabajó en la cervecería Guinness y publicó varios artículos bajo el seudónimo “Student”, motivo por el cual la distribución que desarrolló lleva ese nombre.

Definición 5.6

Sean Z y X^2 dos variables aleatorias independientes con distribuciones normal estándar y Ji-cuadrada respectivamente, es decir: $Z \sim N(0,1)$, $X^2 \sim \chi^2_{(\nu)}$

entonces la variable aleatoria T definida como $T = \dfrac{Z}{\sqrt{\dfrac{X^2}{\nu}}}$

tiene una distribución t de Student con ν grados de libertad y función de densidad dada por

$$f_T(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi\nu}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} \qquad -\infty < t < \infty$$

$$\nu > 0$$

Y se denota por $T \sim t_{(\nu)}$.

Como puede observarse la distribución t de Student posee el parámetro ν, que al igual que para la distribución ji-cuadrada, recibe el nombre de grados de libertad. La notación es similar a la de la ji-cuadrada:

T	Es la variable aleatoria con t de Student.
t	Es un valor que toma la variable aleatoria t de Student.
$t_{(\nu)}$	Es la distribución t de Student con ν (ni) grados de libertad.

Teorema 5.12

Si T es una variable aleatoria que tiene distribución t de Student con ν grados de libertad, entonces

$$E[T] = 0$$

$$Var[T] = \frac{\nu}{\nu - 2} \qquad \nu > 2$$

La función de densidad t de Student es simétrica y unimodal al igual que la normal, pero siempre está centrada en cero; es muy parecida a la distribución normal estándar y tiende a ella cuando el número de grados de libertad tiende a infinito.

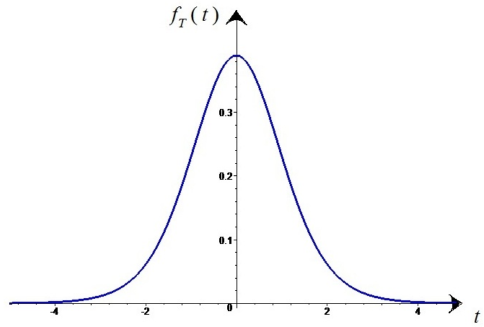


Fig 5.9. Distribución t de Student

Los valores de la distribución t se obtienen de tablas. La gran mayoría de las tablas proporcionan valores de cola derecha, al igual que la distribución Ji-cuadrada.

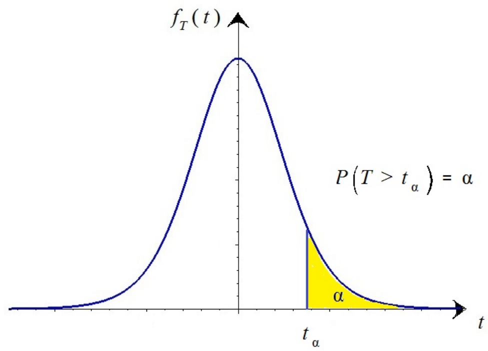


Fig 5.10. Probabilidades de cola derecha, t de Student

Ejemplo 5.11

Sea T una variable aleatoria con distribución t de Student, con 19 grados de libertad, obtener:

a) $P(T > 1.33)$

b) El valor de t, tal que $P(T < t) = 0.01$

Resolución

a) Del enunciado $T \sim t_{(19)}$, por lo que, directamente de tablas en el renglón de 19 grados de libertad se busca el valor más cercano a 1.33 y se lee en la parte superior el valor de α, teniéndose: $P(T > 1.33) = 0.1$

b) De tablas, y recordando que la distribución t es simétrica, se tiene que
$0.01 = P(T < t) = P(T > -t)$
Al localizar $\alpha = 0.01$ y $\nu = 19$ se obtiene que: $-t = 2.54$
Finalmente: $t = -2.54$

DISTRIBUCIÓN F DE FISHER

El estadístico inglés Ronald A. Fisher (1890-1962) y el matemático estadounidense George Snedecor (1881-1974) desarrollaron la distribución F, tambien llamada F de Fisher, F de Snedecor o F de Fisher-Snedecor.

La distribución F tiene su principal aplicación cuando se desea hacer inferencias con respecto a las variancias de dos distribuciones normales independientes a partir de muestras aleatorias de cada distribución, como se verá en el curso de Estadística.

Definición 5.7

Si X y Y son dos variables aleatorias independientes con distribuciones ji cuadrada con parámetros u y v, es decir:

$$X \sim \chi^2_{(u)} \quad , \quad Y \sim \chi^2_{(v)}$$

Entonces la variable aleatoria F definida como

$$F = \frac{\frac{X}{u}}{\frac{Y}{v}}$$

tiene una distribución F de Fisher con u grados de libertad en el numerador, v grados de libertad en el denominador y función de densidad dada por

$$f_F(f) = \begin{cases} \dfrac{\Gamma\left(\dfrac{u+v}{2}\right) u^{\frac{u}{2}} v^{\frac{v}{2}}}{\Gamma\left(\dfrac{v}{2}\right)\Gamma\left(\dfrac{u}{2}\right)} f^{\frac{v-2}{2}} (v + uf)^{-\frac{u+v}{2}} & f > 0 \\ 0 & en \quad otro \quad caso \end{cases}$$

Se denota $F \sim F_{(u,v)}$.

La notación es:

- F Es la variable aleatoria con distribución F.
- f Es un valor que toma la variable aleatoria con distribución F.
- $F_{(u,v)}$ Distribución F con u grados de libertad den el numerador y v grados de libertad en el denominador.

La media y la variancia de la distribución F se proporcionan en el siguiente teorema.

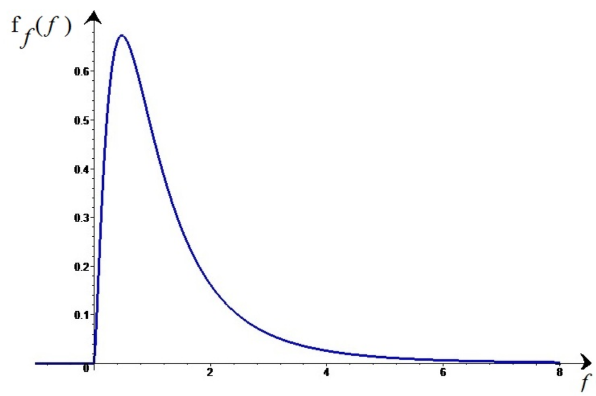
Teorema 5.13

Si F es una variable aleatoria con distribución F de Fisher con u grados de libertad en el numerador y v grados de libertad en el denominador, entonces:

$$\mathrm{E}[F] = \frac{v}{v-2} \qquad v > 2$$

$$\mathrm{Var}[F] = \frac{2v^2(u+v-2)}{u(v-2)^2(v-4)} \qquad v > 4$$

La distribución F es asimétrica y sesgada hacia la derecha (sesgo positivo) según se observa en la siguiente figura



Para obtener probabilidades con la distribución F, se utilizan tablas que proporcionan puntos porcentuales para distintos valores de u y v. Las tablas proporcionan los valores de f como una función de u, v y α según la expresión $P(F \geq f_\alpha) = \alpha$ y cuya interpretación gráfica corresponde al valor sobre el eje, para el cual se obtiene una probabilidad de cola derecha igual a α.

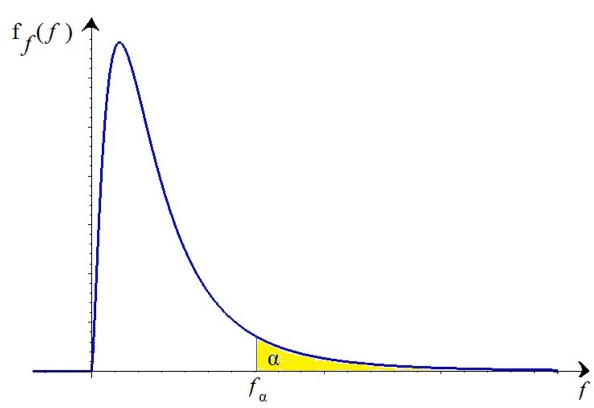


Fig. 5.12. Valores de α de la distribución F

Los puntos porcentuales de la cola inferior (o izquierda) se obtienen mediante la expresión

$$f_{1-\alpha,\,(u,v)} = \frac{1}{f_{\alpha,\,(v,u)}}$$

puesto que si $F \sim F_{(u,v)}$, entonces $\frac{1}{F} \sim F_{(v,u)}$.

Ejemplo 5.12

Obtener el valor de f para el cual

$$P\,(F \le f_{\alpha,\,(5,10)}) = 0.9$$

Resolución

De tablas y recordando que

$P\,(F \le f_{\alpha,(5,10)}) = 1 - P\,(F > f_{0.1,(5,10)}) = 0.9$

Se tiene: $f = 2.52$

Debe observarse que el valor de α se conoce desde el principio, y es 0.1, pero debe utilizase el teorema del complemento para poder utilizar las tablas.

Algunas tablas proporcionan el valor de f como una función de $1-\alpha$, u y v, según la expresión

$$P\,(F \le f_{1-\alpha,\,(u,v)}) = 1 - \alpha$$

por lo que debe ponerse especial atención en el tipo de tabla que se está manejando.

DISTRIBUCIÓN WEIBULL

La distribución Weibull toma su nombre en honor del matemático sueco Waloddi Weibull (1887-1979), quien la describió detalladamente y aplicó en el área de fatiga de materiales (1951), aunque históricamente se le atribuye al francés Maurice René Fréchet (1927).

Definición 5.8

Si X es una variable aleatoria con distribución Weibull, con parámetros $\lambda > 0$ y $r > 0$, entonces su función de densidad está dada por:

$$f_X(x) = \begin{cases} \lambda\, r\,(\lambda\, x)^{r-1}\, e^{-(\lambda x)^r} & , \quad x \ge 0 \\ 0 & , \quad \textit{en otro caso} \end{cases}$$

$X \sim \textit{Weibull}\,(\lambda, r)$.

Los parámetros λ y r nuevamente determinan la escala y la forma, respectivamente.

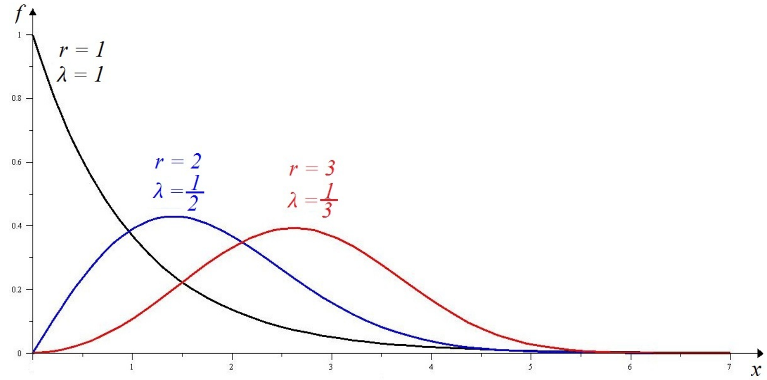


Fig. 5.13. Distribución Weibull

Debe observarse que si $r = 1$, entonces la distribución Weibull se reduce a la distribución exponencial, por lo que no debe sorprender que la distribución Weibull, también se use para modelar tiempo entre ocurrencias, aunque la Weibull se utiliza cuando la tasa de ocurrencias es proporcional a una potencia del tiempo.

En la literatura también se puede encontrar la densidad Weibull escrita mediante los parámetros α y β, con $\alpha = r$ y $\beta = \frac{1}{\lambda}$, obteniendose la siguiente expresión:

$$f_X(x) = \begin{cases} \dfrac{\alpha}{\beta^\alpha} x^{\alpha - 1} e^{-\left(\frac{x}{\beta}\right)^\alpha} & , \quad x \geq 0 \\ 0 & , \quad en\ otro\ caso \end{cases}$$

Teorema 5.14

Si X es una variable aleatoria con distribución Weibull, con parámetros $\lambda > 0$ y $r > 0$, entonces:

$$\mathrm{E}[X] = \frac{1}{\lambda} \Gamma\left(1 + \frac{1}{r}\right)$$

$$\mathrm{Var}[X] = \frac{1}{\lambda^2} \left\{ \Gamma\left(1 + \frac{2}{r}\right) - \left[\Gamma\left(1 + \frac{1}{r}\right) \right]^2 \right\}$$

La media y la varianza de la variable aleatoria Weibull requieren de la función Gamma, lo que dificulta su cálculo; sin embargo, las probabilidades se obtienen con facilidad utilizando la función de distribución.

Teorema 5.15

Si X es una variable aleatoria con distribución Weibull, con parámetros $\lambda > 0$ y $r > 0$, entonces su función de distribución está dada por:

$$F_X(x) = \begin{cases} 1 - e^{-(\lambda x)^r} & , \quad x \geq 0 \\ 0 & , \quad en\ otro\ caso \end{cases}$$

Ejemplo 5.13

Considérese que el tiempo de falla en semanas de un dispositivo electrónico sigue una distribución Weibull, con parámetros $\lambda = \frac{1}{10}$ y $r = 2$.

a) ¿Cuál es el tiempo esperado de falla?
b) Determinar la probabilidad de que el dispositivo falle después de 9 semanas.

Resolución

a) Puesto que $X \sim \textit{Weibull}\left(\lambda = \frac{1}{10}, r = 2\right)$, el valor esperado está dado por:

$$E[X] = \frac{1}{\lambda}\Gamma\left(1 + \frac{1}{r}\right) = 10\,\Gamma\left(\frac{3}{2}\right)$$

De las propiedades de la función Gamma

$$\Gamma(n) = (n-1)\,\Gamma(n-1)$$

$$\Gamma\left(\frac{3}{2}\right) = \left(\frac{3}{2}-1\right)\Gamma\left(\frac{3}{2}-1\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}$$

$$E[X] = \frac{1}{\lambda}\Gamma\left(1 + \frac{1}{r}\right) = 10\,\Gamma\left(\frac{3}{2}\right)$$

Por lo que

$$E[X] = 10\,\frac{\sqrt{\pi}}{2} \approx 8.8623$$

8.86 semanas en promedio

b) $$P(X>9) = 1 - F_X(9) = 1 - \left(1 - e^{-\left(\frac{9}{10}\right)^2}\right) = e^{-\frac{81}{100}} \approx 0.4449$$

DISTRIBUCIÓN LOGNORMAL

Definición 5.9

Si X es una variable aleatoria positiva con distribución

$$f_X(x) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\,x\,\sigma}\, e^{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}} & , \quad x > 0 \\ 0 & , \quad \textit{en otro caso} \end{cases}$$

Entonces X tiene una distribución lognormal con parámetros μ y σ, y está relacionada con la distribución normal mediante el cambio de variable $Y = \ln(X)$, en donde Y tiene distribución normal.

Se denota $X \sim \textit{lognormal}(\mu, \sigma)$.

Debe observarse que los parámetros μ y σ, son la media y la desviación estándar de la variable normal $Y = \ln(X)$, mientras que la media y la varianza para la distribución lognormal están dadas en el siguiente teorema.

Teorema 5.16

Si X es una variable aleatoria lognormal con parámetros μ y σ, entonces:

$$\mathrm{E}[X] = e^{\mu + \frac{\sigma^2}{2}}$$

$$\mathrm{Var}[X] = \left(e^{\sigma^2} - 1\right) e^{2\mu + \sigma^2}$$

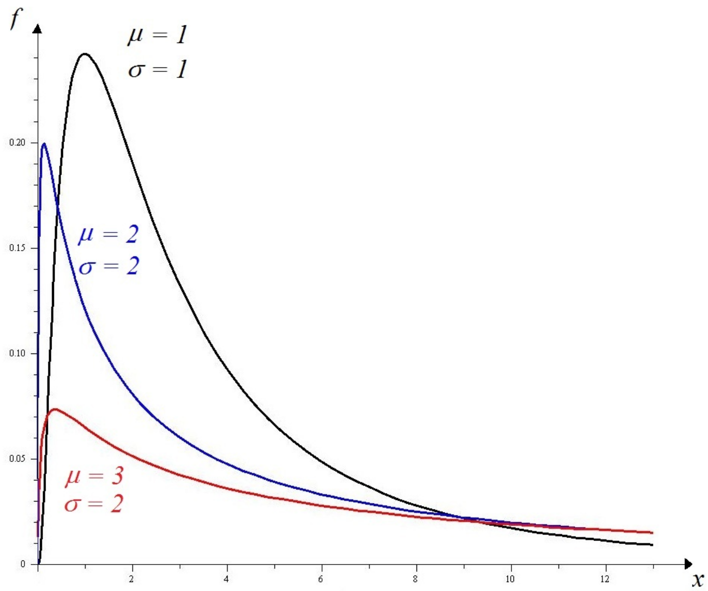


Fig. 5.14. Distribución Lognormal

Para obtener probabilidades de variables aleatorias con distribución lognormal, simplemente debe realizarse un cambio de variable para transformar a una variable aleatoria normal, y posteriormente estandarizar.

Teorema 5.17

Si X es una variable aleatoria lognormal con parámetros μ y σ; y Z es una variable aleatoria normal estándar, entonces:

$$P(X \le x) = P\left(\ln(X) \le \ln(x)\right) = P\left(Z \le \frac{\ln(x) - \mu}{\sigma}\right)$$

Ejemplo 5.14

La vida útil en horas de cierta broca industrial sigue una distribución lognormal con parámetros $\mu = 3$ y $\sigma = 1$.
Determinar:
a) La media y la desviación estándar de la vida útil de la broca.
b) La probabilidad de que la duración sea a lo sumo 40 horas.

Resolución

a) Sea X la variable aleatoria que represente la vida útil en horas de cierta broca industrial.

$X \sim \textit{lognormal}\,(\mu = 3\,,\sigma = 1)$

$$\mathrm{E}(X) = e^{3+\frac{1^2}{2}} = e^{\frac{7}{2}} = \approx 33.16$$

33.16 horas en promedio
Y la desviación estándar es:

$$\sigma_X = \sqrt{\mathrm{Var}(X)} = \sqrt{\left(e^{1} - 1\right) e^{2(3)+1^2}} = \sqrt{(e-1)\,e^{7}} \approx 43.41$$

b) $$P(X \leq 40) = P\big(\ln(X) \leq \ln(40)\big) = P\left(Z \leq \frac{\ln(40) - 3}{1}\right)$$

$$= P(Z \leq 0.69) \approx 0.7549$$

GENERACIÓN DE NÚMEROS ALEATORIOS Y MÉTODO DE LA TRANSFORMADA INVERSA

La generación de números aleatorios permite la simulación digital, que es una técnica muy utilizada en la toma de decisiones cuando los modelos probabilísticos son muy complejos para obtener la solución analítica.

El desarrollo tecnológico de los sistemas computacionales ha permitido que la simulación de procesos sea cada vez más utilizada. En esta sección se mostrará la forma de generar variables aleatorias que sigan una distribución en particular, utilizando el método de la transformada inversa, y posteriormente se ilustrarán unas aplicaciones muy sencillas de simulación.

Método de la Transformada inversa

El método de la transformada inversa consiste en relacionar la función de distribución de una variable aleatoria con un número aleatorio con distribución uniforme entre cero y uno. De las propiedades de la función de distribución se sabe que

$$0 \le F_X(x) \le 1 \quad , \quad \forall x$$

por lo que se iguala el número aleatorio con distribución uniforme entre cero y uno con la función de distribución acumulativa para despejar de esa ecuación la variable aleatoria X, que tendrá la distribución buscada.

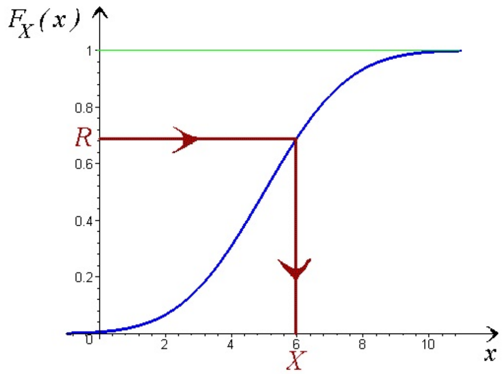


Fig. 5.15. Método de la transformada inversa

Esto es, se iguala

$$r = F_X(x)$$

y de ahí se despeja a X , teniéndose

$$x = F_X^{-1}(r)$$

y al escribir a x como variable aleatoria se tiene

$$X = F_X^{-1}(R)$$

Ejemplo 5.15
Obtener una expresión para generar números aleatorios con distribución exponencial utilizando el método de la transformada inversa.

Resolución
Si $X \sim \ exponencial(\lambda)$
entonces, su función de densidad es

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & ; \quad x > 0 \\ 0 & ; \quad en\ otro\ caso \end{cases}$$

y su función de distribución acumulativa es

$$F_X(x) = \begin{cases} 0 & ; \quad x \le 0 \\ 1 - e^{-\lambda x} & ; \quad x > 0 \end{cases}$$

Al utilizar la variable aleatoria R', con $R' \sim \ uniforme(0\,,1)$ se tiene:
$r' = 1 - e^{-\lambda x}$
De donde

$r' - 1 = -e^{-\lambda x}$

$1 - r' = e^{-\lambda x}$

pero si $R' \sim \textit{uniforme}(0\,,1)$, entonces $1 - R'$ es su simétrico en el intervalo de cero a uno y también es una variable aleatoria uniforme, por lo que $R = 1 - R' \sim \textit{uniforme}(0\,,1)$ y se puede escribir:

$r = e^{-\lambda x}$

de donde

$x = -\dfrac{1}{\lambda} \ln r$

y escribiendo en términos de variables aleatorias

$X = -\dfrac{1}{\lambda} \ln R$

El método de la transformada inversa también puede aplicarse para variables aleatorias discretas. El siguiente ejemplo muestra la forma de hacerlo.

Ejemplo 5.16

Un vendedor de periódicos sabe que la demanda diaria de los periódicos que tiene para venta está dada por:

x	0	1	2	3	4	5
$f_X(x)$	0.1	0.2	0.3	0.2	0.15	0.05

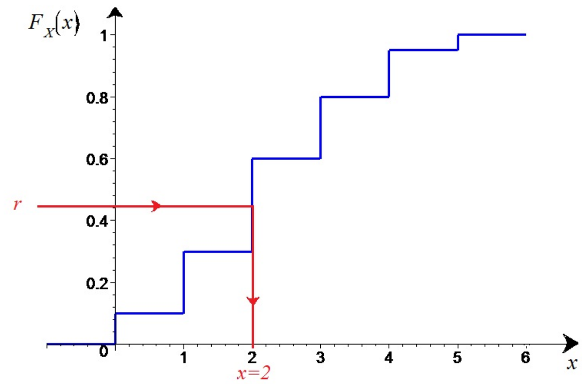
Utilizar los números aleatorios 0.4501, 0.0364, 0.5778, 0.8066 y 0.9591 para determinar la demanda de cada uno de los siguientes 5 días.

Resolución

La función de distribución acumulada de la demanda diaria está dada por

x	0	1	2	3	4	5
$F_X(x)$	0.1	0.3	0.6	0.8	0.95	1

Utilizando la función de distribución acumulada, e ilustrando para $r = 0.4501$ se tiene:



Por lo que en el primer día la demanda es de 2 periódicos.
La regla de asignación es:

$$x = \begin{cases} 0 & ; \quad 0 \leq r \leq 0.1 \\ 1 & ; \quad 0.1 < r \leq 0.3 \\ 2 & ; \quad 0.3 < r \leq 0.6 \\ 3 & ; \quad 0.6 < r <= 0.8 \\ 4 & ; \quad 0.8 < r \leq 0.95 \\ 5 & ; \quad 0.95 < r \leq 1 \end{cases}$$

Utilizando la regla de correspondencia, las demandas serían:

Día	R	Demanda
1	0.4501	2
2	0.0364	0
3	0.5778	2
4	0.8066	4
5	0.9591	5

El método de la transformada inversa sirve para obtener una fórmula que permita generar los números aleatorios con la distribución deseada, a dicha fórmula se le llama *generador*. Una vez que se tiene el generador, entonces se debe utilizar una computadora o una calculadora que permita el manejo de hojas de cálculo o de programación para obtener todos los números aleatorios deseados y de ahí la simulación del experimento.

■

AUTOEXAMEN TEMA V

1.- Ciertas resistencias se fabrican con una tolerancia de $\pm\ 10\%$. Si se considera que la resistencia real está distribuida uniformemente dentro de dicho intervalo, la probabilidad de que una resistencia con valor nominal de 1000 Ω tenga una resistencia real entre 990 y 1010 Ω es: ☐

A) 0.8 B) 0.6 C) 0.4 D) 0.2 E) 0.1

2.- El precio que se pide por cierto artículo se distribuye normalmente con media de \$50.00 y desviación estándar de \$5.00. Los compradores están dispuestos a pagar una cantidad que también se distribuye normalmente con media de \$45.00 y desviación estándar de \$2.50. La probabilidad de que se realice la transacción es: ☐

A) 0 B) 1 C) 0.187 D) 0.813 E) 0.255

3.- Una videocasetera tiene una distribución de tiempo de falla exponencial, con tiempo medio de 20 000 horas. Si la videocasetera ha durado 20 000 horas, entonces la probabilidad de que falle a las 30 000 horas o antes es: ☐

A) 0.606 B) 0.3935 C) Prácticamente cero D) 0.777
E) Ninguna de las anteriores.

4.- De las siguientes curvas normales, con los parámetros que se indican, la que se parece más a la curva con parámetros $\mu = 10$ y $\sigma = 5$ es: ☐

A) $\mu = 10\ ,\ \sigma = 10$ B) $\mu = 20\ ,\ \sigma = 10$ C) $\mu = 20\ ,\ \sigma = 2.5$
D) $\mu = 5\ ,\ \sigma = 2.5$ E) $\mu = 20\ ,\ \sigma = 5$

5.- El tiempo de servicio en la ventanilla de cierto banco sigue una distribución exponencial con una media de 2 minutos. La probabilidad de que el tiempo de servicio para el siguiente cliente sea de 3 minutos o más es: ☐

A) 0.25 B) 0.2231 C) 0.0024 D) 0.7769
E) Ninguna de las anteriores.

6.- Si para una variable aleatoria normal se sabe que $P(X \le -1) = 0.1587$ y $P(X > 7.88) = 0.025$ entonces la media y la variancia de X son, respectivamente: ☐

A) 3 , 2 B) 2 , 3 C) 2 , 9 D) -2 , 3
E) Ningunos de los anteriores.

7.- Los parámetros de forma, sesgo (α_3) y curtosis (α_4) para una distribución normal deben tomar los valores: ☐

A) $\alpha_3 > 0$, $\alpha_4 > 0$ B) $\alpha_3 = 0$, $\alpha_4 > 3$

C) $\alpha_3 < 0$, $\alpha_4 = 3$ D) $\alpha_3 = 0$, $\alpha_4 > 0$

E) $\alpha_3 = 0$, $\alpha_4 = 3$

8.- Si X es una variable aleatoria con distribución normal con media 2 y variancia 9, y Z es una variable aleatoria normal estándar, entonces $P(0 < X < 4)$ es igual a: ☐

A) $P\left(0 < Z < \frac{2}{3}\right)$ B) $P\left(-\frac{2}{9} < Z < \frac{2}{9}\right)$

C) $P\left(-\frac{9}{2} < Z < -\frac{5}{2}\right)$ D) $2\,P\left(0 < Z < \frac{2}{3}\right)$

E) Ninguna de las anteriores.

9.- La precipitación anual por lluvias en cierta región tiene una distribución normal con $\mu = 40\ [\text{cm}]$ y $\sigma^2 = 16\ [\text{cm}^2]$. La probabilidad de que en dos de los próximos cuatro años la precipitación sea de más de 44 $[\text{cm}]$ es: ☐

A) 0.004 B) 0.1587 C) 0.4761
D) 0.1069 E) Ninguna de las anteriores.

10.- Una llamada telefónica llegó a un conmutador en un tiempo aleatorio dentro de un minuto. El conmutador estuvo completamente ocupado durante 15 segundos de ese período de un minuto. La probabilidad de que la llamada llegara cuando el conmutador no estaba ocupado es: ☐

A) 0.75 B) 0.99 C) 0.25 D) 1
E) Ninguna de las anteriores.

BIBLIOGRAFÍA

Hines, William W. y Montgomery, Douglas C. - Probabilidad y Estadística para ingeniería, cuarta edición..- CECSA.- México, 2005.

Milton, Susan J. Y Arnold, Jesse C.- Probabilidad y Estadística para con aplicaciones para ingeniería y ciencias computacionales, cuarta edición.- McGraw-Hill.- México, 2004.

Devore, Jay L.- Probabilidad y Estadística para ingeniería y ciencias, séptima edición.- Cengage Learning.- México, 2008.

Mendenhall, William III. et al.- Introducción a la Probabilidad y Estadística.- Décimo cuarta edición.- Cengage Learning.- México 2015.

Wackerly Dennis D.- Mendenhall, William, *et al*.- Estadística Matemática con Aplicaciones, sexta edición.- Editorial Thomson.- México, 2002.

Walpole, Ronald E., *et al.*- Probability and Statistics for Engineers and Scientists.- Pearson.- USA, 2007.

Montgomery, Douglas C. y Runger, George C.-Probabilidad y Estadística aplicadas a la Ingeniería, segunda edición.- Limusa-Wiley.- México, 2002.

Scheaffer, Richard L. y McClave, James T.- Probabilidad y Estadística para Ingeniería.- Grupo Editorial Iberoamérica.- México, 1993.

Canavos, George C.- Probabilidad y Estadística Aplicaciones y Métodos.- McGraw-Hill.- México, 1988.

Meyer, Paul L.- Probabilidad y Aplicaciones Estadísticas.- Addison Wesley Iberoamericana.- México, 1992.

Spiegel, Murray R. et al.- Probabilidad y Estadística, cuarta edición.- Mc Graw-Hill.-México 2013.

Borras García, Hugo E., *et al*.- Apuntes de Probabilidad y Estadística.-Facultad de Ingeniería.- México, 1985.

Rosenkrantz, Walter A.- Introduction to Probability and Statistics for Scientists and Engineers.- McGraw-Hill.- EE.UU., 1997.

Ziemer, Rodger E.- Elements of Engineering Probability & Statistics.- Prentice Hall.- USA 1997.

García, Francisco, et. al.- Simulación de Sistemas para Administración e Ingeniería.- CECSA.- México, 2005.

Coss Bu, Raúl.- Simulación un enfoque práctico.- Limusa.-México, 1993.