centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

PROBABILIDAD Y ESTADISTICA

FUNDAMENTOS Y APLICACIONES

DR. OCTAVIO RASCON CH.

Palacio de Minería
Tacuba 5, primer piso México 1, D F.
Tels.: 521-40-23 521-73-35 512-31-23

PROBABILIDAD Y ESTADISTICA
FUNDAMENTOS Y APLICACIONES

- Feb-17-76 -

PROBABILIDAD: Es una medida de la incertidumbre que se asocia a la ocurrencia de un evento cualquiera.

ESTADISTICA: Es la rama de las matemáticas que se encarga de enseñar las reglas para colectar, presentar y procesar los datos obtenidos al repetir varias veces un experimento asociado a un fenómeno de interés. Proporciona, además, los métodos para el diseño de experimentos y para tomar decisiones cuando aparecen situaciones de incertidumbre.

ESTADISTICA
- Descriptiva: Trata lo concerniente a la obtención, organización, procesamiento y presentación de los datos.
- Inferencial: Trata lo concerniente a los métodos para inferir conclusiones acerca de la población de la cual provienen los datos.

Teoría de conjuntos

Un conjunto es una colección bien definida de objetos

Ejemplos : El conjunto de números anotados en un dado es

$$S = \{1, 2, 3, 4, 5, 6\}$$

El conjunto de los números enteros menores que 5 es

$$S = \{-\infty, \ldots, -3, -2, -1, 0, 1, 2, 3, 4\}$$

o $S = \{x : x \text{ es entero y } x < 4\}$

El conjunto de los números enteros positivos menores que 5 es

$$E = \{0, 1, 2, 3, 4\}$$

El conjunto de los continentes es

$$C = \{\text{Asia, Europa, América, Africa, Oceanía}\}$$

El conjunto de marcas que tiene una moneda es

$$M = \{\text{cara, cruz}\}$$

Conjuntos
- Finitos.- Cuando tienen un número finito de elementos
- Infinitos.- Cuando tienen un número infinito de elementos

Símbolos:

$<$ menor que

$\leq$ menor o igual que

$>$ mayor que

$\geq$ mayor o igual que

Ejemplo: ¿Cómo se puede denotar el conjunto de números mayores de 5 pero menores o iguales que 10?

$$S_1 = \{x : 5 < x \leq 10\}$$

Para expresar que un elemento pertenece a un conjunto se usa el símbolo $\in$. Para expresar que no pertenece se usa el símbolo $\notin$

Ejemplo: En relación al conjunto S_1:

$3 \notin S_1$; $5 \notin S_1$; $8 \in S_1$; $10 \in S_1$

Para expresar que un conjunto está contenido en otro se usa el símbolo $\subset$; si no lo está se usa el símbolo $\not\subset$.

Ejemplo: Sean $E = \{3, 5\}$; $F = \{3, 8\}$; $G = \{7, 9\}$

$E \not\subset S_1$; $F \not\subset S_1$; $G \subset S_1$

(Para que un conjunto esté contenido en otro se requiere que todos sus elementos lo estén; es decir, que todos ... hay conjunto)

DIVISION DE ESTUDIOS SUPERIORES
FACULTAD DE INGENIERIA
CIUDAD UNIVERSITARIA MEXICO 20 D F
APDO POSTAL 70 256

UNIVERSIDAD NACIONAL AUTONOMA

Si un conjunto, B, está contenido en otro, S, se dice que B es SUBCONJUNTO de S.

Ejemplo.- Sean $A=\{x: 8 \leq x \leq 10\}$, $B=\{x: 3 \leq x \leq 8\}$ y $S_1=\{x: 5 < x \leq 10\}$.

En este caso:

$A \subset S_1 \Rightarrow$ A es subconjunto de S_1

$B \not\subset S_1 \Rightarrow$ B no es subconjunto de S_1

Se dice que dos conjuntos son iguales cuando contienen los mismos elementos (no importa el orden en que estos se escriban)

Ejemplo.- $A=\{1, 3, 5, 7\}$; $B=\{7, 5, 1, 3\}$, $C=\{7, 5, 1\}$

$A = B$; $A \neq B$

De la misma manera que existe el cero en los números, en la teoría de conjuntos existe el conjunto VACIO, el cual es tal que no tiene elementos. Usualmente se denota $\emptyset$

Ejemplo.- ¿Cuál es el conjunto de elementos, x, tales que $2x=7$ y x es entero?

Solución.- El conjunto vacío ϕ.

A ϕ se le considera como subconjunto de cualquier conjunto. Así, por ejem., todos los subconjuntos del conjunto $S=\{2, 5, 10\}$ son: $\{2\}$; $\{5\}$; $\{10\}$; $\{2,5\}$; $\{2,10\}$; $\{5,10\}$; $\{2,5,10\}$ y ϕ.

Experimento.- Para fines de este curso, se entenderá por experimento a todo proceso de observación. Por ejemplo, el lanzar una moneda o un dado y observar la cara que queda hacia arriba, es un experimento.

Asociado a un experimento siempre hay un conjunto de resultados posibles; a dicho conjunto se le llama ESPACIO DE EVENTOS

Ejemplos.- El espacio de eventos asociado al experimento de lanzar un dado y anotar la cara que queda hacia arriba es

$S=\{1, 2, 3, 4, 5, 6\}$

El espacio de eventos correspondiente al experimento de lanzar dos dados y anotar los números que quedan hacia arriba es

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

Si en este experimento la observación de interés fuese la suma de los dos números observados, entonces el espacio de eventos sería

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

Al resultado de un experimento se le llama <u>dato</u> u <u>observación</u>; a un conjunto de datos se le llama <u>muestra</u>.

A todo subconjunto de un espacio de eventos se le llama EVENTO. A los eventos que tienen un solo elemento del espacio se le llama EVENTO SIMPLE.

Si al realizar un experimento se observa un elemento del evento A, entonces se dice que ocurrió o se verificó el evento A. Por ejemplo, si $A = \{2, 4\}$ y al lanzar un dado se observa 2 o 4, se dice que ocurrió el evento A; si se observa cualquier otro número, entonces se dice que no ocurrió A.

Espacios de Eventos
- Discretos.- Si sus elementos pueden numerarse o contarse. Tienen un número finito o infinito numerable de elementos.
- Continuos.- Si sus elementos no pueden enumerarse. Tienen un número infinito no numerable de elementos.

Ejemplos.- Los espacios de eventos $S = \{cara, cruz\}$; $S = \{1, 2, 3, 4, 5, 6\}$; $S = \{verde, rojo\}$ son discretos. Los espacios de eventos $S = \{x : -\infty < x \leq 0\}$; $S = \{z : z \geq 3\}$; $S = \{y : 3 \leq y \leq 5\}$ son continuos.

VNIVERSIDAD NACIONAL AVTONOMA

Ejemplos.- ¿Qué tipos de espacios de eventos corresponden a los siguientes experimentos?

a) Conteo del número de granos de una mazorca de maíz.- Discreto e infinito: $S = \{0, 1, 2, 3, \ldots, \infty\}$

b) Medición de la longitud de una espiga de trigo.- Continuo e infinito: $S = \{x : 0 < x < \infty\}$, x en cm

c) Medición del efecto de una vacuna, en términos de "exito" o "fracaso".- Discreto y finito: $S = \{\text{éxito, fracaso}\}$

d) Medición del número de miligramos de un antibiótico contenido en una cápsula.- Continuo e infinito

$S = \{y : 0 \leq y < \infty\}$, x en mg.

El complemento de un evento A es otro evento que contiene todos los elementos del espacio de eventos correspondiente que no están en A. Usualmente se denota con una tilde sobre el símbolo que corresponde al evento que complementa, es decir, $\bar{A}$.

Ejemplo: $S = \{1, 2, 3, 4, 5, 6\}$. $A = \{1, 3, 5\} \Rightarrow \bar{A} = \{2, 4, 6\}$

Si $S = \{x : 0 \leq x \leq 58\}$ y

$A = \{x : 3 < x \leq 17\}$, entonces

$\bar{A} = \{x : 0 \leq x \leq 3,\ 17 < x \leq 58\}$

Cuando dos o más eventos no
pueden ocurrir simultáneamen
te al realizar una sola vez un
experimento, se dice que éstos
son <u>MUTUAMENTE</u> <u>EXCLUSIVOS</u>.
Es decir, dos eventos son mutua
mente exclusivos cuando <u>no</u> tienen
ni un solo elemento en común.

<u>Ejemplos</u>.- Cualquier evento y su com
plemento son mutuamente exclusivos.

¿Son $E = \{y : 0 \leq y \leq 25\}$ y $A = \{2, 50, 100\}$
mutuamente exclusivos? <u>NO</u>, porque
tienen el elemento <u>2</u> en común.

La <u>UNION</u> de dos eventos es otro
evento cuyos elementos son todos
los de ambos. La operación de
unión se denota con el símbolo $\cup$.

<u>Ejemplos</u>: Si $A=\{2,4,6\}$ y

$B=\{1,6,12\}$, entonces

$C=A\cup B=\{1,4,6,12,2\}$

¿Son A y B mutuamente exclusivos? <u>NO</u> porque tienen el 6 en común.

Si $D=\{y: 0\le y\le 13\}$ y $E=\{y: 20\le y\le 50\}$,

$D\cup E=\{y: 0\le y\le 13\ ,\ 20\le y\le 50\}$

Si $F=\{y: 8\le y\le 20\}$, entonces

$D\cup F=\{y: 0\le y\le 20\}$.

Si $G=\{y: 3\le y\le 10\}$, entonces

$D\cup G=\{y: 3\le y\le 10\}=G$; obsérvese que en este caso $G\subset D$. En general, si $A\subset B$, entonces $A\cup B=B$

En general, la <u>UNION</u> de <u>varios</u> eventos es otro evento cuyos elementos son todos los de los eventos que aparecen en la unión.

Ejemplo: $A\cup B\cup F=K=\{1,2,4,6,\ y: 8\le y\le 20\}$

La INTERSECCION de dos eventos es el conjunto de elementos que pertenecen simultáneamente a ambos.

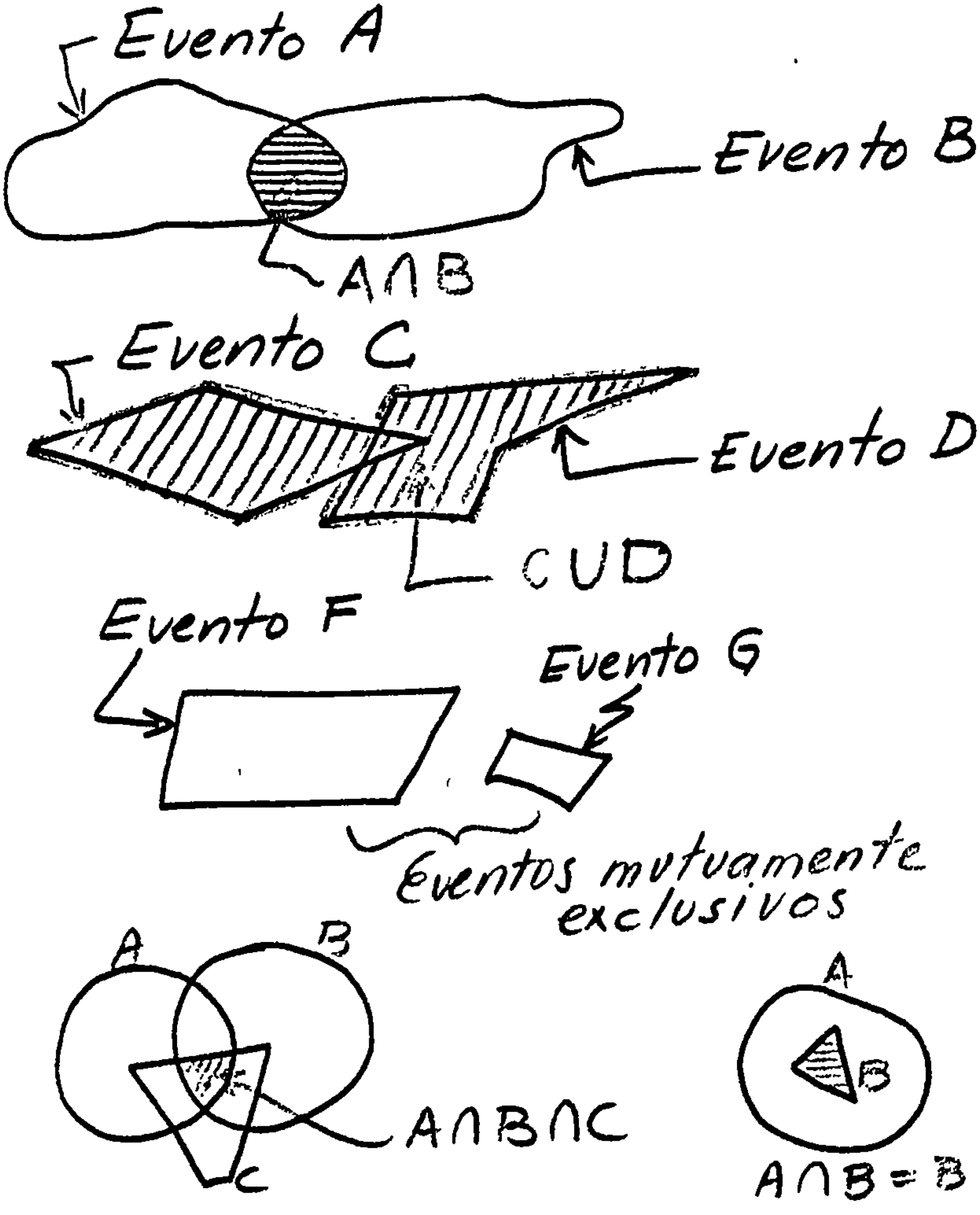
Ejemplos:- Si $A=\{2,3,6\}$ y $B=\{2,6,10\}$, entonces $A \cap B = C = \{2,6\}$ (para denotar la operación de intersección se usa el símbolo $\cap$) Si $D=\{y:\ 4 \leq y \leq 5\}$, $A \cap D = \phi$ (es el conjunto vacío). Obsérvese que A y D son mutuamente exclusivos ya que no tienen ningún elemento en común; cuando dos eventos son mutuamente exclusivos, su intersección es el conjunto vacío.

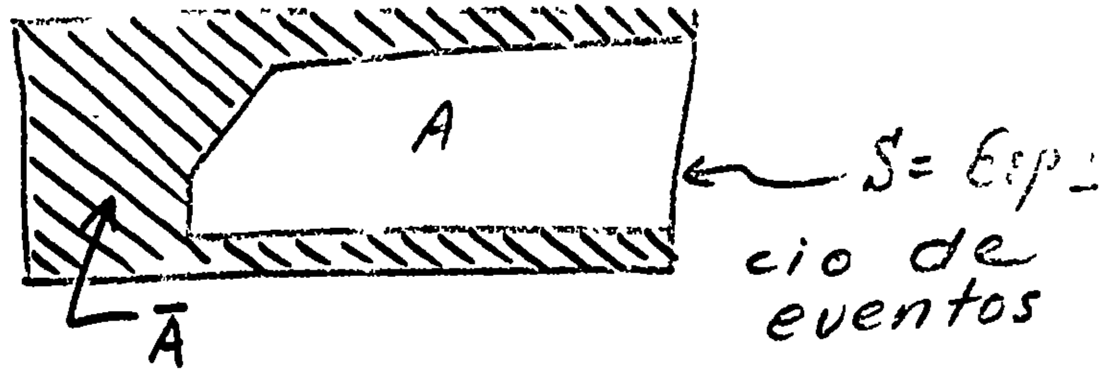
En general, la INTERSECCION de varios eventos es el conjunto de elementos que todos ellos tienen en común.

Ejemplo: Si $A=\{2,3,6,8\}$; $B=\{2,3,10,100\}$; $C=\{y:\ 0 \leq y \leq 5\}$; $D=\{y:\ 2 \leq y \leq 4\}$, entonces $A \cap B \cap C \cap D = E = \{2,3\}$

DIAGRAMAS DE VENN

Una manera de ilustrar gráficamente las operaciones con conjuntos es mediante los diagramas de VENN. En estos, cada conjunto se representa por una curva cerrada que encierra los elementos que le corresponden.





__TEORIA__ __DE__ __PROBABILIDADES__

Al lanzar una moneda no podemos predecir con certeza cuál cara quedará hacia arriba. Lo único que se puede asegurar, si la moneda no está cargada, es que ambas caras tienen la misma oportunidad de salir, es decir, que los eventos simples {cara} y {cruz} tienen la misma __probabilidad__ de ocurrir. En general, "la probabilidad de que ocurra un evento es una medida del grado de confianza que se tiene de que éste ocurra al realizar una vez el experimento correspondiente."

<u>Ejemplo</u>.- De una urna que contiene bolas rojas, blancas y negras, se sacó una bola, anotó su color y se regresó a la urna. Si este experimento se repite 20 veces y los resultados son, ($R=\{Roja\}$, $B=\{Blanca\}$, $N=\{negra\}$):

B, B, N, R, R, R, N, B, R, N, B, B, N, R,

B, R, R, N, R, N

¿que probabilidades le asignaría a los eventos B, N y R, de acuerdo con el método frecuencial?

$$n(B) = 6\,,\ n(N) = 6\,,\ n(R) = 8\,,\ n = 20$$

$$P(B) = \frac{6}{20} = \frac{3}{10}\ ;\ P(N) = \frac{6}{20} = \frac{3}{10}\ ;\ n(R) = \frac{8}{20} = \frac{4}{10}$$

(Nótese que los eventos B, N y R son mutuamente exclusivos y que

$$P(B)+P(N)+P(R) = \frac{3}{10} + \frac{3}{10} + \frac{4}{10} = 1\,)$$

- La definición clásica de probabilidades dice que si $n(A)$ es el número de maneras <u>igualmente probables</u> en que puede ocurrir el evento A y n es el número total [illegible] correspondiente, entonces la probabilidad de A es

$$\boxed{P(A) = \frac{n(A)}{n}}$$

Ejemplos.- Si en una urna se tienen 5 bolas blancas y 15 negras, y se va a seleccionar una al azar, ¿cuál es la probabilidad de que sea roja ($A = \{roja\}$)?:

$$n = 5 + 15 = 20 \; ; \; n(A) = 5 \Rightarrow P(A) = \frac{5}{20} = \frac{1}{4}$$

Si se lanzan dos dados, ¿cuál es la probabilidad de que

a.- la suma sea 7 (evento A)

b.- salga un 2 y un 5 (evento B)?

Para el inciso b el espacio de eventos es:

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

cada pareja de números es igualmente probable, si el dado no está cargado. En tal caso $n = 36$ y $n(B) = 2$ (si aparece $(2,5)$ ó $(5,2)$) $\Rightarrow P(B) = 2/36 = 1/18$.

Para el inciso a el espacio de eventos es

$$S_1 = \{2,3,4,5,6,7,8,9,10,11,12\}$$

pero no todos los elementos (eventos simples) son igualmente probables, ya que, por ejemplo, el 2 sólo aparecerá si se observa la pareja (1,1), en cambio el 3 aparecerá si ocurren las parejas (1,2) o (2,1), es decir, el 3 tiene el doble de probabilidad que el 2. Por esto, para calcular la probabilidad de que la suma sea 7 es necesario trabajar con el ~~[illegible]~~ espacio S y contar las posibilidades de que la suma sea 7, lo cual ocurre si se observa cualquiera de las parejas (6,1); (5,2); (4,3); (3,4); (2,5) o (1,6), es decir, hay 6 maneras igualmente probables de que ocurra el evento A. Por lo tanto

$$P(A) = \frac{n(A)}{n} = \frac{6}{36} = \frac{1}{6}$$

Procediendo de esta manera se pueden calcular las probabilida des de que la suma sea 2, 3, 4, etc. Los resultados son:

$P(\{2\})=\frac{1}{36}$; $P(\{3\})=\frac{2}{36}$; $P(\{4\})=\frac{3}{36}$;

$P(\{5\})=\frac{4}{36}$; $P(\{6\})=\frac{5}{36}$; $P(\{7\})=\frac{6}{36}$;

$P(\{8\})=\frac{5}{36}$; $P(\{9\})=\frac{4}{36}$; $P(\{10\})=\frac{3}{36}$;

$P(\{11\})=\frac{2}{36}$ y $P(\{12\})=\frac{1}{36}$

(Obsérvese que $\sum_{i=2}^{12} P(\{i\}) = 1$)

- Con base en un modelo probabilístico, las probabilidades se asignan a partir de un modelo matemático que involucre todos los factores posibles que intervienen en la aleatoriedad del fenómeno.

AXIOMAS

Las probabilidades que se asignan a los diferentes eventos relacionados con un fenómeno aleatorio deben cumplir con los siguientes <u>tres</u> axiomas:

<u>AXIOMA 1</u>: La probabilidad de ocurrencia de un evento A es un número, $P(A)$, que se le asigna a dicho evento, cuyo valor queda en el intervalo

$$\boxed{0 \leq P(A) \leq 1}$$

AXIOMA 2: Si S es un espacio de eventos, entonces

$$\boxed{P(S) = 1}$$

AXIOMA 3: La probabilidad, $P(C)$, de la unión, C, de dos eventos mutuamente exclusivos, A y B, es igual a la suma de probabilidades de éstos, es decir,

$$\boxed{P(A \cup B) = P(C) = P(A) + P(B)}$$

<u>Ejemplos</u>.- En el problema del lanzamiento de un dado, si $A = \{2, 4\}$ y $B = \{5, 6\}$, y suponiendo que el dado no está cargado, se puede asignar a cada número una probabilidad de 1/6. En tal caso, puesto

[illegible]

evento elemental es mutua-
mente exclusivo con cual
quier otro, aplicando el
Axioma 3 se obtienen:

$$P(A)=P(\{2\})+P(\{4\})=\frac{1}{6}+\frac{1}{6}=\frac{2}{6}=\frac{1}{3}$$

$$P(B)=P(\{5\})+P(\{6\})=\frac{1}{6}+\frac{1}{6}=\frac{2}{6}=\frac{1}{2}$$

Si $C=A\cup B$, y dado que A y B son eventos mutuamente exclusivos:

$$P(C)=P(A)+P(B)=\frac{1}{3}+\frac{1}{3}=\frac{2}{3}$$

Además, obsérvese que se cumple con los axiomas 1 y 2 ya que

$$P(S)=P(\{1\})+P(\{2\})+P(\{3\})+P(\{4\})+P(\{5\})+P(\{6\})$$
$$=\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{6}{6}=1$$

Ejemplo.- En el problema del lanzamiento de dos dados, ¿cuál es la probabilidad que al realizar una vez el experimento la suma de los dos números que queden hacia arriba sea 7 u 11? Esto es equivalente a preguntar por la probabilidad de que ocurra el evento

$C=\{7\}\cup\{11\}$. Puesto que $\{7\}$ y $\{11\}$ son eventos mutuamente exclusivos:

$P(C)=P(\{7\})+P(\{11\})=\frac{6}{\ }+\frac{2}{\ }=\frac{8}{\ }=\frac{2}{\ }$

Ejemplo: En un laboratorio se probaron 40 vigas de concreto reforzado nominalmente idénticas y se anotaron las cargas con las cuales falló cada una. De esta sucesión de experimentos se asignaron, en términos de las frecuencias correspondientes, las siguientes probabilidades:

$A = \{x: 0 \le x \le 20 \text{ ton}\}$; $P(A) = 0.17$

$B = \{x: 20 < x \le 40 \text{ ton}\}$; $P(B) = 0.24$

$C = \{x: 40 < x \le 60 \text{ ton}\}$; $P(C) = 0.27$

$D = \{x: 60 < x \le 80\}$; $P(D) = 0.13$

$E = \{x: 80 < x \le 100\}$; $P(E) = 0.11$

$F = \{x: 100 < x\}$; $P(F) = 0.08$

$\sum P(\cdot) = 1.00$

Si se realiza una vez más el experimento, calculemos las siguientes probabilidades:

a.- Que la resistencia sea menor o igual que 80 ton:

$G = A \cup B \cup C \cup D \Rightarrow P(G) = P(A) + P(B)\, P(C) + P(D)$

$P(G) = 0.17 + 0.24 + 0.27 + 0.13 = 0.81$

b.- La probabilidad que resista más de 60 ton:

$H = D \cup E \cup F \Rightarrow P(H) = P(D) + P(E) + P(F)$

$P(D) = 0.13 + 0.11 + 0.08 = 0.32$

c.- La probabilidad que resista más de 40 ton pero cuando mucho 100 ton:

$I = C \cup D \cup E \Rightarrow P(I) = P(C) + P(D) + P(E)$

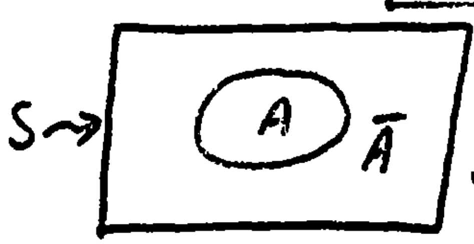
$P(I) = 0.27 + 0.13 + 0.11 = 0.51$

TEOREMAS

Dos teoremas importantes que se deducen a partir de los axiomas son:

Teorema 1.-

Si A es un evento del espacio S, entonces $\boxed{P(\bar{A}) = 1 - P(A)}$



Demostración.- Puesto que A y $\bar{A}$ son mutuamente exclusivos y además $A \cup \bar{A} = S$, entonces, $P(S) = P(A) + P(\bar{A}) = 1$

$$\Rightarrow \quad P(\bar{A}) = 1 - P(A)$$

caso particular: $\bar{S} = \phi \Rightarrow P(\bar{S}) = 1 - P(S) = 0$

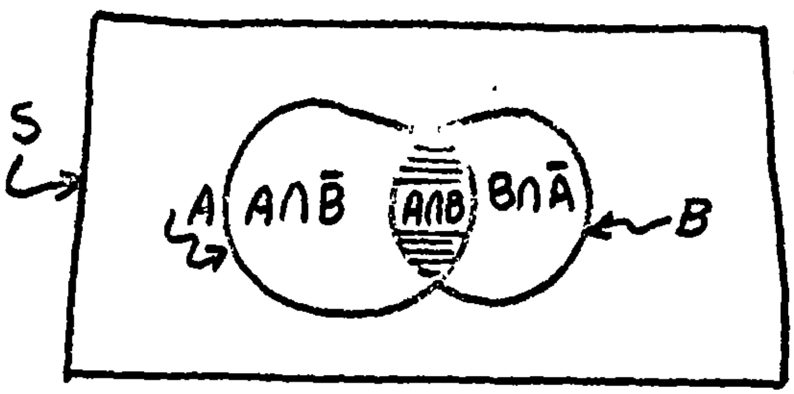
$\therefore \quad P(\phi) = 0$

VNIVERSIDAD NACIONAL AVTONOMA

Teorema 2.- Si A y B son dos eventos cualquiera de S, entonces

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

Demostración: Sea el diagrama de Venn:



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

Puesto que $A \cap B$, $A \cap \bar{B}$ y $B \cap \bar{A}$ son mutuamente exclusivos se tiene que

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

Sumando y restando $P(A \cap B)$ y agrupando términos se obtiene

$$P(A \cup B) = \left[P(A \cap B) + P(A \cap \bar{B})\right] + \left[P(A \cap B) + P(B \cap \bar{A})\right] - P(A \cap B)$$

Pero $A = (A \cap \bar{B}) \cup (A \cap B) \Rightarrow P(A \cap B) + P(A \cap \bar{B}) = P(A)$

y $B = (A \cap B) \cup (B \cap \bar{A}) \Rightarrow P(A \cap B) + P(B \cap \bar{A}) = P(B)$, por lo que

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

VNIVERSIDAD NACIONAL
AVTONOMA

Ejemplo.- En una urna se tienen 28 tiras de papel y en cada una se encuentra anotada una letra distinta del alfabeto. Calcule la probabilidad de que al extraer al azar una tira:

a.- Se obtenga una vocal

b.- Se obtenga a o z

c.- Ocurran C y D, donde $C=\{x,y,z\}$ y $D=\{b,c,y,z\}$

d.- Ocurra C o D

a.- $A=\{a,e,i,o,u\} \Rightarrow P(A)=\frac{5}{28}$

b.- $B=\{a,z\} \Rightarrow P(B)=\frac{2}{28}$

c.- $F=C\cap D=\{y,z\} \Rightarrow P(F)=\frac{2}{28}$

d.- $E=C\cup D=\{b,c,x,y,z\} \Rightarrow P(E)=\frac{5}{28}$

o $P(E)=P(C)+P(D)-P(C\cap D)$

$P(C\cap D)=P(F)=\frac{2}{28} \Rightarrow P(E)=\frac{3}{28}+\frac{4}{28}-\frac{2}{28}=\frac{5}{28}$

<u>NOTA</u>: La ocurrencia de un evento y otro implica la ocurrencia de ambos a la vez, es decir que se verifique la <u>intersección</u>. La ocurrencia de un evento o algún otro, implica la ocurrencia de cualquiera de ellos, es decir de la <u>unión</u>.

REGLAS DE CONTEO

Al asignar probabilidades a los eventos aplicando la teoría clásica es necesario calcular $n(A)$ y n para aplicar la fórmula $P(A)=n(A)/n$.

Sean, por ejemplo, los eventos $A=\{b,c\}$ y $B=\{a,e,i,o,u\}$ con los cuales se forman palabras de dos letras, la primera de A y la segunda de B. El evento que se forma así es

$$C=\{xy : x\in A;\ y\in B\}$$

Si enumeramos los elementos:

$$\left.\begin{array}{l}\text{con la } b\text{: } ba, be, bi, bo, bu\\ \text{con la } c\text{: } ca, ce, ci, co, cu\end{array}\right\}\begin{array}{l}10\\ \text{elementos}\end{array}$$

sin embargo, la solución se puede obtener rápidamente sin necesidad de enumerar todas las posibilidades, observando que la primera letra sólo puede ser de <u>dos</u> tipos <u>b</u> o <u>c</u>, mientras que la segunda, de <u>cinco</u> tipos <u>a</u>, <u>e</u>, <u>i</u>, <u>o</u> o <u>u</u>, por lo que el total de elementos es $2\times5=10$, es decir, el evento C puede ocurrir de 10 maneras distintas e igualmente probables.

En general, si dos eventos, A y B, pueden ocurrir de $n(A)$ y $n(B)$ maneras distintas, respectivamente, entonces el total de maneras en que ambos pueden ocurrir en el orden indicado es $\boxed{n(A) \times n(B)}$. Esta regla se puede generalizar a más de dos eventos.

<u>Ejemplo</u>.- ¿Cuántos números pares de tres cifras se pueden formar utilizando los dígitos 5, 6, 7, 8 y 9, sin que se use el mismo dígito en las decenas y las centenas?

<u>Solución</u>.- Sean los eventos

$$A = \{x : x \text{ está en las centenas}\}$$

$$B = \{y : y \text{ está en las decenas}\}$$

$$C = \{z : z \text{ está en las unidades y es par}\}$$

$$D = \{xyz : x \in A;\ y \in B;\ z \in C\}$$

Puesto que <u>no</u> se permite repetición de dígitos, $n(A) = 5$ y $n(B) = 4$. Además, puesto que el número debe ser par, $n(C) = 2$. Por lo tanto

$$N(D) = 5 \times 4 \times 2 = 40$$

Si el último dígito no tuviese que ser par: $S = \{xyz : x \in A;\ y \in B;\ z \in E\}$

donde $F = \{z : z$ está en las unidades$\}$

entonces $n(F) = 5$ y

$N(S) = 5 \times 4 \times 5 = 100.$

Con esto, calculemos la probabilidad de que si el espacio de eventos es S y se anotan todos los números del mismo en una tira de papel, al sacar uno al azar de una urna el número sea par:

$$P(A) = \frac{n(A)}{n} = \frac{n(A)}{n(S)} = \frac{40}{100} = 0.4$$

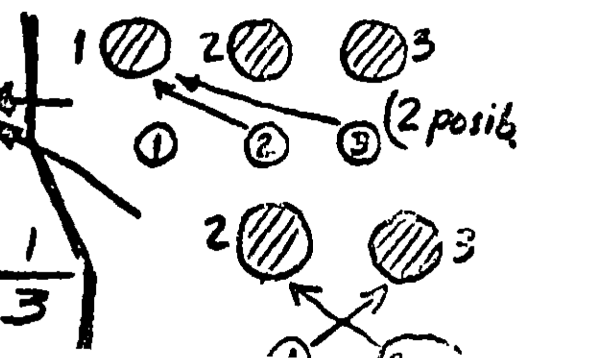
$$= 40\,\%$$

<u>Ejemplo</u>.- En una caja se tienen tres perforaciones numeradas del uno al tres. Si se hechan en ella tres bolas también numeradas del 1 al 3 y se agita la caja, calcular la probabilidad de que ninguna bola caiga en la perforación que tiene su número (evento A)

$$n(A) = 2 \times 1 \times 1 = 2$$

$$n = 3 \times 2 \times 1 = 6$$

$$P(A) = \frac{n(A)}{n} = \frac{2}{6} = \frac{1}{3}$$



Se dispone de tres banderas: una blanca, una negra y una verde.

UNIVERSIDAD NACIONAL AUTÓNOMA

1.- Si cada pareja de banderas de distinto color constituye una señal, ¿Cuántas señales se pueden hacer si el orden de colocación de las banderas es importante (evento A)?

$n(A) = 3 \times 2 = 6$

2.- Si tres banderas también constituyen una señal cuando todas son de diferente color ¿cuántas señales podemos hacer con las 3 banderas a la vez (evento B)?

$n(B) = 3 \times 2 \times 1 = 6$

3.- ¿Cuántas señales se pueden hacer con dos o tres banderas en las condiciones anteriores (evento C)?

$C = A \cup B$

$n(C) = n(A) + n(B) = 6 + 6 = 12$

4.- Si cada señal del evento C se dibuja en una tira de papel y luego se colocan en una urna, ¿cuál es la probabilidad de que si se toma una al azar

a.- Salga una señal específica (evento F).

$P(F) = n(F)/n(C) = 1/12$

b.- Una señal con dos banderas por lo menos (evento G):

$G = C \Rightarrow n(G) = 12 \Rightarrow P(G) = \frac{12}{12} = 1$

c.- Una señal con dos banderas, una de ellas verde (evento H):

$n(H) = 1 \times 2 + 2 \times 1 = 4 \Rightarrow P(H) = \frac{4}{12} = \frac{1}{3}$

d.- Una señal con tres banderas, una de ellas verde (evento I):

$n(I) = 1 \times 2 \times 1 + 1 \times 1 \times 2 + 2 \times 1 \times 1 = 6$

$P(I) = 6/12 = 1/2 = 50\%$

e.- Una señal con dos o tres banderas en que se use una verde (evento J)

$J = H \cup I \Longrightarrow n(J) = n(H) + n(I) = 4 + 6 = 10$

$P(J) = 10/12 = 5/6$

PERMUTACION.- El arreglo de N objetos en cierto orden se denomina Permutación.

Por ejemplo, todas las permutaciones que pueden hacerse con las letras a, b, c son: abc, acb, bac, bca, cab, cba. El total, es $3 \times 2 \times 1 = 6$ permutaciones ($N=3$).

En general, el número de permutaciones es $N(N-1)(N-2)(N-3) \times \cdots \times 1 = N!$

Ejemplo.- ¿Cuántas permutaciones se pueden hacer con 5 objetos?

$\ldots \times 1 = 120$

Ejemplo: En un [illegible] colocarán al azar 7 libros. Calculemos la probabilidad de que el de historia y el de geografía queden juntos (evento A).

$$P(A) = n(A)/n$$

$$n = 7! = 7 \times (6 \times 5 \times 4 \times 3 \times 2 \times 1)$$
$$= 7 \times 6!$$

$$n(A) = 2! \times 6! \implies P(A) = \frac{2! \times 6!}{7!} = \frac{2! \times 6!}{7 \times 6!} = \frac{2}{7}$$

Ejemplo.- En una urna se tienen 6 esferas, tres blancas y tres negras. Si las seis se extraen al azar, una tras otra, la probabilidad de que las 3 blancas salgan en forma consecutiva (evento F) es: $n(F) = 3! \times 4!$; $n = 6!$

$$\implies P(F) = n(F)/n = \frac{3! \times 4!}{6!} = \frac{3! \times 4!}{6 \times 5 \times 4!} =$$

$$P(F) = \frac{3 \times 2 \times 1}{6 \times 5} = \frac{1}{5}$$

Ejemplo.- En una urna se tienen 7 sobres idénticos y cada uno contiene un billete de diferente denominación (1, 5, 10, 20, 50, 100 y 500 pesos). ¿Cuál es la probabilidad de 1, 5, 50 y 500 pesos salgan consecutivamente en cualquier orden, si se sacan los siete al azar, uno tras otro (evento A)?

$n(A) = 4! \times 4!$, $n = 7!$

$$P(A) = \frac{4! \times 4!}{[illegible]} = \frac{4! \times 4!}{[illegible]} = \frac{4 \times 3 \times 2 \times 1}{7 \times 6 \times 5} = \frac{4}{35}$$

El número de maneras en que se pueden ordenar N objetos tomando de r en r es:

$$\boxed{{}_NP_r = \frac{N!}{(N-r)!}}$$

Esto es equivalente a decir que ${}_nP_r$ es el número de diferentes maneras en que r objetos pueden ser seleccionados de N objetos $(r \leq N)$ sin reemplazar ninguno de ellos al lote antes de sacar el siguiente.

Obsérvese que si $r=N$:

$$ {}_NP_N = \frac{n!}{(N-N)!} = \frac{N!}{0!} = N! $$

Ejemplo.- Si se tienen las letras a, b, c, d, el número de maneras en que se pueden ordenar tomando de 2 en 2 es

$$ {}_NP_2 = \frac{4!}{(4-2)!} = \frac{4\times 3\times 2!}{2!} = 12 $$

El conjunto ~~[illegible]~~ de estas posibilidades es:

$$ S = \{ab, ac, ad, ba, bc, bd, ca, cb, cd, da, db, dc\} $$

Obsérvese que cuando el orden es importante, ac no es lo mismo que ca, etc.

Cuando el orden no es importante, ..

decir, si el agrupamiento ca es el mismo que el ac, a los agrupamientos se les denomina COMBINACIONES. Por ejemplo si se formará una comisión de dos individuos de un grupo de 8 tomando sus nombres al azar de una urna, y deseamos saber cuántos comités de 2 miembros podrían formarse como resultado del proceso, entonces los resultados Pedro, José y José, Pedro constituirían el mismo comité, es decir, no importaría en qué orden se sacaran sus nombres de la urna.

Así, se puede demostrar que el número de combinaciones posibles de formar de N objetos tomando de r en r es:

$$\boxed{{}_NC_r = \binom{N}{r} = \frac{N!}{(N-r)!\,r!}}$$

Esto equivale a decir que ${}_NC_r$ es el número de maneras distintas en que r objetos pueden seleccionarse de N $(r \leq N)$ <u>sin reemplazo</u> y <u>sin importar</u> el <u>orden</u> en que aparezcan.

Cuando uno se enfrenta a un problema que implica la repetición de una acción, es necesario determinar si hay o no reemplazo de las observaciones. Por ejemplo, el repetir el lanzamiento de un dado y observar cada vez el número que queda hacia arriba lleva implícito que hay reemplazo.

Ejemplo.- Una caja contiene 10 focos, de los cuales 3 son defectuosos. Si seleccionamos 4 al azar sin remplazo

a.- ¿Cuántos ~~focos~~ son los resultados posibles, es decir, cuantos elementos tiene el espacio de eventos?

$$n(S) = {}_{10}P_4 = \frac{10!}{(10-4)!} = 10 \times 9 \times 8 \times 7 = 5040$$

b.- ¿Cuántos elementos de S tienen como primer resultado un foco ~~bueno y~~ defectuoso y tres focos buenos en los otros tres?

$$n(A) = {}_3P_1 \times {}_7P_3 = \frac{3!}{(3-1)!} \times \frac{7!}{(7-3)!} = \frac{3!}{2!} \times \frac{7!}{4!}$$

$$n(A) = 3 \times (7 \times 6 \times 5) = 630$$

e.- ¿Cuántos elementos de S tienen un foco defectuoso y 3 buenos?

$$n(B) = 4 \times {}_3P_1 \times {}_7P_3 = 4 \times 630 = 2520$$

$$\Rightarrow P(B) = \frac{2520}{5040} = \frac{1}{2}$$

PERMUTACIONES DE GRUPOS DE OBJETOS

"El número de permutaciones posibles de N objetos de los cuales se tienen n_1 iguales entre sí en el primer grupo, n_2 iguales entre sí en el segundo grupo, etc, hasta n_k iguales en el k-ésimo grupo (los grupos son distinguibles entre sí), de manera que $n_1 + n_2 + \cdots + n_k = N$ queda dado por la fórmula:

$$\boxed{{}_NP_{n_1, n_2, \ldots, n_k} = \frac{N!}{n_1!\, n_2! \cdots n_k!}}$$

Caso particular.- Si se tienen dos grupos solamente, entonces, si $n_1 = r$, se tiene que $r + n_2 = N$ o $n_2 = N - r$, por lo que

$${}_NP_{r, N-r} = \frac{N!}{r! \times (N-r)!} = {}_NC_r$$

<u>Ejemplo</u>: Una caja contiene 25 piezas de las cuales 3 son defectuosas. ¿Cuál es la probabilidad de que, si se extraen 5 al azar sin reemplazo,

a.- se obtengan las 3 defectuosas
b.- se obtengan sólo 2 defectuosas
c.- se obtenga sólo 1 defectuosa
d.- no se obtenga ninguna defectuosa?

<u>Solución</u>

a.- $n(S) = {}_{25}P_5 = \dfrac{25!}{(25-5)!} = \dfrac{25!}{20!}$

$$\underbrace{d\ d\ d}_{{}_3P_3}\ \underbrace{b\ b}_{{}_{22}P_2}$$

d = defectuosa
b = buena

$${}_5P_{3,2}$$

$n(A) = {}_3P_3 \times {}_{22}P_2 \times {}_5P_{3,2} = 3! \dfrac{22!}{(22-2)!} \dfrac{5}{3!(5-3)!}$

$n(A) = 60 \dfrac{22!}{20!}$

$$P(A) = \frac{60 \frac{22!}{20!}}{\frac{25!}{20!}} = 60 \frac{22!}{25!} = \frac{60}{13800}$$

VNIVERSIDAD NACIONAL AVTONOMA

b.-

$$\underbrace{d\,d}_{{}_3P_2}\ \underbrace{b\,b\,b}_{{}_{22}P_3} \rightarrow {}_5P_{2,3}$$

$$n(B) = {}_3P_2 \times {}_{22}P_3 \times {}_5P_{2,3} = \frac{3!}{(3-2)!}\,\frac{22!}{(22-3)!}\,\frac{5!}{2!\times 3!}$$

$$n(B) = 3!\,\frac{22!}{19!}\,\frac{5\times 4}{2} = 60\,\frac{22!}{19!}$$

$$P(B) = \frac{60\,\frac{22!}{19!}}{\frac{25!}{20!}} = 60\,\frac{22!}{25!}\,\frac{20\times 19!}{19!} = \frac{1200}{13800}$$

c.-

$$\underbrace{d}_{3}\ \underbrace{b\,b\,b\,b}_{{}_{22}P_4} \rightarrow {}_5P_{1,4}$$

$$n(C) = 3\times {}_{22}P_4 \times {}_5P_{1,4}$$

$$n(C) = 3\,\frac{22!}{(22-4)!}\times\frac{5!}{1!\times 4!}$$

$$n(C) = 3\,\frac{22!}{18!}\,\frac{5\times 4!}{1\times 4!} = 15\,\frac{22!}{18!}$$

$$P(C) = \frac{15\,\frac{22!}{18!}}{\frac{25!}{20!}} = \frac{15\times 20\times 19\times 18!}{\frac{25!}{22!}\times 18!} = \frac{5700}{13800}$$

d.-

$$\underbrace{b\,b\,b\,b\,b}_{{}_{22}P_5}$$

$$n(D) = {}_{22}P_5 = \frac{22!}{(22-5)!} = \frac{22!}{17!}$$

$$P(D) = \frac{22!/17!}{\frac{25!}{20!}} = \frac{20\times 19\times 18\times 17!}{\frac{25!}{20!}\times 17!}$$

$$P(D) = 6840/13800$$

Obsérvese que en este ejemplo hemos calculado las probabilidades de todos los elementos del espacio de eventos correspondiente al "número de defectuosos que se pueden observar en una selección al azar de 5 elementos", en la cual sólo se pueden tener 0, 1, 2 ó 3 defectuosos, es decir,

$$H = \{0, 1, 2, 3\}$$

Verifiquemos que, en efecto, $P(H) = 1$:

$$P(H) = P(\{0\}) + P(\{1\}) + P(\{2\}) + P(\{3\})$$

$$= \frac{6840}{13800} + \frac{5700}{13800} + \frac{1200}{13800} + \frac{60}{13800}$$

$$= \frac{13800}{13800} = \underline{\underline{1}} \quad \underline{\underline{\text{O.K.}}}$$

PROBABILIDAD CONDICIONAL

La probabilidad condicional, $P(A|B)$ del evento A, dado que el B ha ocurrido se calcula con la fórmula

(A) -- $$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)}} \; ; \; P(B) > 0$$

Si dos eventos, A y B, son independientes, la probabilidad de A no se altera si ocurre el evento B; es decir, dos eventos son <u>independientes</u> si

$$\boxed{P(A|B) = P(A)}$$

En tal caso, de la ecuación (A):

$$\boxed{P(A \cap B) = P(A) \times P(B)}$$

En general, los eventos $A_1, A_2, \ldots, A_m$ son independientes si, y sólo si,

$$\boxed{P(A_{k_1} \cap A_{k_2} \cap \cdots A_{k_r}) = P(A_{k_1}) \times P(A_{k_2}) \times \cdots \times P(A_{k_r})}$$

para <u>cualquier</u> grupo de enteros $k_1, k_2, \ldots, k_r$, con $k_r \leq m$.

Por ejemplo, si $m = 3$, A_1, A_2 y A_3 son independientes si, y sólo si,

$$P(A_1 \cap A_2) = P(A_1)P(A_2)$$
$$P(A_1 \cap A_3) = P(A_1)P(A_3)$$
$$P(A_2 \cap A_3) = P(A_2)P(A_3)$$
$$P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$$

Si $m = 4$, para que sean independientes se requiere que se cumpla que

$$P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(A_1)P(A_2)P(A_3)P(A_4)$$

todas las combinaciones de 3 elementos

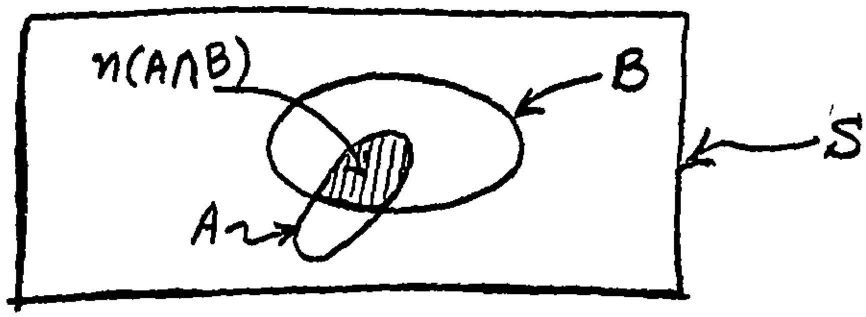
$$\left\{\begin{array}{l} P(A_1 \cap A_2 \cap A_3) = P(A_1)\,P(A_2)\,P(A_3) \\ \vdots \end{array}\right.$$

todas las de 2 elementos

$$\left\{\begin{array}{l} P(A_1 \cap A_2) = P(A_1)\,P(A_2) \\ \vdots \end{array}\right.$$

Puesto que $P(A \cap B) = n(A \cap B)/n(S)$ y $P(B) = n(B)/n(S)$ la ecuación A se puede escribir como

$$(B)\cdots \boxed{P(A|B) = \frac{\frac{n(A\cap B)}{n(S)}}{\frac{n(B)}{n(S)}} = \frac{n(A\cap B)}{n(B)}}$$



El trabajar con la ecuación B equivale a emplear un espacio de eventos reducido de S a B.

<u>Ejemplo</u>.- En una urna hay 10 transistores buenos y 10 defectuosos (malos). ¿Cuál es la probabilidad de sacar uno bueno y uno malo (en cualquier orden) al realizar <u>dos</u> extracciones al azar, si hay reemplazo del primer transistor observado?

Hay varias formas de resolver el problema:

1.- $S = \{(d,d), (d,b), (b,b), (b,d)\}$

$A = \{(d,b), (b,d)\}$; $n(S) = 4$, $n(A) = 2$

$\therefore$ $P(A) = 2/4 = 1/2$ (son igualmente probables los eventos simples porque el número de defectuosos es igual al de buenos.)

2.- $n(A) = (10 \times 10) \times 2 = 200$

$n(S) = 20 \times 20 = 400$

$P(A) = 200/400 = 1/2$

3.- $B = \{$sale primero el bueno y luego el malo$\}$

$C = \{$sale primero el malo y luego el bueno$\}$

$B = D \cap E$ donde

$D = \{$sale primero el bueno$\}$

$E = \{$sale segundo el malo$\}$ y $F = O \cap R$,

donde $O = \{$sale primero el malo$\}$

$R = \{$sale segundo el bueno$\}$

- Luego $A = \{\text{sale uno bueno y uno malo}\}$
$$= B \cup F$$
Por lo que $P(A) = P(B) + P(F)$
$$P(B) = P(D \cap E) = \frac{10}{20} \times \frac{10}{20} = \frac{100}{400} = \frac{1}{4}$$
$$P(F) = P(O \cap R) = \frac{10}{20} \times \frac{10}{20} = \frac{1}{4}$$
$$\therefore \quad P(A) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

- De esto se puede concluir que el formar un evento combinado por los resultados de repeticiones sucesivas de un experimento, al decir y equivale a <u>intersección</u>; al decir o equivale a <u>unión</u>:
$$y = \cap$$
$$o = \cup$$
Resolvamos ahora este problema si NO hay reemplazo:

- $P(D \cap E) = \frac{10}{20} \times \frac{10}{19} = \frac{10}{38}$; $P(O \cap R) = \frac{10}{20} \times \frac{10}{19} = \frac{10}{38}$

$$\therefore \; P(A) = \frac{10}{38} + \frac{10}{38} = \frac{10}{19}$$

En un estudio sociológico se interrogaron 1200 personas de una colonia residencial, y se obtuvieron los siguientes datos:

Gusto por la música clásica	Título Universitario		Sin título universitario		
	Varones	Damas	Varones	Damas	
Alto	100	50	200	250	600
Bajo	150	100	150	200	600
	250	150	350	450	

Si $A=\{\text{varón}\}$, $B=\{\text{con título}\}$

$C=\{\text{Gusto alto}\}$

¿Cuál es la probabilidad de que si se selecciona un ciudadano al azar (de misma colonia) éste sea varón, tenga título y gusto alto por la música?

Por el método frecuencial:

$P(A)= 600/1200 = 1/2$; $P(B)=400/1200=\frac{25}{120}$

$P(C)= 600/1200 = 1/2$

$D=A\cap B\cap C \Rightarrow P(D)=\frac{1}{2}\times\frac{40}{120}\times\frac{1}{2}=\frac{40}{480}=\frac{1}{1\text{[illegible]}}$

Si se supone que A, B y C son independientes.

- De la ecuación (A):

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

Esta ecuación se puede generalizar a más de dos eventos así:

$$\text{C)}\cdots \boxed{P(E_1\cap E_2\cap\cdots\cap E_k) = P(E_1)\,P(E_2|E_1)\cdots P(E_k|E_1,E_2,\cdots E_{k-1})}$$

Ley general de multiplicación

Por ejemplo, si $k=4$

- $$P(E_1\cap E_2\cap E_3\cap E_4) = P(E_1)\,P(E_2|E_1)\,P(E_3|E_1\cap E_2)\times$$
$$\times P(E_4|E_1\cap E_2\cap E_3)$$

Ejemplo.- ¿Cuál es la probabilidad de que al extraer sin reemplazo cuatro cartas al azar de um paquete de 52, las dos primeras sean diamantes y las 2 últimas sean corazones (evento E)?

$A = \{$la 1ª es diamante$\}$, $B = \{$la 2ª es diamante$\}$

$C = \{$la 3ª es corazón$\}$, $D = \{$la 4ª es corazón$\}$

$$E = A\cap B\cap C\cap D$$

- $P(A) = 13/52$, $P(B|A) = 12/51$, $P(C|A,B) = 13/50$

$P(D|A,B,C) = 12/49 \Rightarrow P(E) = \frac{13}{52}\frac{12}{51}\frac{13}{50}\frac{12}{49} = \frac{78}{20825}$

Si los eventos E_i que aparecen en la ecuación (c) son independientes, entonces

$$\boxed{P(E_1 \cap E_2 \cap \cdots \cap E_k) = P(E_1) \times P(E_2) \times \cdots \times P(E_k)}$$

El axioma 3 también se generaliza a:

$$\boxed{P(E_1 \cup E_2 \cup \cdots \cup E_k) = P(E_1) + P(E_2) + \cdots + P(E_k)}$$

si todos los eventos E_i son mutuamente exclusivos entre sí.

2.4.1 Teorema de Bayes

Se dice que un grupo de eventos es *colectivamente exhaustivo* si la union de todos ellos es el espacio de eventos correspondiente.

De la ec 2 4, que define las probabilidades condicionales, se puede obtener un resultado importante En un grupo de eventos colectivamente exhaustivos y mutuamente exclusivos, $B_1, B_2, \ldots, B_n$, si A es un evento cualquiera definido en el mismo espacio (fig 8), entonces, aplicando el axioma 3, resulta

$$P[A] = P[A \cap B_1] + P[A \cap B_2] + \cdots + P[A \cap B_n] = \sum_{i=1}^{i=n} P[A \cap B_i]$$

ya que los eventos $A \cap B_i$ son mutuamente exclusivos

Tomando en cuenta que $P[A \cap B_i] = P[B_i]\, P[A \mid B_i]$, se obtiene finalmente la ecuación

$$\boxed{P[A] = \sum_{i=1}^{i=n} P[B_i]\, P[A \mid B_i]}$$

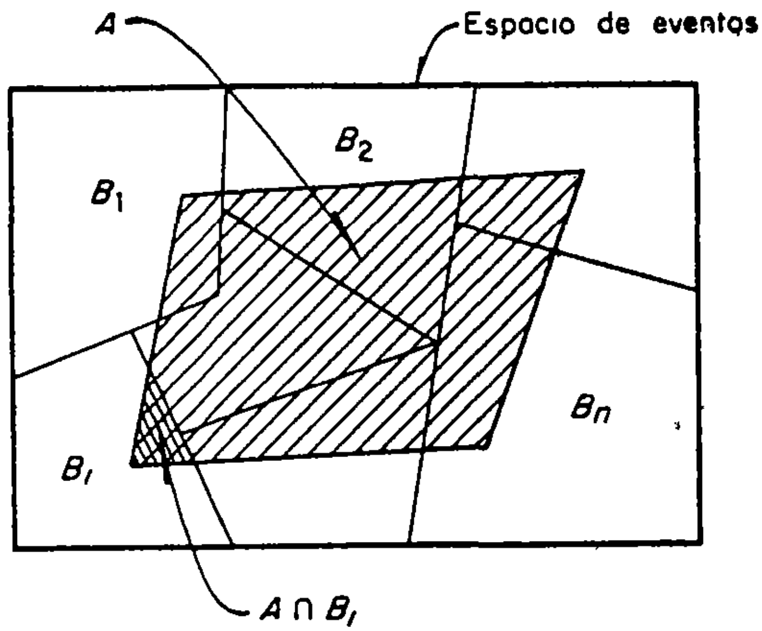


Fig 8 Eventos colectivamente exhaustivos

con la cual se define el llamado *teorema de la probabilidad total*

Considerando que $P[B_j \cap A] = P[A \cap B_j]$, se tiene que

$$P[B_j \mid A] = \frac{P[B_j \cap A]}{P[A]} = \frac{P[A \cap B_j]}{P[A]}$$

de donde

$$\boxed{P[B_j \mid A] = \frac{P[B_j]\, P[A \mid B_j]}{\sum_{i=1}^{i=n} P[B_i]\, P[A \mid B_i]}} \qquad (2\,9)$$

Este resultado se conoce como *teorema de Bayes* A las probabilidades $P[B_j]$ que se asignan a los eventos B_j antes de observar el evento A, se les denomina a priori o previas, a las probabilidades $P[B_j|A]$ que se obtienen después de observar el evento A, se les llama a posteriori o posteriores

Ejemplos: En una fábrica se reciben reguladores de voltaje de dos proveedores, B_1 y B_2, en proporción de 3 a 1; es decir, la probabilidad de que un regulador tomado al azar provenga del proveedor B_1 es $P(B_1)=\frac{3}{4}$, y del B_2 es $P(B_2)=\frac{1}{4}$.
Supongamos además que el control de calidad del proveedor B_1 es mejor que el de B_2, de manera que el 95% de los reguladores de B_1 trabajan bien, y sólo el 80% de los de B_2 funcionan correctamente. Calculemos la probabilidad de que un regulador tomado al azar funcione bien (evento A).

$$P(A|B_1)=0.95 \; ; \; P(A|B_2)=0.80$$

Del teorema de la probabilidad total:

$$P(A)=P(A|B_1)\,P(B_1)+P(A|B_2)\,P(B_2)$$

$$=0.95\times\frac{3}{4}+0.80\times\frac{1}{4}=0.9125$$

Supongamos ahora que la pregunta del problema se cambia a: ¿cuál es la probabilidad de que un regulador tomado al azar provenga del proveedor B_1 si se hizo una prueba del regulador y se observó que funciona correctamente?

$$P(B_1|A)=\frac{P(B_1)\,P(A|B_1)}{P(B_1)P(A|B_1)+P(B_2)P(A|B_2)}$$

$$=\frac{\frac{3}{4}\times 0.95}{\frac{3}{4}\times 0.95+\frac{1}{4}\times 0.80}=\frac{2.85}{3.65}=0.78$$

Además:

$$P(B_2|A)=\frac{P(B_2)\,P(A|B_2)}{\frac{3.65}{4}}=\frac{\frac{1}{4}\times 0.80}{\frac{3.65}{4}}=0.22$$

Obsérvese que

$$P(B_1/A)+P(B_2|A)=0.78+0.22=\underline{\underline{1.00}}$$

<u>Ejemplo</u>.- Supóngase que una prueba para detectar diabetes tiene una eficiencia del 95%, es decir, sólo en el 95% de los casos detecta la diabetes en una persona que la padece. Supongase tam-

... que ... pruebas
que resultan positivas son de gente sana, y que el 3% de la población de una región de México padece esta enfermedad.

a.- ¿Cuál es la probabilidad de que una persona seleccionada al azar pueda ser declarada diabética por la prueba?

b.- Si la prueba dice que sí es diabética, ¿cuál es la probabilidad de que realmente lo sea?

$B_1 = \{\text{tiene diabetes}\}$; $B_2 = \{\text{no tiene diabetes}\}$

$E = \{\text{la prueba detecta diabetes}\}$

$P(B_1) = 0.03$, $P(B_2) = 0.97$

$P(E|B_1) = 0.95$, $P(E|B_2) = 0.02$

a.- $$P(E) = P(E|B_1)P(B_1) + P(E/B_2)P(B_2)$$
$$= 0.95 \times 0.03 + 0.02 \times 0.97 = 0.0479$$

b.- $$P(B_1|E) = \frac{P(B_1)P(E|B_1)}{P(B_1)P(E|B_1) + P(B_2)P(E|B_2)}$$
$$= \frac{0.03 \times 0.95}{0.03 \times 0.95 + 0.97 \times 0.02} = 0.59$$

Ejemplo

Existe un edificio de concreto reforzado para el cual se investiga su capacidad de carga de diseño Un ingeniero, con base en su experiencia personal, decide que la resistencia nominal, f'_c , del concreto pudo ser de 150 kg/cm², 200 kg/cm² o 250 kg/cm², con las siguientes probabilidades previas

Resistencia nominal, en kg/cm²	Probabilidad
150	0.3
200	0.6
250	0.1

Para predecir la resistencia nominal real, es necesario realizar un experimento que consiste en extraer corazones (muestras) del concreto de la estructura y probarlos a compresion simple. El ingeniero decide que la resistencia, S, de un solo corazón dará una predicción confiable, y define resistencias intermedias de 175 y 225 kg/cm² para tomar en cuenta que las obtenidas de los corazones serían mayores que las que se hubieran logrado en una prueba a los 28 dias de edad del concreto De acuerdo con esto, el ingeniero asigna las siguientes probabilidades condicionales a cada ~~resistencia nominal dado el resultado,~~ evento Z_i, ~~del experimento~~ dada la resistencia nominal, f'_c:

Resistencia del corazon, S, en kg/cm²	$P(Z_i \mid f'_c)$		
	$(f'_c)_1 = 150$	$(f'_c)_2 = 200$	$(f'_c)_3 = 250$
Evento Z_1 $\{S < 175\}$	0.7	0.2	0
Evento Z_2 $\{175 < S < 225\}$	0.3	0.6	0.3
Evento Z_3 $\{S > 225\}$	0	0.2	0.7

Supóngase que se saca un corazón y que su resistencia resulta ser de 164 kg/cm², es decir, que ocurre el evento Z_1 Las probabilidades a posteriori de las resistencias nominales son, entonces

$$P[150 \mid Z_1] = \frac{(0.7)\,(0.3)}{(0.7)\,(0.3) + (0.2)\,(0.6) + (0)\,(0.1)} = \frac{0.21}{0.33} = 0.635$$

$$P[200 \mid Z_1] = \frac{(0.2)\ (0.6)}{0\ 33} = 0\ 365$$

$$P[250 \mid Z_1] = \frac{(0)\ (0\ 1)}{0\ 33} = 0$$

Es decir, este resultado elimina la posibilidad de que $f'_c = 250$ kg/cm²

Supóngase ahora que en vez de un solo corazón, el ingeniero hubiese decidido obtener dos, situados en diferentes niveles de la estructura, y que al probarlos en uno ocurrió Z_1 y en el otro Z_2. La probabilidad de que ocurran ambos eventos, (Z_1, Z_2) si f'_c es realmente 150, 200 o 250 kg/cm², será el producto de dos probabilidades condicionales, puesto que Z_1 y Z_2 son independientes

$$P[(Z_1, Z_2) \mid 150] = P[Z_1 \mid 150]\ P[Z_2 \mid 150] = (0\ 7)\ (0\ 3) = 0\ 21$$

$$P[(Z_1, Z_2) \mid 200] = P[Z_1 \mid 200]\ P[Z_2 \mid 200] = (0.2)\ (0.6) = 0\ 12$$

$$P[(Z_1, Z_2) \mid 250] = P[Z_1 \mid 250]\ P[Z_2 \mid 250] = (0.0)\ (0\ 3) = 0\ 0$$

Las probabilidades a posteriori son, entonces

$$P[150 \mid (Z_1, Z_2)] = \frac{(0\ 3)\ (0.21)}{(0.3)\ (0.21) + (0.6)\ (0.12) + (0.1)\ (0.0)} =$$

$$= \frac{0\ 063}{0.063 + 0\ 072 + 0\ 0} = \frac{0\ 063}{0.135} = 0.47$$

$$P[200 \mid (Z_1, Z_2)] = \frac{(0.6)\ (0.12)}{0\ 135} = \frac{0\ 072}{0\ 135} = 0\ 53$$

$$P[250 \mid (Z_1, Z_2)] = \frac{(0.1)\ (0.0)}{0\ 135} = 0$$

Los mismos resultados se habrían obtenido si el ingeniero, despues de extraer el primer corazón y de calcular las probabilidades posteriores correspondientes, hubiera decidido sacar el segundo y recalcular dichas probabilidades con base en las obtenidas para el primero, es decir, las probabilidades previas en el segundo calculo serian 0 635, 0 365 y 0 para 150 , 200 y 250 kg/cm², respectivamente En tal caso, las probabilidades posteriores, dado que ocurrió Z_2, son

$$P[150 \mid Z_2] = \frac{(0.3)\ (0.635)}{(0\ 3)\ (0\,635) + (0\ 365)\ (0\ 6) + (0\ 1)\ (0)} =$$

$$= \frac{0\ 19}{0\ 14 + 0\ 22 + 0\ 0} = \frac{0\ 19}{0\ 41} = 0\ 47$$

$$P[200 \mid Z_2] = \frac{0\ 22}{0\ 41} = 0\ 53$$

$$P[250 \mid Z_2] = \frac{0}{0\ 41} = 0$$

que son iguales a las anteriores

De aqui se concluye que las probabilidades se pueden actualizar conforme se va obteniendo nueva informacion experimental

2 5 Variables aleatorias

Una *variable aleatoria* es una variable tal que no puede predecirse con certidumbre el valor que asumira antes de realizar un experimento Por ejemplo, los valores de la tercera columna de la tabla 2 corresponden a resultados experimentales asociados a la resistencia ultima de las vigas, en este caso, la variable aleatoria es precisamente la resistencia última, ya que antes de romper una viga no se puede precisar cual será su resistencia A fin de evitar confusion, se empleará una letra mayúscula para denotar una variable aleatoria, y la minuscula correspondiente para los valores que puede asumir

Cuando el numero de valores que una variable aleatoria puede tomar está restringido a un numero finito o infinito, pero numerable, dicha variable se llama *discreta o discontinua* En caso contrario se designa *variable continua*

El comportamiento de una variable aleatoria se describe mediante su *ley de probabilidades*, la cual a su vez puede definirse de diferentes formas La manera más común de hacerlo es mediante su *distribución* o *densidad de probabilidades*

Si la variable aleatoria X es discreta y puede asumir los valores x_i, su densidad de probabilidades será el conjunto de las probabilidades

$$\boxed{P_X(x_i) = P[X = x_i]} \qquad (2.10)$$

la cual se lee "probabilidad de que $X = x_i$"

La <u>distribución de probabilidades acumuladas</u> o <u>función de distribución</u>, que es otra forma de especificar la ley de probabilidades de una variable aleatoria, es el conjunto de las sumas parciales de $P_X[x_i]$ para todos los valores de X menores que x_i. Esta distribución permite conocer la probabilidad de que la variable aleatoria tome valores menores o iguales que un numero dado x_m, es decir

$$\boxed{F_X(x_m) = P[X \leq x_m] = \sum_{i=1}^{i=m} P_X[x_i]} \qquad (2.11)$$

Ejemplo:-

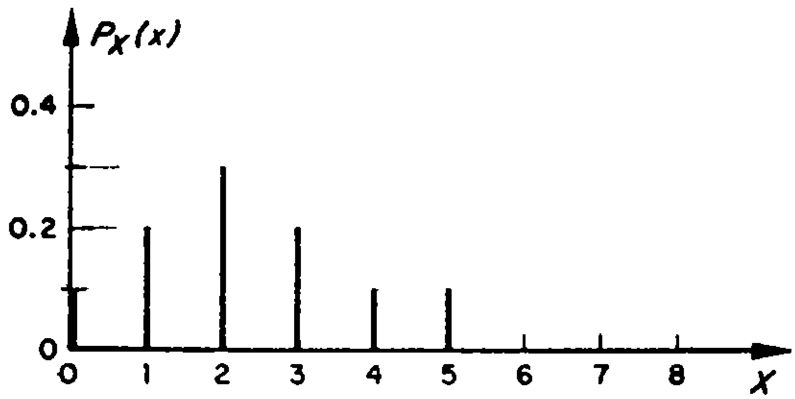
Sea X la variable aleatoria discreta "número total de carros que se detienen en una esquina debido a la luz roja de un semáforo". Si las probabilidades asociadas a cada evento son

$$P_X(x) = \begin{cases} 0.1 & \text{si} \quad x = 0 \\ 0.2 & \text{si} \quad x = 1 \\ 0.3 & \text{si} \quad x = 2 \\ 0.2 & \text{si} \quad x = 3 \\ 0.1 & \text{si} \quad x = 4 \\ 0.1 & \text{si} \quad x = 5 \\ 0 & \text{si} \quad x \geq 6 \end{cases}$$

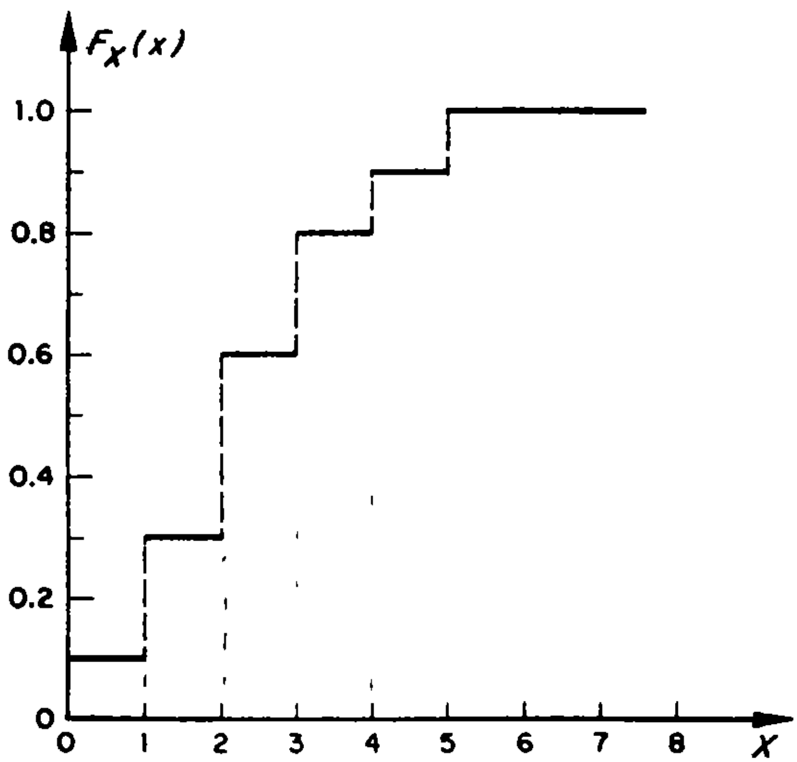
la distribución de probabilidades acumuladas correspondiente será

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 0.1 & \text{si } 0 \leq x < 1 \\ 0.1 + 0.2 = 0.3 & \text{si } 1 \leq x < 2 \\ 0.3 + 0.3 = 0.6 & \text{si } 2 \leq x < 3 \\ 0.6 + 0.2 = 0.8 & \text{si } 3 \leq x < 4 \\ 0.8 + 0.1 = 0.9 & \text{si } 4 \leq x < 5 \\ 0.9 + 0.1 = 1.0 & \text{si } 5 \leq x \end{cases}$$

Las gráficas de estas distribuciones se presentan en la fig. 9



a) Distribución de probabilidades



b) Función de distribución

Fig 9 Ley de probabilidades del ejemplo del tráfico

Para que una densidad de probabilidades satisfaga los tres axiomas de la teoría de probabilidades, se deben cumplir los siguientes requisitos

$$\left.\begin{array}{ll} \text{a)} & 0 \leq P_X(x) \leq 1 \text{ para toda } x_i \\ \text{b)} & \sum_{i=1}^{n} P_X(x_i) = 1, \text{ donde } n \text{ es el numero total de valores que puede asumir } X \\ \text{c)} & P[x_m \leq X \leq x_r] = \sum_{i=m}^{i=r} P_X(x_i) \end{array}\right\} \quad (2\ 12)$$

En el caso de una variable aleatoria continua, X, la probabilidad de que esta tome un valor comprendido entre x y $x + dx$ está dada por $f_X(x)\,dx$, donde $f_X(x)$ es la densidad de probabilidades de X. Por tanto, la probabilidad de que X asuma valores comprendidos en el intervalo $x_1 \leq X \leq x_2$ es

$$\boxed{P[x_1 \leq X \leq x_2] = \int_{x_1}^{x_2} f_X(x)\,dx} \qquad (2\ 13)$$

Graficamente, esta probabilidad es igual al área bajo la curva de $f_X(x)$ comprendida entre x_1 y x_2 como se muestra en la fig 10a para la variable Y.

Puesto que $F_X(x) = P(X \leq x) = P(-\infty \leq X \leq x)$, y en virtud de la ec 2.13, se tiene que

$$\boxed{F_X(x) = \int_{-\infty}^{x} f_X(u)\,du} \qquad (2\ 14)$$

donde u es solo una variable muda de integración El valor de esta integral es igual al área bajo la curva de $f_X(x)$ a la izquierda de x De la ec 2.14 se concluye que

$$\boxed{\frac{dF_X(x)}{dx} = \frac{d}{dx}\left[\int_{-\infty}^{x} f_X(u)\,du\right] = f_X(x)} \qquad (2.15)$$

Las propiedades generales de $F_X(x)$ son:

$$0 \leq F_X(x) \leq 1$$
$$F_X(-\infty) = 0$$
$$F_X(\infty) = 1 \qquad (2.16)$$
$$F_X(x + \epsilon) \geq F_X(x), \text{ si } \epsilon \geq 0$$
$$F_X(x_2) - F_X(x_1) = P[x_1 \leq X \leq x_2]$$

Para satisfacer los axiomas de la teoría de probabilidades se necesita que

$$\boxed{\begin{gathered} f_X(x) \geq 0 \text{ para toda } x \\ \int_{-\infty}^{\infty} f_X(x)\, dx = 1 \end{gathered}} \qquad (2.17)$$

Por ejemplo, si la densidad de probabilidades de una variable aleatoria continua es de forma triangular (fig 10a) y está dada por

$$f_Y(y) = \begin{cases} (y - 35)/60 & \text{si } 35 \leq y \leq 41 \\ (55 - y)/140 & \text{si } 41 \leq y \leq 55 \\ 0 & \text{si } y \leq 35 \text{ o } y > 55 \end{cases}$$

entonces el cálculo de la distribución de probabilidades acumuladas se efectúa aplicando la ec 2.14 a los diversos intervalos de Y en donde cambia la ecuación de $f_Y(y)$, de la manera siguiente

Si $y < 35$

$$F_Y(y) = 0$$

ya que $f_Y(y) = 0$ si $y < 35$

Si $35 \leq y \leq 41$

$$F_Y(y) = P[Y \leq y] = \int_{-\infty}^{y} f_Y(u)\, du = \frac{1}{60} \int_{35}^{y} (u-35)\, du = (y-35)^2/120$$

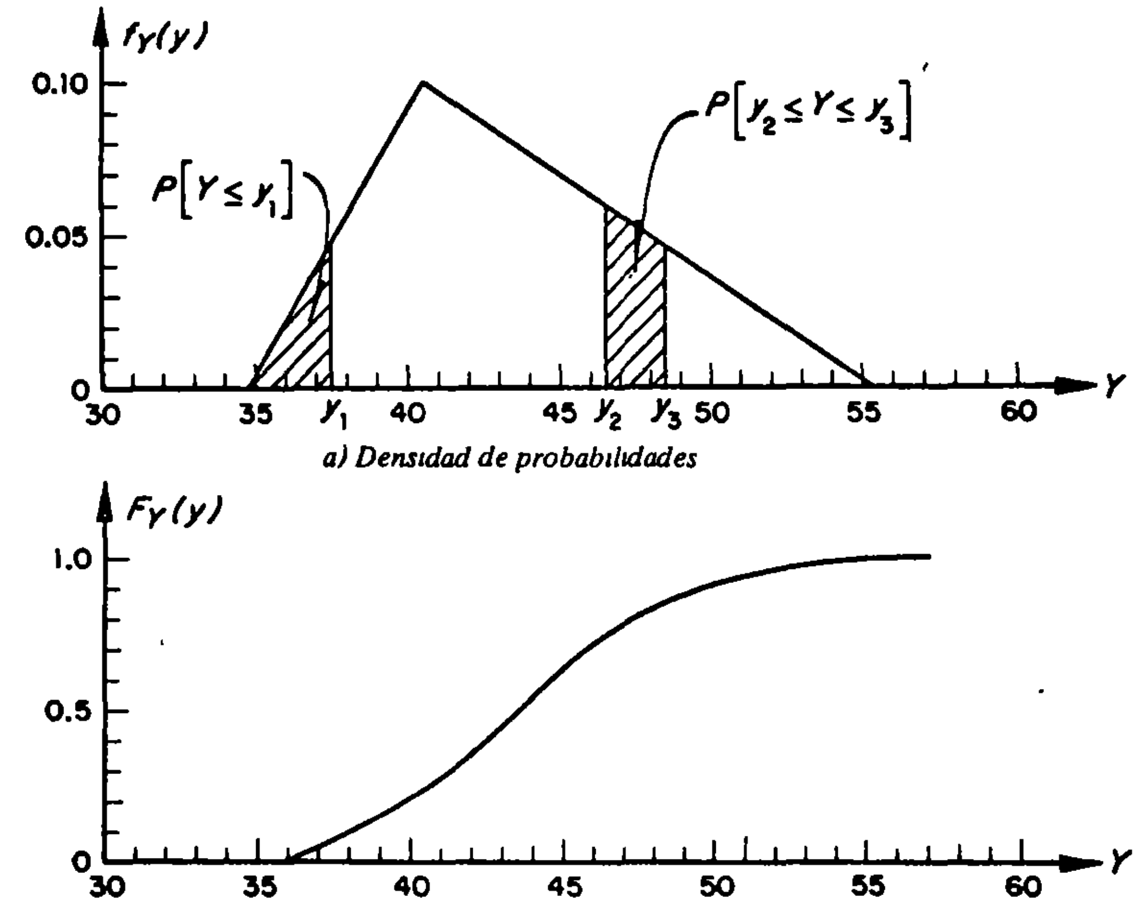
Si $41 < y < 55$

$$F_Y(y) = (41-35)^2/120 + \frac{1}{140}\int_{41}^{y}(55-u)\,du = 3/10 + [55\,(y-41) - \frac{1}{2}(y^2 - 41^2)]/140$$

Si $y > 55$

$$F_Y(y) = 1$$

La fig 10b representa la gráfica de esta función de distribución



a) Densidad de probabilidades

b) Función de distribución

Fig 10. Ley de probabilidades de una variable aleatoria continua

Si se quiere calcular la probabilidad de obtener un valor de la variable comprendido en el intervalo de y_2 a y_3, donde $y_2 > 41$ y $y_3 < 55$, entonces

$$P[y_2 < Y < y_3] = \frac{1}{140}\int_{y_2}^{y_3}(55-y)\,dy = \frac{55\,(y_3 - y_2) - \frac{1}{2}(y_3^2 - y_2^2)}{140}$$

Ejemplo

Un ingeniero esta interesado en diseñar una torre que resista las cargas debidas al viento. De una serie de observaciones de la máxima velocidad anual del viento cerca del sitio de interés, se encuentra que el histograma puede ajustarse, desde un punto de vista estadístico, mediante una distribución de probabilidades exponencial de la forma

$$f_X(x) = Ke^{-\lambda x}\ ,\ x > 0$$

donde X es la máxima velocidad del viento, λ es una constante y K es otra constante tal que obliga a que $f_X(x)$ satisfaga la ec 2 17. Por tanto

$$\int_0^{\infty} Ke^{-\lambda x}\, dx = \frac{-K}{\lambda}\ [e^{-\lambda x}]_0^{\infty} = \frac{K}{\lambda} = 1$$

de donde

$$K = \lambda$$

por tanto

$$\underline{f_X(x) = \lambda e^{-\lambda x}}\ ,\ x > 0$$

Distribución exponencial

La función de distribución será

$$F_X(x) = \int_0^x f_X(u)\, du = \int_0^x \lambda e^{-\lambda u}\, du = [-e^{-\lambda u}]_0^x = 1 - e^{-\lambda x}\ ,\ x > 0$$

El valor de λ se puede tomar, por ejemplo, de manera que $F_X(x)$ se ajuste para que coincida con un valor empírico. Así, si la frecuencia relativa del evento $A = \left\{X < 70 \text{ km/h}\right\}$ es 0.9, entonces

$$P\,[0 < X < 70] = 0\,9$$

de donde

$$0\,9 = 1 - e^{-70\lambda}$$

por lo cual $\lambda = 0.033$. En la fig 11 se presentan las gráficas de $f_X(x)$ y $F_X(x)$.

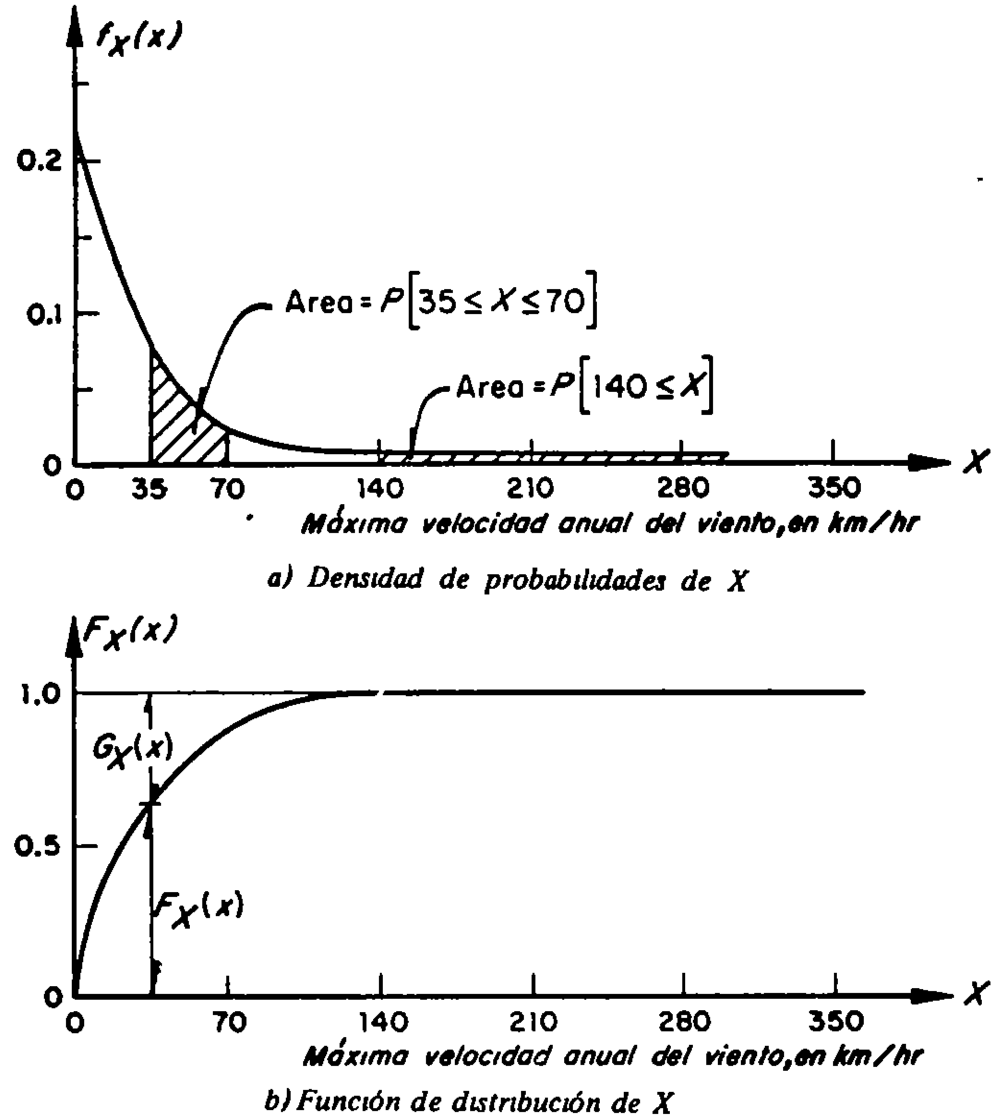


Fig 11 Ley de probabilidades correspondiente al ejemplo de la máxima velocidad anual del viento

El complemento de $F_X(x)$ se utiliza cuando las decisiones se toman con base en probabilidades de que se exceda un valor dado de la variable.

La función de distribución complementaria se define como

$$\boxed{G_X(x) = P[X > x] = 1 - F_X(x)} \qquad (2.18)$$

Así, para el ejemplo anterior, se tiene

$$G_X\ (x) = e^{-0.033x}$$

2.6 Esperanzas

En el capítulo anterior se indicó como calcular el promedio aritmético y la variancia de una muestra. Estos parámetros constituyen, respectivamente, la media y la variancia de una variable aleatoria, los cuales se definen a continuación

La media, m_X, o *esperanza, E* $[X]$, de una variable aleatoria discreta X, es

$$\boxed{m_X = E\,[X] = \sum_{i=1}^{i=N} x_i\, P_X\ (x_i)} \qquad (2.19)$$

donde N es el total de valores que X puede asumir

Para el caso de una variable aleatoria continua, X, la media es

$$\boxed{m_X = \mathrm{E}\,[X] = \int_{-\infty}^{\infty} x\, f_X\,(x)\, dx} \qquad (2.20)$$

Recuérdese que el promedio aritmético de una muestra esta dado por

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{i=n} x_i$$

Si al realizar un experimento se encuentran n_j valores de x_j, la ecuación anterior puede escribirse

$$\bar{x} = \sum_{j=1}^{j=N} \frac{n_j}{n}\, x_j = \sum_{j=1}^{j=N} f_j\ x_j \qquad (2.21)$$

donde f_j es la frecuencia relativa de x_j observada en la muestra, y N es el numero de valores x_j distintos. Es evidente el gran parecido en la forma de las ecs 2.19 y 2.21, la diferencia estriba en que f_j se valua a partir de una muestra de la variable aleatoria X, y $P_X(x)$ es la distribucion de probabilidades que la define Cuando el total de observaciones es grande, $\bar{x}$ tendra un valor cercano a m_X, ya que se puede demostrar que conforme n aumenta, f_j tiende a la probabilidad de que ocurra el valor x_j. En general, $\bar{x}$ se toma como una estimación de m_X.

Ejemplo

Si la densidad de probabilidades de la variable aleatoria X correspondiente a los errores en una nivelación es la de la segunda columna de la siguiente tabla, la media de dicha variable resulta ser 4 167 micras Los cálculos correspondientes se localizan en la tercera columna

x_i, en micras	$P_X(x_i)$	$x_i\ P_X(x_i)$, en micras
0	6/60	0
1 000	2/60	2 000/60
2 000	4/60	8 000/60
3 000	8/60	24 000/60
4 000	13/60	52 000/60
5 000	12/60	60 000/60
6 000	7/60	42 000/60
7 000	4/60	28 000/60
8 000	2/60	16 000/60
9 000	2/60	18 000/60
TOTAL		$m_X = 250\,000/60 = 4\,167$ micras

Una medida muy común de la dispersión de los valores que puede asumir una variable aleatoria es la *variancia*, la cual se denota como S^2_X o Var $[X]$, definiéndose para una variable aleatoria discreta, como

$$\boxed{S^2_X = \text{Var}\,[X] = \sum_{i=1}^{i=N} (x_i - m_X)^2\ P_X(x_i)} \qquad (2.22)$$

y para una continua

$$\boxed{S^2_X = \text{Var}\,[X] = \int_{-\infty}^{\infty} (x - m_X)^2\ f_X(x)\,dx} \qquad (2\ 23)$$

Se puede realizar un razonamiento parecido al del caso de la esperanza de X para mostrar la semejanza de las ecs 2 22 y 2 23

Por definición, la *desviación estándar*, S_X, de la variable aleatoria X, es igual a la raíz cuadrada de la variancia, y el *coeficiente de variación* es

$$\boxed{v_X = S_X / m_X} \qquad (2\ 24)$$

En la siguiente tabla se calcula la variancia de la variable aleatoria, cuya densidad de probabilidades se presento en el ejemplo anterior

$x_i - m_X$, en micras	$P_X(x_i)$	$(x_i - m_X)^2 P_X(x_i)$, en micras^2
−4 167	6/60	1 740 000
−3 167	2/60	333 000
−2 167	4/60	313 000
−1 167	8/60	181 000
− 167	13/60	6 000
833	12/60	139 000
1 833	7/60	390 000
2 833	4/60	531 000
3 833	2/60	487 000
4 833	2/60	687 000
	TOTAL	4 798 000 micras^2 = S_X^2

La desviación estándar y el coeficiente de variación son, respectivamente

$$S_X = \sqrt{4\,798\,000} = 2\,200 \text{ micras, y } v_X = S_X / m_X = \frac{2\,200}{4\,167} = 0\ 528$$

La *esperanza* de una función $g(X)$, de la variable aleatoria continua, X, es, por definición

$$\boxed{E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx} \qquad (2.25)$$

o para una variable aleatoria discreta

$$\boxed{E[g(X)] = \sum_{i=1}^{i=N} g(x_i) P_X(x_i)} \qquad (2\ 26)$$

En particular, si $g(X) = X$, se tiene

$$E[X] = \int_{-\infty}^{\infty} x\, f_X(x)\, dx$$

que es la esperanza o media de X

Si $g(X) = (X - m_X)^2$, entonces

$$E[(X - m_X)^2] = \int_{-\infty}^{\infty} (x - m_X)^2 f_X(x)\, dx$$

que es la variancia de X

La esperanza de $g(X) = x^2$ es

$$\boxed{E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx} \qquad (2.27)$$

y se denomina valor medio cuadratico

Se puede demostrar que la esperanza, la variancia y el valor medio cuadratico no son independientes, y que conociendo dos de ellos es factible calcular el tercero La relación entre estos parametros es

$$\boxed{S_X^2 = E[X^2] - (E[X])^2} \qquad (2\ 28)$$

La media y la variancia juegan el mismo papel para las distribuciones de probabilidades que el centroide y el momento de inercia para las distribuciones de masa. Al respecto, considérese una barra de *masa unitaria*, con densidad de masa variable (masa por unidad de longitud) $\rho(x)$, identica a una densidad de probabilidades $f_X(x)$ La distancia, x_c, del origen al centroide de la barra se calcula mediante una ecuacion similar a la 2 20. El momento de inercia respecto al centroide, se obtiene usando una ecuación semejante a la 2 28 para el calculo de la variancia de X

Entre las propiedades de la esperanza de $g(X)$ se cuentan las siguientes

1 Si $g(X) = \text{constante} = c$

$$E[c] = c \int_{-\infty}^{\infty} f_X(x)\, dx = c$$

$$\text{Var}[c] = 0$$

2. Si $g(X) = cx$

$$E[cx] = \int_{-\infty}^{\infty} x f_X(x)\, dx = c E[X]$$

$$\text{Var}[cX] = c^2 \text{ Var}[X]$$

3 Si $g(X) = a + bx$

$$E[a + bx] = a \int_{-\infty}^{\infty} f_X(x)\, dx + b \int_{-\infty}^{\infty} x f_X(x) dx = a + bE[X]$$

$$\text{Var}[a + bx] = b^2 \quad \text{Var}[X]$$

4. Si $g(X) = g_1(X) + g_2(X)$

$$E[g_1(X) + g_2(X)] = \int_{-\infty}^{\infty} g_1(x) f_X(x)\, dx + \int_{-\infty}^{\infty} g_2(x) f_X(x)\, dx$$

$$= E[g_1(X)] + E[g_2(X)]$$

5 $E[X + Y] = E[X] + E[Y]$

$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$, si X y Y son estadísticamente independientes.

2.7 Distribuciones particulares

2.7.1 Distribución binomial o de Bernoulli

La distribución binomial es una de las más útiles para variables aleatorias discretas, y corresponde a un experimento en el que solo hay dos resultados posibles.

Sea p la probabilidad de ocurrencia (éxito) de un evento, y $q = 1 - p$ la de que este no ocurra (fracaso) en una realización de un experimento. Si X es una variable aleatoria que corresponde al número de éxitos en n intentos independientes, se puede demostrar que X tiene *distribución binomial*, es decir, que su densidad de probabilidades es

$$\boxed{f_X(x) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, 2, \ldots, n} \tag{2 29}$$

donde $\binom{n}{x}$ denota el número de combinaciones de n elementos tomados de x en x, y se calcula con la expresión

$$\binom{n}{x} = \frac{n!}{x!\ (n-x)!} \tag{2.30}$$

Se puede demostrar que la esperanza y la varianza para esta distribución son

$$E(x) = \sum_{x=0}^{n} x \binom{n}{x} p^{x} (1-p)^{n-x}$$

$$E[X] = np$$

$$Var[X] = npq$$

$$Var[x] = \sum_{x=0}^{n} (x-np)^{2} \binom{n}{x} p^{x} q^{n-x}$$

Ejemplo

Si se lanza al aire seis veces una moneda homogenea,

a) ¿Cuál es la probabilidad de obtener dos "caras"?

b) ¿Cuál es la probabilidad de obtener por lo menos cuatro "caras" ($X > 4$)?

a) Puesto que la moneda es homogénea se tiene $p = 1/2$ y $q = 1 - 1/2 = 1/2$, donde p es la probabilidad de observar "cara" en un lanzamiento. Por tanto

$$P[X = 2] = f_X(2) = \binom{6}{2} (1/2)^2 (1/2)^{6-2} = \frac{6!}{2!\ 4!} (1/2)^6 = \frac{15}{64}$$

$$E[X] = np = 6\ (1/2) = 3$$

$$\text{Var}[X] = npq = 6\ (1/2)\ (1/2) = 3/2$$

b) Para que se cumpla $X > 4$ en seis lanzamientos, se necesita que se observen 4, 5 o 6 caras. Puesto que estos tres eventos son mutuamente exclusivos, se tiene

$$P[X > 4] = f_X(4) + f_X(5) + f_X(6)$$

Calculando los tres sumandos como en la pregunta anterior, resulta

$$P[X > 4] = \frac{6!}{4!\ 2!}(1/2)^4 (1/2)^{6-4} + \frac{6!}{5!\ 1!}(1/2)^5 (1/2)^{6-5} + \frac{6!}{6!\ 0!}(1/2)^6 (1/2)^{6-6} = \frac{15}{64} +$$

$$+ \frac{6}{64} + \frac{1}{64} = \frac{11}{32}$$

DISTRIBUCION GEOMETRICA

Sea p la probabilidad de éxito en un experimento. Si se repite el experimento sucesivamente hasta que se observa un éxito se tendrá la variable aleatoria X = número de repeticiones del experimento hasta que se observa el primer éxito. Obtengamos la densidad de probabilidades de X.

El 1er éxito ocurrirá en el experimento número x si, y sólo si, en los $x-1$ anteriores hubo puros fracasos. La probabilidad de este evento, dado que los experimentos son independientes, es

$$\boxed{f_X(x) = (1-p)^{x-1}\, p}$$

distribución geométrica

Se puede demostrar que la distribución de probabilidades acumuladas es

$$\boxed{F_X(x) = 1-(1-p)^{n}}$$

y que $\boxed{E[X] = 1/p}$ y $\boxed{\sigma_X^2 = \dfrac{1-p}{p^2}}$

$$E[X] = \sum_{x=0}^{\infty} x\,(1-p)^{x-1} p \qquad \sigma_X^2 = \sum_{x=0}^{\infty} \left(x-\frac{1}{p}\right)^2 (1-p)^{x-1} p$$

DISTRIBUCION HIPERGEOMETRICA

Cuando se tiene una variable aleatoria discreta cuyo espacio de eventos tiene sólo 2 elementos, digamos $S=\{\text{éxito, fracaso}\}$, y se realiza un muestreo sin reemplazo, entonces los resultados de cada experimento no son independientes ni la probabilidad de éxito permanece constante, como en la distribución binomial, por lo que ésta última no es aplicable.

En este caso, si x es el número de "éxitos" observados de un lote que tiene N objetos de los cuales M son "éxitos", la probabilidad de observar x en n repeticiones del experimento es:

$$\overbrace{\underbrace{\bullet\bullet\bullet\bullet\bullet\bullet\bullet}_{\substack{M \\ \text{"éxitos"}}}\;\underbrace{\bullet\bullet\bullet\bullet\bullet\bullet}_{\substack{N-M \\ \text{"fracasos"}}}}^{N} \Rightarrow n(S)=\binom{N}{n}={}_N C_n$$

$$\text{"éxitos"} \downarrow \; {}_M C_x = \binom{M}{x} \qquad \text{"fracasos"} \downarrow \; {}_{N-M} C_{n-x} = \binom{N-M}{n-x}$$

Si $A=\{X=x\}$

$$n(A)=\binom{M}{x}\times\binom{N-M}{n-x}$$

$$\boxed{P(X=x)=\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}}$$

$x=0,1,2,\ldots,M$

$$\boxed{E(x)=nM/N}$$

$$\boxed{\sigma^2(x)=\frac{M(N-M)n(N-n)}{N^2(N-1)}}$$

Ejemplo.- En un problema de control estadístico de colidad, se tiene un lote de 100 transformadores de corriente eléctrica, de los cuales 40 son defectuosos (no cumplen las normas de fabricación). ¿Cuál es la probabilidad de obtener uno defectuoso de tres seleccionados al azar sin reemplazo?

$$P[X=3] = \frac{\binom{40}{1}\binom{100-40}{3-1}}{\binom{100}{3}} = \frac{\binom{40}{1}\binom{60}{2}}{\binom{100}{3}}$$

$$= \frac{\dfrac{40!}{39! \times 1!} \times \dfrac{60!}{58! \times 2!}}{\dfrac{100!}{97! \times 3!}} = 0.438$$

Cuando <u>N es grande</u> y <u>n pequeño</u>, la distribución binomial se puede usar <u>como aproximación</u> de la hipergeométrica, cuando los cálculos con esta distribución resultan tediosos.

2.7.2 Distribución de Poisson

Una distribución de probabilidades para una variable aleatoria discreta, X, de la forma

$$\boxed{f_X(x) = \frac{\lambda^x \, e^{-\lambda}}{x!}, \quad x = 0, 1, 2, \ldots} \tag{2.32}$$

se llama *distribución de Poisson*, en la ec 2 32, λ es una constante Se puede demostrar que la media y la varianca para esta distribución quedan dadas por

$$\boxed{\begin{aligned} E[X] &= \lambda \\ \text{Var}[X] &= \lambda \end{aligned}} \tag{2.33}$$

Ejemplo

Si la probabilidad de que falle una varilla de acero al aplicarle una determinada fuerza de tensión es de 0.001, ¿cual es la probabilidad de que de 2 000 varillas probadas fallen a) tres, b) más de dos, si se supone que la ~~resistencia~~ falla de las varillas tiene una distribución de Poisson con λ = 2000x0·001=2 ?

a) $$P[X = 3] = \frac{\lambda^3 \, e^{-\lambda}}{3!}$$

$$P[X = 3] = \frac{2^3 \, e^{-2}}{3!} = 0.18$$

b) $$P[X > 2] = 1 - P[X \le 2] = 1 - F_X(2) = 1 - \{P[X = 0] +$$

$$+ P[X = 1] + P[X = 2]\} = 1 - \frac{2^0 \, e^{-2}}{0!} - \frac{2^1 \, e^{-2}}{1!} - \frac{2^2 \, e^{-2}}{2!} =$$

$$= 1 - \frac{1}{e^2} - \frac{2}{e^2} - \frac{2}{e^2} = 1 - \frac{5}{e^2} = 0.323$$

Es posible demostrar que la distribución de Poisson puede emplearse como una aproximación de la de Bernoulli cuando n es grande y p pequeña, pero de tal manera que $npq >> 1$, tomando $\lambda = np$ Al respecto, si $n = 20$ y $p = 0.05$, entonces npq vale 0.95, y aun cuando no cumple con la última condicion, el error que se tiene al usar dicha aproximación es menor de 3 por ciento para valores de X menores de 3, aun cuando npq es casi uno, para $X = 4$ y $X = 5$ los errores respectivos son 15 y 41 por ciento

<u>Ejemplo</u>.- Una compañía aseguradora después de muchos años de experiencia ha estimado que el 0.004 % de la población fallece por un tipo específico de accidente. Si esta compañía tiene 40,000 asegurados, ¿cuál es la probabilidad de que 2 de ellos muera en un año por tal tipo de accidente?

Sea X el número de personas que mueren anualmente entre los asegurados, por accidente. La media de X es

$$E[X] = 0.00004 \times 40{,}000 = 1.6 = \lambda$$

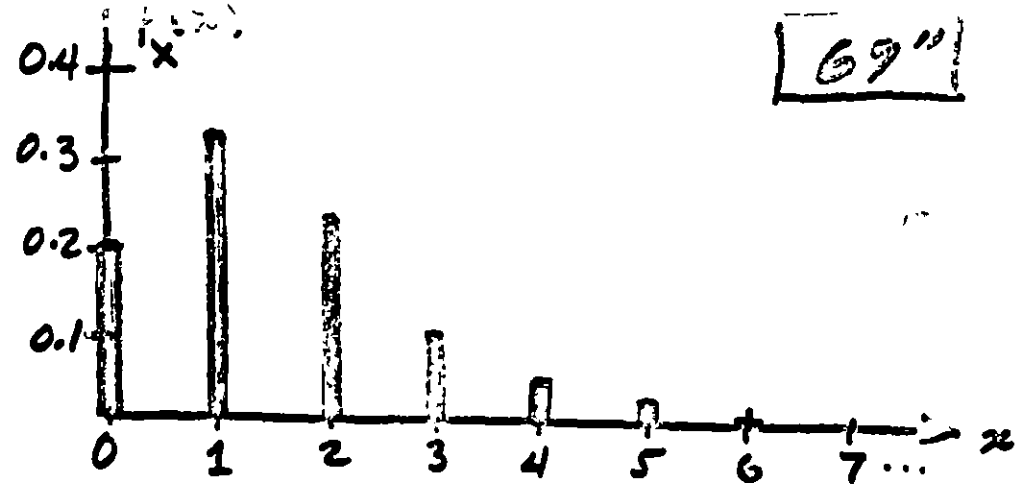
$$P[X=x] = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{(1.6)^x e^{-1.6}}{x!} \quad ; \; x = 0, 1, 2, \ldots$$

$$\therefore \; P[X=2] = \frac{(1.6)^2 e^{-1.6}}{2!} = \frac{0.2019 \times 2.56}{2} = 0.26$$

Para este problema la distribución es:

x	0	1	2	3	4	5	6	...
$f_X(x)$	0.202	0.323	0.258	0.138	0.055	0.018	0.005	...

La gráfica de esta distribución es:



Ejemplo.- En el carril para dar vuelta a la izquierda en una avenida, sólo hay capacidad para 3 autos como máximo esperando la flecha luminosa del semáforo. En un estudio estadístico del tránsito en ese lugar se encontró que en cada ciclo del semáforo hay en promedio 6 autos que van a dar vuelta. ¿Cuál es la probabilidad de que en un ciclo del semáforo tomado al azar se congestione el tránsito por excederse la capacidad del carril? $P[X > 3] = ?$

Si $A = \{X > 3\}$, $\bar{A} = \{X \leq 3\}$

$$P(\bar{A}) = 1 - P(A) \therefore P(A) = 1 - P(\bar{A}); \ \lambda = 6$$

$$P(\bar{A}) = P[X \leq 3] = \sum_{x=0}^{x=3} f_X(x) = \sum_{x=0}^{x=3} \frac{e^{-6} 6^x}{x!}$$

$$P(\bar{A}) = e^{-6}\left(1 + 6 + \frac{6^2}{2} + \frac{6^3}{6}\right) = 61 e^{-6} = 0.152$$

$$P[A] = P[X > 3] = 1 - 0.152 = \underline{\underline{0.848}}$$

PROCESO DE POISSON

Con base en la distribución de Poisson se puede deducir que la distribución de probabilidades del número de ocurrencias de un evento durante un periodo t queda dada por

$$f_X(x) = P[X = x \text{ en un lapso } t]$$

$$\boxed{f_X(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}} \; ; \; x = 0, 1, 2, \ldots$$

donde λ = número medio de ocurrencias por unidad de tiempo. Para que esta distribución se aplique se requiere que el~~os~~ evento~~s~~ cada vez ~~sean tales~~ ocurra de manera independiente de las ocurrencias previas y que λ sea constante.

<u>Ejemplo</u>.- En una central de comunicaciones se tiene una demanda media del servicio de 8 cada minuto. Calcular la probabilidad de que en 2 minutos no se solicite el servicio.

$$\lambda = 8,\; t = 2,\; f_X(0) = P[X=0] = \frac{(\lambda t)^0 e^{-8\times 2}}{0!} = e^{-16}$$

$$= \underline{0.00004}$$

Además:

$$f_X(1) = \frac{16^1 e^{-16}}{1!} = 0.00064 \; ; \; P[X > 1] = 1 - (0.00004 + 0.00064)$$

$$= \underline{0.99932}$$

70

2.7.3 Distribución uniforme

Se dice que una variable aleatoria continua, X, tiene *distribución uniforme* entre $X = a$ y $X = b$ $(b > a)$ si

$$\boxed{f_X(x) = \text{constante} = \frac{1}{b - a}} \qquad (2.34)$$

lo que significa que la probabilidad de obtener un valor entre x y $x + dx$ es la misma para cualquier x comprendida entre a y b. La gráfica de dicha distribución se presenta en la fig 12.

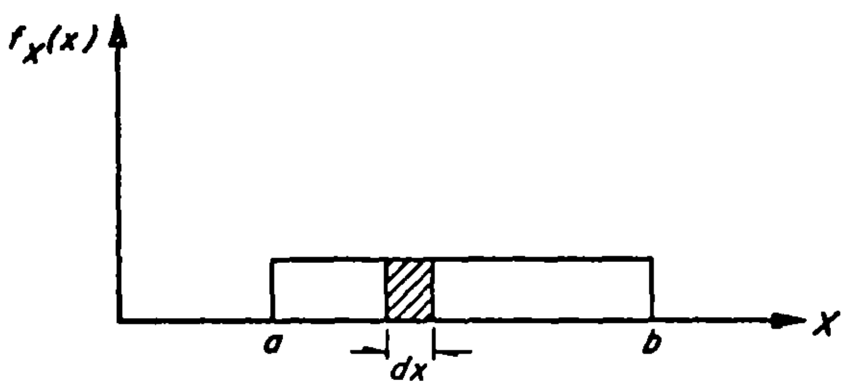


Fig 12. Distribución uniforme de una variable aleatoria continua

La esperanza y la variancia de la distribución uniforme se calculan de la siguiente manera

$$E[X] = \int_a^b x \frac{1}{b - a}\, dx = \left[\frac{x^2}{2(b - a)}\right]_a^b = \frac{b^2 - a^2}{2(b - a)} = (b + a)/2 \qquad (2.35)$$

$$Var[x] = \int_a^b (x - E[X])^2 \frac{1}{b-a}\, dx = \int_a^b \frac{x^2}{b-a}\, dx + \int_a^b \frac{(E[x])^2}{b-a}\, dx - \int_a^b \frac{2x\, E[x]}{b-a}\, dx$$

$$= \left[\frac{x^3}{3\,(b-a)}\right]_a^b + \left[\frac{(E\,[X])^2}{b-a}\, x\right]_a^b - \left[\frac{2\;E\,[X]}{b-a}\;\frac{x^2}{2}\right]_a^b =$$

$$= \frac{b^3 - a^3}{3\,(b-a)} + (E\,[X])^2 - E[X]\,(b+a) = \underline{\underline{\frac{(b-a)^2}{12}}} \qquad (2\ 36)$$

2.7.4 Distribución normal

Una de las distribuciones de variables aleatorias continuas más útil es la ***distribución normal*** o de *Gauss*, definida por la ecuación

$$\boxed{f_X(x) = \frac{1}{\sigma_X\sqrt{2\pi}}\; e^{-(x-\mu)^2/2\sigma_X^2}} \qquad (2.37)$$

donde $\underline{\mu = E\,[X]}$ y $\underline{\sigma_X = \sqrt{\mathrm{Var}\,[X]}}$

Si se hace la transformación

$$\boxed{Z = (X - \mu)/\sigma_X} \qquad (2\ 38)$$

entonces la ec 2.37 se reduce a la llamada *forma estandar*, cuya ecuación es

$$\boxed{f_Z(z) = \frac{1}{\sqrt{2\pi}}\; e^{-z^2/2}} \qquad (2.39)$$

En este caso, Z tiene distribución normal con media igual a cero y varianza igual a 1

Existen tablas para calcular las probabilidades de una variable asociada a una distribución normal estándar, semejantes a la tabla 5. En la fig 13 se muestra la forma de campana de esta distribución, observándose la simetría respecto a $Z = E[Z] = 0$.

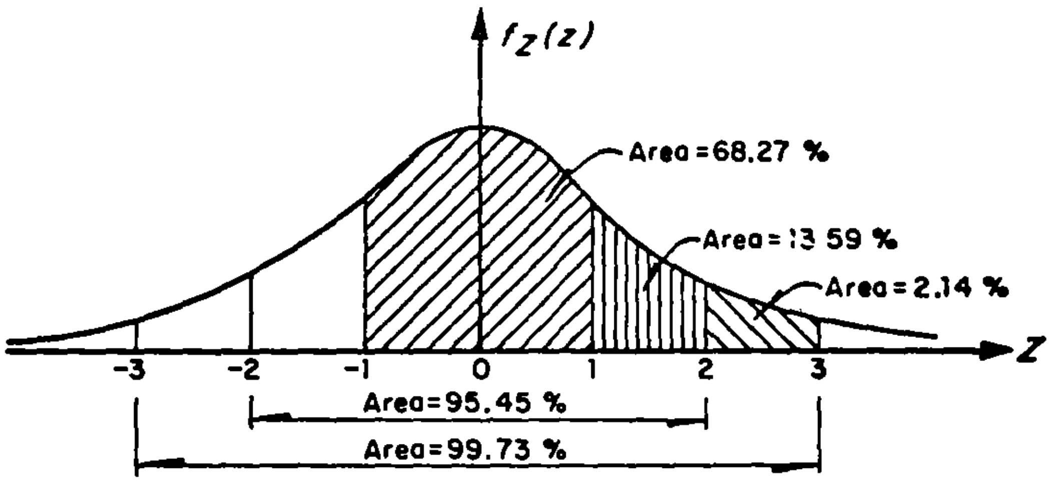


Fig 13. Distribución normal de una variable aleatoria continua

Ejemplo

Al probarse a compresión simple treinta cilindros de concreto, se obtuvieron resultados con un promedio aritmético de 240 kg/cm² y una desviación estándar de 30 kg/cm².

a) ¿Cuál es la probabilidad de que otro cilindro tomado al azar resista menos de 240 kg/cm²?

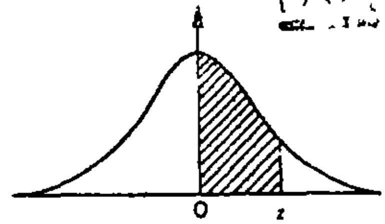
b) ¿Cuál es la probabilidad de que resista más de 330 kg/cm²?

c) ¿Cuál es la probabilidad de que su resistencia esté en el intervalo de 210 a 240 kg/cm²?

Supóngase que la distribución de probabilidades es normal.

a) Para emplear las tablas de distribución normal es necesario estandarizar la variable X mediante la transformación indicada en la ec 2.38.

TABLA 5. AREAS BAJO LA CURVA NORMAL ESTANDAR ENTRE 0 Y z



z	0	1	2	3	4	5	6	7	8	9
0 0	0000	0040	0080	0120	0160	0199	0239	0279	0319	0359
0 1	0398	0438	0478	0517	0557	0 596	0636	0675	0714	0754
0 2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0.3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0 4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0 5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0 6	2258	2291	2324	2357	2389	2422	2454	2486	2518	2549
0 7	2580	2612	2642	2673	2704	2734	2764	2794	2823	2852
0 8	2881	2910	2939	2967	2996	3023	3051	3078	3106	3133
0 9	3159	3189	3212	3238	3264	3289	3315	3340	3365	.3389
1 0	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621
1 1	3643	3665	3686	3708	3729	3749	3770	3790	3810	3830
1.2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1 3	4032	4049	4066	4082	4099	4115	4131	4147	4162	4177
1 4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1.5	4332	4345	4357	4370	4382	4394	4406	4418	.4429	4441
1 6	4452	4463	4474	4484	4495	4505	4515	.4525	.4535	4545
1 7	4554	4564	4573	4582	4591	4599	4608	4616	4625	46.33
1 8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706
1.9	4713	4719	4726	4732	4738	4744	4750	4756	4761	4767
2.0	4772	4778	4783	4788	4793	4798	4803	4808	4812	4817
2 1	4821	4826	4830	4834	4838	4842	4846	4850	4854	4857
2 2	4861	4864	4868	4871	4875	4878	4881	4884	.4887	4890
2 3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916
2 4	4918	4920	4922	4925	4927	4929	4931	4932	4934	4936
2 5	4938	4940	4941	4943	4945	4946	4948	4949	4951	4952
2 6	4953	4955	4956	4957	4959	4960	4961	4962	4963	4964
2 7	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974
2.8	4974	4975	4976	4977	4977	4978	4979	4979	4980	4981
2 9	4981	4982	4982	4983	4984	4984	.4985	.4985	4986	4986
3 0	.4987	4987	4987	4988	4988	4989	4989	4989	4990	4990
3.1	4990	4991	4991	4991	4992	4992	4992	4992	4993	4993
3 2	4993	4993	4994	4994	4994	4994	4994	4995	.4995	.4995
3 3	4995	.4995	4995	4996	4996	4996	4996	.4996	4996	4997
3 4	4997	4997	4997	4997	4997	4997	4997	.4997	.4997	4998
3 5	4998	4998	4998	4998	4998	4998	4998	4998	4998	4998
3 6	4998	4998	4999	4999	4999	4999	4999	4999	4999	4999
3 7	4999	4999	4999	4999	4999	4999	4999	4999	4999	4999
3 8	.4999	4999	4999	4999	4999	4999	4999	4999	4999	4999
3 9	5000	5000	5000	5000	5000	5000	5000	5000	5000	5000

Si el promedio aritmético se toma como estimación de μ y la desviación estándar como estimación de σ_X, para $X = 240$ se tiene

$$z = \frac{240 - 240}{30} = 0$$

Recurriendo a la tabla 5 se obtiene

$$P[X < 240] = P[Z < 0] = 0.5$$

o sea la probabilidad que corresponde al área sombreada de la fig 14.

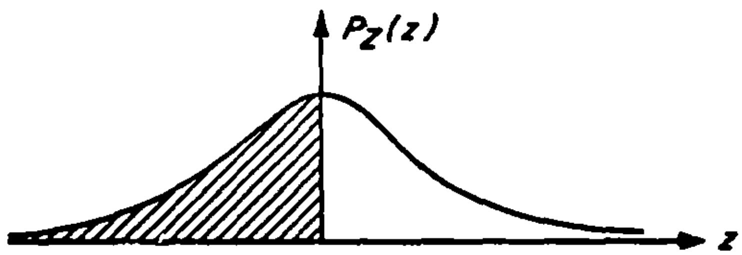


Fig 14. Distribución normal correspondiente al inciso a del ejemplo

b) El valor estandarizado de la variable es

$$z = \frac{330 - 240}{30} = 3$$

por lo que

$$P[X > 330] = P[Z > 3] = 1-0.9987 = 0.0013$$

que es el área sombreada de la fig 15.

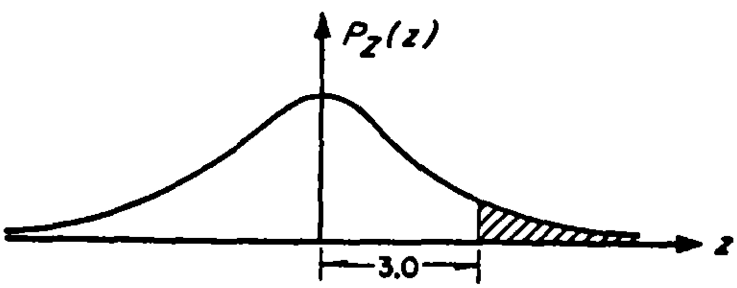


Fig 15. Distribución normal correspondiente al inciso b del ejemplo

c) Los valores estandarizados de la variable son (fig 16)

$$z_1 = \frac{210 - 240}{30} = -1$$

$$z_2 = \frac{240 - 240}{30} = 0$$

por lo que

$$P\{210 < X < 240\} = P[-1 < Z < 0] = 0.3413$$

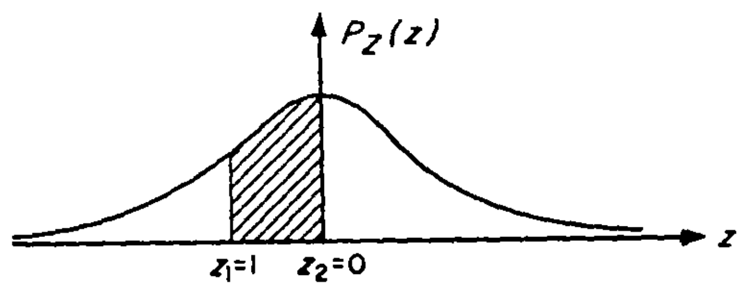


Fig 16 Distribucion normal correspondiente al inciso c del ejemplo

Sean $X_1, X_2, \ldots, X_k$, variables aleatorias con densidades de probabilidades arbitrarias cuya suma se denotará como W, es decir

$$W = X_1 + X_2 + \ldots + X_k$$

Es posible demostrar el teorema denominado *teorema del límite central*, cuyo enunciado indica que conforme aumenta el número de variables involucradas en la suma anterior (al aumentar K), la densidad de probabilidades de W tiende a ser la distribución normal Además se puede demostrar que si X_1, X_2, $\ldots$, X_k tienen distribución normal, entonces, rigurosamente, W también la tiene, independientemente del número de variables que aparezcan en la suma

A partir del teorema del límite central se demuestra que la distribución de Bernoulli se puede aproximar mediante la normal cuando el número de repeticiones del experimento es grande (30 o mas), con lo cual se logra un ahorro considerable de labor numérica en la solución de algunos problemas Para mejorar esta aproximacion, conviene efectuar una correccion por continuidad, la cual se justifica por usar una distribucion continua en vez de una discreta, sumando o restando, según sea el caso, 0 5 al valor de X que se use Por ejemplo, si se desea cuantificar la probabilidad de que de 2 000 ensayes se logren de 3 a 6 éxitos, los limites reales que se deben usar al aplicar la distribución continua son 2 5 y 6 5

Ejemplo

Si la probabilidad de que falle una varilla de acero al aplicarle cierta carga es de 0 001, determinar la probabilidad de que en 2 000 varillas probadas fallen más de dos

Usando la distribucion de Bernoulli se obtiene

$$P[X > 2] = 1 - P[X \leq 2] = 1 - \left\{ P[X = 0] + P[X = 1] + P[X = 2] \right\} =$$

$$= 1 - \left\{ \frac{2\ 000!}{2\ 000!\ 0!} (0\ 001)^0 (0\ 999)^{2000} + \frac{2\ 000!}{1\ 999!\ 1!} (0\ 001)^1 (0\ 999)^{1999} + \right.$$

$$\left. + \frac{2\ 000!}{1\ 998!\ 2!} (0\ 001)^2 (0\ 999)^{1998} \right\} = 0\ 3255$$

Los cálculos necesarios para obtener la solucion son bastante mas tediosos que los que deben efectuarse aprovechando que el numero de repeticiones del experimento es grande, a fin de utilizar la distribucion normal En estas circunstancias, la probabilidad de que $X < 2$ en el caso discreto, equivale a la de que $X < 2\ 5$ en el continuo, así

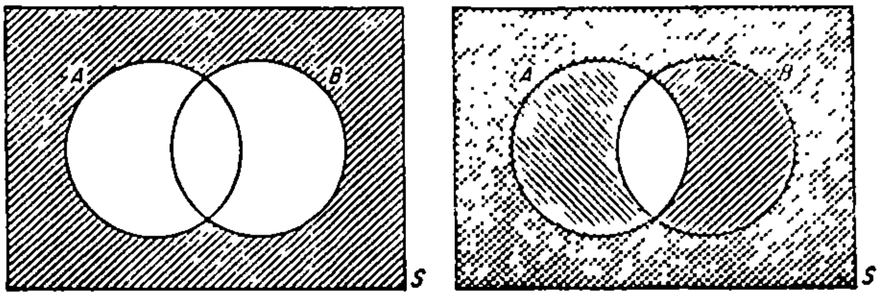
$$\mu = np = 2\ 000 \times 0\ 001 = 2$$

$$\sigma_X = \sqrt{npq} = \sqrt{2\ 000 \times 0\ 001 \times 0\ 999} = 1\ 41$$

$$P[X < 2\ 5] = P\left[Z < \frac{2\ 5 - 2}{1\ 41}\right] = P[Z < 0\ 355] = 0\ 6387$$

de donde

$$P[X > 2] = 1 - P[X < 2\ 5] = 1 - 0\ 6387 = 0\ 3613$$



(a) $(A \cup B)'$ (b) $A' \cap B'$

FIG 2 9 $(A \cup B)' = A' \cap B'$

and B', and it can be seen that it is identical with the shaded region of Figure 2 9(a)

EXERCISES

1. A tire manufacturer wishes to test four different tread designs on three different kinds of road surfaces and at five different speeds How many different test runs are required?

2. An experimenter has four different protective coatings to apply to both sides of a sheet of steel In how many ways can he coat both sides of the sheet of steel if
   (a) both sides must be coated with the same material,
   (b) the two sides can but need not be coated with the same material,
   (c) both sides cannot be coated with the same material

3. An experiment consists of rolling a die and then flipping a coin if and only if the die came up 1, 3, or 5 Draw a tree diagram and count the number of possible outcomes

4. Construct a tree diagram to determine the number of ways that a coin can be flipped four times in succession such that throughout the series of flips the number of heads is always greater than or equal to the number of tails

5. By enumerating all possible outcomes, verify the statement on page 8 that there are 21 distinct outcomes if we are interested only in knowing how many of the five television sets fit each of the three descriptions

6. In each of the following experiments decide whether it would be appropriate to use a sample space which is finite, countably infinite, or continuous
   (a) One of twelve vice-presidents is to be chosen president of a company

(b) An experiment is conducted to measure the coefficient of expansion of certain refractory bricks

(c) Measurements of radiation intensity are made with a Geiger counter

(d) A policeman measures the alcohol content of a driver's blood

(e) A coin is flipped until the first head appears

(f) A traffic survey is made to estimate the number of cars in the county with defective headlights

7 Among six applicants for a job Mr Andrews is married, does not play golf, and is not a home owner, Mr Bailey is single, does not play golf, and is a home owner, Mr Clark is married, plays golf, and is a home owner, Mr Dodds is single, plays golf, and is not a home owner, Mr Edwards is married, does not play golf, and is a home owner, and Mr Fox is single, plays golf, and is a home owner One of these applicants is to get the job and the event that the job is given to a golf player, for example, is denoted {Clark, Dodds, Fox} Indicate in a similar manner sets denoting the events that

(a) the position is given to a home owner,

(b) the position is given to a married golf player,

(c) the position is given to an unmarried person who is not a home owner,

(d) the position is given to a man who is either married or a single golf player

8 Referring to Exercise 7, indicate

(a) the complement of the set given in part (a),

(b) the union of the sets of parts (a) and (b),

(c) the intersection of the sets of parts (a) and (d),

(d) the intersection of the sets of parts (b) and (c).

9 A manufacturer buys parts from four different vendors numbered 1, 2, 3, and 4 Referring to orders placed on two successive days, (1, 4) denotes the event that on the first day the order was given to vendor 1 and on the second day it was given to vendor 4 Letting A represent the event that vendor 1 gets at least one of these two orders, B the event that the same vendor gets both orders, and C the event that vendors 1 and 3 do not get either order, list the elements of

(a) the entire sample space,

(b) A,

(c) B,

(d) C,

(e) A'

(f) $B \cup C$,

(g) $A \cap B$,

(h) $A \cap C$.

10 Use Venn diagrams to verify that

(a) $(A \cap B)' = A' \cup B'$,

(b) $A \cup (A \cap B) = A$,

(c) $(A \cap B) \cup (A \cap B') = A$,

(d) $A \cup B = (A \cap B) \cup (A \cap B') \cup (A' \cap B)$,

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

2.3 Probability

In this section we shall define probability using the concept of a *set function*, an *additive set function* to be exact. Since the reader is probably most familiar with functions for which the elements of the domain and the range are numbers, let us first consider a very simple example where the elements of the domain are sets while the elements of the range are real numbers In other words, we shall study a function, a correspondence, which assigns real numbers to the subsets of a given set (to the subsets of a sample space if we are concerned with the outcomes of a given experiment) The set function which we shall consider is the one which assigns to each subset A of a given finite set the *number of elements* in A, written $N(A)$ Suppose, then, that a company has 50 employees who are classified according to their marital status (married M and not married M') and according to whether they are college graduates or not (G or G'). The distribution of these 50 employees among the various categories is as shown in the Venn diagram of Figure 2 10 and, using this

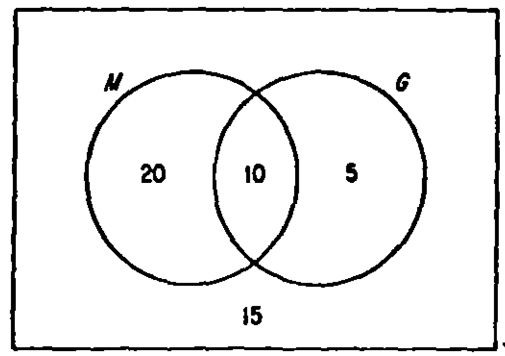


FIG 2 10 Distribution of elements in S

diagram, we can now determine the value $N(A)$ for any one of the 16 subsets (categories) into which the employees can be classified (In Exercise 1 on page 22 the reader will be asked to verify that there are, indeed, 16 subsets including the set S of all 50 employees and the empty set $\varnothing$) The numbers in Figure 2 10 are the number of married employees who are not college graduates, the number of married employees who are college graduates, the number of unmarried employees who are college graduates, It is instructive to note that if we had made the mistake of using the special law of addition in this example, we would have gotten the absurd result that the desired probability equals 1.400.

EXERCISES

1 List the 16 subsets into which the 50 employees of the example on page 14 can be classified.

2 Referring to Figure 2 10, find
(a) $N(M \cup G')$,
(b) $N(M' \cup G)$,
(c) $N[(M \cap G)']$.

3 Among 100 engineering students 15 are studying to be chemical engineers, 60 are undergraduates, and 5 are undergraduates studying to be chemical engineers. How many of these students
(a) are not studying to be chemical engineers,
(b) are not undergraduates,
(c) are studying to be chemical engineers, are undergraduates, or both,
(d) are undergraduates not studying to be chemical engineers,
(e) are nonundergraduates studying to be chemical engineers,
(f) are nonundergraduates not studying to be chemical engineers.

4 A dealer's stock of 75 cars includes cars with power steering (A), compacts (B), and cars with automatic transmissions (C). Using the information given in Figure 2 12, find
(a) $N(A)$,
(b) $N(B)$,
(c) $N(C)$,
(d) $N(A \cap B)$,
(e) $N(A \cap C)$,
(f) $N(A \cap B \cap C)$,
(g) $N(A \cup B)$,
(h) $N(B \cup C)$,
(i) $N(A' \cup B' \cup C)$,
(j) $N[B \cap (A \cup C)]$.

5 An experiment has exactly four distinct outcomes A, B, C, and D Check whether the following assignments of probabilities are permissible:
(a) $P(A) = 0\,36$, $P(B) = 0\,18$, $P(C) = 0\,21$, $P(D) = 0\,25$,
(b) $P(A) = 0\,29$, $P(B) = 0\,35$, $P(C) = 0\,18$, $P(D) = 0\,18$,
(c) $P(A) = 0\,42$, $P(B) = 0\,17$, $P(C) = -0\,08$, $P(D) = 0\,49$,
(d) $P(A) = 17/80$, $P(B) = 11/40$, $P(C) = 1/2$, $P(D) = 1/80$.

6 Referring to the sample space illustrated in Figure 2 5, denote by (i, j) the outcome that there are i defective components in the first subassembly and j defective components in the second subassembly ($i = 0, 1, 2, 3, 4$, $j = 0, 1, 2, 3$) Assuming that the probability of outcome (0, 0) is 1/2 and the probabilities of the remaining outcomes are inversely proportional to $i + j$, the total number of defective components in the whole assembly, find the probability of each outcome in the sample space.

7 Referring to the results of Exercise 6, find the following probabilities:
(a) the first subassembly has at most one defective component,
(b) the second subassembly has at least two defective components,
(c) the entire assembly has at most one defective component,
(d) the second subassembly has more defective components than the first subassembly.

8 In the example on page 18 it was assumed that each of the 50 employees had the same chance of being elected to a labor-management committee. If, instead, a college graduate has twice the probability of being elected as a non-college graduate, find (a) $P(M)$, (b) $P(G)$, (c) $P(M \cap G)$, and (d) $P(M \cup G)$.

9. Referring to the table on page 18, find
(a) $P(P_3)$,
(b) $P(P_2)$,
(c) $P(A_2)$,
(d) $P(P_1 \cap A_2)$,
(e) $P(P_1 \cup A_2)$,
(f) $P(P_3 \cup A_1)$.

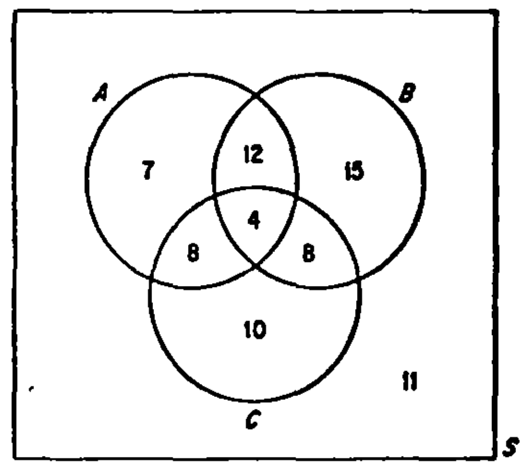


FIG 2 12 Exercise 4

10 What is the probability of rolling (a) 7 (b) 11, (c) 7 or 11, (d) 2 or 3 or 12, with a pair of balanced dice?

11 What is the probability of drawing (a) a black queen, (b) a red card, (c) 5, 6, or 7, (d) a black ace or a red king, from a well-shuffled standard deck of 52 playing cards?

12 A lottery sells tickets numbered from 1 to 10,000 What is the probability that the number drawn is divisible by 20?

13 Referring to Exercise 4, suppose that one of the cars on the dealer's lot was damaged in a windstorm Assuming equal probabilities, find the probability that

(a) the damaged car is a compact

(b) the damaged car has power steering

(c) the damaged car is a compact without automatic transmission

(d) the damaged car is not a compact but has an automatic transmission and power steering

14 If A and B are mutually exclusive events, $P(A) = 0.20$, and $P(B) = 0.55$, find

(a) $P(A')$,
(b) $P(A \cup B)$,
(c) $P(A \cap B)$,
(d) $P(A' \cap B')$.

15 Given $P(A) = 0.30$, $P(B) = 0.78$, and $P(A \cap B) = 0.16$, find

(a) $P(A \cup B)$,
(b) $P(A \cap B')$,
(c) $P(A' \cup B')$,
(d) $P(A' \cap B)$.

16 The probability that at least one of three events A, B, and C will occur is given by

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

Verify this formula with the probabilities shown in Figure 2 13.

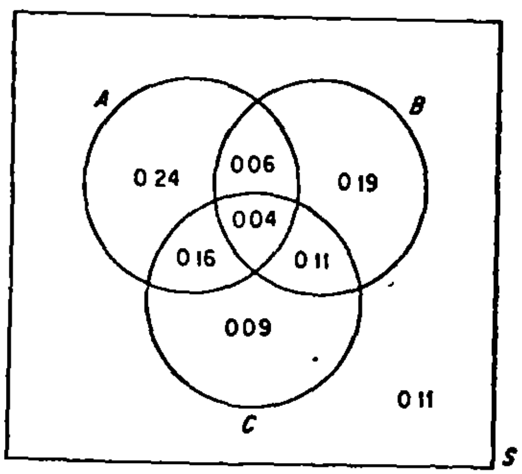


FIG 2 13 Exercise 16

17 In controlling the quality of mass-produced glass bricks, the probability that an inspector will get one which is cracked is 0 0025, the probability that he will get one with air bubbles is 0 0020, the probability that he will get one with discoloration is 0 0020, the probability that he will get one which has an air bubble and is cracked is 0 0006, the probability that he will get one which is cracked and discolored is 0 0005, the probability that he will get one which has air bubbles and is discolored is 0 0004, and the probability that he will get a glass brick with all three of these imperfections is 0 0001 What is the probability that he will get a glass brick with at least one of these imperfections?

18 Show that (a) $P(A) \geq P(A \cap B)$, and (b) $P(A) \leq P(A \cup B)$ [*Hint:* Refer to parts (c) and (d) of Exercise 10 on page 14.]

19 Given that A, B, and C are mutually exclusive events, explain why each of the following is not a permissible assignment of probabilities:

(a) $P(A) = 0.4$, $P(B) = 0.4$, $P(A \cup C) = 0.2$,

(b) $P(A) = 0.7$, $P(B) = 0.1$, $P(B \cap C) = 0.3$,

(c) $P(A) = 0.6$, $P(A \cap B') = 0.5$.

2.5 Conditional Probability

As we have defined probability, it is meaningful to ask for the probability of an event only if we refer to a given sample space S. To ask for the probability that an engineer has a salary of \$10,000 or more is meaningless unless we specify whether we are referring to the entire United States, a particular industry, a given plant, etc., and unless we specify how the selection is to be made. Thus, when we use the symbol $P(A)$ for the probability of A, we really mean the probability of A *given some sample space* S Since the choice of S is by no means always evident, and since there are problems in which we are interested in the probabilities of A with respect to more sample spaces than one, the notation $P(A \mid S)$ is used to make it clear that we are referring to the particular sample space S. We read $P(A \mid S)$ as "the *conditional probability* of A relative to S" and every probability is, thus, a conditional probability Of course, we use the simplified notation $P(A)$ whenever the choice of S is clearly understood.

To illustrate some of the ideas connected with conditional probabilities, let us again consider the set of 50 persons among whom some are married and some are college graduates as shown in Figure 2 10 Assuming equal probabilities, we saw on page 19 that the probability that a college graduate is elected to the labor-management committee is $P(G) = 3/10$, let us now

BIBLIOGRAFIA SELECCIONADA SOBRE TEORIA DE LA PROBABILIDAD, ESTADISTICA MATEMATICA, ESTADISTICA APLICADA, MUESTREO, Y PROCESOS ESTOCASTICOS, Y SUS APLICACIONES A LA INGENIERIA, DISPONIBLE EN LA BIBLIOTECA DE LAS DIVISIONES DE INVESTIGACION Y DE ESTUDIOS SUPERIORES.

Clasificación en biblioteca

1. Anderson, T., and D. Darling, Asymptotic theory of certain goodness of fit criteria based on stochastic processes. C-1097, A

2. Arrow, K., et al, Bayes and minimax solutions of sequential decision problems. C-1101, A

3. Bair, D., Experimentation an introduction to measurement theory and experiment design, Prentice Hall, 1962. QC39 B2

4. Barucha-Reid, A., ed., Probabilistic methods in applied mathematics, Academic Press, 1968. QA273 B42

5. Bellman, R., Programmed statistics; with chapters on probability, computer theory, and programmed instruction, Holt, Rinehart and Winston, 1970. HA29 B45

6. Benjamin, J., Probability, statistics, and decision for civil engineers, McGraw-Hill, 1970. QA273 B46

7. Bhat, U., Elements of applied process, Wiley, 1972. QA274 B42

8. Blackman, R., Linear data-smoothing and prediction in theory and practice, Addison-Wesley, 1965. QA275 B55

9. Blackwell, D., Another countable Markow process with only instantaneous states. C-1589, B.

10. Box, G., A bayesian approach to some outlier problems. C-1586, B

11. Box, G., Time series analysis, forecasting and control, Holden-Day, 1970. QA276 B6

12. Box, G., Evolutionary operation, a statistical method for process improvement, Wiley, 1969. TP155.7 B68

13. Breiman, L., Probability and stochastic processes, Houghton Mifflin, Co., 1969. QA273 B73

14. Breipohl, A., Probabilistic system analysis; an introduction to probabilistic models, decisions, and applications of random processes, Wiley, 1970. QA273 B746

15. Brown, R., Smoothing, forecasting and prediction of discrete time series, Prentice-Hall, 1963. TA168 B6

16. Brown, W., and C. Palermo, Random processed, communications, and radar, McGraw-Hill, 1969. TK5101 B75

17. Brownlee, K., Statistical theory and methodology in science and engineering, Wiley, 1960. QA276 B77

18. Bruning, J. and B. Kintz, Computational handbook of statistics, Scott, Foreman and Co., 1968. HA29 B77

19. Bush, R., A stochastic model with applications to learning. C-1590, B

20. Castro, G. de, Introducacio ao curso sobre instrumentos matematicos da estatística. F-2164, C

21. Clarke, B., Probability and random processes for engineers and scientists, Wiley, 1970. QA273 C43

22. Cochran, W., Experimental designs, Wiley, 1957. Q180 A166

23. Conover, W., Practical nonparametric statistics, Wiley, 1971. QA278.8 C65

24. Cooper, G. Probabilistic methods of signal and systems analysis, Holt, Rinehart and Winstons, 1971. TK454.2 C664

25. Cornell, C., A probability-based structural code, ACI, 1968.

26. Derman, C., Finite state markovian decision processes, Academic Press, 1970.

27. Deutsch, R., Nonlinear transformation of random processes. 517.7 D.

28. Dubes, R., The theory of applied probability, Prentice-Hall, 1968. TK5101 D8

29. Dynkin, E., Markow processes; theorems and problems, Plenum Press, 1969. QA273 D894

30. Edwards, A., Experimental design in psicological research, Holt-Rinehart and Winsont, 1968. BF59 E37

31. Ehrenfeld, S. and S. Littauer, Introduction to statistical method, McGraw-Hill, 1964. 519.9 E

32. Esteva, L., Consideraciones práctias en la estimación bayesiana de riesgo sísmico, México, UNA. Instituto de Ingeniería, 1970. F-8843, E

33. Feller, W., Introducción a la teoría de probabilidades y sus aplicaciones, Limusa-Wilesy, 1973. QA273 F3714

34. Forcadas, J., Estadística aplicada a la ingeniería, 1969. QA276 F66

35. Freeman, H., Introducción a la inferencia estadística, Trillas, 1970. QA276 F6845

36. Freund, J., Mathematical statistics, Prentice Hall, 1971. QA276 F692

37. Freudenthal, H., Probability and statistics, Elsevier, 1965. QA273 F75

38. García, A., Elementos de método estadístico, México, UNA, 1966. HA29 G146

39. Gotkin, L., Estadística descriptiva; texto programado, Limusa-Wiley, 1967. HA29 G695

40. Greenwood, A. and H. Hartley, Guide to tables in mathematical statistics, Princenton University Press, 1962. 519.9 G

41. Grenander, U., Statistical spectral analysis of time series arising from stationary stochastic processes. C-1588, G

42. Hanman, H., Modern factor analysis, University of Chicago Press, 1967. QA276 H38

43. Hanna, E., Time series analysis, Methuen, 1960. QA76 H32

44. Hays, W., Statistics· probability inference and decisions, Holt-Rinehart and Winston, 1970. QA276 H39

45. Hoel, P., Elementary statistics, Wiley, 1971. HQ29 H66

46. Hoel, P., et al, Introduction to stochastic processes, Houghton, Mifflin, 1972. QA274 H63

47. Howard, R., Dynamic probabilistic systems, Wiley, 1971 T57.95 H66

48. Iosifoscu, M., Random process and learning, Springer Verlag, 1969. QA273 I64

49. Jazwinski, A., Stochastic process and filtering theory, Academic Press, 1970. QA276 J38

50. Kefer, J. and J. Wolfowitz, Sotchastic estimation of the maximum of a regression function. C-1481, K

51. Kemeny, J., Finite Markow chains, Van Nostrand, 1960. QA273 K33

52. Kish, L., Muestreo de encuestas, Trillas, 1972. HN29 K5

53. Kozin, F. and J. Bogdanoff, An introduction to random function for engineer, 1963. F-1930, K

54. Kyburg, H., Probability theory, Prentice-Hall, 1969. QA273 K92

55. Lange, F., Correlation techniques· Foundation and applications of correlation analysis in modern communications, measurement and control, Iliffe, 1967 TK7870 L34

56. Larson, H., Introduction to probability theory and statistical inference, Wiley, 1969. QA273 L37

57. La Valle, I., An introduction to probability decision, and inference, 1970. QA276 L36

58. Lecture notes in mathematics, probability and information theory, Springer Verlag, 1964. QA1 L42

59. Lee, T., et al, Estimating the parameters of the Markow probability model, North Holland, 1970. HB74 M3L43

60. Lindgren, B., Statistical theory, Macmillan, 1968. QA276 L55

61. Lin, Y., Random processes, F-8709, L.

62. Prelewicz, D., Range of validity of the method of averaging. F-9116, P

63. Prohorov, Y., Probability theory; basic concept, limit theorems, random processes, Springer Verlag, 1969. QA273 P75

4. Raiffa, H. and R. Schlaifer, Applied statistical decision theory. Harvard University Press. 1971. QA279.4 R34

65. Raj, D., Sampling theory, McGraw-Hill, 1968. QA276.6 R33

66. Raj, D., Design of sample survey, McGraw-Hill, 1972. QA372 R33

67. Rascón, O., Introducción a la teoría de probabilidades. México, UNA, 1971. QA273 R37

68. Rascón, O., Introducción a la estadística descriptiva, 1970. HA31 R37

69. Ravindra, M., Probabilistic evaluation of safety factors, University of Waterloo, 1969. F-8680, R

70. Robinson, E., Multichanel time series analysis with digital computer programs, Holden-Day, 1967. QA276 R633

71. Rosenblueth, E., Current research on probabilistic methods at the National University. C-1517, R

72. Ross, S., Applied probability models with optimization applications, Holden-Day, 1970. QA273 R67

73. Sage, A., Estimation theory with applications to communications and control, McGraw-Hill, 1971. QA276.8 S33

74. Savage, L., The theory of static decision, C-1098, S

75. Sawaragi, Y., Statistical decision theory in adaptive control system, Academic Press, 1967. QA402.3 S37

76. Schäl, M., Markow renewal processes with auxiliary phats. C-1617, S

77. Schlaifer, R., Probability and statistics for business decisions and introduction to mangerial economics under uncertainty, McGraw-Hill, 1959. HD38 S35

78. Schwarts, J., Statistical methods in traffic engineering, The Ohio State University, 1967. HE336 S7S3

79. Searle, S., Linear models, Wiley, 1971, QA279 S

80. Sengupta, J., Stochastic programming methods and applications, North Holland, 1973. T57.79 S44

81. Sengupta, S. and S. Jain, A representative theory for measurable random variable. F-5967, S

82. Sheppard, R., Multidimensional scaling, Seminar Press, 1972. CF39 M84

83. Spiegel, M., Theory and problems of statistics, Schaum's Publishing, Co., 1960. HA29 S65

84. Sterling, T., Introduction to statistical data processing. Prnetice-Hall, 1968. QA276.4 S82

85. Stratonovich, R., Conditional Markow processes and their application to the theory of optimal control, Elsevier, 1968. QA273 S76

86. Symposium on time series analysis proceedings, held at Brown University, June 11-14, 1962. Ed. by Murray Rosenblatt. New York, John Wiley and Sons, 1963. 519.9 S

87. Tannur, M., Statistics: a guide to the unknew, Holden Day, Inc., 1972. QA276 S82

88. Theil, H., Statistical decomposition analysis with applications, North Holland, 1972. H61 T43

89. Thomas J., An introduction to statistical communication theory, Wiley, 1969. TK5102.5 T45

90. Tintner, G., Stochastic economics, stochastic processes, control and programming. Academic Press, 1972. QA274 T54

91. Turkstra, C., Applications of bayesian decision theory, University of Waterloo, 1969. F-8619, T

92. Turkstra, C., Elements of probability theory. F-9105, T

93. Turkstra, C., Theoretical distribution functions. F-9104, T

94. Valdes, R., Nociones de cálculo de probabilidades y estadística, Limusa-Wiley, 1970. QA273 V34

95. Van der Gerr, J., Introduction to multivariate analysis for the social sciences, Freeman, 1971. QA278 V34

96. Vere-Jones, D., Stochastic models for earthquake occurence; discussion. F-9247, V

97. Waerden, B., Mathematical statistics, Springer-Verlag, 1957, QA276 W3

98. Wold, H., ed., Bibliography on time series stochastic processes· an international team project. Edinburgh, Oliver and Boyd, 1965.

99. Yaglom, A., An introduction to the theory of stationar random function, Prentice-Hall, 1962. 517.5 Y

100. R. L. Winkler, "Introduction to Bayesian inference and decision", Holt, Rinehart and Winston, Inc., 1972.

101. A. D. Carter, "Mechanical Reliability", John Wiley, 1972.

BIBLIOTECA DE LAS DIVISIONES
DE INVESTIGACION Y DE ESTUDIOS SUPERIORES DE LA FACULTAD DE INGENIERIA
U N A M

Apartado Postal 70156
Mexico 20, D F México

1º de Mayo 1970

LISTA DE PUBLICACIONES SOBRE PROBABILIDAD Y ESTADISTICA RECIBIDA
P R [illegible] RO [illegible]

1.- [illegible], P.L. "The description of a random field by means of conditional probabilities and conditions of its regularity", Theory of Probability v. 13, n.2, 1968.

2.- PIROGOV, [illegible]. "Probabilities of compound events and linear programming" Theory of Probability, v.13, n.2, 1968.

3.- VENTTSEL, A.D. "On the absolute continuity of transition probabilities of multidimensional generalized Brownian motion, Theory of Probability, v.13, n.1, 1968.

4.- YASHIN, A.I. "Filtering a Discontinuous markov process with unknown — Probability characteristics with the presence of additive noise" Automation and Remote control, n. 12, 1968

5.- [illegible], G. "Determination of Critical flashover voltage and standard — deriviation from flashover probability data", IEEE Transactions on power apparatus and systems, v.88, n.3, Mar 1969.

6. [illegible], [illegible].V."An estimation to the rate of convergence for absorption — probability in case of zero expectation", Theory of probability v.33, n.1, 1968.

7.- TUTUBALIN, V.N."Approximation of probability measures in variation and products of random matrices", Theory of Probability v.3[illegible], n.1, 1968

8.- OCIL [illegible], J.M. "Probable descentraje del nucleo terrestre Anales — del Instituto de Geofísica, v.13, 196 .

9.- KABA[illegible]EVSKII, A.N. "Statistical design of a Random function generator of the CSP, 2 type", Automation and remote n.12,dic. 1968.

10.-DYN[illegible]IN, E. Sufficient statistics for the optimal stopping problem", Theory of Probability v.13, n.2, 1968.

11.-[illegible] a la orientacion "Informacion estadistica necesaria para orientar oportunamente Ferrocarriles [illegible]on del ferrocarril", n.11, En. Feb. Mar. 1969.

BIBLIOTECA DE LAS DIVISIONES
DE INVESTIGACION Y DE ESTUDIOS SUPE-
RIORES DE LA FACULTAD DE INGENIERIA
U N A M

Apartado Postal 70156
Mexico 20, D F México

[illegible] SOBRE [illegible] PEDIDA
[illegible]

1.- The Weber Problem with probabilistic Demands,
F-7741

2.- Probabilistic information processing [illegible] and evaluation
F-7701

3.- [illegible], S., [illegible] "A probabilistic interpretation of [illegible]", Siam Journal on Applied Mathematics v.16, pts 1, Marzo 1969.

4.- [illegible] "Discusion probabilistic structural analysis and design", Journal of the Urban Planning and development Division, v.95, n.1, Abril 1969

5.- [illegible]ILMAR, R. J."Discussion probabilistic structural analysis and design ", Journal of the Structural Division, v.95,n.7, Jul 1969.

6.- [illegible] "A bayesian approach to the problem of [illegible] binomial probabilities" Siam Journal on applied Mathematics v.16, n.1,2, 3, 4, 5, 6, 1968.

7.- Joint probabilities in the rainfall runoff relation
F- 7726.

8.- KOFO [illegible] "Chernoff bounds for the [illegible] probabilities of the ..." [illegible] Transactions on information Theory, v.15, n.2, Marzo 1969.

9.- SKIDINSKY, M. "Minimax estimation of a random probability , ---- Siam Journal on Applied Mathematics, v. 16, n. 1,2,3,4,5,6 1968

10.- Y[illegible]G, J. 'On the bound of first excursion probability", Journal of the Engineering Mechanics Division, v.93, n.2, Abril 1969

11.-A[illegible] "Safety factors and probability in structural design", — Journal of the Structural Division, v.95, n.7, Julio 1969.

12.-WII[illegible]CT, R. "Analysis of a simulated radio simulator for measuring [illegible]", I[illegible] Transactions [illegible] instrumentation and measurement machine', I[illegible] Transactions on instrumentation and measurement, V.18, n.1, Marzo 1969.

13, [illegible]RL, I. C[illegible]TTICK, I. "Rank permutation group [illegible] based on Kendall's correlation I[illegible] transactions on information theory v.15, n.2, Marzo 1969

BIBLIOTECA DE LAS DIVISIONES DE INVESTIGACION Y DE ESTUDIOS SUPERIORES DE LA FACULTAD DE INGENIERIA
U N A M

Apartado Postal 70 156
Mexico 20, D F México

14.- ROST, I... "Statistical optimization ... of the tensor of viscosity coefficients", PMM Applied Mathematics and Mechanics, v.32 n.5, Mayo 1968.

15.- TO.ROJA, E. "Statistical mechanics of gaseous suspension dynamics spectral equations", PMM Applied Mathematics and Mechanics v.3 , n.5, Mayo 1968.

16.- ... "Some statistical methods in the analysis hydrologic series", Soviet Hydrology, n.1, Jul 1969.

17.- ..., P.F.; COR..., R.D. "Discussion statistical model for live floor loads" Journal of Structural Division , v.95, n.7, Jul 1969.

18.- ..., "The eigenvalue problem for structural systems with statistical properties, AIAA Journal v.7, n.4, Abr. 1969

19.- DITZIAN, Z. "Statistical properties of the ... of a graph Siam Journal on Applied Mathematics, v.16, n.1,2,3,4,5,6 1968.

20.- ..., .. "A statistical theory of brittle fractures ... elastic continua" Hydrotechnical construction, n.12, Dic. ...

21.- WILLIAMS H, R. "The vital statistics of an audio amplifier", J. of the Soil Mechanics and Foundations Division, v.95, n.... Julio 1969.

22.- LIU, S. "Earthquake response statistics of ... shear systems", Journal of the Engineering Mechanics Division, v.95, n.2, Abril 1969.

23.- M...COUVRIER, J. "Application de analyse statistique a la detection des accidents ..., Bull de liaison et des laboratoires routiers, n.38, Mar-Jun. 1969.

24.- Interpretation statistique des resultats d'essais in flexion simple et en flexion composee
F/ 7757

BIBLIOTECA DE LAS DIVISIONES
DE INVESTIGACION Y DE ESTUDIOS SUPE
RIORES DE LA FACULTAD DE INGENIERIA
U N A M.

Apartado Postal 70156
Mexico 20, D. F México

LISTA DE PUBLICACIONES EXISTENTES EN LA BIBLIOTECA SOBRE
PROBABILIDAD Y ESTADISTICA SOLICITADA POR EL INGENIERO

1.- FLAATE, .. "A statistical analysis of some methods for shear strength determinations in soil mechanics", Norwegian Geotechnical Institute, n.62, 1965.

2.- STANCU ..., G. IONESCU, ". "Statistical analysis of the distance between cracks and the IR openings in reinforced concrete ... elements", Revue Roumaine des Sciences Techniques, Mecanique Appliquee, v.1, n.6, 1965.

3.- PAVEL..CU, A, "Aspects of wear evolution related to the statistical ... character radioactive isotope measuring techniques", Revue Roumaine des Sciences techniques, Mecanique Appliquee, v.1, n.6, 196

4.- TUDY, S. F. "Statistical determination of certain mathematical constants, and functions using computers", Journal of the Association for computing machinery, v.13, n.4, Oct. 1966

5.- FR..DMAN, "; 'The little variable factor, A statistical discussion of the reading of seismograms", Bulletin of the Seismological Society of America, v.56, n.2, Abril 1966.

6.- FUJITA, "Thermodynamic evolution equation for a quantum statistical Gas.Journal of Mathematical Physics, v.6, n.12, Dic 1965.

7.- BARNES, J. ALLAN, C. "A statistical model of flicker noise, Proc. of the IEEE, v.54, n.2, Feb 1966.

8.- MUO..ERJEE, A, "Applications of statistical techniques in power study of stripping shovels", Mining and Minerals engineering, v.1, n.14, Oct. 1965.

9.- "Introduction to statistical theory of land locomotion II ground". Mining and Mineral v.1, n.14, Oct. 1965.

10.-"Introduction to statistical theory of land locomotion", Jl. of Terramechanics, v.2, n.2, 1965.

11.- MOGAM I.T. A statistical theory of mechanics of granular materials Journal of the Faculty of Engineering University of Tokyo, Serie B, v.28, n.2, Sep 1965.

12.- BELL, "Statistics in electronics", Wireless World, v.72, n.1, Ene. - 1966.

13.- TARZIAN, R: "On the statistics of a product, Proc. of the IEEE, v.54 n.2, Feb 1966.

BIBLIOTECA DE LAS DIVISIONES DE INVESTIGACION Y DE ESTUDIOS SUPERIORES DE LA FACULTAD DE INGENIERIA
U N A M

Apartado Postal 70156
Mexico 20 D F México

14.- ALLAN, "Statistics of atomic frequency standards, Proceedings of the IEEE, v.54, n.2, Sep. 1966.

15.- BOURSMA, S. "Statistics of lassor . diation", Proc. of the IEEE, v.54, n.8, A. 1966.

16.- FAGIEFIELD, C. "Statistics of laser radiation", Proc. of the IEEE, v.54, n.1, Enero1966.

17.- BARNES, J. "Atomic timekeeping and the statistics of precision signal generation", Proc. of the IEEE, v.54 n.2, Feb. 1966

18.- BRUNEAU, C.M. "Description statistique de [illegible] electroniques pour servir de modele a [illegible]", Revue Francaise de Traitement de l'information n.3, Jul-Ag-Sep. 1966.

19.- HULL, [illegible], J. "Tests of probabilistic models for propagation of roundoff errors, Communications of the ACM, v.9, n.2, Feb. 1966.

20.- PETIGNY, B. "Le calcul des probabilities et la circulation des vehicles extension de la distribution de Schul", Annales des ponts et Chaussees, v. 136, n.11, Mar-Apr. 1966.

21.- "Sundu tions channel model to predict error probabilities of . ." Research and Technology Briefs, v.4, Jul 1966.

22.- VICKOROV, "Asymptotic normality on a problem of balls when the probabilities of falling into different boxes are different
C Contractors of Engineering Magazine, v.8, Ag. 1966.

23.- [illegible] RK.- "On the false alarm probabilities of filtered noice", Proc of the IEEE, v.54, n.2, Feb 1966,

24.- [illegible], I; "Probability bounds on the size of storage areas, Operation Research, v. 14, n.5, Sep-Oct. 1966

25.- KOZNIEWSKI, J. "Stochastic processes in linear configurations in which the Transition probability densities do not depend on states of particular components, Bulletin de la Academy Polonaise des Sciences v. 13, n.2, Junio 1966.

26.- FELSENBY, [illegible]. THOMSON. "Probability distribution of cylinder systems response to impulse excitation, Journal of Applied Mechanics, v33, n.2, Junio 1966.

27.- JOHNSON, N. "Cumulative detection probability for swerling III and IV, targets", Proc. of the IEEE, v.54, n.11, Nov. 1966.

BIBLIOTECA DE LAS DIVISIONES DE INVESTIGACION Y DE ESTUDIOS SUPERIORES DE LA FACULTAD DE INGENIERIA
U N A M

Apartado Postal 70156
Mexico 20, D F México

28.- VIKEN, LANCFORD, "Failure probability formulas for systems with spares, Operation Research, v.14, n.4, Jul-Ag. 1966.

29.- SAZONOV, V., "On the limiting behavoir of the maximal probability in the infinite dimensional case", Theory of Probability and its Applications, v.10, n.4, 1965.

30.- "Dynamic range centering for minimum probability of excluding a Rayleigh distributed signal", Proc. of the IEEE, v.54, n.1 Ene. 1966.

31.- LEVINE, A. The probability of an excesive nonfunctioning interval, Operation Research, v.14, n.5, Sep-Oct 1966.

32.- "Asymptotic, behavoir of the Optimal probability of error for the transmission of information in a chanel without memory Contractors and Engineers Magazine, v.8, Ag. 1966.

33.- ZEGER, A. "Dynamical range positioning for minimal probability of excluding a randomly distributed signal, Proc. of the IEEE, v.54,n.6, Jun 1966.

34.- DENHAM, W, MOHAM, N. "Choosing trajectory parameters to maximize mission probability of succes, AIAA, Journal, v.4, n.7, Jul 1966.

35,-AVAZYAN, S. A; "Distinguishing neighboring hypothese about frequency distribu ting in the scheme of the generalized probability ratio test, th ory of Probability and its applications v.10, n.4, 1965.

36.- "Information on the conference on limit theorems on probability" Theory of Probability and its Applications", v.10, n.4, 1965.

LISTA DE PUBLICACIONES PERIODICAS EXISTENTES EN LA BIBLIOTECA SOBRE PROBABILIDAD Y ESTADISTICA QUE FUE SOLICITADA POR EL ING.

1.- Theory of probability and its applications v.9, n.1, 1964

"Reservoir yield probability applied", Proc. of the Institution of Civil Engineers, v. 30, Ene. 1965.

3.- Technique for probability density function measurement"-- Electronic Engineering v.37, n. 447 Mayo 1965.

4.- "Infinite Integrals probability density function" Proc. of the IEEE, v.53, n11, nov. 1965.

5.- "Probability detecting region by linear research", Theory of probability and its applications, v.9, n.2, 1964.

6.- Estimating probability distribution drought flows, water and sewage works, v. 112, n.5, Mayo 1965.

7.- "Selection probability distribution random variable" Proc. of the IEEE v.53, n.9, Sep 1965.

8.- "Signal method for determining probability distribution random" The review of Scientific Instruments, v.36, n.12, Dic. 1965

9.- "Signal method determining probability distribution random", The review of Scientific Instruments, v.36, n.12, Dic. 1965

10. "Error probability due atmospheric noise flat fading in HF" Proc. of the IEEE, v.53, n.3, Mar. 1965.

11.- Probability isotropic link connecting using comsat elliptic" Proc. of the IEEE, v.53, n.5, Mayo 1965

12.- "Probability method gould probability moran steady state" Jnl. of The Institution of Water, v.19, n.4, Junio 1965

13.- "Expected error probability misclassif. ", Proc. of the IEEE v.53, n.6, Jun. 1965

14.-"Expected error in probability misclassif." Proc. of the IEEE

15.- "Probability of markov point falling into a plane region" Theory of probability and its applications, v.9, n.2, 1964.

16.- "Probability reservoir yield failure using moran steady statistics, Journal of the Intitution of Water v.19, n.4, Junio 1965

17.- "State probability method gould probability routing method" Journal of the Institution of Water, v.19, n.4, Junio de 1965.

18.- "Betting systems wich minimize probability" Journal of Fluid ~~Mechanics~~ v.23, pte2, Oct. 1968

19.- 2Transmision probability some arii laser lines" Journal of applied physics v.36, n.7, Julio 1965.

20.- "Concept entropy in probability theory" Theory of Probability and its applications, v. 9, n.2, 1964.

21.- "Free molecule transmission probability" Journal of applied physics , V.36, n.10, 1965.

22.- "Gamma functions evaluating erlang process", Mathematics of computation, v.17, n.81, Ene. 1963,

23.- 2Schmitt circuit for determining probability" Proc. of the IEEE v.53,n.10, 1965.

BIBLIOTECA DE LAS DIVISIONES
DE INVESTIGACION Y DE ESTUDIOS SUPE-
RIORES DE LA FACULTAD DE INGENIERIA
U N A M

Apartado Postal 70.[illegible]56
Mexico 20, D. F. México

17-IV/1970

LISTA DE PUBLICACIONES QUE ESTAN EN LA BIBLIOTECA
SOBRE PROBABILIDAD Y ESTADISTICA QUE FUE SOLICITA/
DA POR EL INGENIERO [illegible]

1.-Curso Estadística, Histograma de Distribuciones Empíricas
F-4068

2.-Mecánica Estadística Clasica
F-2765

3.-"Statistical analysis reasons for varying magnetization gra--
nite", International Geology Review, v.7, n.11, Nov. 1965

4.-"Quantom statistic of multicompound systems", Journal of the
Mathematical Physics, v.6, n.3, Mar.20, 1965.

5.-"Statistical analysis of nonlinear time-variant control", --
Technical Report of the Engineering Research Institute
v.115, 1965.

6.-"Statistical analysis spacecraft replenishm", Proc. of the -
IEEE, v.53, n.12, Dec.1965.

7.-"Application statistical analysis to quality control for Road"
Bulletin, Congres de la Route, n.176, 1965.
vol 51-60 [illegible]

8.-Statistical Vocabulary
F-2827.

9.-Reasonable application statistically", Journal of the Assn. -
for computing machinery, v.12, n.4, Nov. 1965.

10.- Rail failures statistics all failures transverse fissures",
American Railway engineering association, v.66, n.591,
1965.

11.-"Use social statistics estimating autownership", Highway Res.
News, n.16, 1964,

12.-"Mathematic and statistical technique, The Computer Bulletin
v.8, n.4, 1965.

13.- New merenne primes statistical theory", Mathematical of ---
Computation, v.18, n.85, Enero 1964.

14.-"Statistics rastler multimode radiation ensemble", Proc. of
the IEEE, v.53, n.11, Nov. 1965.

BIBLIOTECA DE LAS DIVISIONES
DE INVESTIGACION Y DE ESTUDIOS SUPERIORES DE LA FACULTAD DE INGENIERIA
U N A M

Apartado Postal 70156
Mexico 20, D F. México

15.-"New statistical dynamic methods numericàl wather" OAR Research Review, v.4, n.6, Ag. 1965.

16.-"Statistical approach analysis fatigue failure of prestressed.. Journal of the ACI, n.11, Dic. 1965.

17.-"Systemes Reliability enginering statistical aspect", American Scientist, v.53n.3, Sep 1965.

18.-"Statistical complexity algorithms for boolean functions Journal of the ACM v.12, n.3, Julio de 1965.

19.-"Statistical ensembles of complex quaternion and real matrix, Journal of the Mathematical Physics, v.6, n.3, Marzo 1965.

20.-"Structure cluster expansion statistical mechanics" --- Journal of the Mathematical Physics, v.6, n.8, Ag. 1965.

21.-"Inequality application statistical mechanics", Journal of the Mathematical Physics, v.6, n.11, Nov. 1965

22.-"Microcanon ensemble in quantom statistical mechanics". Journal of the Mathematical Physics, v.6, n.10, Oct. 1965.

23.-"Statistical mechanics, assembli escoupled oscillator", Journal of the Mathematical Physics, v.6, n.4, -- Abril 1965.

24.-"Statistical specifications in Highway", California --- Highways and Public works, v.44, n.9-10 Sep-Oct. 1965.

25.-"Statistical tables and approach", Communications of the A.C.M, v.8, n.6, Junio 1965.

26.

LISTA DE PUBLICACIONES EXISTENTES EN LA BIBLIOTECA SOBRE PROBABILIDAD Y ESTADISTICA, QUE FUE SOLICITADA POR EL ING.

No. de Clasif.	Autor.	Título.
519.9-W48	Wetherill, G. Barrie.	Sequential methods in statistics.
519.9-W43	Weatherburn, C. E.	A first course in mathematical statistics.
519.9-W38	Waugh, Albert E.	Elements of statistical methods.
519.9-W	Waerden, Bartel L.	Mathematische Statistik.
519.9-W	Wadsworth, George P.	Introduction to probability and random variables.
519.9-T	Thrall, R. M. ed.	Decision processes.
519.9-T	Tippet, L. H. C.	Technological applications of statistics.
519.9-T	Tippet, L. H. C.	The methods of statistics.
519.9-T	-.-	Transactions of the second conference on information theory statistical decion functions, random processes.
519.9-T	-.-	Transactions of the symposium on computing mechanics, statistics and partial differential equations.
519.9-S	-.-	Symposium on time series analysis
519.9-R	Rufener, Ernest.	La mise équation des résultats d'experiences.
519.9-O	Obras Públicas, Sría. de.	Servicio estadístico anuario.
519.9-N	Neyman, Jersy ed.	Proceedings of the third Berkeley symposium on mathematical statistic and probability.
519.9-N	Natrella, Mary G.	Experimental statistics.
519.9-M66	Mood, Alexander M.	Introduction to the theory of statistics.
519.93-M	Meyer, Herbert A. ed.	Symposium on Monte Carlo methods
519.9-K	Kendall, Maurice G.	The advanced theory of statistics.
519.9-J	Jeffreys, Harold.	Scientific inference.

No. de Clasif.	Autor.	Título.
519.9-H	Hogg, Robert V.	Introduction to mathematical statistics.
519.9-H	Hald, A.	Statistical theory with engineering applications.
519.9-G	Grenander, Ulf, ed.	Probability and statistics.
519.9-G	Greenwood, Arthur J.	Guide to tables in mathematical statistics.
519.9-F	Friedlander, S. K. ed.	Turbulance classic papers on statistical theory.
519.9-F74	Freeman, Harold.	Introduction to statistical inference.
519.9-F	Fisher, R. A.	statistical methods and scientific inference.
519.9-F	Fisher, Ronald A.	Statistical method for research workers.
519.9-F47	Ferguson, Thomas S.	Mathematical statistics; a decision theoretic approach.
519.9-F	Fisher, R. A.	Contributions to mathematical statistics.
519.9-F	Feller, William.	An introduction to probability theory and its applications.
519.9-F	Federer, Walter T.	Experimental design theory and application.
519.9-F	Fano, Robert.	Transmission of information a statistical theory of communications
519.9-C	Cramer, Harald.	Mathematical methods of statistics.
519.9-C	Cochran, William G.	Experimental designs.
519.9-B	Brownlee, K. A.	Statistical theory and methodology in science and engineering.
519.9-B	Bogdanoff, John L. ed.	Proceedings of the first symposium on engineering applications of rando theory and probability.
519.9-B	Bendat, Julius S.	Principles and applications of random noise theory.
517.7-B47	-.-	Berkeley symposium on mathematical statistics and probability, 5.

No. de Clasif.	Autor.	Título.
519.9-B	Bartlett, M. S.	An introduction to stochastic processes with special reference to methods and applications.
519.9-A	Anderson, R. L.	Statistical theory in research.
519.9-A	Arley, Niels.	Introduction to the theory of probability and statistics.
519.9-A	Arrow, Kenneth J.	Studies in the mathematical theory of inventory and production.
519.9-A54	Alexander, Howard W.	Elements of mathematical statistics.
624.17-L55	Lin, Y. K.	Probability theory of structural dynamics.
519.9-S	Schrödinger, Erwin.	Statistical thermodynamics.
519.9-C	Cramer, Harald.	Métodos matemáticos de estadística.
519-S46	Skorokhod, Anatolii V.	Studies in the theory of random processes.
519-G54	Gnedenko, B. V.	The theory of probability.
510.83-S	Savinskii, I.	Probability tables for locating elliptical.
519.93-U	Uspensky, J. V.	Introduction to mathematical probability.
519.9-F	Fraser, D. A. S.	Nonparametric methods in statistics.

No. Clasif.	Autor	titulo
331.2 A	Am. Statistical Assn.	Proc. of the business and economic statistic section
519.9 W	Wolf, Frank	Elements of probability and statistics
310 W53	Wallis, W. Allen	Statistics a new approach
519.9 S65	Spiegel, Murray	Statistics
519.9 S47	Schreider, Yu A	Method of statistical testing Monte Carlo Method
311.2 O	Ostle, Bernard	Statistics in Researdh, Basic concepts and techniques for research workers.
519.7 K85	Kullback, Solomon	Information Theory and Statisties
531 K82	Kubo, Ryogo	Statistical mechanis and advanced course with problems and solutions
310 K	Kelley, Truman Lee	Fundamentals of Statistics
519.9 H33	Hadley, G.	Introduction to probability and statistical decision Theory
311.2 F	Fisher, R.	Les methodes statistiques -- adaptées a la recherche schientifique.
519.9 E	E.U.A. Nat. Science Foundation	Methodological Aspects of Statistics on research and development costs and manpower
519.9 E	Ehrenfeld, Sylvain	Introduction to statistical method
310 D59	Dixon, Wilfrid J	Introduction tostatistical analysis
310 D.	Davies, George R	Methods of statistical analysis in the social sciences
310 D	Deming, W Edwards	Statistical adjustment of data
311	Davenport, C B	Statistical methods in biology, medicine and psychology
519.93 C	Chorafas, Dimitris N	Statistical processes and reliability engineering
311.2 C953a	Croxton, Frederick Emory	Applied general statistic-

No. Clasificación	Autor	Título
F-972 U	U.S. Department of Commerce.	Probability tables for the analysis of extreme value data.
F-3615 H	Housner, George W.	Probability of ground shaking and estimates of maximum magnitude.
F-6500	W u, Tien Hsing.	The probability of foundation safety
F-724 V	Venuti, William J.	Probability of fatigue failure of bonded-type prestressed concrete beams.
F-6451 M	Markovic, Radmilo D.	Probability functions of best fit to distributions of annual precipitation and runoff.
F-8468 R	Ruiz, Patricio.	Probabilistic study of the behavior of structures during earthquakes,
F-7701 E	Edwards, Ward.	Probabilistic information processing systems design and evaluation.
F-3630 R	Ravindra, M.	Probabilistic evaluation of safety factors.
F-7242 B	Benjamin, J	A probabilistic basis for a deterministic code.
F-7148 R	Rascón Chávez, Octavio ..	Introducción a probabilidades y estadística.
F-5867 P	Pollock, S tephen M.	Sequential search and detection.
F-5599 B	Bustamante, Jorge I., y Rosenblueth, Emilio.	Probabilidad de cavernas en el subsuelo del conjunto urbano Nonoalco.
F-5893 B	Barlow, R.E., and Marshall, A.W.	Bounds on interval probabilities for restricted families of distributions.
C-1156 L	Lieblein, Julius .	A new method of analyzing extreme value data.
C-239 M	Morgan, P.A.P.	A probability theory of dams and storage, systems.
C-237 M	Moran, P.A.P.	A probability theory of a dam with a continuous release.
C-1287 H	Harsanyi, John C.	Games with incomplete information played by bayesian players, part III. The basic probability distribution of the game.

BIBLIOTECA DE LAS DIVISIONES
DE INVESTIGACION Y DE ESTUDIOS SUPE-
RIORES DE LA FACULTAD DE INGENIERIA
U N A M

Apartado Postal 70156
Mexico 20, D. F. México

519.9 F — Fisher, Ronald A
Tablas estadísticas para investigadores científicos, económicos, demográficos y especialmente biológicos, agrónomicos y médicos

517 S65 — Some aspects of analysis and probability

519.9 S55 — Smith, Walter L
Proceedings of the symposium on congestion theory.

519 R52 — Richter, Hans
Wahrscheinlichkeits-theorie

519 P — Pfeiffer, Paul E
Concepts of probability theory

519.9 P — Peters, Charles C
Statistical procedures and their mathemarical bases

519.1 P36 — Papoulis, Athanasios
Probability random variables and stochastic processes

519 M68 — Mosteller, Frederick
Probability with statistical applications

519 M58 — Mises, Richard von
Probability, statistics and truth

519 M — Manheim, Marvin L
Highway route location as hierarchicallistructured sequential decision process: An experiment in the use of bayesian decision theory for guiding and engineering process

519 M55 — Miller, Irwin
Probability and statistics for engineers

519 L35 — Lamperti, John
Probabilty; a survey of the mathematical theory

519.1 K88 — Kushner, Harlod J
Stochastic stability and control

519.92 G — Good, I J
Probabilty and the wighing of evidence

519 G54 — Gnedenko, B V
The theory of probability

BIBLIOTECA DE LAS DIVISIONES
DE INVESTIGACION Y DE ESTUDIOS SUPE-
RIORES DE LA FACULTAD DE INGENIERIA
U N A. M

Apartado Postal 70156
Mexico 20, D F. México

519 F74	Freudenthal, Hans Probability and statistics
519.1 E68	Epstein, Richard A The Theory of gambling and statistica logic
519.9 D47	Derman, Cyrus Probability and statistical inference for engineers
519.9 C	Clark, Charles E Exponentially distributed random numbers
519 C45	Centre National de la Recherche Scientifique Le calcul des probabilités et ses applications
519.1 B47	Bertrand, J Cálcul des probabilités

BIBLIOTECA DE LAS DIVISIONES
DE INVESTIGACION Y DE ESTUDIOS SUPE-
RIORES DE LA FACULTAD DE INGENIERIA
U. N. A. M.

Apartado Postal 70156
Mexico 20, D. F. México

310.23 C89	Cox, David Roxbee	The statistical analysis of series of events
530.132 B45	Bendat, Julius S	Measurement and analysis of random data

VNIVERSIDAD NACIONAL AVTONOMA

ESTADISTICA DESCRIPTIVA (1)

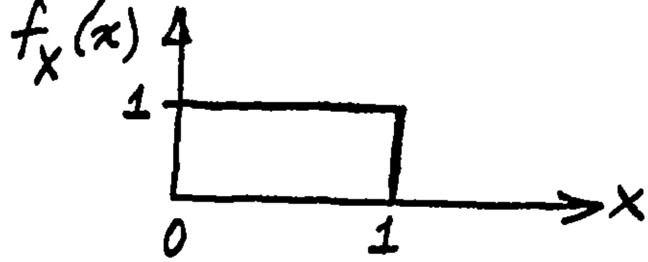
Dato y observación.- Es el resultado de realizar un experimento.

Muestra.- Es una colección de datos

Población.- Total de datos que se pueden obtener al realizar una secuencia exhaustiva de experimentos

Muestra aleatoria.- Es una muestra obtenida de tal manera que todos los elementos de la población tienen la misma probabilidad de ser observados y además la observación de un elemento no afecta la probabilidad de observar cualquier otro (si son independientes)

Tabla de números aleatorios.- Es una tabla que contiene números que constituyen una muestra aleatoria obtenida de una distribución de probabilidades uniforme, que generalmente corresponde a una variable aleatoria que puede asumir valores entre 0 y 1.



Las tablas que se usen para obtener una muestra aleatoria deben contener números con mayor número de dígitos que los que tiene el total de elementos de la población que se va a muestrear. Por ejemplo, si se va a obtener una muestra aleatoria de un lote de refacciones que tiene 10,000 elementos, la tabla que se use deberá tener números aleatorios con 5 o más dígitos.

Método de muestreo

1.- Se numeran los elementos de la población

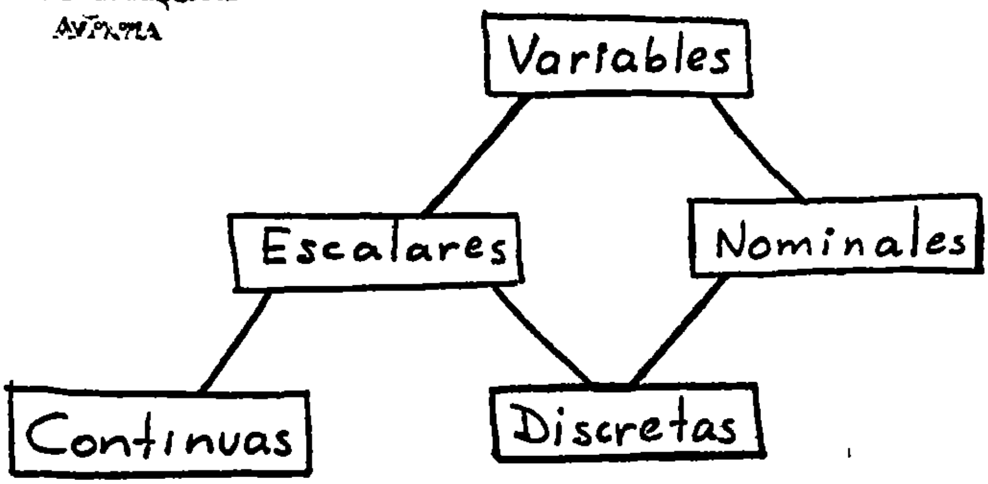
2.- Se fija el criterio de selección de los números aleatorios (por ejemplo, se define qué renglones y qué columnas se van a leer)

3.- Se indica qué dígitos se van a eliminar en caso de que los números de la tabla tengan más dígitos que los necesarios.

4.- Se leen los números, de acuerdo con lo fijado en los puntos 2 y 3, y se extraen del lote los elementos que tienen los números leídos. Estos constituyen la muestra física deseada, con la cual realizar los experimentos. Las observaciones constituirán la muestra deseada.

Universidad Nacional Autónoma

Variables
- **Escalares.-** Son las que asumen valores numéricos.
- **Nominales.-** No asumen valoras numéricos, sino nombres o atributos.



Agrupamiento de datos

Con objeto de facilitar la interpretación de los datos que se tienen en una muestra, es conveniente agruparlos ya sea por valores o por intervalos. Para facilitar el agrupamiento, es útil ordenarlos en forma creciente (o decreciente) de valores, cuando se trata de datos de variables escalares, formando así una tabla de datos ordenados.

Frecuencia de un evento.- Es el número de veces que ocurre el evento al obtener una muestra de la población correspondiente.

UNIVERSIDAD NACIONAL AUTÓNOMA

Frecuencia relativa de un evento.- Es el cociente de la frecuencia entre el total de elementos (tamaño) de la muestra.

Frecuencia relativa acumulada.- Es la acumulación (suma) de las frecuencias relativas hasta un valor dado.

Ejemplo.- En una escuela secundaria se les aplicó a 30 profesores un examen sobre pedagogía. Las calificaciones (datos) que se obtuvieron fueron (ya están ordenados)

$\underbrace{57, 59}_{A}, \underbrace{65, 67, 67, 67, 69}_{B}, \underbrace{72, 73, 73, 77, 78, 78}_{C},$
$\underbrace{81, 81, 83, 83, 83, 84, 84, 87, 88, 89, 89}_{D}, \underbrace{91, 91, 93,}_{E}$
$\underbrace{95, 97, 99}_{E}$

Calificación	frecuencia	Frecuencia relativa	Frecuencia relativa acumulada
57	1	1/30	1/30
59	1	1/30	2/30
65	1	1/30	3/30
67	3	3/30	6/30
69	1	1/30	7/30
72	1	1/30	8/30
73	2	2/30	10/30
77	1	1/30	11/30
78	2	2/30	13/30
81	2	2/30	15/30
83	3	3/30	18/30

(Continuación)

Calif.	Frec.	Frec. Rel.	Frec. Rel. Acum.
84	2	2/30	20/30
87	1	1/30	21/30
88	1	1/30	22/30
89	2	2/30	24/30
91	2	2/30	26/30
93	1	1/30	27/30 ←
95	1	1/30	28/30
97	1	1/30	29/30
99	1	1/30	30/30 = 1
	$\Sigma = 30$	$\Sigma = 30/30 = 1$	

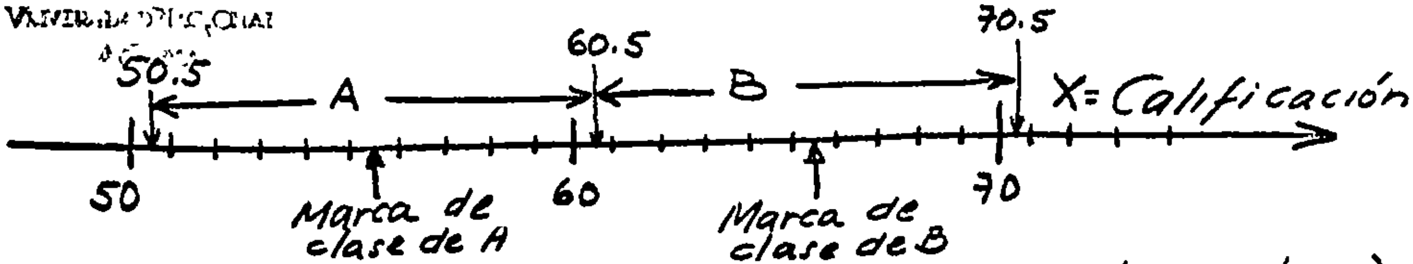
¿Cuál es la frecuencia de valores $\leq$ que 93?: 27/30

Con objeto de facilitar aún más la interpretación de los datos, es conveniente a menudo agruparlos por intervalos de valores.

Evento (intervalo de calificaciones)	Elementos observados	Frec.	Frec. Rel.
A: 51 – 60	57, 59	2	2/30
B: 61 – 70	65, 67, 67, 67, 69	5	5/30
C: 71 – 80	72, 73, 73, 77, 78, 78	6	6/30
D: 81 – 90	81, 81, 83, 83, 83, 84, 84, 87, 88, 89, 89	11	11/30
E: 91 – 100	91, 91, 93, 95, 97, 99	6	6/30
	$\Sigma =$	30	30/30 = 1

Límites inferiores de clase

Límites superiores de clase



En una escala continua (variable continua):

$$A = \{X : 50.5 < X \leq 60.5\}$$

$$B = \{X : 60.5 < X \leq 70.5\}$$

$$C = \{X : 70.5 < X \leq 80.5\}$$

$$D = \{X : 80.5 < X \leq 90.5\}$$

$$E = \{X : 90.5 < X \leq 100.5\}$$

Límites reales inferiores de clase

Límites reales superiores de clase

A mayor número de datos se requiere mayor número de intervalos, pero se recomienda que éste número esté entre 5 y 20, suponiendo que en promedio caigan unos 5 elementos en cada intervalo. Así, si se tienen 30 datos, se recomienda usar $30/5=6$ intervalos. Los límites reales deben tener una cifra decimal más que los datos.

Ejemplo

En un estudio antropológico se obtuvo una muestra de 30 estaturas de los varones residentes en una región. Los datos, ordenados en forma creciente de valores, fueron los siguientes:
160, 161, 163, 163, 163, 167, 167, 167, 167, 168, 168,
168, 169, 169, 170, 171, 171, 173, 174, 175, 175,
175, 178, 179, 181, 181, 183, 184, 187, 191 cm. Obtenga la tabla de frecuencias.

Solución

1.- Determinación del rango

rango = valor máximo − valor mínimo = 191 − 160 = 31 cm

2.- Determinación del número de intervalos

$$\text{núm. de intervalos} \doteq \frac{30}{5} = 6$$

(para continuar con el ejemplo tomaremos 6 intervalos)

3.- Determinación de los límites de clase

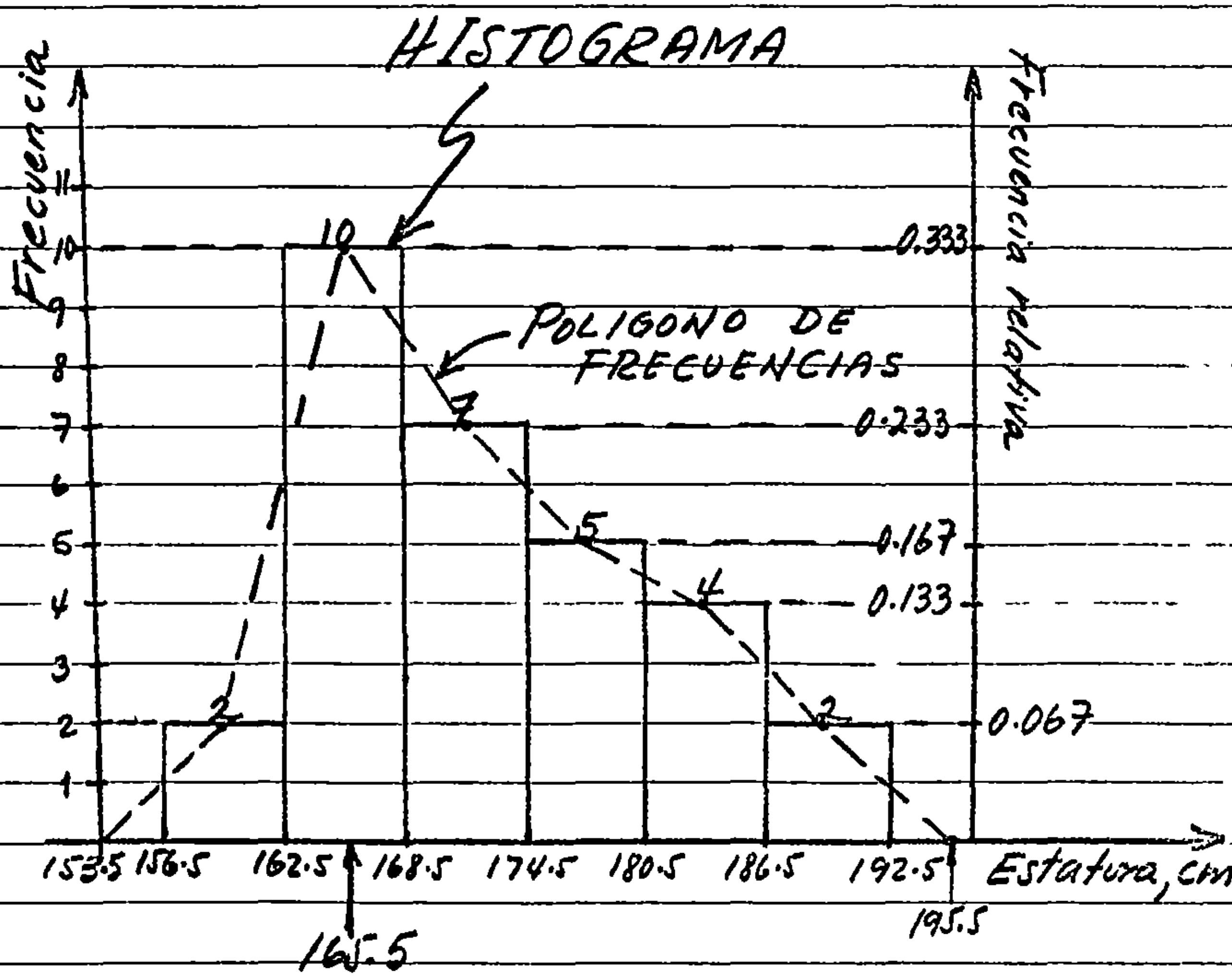
$$\text{Ancho de los intervalos} \doteq \frac{\text{rango}}{\text{número}} = \frac{31}{6} = 5.1 \Rightarrow 6$$

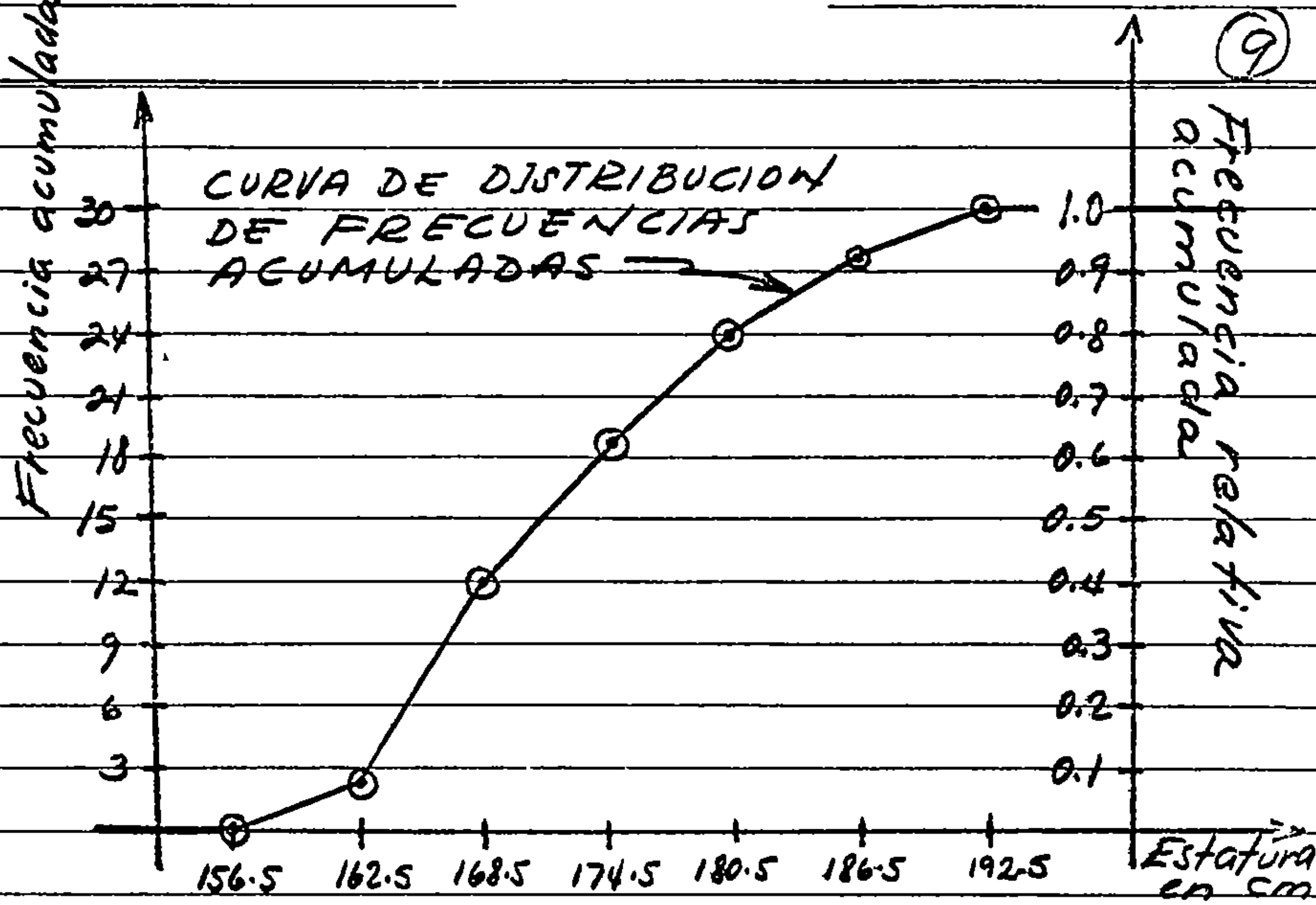
157 − 162, 163 − 168, 169 − 174, 175 − 180,
181 − 186, 187 − 192

4.- Integración de la tabla:

Intervalo	Límites Reales Inf.	Límites Reales Sup.	Frec.	Frec. Rel.	Frec. Acum.	Frec. Rel. Acum.
157 - 162	156.5	162.5	2	$\frac{2}{30} = 0.067$	2	0.067
163 - 168	162.5	168.5	10	$\frac{10}{30} = 0.333$	12	0.400
169 - 174	168.5	174.5	7	$\frac{7}{30} = 0.233$	19	0.633
175 - 180	174.5	180.5	5	$\frac{5}{30} = 0.167$	24	0.800
181 - 186	180.5	186.5	4	$\frac{4}{30} = 0.133$	28	0.933
187 - 192	186.5	192.5	2	$\frac{2}{30} = 0.067$	30	1.000
$\Sigma =$			30	1.000		

Presentación gráfica de las distribuciones de frecuencias





Ya vimos que la frecuencia acumulada corresponde a los valores menores o iguales que algún valor dado de la variable. De igual manera la <u>frecuencia</u> <u>acumulada</u> <u>complementaria</u> corresponde a los valores mayores que ese valor dado. Así:

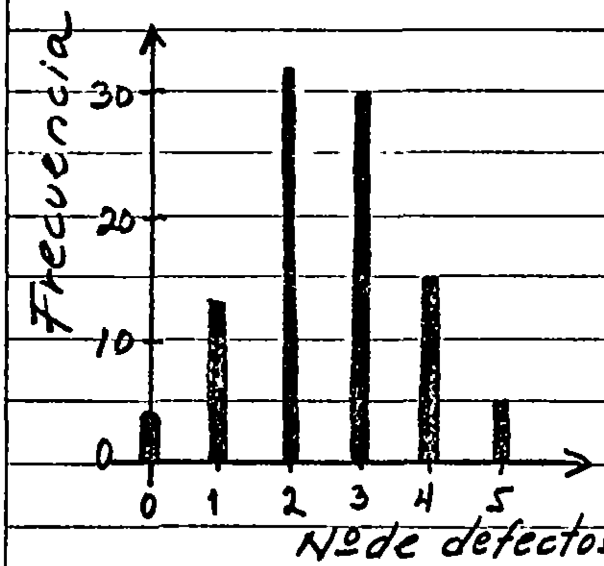
Frecuencia acumulada complementaria = número de datos − frecuencia acumulada

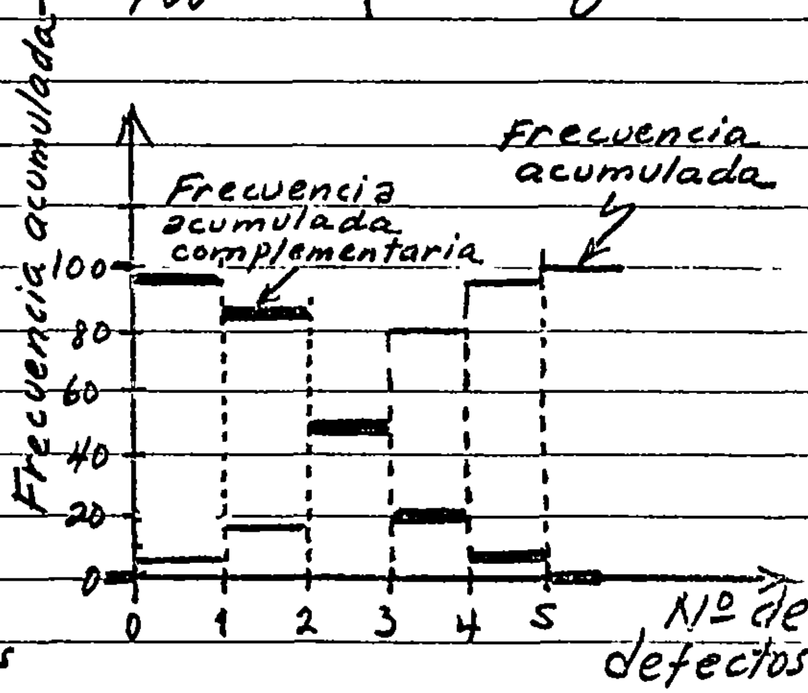
¿Cuál es la frecuencia de valores mayores que 180.5? : 30 − 24 = 6

La frecuencia acumulada relativa complementaria es: 1 − 0.800 = 0.200 (20%)

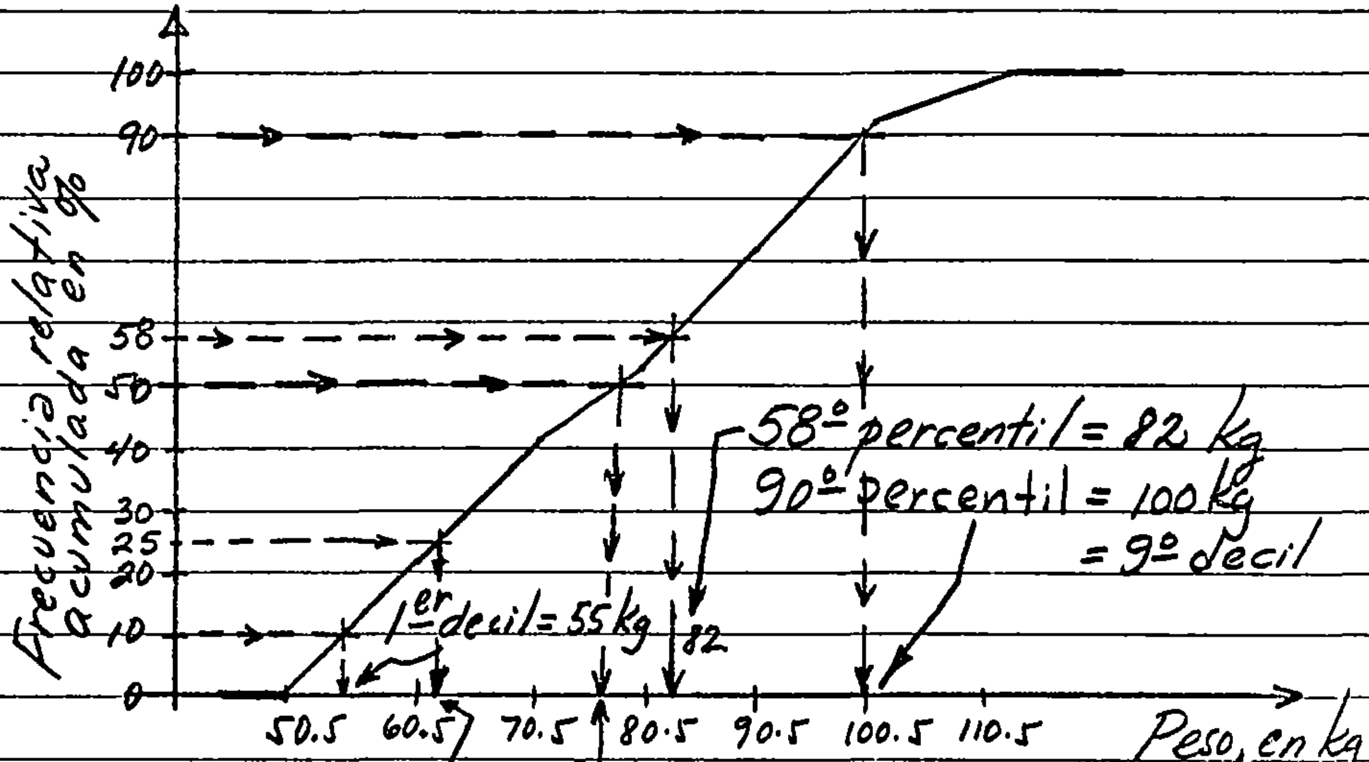
Ejemplo.- En un estudio sobre la calidad de los monobloks producidos por una fábrica, se obtuvo una muestra aleatoria de 100 elementos, a los cuales se les contó el número de defectos de fabricación. La tabla de frecuencias es la siguiente:

Número de defectos	Frecuencia	Frecuencia acumulada	Frecuencia acumulada complementaria
0	4	4	96
1	13	17	83
2	33	50	50
3	30	80	20
4	15	95	5
5	5	100	0
	100		





Percentiles – Son los valores de la variable correspondientes a frecuencias relativas acúmuladas que son múltiplos de 1 por ciento.



25º percentil = 1er cuartil = 62 kg ; Mediana = 76 kg = 5º decil = 2º cuartil

Deciles – Son los valores de la variable correspondientes a frecuencias relativas acumuladas que son múltiplos de 10 por ciento.

Cuartiles – Son los valores de la variable correspondientes a frecuencias relativas acumuladas que son múltiplos de 25 por ciento.

Mediana – Valor de la variable correspondiente a la frecuencia relativa acumulada de 50%.

Medidas representativas de los datos.

Medidas de TENDENCIA CENTRAL

Valor medio o promedio aritmético

$$\boxed{\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i}$$

donde x_i son los valores de los datos y n es el tamaño de la muestra.
Si los datos están agrupados y f_j es la frecuencia del intervalo j-ésimo y x_j es la marca de clase correspondiente, entonces

$$\boxed{\bar{x} = \frac{1}{n} \sum_{j=1}^{K} f_j x_j}$$

Ejemplo.- Sea el ejemplo enunciado en la mica Nº 10 de los defectos en monoblocks. Se tenía:

j	Nº de defectos x	Frecuencia f	fx
1	0	4	4x0 = 0
2	1	13	13x1 = 13
3	2	33	33x2 = 66
4	3	30	30x3 = 90
5	4	15	15x4 = 60
K=6	5	5	5x5 = 25
		$\Sigma = 100$	$\sum_{j=1}^{6} = 254$

$\bar{x} = \frac{254}{100} =$

$\bar{x} = 2.54$ defectos por monoblock

Modo.- Es el valor de la variable que aparece con mayor frecuencia en una muestra. Si los datos están agrupados, el modo es la marca de clase del intervalo que tiene la mayor frecuencia.

Ejemplo.- En el problema de los monoblocks el modo es 2. En el problema de las estaturas el modo es 165.5 cm

Mediana.- Es el valor de la variable que corresponde al 50% de la frecuencia relativa acumulada.
Si los datos están agrupados por intervalos la mediana se puede calcular con la fórmula (además de gráficamente, como ya se vio)

$$\boxed{\text{Mediana} = M = L_M + \frac{\frac{n}{2} - F_M}{f_M} d_M}$$

donde L_M = límite inferior real del intervalo que contiene a la mediana

f_M y d_m = respectivamente a la frecuencia y ancho del intervalo del intervalo que contiene a la mediana

F_M = frecuencia acumulada hasta el intervalo que contiene a la mediana exclusive

n = Tamaño de la muestra

Ejemplo- En un estudio para determinar los tiempos en que una muestra aleatoria de individuos reaccionaba a ciertos estímulos psicológicos se obtuvo lo siguiente:

j	Marca de clase x, en seg	Frecuencia f	Frecuencia acumulada, F	fx, seg	Límites reales
1	0.10	2	2	0.20	0.075-0.125
2	0.15	7	9	1.05	0.125-0.175
3	0.20	14	23	2.80	0.175-0.225
4	0.25	4	27	1.00	0.225-0.275
$K=5$	0.30	3	30	0.90	0.275-0.325
		$\Sigma = 30$		$\sum_{j=1}^{5} f_i x_i = 5.95$	

$$\bar{x} = \frac{5.95}{30} = \underline{0.198} \text{ seg}$$

$$\text{Modo} = \underline{0.20} \text{ seg}$$

$$d_M = 0.05, \quad L_M = 0.20 - \frac{0.05}{2} = 0.175, \quad F_M = 9$$

$$n/2 = 30/2 = 15, \quad f_M = 14$$

$$\text{Mediana} = M = 0.175 + \frac{15-9}{14} 0.05$$

$$M = 0.175 + \frac{0.30}{14} = 0.175 + 0.021 = \underline{0.196} \text{ seg}$$

BLOCK TEYCO

MEDIDAS DE DISPERSION

Rango = Máximo valor observado − Mínimo valor observado

Variancia: Si los datos no están agrupados:

$$\boxed{S_x^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Si los datos están agrupados

$$\boxed{S_x^2 = \frac{1}{n}\sum_{j=1}^{k}(x_j - \bar{x})^2}$$

donde las x_j son los valores de las marcas de clase de los intervalos.

Desviación estándar

$$\boxed{S_x = \sqrt{S_x^2}}$$

Coeficiente de variación

$$\boxed{v_x = S_x / \bar{x}}$$

Ejemplo: En un estudio sobre la temperatura máxima diaria en una ciudad se obtuvo lo siguiente para una primavera:

j	Intervalos de temperatura °F	Marca de clase, °F	Frecuencia f	xf	$x-\bar{x}$	$(x-\bar{x})^2$	$(x-\bar{x})^2 f$
1	55 – 63	59	2	118	-21.3	453.7	907.4
2	64 – 72	68	6	408	-12.3	151.3	907.8
3	73 – 81	77	7	539	-3.3	10.9	76.3
4	82 – 90	86	9	774	5.7	32.5	292.5
5	91 – 99	95	6	570	14.7	216.1	1296.6
			30	2409			3480.6

$$\bar{x} = \frac{2409}{30} = 80.3\ °F\ ;\ S_x^2 = \frac{3480.6}{30} = 116\ °F^2$$

$$S_x = \sqrt{116} = 10.8\ °F,\ V_x = \frac{10.8}{80.3} = 0.134\ (13.4\%)$$

Modo = 86

$$d_M = 9,\ L_M = 72.5,\ f_M = 7,\ F_M = 8,\ \frac{n}{2} = \frac{30}{2} = 15$$

$$\text{Mediana} = M = 72.5 + \frac{15-8}{7}9 =$$

$$= 72.5 + 9 = 81.5\ °F$$

BIBLIOGRAFIA DE CONTROL DE CALIDAD

Clave de biblioteca de la DESFI

1. Engineering Statistics 519.9
Bowker and Lieberman
Prentice-Hall, Inc. Englewood Cliffs, N. J., 1959

2. Stastistics, a new approach 310 w 35
Wallis and Roberts
(The frees press, New York), 1956

3. Teoría de probabilidades y aplicaciones 519 c73
Cramer
Aguilar S.A. de Ediciones Madrid, 1963

4. Elements of mathemathical statistics 519.9 A.54
Alexander
John Wiley & Sons, Inc., 1961

5. Statistical processes and reliability engineering 519.93 c
Dimitris N. Chorafas
D. van Nostrand Company Inc. Princeton New Yersey, 1960

6. Experimental statistics, Handbook 91 519.9 n
M.G. Natrella
U.S. Department of Commerce (Superintendent of Documents
U.S. Government printing office,
Washington 25, D.C.)

7. Statistical Theory with Engineering Applications 519.9 H
H. Hald
John Wiley & Sons. Inc., 1952

8. Quality control, highway research record F-6428 H
No. 184
Ed. Highway Research Board, 1967

9. Statistics theory and problems HA.29 S65
Murray R. Spiegel
Schaum Publishing Company, 1961

10. Estadística aplicada HA29 081
Bernard Ostle
Editorial Limusa-Wiley, S.A., 1973

11. The design of sample surveys HA31.2 R33
DES RAJ
McGraw Hill Book Company, 1972

12. Mathemathical methods of statistical quality control TS156 S35
Sarkadi Vince
Academic Press, New York and London, 1974

13. Sensibilidad estadística de cartas de control de calidad FUH 7-17
Universidad de La Habana - Centro de Información Científica y Técnica, 1971

14. Curso sobre control de calidad TS15 M48
Centro de Educacion Continua
Editado por Centro de Educación Continua, 1973

15. Control de calidad, Seminario F-9158 S
Memoria del Instituto Mexicano del Petróleo
Publicación No. 69 JF/050
Ed. Instituto Mexicano del Petróleo, 1969

16. Scientific method in production management TS 155 G43
G.R. Gedye
Oxford University Press, 1969

17. Manual de la ingeniería de la producción industrial T56 M38
H.B. Maynard
Editorial Reverté, 1960

18. Control de calidad estadístico
E.L. Grant
Ed. CECSA, 1974

19. Control de calidad (Proper Quality Control)
Johnson, Sidney M.
Ed. ASCE Structural Engineering. Conference / Seattle Washington May (8-12) 1967

20. Control total de la calidad
A.V. Feigenbaum
Editado por C.E.C.S.A., 1975

21. Practical Quality Control
Simmons A. David
Ed. Addison-Wesley Publishing Company, 1970

22. Bulletin Congress de la Route
No. 176, 1965, Vol. 54

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

PROBABILIDAD Y ESTADISTICA: FUNDAMENTOS Y APLICACIONES

TAREA No. 1
Revisión

DR. OCTAVIO RASCON CHAVEZ
Marzo, 1976

Palacio de Minería
Tacuba 5, primer piso México 1, D F
Tels 521-40-23 521-73-35 512-31-23

TAREA Nº 1 (9) 29/Feb/76.

Jesús Rios Alvarado.

Problema # 6 (pág 12).- En cualquiera de los experimentos siguientes decida cuando será apropiado utilizar un espacio de puntos finito, enumerable infinito o continuo:

a).- Uno de dos vicepresidentes va a elegirse presidente de una compañía.

b).- Se realiza un experimento para medir el coeficiente de expansión de los ladrillos de cierta fábrica.

c).- Con un contador Geiger se hacen mediciones de intensidad de radiación.

d).- Un policía mide el contenido de alcohol en la sangre de un conductor.

e).- Una moneda es lanzada hasta que aparece la primera cara.

f).- Se efectúa un estudio de tráfico para estimar el número de automóviles en el condado con defectos en las luces.

Respuestas:

a).- finito ✓

b).- continuo ✓

c).- continuo ✓

d).- continuo ✓

e).- enumerable infinito ✓

f).- enumerable infinito ✓

Problema # 7 (pág. 13) - Entre 6 aplicantes a un trabajo

- el Sr. Andrews es casado, no juega golf y no es propietario de casa.
- el Sr. Bailey es soltero, no juega golf y es propietario de casa
- el Sr. Clark es casado, juega golf y es propietario de casa
- el Sr. Dodds es soltero, juega golf y no es propietario de casa
- el Sr. Edwards es casado, no juega golf y es propietario de casa y
- el Sr. Fox es soltero, juega golf y es propietario de casa

Uno de estos aplicantes va a obtener el trabajo y el evento de que el trabajo, por ejemplo se asigne a un jugador de golf se denota {Clark, Dodds, Fox}. Indique, en una manera similar, conjuntos que denoten los eventos

(a) La posición se asigna a un propietario de casa, (b) la posición se asigna a un jugador de golf casado, (c) la posición se asigna a una persona soltera que no es propietario de casa, (d) la posición se asigna a un hombre que no es ni casado ni un jugador de golf soltero.

NOMBRE	Casado	Juega Golf.	Propietario de casa
Andrews	Si	No	No
Bailey	No	No	Si
Clark	Si	Si	Si
Dodds	No	Si	No
Edwards	Si	No	Si
Fox	No	Si	Si

Respuesta: (a) = {Bailey, Clark, Edward, Fox} (b) = {Clark}
(c) = {Dodds} (d) = {Bailey}

TAREA N° 1 — 24/Feb/76

Jesús Ríos Alvarado

3/4

Problema #8 (pág. 13).- Refiriéndose al ejercicio #7, indicar:

(a) el complemento del conjunto dado en la parte (a),

(b) la unión de los conjuntos de las partes (a) y (b),

(c) la intersección de los conjuntos de las partes (a) y (d),

(d) la intersección de los conjuntos de las partes (b) y (c).

Respuestas.-

(a) {Dodds, Andrews} (b) {Bailey, Clark, Edwards, Fox}

(c) {Bailey} (d) {ϕ} ϕ

7/8

Problema #9 (pág. 13).- Un fabricante compra partes de cuatro diferentes proveedores numerados 1, 2, 3, y 4. Refiriéndose a órdenes colocadas en dos días consecutivos, (1,4) denota el evento en que el primer día la orden fue otorgada al vendedor 1 y el segundo día fue otorgada al vendedor 4. Si A representa el evento en que el vendedor 1 obtiene cuando menos una de las órdenes, B el evento en que el mismo vendedor obtiene ambas órdenes y C el evento en que los vendedores 1 y 3 no obtienen ninguna orden, enliste los elementos de

(a) el espacio de eventos, (b) A, (c) B, (d) C.

(e) $\bar{A}$, (f) $B \cup C$, (g) $A \cap B$, (h) $A \cap C$.

Respuestas.-

(a)

$$S \begin{Bmatrix} (1,1) & (1,2) & (1,3) & (1,4) \\ (2,1) & (2,2) & (2,3) & (2,4) \\ (3,1) & (3,2) & (3,3) & (3,4) \\ (4,1) & (4,2) & (4,3) & (4,4) \end{Bmatrix}$$

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

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

(d) $C = \{(2,2),(2,4),(4,2),(4,4)\}$ $(2,4)$ (e) $\bar{A}\ \{(2,2),(2,3),(2,4),(3,2),(3,3),(3,4),(4,2),(4,3),(4,4)\}$

(f) $B \cup C = \{(1,1),(2,2),(2,3),(3,2),(4,4)\}$ (g) $A \cap B = \{(1,1)\}$

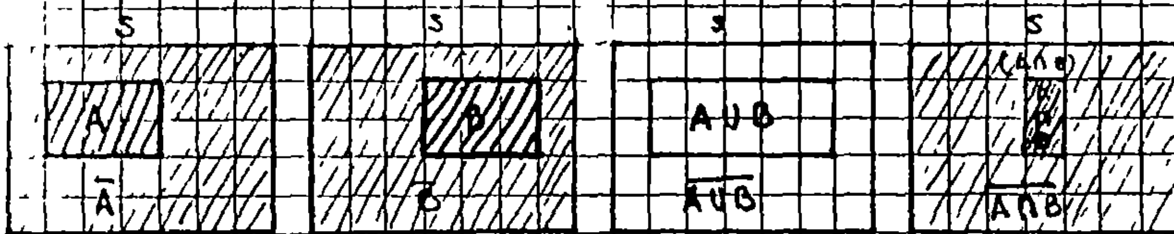
(h) $A \cap C = \cancel{\{A\}}\ \phi$

Problema # 10 (pag. 13) - Usa un diagrama de Venn para verificar que

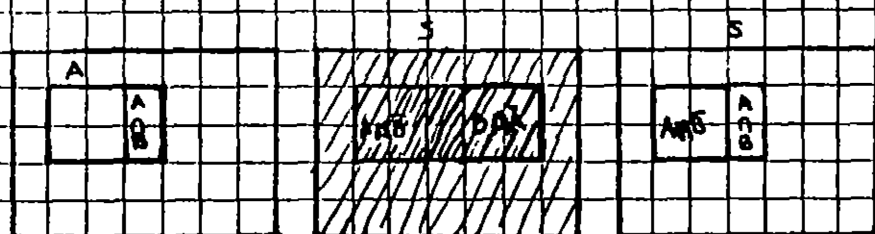
(a) $\overline{(A \cap B)} = \bar{A} \cup \bar{B}$ (b) $A \cup (A \cap B) = A$ (c) $(A \cap B) \cup (A \cap \bar{B}) = A$

(d) $A \cup B = (A \cap \bar{B}) \cup (A \cap B) \cup (\bar{A} \cap B)$

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

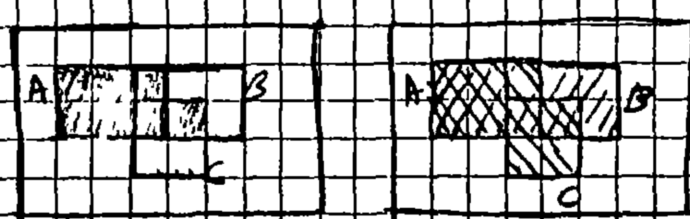


(a) $\overline{A \cap B} = \bar{A} \cup \bar{B}$



(b) $A \cup (A \cap B) = A$ (c) $(A \cap B) \cup (A \cap \bar{B}) = A$

(d) $(A \cap \bar{B}) \cup (A \cap B) \cup (B \cap \bar{A}) = (A \cup B)$



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

TAREA Nº 1 29/Feb./76

Jesús Rios Alvarado

5 ✓ 6

Problema #3 (pág 22).- Entre 100 estudiantes de ingeniería 15 están estudiando para ser ingenieros químicos, 60 son no-graduados y 5 son no graduados estudiando ingeniería química. Cuantos de esos estudiantes

(a) no están estudiando ingeniería química

(b) no son no-graduados (son graduados?)

(c) están estudiando ingeniería química, son no-graduados o ambas cosas.

(d) no son graduados y no están estudiando ingeniería química.

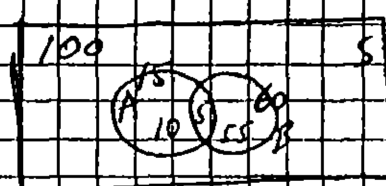
(e) no son graduados estudiando ingeniería química

(f) no son no-graduados que no estudian para ingenieros químicos.

Respuestas.-

S = 100 estudiantes de ingeniería ✓

A = 15 estudiantes de ingeniería química ✓



B = 60 estudiantes no graduados ✓

$C = A \cap B$ = 5 estudiantes no graduados estudiando ingeniería química ✓

(a) $\bar{A}$ = estudiantes que no estudian ingeniería química = 85 ✓

(b) $\bar{B}$ = estudiantes graduados = 40 ✓

(c) ~~$A \cap B$~~ $A \cup B = 70$ ✓

(d) $B \cap \bar{A}$ = 55 ✓ estudiantes no graduados y no estudiando ingeniería química

X (e) $A \cap \bar{B}$ = 10 estudiantes no graduados estudiando ingeniería química

(f) $\bar{A} \cap \bar{B} = 30$ ✓

3

Problema 4.8 (pág 22).- Un experimento tiene exactamente cuatro resultados distintos. A, B, C y D. Verificar ~~cuando~~ si son permisibles las asignaciones de probabilidades siguientes:

(a) $P(A) = 0.36$, $P(B) = 0.18$, $P(C) = 0.21$, $P(D) = 0.25$ R.- SI

(b) $P(A) = 0.25$, $P(B) = 0.35$, $P(C) = 0.18$, $P(D) = 0.15$ R.- NO

(c) $P(A) = 0.42$, $P(B) = 0.17$, $P(C) = -0.08$, $P(D) = 0.49$ R.- NO

(d) $P(A) = 17/80$, $P(B) = 11/40$, $P(C) = \frac{1}{2}$, $P(D) = 1/80$ R.- NO SI

1/2

Problema 4.11 (pág 23).- Cual es la probabilidad de extraer (a) reina negra, (b) una carta roja, (c) 5, 6 ó 7, (d) un as negro o un rey rojo, de un mazo de 52 cartas bien mezclado?

(a) Respuestas.- S = {Mazo de 52 cartas americanas}

(b) A = {carta roja} $p(A) = 0.5$

(a) B = {reina negra} $p(B) = 1/52$ X $2/52$ $P(B) = \frac{1}{2} \times \frac{4}{52} = \frac{2}{52}$

(c) C = {5, 6, 7} $p(C) = 0.23$

(d) D = {as negro, rey ~~rojo~~} $p(D) = 1/13$ X $\frac{4}{52} = 1/13$

Problema 4.14 (pág 24).- Si A y B son eventos mutuamente exclusivos, $P(A) = 0.20$, $P(B) = 0.55$ encuentre:

(a) $p(\bar{A})$ (b) $p(A \cup B)$ (c) $p(A \cap B)$ (d) $P(\bar{A} \cap \bar{B})$

Respuesta.-

(a) $p(\bar{A}) = 0.80 = 1 - p(A)$ (b) $P(A \cup B) = 0.75 = P(A) + P(B) - \phi$

(c) $P(A \cap B) = 0 = \phi$ (d) $P(\bar{A} \cap \bar{B}) = 0.25 = P(\overline{A \cup B})$

PROBABILIDAD Y ESTADISTICA

Marzo 2/76

Aguilar Anibal Gustavo Homberto	7.3
Cabrera Cabazos Pedro C.	7.0
Calderón Caumiers Armando	6.6
Gamez Arista Ramón	4.2
Coello Lemarroy Mario A.	5.0
Maldonado Perez Maria de la Cruz	7.5
Martinez Omaña Guillermo A.	6.5
Eduardo Ramirez S.	6.9
Rios Alvarado Jesus	9.0
Zychlnski Zychlinska Benito	4.9

m. a. f. l.

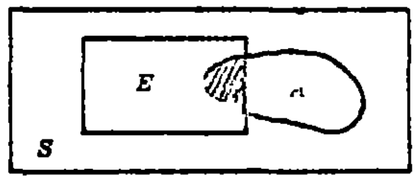
Capítulo 4

Probabilidad condicional e independencia

PROBABILIDAD CONDICIONAL

Sea E un evento arbitrario de un espacio muestral S con $P(E) > 0$ La probabilidad de que un evento A suceda una vez que E haya sucedido o, en otras palabras, la *probabilidad condicional* de A dado E, escrito $P(A \mid E)$, se define como sigue

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



Como se aprecia en el diagrama de Venn expuesto, $P(A \mid E)$ en cierto sentido mide la probabilidad relativa de A con relación al espacio reducido E

En particular, si S es un espacio finito equiprobable y $|A|$ denota el número de elementos de un evento A, entonces

$$P(A \cap E) = \frac{|A \cap E|}{|S|}, \quad P(E) = \frac{|E|}{|S|} \quad \text{y así} \quad P(A \mid E) = \frac{P(A \cap E)}{P(E)} = \frac{|A \cap E|}{|E|}$$

Esto es,

Teorema 4.1: Sea S un espacio finito equiprobable con eventos A y E Entonces

$$P(A \mid E) = \frac{\text{número de elementos de } A \cap E}{\text{numero de elementos de } E}$$

o

$$P(A \mid E) = \frac{\text{número de maneras en que } A \text{ y } E \text{ pueden suceder}}{\text{numero de maneras en que } E \text{ puede suceder}}$$

Ejemplo 4.1. Sea el caso de lanzar un par de dados corrientes Si la suma es 6, hallar la probabilidad de que uno de los dados sea 2 En otras palabras, si

$$E = \{\text{suma es } 6\} = \{(1,5), (2,4), (3,3), (4,2), (5,1)\}$$

y

$$A = \{\text{un 2 aparece por lo menos en un dado}\}$$

hallar $P(A \mid E)$

Ahora E consta de cinco elementos y dos de ellos, $(2,4)$ y $(4,2)$, pertenecen a A $A \cap E = \{(2,4)$ $(4,2)\}$ Entonces $P(A \mid E) = \frac{2}{5}$

Por otra parte, puesto que A consta de nueve elementos,

$$A = \{(2,1), (2,2), (2,3), (2,4), (2,5), (2,6), (1,2), (3,2), (4,2), (5,2), (6,2)\}$$

y S consta de 36 elementos, $P(A) = \frac{11}{36}$

Ejemplo 4.2: Un hombre visita a un matrimonio que tiene dos hijos Uno de los hijos, un niño, entra en la sala Hallar la probabilidad p de que el otro sea tambien nino si, (i) se sabe que el otro hijo (o hija) es menor, (ii) no se sabe nada del otro hijo

El espacio muestral para el sexo de los dos hijos es $S = \{bb\ bg\ gb, gg\}$ con probabilidad $\frac{1}{4}$ para cada muestra (Aquí la serie de cada punto corresponde a la serie de nacimientos)

(i) El espacio muestral reducido consta de dos elementos, $\{bb\ bg\}$, o sea $p = \frac{1}{2}$

(ii) El espacio muestral reducido consta de tres elementos, $\{bb, bg\ gb\}$, o sea $p = \frac{1}{3}$

TEOREMA DE LA MULTIPLICACION PARA PROBABILIDAD CONDICIONAL

Si multiplicamos en cruz la ecuación anterior que define la probabilidad condicional y usamos el hecho de que $A \cap E = E \cap A$, obtenemos la siguiente formula util

Teorema 4.2: $P(E \cap A) = P(E)\ P(A \mid E)$

Este teorema puede extenderse por inducción matemática como sigue:

Corolario 4.3. Para los eventos $A_1, A_2, \ldots, A_n$,

$$P(A_1 \cap A_2 \cap \cdots \cap A_n)$$
$$= P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \cdots P(A_n \mid A_1 \cap A_2 \cap \cdots \cap A_{n-1})$$

Ahora aplicamos el teorema anterior que es llamado, apropiadamente, el *teorema de la multiplicacion*

Ejemplo 4.3. Un lote de 12 artículos tiene 4 defectuosos Se toman al azar tres artículos del lote uno tras otro Hallar la probabilidad p de que todos los tres estén buenos

La probabilidad de que el primer articulo no sea defectuoso es $\frac{8}{12}$ puesto que 8 entre los 12 no son defectuosos Si el primero no es defectuoso entonces la probabilidad de que el próximo artículo no sea defectuoso es $\frac{7}{11}$ puesto que solamente 7 de los 11 sobrantes no son defectuosos Si los dos primeros artículos no son defectuosos, entonces la probabilidad de que el último no sea defectuoso es $\frac{6}{10}$ puesto que solamente 6 entre los 10 que quedan no son defectuosos Así por el teorema de la multiplicación,

$$p = \frac{8}{12} \cdot \frac{7}{11} \cdot \frac{6}{10} = \frac{14}{55}$$

PROCESOS ESTOCASTICOS FINITOS Y DIAGRAMAS DE ARBOL

Una sucesión (finita) de experimentos en los cuales cada experimento tiene un número finito de resultados con probabilidades dadas se llama un *proceso estocastico (finito)* Una manera conveniente de describir tal proceso y calcular la probabilidad de un evento se obtiene por el *diagrama de arbol* como se ilustra en la figura siguiente, el teorema de la multiplicación de la sección anterior se usa para calcular la probabilidad de que el resultado representado por una trayectoria determinada del árbol suceda. suceda

Ejemplo 4.4 Tomemos las tres cajas siguientes

Caja I contiene 10 lámparas de las cuales 4 son defectuosas

Caja II contiene 6 con 1 defectuosa

Caja III contiene 8 con 3 defectuosas

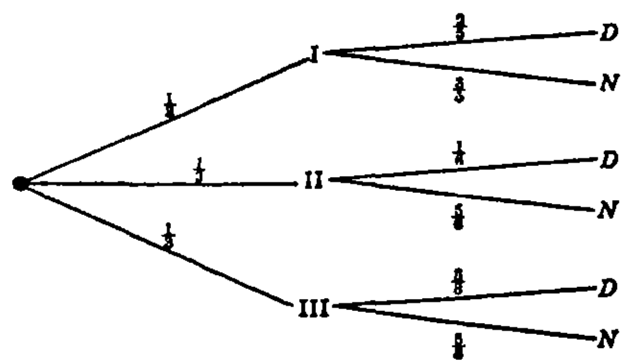
Escogemos al azar una caja y luego sacamos al azar una lámpara ¿Cuál es la probabilidad p de que la lámpara sea defectuosa?

Aquí realizamos una serie de dos experimentos

(i) escoger una de las tres cajas,

(ii) escoger una lampara que sea o defectuosa (D) o no defectuosa (N)

El diagrama de árbol siguiente describe el proceso y da la probabilidad de cada rama del árbol



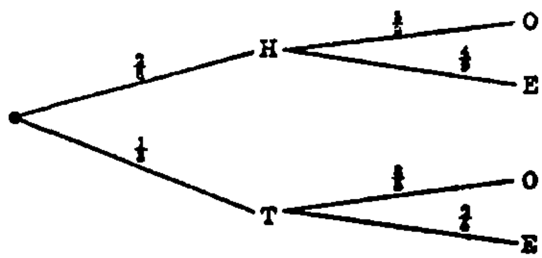
La probabilidad de que una trayectoria determinada del árbol suceda es, según el teorema de la multiplicación, el producto de las probabilidades de cada rama de la trayectoria, o sea que la probabilidad de escoger la caja 1 y luego una lámpara defectuosa es $\frac{1}{3}\cdot\frac{2}{5} = \frac{2}{15}$.

Ahora como hay tres trayectorias mutuamente exclusivas que conducen a una lámpara defectuosa, la suma de las probabilidades de estas trayectorias es la probabilidad buscada

$$p = \frac{1}{3}\cdot\frac{2}{5} + \frac{1}{3}\cdot\frac{1}{6} + \frac{1}{3}\cdot\frac{3}{8} = \frac{113}{360}$$

Ejemplo 4.5 Se lanza una moneda cargada de modo que $P(C) = \frac{2}{3}$ y $P(S) = \frac{1}{3}$. Si sale cara, se escoge al azar un número de 1 a 9; si sale sello, se escoge al azar un número de 1 a 5. Hallar la probabilidad p de que se escoja un número par.

El diagrama de árbol con las probabilidades respectivas es.

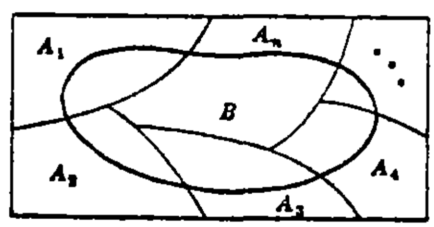


Obsérvese que la probabilidad de escoger un número par de 1 a 9 es $\frac{4}{9}$ puesto que hay 4 pares entre los 9 números, mientras que la probabilidad de escoger un par de 1 a 5 es $\frac{2}{5}$ puesto que hay 2 números pares entre los 5. Dos de las trayectorias conducen a un número par: CP y SP. Así

$$p = P(E) = \frac{2}{3}\cdot\frac{4}{9} + \frac{1}{3}\cdot\frac{2}{5} = \frac{58}{135}$$

PARTICIONES Y TEOREMA DE BAYES

Supongamos que los eventos $A_1, A_2, \ldots, A_n$ forman una partición de un espacio muestral S, esto es, que los eventos A_i son mutuamente exclusivos y su unión es S. Ahora sea B otro evento. Entonces



B sombreado

$$\begin{aligned} B &= S \cap B = (A_1 \cup A_2 \cup \cdots \cup A_n) \cap B \\ &= (A_1 \cap B) \cup (A_2 \cap B) \cup \cdots \cup (A_n \cap B) \end{aligned}$$

donde las $A_i \cap B$ son eventos mutuamente exclusivos. En consecuencia

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + \cdots + P(A_n \cap B)$$

Luego por el teorema de la multiplicación,

$$P(B) = P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + \cdots + P(A_n)P(B \mid A_n) \qquad (1)$$

Por otra parte, para cualquier i, la probabilidad condicional de A_i dado B se define por

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

En esta ecuación usamos (*1*) para remplazar $P(B)$ y usamos $P(A_i \cap B) = P(A_i)\,P(B \mid A_i)$ para remplazar $P(A_i \cap B)$, obteniendo así el

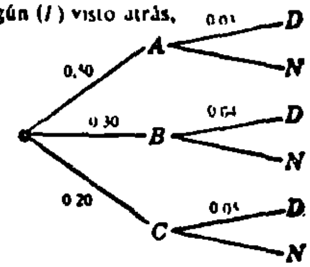
Teorema de Bayes 4.4 Supongase que $A_1, A_2, \ldots, A_n$ es una partición de S y que B es cualquier evento. Entonces para cualquier i,

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{P(A_1)P(B \mid A_1) + P(A_2)P(B \mid A_2) + \cdots + P(A_n)P(B \mid A_n)}$$

Ejemplo 4 6 Tres maquinas A B y C producen respectivamente 50% 30% y 20% del numero total de articulos de una fabrica Los porcentajes de desperfectos de produccion de estas maquinas son 3%, 4% y 5% Si se selecciona al azar un articulo hallar la probabilidad de que el articulo sea defectuoso

Sea X el evento de que un articulo es defectuoso Entonces según (I) visto atrás,

$$\begin{aligned} P(X) &= P(A)\,P(X\,|\,A) + P(B)\,P(X\,|\,B) \\ &\qquad + P(C)\,P(X\,|\,C) \\ &= (0\,50)(0\,03) + (0\,30)(0\,04) + (0\,20)(0\,05) \\ &= 0{,}037 \end{aligned}$$



Obsérvese que tambien podemos considerar este problema como un proceso estocastico que tiene el diagrama de arbol adjunto

Ejemplo 4 7 Considerese la fabrica del ejemplo anterior Supongase que se selecciona un articulo al azar y resulta ser defectuoso Hallar la probabilidad de que el articulo fue producido por la maquina A esto es hallar $P(A\,|\,X)$

Por el teorema de Bayes,

$$P(A\,|\,X) = \frac{P(A)\,P(X\,|\,A)}{P(A)\,P(X\,|\,A) + P(B)\,P(X\,|\,B) + P(C)\,P(X\,|\,C)}$$

$$= \frac{(0\,50)(0\,03)}{(0{,}50)(0\,03) + (0\,30)(0{,}04) + (0{,}20)(0{,}05)} = \frac{15}{37}$$

En otras palabras, dividimos la probabilidad de la trayectoria pedida por la probabilidad del espacio muestral reducido, o sea aquellas trayectorias que conducen a un artículo defectuoso

INDEPENDENCIA

Se dice que un evento B es *independiente* de un evento A si la probabilidad de que B suceda no esta influenciada porque A haya o no sucedido En otras palabras, si la probabilidad de B iguala la probabilidad condicional de B dado A $P(B) = P(B\,|\,A)$ Ahora sustituyendo $P(B)$ por $P(B|A)$ en el teorema de la multiplicacion $P(A \cap B) = P(A)\;P(B\,|\,A)$, obtenemos

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

Usamos la ecuación anterior como nuestra definicion formal de independencia

Definicion: A y B son eventos independientes si $P(A \cap B) = P(A)\;P(B)$, de otro modo son dependientes

Ejemplo 4 8 Láncese una moneda corriente tres veces obtenemos el espacio equiprobable

$$S = \{\text{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}\}$$

Consideremos los eventos

$A = \{$ primeros lanzamientos son caras $\}$ $\quad B = \{$ segundos lanzamientos son caras $\}$

$C = \{$ exactamente se lanzan dos caras seguidas $\}$

Claramente A y B son eventos independientes este hecho se verifica en seguida Por otra parte, la relación entre A y C o B y C no es obvia Insistimos en que A y C son independientes pero que B y C son dependientes Tenemos

$$P(A) = P(\{\text{HHH, HHT, HTH, HTT}\}) = \tfrac{4}{8} = \tfrac{1}{2}$$

$$P(B) = P(\{\text{HHH, HHT, THH, THT}\}) = \tfrac{4}{8} = \tfrac{1}{2}$$

$$P(C) = P(\{\text{HHT, THH}\}) = \tfrac{2}{8} = \tfrac{1}{4}$$

Entonces

$$P(A \cap B) = P(\{\text{HHH, HHT}\}) = \tfrac{1}{4}, \quad P(A \cap C) = P(\{\text{HHT}\}) = \tfrac{1}{8},$$

$$P(B \cap C) = P(\{\text{HHT, THH}\}) = \tfrac{1}{4}$$

En consecuencia

$$P(A)\,P(B) \;=\; \frac{1}{2}\cdot\frac{1}{2} \;=\; \frac{1}{4} \;=\; P(A\cap B), \quad \text{y así } A \text{ y } B \text{ son independientes,}$$

$$P(A)\,P(C) \;=\; \frac{1}{2}\cdot\frac{1}{4} \;=\; \frac{1}{8} \;=\; P(A\cap C), \quad \text{y así } A \text{ y } B \text{ son independientes,}$$

$$P(B)\,P(C) \;=\; \frac{1}{2}\cdot\frac{1}{4} \;=\; \frac{1}{8} \;\neq\; P(B\cap C), \quad \text{y así } B \text{ y } C \text{ son dependientes}$$

Frecuentemente, postularemos que dos eventos son independientes, o que será claro por la naturaleza del experimento que dos eventos son independientes

Ejemplo 4.9 La probabilidad de que A dé en el blanco es $\frac{1}{4}$ y la de B es $\frac{2}{5}$. Si A y B disparan, ¿cuál es la probabilidad de que se pegue al blanco?

Sabemos que $P(A) = \frac{1}{4}$ y $P(B) = \frac{2}{5}$ y buscamos $P(A \cup B)$. Además, la probabilidad de que A o B den en el blanco no depende de que el otro dé; esto es, el evento de que A dé en el blanco es independiente del evento de que B dé en el blanco: $P(A \cap B) = P(A)\,P(B)$. Así

$$\begin{aligned} P(A\cup B) \;&=\; P(A) - P(B) - P(A\cap B) \;=\; P(A) + P(B) - P(A)\,P(B) \\ &=\; \frac{1}{4} + \frac{2}{5} - \frac{1}{4}\cdot\frac{2}{5} \;=\; \frac{11}{20} \end{aligned}$$

Tres eventos A, B y C son *independientes* si

(i) $P(A\cap B) = P(A)\,P(B)$, $P(A\cap C) = P(A)\,P(C)$ y $P(B\cap C) = P(B)\,P(C)$

esto es, si los eventos son independientes dos a dos, y

(ii) $P(A\cap B\cap C) = P(A)\,P(B)\,P(C)$.

El próximo ejemplo muestra que la condición (ii) no se desprende de la condición (i), en otras palabras, tres eventos pueden ser dos a dos independientes pero no independientes entre sí

Ejemplo 4.10 Sea el caso de lanzar un par de monedas corrientes; aquí $S = \{HH, HT, TH, TT\}$ es un espacio equiprobable. Consideremos los eventos

$$\begin{aligned} A &= \{\text{caras en la primera moneda}\} &&= \{HH, HT\} \\ B &= \{\text{caras en la segunda moneda}\} &&= \{HH, TH\} \\ C &= \{\text{caras en una moneda exactamente}\} &&= \{HT, TH\} \end{aligned}$$

Entonces $P(A) = P(B) = P(C) = \frac{2}{4} = \frac{1}{2}$ y

$$P(A\cap B) \;=\; P(\{HH\}) \;=\; \frac{1}{4}, \quad P(A\cap C) \;=\; P(\{HT\}) \;=\; \frac{1}{4}, \quad P(B\cap C) \;=\; (\{TH\}) \;=\; \frac{1}{4}$$

Así la condición (i) se satisface, o sea, los eventos son independientes dos a dos. Sin embargo, $A \cap B \cap C = \emptyset$ y así

$$P(A\cap B\cap C) \;=\; P(\emptyset) \;=\; 0 \;\neq\; P(A)\,P(B)\,P(C)$$

En otras palabras, la condición (ii) no se satisface y por tanto los tres eventos no son independientes

PRUEBAS REPETIDAS O INDEPENDIENTES

Hemos discutido previamente espacios de probabilidad que estaban relacionados con un experimento repetido un número finito de veces, tal como el lanzamiento de una moneda tres veces. Este concepto de repetición se formaliza como sigue

Definición: Sea S un espacio finito de probabilidad. Por *n pruebas repetidas o independientes*, significamos el espacio de probabilidad T que consta de n-uplas o elementos de S con la probabilidad de una n-upla definida como el producto de las probabilidades de sus componentes

$$P((s_1, s_2, \ldots, s_n)) \;=\; P(s_1)\,P(s_2) \cdots P(s_n)$$

Ejemplo 4.11 Tres caballos a, b y c corren juntos sus probabilidades de ganar son respectivamente $\frac{1}{2}$, $\frac{1}{3}$ y $\frac{1}{6}$. En otras palabras $S = \{a, b, c\}$ con $P(a) = \frac{1}{2}$, $P(b) = \frac{1}{3}$ y $P(c) = \frac{1}{6}$. Si los caballos corren dos veces entonces el espacio muestral de las dos pruebas repetidas es

$$T = \{aa, ab, ac, ba, bb, bc, ca, cb, cc\}$$

Por conveniencia en la notacion escribimos ac en lugar de la pareja ordenada (a, c). La probabilidad de cada punto de T es

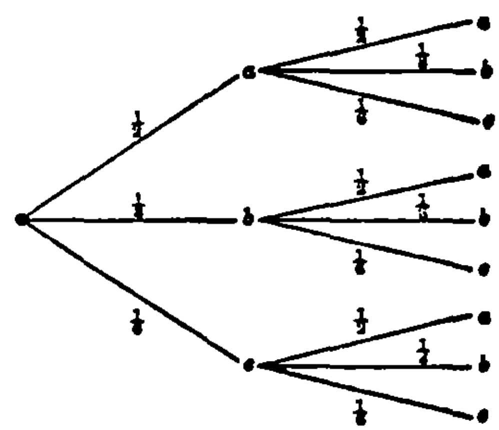
$$P(aa) = P(a)\,P(a) = \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4} \qquad P(ba) = \frac{1}{6} \qquad P(ca) = \frac{1}{12}$$

$$P(ab) = P(a)\,P(b) = \frac{1}{2}\cdot\frac{1}{3} = \frac{1}{6} \qquad P(bb) = \frac{1}{9} \qquad P(cb) = \frac{1}{18}$$

$$P(ac) = P(a)\,P(c) = \frac{1}{2}\cdot\frac{1}{6} = \frac{1}{12} \qquad P(bc) = \frac{1}{18} \qquad P(cc) = \frac{1}{36}$$

Así la probabilidad de que c gane la primera carrera y de que a gane la segunda es $P(ca) = \frac{1}{12}$.

Desde otro punto de vista un proceso de pruebas repetidas es un proceso estocástico, cuyo diagrama de árbol tiene las siguientes propiedades (i) cada ramal tiene los mismos resultados, (ii) la probabilidad es la misma en cada rama que conduce a un mismo final. Por ejemplo, el diagrama de árbol del proceso de pruebas repetidas del experimento anterior es tal como se muestra en la figura adjunta.



Obsérvese que cada ramal tiene los resultados a, b y c, y cada rama que conduce al resultado a tiene probabilidad $\frac{1}{2}$, las que conducen a b tienen probabilidad $\frac{1}{3}$ y cada una de las que conducen a c tienen probabilidad $\frac{1}{6}$.

Problemas resueltos

PROBABILIDAD CONDICIONAL EN ESPACIOS FINITOS EQUIPROBABLES

4.1. Se lanza un par de dados corrientes. Hallar la probabilidad p de que la suma de sus numeros sea 10 o mayor si, (i) aparece un 5 en el primer dado, (ii) aparece un 5 en uno de los dados por lo menos.

(i) Si aparece un 5 en el primer dado, entonces el espacio muestral reducido es

$$A = \{(5,1), (5,2), (5,3), (5,4), (5,5), (5,6)\}$$

La suma es 10 o mayor en dos de los seis resultados (5, 5), (5, 6). Por lo tanto $p = \frac{2}{6} = \frac{1}{3}$.

(ii) Si aparece un 5 por lo menos en un dado, entonces el espacio muestral reducido tiene once elementos

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

La suma es 10 o mayor en tres de los once resultados (5, 5), (5, 6), (6, 5). Por lo tanto $p = \frac{3}{11}$.

4.2. Se lanzan tres monedas corrientes. Hallar la probabilidad p de que sean todas caras si, (i) la primera de las monedas es cara, (ii) una de las monedas es cara.

El espacio muestral tiene ocho elementos: S = {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}.

(i) Si la primera moneda es cara, el espacio muestral reducido es A = {HHH, HHT, HTH, HTT}. Puesto que las monedas son todas caras en 1 de 4 casos, $p = \frac{1}{4}$.

(ii) Si una de las monedas es cara, el espacio muestral reducido es B = {HHH, HHT, HTH, HTT, THH, THT, TTH}. Puesto que las monedas son todas caras en 1 de 7 casos, $p = \frac{1}{7}$.

4.3. Se lanza un par de dados corrientes. Si los dos numeros que aparecen son diferentes, hallar la probabilidad p de que, (i) la suma sea seis, (ii) aparezca un as, (iii) la suma sea menor o igual a 4.

De las 36 maneras que se puede lanzar el par de dados, 6 contienen numeros repetidos: (1, 1), (2, 2), ..., (6, 6). Así el espacio muestral reducido constara de $36 - 6 = 30$ elementos.

(i) La suma 6 puede suceder de 4 maneras: (1, 5), (2, 4), (4, 2), (5, 1). (No incluimos (3, 3) puesto que los numeros son iguales.) Entonces $p = \frac{4}{30} = \frac{2}{15}$.

(ii) Un as puede aparecer de 10 maneras: (1, 2), (1, 3), ..., (1, 6) y (2, 1), (3, 1), ..., (6, 1). Entonces $p = \frac{10}{30} = \frac{1}{3}$.

(iii) La suma menor o igual a 4 puede suceder de 4 maneras: (3, 1), (1, 3), (2, 1), (1, 2). Así $p = \frac{4}{30} = \frac{2}{15}$.

4.4. Se escogen al azar dos digitos desde 1 hasta 9. Si la suma es par, hallar la probabilidad p de que ambos numeros sean impares.

La suma es par si los numeros son impares o si son pares. Hay 4 pares (2, 4, 6, 8), por tanto hay $\binom{4}{2} = 6$ maneras de escoger dos numeros pares. Hay 5 impares (1, 3, 5, 7, 9), o sea que hay $\binom{5}{2} = 10$ maneras de escoger dos numeros impares. Así hay $6 + 10 = 16$ maneras de escoger dos números tales que su suma sea par; puesto que 10 de estas maneras suceden cuando los dos numeros son impares, $p = \frac{10}{16} = \frac{5}{8}$.

4.5. A un hombre se reparten 4 espadas de una baraja corriente de 52 cartas. Si se le dan tres cartas mas, hallar la probabilidad p de que por lo menos una de las cartas adicionales sea tambien espada.

Puesto que recibió 4 espadas, quedan $52 - 4 = 48$ cartas de las cuales $13 - 4 = 9$ son espadas. Hay $\binom{48}{3} =$ 17 296 maneras en las que puede recibir tres cartas más. Puesto que hay $48 - 9 = 39$ cartas que no son espadas, hay $\binom{39}{3} = 9139$ maneras en que puede recibir tres cartas que no son espadas. Así la probabilidad q de que no reciba espadas es $q = \frac{9139}{17\,296}$; por lo tanto $p = 1 - q = \frac{8157}{17\,296}$.

4.6. Se reparten 13 cartas de una baraja corriente de 52 cartas a cuatro personas que denominamos Norte, Sur, Este y Oeste.

(i) Si S no tiene ases, hallar la probabilidad p de que su compañero N tenga exactamente dos ases.

(ii) Si N y S juntos tienen nueve corazones, hallar la probabilidad p de que E y O tengan cada uno dos corazones.

(i) Hay 39 cartas, contando los 4 ases, repartidas entre N, E y O. Hay $\binom{39}{13}$ maneras de que N reciba 13 de las 39 cartas. Hay $\binom{4}{2}$ maneras de que pueda recibir 2 de los cuatro ases, y $\binom{35}{11}$ maneras de que pueda recibir 11 cartas de las $39 - 4 = 35$ cartas que no son ases. Así

$$p = \frac{\binom{4}{2}\binom{35}{11}}{\binom{39}{13}} = \frac{6 \cdot 12 \cdot 13 \cdot 25 \cdot 26}{36 \cdot 37 \cdot 38 \cdot 39} = \frac{650}{2109}$$

(ii) Hay 26 cartas incluyendo 4 corazones repartidos entre E y O. Hay $\binom{26}{13}$ maneras de que, por ejemplo, E pueda recibir 13 cartas. (Necesitamos solamente analizar las 13 cartas de E puesto que O debe tener el resto.) Hay $\binom{4}{2}$ maneras para que E pueda recibir 2 corazones de los 4 y $\binom{22}{11}$ maneras para que el mismo pueda recibir 11 no-corazones de $26 - 4 = 22$ no-corazones. Así

$$p = \frac{\binom{4}{2}\binom{22}{11}}{\binom{26}{13}} = \frac{6\cdot12\cdot13\cdot12\cdot13}{23\cdot24\cdot25\cdot26} = \frac{234}{575}$$

TEOREMA DE LA MULTIPLICACION

4.7. Una clase tiene 12 niños y 4 niñas. Si se escogen tres estudiantes de la clase al azar, ¿cuál es la probabilidad p de que sean todos niños?

La probabilidad de que el primer estudiante escogido sea un niño es $12/16$ puesto que hay 12 niños entre los 16 estudiantes. Si el primero es un niño, entonces la probabilidad de que el segundo sea niño es $11/15$ puesto que hay 11 niños entre los 15 restantes. Finalmente, si los primeros dos escogidos son niños, entonces la probabilidad de que el tercero sea niño es $10/14$ puesto que quedan 10 niños entre 14. Así, por el teorema de la multiplicación, la probabilidad de que todos tres sean niños es

$$p = \frac{12}{16}\cdot\frac{11}{15}\cdot\frac{10}{14} = \frac{11}{28}$$

Otro método. Hay $\binom{16}{3} = 560$ maneras de escoger 3 estudiantes entre 16, y $\binom{12}{3} = 220$ maneras de escoger 3 niños entre 12, por lo tanto $p = \frac{220}{560} = \frac{11}{28}$.

Un tercer método. Si los estudiantes se escogen uno después del otro, entonces hay $16\cdot15\cdot14$ maneras de escoger tres estudiantes, y $12\cdot11\cdot10$ maneras de escoger tres niños, por consiguiente $p = \frac{12\cdot11\cdot10}{16\cdot15\cdot14} = \frac{11}{28}$.

4.8. A un jugador le reparten 5 cartas una tras otra de una baraja corriente de 52 cartas. ¿Cuál es la probabilidad p de que todas sean espadas?

La probabilidad que la primera carta sea espada es $13/52$, la segunda sea espada es $12/51$, la tercera $11/50$, la cuarta $10/49$, y la última $9/48$. (Suponemos en cada caso que las cartas anteriores fueron espadas.) Así

$$p = \frac{13}{52}\cdot\frac{12}{51}\cdot\frac{11}{50}\cdot\frac{10}{49}\cdot\frac{9}{48} = \frac{33}{66.640}$$

4.9. Una urna contiene 7 bolas rojas y 3 bolas blancas. Se sacan 3 bolas de la urna una tras otra. Hallar la probabilidad p de que las dos primeras sean rojas y la tercera blanca.

La probabilidad de que la primera bola sea roja es $7/10$ puesto que hay 7 rojas entre las 10 bolas. Si la primera bola es roja, entonces la probabilidad de que la segunda bola sea roja es $6/9$ puesto que quedan 6 rojas entre las 9 bolas restantes. Si las dos primeras son rojas, entonces la probabilidad de que la tercera sea blanca es $3/8$ puesto que quedan 3 blancas entre las 8 bolas restantes en la urna. Entonces por el teorema de la multiplicación,

$$p = \frac{7}{10}\cdot\frac{6}{9}\cdot\frac{3}{8} = \frac{7}{40}$$

4.10. Los estudiantes de una clase se escogen al azar, uno tras otro, para presentar un examen. Hallar la probabilidad p de que niños y niñas queden alternados si (i) la clase consta de 4 niños y 3 niñas, (ii) la clase consta de 3 niños y 3 niñas.

(i) Si los niños y las niñas se alternan, el primer estudiante examinado debe ser niño. La probabilidad de que el segundo sea niña es $3/6$ puesto que hay 3 niñas entre los 6 restantes. Continuando en esta forma, obtenemos que la probabilidad de que el tercero sea niño es $3/5$, que el cuarto sea niña es $2/4$, que el quinto sea niño es $2/3$, que el sexto sea niña es $1/2$, y que el último sea niño es $1/1$. Así

$$p = \frac{4}{7}\cdot\frac{3}{6}\cdot\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot\frac{1}{1} = \frac{1}{35}$$

(ii) Hay dos casos mutuamente exclusivos: el primer estudiante es un niño y el primero es una niña. Si el primer estudiante es un niño, entonces por el teorema de la multiplicación la probabilidad p_1 de que los estudiantes se alternen es

$$p_1 = \frac{3}{6}\cdot\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot\frac{1}{1} = \frac{1}{20}$$

Si el primer estudiante es una niña, entonces por el teorema de la multiplicación la probabilidad p_2 de que los estudiantes se alternen es

$$p_2 = \frac{3}{6}\cdot\frac{3}{5}\cdot\frac{2}{4}\cdot\frac{2}{3}\cdot\frac{1}{2}\cdot\frac{1}{1} = \frac{1}{20}$$

Así $p = p_1 + p_2 = \frac{1}{20} + \frac{1}{20} = \frac{1}{10}$.

PROBLEMAS VARIOS SOBRE PROBABILIDAD CONDICIONAL

4.11. En cierta facultad 25% de los estudiantes perdieron matemáticas, 15% perdieron química y 10% perdieron las dos. Se selecciona un estudiante al azar.

(i) Si perdió química, ¿cuál es la probabilidad de que perdió matemáticas?

(ii) Si perdió matemáticas, ¿cuál es la probabilidad de que perdió química?

(iii) ¿Cuál es la probabilidad de que perdió matemáticas o química?

Sea M = {estudiantes que perdieron matemáticas} y C = {estudiantes que perdieron química}, entonces

$$P(M) = 0.25 \qquad P(C) = 0.15 \qquad P(M\cap C) = 0.10$$

(i) La probabilidad de que el estudiante perdiera matemáticas, dado que haya perdido química es

$$P(M\mid C) = \frac{P(M\cap C)}{P(C)} = \frac{0.10}{0.15} = \frac{2}{3}$$

(ii) La probabilidad de que el estudiante perdiera química, dado que haya perdido matemáticas es

$$P(C\mid M) = \frac{P(C\cap M)}{P(M)} = \frac{0.10}{0.25} = \frac{2}{5}$$

(iii)
$$P(M\cup C) = P(M) + P(C) - P(M\cap C) = 0.25 + 0.15 - 0.10 = 0.30 = \frac{3}{10}$$

4.12. Sean los eventos A y B con $P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$ y $P(A\cap B) = \frac{1}{4}$. Hallar,
(i) $P(A\mid B)$, (ii) $P(B\mid A)$, (iii) $P(A\cup B)$, (iv) $P(A^c\mid B^c)$, (v) $P(B^c\mid A^c)$.

(i) $P(A\mid B) = \dfrac{P(A\cap B)}{P(B)} = \dfrac{\frac{1}{4}}{\frac{1}{3}} = \dfrac{3}{4}$ (ii) $P(B\mid A) = \dfrac{P(B\cap A)}{P(A)} = \dfrac{\frac{1}{4}}{\frac{1}{2}} = \dfrac{1}{2}$

(iii) $P(A\cup B) = P(A) + P(B) - P(A\cap B) = \frac{1}{2} + \frac{1}{3} - \frac{1}{4} = \frac{7}{12}$

(iv) Primero calculamos $P(B^c)$ y $P(A^c\cap B^c)$. $P(B^c) = 1 - P(B) = 1 - \frac{1}{3} = \frac{2}{3}$. Por la ley de De Morgan $(A\cup B)^c = A^c\cap B^c$, por lo tanto $P(A^c\cap B^c) = P((A\cup B)^c) = 1 - P(A\cup B) = 1 - \frac{7}{12} = \frac{5}{12}$.

Así $P(A^c\mid B^c) = \dfrac{P(A^c\cap B^c)}{P(B^c)} = \dfrac{\frac{5}{12}}{\frac{2}{3}} = \dfrac{5}{8}$

(v) $P(A^c) = 1 - P(A) = 1 - \frac{1}{2} = \frac{1}{2}$. Luego $P(B^c\mid A^c) = \dfrac{P(B^c\cap A^c)}{P(A^c)} = \dfrac{\frac{5}{12}}{\frac{1}{2}} = \dfrac{5}{6}$.

4.13. Sean los eventos A y B con $P(A) = \frac{3}{8}$, $P(B) = \frac{5}{8}$ y $P(A\cup B) = \frac{3}{4}$. Hallar $P(A\mid B)$ y $P(B\mid A)$.

Primero calculemos $P(A\cap B)$ usando la fórmula $P(A\cup B) = P(A) + P(B) - P(A\cap B)$.

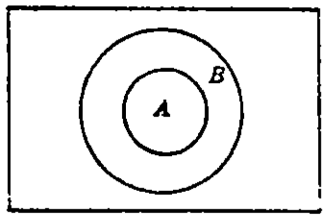
$$\frac{3}{4} = \frac{3}{8} + \frac{5}{8} - P(A\cap B) \quad \text{o} \quad P(A\cap B) = \frac{1}{4}$$

Así $P(A\mid B) = \dfrac{P(A\cap B)}{P(B)} = \dfrac{\frac{1}{4}}{\frac{5}{8}} = \dfrac{2}{5}$ y $P(B\mid A) = \dfrac{P(B\cap A)}{P(A)} = \dfrac{\frac{1}{4}}{\frac{3}{8}} = \dfrac{2}{3}$

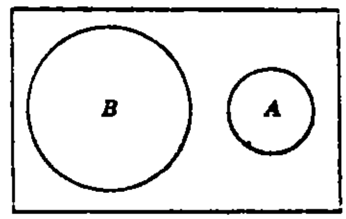
4.14. Hallar $P(B \mid A)$ si, (i) A es un subconjunto de B (ii) A y B son mutuamente exclusivos

(i) Si A es un subconjunto de B entonces siempre que A suceda B debe suceder, por lo tanto $P(B \mid A) = 1$. A su turno, si A es un subconjunto de B entonces $A \cap B = A$, entonces

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A)}{P(A)} = 1$$



(i)



(ii)

(ii) Si A y B son mutuamente exclusivos, es decir, disyuntos, entonces siempre que A suceda B no puede suceder, por lo tanto $P(B \mid A) = 0$. Alternadamente, si A y B son mutuamente exclusivos entonces $A \cap B = \emptyset$, por lo tanto

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{P(\emptyset)}{P(A)} = \frac{0}{P(A)} = 0$$

4.15. Tres maquinas A, B y C producen respectivamente 60%, 30% y 10% del numero total de artículos de una fabrica. Los porcentajes de desperfectos de producción de estas maquinas son respectivamente 2%, 3% y 4%. Seleccionado un articulo al azar resulto defectuoso. Hallar la probabilidad de que el articulo hubiera sido producido por la maquina C.

Sea $X = \{$articulos defectuosos$\}$. Buscamos $P(C \mid X)$, probabilidad de que un articulo sea producido por la máquina C dado que el articulo sea defectuoso. Por el teorema de Bayes,

$$P(C \mid X) = \frac{P(C)\,P(X \mid C)}{P(A)\,P(X \mid A) + P(B)\,P(X \mid B) + P(C)\,P(X \mid C)}$$

$$= \frac{(0.10)(0.04)}{(0.60)(0.02) + (0.30)(0.03) + (0.10)(0.04)} = \frac{4}{25}$$

4.16. En cierta facultad, 4% de los hombres y 1% de las mujeres tienen más de 6 pies de estatura. Además, 60% de los estudiantes son mujeres. Ahora bien si se selecciona al azar un estudiante y es más alto que 6 pies, ¿cuál es la probabilidad que el estudiante sea mujer?

Sea $A = \{$estudiantes de más de 6 pies$\}$. Buscamos $P(W \mid A)$, probabilidad de que el estudiante sea una mujer dado que el estudiante es de más de 6 pies. Por el teorema de Bayes,

$$P(W \mid A) = \frac{P(W)\,P(A \mid W)}{P(W)\,P(A \mid W) + P(M)\,P(A \mid M)} = \frac{(0.60)(0.01)}{(0.60)(0.01) + (0.40)(0.04)} = \frac{3}{11}$$

4.17. Sea E un evento para el cual $P(E) > 0$. Comprobar que la funcion de probabilidad condicional $P(* \mid E)$ satisface los axiomas de un espacio de probabilidad, esto es,

[P_1] Para un evento A, $0 \leq P(A \mid E) \leq 1$.

[P_2] Para el evento cierto S, $P(S \mid E) = 1$.

[P_3] Si A y B son mutuamente exclusivos, entonces $P(A \cup B \mid E) = P(A \mid E) + P(B \mid E)$.

[P_4] Si $A_1, A_2, \ldots$ es una sucesión de eventos mutuamente exclusivos, entonces

$$P(A_1 \cup A_2 \cup \cdots \mid E) = P(A_1 \mid E) + P(A_2 \mid E) + \cdots$$

(i) Tenemos $A \cap E \subset E$, por lo tanto $P(A \cap E) \leq P(E)$. Así $P(A \mid E) = \dfrac{P(A \cap E)}{P(E)} \leq 1$ y es no negativo también. Esto es $0 \leq P(A \mid E) \leq 1$ y así [P_1] cumple.

(ii) Tenemos $S \cap E = E$ por consiguiente $P(S \mid E) = \dfrac{P(S\cap E)}{P(E)} = \dfrac{P(E)}{P(E)} = 1$. Así $[\mathbf{P_2}]$ satisface.

(iii) Si A y B son eventos mutuamente exclusivos entonces así son $A \cap E$ y $B \cap E$. Además $(A\cup B)\cap E = (A\cap E)\cup(B\cap E)$. Así

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

y por ende

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

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

O sea que $[\mathbf{P_3}]$ satisface.

(iv) Similarmente si $A_1, A_2, \ldots$ son mutuamente exclusivos también lo son $A_1\cap E, A_2\cap E, \ldots$. Así

$$P((A_1\cup A_2\cup\cdots)\cap E) = P((A_1\cap E)\cup(A_2\cap E)\cup\cdots) = P(A_1\cap E) + P(A_2\cap E) + \cdots$$

y por tanto

$$P(A_1\cup A_2\cup\cdots \mid E) = \frac{P((A_1\cup A_2\cup\cdots)\cap E)}{P(E)} = \frac{P(A_1\cap E) + P(A_2\cap E) + \cdots}{P(E)}$$

$$= \frac{P(A_1\cap E)}{P(E)} + \frac{P(A_2\cap E)}{P(E)} + \cdots = P(A_1 \mid E) + P(A_2 \mid E) + \cdots$$

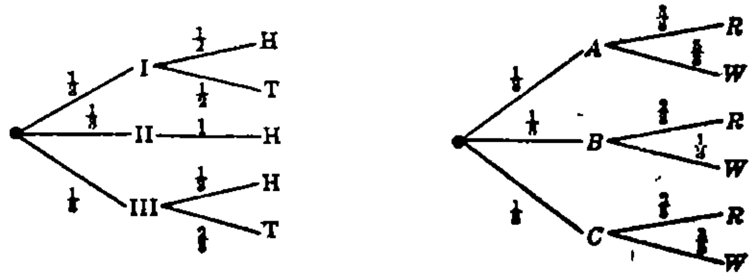
Esto es, $[\mathbf{P_4}]$ satisface.

PROCESOS ESTOCASTICOS FINITOS

4.18. Una caja contiene tres monedas; una moneda es corriente, una moneda tiene dos caras y una moneda está cargada de modo que la probabilidad de obtener cara sea $\frac{1}{3}$. Se selecciona una moneda al azar y se lanza. Hallar la probabilidad p de que salga cara.

Construimos el diagrama de árbol como se muestra en la figura (*a*) siguiente. Obsérvese que I se refiere a la moneda corriente, II a la de doble cara y III a la moneda cargada. Ahora las caras aparecen a lo largo de tres de las trayectorias, por lo tanto

$$p = \frac{1}{3}\cdot\frac{1}{2} + \frac{1}{3}\cdot 1 + \frac{1}{3}\cdot\frac{1}{3} = \frac{11}{18}$$



(*a*) (*b*)

4.19. Se nos dan tres urnas como sigue:

Una urna A contiene 3 bolas rojas y 5 blancas.
Una urna B contiene 2 bolas rojas y 1 blanca.
Una urna C contiene 2 bolas rojas y 3 blancas.

Se selecciona una urna al azar y se saca una bola de la urna. Si la bola es roja, ¿cuál es la probabilidad de que proceda de la urna A?

Se construye el diagrama de árbol como se muestra en la figura (*b*) anterior.

Buscamos la probabilidad de que se seleccione A dado que la bola es roja, esto es, $P(A \mid R)$. Con el fin de hallar $P(A \mid R)$, es necesario calcular primero $P(A \cap R)$ y $P(R)$.

La probabilidad de que se seleccione A y se saque una bola roja es $\frac{1}{3}\cdot\frac{3}{8} = \frac{1}{8}$, esto es $P(A\cap R) = \frac{1}{8}$. Puesto que hay tres trayectorias que conducen a bola roja $P(R) = \frac{1}{3}\cdot\frac{3}{8} + \frac{1}{3}\cdot\frac{2}{3} + \frac{1}{3}\cdot\frac{2}{5} = \frac{173}{360}$. Entonces

$$P(A \mid R) = \frac{P(A \cap R)}{P(R)} = \frac{\frac{1}{8}}{\frac{173}{360}} = \frac{45}{173}$$

Alternadamente por el teorema de Bayes

$$P(A \mid R) = \frac{P(A)\,P(R \mid A)}{P(A)\,P(R \mid A) + P(B)\,P(R \mid B) + P(C)\,P(R \mid C)}$$

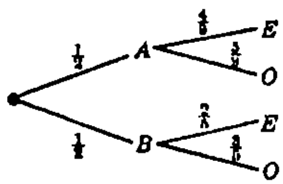
$$= \frac{\frac{1}{3}\cdot\frac{3}{8}}{\frac{1}{3}\cdot\frac{3}{8} + \frac{1}{3}\cdot\frac{2}{5} + \frac{1}{3}\cdot\frac{2}{3}} = \frac{45}{173}$$

4.20. La caja A contiene nueve cartas numeradas de 1 a 9, y la caja B contiene cinco cartas numeradas de 1 a 5. Se escoge una caja al azar y se saca una carta. Si el numero es par, hallar la probabilidad de que la carta proceda de la caja A.

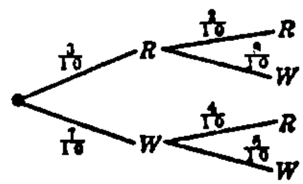
El diagrama de árbol del proceso se muestra en la figura (*a*) siguiente.

Buscamos $P(A \mid E)$, probabilidad de que se seleccione A dado que el numero es par. La probabilidad de que se escoja la caja A y un numero par es $\frac{1}{2}\cdot\frac{4}{9} = \frac{2}{9}$, esto es $P(A \cap E) = \frac{2}{9}$. Puesto que hay dos trayectorias que conducen a un numero par $P(E) = \frac{1}{2}\cdot\frac{4}{9} + \frac{1}{2}\cdot\frac{2}{5} = \frac{19}{45}$. Así

$$P(A \mid E) = \frac{P(A \cap E)}{P(E)} = \frac{\frac{2}{9}}{\frac{19}{45}} = \frac{10}{19}$$



(*a*)



(*b*)

4.21. Una urna contiene 3 bolas rojas y 7 blancas. Se saca una bola de la urna y se remplaza por una del otro color. Se saca de la urna una segunda bola.

(i) Hallar la probabilidad p de que la segunda bola sea roja.

(ii) Si ambas bolas son del mismo color, ¿cuál es la probabilidad p de que las dos sean blancas?

Construimos el diagrama de árbol como se indica en la figura (*b*) anterior.

(i) Dos trayectorias del diagrama de árbol conducen a bola roja: $p = \frac{3}{10}\cdot\frac{2}{10} + \frac{7}{10}\cdot\frac{4}{10} = \frac{17}{50}$.

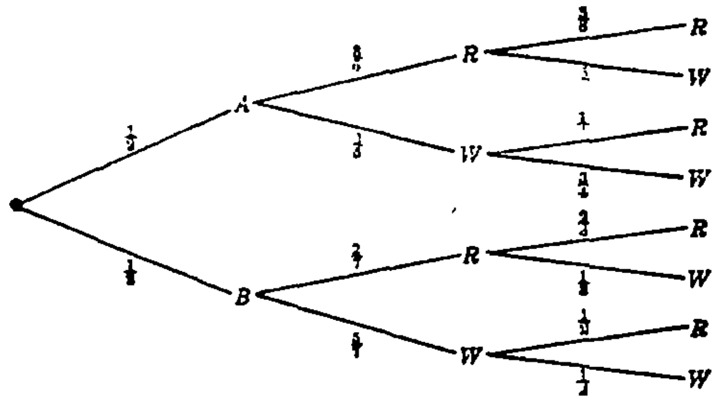
(ii) La probabilidad de que ambas bolas fueran blancas es $\frac{7}{10}\cdot\frac{6}{10} = \frac{21}{50}$. La probabilidad de que ambas bolas fueran del mismo color, esto es la probabilidad del espacio muestral reducido, es $\frac{3}{10}\cdot\frac{2}{10} + \frac{7}{10}\cdot\frac{6}{10} = \frac{12}{25}$. Por lo tanto la probabilidad condicional $p = \frac{21}{50}/\frac{12}{25} = \frac{7}{8}$.

4.22. Se nos dan dos urnas como sigue:

La urna A contiene 3 bolas rojas y 2 blancas.
La urna B contiene 2 bolas rojas y 5 blancas.

Se selecciona al azar una urna, se saca una bola y se coloca en la otra urna, luego se saca una bola de la segunda urna. Hallar la probabilidad p de que las dos bolas sacadas sean del mismo color.

Construimos el siguiente diagrama de árbol.



Nótese que si se selecciona la urna A y se saca una bola roja y se coloca en la urna B, entonces la urna B tiene 3 bolas rojas y 5 blancas.

Puesto que hay cuatro trayectorias que conducen a dos bolas del mismo color

$$p \;=\; \frac{1}{2}\cdot\frac{3}{5}\cdot\frac{3}{8}+\frac{1}{2}\cdot\frac{2}{5}\cdot\frac{3}{4}+\frac{1}{2}\cdot\frac{2}{7}\cdot\frac{2}{3}+\frac{1}{2}\cdot\frac{5}{7}\cdot\frac{1}{2} \;=\; \frac{901}{1680}$$

INDEPENDENCIA

4.23. Sea A = al evento de que una familia tenga niños de ambos sexos y sea B = al evento de que una familia tenga a lo sumo un niño. (i) Comprobar que A y B son eventos independientes si una familia tiene tres hijos. (ii) Comprobar que A y B son eventos dependientes si una familia tiene dos hijos.

(i) Tenemos el espacio equiprobable $S = \{bbb, bbg, bgb, bgg, gbb, gbg, ggb, ggg\}$. Aquí

$$A = \{bbg, bgb, bgg, gbb, gbg, ggb\} \quad \text{y así} \quad P(A) = \tfrac{6}{8} = \tfrac{3}{4}$$

$$B = \{bgg, gbg, ggb, ggg\} \quad \text{y así} \quad P(B) = \tfrac{4}{8} = \tfrac{1}{2}$$

$$A\cap B = \{bgg, gbg, ggb\} \quad \text{y así} \quad P(A\cap B) = \tfrac{3}{8}$$

Puesto que $P(A)\,P(B) = \frac{3}{4}\cdot\frac{1}{2} = \frac{3}{8} = P(A\cap B)$, A y B son independientes.

(ii) Tenemos el espacio equiprobable $S = \{bb, bg, gb, gg\}$. Aquí

$$A = \{bg, gb\} \quad \text{y así} \quad P(A) = \tfrac{1}{2}$$

$$B = \{bg, gb, gg\} \quad \text{y así} \quad P(B) = \tfrac{3}{4}$$

$$A\cap B = \{bg, gb\} \quad \text{y así} \quad P(A\cap B) = \tfrac{1}{2}$$

Puesto que $P(A)\,P(B) \neq P(A\cap B)$, A y B son dependientes.

4.24. Probar si A y B son eventos independientes, entonces A^c y B^c son eventos independientes.

$$\begin{aligned} P(A^c\cap B^c) &= P((A\cup B)^c) = 1-P(A\cup B) = 1-P(A)-P(B)+P(A\cap B) \\ &= 1-P(A)-P(B)+P(A)\,P(B) = [1-P(A)][1-P(B)] = P(A^c)\,P(B^c) \end{aligned}$$

4.25. La probabilidad de que un hombre vivirá 10 años más es $\frac{1}{4}$, y la probabilidad de que su esposa vivirá 10 años más es $\frac{1}{3}$. Hallar la probabilidad de que, (i) ambos estén vivos dentro de 10 años, (ii) al menos uno estará vivo a los 10 años, (iii) ninguno estará vivo a los 10 años, (iv) solamente la esposa estará viva a los 10 años.

Sea A = al evento de que el hombre viva a los 10 años, y B = al evento de que su esposa esté viva a los 10 años; entonces $P(A) = \frac{1}{4}$ y $P(B) = \frac{1}{3}$.

(i) Buscamos $P(A\cap B)$. Puesto que A y B son independientes, $P(A\cap B) = P(A)\,P(B) = \frac{1}{4}\cdot\frac{1}{3} = \frac{1}{12}$.

(ii) Buscamos $P(A\cup B)$. $P(A\cup B) = P(A) + P(B) - P(A\cap B) = \frac{1}{4} + \frac{1}{3} - \frac{1}{12} = \frac{1}{2}$.

(iii) Buscamos $P(A^c\cap B^c)$. Ahora $P(A^c) = 1 - P(A) = 1 - \frac{1}{4} = \frac{3}{4}$ y $P(B^c) = 1 - P(B) = 1 - \frac{1}{3} = \frac{2}{3}$. Además puesto que A^c y B^c son independientes, $P(A^c\cap B^c) = P(A^c)\,P(B^c) = \frac{3}{4}\cdot\frac{2}{3} = \frac{1}{2}$.

Alternadamente, puesto que

$$(A\cup B)^c = A^c\cap B^c,\quad P(A^c\cap B^c) = P((A\cup B)^c) = 1 - P(A\cup B) = 1 - \frac{1}{2} = \frac{1}{2}.$$

(iv) Buscamos $P(A^c\cap B)$. Puesto que $P(A^c) = 1 - P(A) = \frac{3}{4}$ y A^c y B son independientes (ver problema 4.56), $P(A^c\cap B) = P(A^c)\,P(B) = \frac{1}{4}$.

4.26. La caja A contiene 8 artículos de los cuales 3 son defectuosos, y la caja B contiene 5 artículos de los cuales 2 son defectuosos. Se saca al azar un artículo de cada caja.

(i) ¿Cuál es la probabilidad p de que ambos artículos sean defectuosos?

(ii) ¿Cuál es la probabilidad p de que un artículo sea defectuoso y otro no?

(iii) Si un artículo es defectuoso y otro no, ¿cuál es la probabilidad p de que el artículo defectuoso proceda de la caja A?

(i) La probabilidad de escoger un artículo no defectuoso de A es $\frac{5}{8}$ y de B es $\frac{3}{5}$. Puesto que los eventos son independientes, $p = \frac{5}{8}\cdot\frac{3}{5} = \frac{3}{8}$.

(ii) Método 1. La probabilidad de escoger dos artículos defectuosos es $\frac{3}{8}\cdot\frac{2}{5} = \frac{3}{20}$. De (i) la probabilidad de que ambos sean no defectuosos es $\frac{3}{8}$. Por lo tanto $p = 1 - \frac{3}{8} - \frac{3}{20} = \frac{19}{40}$.

Método 2. La probabilidad p_1 de escoger un artículo defectuoso de A y uno no defectuoso de B es $\frac{3}{8}\cdot\frac{3}{5} = \frac{9}{40}$. La probabilidad p_2 de escoger un artículo no defectuoso de A y uno defectuoso de B es $\frac{5}{8}\cdot\frac{2}{5} = \frac{1}{4}$. Por lo tanto $p = p_1 + p_2 = \frac{9}{40} + \frac{1}{4} = \frac{19}{40}$.

(iii) Consideremos los eventos $X = \{$artículos defectuosos de $A\}$ y $Y = \{$un artículo es defectuoso y otro no$\}$. Buscamos $P(X\mid Y)$. Por (ii), $P(X\cap Y) = p_1 = \frac{9}{40}$ y $P(Y) = \frac{19}{40}$. Por consiguiente

$$p = P(X\mid Y) = \frac{P(X\cap Y)}{P(Y)} = \frac{\frac{9}{40}}{\frac{19}{40}} = \frac{9}{19}$$

4.27. Las probabilidades de que tres hombres peguen en el blanco son, respectivamente, $\frac{1}{6}$, $\frac{1}{4}$ y $\frac{1}{3}$. Cada uno dispara una vez al blanco. (i) Hallar la probabilidad p de que exactamente uno de ellos pegue en el blanco. (ii) Si solamente uno pega en el blanco, ¿cuál es la probabilidad de que sea el primer hombre?

Consideremos los eventos $A = \{$el primer hombre pega en el blanco$\}$, $B = \{$el segundo hombre pega en el blanco$\}$ y $C = \{$el tercer hombre pega en el blanco$\}$; entonces $P(A) = \frac{1}{6}$, $P(B) = \frac{1}{4}$ y $P(C) = \frac{1}{3}$. Los tres eventos son independientes y $P(A^c) = \frac{5}{6}$, $P(B^c) = \frac{3}{4}$, $P(C^c) = \frac{2}{3}$.

(i) Sea $E = \{$exactamente un hombre pega en el blanco$\}$. Entonces

$$E = (A\cap B^c\cap C^c)\cup(A^c\cap B\cap C^c)\cup(A^c\cap B^c\cap C)$$

En otras palabras, si solamente uno pegó en el blanco, entonces fue o únicamente el primer hombre, $A\cap B^c\cap C^c$, o únicamente el segundo hombre, $A^c\cap B\cap C^c$, o únicamente el tercer hombre, $A^c\cap B^c\cap C$. Como los tres eventos son mutuamente exclusivos, obtenemos (usando el problema 4.62)

$$\begin{aligned} p = P(E) &= P(A\cap B^c\cap C^c) + P(A^c\cap B\cap C^c) + P(A^c\cap B^c\cap C) \\ &= P(A)\,P(B^c)\,P(C^c) + P(A^c)\,P(B)\,P(C^c) + P(A^c)\,P(B^c)\,P(C) \\ &= \frac{1}{6}\cdot\frac{3}{4}\cdot\frac{2}{3} + \frac{5}{6}\cdot\frac{1}{4}\cdot\frac{2}{3} + \frac{5}{6}\cdot\frac{3}{4}\cdot\frac{1}{3} = \frac{1}{12} + \frac{5}{36} + \frac{5}{24} = \frac{31}{72} \end{aligned}$$

(ii) Buscamos $P(A\mid E)$, la probabilidad de que el primer hombre pegue en el blanco dado que solamente un hombre pega en el blanco. Ahora $A\cap E = A\cap B^c\cap C^c$ es el evento en que solamente el primer hombre pega en el blanco.

Por (i) $P(A\cap E) = P(A\cap B^c\cap C^c) = \frac{1}{12}$ y $P(E) = \frac{31}{72}$, o sea

$$P(A\mid E) = \frac{P(A\cap E)}{P(E)} = \frac{\frac{1}{12}}{\frac{31}{72}} = \frac{6}{31}$$

PRUEBAS INDEPENDIENTES

4.28. Cierto tipo de proyectil da en el blanco con probabilidad 0,3. ¿Cuántos proyectiles deberán ser disparados para que haya al menos un 80% de probabilidad de pegar en el blanco?

La probabilidad de que un proyectil falle su blanco es 0,7; por lo tanto la probabilidad de que n proyectiles fallen el blanco es $(0,7)^n$. Así buscamos el menor n para el que

$$1-(0,7)^n > 0,8 \quad \text{o equivalentemente} \quad (0,7)^n < 0,2$$

Calculamos: $(0,7)^1 = 0,7$, $(0,7)^2 = 0,49$, $(0,7)^3 = 0,343$, $(0,7)^4 = 0,2401$, $(0,7)^5 = 0,16807$. Así por lo menos 5 proyectiles deben ser disparados.

4.29. Cierto equipo de balompié gana (W), con probabilidad 0,6, pierde (L), con probabilidad 0,3 y empata (T), con probabilidad 0,1. El equipo juega tres encuentros durante el fin de semana. (i) Determinar los elementos del evento A en que el equipo gana por lo menos dos y no pierde, y hallar $P(A)$. (ii) Determinar los elementos del evento B en que el equipo gana, pierde y empata, y hallar $P(B)$.

(i) A consta de todas las ternas con al menos dos W (juegos ganados) y ningún L (juego perdido). Así

$$A = \{WWW, WWT, WTW, TWW\}$$

Además

$$\begin{aligned} P(A) &= P(WWW) + P(WWT) + P(WTW) + P(TWW) \\ &= (0,6)(0,6)(0,6) + (0,6)(0,6)(0,1) + (0,6)(0,1)(0,6) + (0,1)(0,6)(0,6) \\ &= 0,216 + 0,036 + 0,036 + 0,036 = 0,324 \end{aligned}$$

(ii) Aquí $B = \{LWT, WTL, LWT, LTW, TWL, TLW\}$. Puesto que cada elemento de B tiene probabilidad $(0,6)(0,3)(0,1) = 0,018$, $P(B) = 6(0,018) = 0,108$.

4.30. Sea S un espacio finito de probabilidad y sea T el espacio de probabilidad de n pruebas independientes de S. Comprobar que T está bien definido, esto es, mostrar, (i) la probabilidad de que cada elemento de T es no-negativo, y (ii) la suma de las probabilidades es 1.

Si $S = \{a_1, \ldots, a_r\}$, entonces T puede representarse por

$$T = \{a_{i_1} \cdots a_{i_n} : i_1, \ldots, i_n = 1, \ldots, r\}$$

Puesto que $P(a_i) \geq 0$, tenemos

$$P(a_{i_1} \cdots a_{i_n}) = P(a_{i_1}) \cdots P(a_{i_n}) \geq 0$$

para un elemento típico $a_{i_1} \cdots a_{i_n}$ de T, lo cual prueba (i).

Probamos (ii) por inducción en n. Esto es cierto obviamente para $n = 1$. Por lo tanto consideremos $n > 1$ y aceptamos que (ii) ha sido probado para $n - 1$. Entonces

$$\begin{aligned} \sum_{i_1, \ldots, i_n = 1}^{r} P(a_{i_1} \cdots a_{i_n}) &= \sum_{i_1, \ldots, i_n = 1}^{r} P(a_{i_1}) \cdots P(a_{i_n}) = \sum_{i_1, \ldots, i_{n-1} = 1}^{r} P(a_{i_1}) \cdots P(a_{i_{n-1}}) \sum_{i_n = 1}^{r} P(a_{i_n}) \\ &= \sum_{i_1, \ldots, i_{n-1} = 1}^{r} P(a_{i_1}) \cdots P(a_{i_{n-1}}) = \sum_{i_1, \ldots, i_{n-1} = 1}^{r} P(a_{i_1} \cdots a_{i_{n-1}}) = 1 \end{aligned}$$

por la hipótesis inductiva, lo cual prueba (ii) para n.

4. Given that

$$f(x) = k(1/3)^x$$

is a probability distribution for a random variable assuming the values $x = 0, 1, 2, \ldots$, find k and also find an expression for the corresponding cumulative probabilities $F(x)$

5. Prove parts (a) and (b) of Theorem 3.1 on page 41

6. Prove parts (c) and (d) of Theorem 3.1 on page 41

7. Prove that

$$\frac{b(x+1, n, p)}{b(x; n, p)} = \frac{p(n-x)}{(1-p)(x+1)}$$

for $x = 0, 1, 2, \ldots, n-1$.

8. Use the recursion formula of Exercise 7 to calculate the probabilities of the binomial distribution with $n = 5$ and $p = 0.30$, and draw its histogram. Verify your results by referring to Table I.

9. Using Table I find
   (a) $B(5, 15, 0.40)$,
   (b) $b(5; 15, 0.40)$,
   (c) $B(8, 12, 0.70)$,
   (d) $b(8, 12, 0.70)$,
   (e) $\sum_{k=8}^{20} b(k; 20, 0.10)$,
   (f) $\sum_{k=3}^{9} b(k, 9, 0.80)$.

10. Using Table I find
   (a) $B(8, 18, 0.45)$,
   (b) $b(8, 18, 0.45)$,
   (c) $B(9, 10, 0.90)$,
   (d) $b(9, 10, 0.90)$,
   (e) $\sum_{k=3}^{10} b(k, 10, 0.30)$,
   (f) $\sum_{k=2}^{4} b(k, 8, 0.40)$.

11. The probability that a car driving the entire length of a certain turnpike will have a blowout is 0.05. Find the probability that among 17 cars traveling the length of this turnpike
   (a) exactly one will have a blowout,
   (b) at most 3 will have a blowout,
   (c) 2 or more will have a blowout.

12. The probability that a certain kind of vacuum tube will survive a given thermal shock test is 0.85. Find the probability that among 20 such tubes
   (a) exactly 17 will survive,
   (b) at least 15 will survive,
   (c) at least 3 will not survive

13. A manufacturer claims that at most 10 per cent of his product is defective. To test this claim, 18 units are inspected and his claim is accepted if among these 18 units at most 2 are defective. Find the probability that the manufacturer's claim will be accepted if the actual probability that a unit is defective is
   (a) 0.05,
   (b) 0.10,
   (c) 0.15,
   (d) 0.20

3.3 The Hypergeometric Distribution

Suppose a sample of n units is to be drawn from a lot containing N units, of which a are defective. The sample is to be drawn in such a way that at each successive drawing whatever units are left in the lot have the same chance of being included in the sample. Thus, for the first drawing the probability of obtaining a defective unit is a/N, but for the second drawing it remains a/N *only if the first unit drawn is replaced.* Otherwise, this probability is $\frac{a-1}{N-1}$ or $\frac{a}{N-1}$ depending on whether or not a defective unit was obtained in the first drawing. *Assumption 4 underlying the binomial distribution is, therefore, met only if the sample is drawn with replacement.* In practice, this is seldom the case, and we shall, thus, have to derive a new probability distribution, called the *hypergeometric distribution*, which applies to this kind of situation.

To solve this problem of "sampling without replacement," let us refer to the result obtained on page 40, namely, that a subset of x objects can be selected from a set of n objects in

$$\binom{n}{x} = \frac{n!}{x!(n-x)!}$$

ways. We proved this in connection with the problem of the distribution of x successes among n trials, that is, choosing subsets on which the x successes are to occur. We referred to $\binom{n}{x}$ as a binomial coefficient, and it is also called the "number of combinations of n objects taken x at a time."

Returning now to the problem of finding the probability of getting x defectives in a sample of n units taken without replacement, note first that the sample space for this experiment has $\binom{N}{n}$ possible outcomes, namely, the number of ways in which a subset of n objects can be selected from among a set of N objects. Furthermore, the x defectives can be selected from the a defectives in $\binom{a}{x}$ ways, the $n-x$ nondefective units in the sample can be selected from the $N-a$ nondefective units in the lot in

9 Using Table II, find

(a) $F(9, 15)$, (c) $\sum_{k=3}^{13} f(k, 8.5)$.

(b) $f(9, 15)$,

10 It was pointed out on page 49 that values of the binomial distribution can sometimes be approximated with corresponding values of the Poisson distribution. Verify this statement by comparing $b(4, 50, 0.10)$ with $f(4, 5)$.

11. In the inspection of tin plate produced by a continuous electrolytic process, the probability of spotting an imperfection in a very small interval of time Δt is $(0.2)\,\Delta t$, time being measured in minutes. What are the probabilities of spotting, respectively, 0, 1, and 2 imperfections in 5 minutes?

12. The number of weekly breakdowns of a computer is a random variable having a Poisson distribution with $\lambda = 0.4$. What is the probability that the computer will operate without a breakdown for two consecutive weeks?

13. The number of gamma rays emitted per second by a certain radioactive substance is a random variable having the Poisson distribution with $\lambda = 5.6$. If a recording instrument becomes inoperative when there are more than 12 rays per second, what is the probability that this instrument becomes inoperative during any given second?

14. The number of hurricanes reaching the East Coast of the United States per year is a random variable having the Poisson distribution with $\lambda = 1.9$. Find the probabilities that in a given year the number of hurricanes reaching the East Coast of the United States is (a) zero, (b) exactly 3, (c) 2 or more

3.5 The Mean and Variance of a Probability Distribution

In Section 3.2 we introduced one important characteristic of a probability distribution (or the corresponding histogram), namely, that of the symmetry or skewness of a distribution. Two other important characteristics are illustrated in Figure 3.7, which show the histograms of two binomial distributions, one with the parameters $n = 4$ and $p = 1/2$ (unshaded) and the other with the parameters $n = 16$ and $p = 1/2$ (shaded). Essentially, these two distributions differ in two respects. The first distribution is "centered" about $x = 2$ while the second is "centered" about $x = 8$, and we thus say that the two distributions differ in their *location*. Another distinction between the two distributions is that the histogram of the second is more "spread out," the probabilities represented by the rectangles are more "dispersed," and we say that the two distributions differ in *variation*. In this section we shall introduce two of the foremost measures describing the location and the variation of a probability distribution; they are called the *mean* and the *variance*.

Table I

BINOMIAL DISTRIBUTION FUNCTION

$$B(x, n, p) = \sum_{k=0}^{x} \binom{n}{k} p^k (1-p)^{n-k}$$

n	x	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
						p					
2	0	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600	0.3025	0.2500
	1	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8775	0.8400	0.7975	0.7500
3	0	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160	0.1664	0.1250
	1	0.9928	0.9720	0.9392	0.8960	0.8438	0.7840	0.7182	0.6480	0.5748	0.5000
	2	0.9999	0.9990	0.9966	0.9920	0.9844	0.9730	0.9571	0.9360	0.9089	0.8750
4	0	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296	0.0915	0.0625
	1	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.5630	0.4752	0.3910	0.3125
	2	0.9995	0.9963	0.9880	0.9728	0.9492	0.9163	0.8735	0.8208	0.7585	0.6875
	3	1.0000	0.9999	0.9995	0.9984	0.9961	0.9919	0.9850	0.9744	0.9590	0.9375
5	0	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778	0.0503	0.0313
	1	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.4284	0.3370	0.2562	0.1875
	2	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369	0.7648	0.6826	0.5931	0.5000
	3	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692	0.9460	0.9130	0.8688	0.8125
	4	1.0000	1.0000	0.9999	0.9997	0.9990	0.9976	0.9947	0.9898	0.9815	0.9688
6	0	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0754	0.0467	0.0277	0.0156
	1	0.9672	0.8857	0.7765	0.6554	0.5339	0.4202	0.3191	0.2333	0.1636	0.1094
	2	0.9978	0.9842	0.9527	0.9011	0.8306	0.7443	0.6471	0.5443	0.4415	0.3438
	3	0.9999	0.9987	0.9941	0.9830	0.9624	0.9295	0.8826	0.8208	0.7447	0.6563
	4	1.0000	0.9999	0.9996	0.9984	0.9954	0.9891	0.9777	0.9590	0.9308	0.8906
	5	1.0000	1.0000	1.0000	0.9999	0.9998	0.9993	0.9982	0.9959	0.9917	0.9844
7	0	0.6983	0.4783	0.3206	0.2097	0.1335	0.0824	0.0490	0.0280	0.0152	0.0078
	1	0.9556	0.8503	0.7166	0.5767	0.4449	0.3294	0.2338	0.1586	0.1024	0.0625
	2	0.9962	0.9743	0.9262	0.8520	0.7564	0.6471	0.5323	0.4199	0.3164	0.2266
	3	0.9998	0.9973	0.9879	0.9667	0.9294	0.8740	0.8002	0.7102	0.6083	0.5000
	4	1.0000	0.9998	0.9988	0.9953	0.9871	0.9712	0.9444	0.9037	0.8471	0.7734
	5	1.0000	1.0000	0.9999	0.9996	0.9987	0.9962	0.9910	0.9812	0.9643	0.9375
	6	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9994	0.9984	0.9963	0.9922
8	0	0.6634	0.4305	0.2725	0.1678	0.1001	0.0576	0.0319	0.0168	0.0084	0.0039
	1	0.9428	0.8131	0.6572	0.5033	0.3671	0.2553	0.1691	0.1064	0.0632	0.0352
	2	0.9942	0.9619	0.8948	0.7969	0.6785	0.5518	0.4278	0.3154	0.2201	0.1445
	3	0.9996	0.9950	0.9786	0.9437	0.8862	0.8059	0.7064	0.5941	0.4770	0.3633
	4	1.0000	0.9996	0.9971	0.9896	0.9727	0.9420	0.8939	0.8263	0.7396	0.6367
	5	1.0000	1.0000	0.9998	0.9988	0.9958	0.9887	0.9747	0.9502	0.9115	0.8555
	6	1.0000	1.0000	1.0000	0.9999	0.9996	0.9987	0.9964	0.9915	0.9819	0.9648
	7	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998	0.9993	0.9983	0.9961
9	0	0.6302	0.3874	0.2316	0.1342	0.0751	0.0404	0.0207	0.0101	0.0046	0.0020
	1	0.9288	0.7748	0.5995	0.4362	0.3003	0.1960	0.1211	0.0705	0.0385	0.0195
	2	0.9916	0.9470	0.8591	0.7382	0.6007	0.4628	0.3373	0.2318	0.1495	0.0898
	3	0.9994	0.9917	0.9661	0.9144	0.8343	0.7297	0.6089	0.4826	0.3614	0.2539
	4	1.0000	0.9991	0.9944	0.9804	0.9511	0.9012	0.8283	0.7334	0.6214	0.5000
	5	1.0000	0.9999	0.9994	0.9969	0.9900	0.9747	0.9464	0.9006	0.8342	0.7461
	6	1.0000	1.0000	1.0000	0.9997	0.9987	0.9957	0.9888	0.9750	0.9502	0.9102
	7	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996	0.9986	0.9962	0.9909	0.9805
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9992	0.9980

Table I

BINOMIAL DISTRIBUTION FUNCTION (*Continued*)

n	x	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
						p					
10	0	0.5987	0.3487	0.1969	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010
	1	0.9139	0.7361	0.5443	0.3758	0.2440	0.1493	0.0860	0.0464	0.0233	0.0107
	2	0.9885	0.9298	0.8202	0.6778	0.5256	0.3828	0.2616	0.1673	0.0996	0.0547
	3	0.9990	0.9872	0.9500	0.8791	0.7759	0.6496	0.5138	0.3823	0.2660	0.1719
	4	0.9999	0.9984	0.9901	0.9672	0.9219	0.8497	0.7515	0.6331	0.5044	0.3770
	5	1.0000	0.9999	0.9986	0.9936	0.9803	0.9527	0.9051	0.8338	0.7384	0.6230
	6	1.0000	1.0000	0.9999	0.9991	0.9965	0.9894	0.9740	0.9452	0.8980	0.8281
	7	1.0000	1.0000	1.0000	0.9999	0.9996	0.9984	0.9952	0.9877	0.9726	0.9453
	8	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9983	0.9955	0.9893
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9990
11	0	0.5688	0.3138	0.1673	0.0859	0.0422	0.0198	0.0088	0.0036	0.0014	0.0005
	1	0.8981	0.6974	0.4922	0.3221	0.1971	0.1130	0.0606	0.0302	0.0139	0.0059
	2	0.9848	0.9104	0.7788	0.6174	0.4552	0.3127	0.2001	0.1189	0.0652	0.0327
	3	0.9984	0.9815	0.9306	0.8389	0.7133	0.5696	0.4256	0.2963	0.1911	0.1133
	4	0.9999	0.9972	0.9841	0.9496	0.8854	0.7897	0.6683	0.5328	0.3971	0.2744
	5	1.0000	0.9997	0.9973	0.9883	0.9657	0.9218	0.8513	0.7535	0.6331	0.5000
	6	1.0000	1.0000	0.9997	0.9980	0.9924	0.9784	0.9499	0.9006	0.8262	0.7256
	7	1.0000	1.0000	1.0000	0.9998	0.9988	0.9957	0.9878	0.9707	0.9390	0.8867
	8	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9980	0.9941	0.9852	0.9673
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9993	0.9978	0.9941
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9995
12	0	0.5404	0.2824	0.1422	0.0687	0.0317	0.0138	0.0057	0.0022	0.0008	0.0002
	1	0.8816	0.6590	0.4435	0.2749	0.1584	0.0850	0.0424	0.0196	0.0083	0.0032
	2	0.9804	0.8891	0.7358	0.5583	0.3907	0.2528	0.1513	0.0834	0.0421	0.0193
	3	0.9978	0.9744	0.9078	0.7946	0.6488	0.4925	0.3467	0.2253	0.1345	0.0730
	4	0.9998	0.9957	0.9761	0.9274	0.8424	0.7237	0.5833	0.4382	0.3044	0.1938
	5	1.0000	0.9995	0.9954	0.9806	0.9456	0.8822	0.7873	0.6652	0.5269	0.3872
	6	1.0000	0.9999	0.9993	0.9961	0.9857	0.9614	0.9154	0.8418	0.7393	0.6128
	7	1.0000	1.0000	0.9999	0.9994	0.9972	0.9905	0.9745	0.9427	0.8883	0.8062
	8	1.0000	1.0000	1.0000	0.9999	0.9996	0.9983	0.9944	0.9847	0.9644	0.9270
	9	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9992	0.9972	0.9921	0.9807
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9989	0.9968
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9998
13	0	0.5133	0.2542	0.1209	0.0550	0.0238	0.0097	0.0037	0.0013	0.0004	0.0001
	1	0.8646	0.6213	0.3983	0.2336	0.1267	0.0637	0.0296	0.0126	0.0049	0.0017
	2	0.9755	0.8661	0.6920	0.5017	0.3326	0.2025	0.1132	0.0579	0.0269	0.0112
	3	0.9969	0.9658	0.8820	0.7473	0.5843	0.4206	0.2783	0.1686	0.0929	0.0461
	4	0.9997	0.9935	0.9658	0.9009	0.7940	0.6543	0.5005	0.3530	0.2279	0.1334
	5	1.0000	0.9991	0.9925	0.9700	0.9198	0.8346	0.7159	0.5744	0.4268	0.2905
	6	1.0000	0.9999	0.9987	0.9930	0.9757	0.9376	0.8705	0.7712	0.6437	0.5000
	7	1.0000	1.0000	0.9998	0.9988	0.9944	0.9818	0.9538	0.9023	0.8212	0.7095
	8	1.0000	1.0000	1.0000	0.9998	0.9990	0.9960	0.9874	0.9679	0.9302	0.8666
	9	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9975	0.9922	0.9797	0.9539
	10	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9987	0.9959	0.9888
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9983
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
14	0	0.4877	0.2288	0.1028	0.0440	0.0178	0.0068	0.0024	0.0008	0.0002	0.0001
	1	0.8470	0.5846	0.3567	0.1979	0.1010	0.0475	0.0205	0.0081	0.0029	0.0009

Table I

BINOMIAL DISTRIBUTION FUNCTION (*Continued*)

n	x	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
14	2	0.9699	0.8416	0.6479	0.4481	0.2811	0.1608	0.0839	0.0398	0.0170	0.0065
	3	0.9958	0.9559	0.8535	0.6982	0.5213	0.3552	0.2205	0.1243	0.0632	0.0287
	4	0.9996	0.9908	0.9533	0.8702	0.7415	0.5842	0.4227	0.2793	0.1672	0.0898
	5	1.0000	0.9985	0.9885	0.9561	0.8883	0.7805	0.6405	0.4859	0.3373	0.2120
	6	1.0000	0.9998	0.9978	0.9884	0.9617	0.9067	0.8164	0.6925	0.5461	0.3953
	7	1.0000	1.0000	0.9997	0.9976	0.9897	0.9685	0.9247	0.8499	0.7414	0.6047
	8	1.0000	1.0000	1.0000	0.9996	0.9978	0.9917	0.9757	0.9417	0.8811	0.7880
	9	1.0000	1.0000	1.0000	1.0000	0.9997	0.9983	0.9940	0.9825	0.9574	0.9102
	10	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9989	0.9961	0.9886	0.9713
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9978	0.9935
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9991
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
15	0	0.4633	0.2059	0.0874	0.0352	0.0134	0.0047	0.0016	0.0005	0.0001	0.0000
	1	0.8290	0.5490	0.3186	0.1671	0.0802	0.0353	0.0142	0.0052	0.0017	0.0005
	2	0.9638	0.8159	0.6042	0.3980	0.2361	0.1268	0.0617	0.0271	0.0107	0.0037
	3	0.9945	0.9444	0.8227	0.6482	0.4613	0.2969	0.1727	0.0905	0.0424	0.0176
	4	0.9994	0.9873	0.9383	0.8358	0.6865	0.5155	0.3519	0.2173	0.1204	0.0592
	5	0.9999	0.9978	0.9832	0.9389	0.8516	0.7216	0.5643	0.4032	0.2608	0.1509
	6	1.0000	0.9997	0.9964	0.9819	0.9434	0.8689	0.7548	0.6098	0.4522	0.3036
	7	1.0000	1.0000	0.9994	0.9958	0.9827	0.9500	0.8868	0.7869	0.6535	0.5000
	8	1.0000	1.0000	0.9999	0.9992	0.9958	0.9848	0.9578	0.9050	0.8182	0.6964
	9	1.0000	1.0000	1.0000	0.9999	0.9992	0.9963	0.9876	0.9662	0.9231	0.8491
	10	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9972	0.9907	0.9745	0.9408
	11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9981	0.9937	0.9824
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9989	0.9963
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
16	0	0.4401	0.1853	0.0743	0.0281	0.0100	0.0033	0.0010	0.0003	0.0001	0.0000
	1	0.8105	0.5147	0.2839	0.1407	0.0635	0.0261	0.0098	0.0033	0.0010	0.0003
	2	0.9571	0.7892	0.5614	0.3518	0.1971	0.0994	0.0451	0.0183	0.0066	0.0021
	3	0.9930	0.9316	0.7899	0.5981	0.4050	0.2459	0.1339	0.0651	0.0281	0.0106
	4	0.9991	0.9830	0.9209	0.7982	0.6302	0.4499	0.2892	0.1666	0.0853	0.0384
	5	0.9999	0.9967	0.9765	0.9183	0.8103	0.6598	0.4900	0.3288	0.1976	0.1051
	6	1.0000	0.9995	0.9944	0.9733	0.9204	0.8247	0.6881	0.5272	0.3660	0.2272
	7	1.0000	0.9999	0.9989	0.9930	0.9729	0.9256	0.8406	0.7161	0.5629	0.4018
	8	1.0000	1.0000	0.9998	0.9985	0.9925	0.9743	0.9329	0.8577	0.7441	0.5982
	9	1.0000	1.0000	1.0000	0.9998	0.9984	0.9929	0.9771	0.9417	0.8759	0.7728
	10	1.0000	1.0000	1.0000	1.0000	0.9997	0.9984	0.9938	0.9809	0.9514	0.8949
	11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9987	0.9951	0.9851	0.9616
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9991	0.9965	0.9894
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9979
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
17	0	0.4181	0.1668	0.0631	0.0225	0.0075	0.0023	0.0007	0.0002	0.0000	0.0000
	1	0.7922	0.4818	0.2525	0.1182	0.0501	0.0193	0.0067	0.0021	0.0006	0.0001
	2	0.9497	0.7618	0.5198	0.3096	0.1637	0.0774	0.0327	0.0123	0.0041	0.0012
	3	0.9912	0.9174	0.7556	0.5489	0.3530	0.2019	0.1028	0.0464	0.0184	0.0064
	4	0.9988	0.9779	0.9013	0.7582	0.5739	0.3887	0.2348	0.1260	0.0596	0.0245

Table I

BINOMIAL DISTRIBUTION FUNCTION (*Continued*)

n	x	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
17	5	0.9999	0.9953	0.9681	0.8943	0.7653	0.5968	0.4197	0.2639	0.1471	0.0717
	6	1.0000	0.9992	0.9917	0.9623	0.8929	0.7752	0.6188	0.4478	0.2902	0.1662
	7	1.0000	0.9999	0.9983	0.9891	0.9598	0.8954	0.7872	0.6405	0.4743	0.3145
	8	1.0000	1.0000	0.9997	0.9974	0.9876	0.9597	0.9006	0.8011	0.6626	0.5000
	9	1.0000	1.0000	1.0000	0.9995	0.9969	0.9873	0.9617	0.9081	0.8166	0.6855
	10	1.0000	1.0000	1.0000	0.9999	0.9994	0.9968	0.9880	0.9652	0.9174	0.8338
	11	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9970	0.9894	0.9699	0.9283
	12	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9975	0.9914	0.9755
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9981	0.9936
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9997	0.9988
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
18	0	0.3972	0.1501	0.0536	0.0180	0.0056	0.0016	0.0004	0.0001	0.0000	0.0000
	1	0.7735	0.4503	0.2241	0.0991	0.0395	0.0142	0.0046	0.0013	0.0003	0.0001
	2	0.9419	0.7338	0.4797	0.2713	0.1353	0.0600	0.0236	0.0082	0.0025	0.0007
	3	0.9891	0.9018	0.7202	0.5010	0.3057	0.1646	0.0783	0.0328	0.0120	0.0038
	4	0.9985	0.9718	0.8794	0.7164	0.5187	0.3327	0.1886	0.0942	0.0411	0.0154
	5	0.9998	0.9936	0.9581	0.8671	0.7175	0.5344	0.3550	0.2088	0.1077	0.0481
	6	1.0000	0.9988	0.9882	0.9487	0.8610	0.7217	0.5491	0.3743	0.2258	0.1189
	7	1.0000	0.9998	0.9973	0.9837	0.9431	0.8593	0.7283	0.5634	0.3915	0.2403
	8	1.0000	1.0000	0.9995	0.9957	0.9807	0.9404	0.8609	0.7368	0.5778	0.4073
	9	1.0000	1.0000	0.9999	0.9991	0.9946	0.9790	0.9403	0.8653	0.7473	0.5927
	10	1.0000	1.0000	1.0000	0.9998	0.9988	0.9939	0.9788	0.9424	0.8720	0.7597
	11	1.0000	1.0000	1.0000	1.0000	0.9998	0.9986	0.9938	0.9797	0.9463	0.8811
	12	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9986	0.9942	0.9817	0.9519
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9987	0.9951	0.9846
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9990	0.9962
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
19	0	0.3774	0.1351	0.0456	0.0144	0.0042	0.0011	0.0003	0.0001	0.0000	0.0000
	1	0.7547	0.4203	0.1985	0.0829	0.0310	0.0104	0.0031	0.0008	0.0002	0.0000
	2	0.9335	0.7054	0.4413	0.2369	0.1113	0.0462	0.0170	0.0055	0.0015	0.0004
	3	0.9868	0.8850	0.6841	0.4551	0.2630	0.1332	0.0591	0.0230	0.0077	0.0022
	4	0.9980	0.9648	0.8556	0.6733	0.4654	0.2822	0.1500	0.0696	0.0280	0.0096
	5	0.9998	0.9914	0.9463	0.8369	0.6678	0.4739	0.2968	0.1629	0.0777	0.0318
	6	1.0000	0.9983	0.9837	0.9324	0.8251	0.6655	0.4812	0.3081	0.1727	0.0835
	7	1.0000	0.9997	0.9959	0.9767	0.9225	0.8180	0.6656	0.4878	0.3169	0.1796
	8	1.0000	1.0000	0.9992	0.9933	0.9713	0.9161	0.8145	0.6675	0.4940	0.3238
	9	1.0000	1.0000	0.9999	0.9984	0.9911	0.9674	0.9125	0.8139	0.6710	0.5000
	10	1.0000	1.0000	1.0000	0.9997	0.9977	0.9895	0.9653	0.9115	0.8159	0.6762
	11	1.0000	1.0000	1.0000	1.0000	0.9995	0.9972	0.9886	0.9648	0.9129	0.8204
	12	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9969	0.9884	0.9658	0.9165
	13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9969	0.9891	0.9682
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9994	0.9972	0.9904
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9995	0.9978
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9996
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table I

BINOMIAL DISTRIBUTION FUNCTION (*Continued*)

n	x	p = 0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
20	0	0.3585	0.1216	0.0388	0.0115	0.0032	0.0008	0.0002	0.0000	0.0000	0.0000
	1	0.7358	0.3917	0.1756	0.0692	0.0243	0.0076	0.0021	0.0005	0.0001	0.0000
	2	0.9245	0.6769	0.4049	0.2061	0.0913	0.0355	0.0121	0.0036	0.0009	0.0002
	3	0.9841	0.8670	0.6477	0.4114	0.2252	0.1071	0.0444	0.0160	0.0049	0.0013
	4	0.9974	0.9568	0.8298	0.6296	0.4148	0.2375	0.1182	0.0510	0.0189	0.0059
	5	0.9997	0.9887	0.9327	0.8042	0.6172	0.4164	0.2454	0.1256	0.0553	0.0207
	6	1.0000	0.9976	0.9781	0.9133	0.7858	0.6080	0.4166	0.2500	0.1299	0.0577
	7	1.0000	0.9996	0.9941	0.9679	0.8982	0.7723	0.6010	0.4159	0.2520	0.1316
	8	1.0000	0.9999	0.9987	0.9900	0.9591	0.8867	0.7624	0.5956	0.4143	0.2517
	9	1.0000	1.0000	0.9998	0.9974	0.9861	0.9520	0.8782	0.7553	0.5914	0.4119
	10	1.0000	1.0000	1.0000	0.9994	0.9961	0.9829	0.9468	0.8725	0.7507	0.5881
	11	1.0000	1.0000	1.0000	0.9999	0.9991	0.9949	0.9804	0.9435	0.8692	0.7483
	12	1.0000	1.0000	1.0000	1.0000	0.9998	0.9987	0.9940	0.9790	0.9420	0.8684
	13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9985	0.9935	0.9786	0.9423
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9984	0.9936	0.9793
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9985	0.9941
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9987
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table II

POISSON DISTRIBUTION FUNCTION*

$$F(x, \lambda) = \sum_{k=0}^{x} e^{-\lambda} \frac{\lambda^k}{k!}$$

λ \ x	0	1	2	3	4	5	6	7	8	9
0.02	0.980	1.000								
0.04	0.961	0.999	1.000							
0.06	0.942	0.998	1.000							
0.08	0.923	0.997	1.000							
0.10	0.905	0.995	1.000							
0.15	0.861	0.990	0.999	1.000						
0.20	0.819	0.982	0.999	1.000						
0.25	0.779	0.974	0.998	1.000						
0.30	0.741	0.963	0.996	1.000						
0.35	0.705	0.951	0.994	1.000						
0.40	0.670	0.938	0.992	0.999	1.000					
0.45	0.638	0.925	0.989	0.999	1.000					
0.50	0.607	0.910	0.986	0.998	1.000					
0.55	0.577	0.894	0.982	0.998	1.000					
0.60	0.549	0.878	0.977	0.997	1.000					
0.65	0.522	0.861	0.972	0.996	0.999	1.000				
0.70	0.497	0.844	0.966	0.994	0.999	1.000				
0.75	0.472	0.827	0.959	0.993	0.999	1.000				
0.80	0.449	0.809	0.953	0.991	0.999	1.000				
0.85	0.427	0.791	0.945	0.989	0.998	1.000				
0.90	0.407	0.772	0.937	0.987	0.998	1.000				
0.95	0.387	0.754	0.929	0.984	0.997	1.000				
1.00	0.368	0.736	0.920	0.981	0.996	0.999	1.000			
1.1	0.333	0.699	0.900	0.974	0.995	0.999	1.000			
1.2	0.301	0.663	0.879	0.966	0.992	0.998	1.000			
1.3	0.273	0.627	0.857	0.957	0.989	0.998	1.000			
1.4	0.247	0.592	0.833	0.946	0.986	0.997	0.999	1.000		
1.5	0.223	0.558	0.809	0.934	0.981	0.996	0.999	1.000		
1.6	0.202	0.525	0.783	0.921	0.976	0.994	0.999	1.000		
1.7	0.183	0.493	0.757	0.907	0.970	0.992	0.998	1.000		
1.8	0.165	0.463	0.731	0.891	0.964	0.990	0.997	0.999	1.000	
1.9	0.150	0.434	0.704	0.875	0.956	0.987	0.997	0.999	1.000	
2.0	0.135	0.406	0.677	0.857	0.947	0.983	0.995	0.999	1.000	

* Reprinted by kind permission from E. C. Molina, *Poisson's Exponential Binomial Limit*, D. Van Nostrand Company, Inc., Princeton, N.J., 1947.

Table II

POISSON DISTRIBUTION FUNCTION (*Continued*)

λ \ x	0	1	2	3	4	5	6	7	8	9
2.2	0.111	0.355	0.023	0.819	0.928	0.975	0.993	0.998	1.000	
2.4	0.091	0.308	0.570	0.779	0.904	0.964	0.988	0.997	0.999	1.000
2.6	0.074	0.207	0.518	0.736	0.877	0.951	0.983	0.995	0.999	1.000
2.8	0.061	0.231	0.469	0.692	0.848	0.935	0.976	0.992	0.998	0.999
3.0	0.050	0.199	0.423	0.647	0.815	0.916	0.966	0.988	0.996	0.999
3.2	0.041	0.171	0.380	0.603	0.781	0.895	0.955	0.983	0.994	0.998
3.4	0.033	0.147	0.340	0.558	0.744	0.871	0.942	0.977	0.992	0.997
3.6	0.027	0.126	0.303	0.515	0.706	0.844	0.927	0.969	0.988	0.996
3.8	0.022	0.107	0.209	0.473	0.668	0.816	0.909	0.960	0.984	0.994
4.0	0.018	0.092	0.238	0.433	0.629	0.785	0.889	0.949	0.979	0.992
4.2	0.015	0.078	0.210	0.395	0.590	0.753	0.867	0.936	0.972	0.989
4.4	0.012	0.066	0.185	0.359	0.551	0.720	0.844	0.921	0.964	0.985
4.6	0.010	0.056	0.163	0.326	0.513	0.686	0.818	0.905	0.955	0.980
4.8	0.008	0.048	0.143	0.294	0.476	0.651	0.791	0.887	0.944	0.975
5.0	0.007	0.040	0.125	0.265	0.440	0.616	0.762	0.867	0.932	0.968
5.2	0.006	0.034	0.109	0.238	0.406	0.581	0.732	0.845	0.918	0.960
5.4	0.005	0.029	0.095	0.213	0.373	0.546	0.702	0.822	0.903	0.951
5.6	0.004	0.024	0.082	0.191	0.342	0.512	0.670	0.797	0.886	0.941
5.8	0.003	0.021	0.072	0.170	0.313	0.478	0.638	0.771	0.867	0.929
6.0	0.002	0.017	0.062	0.151	0.285	0.446	0.606	0.744	0.847	0.916

λ \ x	10	11	12	13	14	15	16
2.8	1.000						
3.0	1.000						
3.2	1.000						
3.4	0.999	1.000					
3.6	0.999	1.000					
3.8	0.998	0.999	1.000				
4.0	0.997	0.999	1.000				
4.2	0.996	0.999	1.000				
4.4	0.994	0.998	0.999	1.000			
4.6	0.992	0.997	0.999	1.000			
4.8	0.990	0.996	0.999	1.000			
5.0	0.986	0.995	0.998	0.999	1.000		
5.2	0.982	0.993	0.997	0.999	1.000		
5.4	0.977	0.990	0.996	0.999	1.000		
5.6	0.972	0.988	0.995	0.998	0.999	1.000	
5.8	0.965	0.984	0.993	0.997	0.999	1.000	
6.0	0.957	0.980	0.991	0.996	0.999	0.999	1.000

Table II

POISSON DISTRIBUTION FUNCTION (*Continued*)

λ \ x	0	1	2	3	4	5	6	7	8	9
6.2	0.002	0.015	0.054	0.134	0.259	0.414	0.574	0.716	0.826	0.902
6.4	0.002	0.012	0.046	0.119	0.235	0.384	0.542	0.687	0.803	0.886
6.6	0.001	0.010	0.040	0.105	0.213	0.355	0.511	0.658	0.780	0.869
6.8	0.001	0.009	0.034	0.093	0.192	0.327	0.480	0.628	0.755	0.850
7.0	0.001	0.007	0.030	0.082	0.173	0.301	0.450	0.599	0.729	0.830
7.2	0.001	0.006	0.025	0.072	0.156	0.276	0.420	0.569	0.703	0.810
7.4	0.001	0.005	0.022	0.063	0.140	0.253	0.392	0.539	0.676	0.788
7.6	0.001	0.004	0.019	0.055	0.125	0.231	0.365	0.510	0.648	0.765
7.8	0.000	0.004	0.016	0.048	0.112	0.210	0.338	0.481	0.620	0.741
8.0	0.000	0.003	0.014	0.042	0.100	0.191	0.313	0.453	0.593	0.717
8.5	0.000	0.002	0.009	0.030	0.074	0.150	0.256	0.386	0.523	0.653
9.0	0.000	0.001	0.006	0.021	0.055	0.116	0.207	0.324	0.456	0.587
9.5	0.000	0.001	0.004	0.015	0.040	0.089	0.165	0.269	0.392	0.522
10.0	0.000	0.000	0.003	0.010	0.029	0.067	0.130	0.220	0.333	0.458

λ \ x	10	11	12	13	14	15	16	17	18	19
6.2	0.949	0.975	0.989	0.995	0.998	0.999	1.000			
6.4	0.939	0.969	0.986	0.994	0.997	0.999	1.000			
6.6	0.927	0.963	0.982	0.992	0.997	0.999	0.999	1.000		
6.8	0.915	0.955	0.978	0.990	0.996	0.998	0.999	1.000		
7.0	0.901	0.947	0.973	0.987	0.994	0.998	0.999	1.000		
7.2	0.887	0.937	0.967	0.984	0.993	0.997	0.999	0.999	1.000	
7.4	0.871	0.926	0.961	0.980	0.991	0.996	0.998	0.999	1.000	
7.6	0.854	0.915	0.954	0.976	0.989	0.995	0.998	0.999	1.000	
7.8	0.835	0.902	0.945	0.971	0.986	0.993	0.997	0.999	1.000	
8.0	0.816	0.888	0.936	0.966	0.983	0.992	0.996	0.998	0.999	1.000
8.5	0.763	0.849	0.909	0.949	0.973	0.986	0.993	0.997	0.999	0.999
9.0	0.706	0.803	0.876	0.926	0.959	0.978	0.989	0.995	0.998	0.999
9.5	0.645	0.752	0.836	0.898	0.940	0.967	0.982	0.991	0.996	0.998
10.0	0.583	0.697	0.792	0.864	0.917	0.951	0.973	0.986	0.993	0.997

λ \ x	20	21	22
8.5	1.000		
9.0	1.000		
9.5	0.999	1.000	
10.0	0.998	0.999	1.000

Table II

POISSON DISTRIBUTION FUNCTION (*Continued*)

λ \ x	0	1	2	3	4	5	6	7	8	9
10.5	0.000	0.000	0.002	0.007	0.021	0.050	0.102	0.179	0.279	0.397
11.0	0.000	0.000	0.001	0.005	0.015	0.038	0.079	0.143	0.232	0.341
11.5	0.000	0.000	0.001	0.003	0.011	0.028	0.060	0.114	0.191	0.289
12.0	0.000	0.000	0.001	0.002	0.008	0.020	0.046	0.090	0.155	0.242
12.5	0.000	0.000	0.000	0.002	0.005	0.015	0.035	0.070	0.125	0.201
13.0	0.000	0.000	0.000	0.001	0.004	0.011	0.026	0.054	0.100	0.166
13.5	0.000	0.000	0.000	0.001	0.003	0.008	0.019	0.041	0.079	0.135
14.0	0.000	0.000	0.000	0.000	0.002	0.006	0.014	0.032	0.062	0.109
14.5	0.000	0.000	0.000	0.000	0.001	0.004	0.010	0.024	0.048	0.088
15.0	0.000	0.000	0.000	0.000	0.001	0.003	0.008	0.018	0.037	0.070
	10	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
10.5	0.521	0.639	0.742	0.825	0.888	0.932	0.960	0.978	0.988	0.994
11.0	0.460	0.579	0.689	0.781	0.854	0.907	0.944	0.968	0.982	0.991
11.5	0.402	0.520	0.633	0.733	0.815	0.878	0.924	0.954	0.974	0.986
12.0	0.347	0.462	0.576	0.682	0.772	0.844	0.899	0.937	0.963	0.979
12.5	0.297	0.406	0.519	0.628	0.725	0.806	0.869	0.916	0.948	0.969
13.0	0.252	0.353	0.463	0.573	0.675	0.764	0.835	0.890	0.930	0.957
13.5	0.211	0.304	0.409	0.518	0.623	0.718	0.798	0.861	0.908	0.942
14.0	0.176	0.260	0.358	0.464	0.570	0.669	0.756	0.827	0.883	0.923
14.5	0.145	0.220	0.311	0.413	0.518	0.619	0.711	0.790	0.853	0.901
15.0	0.118	0.185	0.268	0.363	0.466	0.568	0.664	0.749	0.819	0.875
	20	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**
10.5	0.997	0.999	0.999	1.000						
11.0	0.995	0.998	0.999	1.000						
11.5	0.992	0.996	0.998	0.999	1.000					
12.0	0.988	0.994	0.997	0.999	0.999	1.000				
12.5	0.983	0.991	0.995	0.998	0.999	0.999	1.000			
13.0	0.975	0.986	0.992	0.996	0.998	0.999	1.000			
13.5	0.965	0.980	0.989	0.994	0.997	0.998	0.999	1.000		
14.0	0.952	0.971	0.983	0.991	0.995	0.997	0.999	0.999	1.000	
14.5	0.936	0.960	0.976	0.986	0.992	0.996	0.998	0.999	0.999	1.000
15.0	0.917	0.947	0.967	0.981	0.989	0.994	0.997	0.998	0.999	1.000

Table II

POISSON DISTRIBUTION FUNCTION (*Continued*)

λ \ x	4	5	6	7	8	9	10	11	12	13
16	0.000	0.001	0.004	0.010	0.022	0.043	0.077	0.127	0.193	0.275
17	0.000	0.001	0.002	0.005	0.013	0.026	0.049	0.085	0.135	0.201
18	0.000	0.000	0.001	0.003	0.007	0.015	0.030	0.055	0.092	0.143
19	0.000	0.000	0.001	0.002	0.004	0.009	0.018	0.035	0.061	0.098
20	0.000	0.000	0.000	0.001	0.002	0.005	0.011	0.021	0.039	0.066
21	0.000	0.000	0.000	0.000	0.001	0.003	0.006	0.013	0.025	0.043
22	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.008	0.015	0.028
23	0.000	0.000	0.000	0.000	0.000	0.001	0.002	0.004	0.009	0.017
24	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.003	0.005	0.011
25	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.001	0.003	0.006
	14	**15**	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**
16	0.368	0.467	0.566	0.659	0.742	0.812	0.868	0.911	0.942	0.963
17	0.281	0.371	0.468	0.564	0.655	0.736	0.805	0.861	0.905	0.937
18	0.208	0.287	0.375	0.469	0.562	0.651	0.731	0.799	0.855	0.899
19	0.150	0.215	0.292	0.378	0.469	0.561	0.647	0.725	0.793	0.849
20	0.105	0.157	0.221	0.297	0.381	0.470	0.559	0.644	0.721	0.787
21	0.072	0.111	0.163	0.227	0.302	0.384	0.471	0.558	0.640	0.716
22	0.048	0.077	0.117	0.169	0.232	0.306	0.387	0.472	0.556	0.637
23	0.031	0.052	0.082	0.123	0.175	0.238	0.310	0.389	0.472	0.555
24	0.020	0.034	0.056	0.087	0.128	0.180	0.243	0.314	0.392	0.473
25	0.012	0.022	0.038	0.060	0.092	0.134	0.185	0.247	0.318	0.394
	24	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**
16	0.978	0.987	0.993	0.996	0.998	0.999	0.999	1.000		
17	0.959	0.975	0.985	0.991	0.995	0.997	0.999	0.999	1.000	
18	0.932	0.955	0.972	0.983	0.990	0.994	0.997	0.998	0.999	1.000
19	0.893	0.927	0.951	0.969	0.980	0.988	0.993	0.996	0.998	0.999
20	0.843	0.888	0.922	0.948	0.966	0.978	0.987	0.992	0.995	0.997
21	0.782	0.838	0.883	0.917	0.944	0.963	0.976	0.985	0.991	0.994
22	0.712	0.777	0.832	0.877	0.913	0.940	0.959	0.973	0.983	0.989
23	0.635	0.708	0.772	0.827	0.873	0.908	0.936	0.956	0.971	0.981
24	0.554	0.632	0.704	0.768	0.823	0.868	0.904	0.932	0.953	0.969
25	0.473	0.553	0.629	0.700	0.763	0.818	0.863	0.900	0.929	0.950
	34	**35**	**36**	**37**	**38**	**39**	**40**	**41**	**42**	**43**
19	0.999	1.000								
20	0.999	0.999	1.000							
21	0.997	0.998	0.999	0.999	1.000					
22	0.994	0.996	0.998	0.999	0.999	1.000				
23	0.988	0.993	0.996	0.997	0.999	0.999	1.000			
24	0.979	0.987	0.992	0.995	0.997	0.998	0.999	0.999	1.000	
25	0.966	0.978	0.985	0.991	0.994	0.997	0.998	0.999	0.999	1.000

and it is of interest to note that although only 5 per cent of the reports handled by secretary No. 1 are misfiled, over 11 per cent of all misfiled reports are her responsibility.

EXERCISES

1. Referring to Figure 2.10, find $P(M \mid G)$ and $P(M \mid G')$, assume that originally each of the 50 employees had the same probability of being elected.

2. If in the example on page 27 it is also given that $P(L) = 0.60$, find the probability that a shipment was late arriving in Chicago, given that it arrived on time in Los Angeles.

3. Referring to Exercise 13 on page 24 and Figure 2.12, find the following probabilities:
   (a) the damaged car is a compact, given that it has power steering,
   (b) the damaged car has automatic transmission, given that it is a compact,
   (c) the damaged car has power steering or an automatic transmission, given that it is a compact,
   (d) the damaged car is a compact with automatic transmission, given that it does not have power steering,
   (e) the damaged car is not a compact, given that it has power steering and an automatic transmission,
   (f) the damaged car is not a compact, given that it has power steering, an automatic transmission, or both.

4. Referring to the probabilities of Figure 2.13, find
   (a) $P(A \mid B)$,
   (b) $P(B \mid C')$,
   (c) $P(A \cap B \mid C)$,
   (d) $P(B \cup C \mid A')$,
   (e) $P(A \mid B \cup C)$,
   (f) $P(A \mid B \cap C)$,
   (g) $P(A \cap B \cap C \mid B \cap C)$,
   (h) $P(A \cap B \cap C \mid B \cup C)$.

5. Using the results of Exercise 6 on page 22, find
   (a) the probability that the whole assembly has two defective components, given that the first subassembly has one defective component,
   (b) the probability that the first subassembly has at least two defective components, given that the whole assembly has three defective components,
   (c) the probability that the whole assembly has at least three defective components, given that the first subassembly has more defective components than the second subassembly.

6. The probability that a construction job will be finished on time is 17/20, the probability that there will be no strikes is 3/4, and the probability that the

construction job will be finished on time, given that there will be no strikes, is 14/15. Find
   (a) the probability that the construction job will be finished on time and there will be no strikes,
   (b) the probability that there will have been no strikes, given that the construction job will be finished on time.

7. An urn contains 40 white marbles and 10 black marbles. If two marbles are drawn at random (with equal probabilities), find the probability that they are both white if
   (a) the first marble is replaced before the second is drawn,
   (b) the first marble is not replaced before the second is drawn.

8. What is the probability of drawing three aces in succession from a well-shuffled standard deck of 52 playing cards, if cards are not put back into the deck immediately after they have been drawn?

9. Suppose the probability that the Los Angeles Dodgers will win the National League pennant is 0.25 and the probability that the San Francisco Giants will win it is 0.20. Furthermore, the probability that the National League team will win the World Series is 0.45, 0.55, or 0.35, depending on whether the Los Angeles Dodgers, the San Francisco Giants, or some other team wins the National League pennant. What is the probability that a National League team will win the World Series?

10. Referring to Exercise 2 above, find the probability that such an order will arrive on time in Los Angeles, given that it arrived on time in Chicago. [*Hint:* Use the rule of elimination.]

11. Referring to the illustration on page 29, find $P(B_1 \mid A)$ and $P(B_2 \mid \bar{A})$.

12. Given that a National League team wins the World Series, use the probabilities of Exercise 9 above to find the probability that the Los Angeles Dodgers won the National League pennant.

13. The probability that an airplane accident which is due to structural failure is diagnosed correctly is 0.85 and the probability that an airplane accident which is *not* due to structural failure is diagnosed incorrectly as being due to structural failure is 0.35. If 30 per cent of all airplane accidents are due to structural failures, find the probability that an airplane accident is due to structural failure, given that it has been diagnosed as being due to structural failure.

14. A race driver uses a Corvette in 50 per cent of the races he enters, a Jaguar in 30 per cent of the races, and an Alfa Romeo in 20 per cent of the races. Of 25 races he has entered with the Corvette he has won 5, of 15 races he has entered with the Jaguar he has won 4, and of 10 races he has entered with the Alfa Romeo he has won 4. Using these figures to estimate the respective probabilities, what is the probability that this race driver will win in the race for which he is entered at Le Mans?

(d) $P(A \cup B) = 0.70$, $P(A) = 0.30$, and $P(B) = 0.40$
(e) $P(A \cup B) = 0.90$, $P(A|B) = 0.80$, and $P(B) = 0.30$
(f) Two dice are tossed.
  (i) A = first die comes up "4," B = second die comes up "6."
  (ii) A = sum on dice is 4, B = sum on dice is 6.
(g) A coin is flipped twice.
  (i) A = head occurs on first flip, B = tail occurs on first flip
  (ii) A = head occurs on first flip, B = tail occurs on second flip
  (iii) A = head occurs on first flip, B = head occurs on second flip
(h) A coin is flipped five times.
  (i) A = the outcome *hhhtt*, B = the outcome *hthht*.
  (ii) A = the occurrence of three heads and two tails
  B = the occurrence of one head and four tails
  (iii) A = the occurrence of four heads on the first four flips
  B = the occurrence of a head on the fifth flip.
(i) Two objects are randomly selected without replacement from a group of 10 objects (four green and six yellow).
  (i) A = a green object is selected first.
  B = a yellow object is selected second.
  (ii) same—but with replacement.

2. Given that A, B, and C are mutually exclusive events, $P(A) = 0.62$, $P(B) = 0.20$, and $P(C) = 0.15$, find.

(a) $P(A \cap B)$	(b) $P(A\|C)$
(c) $P(A \cup C)$	(d) $P(A \cup B \cup C)$
(e) $P(A \cap B \cap C)$	(f) $P(A' \cap B' \cap C')$
(g) $P(A' \cup B)$	(h) $P(A' \cap B)$

(It sometimes helps to draw the appropriate Venn diagram, filling in the probabilities associated with each region.)

3. Given that A, B, and C are independent events, $P(A) = 0.60$, $P(B) = 0.50$, and $P(C) = 0.90$, find.

(a) $P(A \cap B)$	(b) $P(B\|C)$
(c) $P(A \cup C)$	(d) $P(A' \cap B)$
(e) $P(B' \cap C')$	(f) $P(A \cup B \cup C)$
(g) $P(A \cap B \cap C)$	(h) $P(A \cup B')$

4. Why are events A and B, defined on page 127, neither mutually exclusive nor independent?

6.8 Bayes' Theorem

In this section we shall extend our knowledge of conditional probability by considering a particular type of situation in which the occurrence of a certain event (E) depends upon the occurrence of one of k mutually exclusive events ($F_1, F_2, \ldots, F_k$). For example, the probability that the local tax rate will increase within the next year may depend upon which of three mayoral candidates is elected this year. The event that a new washing machine will need repair within the first two years of use de-

Exercises

1. Let $\{X, Y, Z\}$ be a partition of sample space S and W be a subset of S. We shall assume that $P(X) = 0.20$, $P(Y) = 0.50$, $P(W|X) = 0.30$, $P(W|Y) = 0.60$, and $P(W|Z) = 0.80$. Find.

(a) $P(Z)$	(b) $P(W)$	(c) $P(X\|W)$
(d) $P(Y\|W)$	(e) $P(Z\|W)$	

2. In a college fraternity 20 percent of the membership are seniors, 25 percent are juniors, 30 percent are sophomores and 25 percent are freshmen. None of the seniors are math majors. However, 10 percent of the juniors, five percent of the sophomores, and eight percent of the freshmen are math majors. If a person were selected randomly from the total membership of this fraternity,
(a) what is the probability that he would be a math major?
(b) what is the probability that he would be a sophomore, given that he is a math major?

3. Container A holds five pennies and three nickels, B holds two pennies and four nickels, and C holds three pennies and three nickels. The task is to choose one of the containers and then randomly select a coin from it. The container is chosen by the following method: toss a die—if a one appears, choose A, if a two or three appears, choose B, and if a four, five, or six appears, choose C.
(a) What is the probability that the coin we select will be a nickel?
(b) If someone else performed this experiment and told you, after he was done, that he selected a nickel, what is the probability that he selected a nickel from container B?

4. An old machine produces three defective parts in every 100 parts produced. A newer machine produces only one defective part in every 1000 parts produced. Furthermore, the new machine yields three times the output of the old machine. During a quality control check, a part is selected from the joint output of the two machines and found to be defective. What is the probability that the defective part was produced by the old machine?

5. The probability that John's office door is locked is 0.80. The key to the office is one of four identical-looking keys in John's pocket. If he selects a key at random to try in the door, what is the probability that he can open the door (i.e., that either the key works, or the door was already unlocked)?

6.9 General Application of Classical Probability Theory

In this chapter we have been developing from the *classical* approach a *theory of probability*. It is important to remember one significant limitation of the classical approach, namely, that it assumes that the sample space involved is a set of *equally likely* outcomes. Unless this condition is satisfied, our two basic definitions (Definitions 6.10 and 6.12) cannot be used to find probabilities. The theorems which describe the basic laws governing probability (Theorems 6.1–6.10) are based on the

Se puede demostrar que la relación entre el coeficiente de correlación, la variancia total y la variancia explicada es

$$r_{xy}^2 = \frac{S_{\tilde{y}}}{S_y^2} \qquad (1\ 21)$$

El coeficiente de correlación también se puede calcular despejando r_{xy} de la ec 1.18, lo cual conduce a

$$r_{xy} = \sqrt{1 - \frac{S_{y|x}^2}{S_y^2}} \qquad (1\ 22)$$

Ejemplo

Calcular el coeficiente de correlación para los datos de las cargas vivas de los ejemplos anteriores si se sabe que $S_y^2 = 28\ 561$ kg² y que la ecuación de la recta de regresión es $\tilde{y} = 0.22\ x + 307.3$ (se consideran los datos básicos).

Como la variancia explicada es

$$S_{\tilde{y}}^2 = \frac{\sum_{i=1}^{i=n} (\tilde{y}_i - \bar{y})^2}{n} = 869 \text{ kg}^2$$

entonces

$$r_{xy} = \sqrt{\frac{S_{\tilde{y}}^2}{S_y^2}} = \sqrt{\frac{869}{28\ 561}} = 0.175$$

Obsérvese que anteriormente se obtuvo un valor de 0.18 para r_{xy}, el valor 0.175 de este ejemplo se debió a la falta de exactitud en el cálculo numérico.

Tomando en cuenta que el valor de la variancia de la predicción que se obtuvo para este ejemplo fue $S_{y|x}^2 = 27\ 635$ kg², mediante la ec 1.22 se obtiene

$$r_{xy} = \sqrt{1 - \frac{27\ 635}{28\ 561}} = \sqrt{0.0324} = 0.18$$

Si a la diferencia $y_i - \overline{y}$ se le llama *desviación total* de y_i con respecto al promedio aritmético, $\overline{y}$, la ecuación

$$y_i - \overline{y} = (\widetilde{y}_i - \overline{y}) + (y_i - \widetilde{y}_i)$$

indica que la desviación total es igual a la desviación explicada más $y_i - \widetilde{y}_i$ Las desviaciones $y_i - \widetilde{y}_i$ ocurren al azar, es decir, en forma inexplicable, de ahí que se les conozca con el nombre de *desviaciones inexplicadas* (no explicadas)

Como consecuencia de la ecuación anterior, se puede demostrar que

$$\frac{\sum_{i=1}^{i=n} (y_i - \overline{y})^2}{n} = \frac{\sum_{i=1}^{i=n} (\widetilde{y}_i - \overline{y})^2}{n} + \frac{\sum_{i=1}^{i=n} (y_i - \widetilde{y}_i)^2}{n}$$

El miembro izquierdo de esta ecuación corresponde a la variancia, S_y^2, de los datos de y El segundo término del miembro derecho es precisamente la variancia de la predicción, $S_{y|x}^2$, conocida también como *variancia inexplicada*. El primer término del mismo miembro se denomina *variancia explicada*, y se representa con el símbolo $S_{\widetilde{y}}^2$ En consecuencia, se puede escribir

$$S_y^2 = S_{\widetilde{y}}^2 + S_{y|x}^2 \qquad (1\ 20)$$

Regresando al ejemplo de las cargas vivas observadas sobre los pisos de un edificio, la variancia explicada vale

$$S_{\widetilde{y}}^2 = S_y^2 - S_{y|x}^2 = 28\,561 - 27\,635 = 926 \text{ kg}^2$$

cuando se consideran los datos básicos, o

$$S_{\widetilde{y}}^2 = 28\,561 - 27\,439 = 1\,122 \text{ kg}^2$$

cuando se consideran agrupados en celdas.

Ejemplo

Calcular el error estándar de la estimación para los datos mencionados en los ejemplos 1 y 2 anteriores, si la ecuación de la recta de regresión se calculó

1 A partir de los datos básicos
2. A partir de los datos agrupados en celdas.

1. Para el primer caso, $S_y^2 = 28\,561\ kg^2$ y $r_{xy}^2 = 0\,0324$. Entonces

$$S_{y|x}^2 = S_y^2\,(1 - r_{xy}^2) = 28\,561\,(1 - 0\,0324) = 27\,635\ kg^2$$

y

$$S_{y|x} = \sqrt{S_{y|x}^2} = \sqrt{27\,635} = 166\,2\ kg$$

2. En el segundo caso, como $S_x^2 = 17\,956\ kg^2$ y $m^2 = 0.0625$, resulta

$$S_{y|x}^2 = S_y^2 - m^2\,S_x^2 = 28\,561 - 0.0625\,(17\,956) = 27\,439\ kg^2$$

y de la ec 1.19

$$S_{y|x} = \sqrt{S_{y|x}^2} = \sqrt{27\,439} = 165.6\ kg$$

Se aprecia que para los mismos datos el error de la estimación varía ligeramente, esto se debe a que para las ecuaciones de las rectas de regresión encontradas anteriormente, en un caso se usaron los datos básicos y en otro los agrupados en celdas

Ya se dijo que la diferencia $y_i - \tilde{y}_i$ representa la desviación de una ordenada observada respecto a su ordenada predicha mediante la recta de regresión. Existe otro tipo de desviación: la de las ordenadas predichas mediante la recta de regresión, $\tilde{y}_i$, respecto al promedio aritmético, $\bar{y}$, de las ordenadas observadas, y_i. Esta desviación, indicada como $\tilde{y}_i - \bar{y}$, se llama *desviación explicada*, ya que la ecuación

$$\tilde{y}_i - \bar{y} = m\,(x_i - \bar{x})$$

la cual se puede demostrar algebraicamente, indica que las desviaciones de $\tilde{y}_i$ respecto a $\bar{y}$ se explican exclusivamente por la desviación de x_i respecto a $\bar{x}$.

1.5.3 Variancia de la estimación

Como ya se indicó, el término $y_i - \widetilde{y}_i$ representa la diferencia entre el valor observado de la variable y y el valor predicho (la ordenada de la recta de regresión) correspondiente a x_i. Dicho término se llama *error de predicción* o *de estimación*. Por ejemplo, si para $x_3 = 50$ se observa que $y_3 = 65$, y la ecuación de la recta de regresión es $y = 0.7\ x + 21.9$, el valor predicho resulta $\widetilde{y}_3 = 0.7 \times 50 + 21.9 = 56.9$, y el error de predicción correspondiente es $65 - 56.9 = 8.1$.

La *variancia de la predicción* o *de la estimación*, $S^2_{y|x}$, que es una estimación global del error de predicción para todos los puntos observados, se calcula por medio de la fórmula

$$S^2_{y|x} = \frac{\sum_{i=1}^{i=n} (y_i - \widetilde{y}_i)^2}{n} \qquad (1.16)$$

Además, se puede demostrar que $S^2_{y|x}$ se relaciona con la pendiente de la recta de regresión mediante la ecuación

$$S^2_{y|x} = S^2_y - m^2\ S^2_x \qquad (1.17)$$

Recordando que la relación entre m y r_{xy} es $m = r_{xy} \dfrac{S_y}{S_x}$, la variancia de la estimación se puede expresar como

$$S^2_{y|x} = S^2_y\ (1 - r^2_{xy}) \qquad (1.18)$$

Puesto que la ecuación $y = mx + b$ se obtiene mediante el método de mínimos cuadrados, es decir, si $\sum_{i=1}^{i=n} (y_i - \widetilde{y}_i)^2$ tiene el mínimo valor posible, y como la variancia de la predicción se calcula con la ec 1.16, las predicciones basadas en la recta de mínimos cuadrados son tales que la variancia de la predicción es mínima.

Igual que la desviación estándar de una muestra se define como la raíz cuadrada de la variancia, el *error estandar de la estimación* (o de la predicción), $S_{y|x}$, se define como la raíz cuadrada de la variancia de la estimación, es decir

$$S_{y|x} = \sqrt{S^2_{y|x}} \qquad (1.19)$$

Table 17.5.2 Deformation x (in Millimeters) and Brinell Hardness y (in kilograms/millimeter²) of a Certain Type of Steel (Type 556-5) (K. Schimz, *Industr. Organisation*, 26, 1957, 107)

Deformation x_j (in millimeters)	Brinell Hardness y_j (in kilograms/millimeter²)
6	68
9	67
11	65
13	53
22	44
26	40
28	37
33	34
35	32

Table 17.5.2 was taken. The corresponding graphical representation (Fig. 17.5.1) shows that we may regard the regression curve corresponding to the regression of the Brinell hardness Y on the deformation x as a straight line. (A corresponding test for linearity will be discussed later.) ~~We suppose that the [illegible] assumptions (A1) and (A2), Sec. 17.4, are satisfied, so that we may use the procedure in Table 17.5.1 for determining a confidence interval for the regression coefficient β_1.~~

From the given data we obtain

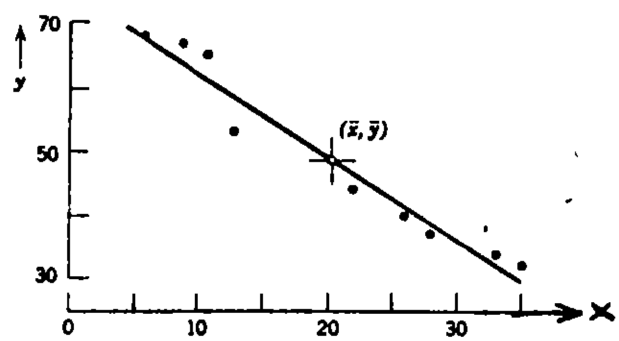
$$\sum x_j = 183 \quad \text{and} \quad \sum y_j = 440$$

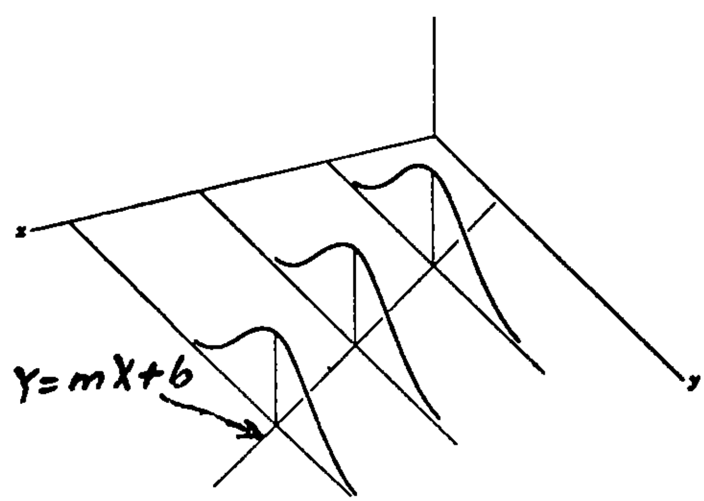
Since $n = 9$, it follows that

$$\bar{x} = \frac{183}{9} = 20.33 \quad \text{and} \quad \bar{y} = \frac{440}{9} = 48.89$$

Furthermore,

$$\sum x_j^2 = 4665, \quad \sum x_j y_j = 7701, \quad \sum y_j^2 = 23{,}232.$$





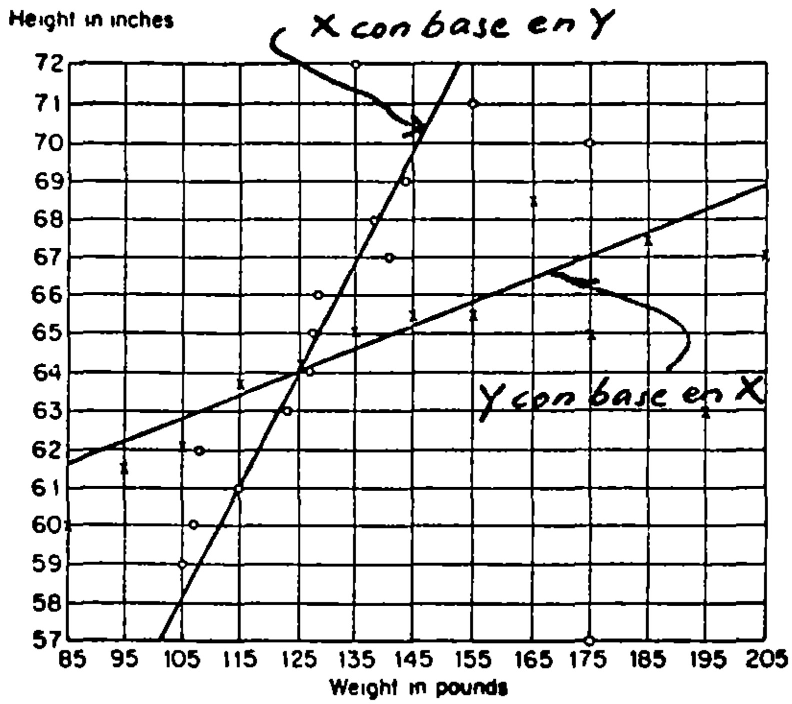


Table 17.2.1 Regression of the Iron Content Y (Percent) of Iron Ore on the Density x (Grams/centimeter³) (H. Bottke, *Bergbauwiss.*, 10, 1963, 377)

	Sample Values		Auxiliary Values	
	x_j	y_j	x_j^2	x_jy_j
	2.8	27	7.84	75.6
	2.9	23	8.41	66.7
	3.0	30	9.00	90.0
	3.1	28	9.61	86.8
	3.2	30	10.24	96.0
	3.2	32	10.24	102.4
	3.2	34	10.24	108.8
	3.3	33	10.89	108.9
	3.4	30	11.56	102.0
Sum	28.1	267	88.03	837.2

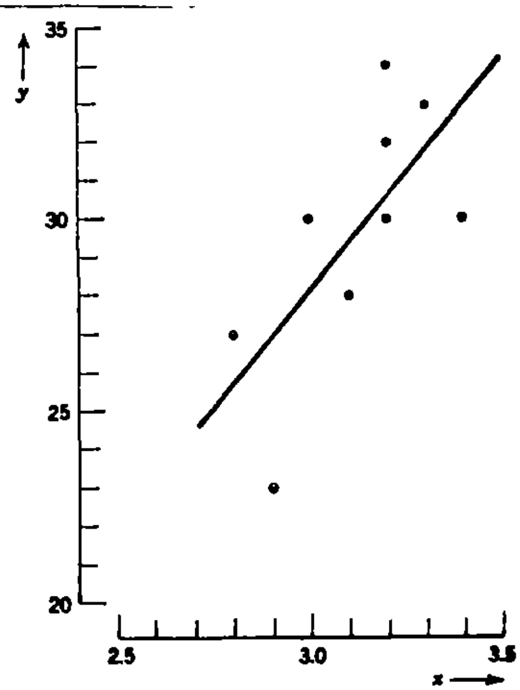


Fig. 17.2.1. Sample values and regression line in Ex. 1

and from this the regression coefficient

$$b = \frac{0.4458}{0.03694} = 12.07$$

Hence the regression line has the representation (cf. Fig. 17.2.1)

$$y - 29.67 = 12.07(x - 3.12)$$

This may also be written

$$y = 12.07x - 7.99$$

muestra en la tabla 4

Puesto que en este caso los datos se encuentran agrupados en celdas, la pendiente, m, de la recta de regresión se debe calcular con la fórmula

$$m = \frac{(1/n) \sum_{i=1}^{i=k} f_{ixy} \; (x_i - \bar{x}) \; (y_i - \bar{y})}{S_x^2} \tag{1.15}$$

Las marcas de clase para los intervalos tratados son 50 5, 150.5, etc

La suma (desde la celda 1 hasta la 60) de los productos de las frecuencias por las desviaciones respecto a los promedios $\bar{x}$ y $\bar{y}$ resulta

$$(1/n) \sum_{i=1}^{i=k} f_{ixy} \; (x_i - \bar{x}) \; (y_i - \bar{y}) = 4\,507 \text{ kg}^2$$

Como $S_x = 134$ kg, entonces $S_x^2 = 17\,956$ kg^2, siendo la pendiente

$$m = \frac{4\,507 \text{ kg}^2}{17\,956 \text{ kg}^2} = 0\,25$$

El paso siguiente es calcular el valor de la ordenada al origen, b

$$b = \bar{y} - m\bar{x} = 368\,3 - 0\,25 \; (277.6) = 298.9$$

Finalmente, la recta de regresion pedida resulta

$$\tilde{y} = 0\,25 \; x + 298\,9$$

cuya grafica aparece en la fig 6

Obsérvese que existe una diferencia entre los valores de los parametros b y m de las rectas de regresion calculados para los mismos datos en los dos ejemplos anteriores. Esto se debe a que en el primer caso el tratamiento se efectuó en forma directa con los datos basicos, haciendo uso de las desviaciones (diferencias) de cada uno respecto al promedio aritmético de los mismos, y en el segundo a que las desviaciones correspondieron a las marcas de clase de las celdas en las que se agruparon los datos.

Ejemplos

1. Obtener la ecuación de la recta de regresión de y con base en x para los datos mencionados en el problema 2 de 1 5 1, referente a las cargas vivas observadas en los pisos de un edificio, teniendo en cuenta que los valores calculados a partir de los datos básicos fueron $\bar{x} = 277.6$ kg, $\bar{y} = 368\ 3$ kg, $S_x = 134$ kg, $S_y = 169$ kg, y $r_{xy} = 0\ 18$

Aplicando las ecs 1 11 y 1.12, la pendiente y la ordenada al origen resultan

$$m = r_{xy} \frac{S_y}{S_x} = (0\ 18) \frac{169}{134} = 0\ 22$$

$$b = \bar{y} - m\bar{x} = 368\ 3 - 0.22\ (277\ 6) = 307\ 3$$

Entonces, de acuerdo con la ec 1 10, la ecuación de la recta de regresión pedida es

$$\tilde{y} = 0\ 22\ x + 307.3$$

cuya gráfica se presenta en la fig 6.

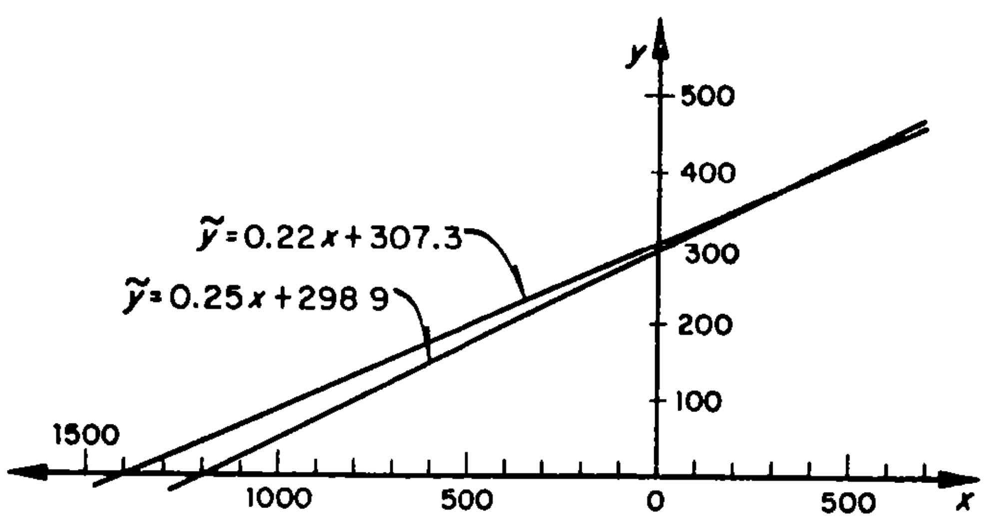


Fig 6 Rectas de regresión para cargas vivas

2. Obtener la ecuación de la recta de regresión de y con base en x para los datos del ejemplo anterior si la distribución conjunta de frecuencias de los mismos es la que se

El método de mínimos cuadrados se basa en suponer que la mejor recta de regresión es aquella para la cual la suma de los cuadrados de las desviaciones de todos los puntos con respecto a la misma, $\sum_{i=1}^{i=n} (y_i - \tilde{y}_i)^2$, es mínima

Se puede demostrar que para datos no agrupados en celdas, la pendiente, m, y la ordenada al origen, b, de la recta de regresión, según el método de mínimos cuadrados, se calculan con las ecuaciones

$$m = \frac{(1/n) \sum_{i=1}^{i=n} (x_i - \bar{x})(y_i - \bar{y})}{S_x^2} \qquad (1.11)$$

y

$$b = \bar{y} - m\bar{x} \qquad (1.12)$$

En estas fórmulas, $\bar{x}$ y $\bar{y}$ denotan los promedios aritméticos de los datos de x y y, respectivamente, y n es el total de parejas de puntos.

Para el caso en el que los datos se encuentren agrupados en celdas, la pendiente de la recta de regresión se calcula mediante la fórmula

$$m = \frac{(1/n) \sum_{i=1}^{i=k} f_{ixy} (x_i - \bar{x})(y_i - \bar{y})}{S_x^2} \qquad (1.13)$$

donde k es el número de celdas definidas por las intersecciones de los intervalos de clase de x y y, x_i y y_i son las marcas de clase de dichos intervalos, y f_{ixy} es la frecuencia de la i-ésima celda

Es posible relacionar el coeficiente de correlación con la pendiente de la recta de regresión si se toma en cuenta que $S_x^2 = S_x S_x$, y que si se divide y multiplica un miembro de una ecuación por un mismo número (en este caso S_y), la ecuación no se altera. Entonces

$$m = \frac{(1/n) \sum_{i=1}^{i=n} (x_i - \bar{x})(y_i - \bar{y})}{S_x^2} = \frac{(1/n) \sum_{i=1}^{i=n} (x_i - \bar{x})(y_i - \bar{y})}{S_x \; S_x} \; \frac{S_y}{S_y} =$$

$$= \frac{(1/n) \sum_{i=1}^{i=n} (x_i - \bar{x})(y_i - \bar{y})}{S_x \; S_y} \; \frac{S_y}{S_x} = r_{xy} \; \frac{S_y}{S_x} \qquad (1.14)$$

2. Calcule el coeficiente de correlación para los datos de las cargas vivas de los pisos 1 y 9 que se presentan en la tabla 1 Considere que las x_i corresponden a los datos del primer piso, y las y_i a los del noveno

1.5.2 Regresión lineal

En muchas ocasiones, los puntos registrados en una grafica de correlación tienden a situarse a lo largo de una línea recta llamada *recta de regresión*, de ecuación general

$$y = mx + b \qquad (1\ 10)$$

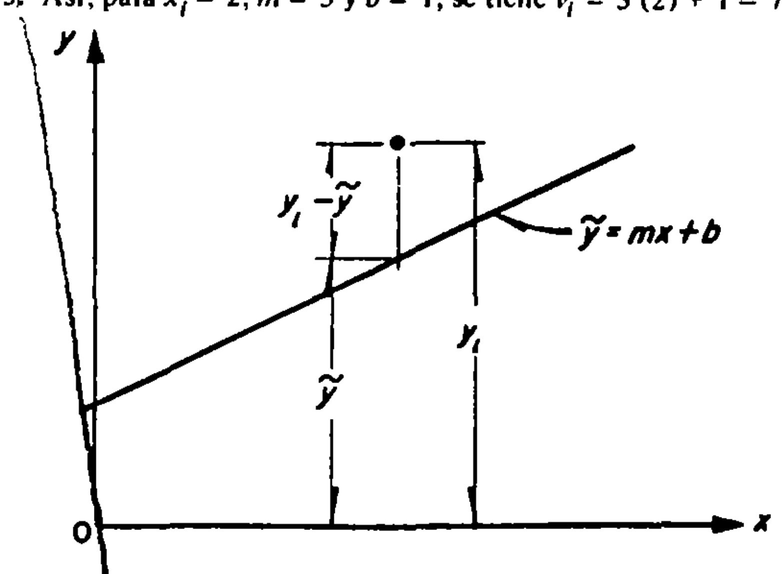
donde m es la pendiente y b la ordenada al origen

En la ecuacion anterior, la variable x sirve de base para encontrar el valor de y, es decir, es la variable independiente Sin embargo, es posible describir a la variable x en términos de y, siendo esta ultima la variable independiente Para este último caso, la ecuación de la recta de regresion toma la forma

$$x = m'y + b'$$

En cualquiera de los dos casos anteriores, la variable cuyo valor se desea encontrar se llama *variable dependiente*

El metodo que se presenta en seguida para calcular la ecuación de la recta de regresion es el llamado *método de minimos cuadrados*, el cual se establece aquí solo para el caso de regresión de y con base en x. En el caso de regresión de x con base en y, el tratamiento es similar, intercambiando las letras x y y En este método se emplea la *desviación respecto a la recta de regresión*, $y_i - \tilde{y}_i$, donde $\tilde{y}_i$ es el valor obtenido de la ecuación de la recta de regresión para $x = x_i$, es decir, $\tilde{y}_i = mx_i + b$, tal como se puede apreciar en la fig 5. Así, para $x_i = 2$, $m = 3$ y $b = 1$, se tiene $\tilde{y}_i = 3\ (2) + 1 = 7$



237 Si cometió algún error corrija su respuesta en la hoja de trabajo 8.12 y luego calcule ρ_{xy} mediante el método corto.

Y \ X	70–75			76–81			82–87			88–93			94–99			f_y	Y'	$f_y Y'$	Y'^2	$f_y Y'^2$	$\sum f_{ixy} X'Y'$
41–50										−2	1	−2				1	−2	−2	4	4	−2
51–60							0	2	0				−2	2	−4	4	−1	−4	1	4	−4
61–70				0	2	0	0	3	0	0	5	0	0	1	0	11	0	0	0	0	0
71–80	−2	2	−4	−1	3	−3	0	3	0	1	7	7	2	2	4	17	1	17	1	17	4
81–90				−2	4	−8	0	1	0	2	3	6	4	2	8	10	2	20	4	40	6
91–100	−6	2	−12	−3	3	−9	0	1	0	3	1	3				7	3	21	9	63	−18
f_x	4			12			10			17			7			50		52		128	−14
X'	−2			−1			0			1			2								
$f_x X'$	−8			−12			0			17			14			11					
X'^2	4			1			0			1			4								
$f_x X'^2$	16			12			0			17			28			73					
$\sum f_{ixy} X'Y'$			−16			−20			0			14			8	−14					

$\overline{X'} = 0.22$; $\overline{Y'} = 1.04$; $S_{x'} = 1.19$; $S_{y'} = 1.22$

$$\sum f_{ixy} X'Y' = 14; \quad \rho_{xy} = \frac{\frac{-14}{50} - 1.04 \times 0.22}{1.19 \times 1.22} = \frac{-0.28 - 0.23}{1.45} = -0.35$$

f_x
X'
$f_x X'$
X'^2
$f_x X'^2$
$\sum f_{1xy} X'Y'$

235 Observe la tabla 8.3 correspondiente a las calificaciones obtenidas por 50 alumnos en un examen de Estadística y a los tiempos que emplearon para resolverlo.
Los datos se presentan por parejas siguiendo un orden decreciente.
Obtenga, en la hoja de trabajo 8.12, la distribución conjunta de frecuencias.

→ Calificación

Tiempo, min ↓

Y \ X	70–75	76–81	82–87	88–93	94–99
41–50				1	
51–60			2		2
61–70		2	3	5	1
71–80	2	3	3	7	2
81–90		4	1	3	2
91–100	2	3	1	1	

$$\rho_{xy} = 0.74; \left[\frac{\frac{63}{30} - 0.7\,(-0.27)}{1.81 \times 1.71} = 0.74 \right]$$

233 Complete los encabezados de las siguientes columnas, las cuales se usan en la tabulación del método corto para calcular el coeficiente de correlación.

f_y	Y'	$f_y Y'$	Y'^2	$f_y Y'^2$	$\sum f_{ixy} X'Y'$

234 Complete los encabezados para los siguientes renglones, los cuales se utilizan en el método corto para calcular ρ_{xy}.

f_x

TABLA 8.3

Calificación, X	Tiempo, Y, en min.	Calificación, X	Tiempo, Y, en min.
97	77	87	83
97	86	87	58
95	60	87	79
95	52	86	60
94	62	85	62
94	86	83	72
94	80	82	68
93	79	82	66
93	92	82	71
93	88	81	70
92	74	80	65
92	43	80	84
92	61	79	82
92	75	79	93
91	79	79	76
90	62	78	71
90	81	77	89
90	80	77	71
90	76	77	98
90	70	76	92
89	67	76	82
88	69	74	98
88	81	72	78
88	83	79	93
87	91	70	78

UNIVERSIDAD NACIONAL AUTÓNOMA

en tal caso

$$\bar{X}' = \frac{\bar{X} - c_1}{c_2} \; ; \; \bar{Y}' = \frac{\bar{Y} - c_3}{c_4}$$

$$S^2_{X'} = S^2_X / C^2_2 \; ; \; S^2_{Y'} = S^2_Y / C^2_4$$

y

$$\boxed{\rho_{xy} = \frac{\sum_{i=1}^{N} (X'_i - \bar{X}')(Y'_i - \bar{Y}')}{N\, S_{X'}\, S_{Y'}} = \frac{\frac{1}{N} \sum_{j=1}^{k} f_{j\,xy}\, X'_j Y'_j - \bar{X}' \bar{Y}'}{S_{X'}\, S_{Y'}}}$$

En términos de datos básicos

En términos de datos agrupados

donde:

N = número de parejas de datos

k = número de celdas

Si los datos están agrupados:

$$S_{XY} = \frac{1}{N} \sum_{j=1}^{K} f_{jxy} (X_j - \bar{X})(Y_j - \bar{Y})$$

donde f_{jxy} es la frecuencia de la celda j, y X_j y Y_j son las marcas de clase de X y Y, respectivamente, correspondientes a la celda j; K es el número de celdas.

METODO CORTO PARA CALCULAR EL COEFICIENTE DE CORRELACION

Sean las transformaciones

$$X' = \frac{X - C_1}{C_2} \quad y \quad Y' = \frac{Y - C_3}{C_4}$$

donde

C_1 = constante con valor cercano a $\bar{X}$ e igual a alguna marca de clase

C_2 = ancho de los intervalos de X

C_3 = Constante con valor igual a alguna marca de clase cercana a $\bar{Y}$

C_4 = ancho de los intervalos de Y

Calificación, X	Tiempo, Y, min
4	100
5	94
6	90
7	83
8	80
9	75
10	71

$S_x = 2$, $\bar{x} = 7$

$S_y = 9.7$ min

$\bar{Y} = 84.7$ min

X	Y	$X-\bar{X}$	$Y-\bar{Y}$	$(X-\bar{X})(Y-\bar{Y})$
4	100	−3	15.3	−45.9
5	94	−2	9.3	−18.6
6	90	−1	5.3	−5.1
7	83	0	−1.7	0
8	80	1	−4.7	−4.7
9	75	2	−9.7	−19.4
10	71	3	−13.7	−41.1
				$\Sigma = -134.8$

$$S_{XY} = \frac{-134.8}{7} = -19.26; \quad \rho_{XY} = \frac{-19.26}{2 \times 9.7} =$$

$$= -0.99$$

BLOCK TEYCO

conclusion cualitativa a que puede llegarse al observar la grafica de correlacion de la fig 4, ya que los puntos se localizan en una franja casi vertical

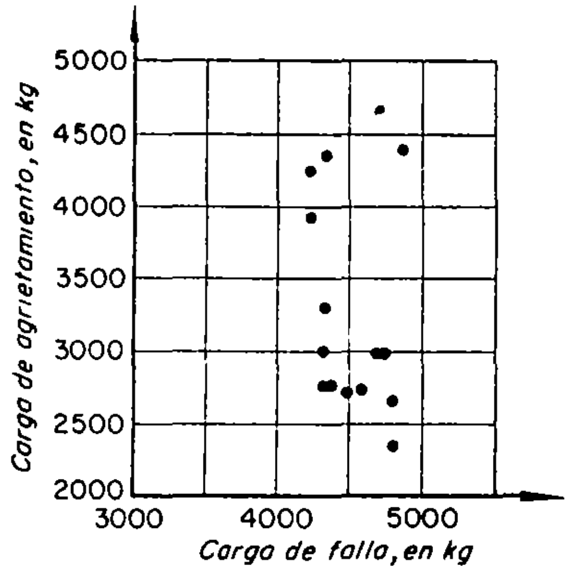


Fig 4 Grafica de correlacion de los datos en las vigas

Como ejercicio, es conveniente que el lector resuelva los siguientes problemas

1 Diez vigas de madera de 2 m de largo se probaron con una carga concentrada en el centro del claro Los resultados fueron

Espécimen	Deflexión máxima al centro del claro, en cm	Carga de falla, en kg
1	0 330	950
2	0 330	1 050
3	0 335	750
4	0 313	900
5	0 343	700
6	0 426	650
7	0 340	950
8	0 406	850
9	0 380	650
10	0 393	900

Calcule el promedio y la variancia, y elabore histogramas y curvas de distribucion de frecuencias acumuladas para cada grupo de datos

$$\rho_{xy} = \frac{S^2_{xy}}{S_x\ S_y} = \frac{1}{n} \sum_{i=1}^{i=n} \frac{(x_i - \bar{x})}{S_x} \frac{(y_i - \bar{y})}{S_y} \qquad (1\ 9)$$

Se puede demostrar que para ρ_{xy} se cumple la desigualdad

$$-1 < \rho_{xy} < 1$$

Si en la grafica de correlación los puntos tienden a agruparse en una franja que asciende de izquierda a derecha (fig 3a), entonces $\rho_{xy} > 0$, pero si por lo contrario, asciende de derecha a izquierda, entonces $\rho_{xy} < 0$ (fig 3b) Si la franja es horizontal o vertical, entonces $\rho_{xy} = 0$

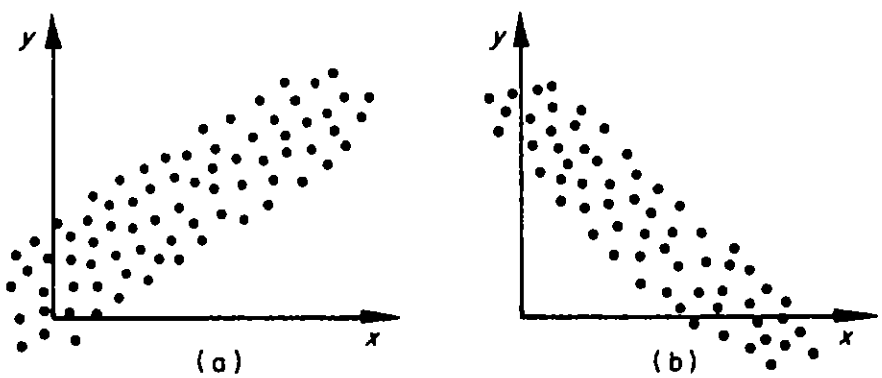


Fig 3 Agrupamiento de puntos en la gráfica de correlación

Si $\rho_{xy} = \pm 1$, se dice que los parámetros están *perfectamente correlacionados*, en cambio si $\rho_{xy} = 0$, se dice que *no están correlacionados*

Para los datos de las vigas del ejemplo que se ha venido empleando se tiene $\bar{x} = 4\,490\ 7$ kg, $\bar{y} = 3\,321\ 3$ kg, $S_x = 199.1$ kg y $S_y = 797.3$ kg, por lo cual

$$S^2_{xy} = (1/n) \sum_{i=1}^{i=n} (x_i - 4\,490.7)\ (y_i - 3\,321\ 3) = -\ 5\,054\ 22\ \text{kg}^2$$

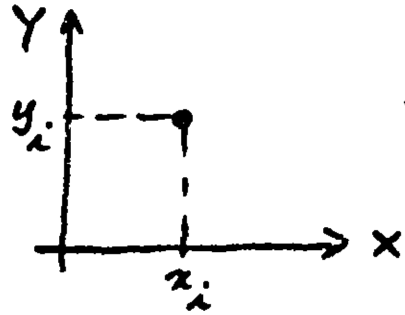
y

$$\rho_{xy} = -\ \frac{5\,054\ 22}{(199\ 1)\ (797\ 3)} = -\ 0\ 032$$

El valor pequeño de ρ_{xy} indica que las cargas de agrietamiento, x, y las de falla, y, casi no están correlacionadas, o que son prácticamente independientes Esto concuerda con la

MEDIDAS DE CORRELACION

$$\text{Covariancia} = S_{xy} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{N}$$



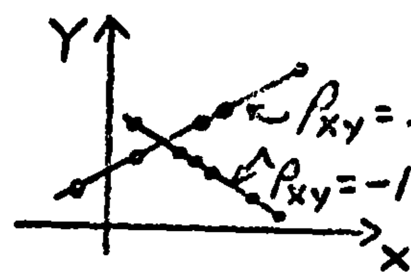
$\bar{y}$ = promedio aritmético de los datos de Y.

$\bar{x}$ = promedio aritmético de los datos de X.

$$\text{Coeficiente de correlación} = \rho_{xy} = \frac{S_{xy}}{S_x\, S_y}$$

S_x = desviación estándar de los datos de X.

S_y = desviación estándar de los datos de Y.



Siempre: $-1 \leq \rho_{xy} \leq 1$

Si $\rho_{xy} = 1$ ó -1 se tiene correlación perfecta (todos los puntos (x_i, y_i) quedan sobre la recta)

Table 13-2

CORRELATION TABLE FOR THE WEIGHTS AND HEIGHTS OF 285 BOSTON UNIVERSITY WOMEN STUDENTS (*Original data*)

Weights in Pounds

Heights in Inches

$x \rightarrow$ / $y \downarrow$	85	95	105	115	125	135	145	155	165	175	185	195	205
57										1			
58													
59	1		4	2									
60		3	8	1				1					
61	1	3	2	4	5	3							
62		3	7	8	11	3	2						
63			7	12	13	6	1	1				1	
64			8	8	14	10	6	2			1		
65			1	10	15	11	4			1			
66		1	2	9	10	6	5	2		1			
67				2	4	8	5	3		1			1
68				1	2	2	2	1	1				
69						4	1	1	1				
70										1			
71					1						1		
72						1							

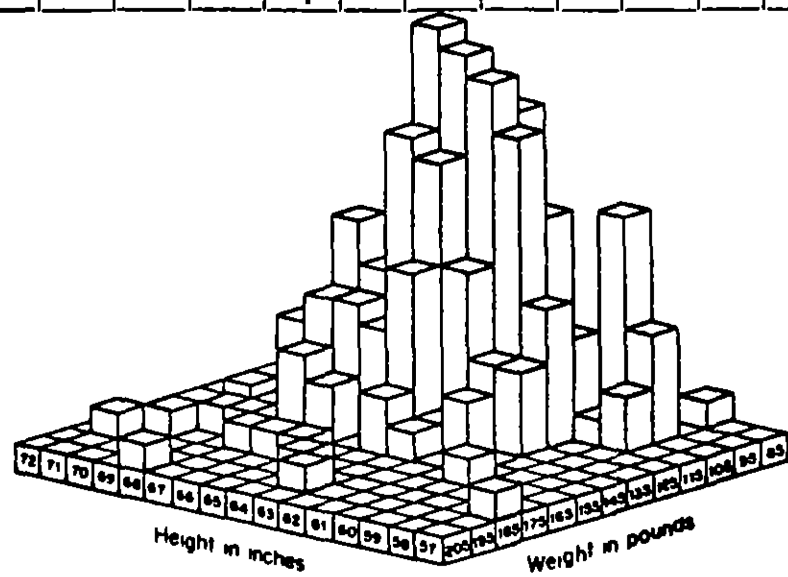


Fig 13-6 Solid histogram for the Correlation Table 13-2, of heights and weights of 285 Boston University women students.

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y(1-\rho^2)^{1/2}} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]} \quad (13.25)$$

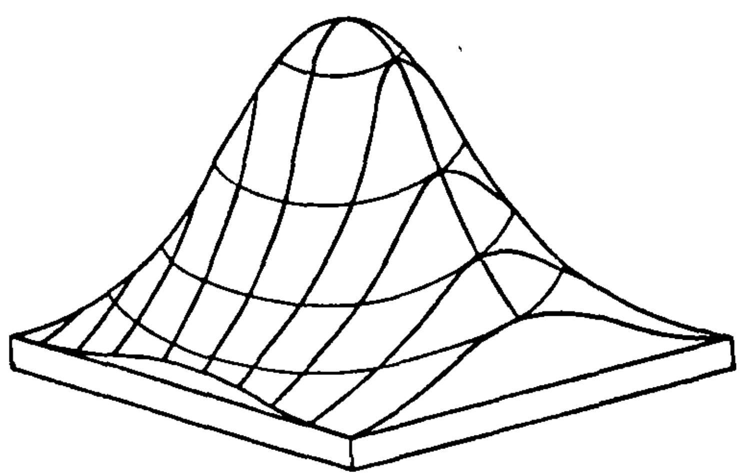


Fig 13-7 A normal frequency surface

we assume that the relationship is *linear* and we proceed to find the equation of a specific straight line which "best fits" the data. This best-fitting line is called the *regression line.**

Table 5 1 Verbal Aptitude Scores (x) and Mathematics Achievement Scores (y) for a Sample of 15 Individuals

	x_i	y_i
1	80	60
2	75	62
3	95	78
4	50	40
5	65	58
6	55	65
7	58	52
8	87	72
9	76	70
10	40	50
11	68	78
12	35	40
13	55	48
14	74	80
15	98	88

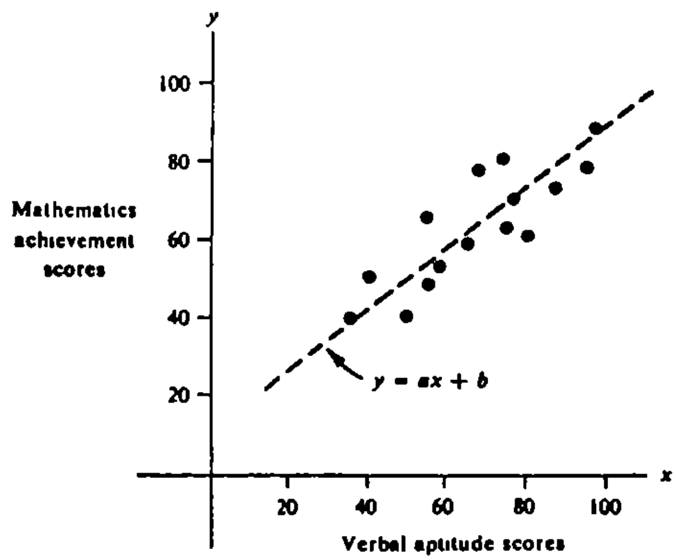


Figure 5 4 Scatterplot verbal aptitude (x) vs mathematics achievement (y)

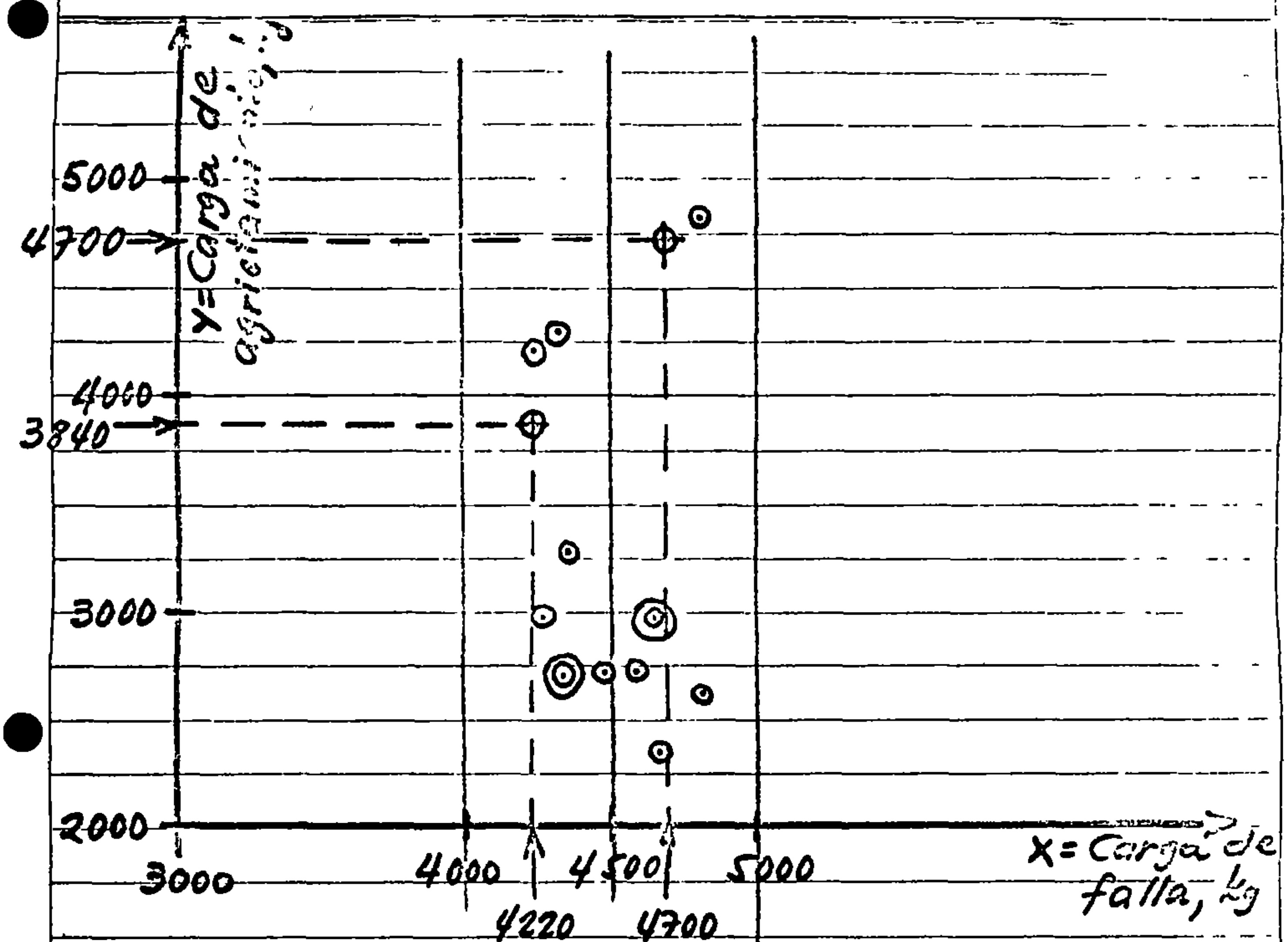
*The term "regression" is somewhat of a misnomer See Walker and Lev (1969, page 210) for an explanation of the origin of the term

TABLA 4. DISTRIBUCION CONJUNTA DE FRECUENCIAS DE LAS CARGAS EN LOS PISOS 1 y 9

→ Cargas en el 9º piso, en ton

Cargas en el 1er piso, en ton ↓

X \ Y	0.5 a 100.5	100.5 a 200.5	200.5 a 300.5	300.5 a 400.5	400.5 a 500.5	500.5 a 600.5	600.5 a 700.5	700.5 a 800.5	800.5 a 900.5	900.5 a 1000.5
0.5 a 100.5				X (1)						
100.5 a 200.5		X (1)	X (1)	X (1)	XX (2)					
200.5 a 300.5		X (1)	XXXX (4)	X (1)			X (1)			
300.5 a 400.5			X (1)	XXXX (4)						X (1)
400.5 a 500.5				X (1)	XX (2)					
500.5 a 600.5				X (1)	XX (2)					



GRAFICA DE CORRELACION DE CARGA DE AGRIETAMIENTO Vs. CARGA DE FALLA (Datos tomados de la tabla 2)

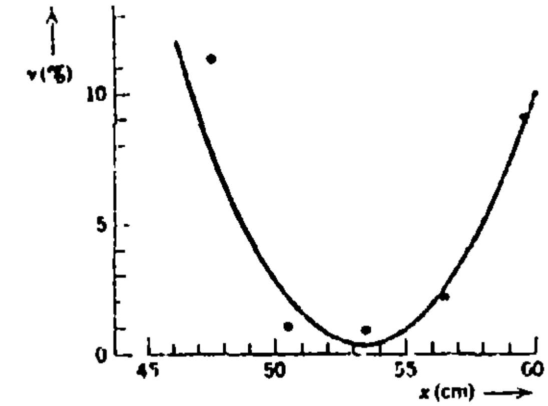
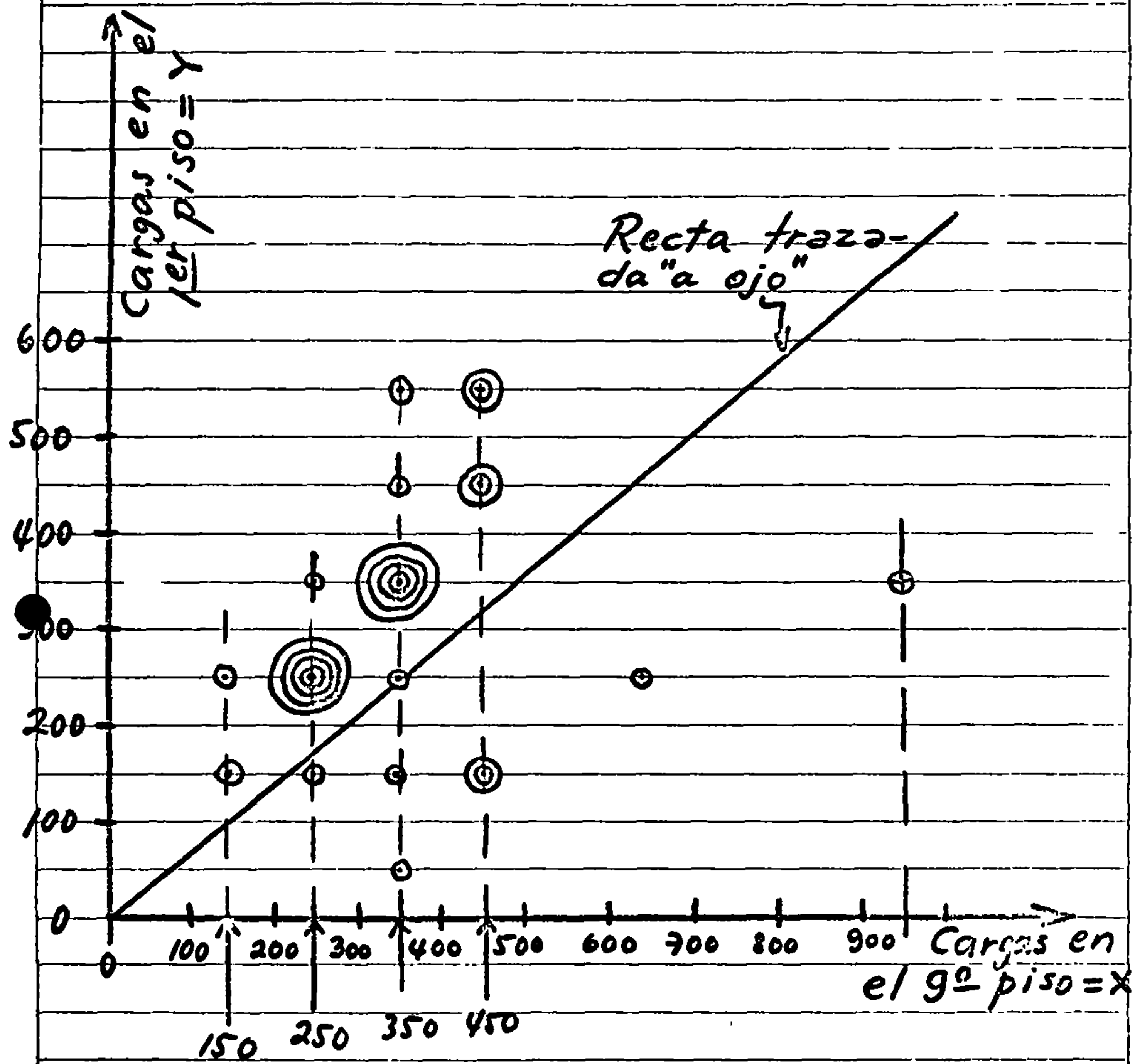


TABLA 2. PRUEBAS DE VIGAS DE CONCRETO REFORZADO

Número de la viga	Carga de agrietamiento, en kg	Carga de falla, en kg
1	4 700	4 700
2	3 840	4 220
3	3 270	4 360
4	2 310	4 680
5	2 950	4 270
6	4 810	4 810
7	2 720	4 590
8	2 720	4 490
9	4 310	4 310
10	2 950	4 630
11	4 220	4 220
12	2 720	4 340
13	2 720	4 340
14	2 630	4 770
15	2 950	4 630

TABLA 3. DATOS ORDENADOS DE CARGAS DE FALLA EN VIGAS DE CONCRETO REFORZADO

Carga de agrietamiento, en kg	Carga de falla, en kg
2 310	4 220
2 630	4 220
2 720	4 270
2 720	4 310
2 720	4 340
2 720	4 340
2 950	4 360
2 950	4 490
2 950	4 590
3 270	4 630
3 840	4 630
4 220	4 680
4 310	4 700
4 700	4 770
4 810	4 810



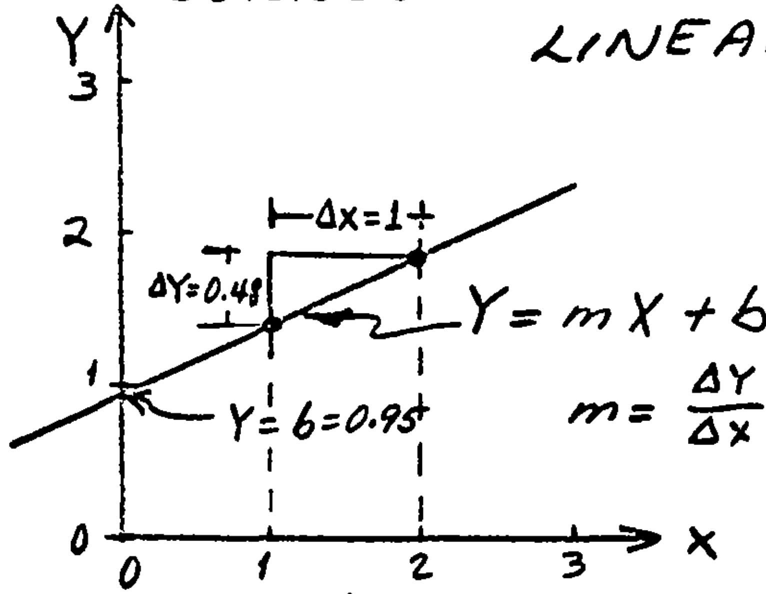
GRAFICA DE CORRELACION DE LAS CARGAS VIVAS DEL 1er PISO Vs. LAS DEL 9º PISO (Datos tomados de la tabla 4.) (Los datos básicos son los de la tabla 1.)

TABLA 1. CARGAS VIVAS OBSERVADAS EN UN EDIFICIO, en kg/m²

Crujía	PISOS									
	Sótano	1o	2o	3o	4o	5o	6o	7o	8o	9o
A	1	38	177	296	312	313	386	432	185	355
B	352	354	314	106	83	237	82	517	279	370
C	1 102	207	292	203	195	271	328	599	221	307
D	269	273	428	289	308	287	331	441	211	270
E	179	127	442	112	212	254	499	350	20	182
F	632	324	677	620	444	229	964	738	768	962
G	657	358	395	760	391	307	736	499	31	222
H	591	519	461	682	745	343	546	850	417	405
I	873	147	215	767	514	425	245	967	423	315
J	384	181	345	405	877	880	296	1 037	353	420
K	461	118	426	393	365	353	640	567	262	484
L	196	114	41	208	212	134	311	90	79	287
M	450	243	248	573	606	646	514	783	140	228
N	486	522	273	668	539	603	451	786	222	470
O	432	236	304	348	650	450	545	332	259	194
P	455	269	394	701	597	899	732	281	33	260
Q	469	268	307	1 115	680	289	547	248	774	679
R	1 043	321	441	969	476	757	798	758	1 121	366
S	672	305	764	752	656	398	949	757	436	358
T	390	335	418	691	492	518	640	769	392	317
U	386	577	617	165	608	385	713	490	478	368
V	121	271	662	275	327	352	515	483	497	284

UNIVERSIDAD NACIONAL AUTÓNOMA

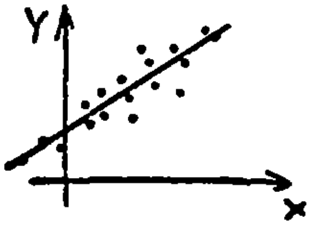
CORRELACION Y REGRESION LINEAL



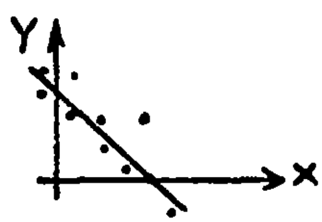
$$m = \frac{\Delta Y}{\Delta X} = \frac{0.48}{1} = 0.48$$

m = pendiente

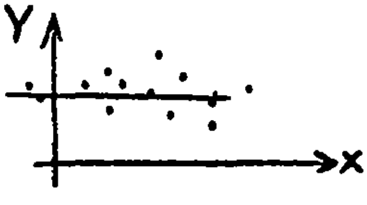
b = ordenada al origen



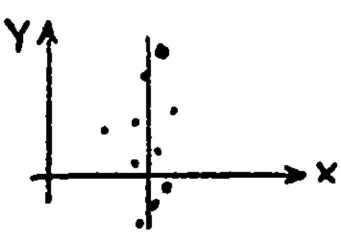
Pendiente positiva ($m > 0$)
Correlación positiva (S_{xy} y $\rho_{xy} > 0$)



Pendiente negativa ($m < 0$)
Correlación negativa (S_{xy} y $\rho_{xy} < 0$)



Pendiente nula ($m = 0$)
Correlación nula (S_{xy} y $\rho_{xy} = 0$)



Pendiente infinita ($m = \infty$
Correlación nula (S_{xy} y $\rho_{xy} = 0$)

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

PROBABILIDAD Y ESTADISTICA. FUNDAMENTOS Y APLICACIONES

ELEMENTOS DE LA TEORIA ESTADISTICA DE DECISIONES

DR. OCTAVIO A. RASCON CHAVEZ

MARZO DE 1976.

Palacio de Minería
Tacuba 5, primer piso. México 1, D. F.
Tels 521-40-23 521-73-35 5123-123

CAPITULO 6

Por Octavio A. Rascón Ch. *

ELEMENTOS DE LA TEORIA ESTADISTICA DE DECISIONES

6.1 *Introducción*

Una de las ramas de la estadística es la *inferencia estadística*; en ella se utiliza únicamente la información que proporciona una muestra de la población bajo estudio para inferir algunas de sus características, tales como la media, la variancia, etc. En el proceso de inferir alguno de estos parámetros suele realizarse una *prueba de hipótesis*, en la cual se *toma la decisión* de adoptar una *acción*: aceptar o rechazar la hipótesis nula con un cierto nivel de confianza. Por ejemplo, después de probar una muestra de un material manufacturado, tal como el acero de refuerzo para concreto, podemos inferir el porcentaje de elementos que no cumplen con la especificación bajo prueba; sin embargo, lo que más interesa es contar con un criterio o *regla de decisión* para aceptar o rechazar, con base en una muestra, un lote de dicho producto.

Situaciones como ésta, y aún más generales, tiene que afrontar a menudo un economista o un ingeniero en su práctica profesional; ejemplos de ellas son:

i) Seleccionar las dimensiones de un vertedor de demasías de una presa, tomando en cuenta que el flujo futuro máximo no se conoce con certeza.

ii) Decidir a qué planta premezcladora de concreto hacer un pedido para construir la carpeta de una carretera,

* Jefe de la División de Estudios Superiores de la Facultad de Ingeniería, UNAM.

tomando en cuenta las diferencias en precio y control de calidad que se tienen en cada una.

iii) Decidir qué cantidad de cierto producto elaborado hay que tener en inventario, si tanto la demanda como la producción son aleatorias.

iv) Decidir la trayectoria de una carretera, tomando en consideración que no se conoce con certeza el potencial económico de los sitios que cruzará, y que su longitud y las distancias a los bancos de materiales serán diferentes en cada caso.

Un factor importante que debe considerarse al tomar una decisión es la evaluación, en una escala cuantitativa, de las *consecuencias* que se tendrían al adoptar las diversas acciones. Puesto que los responsables de tomar una decisión están generalmente interesados en los costos y beneficios (no necesariamente en escala monetaria), el criterio de decisión deberá basarse en el costo o beneficio esperado para cada acción.

En general, la persona que va a tomar una decisión deberá considerar los factores económicos, las diversas técnicas existentes para resolver el problema, las necesidades humanas, sociales o científicas, etc., para seleccionar la acción que optimice la esperanza del beneficio (o utilidad) derivado de la obra.

Resumiendo, los factores que se deben considerar durante el proceso de tomar una decisión son (ref 1):

i) *Un objetivo*. Sin un objetivo no hay nada acerca de qué

...

decidir. Los objetivos varían ampliamente de una situación a otra, y pueden ser tales como: costo inicial, de operación o de mantenimiento, confiabilidad de un sistema, utilidad anual, etc.

ii) *Diversos cursos de acción.* Siempre debe de haber más de una manera de lograr el objetivo, puesto que si sólo hay una no hay necesidad de escoger o de *decidir* por una de ellas. Las diversas acciones pueden tener diferentes costos o beneficios y diferentes probabilidades de tener éxito; el proceso de formular una lista de las diversas acciones o alternativas es un proceso de inventiva que debe desarrollar la persona que toma la decisión, con base en su experiencia o intuición. En ocasiones una de las acciones posibles es "posponer la decisión hasta tener mayor información acerca del problema" lo cual, evidentemente, no puede prolongarse en forma indefinida.

iii) *Factores relevantes.* Los factores relevantes que pueden intervenir en un proceso de decisión pueden ser iguales o diferentes para las diversas acciones, y pueden ser de carácter económico, tecnológico y humano. Dentro de los factores económicos se cuentan: disponibilidad de materiales, financiamiento, instalaciones, organización administrativa, equipo, etc.

Como factores tecnológicos podemos considerar aquellos directamente relacionados con el análisis y diseño de un sistema, tales como resistencia de los materiales, efectos dinámicos, térmicos, magnéticos, etc. Problemas de fatiga, corrosión o flujo plástico de los materiales, disponibilidad de combustible o corriente eléctrica, etc. Los factores humanos no sólo se relacionan con la práctica política o social para diseñar un sistema, sino también con la ética y moral del que decide, tales como prestigio profesional, preferencias y prejuicios personales, confort o bienestar familiar, etc.

6.2 *El modelo de decisión*

Para formular el modelo de decisión (ref. 2) tenemos que definir todos los elementos que intervendrán en él. Es necesario saber que las consecuencias de adoptar una acción dependen de un factor que no conocemos con certeza y que llamaremos *estado de la naturaleza*. Por ejemplo, al diseñar un edificio no conocemos exactamente cuál será el asentamiento total del mismo, de acuerdo con los diversos tipos de cimentación que se utilicen; tampoco conocemos con certeza el volumen máximo de agua que escurrirá hacia una presa que forma parte de un sistema de riego y de control de avenidas durante un período dado, etc. La manera de reconocer que no se sabe cuál es el verdadero estado de la naturaleza consiste en asignarle a cada uno una *probabilidad* de que ocurra.

El proceso de tomar una decisión consiste en escoger una *acción* o alternativa, a, de entre las diversas acciones posibles a_1, a_2, ..., a_n, que constituyen el *espacio de acciones*, A (el espacio de acciones es el conjunto de las diversas acciones, es decir,

...

$A = \{a_1, a_2, \ldots, a_n\}$), bajo la consideración de que no se sabe con certeza (sólo probabilísticamente), cuál de los posibles estados de la naturaleza $\theta_1, \theta_2, \ldots, \theta_m$, es el verdadero (el conjunto $\Omega = \{\theta_1, \theta_2, \ldots, \theta_m\}$ se denomina *espacio de estados de la naturaleza*). Una vez que se ha seleccionado y realizado una acción, con el transcurso del tiempo la persona que tomó la decisión podrá darse cuenta de cuál es el verdadero estado de la naturaleza, con lo cual obtendrá una utilidad $u(a,\theta)$; en otras palabras, $u(a,\theta)$ es la utilidad que se obtiene al escoger la acción a cuando el *verdadero* estado de la naturaleza es θ.

Este modelo de decisiones se muestra gráficamente en la fig. 6.1, mediante el llamado *árbol de decisiones*.

AQUI ENTRA FIG 6.1

Otra manera de representar gráficamente el proceso de decisiones se muestra en la fig 6.2; en ella, aparece en primer término el objetivo que se persigue al tomar la decisión acerca del sistema bajo estudio; en segundo lugar, se tiene la información - acerca del sistema, que incluye la lista de las acciones posibles, la lista de los estados de la naturaleza y sus probabilidades correspondientes, y la lista de los factores relevantes, en tercer lugar, aparece el criterio de decisión que se utilice y, finalmente, la acción óptima o más conveniente.

AQUI ENTRA FIG 6.2

Una manera usual de presentar las utilidades que corresponden a cada combinación acción-estado de la naturaleza consiste en una tabla como la 6.1; por ejemplo, en el renglón de a_2 y la columna de θ_m se tiene la utilidad $u(a_2, \theta_m)$.

Tabla 6.1. *Tabla de utilidades*

ACCIONES	UTILIDADES			
a_i	$u(a_i,\theta_1)$	$u(a_i,\theta_2)$	...	$u(a_i,\theta_m)$
a_1	$u(a_1,\theta_1)$	$u(a_1,\theta_2)$	...	$u(a_1,\theta_m)$
a_2	$u(a_2,\theta_1)$	$u(a_2,\theta_2)$	...	$u(a_2,\theta_m)$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$
a_n	$u(a_n,\theta_1)$	$u(a_n,\theta_2)$	...	$u(a_n,\theta_m)$

La formulación de los problemas de decisión se ilustrará mediante los siguientes ejemplos.

Ejemplo 6.1. Prueba estadística de hipótesis. Al realizar una prueba de hipótesis sólo hay dos acciones:

a_1 : aceptar la hipótesis

a_2 : rechazar la hipótesis

El espacio de acciones es, por lo tanto, $A = \{a_1, a_2\}$. Supongamos que los estados de la naturaleza son θ_0 y θ_1, que corresponden, por ejemplo, a los valores posibles de la media o de la varianza de una variable aleatoria, el espacio de estados de la naturaleza es, entonces, $\Omega = \{\theta_0, \theta_1\}$.

...

Se trata de probar la hipótesis H_0: $\theta = \theta_0$, contra la alternativa H_1: $\theta = \theta_1$. La tabla de utilidades correspondiente es

a_i	$u(a_i, \theta_0)$	$u(a_i, \theta_1)$
a_1	U_1	U_2
a_2	U_3	U_4

donde $U_3 = u(a_2, \theta_0)$ y $U_2 = u(a_1, \theta_1)$, son utilidades con valores negativos, ya que corresponden, a una *pérdida* ocasionada por cometer el error tipo I, de rechazar H_0 (tomar a_2) siendo verdadera ($\theta = \theta_0$), y a cometer el error tipo II, de aceptar H_0 (tomar a_1) siendo falsa ($\theta = \theta_1$), respectivamente. (En algunos textos, como la ref 3, se utiliza la *función de pérdida* $L(a_i, \theta_j)$ en vez de la *función de utilidad*, pero ambos criterios son equivalentes puesto que una pérdida equivale a una utilidad negativa, y viceversa). Las utilidades $U_1 = u(a_1, \theta_0)$ y $U_4 = u(a_2, \theta_1)$ son positivas (hay ganancia), puesto que corresponden a tomar las acciones correctas. El árbol de decisión para este problema es el mostrado en la fig 6.3.

ENTRA FIG 6.3

Ejemplo 6.2. Construcción de la carpeta de una carretera.

Un contratista construirá la carpeta de una carretera en tramos de 50 m; el gobierno aceptará o rechazará cada tramo de acuerdo con una prueba de control de calidad. El contratista tiene la opción de pedir el concreto a una de dos plantas premezcladoras; la planta A cobra 140 pesos/m^3 y la B 160 pesos/m^3, pero el control de calidad que se lleva en la planta B es mejor, lo cual hace más probable que un tramo dado pase favorablemente la prueba de aceptación. Tomando en cuenta que en cada tramo se usan 100 m^3 de concreto y que la probabilidad de que el proveniente de la planta A no pase la prueba de control es 0.10, y la de B es 0.05, el constructor deberá decidirse por cuál planta usar. El árbol de decisiones de este problema es el mostrado en la fig 6.4, donde $P(\theta_1)$ y $P(\theta_2)$ son las probabili-

ENTRA FIG 6.4

dades de que ocurran θ_1 y θ_2, respectivamente. La utilidad $U_1 = u(a_1, \theta_1)$ es la que corresponde a utilizar la planta A y que la carpeta pase la prueba de control de calidad; en este caso la utilidad (negativa) es el costo del concreto (\$14,000.00) más la colocación (supongamos \$100,000.00), por lo cual $U_1 =$ -114,000.00. $U_2 = u(a_1, \theta_2)$ es la que corresponde a usar la planta A y que la carpeta no pase la prueba de calidad; en este caso el constructor deberá demoler y reconstruir el tramo con los siguientes costos:

Carpeta demolida	Pérdida de prestigio	\$ 5,000.00
	Mano de obra de demolición	15,000.00
	Concreto	14,000.00
	Mano de obra de colocación	100,000.00
Reconstrucción	Mano de obra	100,000.00
	Concreto	14,000.00
	T O T A L	\$ 248,000.00

...

De manera similar se obtienen U_3 y U_4, cuyos valores resultan ser U_3 = - \$116,000.00 y U_4 = - \$252,000.00.

Si la decisión se tomara sin considerar las probabilidades de aceptar la carpeta, el constructor se decidiría por la planta A, ya que la pérdida (utilidad negativa) sería menor. Si se toma en cuenta este hecho, y adoptamos como *criterio de decisión* el escoger la planta que conduzca a una *esperanza de pérdida* menor se tendrá (recuerde que la esperanza de la variable aleatoria X, $E[X]$, es $E[X] = \sum_{i=1}^{m} P[X_i]X_i$, donde las X_i son los valores que puede asumir X, y $P[X_i]$ son las probabilidades correspondientes):

Para la planta A:

$$E[U] = 0.90 \times (-114,000) + 0.10 \times (-248,000) = -\$127,400.$$

Para la planta B:

$$E[U] = 0.95 \times (-116,000) + 0.05 \times (-252,000) = -\$122,800.$$

Comparando ambas cifras se concluye que la decisión de comprar el concreto de la planta B conduce a una pérdida esperada menor que la de la planta A, es decir, se escoge la planta B aunque el precio unitario del concreto sea mayor.

Ejemplo 6.3. Un ingeniero industrial debe decidir sobre la conveniencia de instalar un sistema de aire acondicionado en una planta industrial. Con base en los reportes diarios del servicio meteorológico se ha estimado que los estados de la naturaleza posibles se pueden reducir a:

θ_1: el 80% de los días son muy calientes y el 20% son calientes

θ_2: el 50% de los días son muy calientes, el 30% son calientes y el 20% son tibios

θ_3: el 20% de los días son muy calientes, el 30% son calientes y el 50% son tibios

Las acciones posibles son:

a_1: instalar el aire acondicionado

a_2: no instalar el aire acondicionado

Se estima que las utilidades que corresponden a cada combinación acción-estado de la naturaleza son:

$$u(a_1, \theta_1) = -\$30,000 \ , \ u(a_2, \theta_1) = -\$100,000$$
$$u(a_1, \theta_2) = -\$30,000 \ , \ u(a_2, \theta_2) = -\$\ 40,000$$
$$u(a_1, \theta_3) = -\$30,000 \ , \ u(a_2, \theta_3) = -\$\ 10,000$$

Observe que los \$30,000 que se pierden al instalar el aire acondicionado (acción a_1) corresponden al costo inicial y de mantenimiento del sistema; los \$100,000 que se pierden con la acción a_2 cuando la mayoría de los días son muy calientes corresponden al costo de los daños que las altas temperaturas ambientales puedan ocasionar al producto elaborado o a la materia prima; un comentario semejante a este último se aplica a las cifras de -\$40,000 y -\$10,000.00.

La tabla de utilidades que corresponde a este problema es:

a_i	$u(a_i, \theta_1)$	$u(a_i, \theta_2)$	$u(a_i, \theta_3)$
a_1	-30,000	-30,000	-30,000
a_2	-100,000	-40,000	-10,000

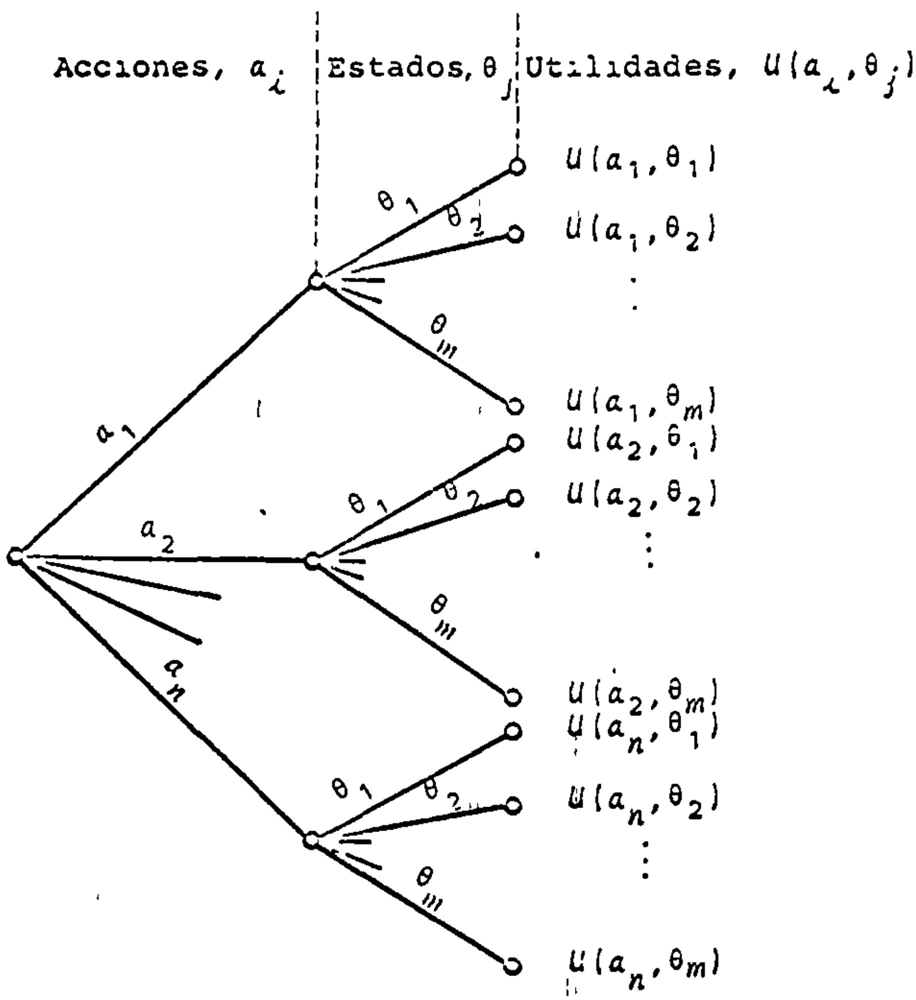


Fig 6.1 Arbol de decisiones

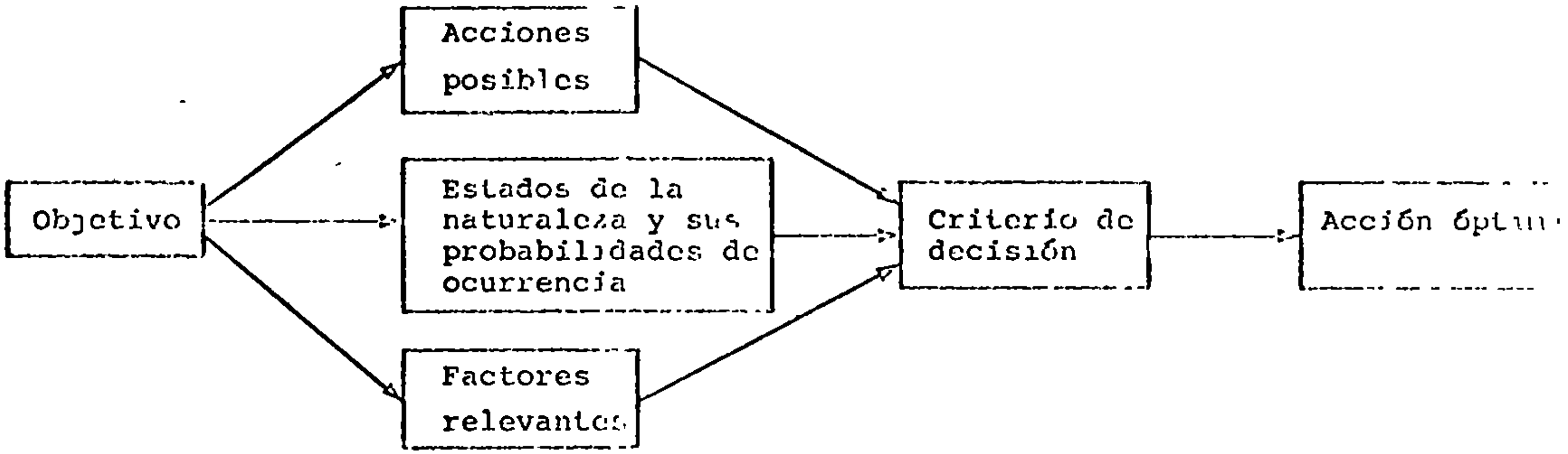


FIG 6.2 DIAGRAMA DEL PROCESO DE DECISION

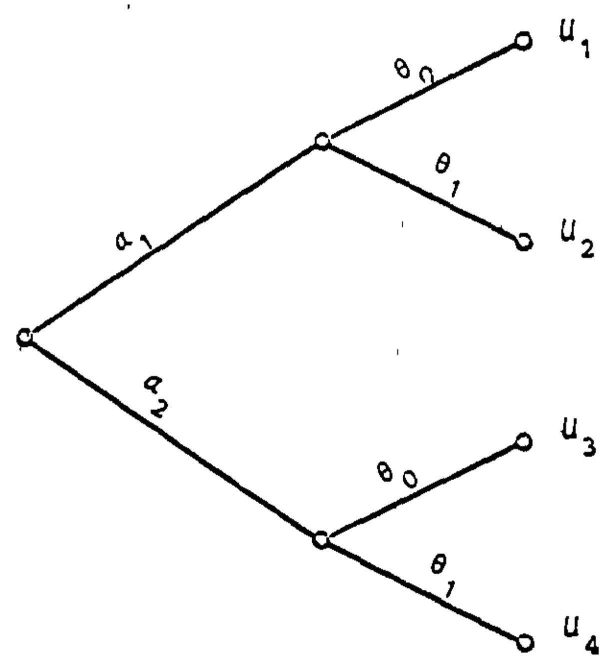


Fig 6.3 Arbol de decisiones del ejemplo 6.1

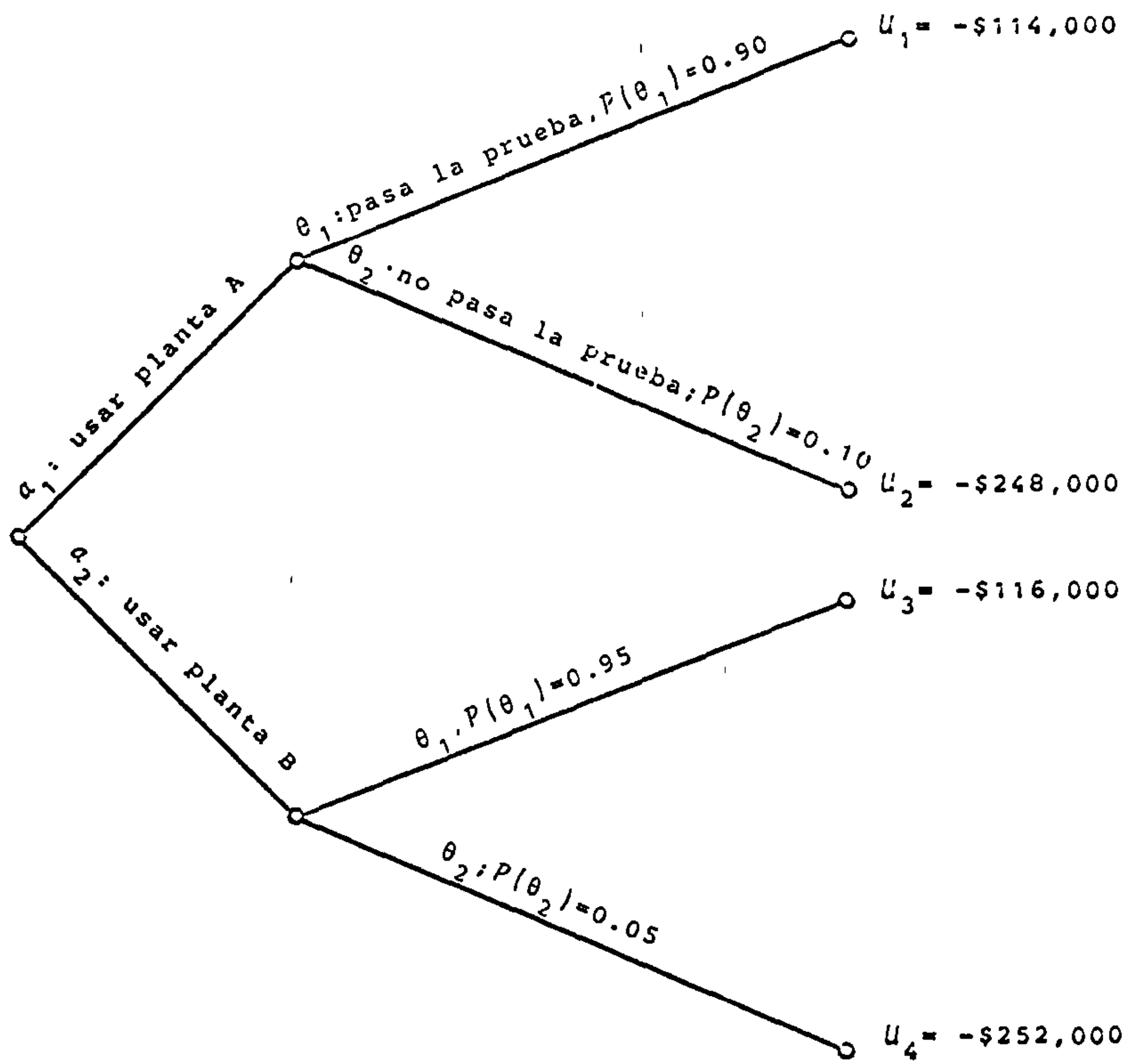


Fig 6.4 Arbol de decisiones del ejemplo 6.2

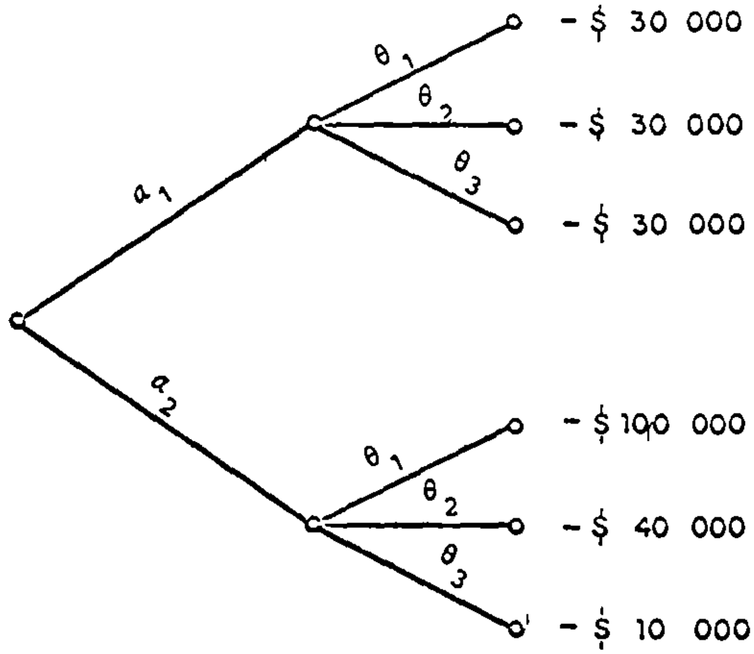


Fig 6.5 Arbol de decisiones del ejemplo 6.3

Prof. M. Tejada

1. Una urna contiene m bolas blancas y n negras. Se sacan k bolas de la urna y se dejan a un lado, sin observar su color. Se toma otra bola de la urna, ¿cuál es la probabilidad de que esta última sea blanca?

2. Calcular la probabilidad de obtener un total de 4 puntos al tirar tres dados.

3. Resolver el problema 2.9 del Papoulis.

4. Si se tiran 6 bolas en 3 canastas, de manera que cada bola tenga la misma probabilidad de caer en cualquier canasta, ¿cuál es la probabilidad de que todas las canastas estén ocupadas?

5. Una urna, A, contiene dos bolas blancas y dos negras. Otra urna, B, tres bolas blancas y dos negras. Pasamos una bola de la urna A a la B. Se saca una bola de la urna B y resulta que es blanca. ¿cuál es la probabilidad de que la bola que pasamos haya sido blanca?

6. Un juguete se arma con tres partes. La probabilidad de que cada parte sea defectuosa es 0.1; calcular la probabilidad de que el juguete esté defectuoso.

7. Tres jugadores (a, b y c) toman sus turnos en un juego de acuerdo con las siguientes reglas:

   Al inicio, a y b juegan, mientras c espera. El perdedor es remplazado por c. El segundo juego es entre c y el ganador, mientras el perdedor del primer juego espera. El juego continúa así hasta que uno de los jugadores gane dos veces consecutivas, resultando por ello el ganador del juego.

   a) Describir el espacio de muestras.
   b) En el espacio anterior, asigne a cada punto, o evento elemental, que contenga k letras, la probabilidad $1/2^k$. Por ejemplo: aa y bb tendrán probabilidad 1/4 cada uno. Demuestre que la suma de las probabilidades asignadas es igual a uno.
   c) Demuestre que P(a gane)=P(b gane)=$\frac{5}{14}$ y P(c gane)=$\frac{2}{7}$

8. El espacio S tiene cinco elementos, S={a, b, c, d, e} Encuentre el mínimo campo de Borel generado por (a,b) y (c).

Probabilidad. Solución a la tarea 1.

1) Nuestra cantidad de información acerca del experimento no ha cambiado, ya que no conocemos el espacio reducido resultante, después de sacar las k bolas. Por lo tanto:

$$P\{blanca\} = \frac{m}{m+n}$$

2) Usando la definición clásica, dado que todos los eventos (a_1, a_2, a_3) donde $\sum_{i=1}^{3} a_i = 4$ se les puede considerar igualmente probables, tenemos que los eventos favorables son $(1,1,2), (1,2,1), (2,1,1)$ y el número de eventos posibles es 6^3, entonces

$$P(\text{obtener cuatro puntos}) = \frac{3}{6^3}$$

3)
a) $$P(a/M) = \frac{P(aM)}{P(M)} = \frac{P(a(M_1+M_2+\cdots+M_n))}{P(M)}$$

$$= \frac{P(aM_1)}{P(M)} + \cdots + \frac{P(aM_n)}{P(M)} \quad \cdots (1)$$

De $P(a/M_i) = \dfrac{P(aM_i)}{P(M_i)}$

$$P(a/M) = \frac{P(a/M_1)\,P(M_1)}{P(M)} + \cdots + \frac{P(a/M_n)\,P(M_n)}{P(M)}$$

b) Falso, ya que la ec. 1 solo es válida si los eventos M_i son mutuamente exclusivos.

4) De la definición clásica

Eventos favorables:

una canasta está ocupada por una bola, otra por dos y la restante por tres

$$a_1 = \binom{6}{1\ 2\ 3} 3!$$

Una canasta está ocupada por cuatro bolas y las otras dos por una bola cada una

$$\partial_2 = \binom{6}{1,1,4} \cdot 3$$

Cada una de las canastas contiene dos bolas

$$\partial_3 = \binom{6}{2,2,2}$$

Eventos posibles : 3^6

$$P\{\text{todas las canastas ocupadas}\} = \frac{\partial_1 + \partial_2 + \partial_3}{3^6} = \frac{20}{29}$$

5) Sean

B = {la bola que pasamos fue blanca}

N = { " " " " " negra }

S_b = { " " " sacamos fue blanca}

Entonces

$$P(B/S_b) = \frac{P(S_b/B)P(B)}{P(S_b/B)P(B) + P(S_b/N)P(N)} = \frac{\frac{4}{6} \times \frac{1}{2}}{\frac{4}{6} \times \frac{1}{2} + \frac{3}{6} \times \frac{1}{2}} = \frac{4}{7}$$

6) Los eventos a, b y c son independientes, por lo que en

$$P(a+b+c) = P(a) + P(b) + P(c) - P(ab) - P(bc) - P(ac) + P(abc)$$

se tiene

$$P(ab) = P(a)P(b)$$

$$P(bc) = P(b)P(c)$$

$$P(ac) = P(a)P(c)$$

$$P(abc) = P(a)P(b)P(c)$$

Entonces

$$P(a+b+c) = 3(0.1) - 3(0.1)(0.1) + (0.1)^3 = 0.271$$

7) a) aa acc acbb acbaa acbacc acbacbb ... I-1-4 (3)
bb bcc bcaa bcabb bcabcc bcabcaa ...

b) De acuerdo con a)

$\frac{1}{4}$ $\frac{1}{8}$ $\frac{1}{16}$ $\frac{1}{32}$ ...

$\frac{1}{4}$ $\frac{1}{8}$ $\frac{1}{16}$ $\frac{1}{32}$ ...

De ahí que

$$\sum P_i = 2\left(\frac{1}{4}+\frac{1}{8}+\frac{1}{16}+\cdots\right) = 2\left(\frac{1}{2^2}+\frac{1}{2^3}+\frac{1}{2^4}+\cdots\right)$$

$$= \frac{2}{2^2}\left(1+\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\cdots\right) = \frac{2}{2^2}\left(\frac{1}{1-\frac{1}{2}}\right) = 1$$

c) $$P(a \text{ gane}) = \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^5} + \frac{1}{2^7} + \frac{1}{2^8} + \frac{1}{2^{10}} + \frac{1}{2^{11}} + \cdots$$

$$+\frac{1}{2^3} \qquad +\frac{1}{2^6} \qquad +\frac{1}{2^9} \qquad +\frac{1}{2^{12}}+\cdots$$

$$-\frac{1}{2^3} \qquad -\frac{1}{2^6} \qquad -\frac{1}{2^9} \qquad -\frac{1}{2^{12}}-\cdots$$

$$= \left(\frac{1}{2^2}+\frac{1}{2^3}+\frac{1}{2^4}+\cdots\right) - \left(\frac{1}{2^3}+\frac{1}{2^6}+\frac{1}{2^9}+\cdots\right)$$

$$= \frac{1}{2^2}\left(\frac{1}{1-\frac{1}{2}}\right) - \frac{1}{2^3}\left(\frac{1}{1-\frac{1}{2^3}}\right) = \frac{5}{14}$$

Del espacio de muestras se ve que $P(a \text{ gane}) = P(b \text{ gane})$

$$P(c \text{ gane}) = 2\left(\frac{1}{2^3}+\frac{1}{2^6}+\frac{1}{2^9}+\cdots\right) = \frac{2}{2^3}\left(\frac{1}{1-\frac{1}{2^3}}\right) = \frac{2}{7} = \frac{2}{7}$$

8) $\{ (a,b), (c), 0, S, (c,d,e), (a,b,d,e), (a,b,c), (d,e) \}$

1. ¿Cuántas diagonales hay en un polígono de doce lados?

2. Hay cinco pistolas que, cuando se apuntan y disparan bien, tienen las siguientes probabilidades de dar en el blanco: 0.5, 0.6, 0.7, 0.8 y 0.9. Se escoge al azar una pistola, se apunta y dispara. Calcular la probabilidad de dar en el blanco.

3. Resolver el problema 3.3 del Papoulis.

4. En un cuarto están reunidas 730 personas. Para $i = 1, 2, \ldots, 730$, sea A_i el evento: la i-ésima persona nació el primero de enero. Considere que los eventos $A_1, \ldots, A_{730}$ son independientes e igualmente probables, con probabilidad $\frac{1}{365}$. Encuentre, en forma aproximada, la probabilidad de que

   i) 2

   ii) 2 ó más personas hayan nacido el primero de enero.

5. Un individuo trata de distraerse mientras espera a su novia y decide caminar de acuerdo con las siguientes reglas. Tira una moneda; si cae águila camina 10 m. hacia el norte; si cae sol camina los 10 m. hacia el sur. En teoría de probabilidad esto se conoce como camino aleatorio. Calcular la probabilidad de que después de haber caminado 100 m.

   a) Esté donde empezó

   b) A 10 m. de donde empezó

   c) A 20 m.

6. Dos bulbos están conectados en serie en un circuito. Encuentre la probabilidad de que un incremento en el voltaje, superior a su valor especificado, interrumpirá el funcionamiento del circuito si, en las condiciones expuestas, la probabilidad de que un bulbo se queme es 0.4, para cada uno de los dos bulbos. I-2-2

7. Un juego consiste en tirar aros a los muñecos de barro, como en las ferias. Un jugador compra seis aros y los tira consecutivamente hasta que ensarta un muñeco y gana. Calcular la probabilidad de que se quede al menos con un aro, si la probabilidad de un tiro afortunado es 0.1.

Probabilidad. Solución a la tarea 2. I-2-3 1/2

1. # diagonales = 54

2. Sean los eventos B = {dar en el blanco}

A_i = {seleccionar la i-ésima pistola}

Del teo. de prob total

$$P(B) = 0.5(0.2) + 0.6(0.2) + 0.7(0.2) + 0.8(0.2) + 0.9(0.2) = 0.7$$

3. $$P(\#\,par) = \sum_{k=0}^{[\frac{n}{2}]} \binom{n}{2k} p^{2k} q^{n-2k} = 1 - \sum_{k=1}^{[\frac{n+1}{2}]} \binom{n}{2k-1} p^{2k-1} q^{n+1-2k}$$

Sumando a ambos miembros de la última igualdad el primero de ellos

$$2\sum_{k=0}^{[\frac{n}{2}]} \binom{n}{2k} p^{2k} q^{n-2k} = 1 + \sum_{k=0}^{[\frac{n}{2}]} \binom{n}{2k} p^{2k} q^{n-2k} - \sum_{k=1}^{[\frac{n+1}{2}]} \binom{n}{2k-1} p^{2k-1} q^{n+1-2k}$$

$$\therefore \sum_{k=0}^{[\frac{n}{2}]} \binom{n}{2k} p^{2k} q^{n-2k} = \frac{1}{2}\left[1 + (q-p)^n\right]$$

4. a) $$P_{730}(2) = \binom{730}{2}\left(\frac{1}{365}\right)^2\left(\frac{364}{365}\right)^{728}$$

Usando el teo. de DeMoivre-Laplace

$$\simeq \frac{1}{\sqrt{730\left(\frac{1}{365}\right)\left(\frac{364}{365}\right)}}\, \phi\left(\frac{2-2}{\sqrt{npq}}\right) = \frac{\sqrt{365}}{\sqrt{2\times 364}}\,\phi(0) = \frac{\sqrt{365}}{\sqrt{2\times 364}}\left(\frac{1}{\sqrt{2\pi}}\right) = 0.27$$

Usando el teo. de Poisson

$$\simeq \frac{e^{-2}(2)^2}{2!} = 2(0.135) = 0.27$$

b) $P\{2 \text{ o más}\} = 1 - P\{0 \text{ o } 1\} = 1 - P(0) - P(1)$

Usando el teo de DeMoivre-Laplace

$$\simeq 1 - \frac{1}{\sqrt{2}}\, \sigma\left(-\frac{2}{\sqrt{2}}\right) - \frac{1}{\sqrt{2}}\, \sigma\left(-\frac{1}{\sqrt{2}}\right) = 1 - 0.7\left[\sigma(-1.4) + \sigma(0.7)\right]$$

$$= 1 - \frac{0.7}{\sqrt{2\pi}}\left(e^{-0.7} + e^{-0.35}\right) = 0.64$$

Usando el teo. de Poisson

$$P\{2 \text{ o más}\} \simeq 1 - \frac{e^{-2}(2)^0}{0!} - \frac{e^{-2}(2)}{1!} = 1 - 0.135 - 0.27 = 0.6$$

5. $n = 10$

a) Para regresar al origen es necesario caminar 5 veces hacia el norte y 5 hacia el sur, o sea

$$P = \binom{10}{5}\left(\frac{1}{2}\right)^5\left(\frac{1}{2}\right)^5 = 0.2461$$

b) No es posible ocupar esa posición en un número par de pruebas, por lo tanto

$$P = 0$$

c) Existen dos posibilidades

Caminar 6 veces hacia el norte y 4 hacia el sur

o caminar 4 hacia el norte y 6 hacia el sur

$$\therefore P = \binom{10}{6}\left(\frac{1}{2}\right)^6\left(\frac{1}{2}\right)^4 + \binom{10}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^6 = 0.41$$

6. De $R_{n,1} = 1 - q^n$

para $n = 2$, $q = 0.6$ $\qquad 1 - (0.6)(0.3)(3)$

$$R_{2,1} = 1 - (0.6)^2 = 0.64$$

7. Tenemos que calcular la probabilidad de ensartar un aro al menos una vez en cinco tiradas

$$R_{5,1} = 1 - (0.9)^5 = 0.41$$

1.- Tiramos un dado hasta que aparezca un dos (2). Sea la v.a. X el número de tiradas necesarias
a).- Encontrar $F_X(x)$
b)- Encontrar $F_X(x)$ si se utilizan dos (2) dados

2.- Calcular el valor de A que hace a $f_X(x)$ una función de densidad

$$f_X(x) = \begin{cases} Ax & 0 \leq x \leq 4 \\ A(8-x) & 4 \leq x \leq 8 \\ 0 & x<0,\ x>8 \end{cases}$$

Determinar y dibujar la función de distribución correspondiente.
Calcular $P\{X \leq 6\}$ y $P\{0 \leq X \leq 6\}$

3.- Definir una función de densidad usando $x(2-x)$, sobre el conjunto $0<x<2$. Encontrar $P(a<X<b)$ si
a- $0<a<b<2$
b.- $a<0,\ 2<b$

4.- Suponiendo que la vida útil, en hrs., de un cierto tipo de bulbos sigue una densidad $f_X(x) = 100/x^2$, si $x>100$ y $f_X(x)=0$ si $x \leq 100$.

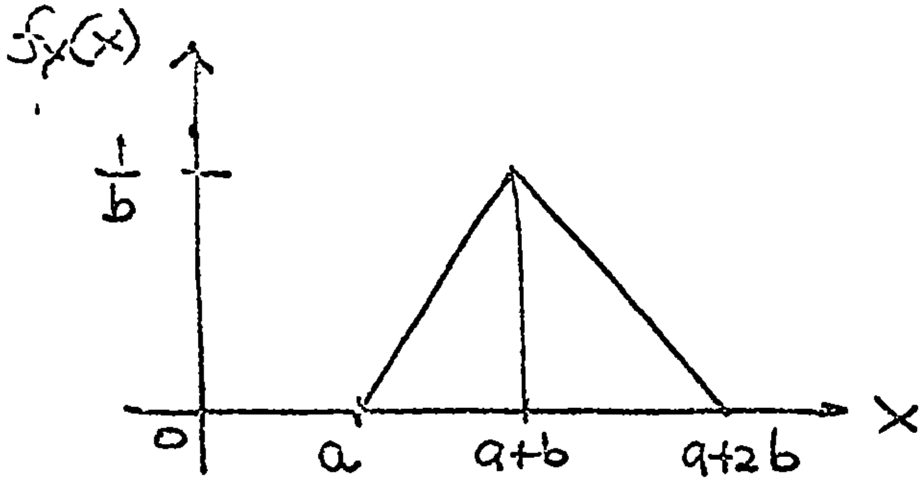
¿Cuál es la probabilidad de que ninguno de los tres bulbos de un radio tenga que ser sustituido durante las primeras 150 horas de operación?

¿Cuál es la probabilidad de que los tres bulbos originales tengan que ser sustituidos durante las primeras 150 horas de operación?

5).- Calcular y dibujar la función de distribución correspondiente a la función de densidad definida como:

$$f_X(x) = \begin{cases} 0 & x<0 \\ 3/4 & x=0 \\ \dfrac{e^{-x}}{4} & x>0 \end{cases}$$

6) Calcular y dibujar la función de distribución correspondiente a la función de densidad:



7) Resolver el problema 4.1 del PAPOULIS.

El espacio del experimento de lanzar un dado es $S=\{f_1, f_2, .. f_6\}$. Consideramos como eventos solo los cuatro conjuntos ϕ, S, {par}, {Non}.

Defina dos funciones $x(f_i)$ y $y(f_i)$ con dominio el conjunto S tal que x sea una v.a. pero y no lo sea.

Probabilidad. Solución a la tarea 3. I-3-3

1.- a)

$\frac{1}{6}$ $\left(\frac{5}{6}\right)\left(\frac{1}{6}\right)$ $\left(\frac{5}{6}\right)^2\left(\frac{1}{6}\right)$

$x=1$ $x=2$ $x=3$

$$F_X(x) = P\{X \le x\} = \sum_{k=1}^{x} \left(\tfrac{5}{6}\right)^{k-1}\left(\tfrac{1}{6}\right)$$

b) Eventos favorables

(2,1), (2,2), (2,3), (2,4), (2,5), (2,6), (1,2) (3,2), (4,2), (5,2), (6,2)

$E.f. = 12$

Eventos posibles $= 3^2$

$P(\text{obtener un } 2) = \frac{1}{3}$

$\left(\frac{1}{3}\right)$ $\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)$ $\left(\frac{2}{3}\right)^2\left(\frac{1}{3}\right)$

$x=1$ $x=2$ $x=3$

$$F_X(x) = \sum_{k=1}^{x} \left(\tfrac{2}{3}\right)^{k-1}\left(\tfrac{1}{3}\right)$$

2.- Valor de A

$$1 = \int_0^4 Ax\,dx + \int_4^8 A(8-x)\,dx = 8A + 32A - 32A + 8A \quad \therefore \quad A = \frac{1}{16}$$

Determinación de $F_X(x)$

$x \ge 8$

$$F_X(x) = P\{X \le x\} = 1$$

$4 \le x \le 8$

$$F_X(x) = \frac{1}{16}\int_0^4 x\,dv + \frac{1}{16}\int_4^x (8-x)\,dx = -\frac{x^2}{32} + \frac{x}{2} - 1$$

$0 \le x \le 4$

$$F_X(x) = \frac{1}{16}\int^x x\,dx = \frac{x^2}{\ldots}$$

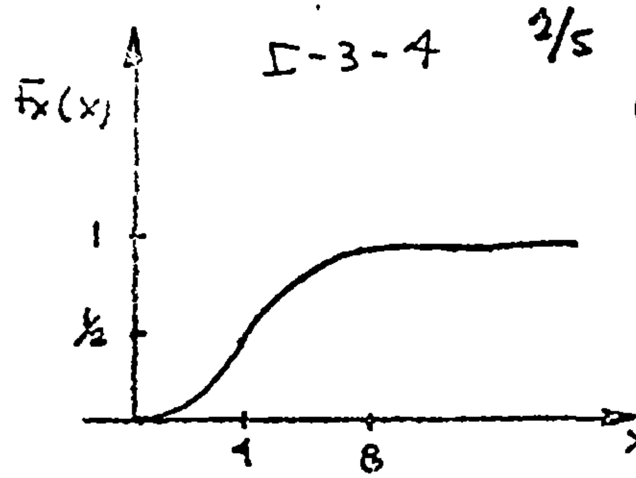
I-3-4 2/5

$X < 0$

$F_X(x) = 0$

o sea

$$F_X(x) \begin{cases} 1 & x = 8 \\ -\frac{x^2}{32} + \frac{x}{2} - 1 & 4 \le x \le 8 \\ \frac{x^2}{32} & 0 \le x \le 4 \\ 0 & x < 0 \end{cases}$$



$$F_X(6) = P\{X \le 6\} = -\frac{6^2}{32} + \frac{6}{2} - 1 = \frac{7}{8}$$

$$P\{0 \le X \le 6\} = F_X(6) - F_X(0) = \frac{7}{8}$$

3.- $A \int_0^2 x(2-x)\,dx = 1 = A\left(4 - \frac{8}{3}\right) \quad \therefore \quad A = 3/4$

$$f_X(x) = \begin{cases} \frac{3}{4} x(2-x) & 0 < x < 2 \\ 0 & x < 0,\ x > 2 \end{cases}$$

a) $P(a < X < b)$

$$F_X(x) = \begin{cases} 3/4\left(x^2 - \frac{x^3}{3}\right) & 0 < x < 2 \\ 0 & x < 0,\ x > 2 \end{cases}$$

Para $0 < a < b < 2$

$$P\{a < x < b\} = F(b) - F(a) = \frac{3}{4}\left(b^2 - \frac{b^3}{3}\right) - \frac{3}{4}\left(a^2 - \frac{a^3}{3}\right) = \frac{b^2(3-b)}{4} - \frac{a^2(3-a)}{4}$$

b) si $a < 0$, $2 < b$

$$P\{a < x < b\} = P\{a < x < 0\} + P\{0 < x < 2\} + P\{2 < x < b\}$$

ya que $P\{a < x < 0\} = P\{2 < x < b\} = 0$

y $P\{0 < X < 2\} = F_X(2) - F_X(0) = 1$

4.- $P\{\text{falla de un bulbo}\} = \int_{100}^{150} \frac{100}{x^2}\,dx = \frac{1}{3}$ I-3-5 3/5

$P\{\text{0 fallas en 3 bulbos}\} = \binom{3}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^3 = \frac{8}{27}$

$P\{\text{3 fallas en 3 bulbos}\} = \binom{3}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^0 = \frac{1}{27}$

Otra forma:

$P\left\{\begin{array}{l}\text{ninguno de los 3 bulbos tenga} \\ \text{que ser sustituido durante} \\ \text{las 1as 150 Hs. de operación}\end{array}\right\} = \left[\int_{150}^{\infty} \frac{100}{x^2}\,dx\right]^3 = \frac{8}{27}$

$P\left\{\begin{array}{l}\text{los 3 bulbos tengan que ser} \\ \text{reemplazados durante las} \\ \text{primeras 150 Hs.}\end{array}\right\} = \left[\int_{100}^{150} \frac{100}{x^2}\,dx\right]^3 = \frac{1}{27}$

5.- $x > 0$

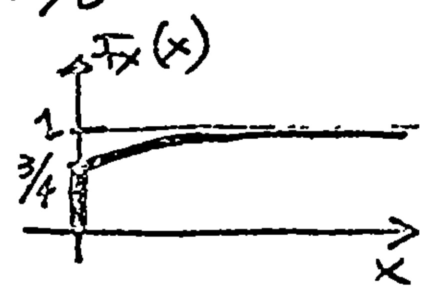
$F_X(x) = \frac{3}{4} + \int_0^x \frac{e^{-t}}{4}\,dt = \frac{3}{4} + \frac{1}{4}\left[-e^{-t}\right]_0^x = 1 - \frac{e^{-x}}{4}$

$x = 0 \quad F_X(x) = \frac{3}{4}$

$x < 0$

$F_X(x) = 0$

$$F_X(x) = \begin{cases} 1 - \dfrac{e^{-x}}{4} & x > 0 \\ \frac{3}{4} & x = 0 \\ 0 & x < 0 \end{cases}$$



6.-

$$f_X(x) = \begin{cases} 0 & x > a+2b \\ \frac{1}{b^2}(a+2b-x) & a+b \le x \le a+2b \\ \frac{1}{b^2}(x-a) & a \le x \le a+b \\ 0 & x < a \end{cases}$$

$x > a+2b$

$F_x(x) = 1$

$a+b \leq x \leq a+2b$

$$F_x(x) = \frac{1}{b^2}\int_a^{a+b}(x-a)\,dx + \frac{1}{b^2}\int_{a+b}^{x}(a+2b-x)\,dx$$

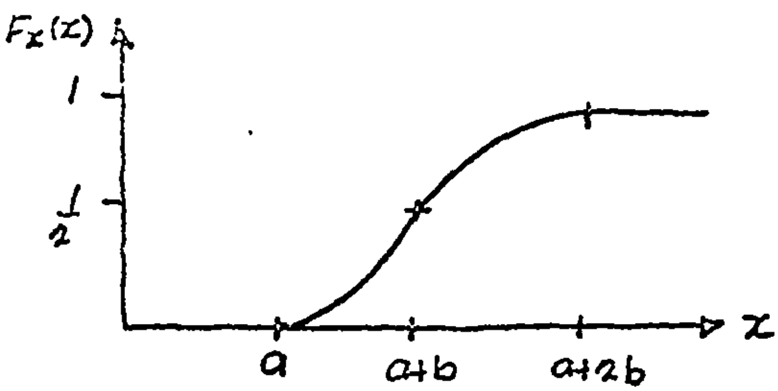
$$= \frac{1}{2} - \frac{a+b}{2b^2}\left[a+3b\right] + \frac{1}{b^2}\left[(a+2b)x - \frac{x^2}{2}\right]$$

$a \leq x \leq a+b$

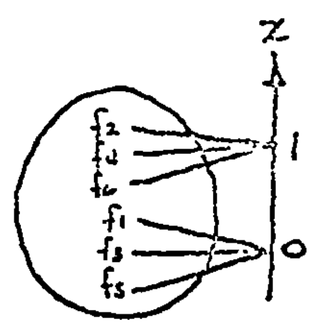
$$F_x(x) = \frac{1}{b^2}\int_a^{x}(x-a)\,dx = \frac{1}{b^2}\left(\frac{x^2}{2} - ax + \frac{a^2}{2}\right)$$

$x < a$

$F_x(x) = 0$



Problema 7.



$F\{\phi, \{pares\}, \{impares\}, S\}$

Sea $x(\xi) = \begin{cases} 0 & \text{si } \xi \text{ impar} \\ 1 & \text{si } \xi \text{ par} \end{cases}$

si $x \geq 1$

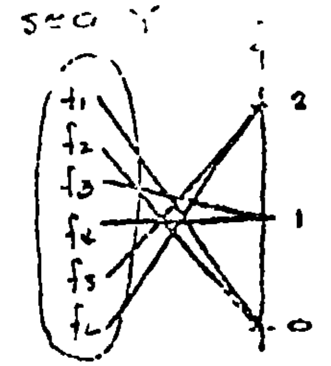
$\{\xi : x(\xi) \leq x\} = S \in F$

si $0 \leq x < 1$

$\{\xi : x(\xi) \leq x\} = \{f_1, f_3, f_5\} \in F$

$0 > x$

$\{\xi : x(\xi) \leq x\} = \phi \in F$ $\qquad$ x es una v.a si $y \geq 2$ $\quad \{\zeta : y(\zeta) \leq y\} = S \in F$ $\quad$ I-3-7

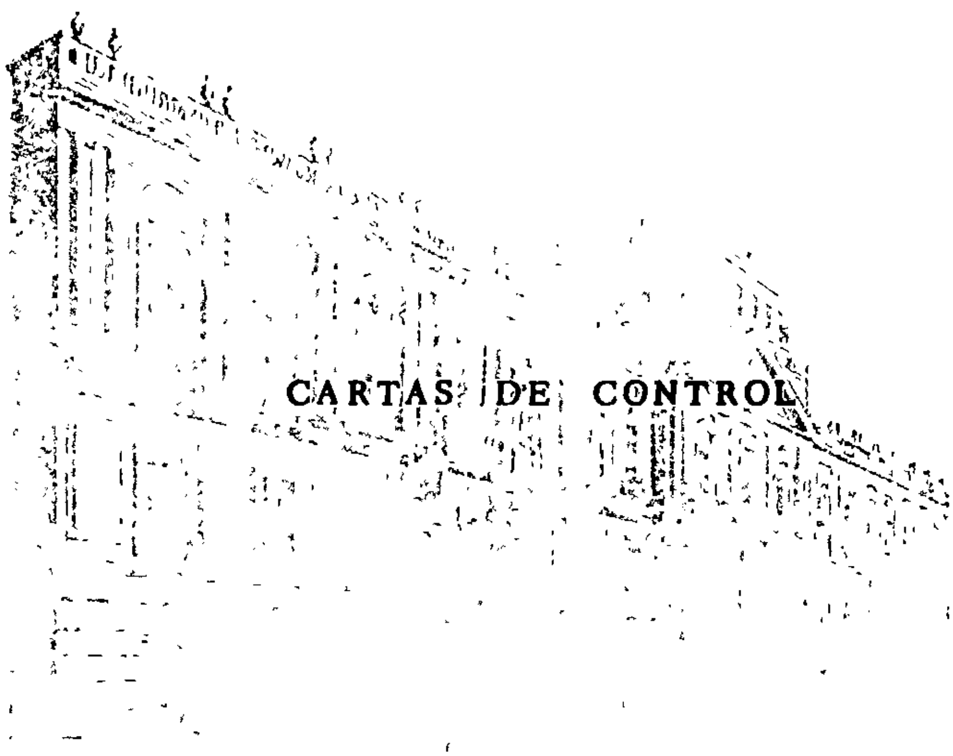


si $1 \leq y < 2$ $\quad \{\zeta : y(\zeta) \leq y\} = \{f_3, f_4\} \notin F$

$\therefore$ y no es una v.a en F.

centro de educación continua
división de estudios superiores
facultad de ingeniería, unam

"PROBABILIDAD Y ESTADISTICA"



CARTAS DE CONTROL

Ing. Augusto Villarreal

Palacio de Minería
Tacuba 5, primer piso México 1, D F
Tels 521-40-23 521-73-35 5123-123

CARTAS DE CONTROL

Por: Augusto Villarreal A*

INTRODUCCION

Aunque existe la tendencia generalizada a pensar que el Control de Calidad es de desarrollo reciente, realmente no existe nada nuevo en la idea básica de elaborar un producto caracterizado por un alto grado de uniformidad.

Durante siglos, hábiles artesanos han procurado elaborar productos que se distingan por su superior calidad, y una vez que han logrado obtener un cierto estándar de calidad óptimo, eliminar dentro de lo posible la variación entre productos que nominalmente deben resultar iguales.

La idea de que la Estadística puede resultar un instrumento muy útil para asegurar un estándar adecuado de calidad para los productos manufacturados se remonta no mas allá del advenimiento de la producción masiva, y el uso extendido de los métodos estadísticos para resolver problemas de control de calidad es aún más reciente.

Muchos problemas que aparecen durante la elaboración de un producto son susceptibles de ser atacados empleando tratamientos estadísticos, por lo que al hablar de Control de Calidad Estadístico, nos estaremos refiriendo esencialmente a las tres técnicas especiales, que se discutirán en el resto del Curso: Uso de las Cartas de Control, muestreo de aceptación y establecimiento de límites de tolerancia.

Conviene mencionar que la palabra calidad, al ser empleada de aquí en adelante, se referirá a alguna propiedad medible o contable de algún producto, tal como el diámetro de un balín de acero, la resistencia de una viga de concreto, el número de defectos en una

* Profesor Investigador, División de Estudios Superiores e Instituto de In-

pieza de tela, la eficiencia de cierta droga, etc.

IDEAS SOBRE CARTAS DE CONTROL

A muchos individuos les puede sorprender el hecho de que dos artículos aparentemente idénticos elaborados bajo condiciones cuidadosamente controladas, de las mismas materias primas, y por una misma máquina con diferencia de pocos segundos, puedan, sin embargo, diferir en muchos aspectos.

En efecto, cualquier proceso de manufactura, aún siendo muy bueno, se encuentra caracterizado por una cierta cantidad de variación que es de naturaleza aleatoria, y que no puede ser eliminada en forma completa.

Cuando la variabilidad presente en un proceso de producción se limita a <u>variación aleatoria</u> se dice que el proceso se encuentra en un estado de <u>control estadístico</u>.

Tal estado se puede alcanzar cuando se eliminan aquellos problemas causados por otro tipo de variación, llamada <u>variación sistemática</u>, que es de naturaleza más bien determinística, y que se puede achacar, por ejemplo, a operadores mal entrenados, materias primas de baja calidad, máquinas en mal estado, etc.

Ya que los procesos de manufactura se encuentran rara vez libres de estos problemas, conviene contar con algún método sistemático para detectar desviaciones serias de un estado de control estadístico cuando ocurren, o inclusive antes de que ocurran tales desviaciones.

Ese método sistemático de detección se puede tener mediante el empleo de las llamadas Cartas de Control.

TIPOS DE CARTAS DE CONTROL

En lo que sigue distinguiremos entre las cartas de control para mediciones o variables ($\bar{X}$, R, σ) y las cartas de control para atributos, (p, c) dependiendo de que las observaciones que estemos analizando sean mediciones o datos contados o calculados, respectivamente.

Un ejemplo del primer caso sería, la longitud de las varillas de acero de una muestra. Como ejemplo del segundo caso tendríamos el número de focos defectuosos en una muestra de tamaño dado.

CONFIGURACION DE LAS CARTAS DE CONTROL

En cualquiera de los casos mencionados, una carta de control consiste de una línea central, correspondiente a la calidad promedio a la que el proceso debe funcionar, y dos líneas que corresponden al Límite Superior de Control (LSC) y al Límite Inferior de Control (LIC), respectivamente.

Estos límites se escogen en forma tal que los valores que se encuentren dentro de ellos se puedan atribuir al azar, en tanto que los valores que caigan fuera de ellos se puedan considerar como indicaciones de falta de control.

No obstante la idea anterior, conviene mencionar que en la figura 1 que se presenta a continuación se pueden considerar otras posibles situaciones de "falta de control" que ameritan investigarse:

1. Cuando dos de tres puntos sucesivos caen en la zona A,

o más allá.

2. Cuando cuatro de cinco puntos sucesivos caen en la zona B o más allá.

3. Cuando ocho puntos sucesivos caen en la zona C o más allá.

Debe hacerse notar que las pruebas arriba mencionadas se aplican a ambas mitades de la carta de control, pero se aplican separadamente para cada mitad, y nunca a las dos mitades en combinación.

EXPLICACION DEL EMPLEO DE LAS CARTAS DE CONTROL

Si se grafican en una carta los resultados obtenidos a partir de muestras tomadas periódicamente a intervalos frecuentes, es posible checar por medio de la carta si el proceso se encuentra bajo control o si se encuentra presente en el proceso la variación sistemática del tipo descrito anteriormente.

Cuando un punto muestra cae fuera de los límites de control se trata de encontrar el problema que causó tal evento dentro del proceso. Pero aún si los puntos caen dentro de los límites mencionados, alguna tendencia, o cierto patrón de los mismos puede indicar que se debe llevar a cabo alguna acción para prevenir y evitar algún problema serio.

La habilidad para "leer" las cartas de control y para determinar a partir de ellas cuál acción correctiva debe llevarse a cabo, se obtiene a partir de la experiencia y del juicio altamente desarrollado.

Un practicante del control estadístico de la calidad debe no tan solo comprender los fundamentos estadísticos de la materia, sino también encontrarse identificado plenamente con los procesos que desea controlar.

CARTAS DE CONTROL PARA MEDICIONES (VARIABLES)

Cuando se requiere establecer control estadístico de la calidad de algún producto en términos de mediciones o variables, es costumbre ejercitar tal control sobre la calidad promedio del proceso, al igual que sobre su variabilidad.

La primera meta se logra al graficar los promedios (medias) de muestras extraídas periódicamente en la llamada carta de control para medias, o simplemente, carta $\bar{x}$. La variabilidad se puede controlar de igual forma si se grafican los rangos o las desviaciones estándar de las muestras en las llamadas Cartas R o cartas σ, respectivamente, dependiendo de cuál estadística se emplee para estimar la desviación estándar de la población.

Si se conocen μ y σ de la población, y es razonable suponer las mediciones obtenidas como muestras extraídas de una población normal, podremos asegurar, con $P = 1 - \alpha$ que la media de una muestra aleatoria de tamaño n se encontrará entre

$$\mu - z_c \frac{\sigma}{\sqrt{n}} \quad y \quad \mu + z_c \frac{\sigma}{\sqrt{n}}$$

o

$$\mu - z_c \, \sigma_{\bar{x}} \quad y \quad \mu + z_c \, \sigma_{\bar{x}}$$

Los dos límites anteriores ($\pm z_c \, \sigma_{\bar{x}}$) proporcionan entonces límites inferiores y superiores de control y, bajo las suposiciones anteriores, permiten al practicante del control de calidad determinar si se debe o no llevar a cabo algún ajuste en el proceso.

Conviene establecer en este momento que realmente al emplear una carta de control lo que estamos haciendo es probar las hipótesis nulas de que a un cierto nivel de confianza $1 - \alpha$ los valores de las estadísticas obtenidas de las muestras sean iguales al parámetro correspondiente de la población (calidad promedio), es decir, estamos realizando prueba de hipótesis en la cual:

$$H_0 : \theta = S$$

$$H_1 : \theta \neq S \qquad \text{(Dos colas)}$$

Abundando más en el asunto, estaremos probando las hipótesis secuenciales de que con cierta probabilidad $1 - \alpha$ la calidad de las muestras sea igual a la calidad promedio del proceso, tal como se puede ver en la figura (4).

Si se consideran problemas prácticos, los valores de μ y σ del proceso se desconocen, y es entonces conveniente estimar sus valores a partir de muestras tomadas mientras el proceso se encuentre "bajo control". En la práctica pues, es difícil llegar a establecer límites de control del tipo $\pm z_c \frac{\sigma}{\sqrt{n}}$ al desconocerse μ y σ, independientemente de que en muchos casos es demasiado arriesgado considerar a las mediciones como muestras aleatorias extraídas de población normal.

En lugar de lo anterior, en el control de calidad industrial se emplean comúnmente los límites de control de "tres desviaciones estándar" ó de "tres sigmas", que se obtienen al sustituir a z_c

por un 3.

Así pues, con los límites de control

$$S \pm 3\frac{\sigma}{\sqrt{n}} \quad ó \quad S \pm 3\sigma_S$$

se puede estar 99.63% seguro de que el proceso no será declarado "fuera de control" cuando de hecho se encuentra "bajo control".

En otras palabras, estos límites de control permiten considerar que la probabilidad máxima de rechazar la hipótesis

$$H_0 : \theta = S$$

cuando debería de ser aceptada (probabilidad de cometer un error de tipo I) es de o.13%.

Carta $\bar{x}$.

a) Caso en que se conocen μ y σ

Línea central ——— μ

Límites de control ——— $\mu \pm 3\frac{\sigma}{\sqrt{n}}$

o

$$\mu \pm A\sigma \text{ , siendo } A = \frac{3}{\sqrt{n}}$$

Notación:

Media de la población = μ = $\bar{x}'$

Desviación estándar = σ = σ'

donde A viene dado en la tabla de la figura II.

Ejemplo: Sea el proceso de elaboración de varillas de acero para las cuales se sabe que el diámetro medio de las mismas es de 2.5 cm, con una desviación estándar de 0.01 cm. Se desea efectuar control del diámetro de las mismas, para lo cual se extraen periódicamente muestras de cinco varillas. Establecer la línea central y los límites de control para una carta $\bar{x}$.

Solución. Siendo $\mu = 2.5$ cm, $\sigma = 0.01$ y $n = 5$ se tiene que

línea central = $\mu = 2.5$

Límites de control:

$$2.5 \pm 3 \frac{\sigma}{\sqrt{n}} = 2.5 \pm \frac{3(0.01)}{\sqrt{5}} = 2.5 \pm 0.0134 \Rightarrow 2.5134,\ 2.4866$$

o, de figura II

$$2.5 \pm A\sigma = 2.5 \pm 1.342(0.01) = 2.5 \pm 0.01342 \Rightarrow 2.51342,\ 2.48658$$

b) Caso en que se desconocen μ y σ

Para este caso, que es el más común, es necesario estimar, como se dijo anteriormente, tales parámetros con base en muestras preliminares. Para el caso, normalmente se acostumbra emplear un mínimo de 20 o 25 muestras de 4 o 5 elementos obtenidos consecutivamente cuando el proceso está "bajo control".

Sin embargo, como veremos más adelante, se pueden emplear procedimientos estadísticos más formales para determinar el número de muestras y de elementos en las mismas más adecuado para las cartas $\bar{x}$.

Así pues, si se utilizan k muestras preliminares, cada una de tamaño n, podemos estimar con adecuada precisión el valor de μ mediante

$$\bar{\bar{X}} = \frac{1}{k} \sum_{i=1}^{k} \bar{x}_i$$

siendo $\bar{\bar{X}}$ un estimador insesgado y consistente de μ, donde $\bar{x}_i$ denota el promedio aritmético de la iésima muestra y $\bar{\bar{X}}$ es la media de las medias de las muestras.

El valor de σ de la población puede ser estimado a partir de las desviaciones estándar o de los rangos de las muestras. Si el tamaño de las mismas es pequeño, usualmente el rango proporciona un estimador eficiente de σ, además de que el proceso de cálculo del mismo es bastante más simple que el de la desviación estándar para las muestras.

Sin embargo, conviene, cuando se requiere bastante precisión en el cálculo de los límites de control, estimar a σ mediante las desviaciones estándar de las muestras.

Tal es el caso por ejemplo de muestras de productos que son caros y que deben destruirse al momento de tomar las mediciones.

b.1) Estimando a σ mediante los rangos de las muestras

Hay que obtener primero el valor $\bar{R}$ que es el rango promedio de los rangos de las k muestras, es decir,

$$\bar{R} = \frac{1}{k} \sum_{i=1}^{k} R_i$$

Puesto que la estadística $\bar{R}$ es un estimador sesgado de σ, es indispensable afectar el valor obtenido en forma tal de convertirlo en un estimador insesgado de σ, para lo cual se hace

$$\text{Estimador insesgado de } \sigma = \frac{\bar{R}}{d_2}$$

DISTRIBUCION EXPONENCIAL

Se dice que una variable aleatoria X tiene distribución exponencial con parámetro λ ($\lambda>0$) si su función de densidad de probabilidades se encuentra dada por

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & ; \quad x \geq 0 \\ 0 & ; \quad x < 0 \end{cases}$$

Se puede demostrar que la esperanza y la varianza de una variable de este tipo están dadas por

$$E[X] = 1/\lambda \quad ; \quad Var(X) = 1/\lambda^2$$

Una propiedad importante de la variable aleatoria distribuida exponencialmente es que no tiene <u>memoria</u>. Es decir, si X es la variable aleatoria "Tiempo de duración de una guitarra", dicha propiedad indica que la probabilidad de que sirva una guitarra por lo menos $t+s$ horas sabiendo que ya ha dado servicio durante t horas es igual a la probabilidad de que dure al menos s horas. Matemáticamente esto se escribe

$$P\{X > s+t \mid X > t\} = P\{X > s\} \quad \forall\, s,t \geq 0$$

En otras palabras, lo que quiere decir lo anterior es que si el instrumento se encuentra en servicio en el tiempo t, entonces la distribución del tiempo de servicio restante es igual a la distribución del tiempo que lleva la guitarra en uso, es decir, el instrumento no recuerda que ya ha "tocado" durante un tiempo t!

PROCESOS DE MARKOV DE PARAMETRO CONTINUO

1º Sistema de cola (línea de espera) con una estación de servicio

1. La gente arriban a una sola estación de acuerdo con un proceso de Poisson con parámetro λ. Es decir

$$E[\text{t de arribo}] = 1/\lambda \quad (\text{variables aleatorias indep de dist expon}$$

2. Los tiempos de servicio sucesivos se suponen variables aleatorias exponenciales con media

$$E[\text{t. de servicio}] = \frac{1}{\mu}$$

3. Nos interesa conocer lo siguiente.

L = número medio de personas en el sistema
L_Q = 〃 〃 〃 〃 esperando en la cola
W = tiempo promedio que la persona está en el sistema
W_Q = 〃 〃 〃 〃 〃 espera en la cola

4. Para calcular lo anterior, definimos

$X(t)$ = número de personas en el sistema en el tiempo

y $$P_n(t) = P\{X(t) = n\}$$

5. $P_n(t)$ se podría obtener mediante ecuaciones diferenciales de difícil solución; por tanto, nos empeñamos en encontrar

$$P_n = \lim_{t \to \infty} P_n(t) = \lim_{t \to \infty} P\{X(t) = n\}$$

(Probabilidad de que a la larga haya en el sist. exactamente $\underline{n}$ p

6. El número medio de personas en el sistema es, de acuerdo a la definición de P_n,

$$L = \sum^{\infty} n P_n$$

7. Siempre que hay n personas en el sistema $(n \geq 1)$, a una de ellas se le da servicio y $n-1$ esperan en la cola; por ello,

$$L_Q = \sum_{n=1}^{\infty} (n-1) P_n$$

8. Para obtener el tiempo medio que está una persona en el sistema condicionemos al número de personas que ya están cuando esa persona llega; es decir

$$W = \sum_{n=0}^{\infty} E\left[\begin{array}{c|c} \text{tiempo que está la} & \text{había } n \text{ personas} \\ \text{persona en el sist} & \text{en el sist. cuando} \\ & \text{arribó} \end{array}\right] \times P\{n \text{ personas en el sist. cuando arriba}\}$$

8.1 Si hay $n = 0$ personas en el sistema cuando arriba un individuo, el tiempo que pasa en el mismo es el tiempo medio en el que le darán servicio, es decir, $1/\mu$.

8.2 Si hay $n \geq 1$ personas en el sistema al arribar el individuo, una de ellas está recibiendo servicio y otras $n-1$ están en la cola. Puesto que el tiempo de servicio es exponencial sin memoria, el tiempo medio que estará en el sistema el individuo es

$$W = \sum_{n=0}^{\infty} \frac{1}{\mu} \Big(\underbrace{n-1}_{\text{en la cola}} + \underbrace{1}_{\text{recibiendo servicio}} + \underbrace{1}_{\text{nuestro individuo}}\Big) P_n$$

$$= \sum_{n=0}^{\infty} \frac{n+1}{\mu} P_n$$

9. El tiempo medio que un individuo está en la cola es igual al tiempo medio que pasa en el sistema menos el tiempo medio en el que recibe servicio es decir,

$$W_Q = \sum_{n=0}^{\infty} \frac{n+1}{\mu} P_n - \frac{1}{\mu}$$

$$= \sum_{n=0}^{\infty} \frac{n}{\mu} P_n$$

10. Para determinar las probabilidades límite P_n, para $n=0,1,2,\ldots$ supongamos que tenemos un número infinito de cuartos numerados $0,1,2,\ldots$, y que pedimos a un individuo que entre al cuarto n cuando estén n personas en el sistema. Por ejemplo.

10.1 El individuo se meterá al cuarto nº 2 siempre que en el sistema estén dos personas.

10.2 Suponiendo que esté en el cuarto 2 y otra persona llegue al sistema, dejará ese cuarto y se irá al 3.

10.3 Si está en el cuarto 2 y se termina de proporcionar un servicio, dejará este cuarto y se irá al 1.

11. Si nuestro individuo a la larga entra al cuarto 1 a razón de 10 veces por hora, también deberá salir de ese cuarto el mismo número de veces por hora (o tal vez una ocasión menos). De lo anterior se desprende que para cada $n \geq 0$, la proporción de veces en que el sistema entra a un estado n es igual a la proporción de veces en que sale de él.

12. Consideremos el estado 0 (ninguna persona en el sistema). El proceso puede dejar ese estado únicamente si arriba una persona al sistema. Siendo P_0 la proporción de tiempo en que el proceso se encuentra en el estado 0 y λ la tasa de arribos, nuestro proceso abandona ese estado con una proporción λP_0.

13. Al estado 0 se puede llegar del estado 1 (una persona en el sistema) si deja el sistema una persona al habérsele proporcionado servicio. Siendo μ la tasa de servicio y P_1 la proporción de tiempo que en el sistema se encuentra una persona, el proceso llega al estado 0 con una proporción μP_1. De acuerdo con el punto 11, se tiene que

$$\lambda P_0 = \mu P_1$$

14. Consideremos ahora el estado 1 (una persona en el sistema). Puesto que este estado puede ser dejado por el arribo de una persona al sistema o porque a otra se le haya dado servicio, el proceso dejará el estado 1 con una tasa $\lambda + \mu$. Por tanto, el proceso abandona dicho estado con una proporción $(\lambda + \mu) P_1$.

15. El proceso puede llegar al estado 1 desde el estado 0 (por el arribo de una persona) o desde el estado 2 (por haberle sido dado el servicio a otra persona). Por tanto, el proceso llega a dicho estado con una proporción $\lambda P_0 + \mu P_2$. De acuerdo con el punto 11, se tiene que

$$(\lambda + \mu) P_1 = \lambda P_0 + \mu P_2 .$$

16. Generalizando, se tiene que

Estado	Proporción en que el proceso llega		Proporción en que el proceso abandona
0	λP_0	$=$	μP_1
1	$(\lambda+\mu) P_1$	$=$	$\lambda P_0 + \mu P_2$
$\vdots$			
n	$(\lambda+\mu) P_n$	$=$	$\lambda P_{n-1} + \mu P_{n+1}$

Ecuaciones de Balanceo

17. Despejando de las ecuaciones anteriores los valores P_n ($n \geq 1$) se tiene

$$P_1 = \frac{\lambda}{\mu} P_0$$

$$\vdots$$

$$P_{n+1} = \frac{\lambda}{\mu} P_n + \left(P_n - \frac{\lambda}{\mu} P_{n-1}\right)$$

18. Poniendo las ecuaciones anteriores en términos de P_0 se llega a

$$P_0 = P_0$$

$$P_1 = \frac{\lambda}{\mu} P_0$$

$$P_2 = \frac{\lambda}{\mu} P_1 + \left(P_1 - \frac{\lambda}{\mu} P_0\right) = \frac{\lambda}{\mu} P_1 = \left(\frac{\lambda}{\mu}\right)^2 P_0$$

$$\vdots$$

$$P_{n+1} = \frac{\lambda}{\mu} P_n + \left(P_n - \frac{\lambda}{\mu} P_{n-1}\right) \equiv \frac{\lambda}{\mu} P_n = \left(\frac{\lambda}{\mu}\right)^{n+1} P_0$$

19. Para determinar P_0 consideramos que $\sum_{n=0}^{\infty} P_n = 1$, es decir,

$$\sum_{n=0}^{\infty} P_n = \sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^n P_0 = \frac{P_0}{1 - \frac{\lambda}{\mu}} = 1$$

por lo que

$$P_0 = 1 - \frac{\lambda}{\mu}$$

$$P_1 = \left(\frac{\lambda}{\mu}\right)\left(1 - \frac{\lambda}{\mu}\right)$$

$$P_2 = \left(\frac{\lambda}{\mu}\right)^2 \left(1 - \frac{\lambda}{\mu}\right)$$

$$\vdots$$

$$P_n = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right)$$

20. Para las ecuaciones anteriores, el cociente λ/μ debe ser tal que

$$\frac{\lambda}{\mu} < 1$$

para que se satisfaga $0 \leq P_n \leq 1$. Ello implica también que el tiempo medio de servicio $\left(\frac{1}{\mu}\right)$ sea menor que el tiempo medio de arribos $\left(\frac{1}{\lambda}\right)$, siendo esta condición necesaria para que el sistema sea óptimo

21. Sustituyendo los valores P_n obtenidos anteriormente en las expresiones para L, L_Q, W y W_Q ya calculados, y empleando la fórmula $\sum_{n=0}^{\infty} n x^n = x/(1-x)^2$, se llega a

$$L = \frac{\lambda}{\mu - \lambda} \qquad W = \frac{1}{\mu - \lambda}$$

$$L_Q = \frac{\lambda^2}{\mu(\mu - \lambda)} \qquad W_Q = \frac{\lambda}{\mu(\mu - \lambda)}$$

Ejemplo:

Supóngase que se tiene una población finita formada por los datos 1, 2, 3, 4, 5. Se desea conocer la media y la desviación estándar de la distribución muestral del promedio aritmético, considerando las muestras de tamaño 3 obtenidas sin remplazo.

Primer procedimiento.-

Siendo la población finita y el muestreo sin remplazo, es posible obtener la distribución muestral correspondiente para calcular después sus parámetros considerando que el número total de muestras distintas de tamaño 3 que pueden obtenerse a partir de una población de 5 elementos es

$$\frac{5!}{3!\,(5-3)!} = 10$$

Dichas muestras son las siguientes, junto con sus promedios aritméticos correspondientes.

	$\bar{x}_i$		$\bar{x}_i$
1, 2, 3	6/3	3, 4, 5	12/3
1, 2, 4	7/3	3, 4, 1	8/3
1, 2, 5	8/3	4, 5, 1	10/3
2, 3, 4	9/3	4, 5, 2	11/3
2, 3, 5	10/3	5, 1, 3	9/3

Para calcular la media y la desviación estándar, se emplea la siguiente tabla

$\bar{x}_i$	6/3	7/3	8/3	8/3	9/3	9/3	10/3	10/3	11/3	12/3
$\bar{x}_i^2$	36/9	49/9	64/9	64/9	81/9	81/9	100/9	100/9	121/9	144/9

$$\sum_{i=1}^{10} \bar{x}_i = 90/3 \quad , \quad \sum_{i=1}^{10} \bar{x}_i^2 = 840/9$$

III (W) 8

$$\mu_{\bar{x}} = \bar{\bar{x}} = \frac{1}{10} \sum_{i=1}^{10} \bar{x}_i = \frac{1}{10} \cdot \frac{90}{3} = 3$$

$$\sigma^2_{\bar{x}} = \frac{1}{10} \sum_{i=1}^{10} \bar{x}_i^2 - \bar{\bar{x}}^2 = \frac{1}{10} \cdot \frac{840}{9} - (3)^2 =$$

$$= 9.333 - 9.000 = 0.333 \qquad \Rightarrow \sigma_{\bar{x}} = \sqrt{0.333} = 0.577$$

Es decir, $\mu_{\bar{x}} = 3$ y $\sigma_{\bar{x}} = 0.577$

Segundo Procedimiento.

Por tratarse de una población finita, se verifica que

$$\mu_{\bar{x}} = \mu \qquad y \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N_p - n}{N_p - 1}}$$

en donde $N_p = 5$, $n = 3$ y $\mu = 3$.

El valor de σ de la población es

$$\sigma^2 = \frac{1+4+9+16+25}{5} - (3)^2 = \frac{55}{5} - 9 = 11 - 9 = 2$$

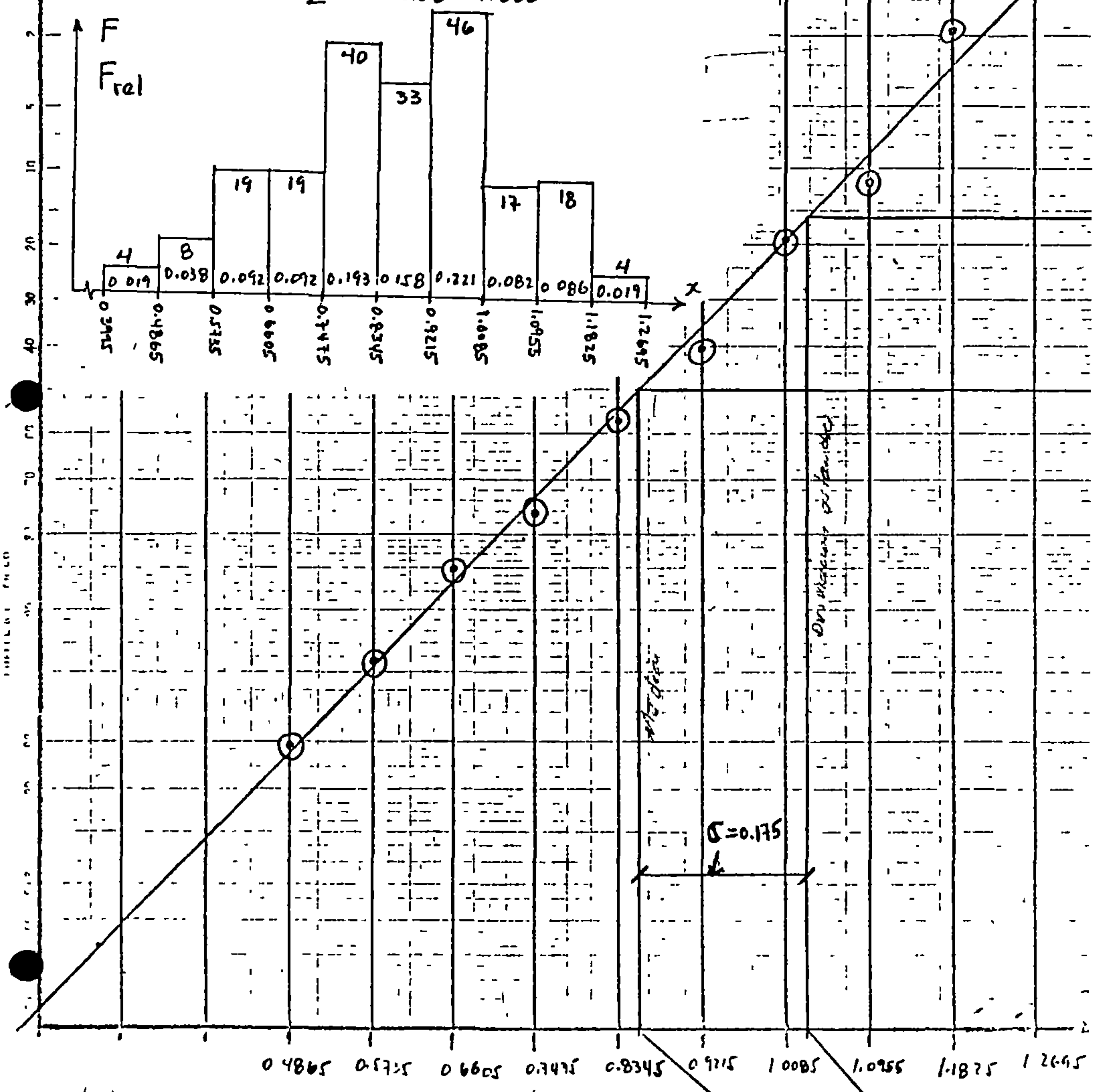
Por lo tanto, $\sigma = \sqrt{2} = 1.4145$ y

$$\sigma_{\bar{x}} = \frac{1.4145}{\sqrt{3}} \sqrt{\frac{5-3}{5-1}} = (0.8164)(0.7071) = 0.577$$

Es decir, $\mu_{\bar{x}} = 3$ y $\sigma_{\bar{x}} = 0.577$

Comparando los resultados, se puede observar que ambos procedimientos conducen a la obtención de los mismos valores de $\mu_{\bar{x}}$ y $\sigma_{\bar{x}}$ para la distribución muestral del promedio aritmético.

LRI	LRS	F	Frel	Frel AC.
0.3995	0.4865	4	0.019	0.019
0.4865	0.5735	8	0.038	0.057
0.5735	0.6605	19	0.092	0.149
0.6605	0.7475	19	0.092	0.241
0.7475	0.8345	40	0.193	0.434
0.8345	0.9215	33	0.158	0.592
0.9215	1.0085	46	0.221	0.813
1.0085	1.0955	17	0.082	0.895
1.0955	1.1825	18	0.086	0.981
1.1825	1.2695	4	0.019	1.000
	Σ	208	1.000	



papel de distribución de probabilidad normal

$$z_1 = \frac{0.3995 - 0.85}{0.17} = -\frac{0.4505}{0.17} = -2.65$$

$$z_2 = \frac{0.4865 - 0.85}{0.17} = -\frac{0.3635}{0.17} = -2.138$$

$$z_3 = \frac{0.5735 - 0.85}{0.17} = -\frac{0.2765}{0.17} = -1.626$$

$$z_4 = \frac{0.6605 - 0.85}{0.17} = -\frac{0.1895}{0.17} = -1.115$$

$$z_5 = \frac{0.7475 - 0.85}{0.17} = -\frac{0.1025}{0.17} = -0.603$$

$$z_6 = \frac{0.8345 - 0.85}{0.17} = -\frac{0.0155}{0.17} = -0.091$$

$$z_7 = \frac{0.9215 - 0.85}{0.17} = \frac{0.0715}{0.17} = 0.420$$

$$z_8 = \frac{1.0085 - 0.85}{0.17} = \frac{0.1585}{0.17} = 0.932$$

$$z_9 = \frac{1.0955 - 0.85}{0.17} = \frac{0.2455}{0.17} = 1.444$$

$$z_{10} = \frac{1.1825 - 0.85}{0.17} = \frac{0.3325}{0.17} = 1.956$$

$$z_{11} = \frac{1.2695 - 0.85}{0.17} = \frac{0.4195}{0.17} = 2.467$$

DATOS DE LA POBLACION

$\mu = 0.85$

$\sigma = 0.17$

Límite Reales de clase, L_i	Z_{L_i}	Area de 0 a Z_{L_i}	Area de Cada clase	Frecuencia Esperada	Frecuencia Observada
0.3995	-2.65	0.4960			
			(-) 0.0125	2.6 (3)	4
0.4865	-2.138	0.4835			
			(-) 0.0355	7.68 (8)	8
0.5735	-1.626	0.4480			
			0.0813	16.91 (17)	19
0.6605	-1.115	0.3667			
			(-) 0.140	29.12 (29)	19
0.7475	-0.603	0.2267			
			0.1903	39.58 (40)	40
0.8345	-0.091	0.0364			
			0.1992	41.43 (41)	33
0.9215	0.420	0.1628			
			0.1615	33.59 (34)	46
1.0085	0.932	0.3243			
			0.1014	21.09 (21)	17
1.0955	1.444	0.4257			
			0.049	10.59 (11)	18
1.1825	1.956	0.4747			
			0.0185	3.85 (4)	4
1.2695	2.467	0.4932			

Frecuencia Esperada = Area de Cada clase × 208

Prueba χ^2 de "bondad de ajuste"

El valor de la estadística χ^2 para la prueba de bondad de ajuste es

$$\chi^2 = \frac{(\theta_1 - e_1)^2}{e_1} + \frac{(\theta_2 - e_2)^2}{e_2} + \cdots + \frac{(\theta_K - e_K)^2}{e_K} = \sum_{j=1}^{K} \frac{(\theta_j - e_j)^2}{e_j}$$

Siendo θ_j los valores observados y e_j los valores esperados (después de realizado el ajuste a alguna distribución teórica)

Siendo K el número de valores esperados, y K el de observados, se debe verificar que

$$\sum_{j=1}^{K} \theta_j = \sum_{j=1}^{K} e_j = N$$

en donde N es el total de datos (o frecuencia total).

Para nuestro ajuste a la distribución normal, tendríamos

$$\chi^2 = \frac{(4-2.6)^2}{2.6} + \frac{(8-7.68)^2}{7.68} + \frac{(19-16.91)^2}{16.91} + \frac{(19-29.12)^2}{29.12} + \frac{(40-39.58)^2}{39.58} + \frac{(33-41.43)^2}{41.43}$$

$$+ \frac{(46-33.59)^2}{33.59} + \frac{(17-21.09)^2}{21.09} + \frac{(18-10.59)^2}{10.59} + \frac{(4-3.85)^2}{3.85}$$

$$= 16.8$$

Para la prueba de hipótesis con nivel de confianza del 99%, se tendría

$\nu = K - 1 - m$ (m = número de parámetros de la población empleados para obtener las e_j)

$= 10 - 1 - 2 = 7$

El valor $\chi^2_{0.99}$ para $\nu = 7$ grados de libertad es (de la tabla 8)

$$\chi^2_{0.99} = 18.5$$

Como el valor χ^2 experimental es menor que el de la tabla ($16.8 < 18.5$) se acepta la hipótesis de normalidad de la población con un nivel de confianza del 99%.

Ejemplo.—

Considérese que de una población X se obtienen tres muestras posibles, cuyos correspondientes promedios aritméticos son 3, 7 y 8. De otra población Y se extraen dos muestras posibles, con promedios 2 y 4, respectivamente. Se pide obtener los parámetros de las distribuciones muestrales de las diferencias y las sumas de los promedios aritméticos.

Primer procedimiento.—

a) Sumas. Todas las posibles sumas de promedios aritméticos de X con los de Y serían

$$\begin{matrix} 3+2 & 7+2 & 8+2 \\ 3+4 & 7+4 & 8+4 \end{matrix} \Rightarrow \begin{matrix} 5 & 9 & 10 \\ 7 & 11 & 12 \end{matrix}$$

Es decir,

$$\mu_{\bar{X}+\bar{Y}} = \frac{5+7+9+10+11+12}{6} = \frac{54}{6} = 9$$

$$\sigma^2_{\bar{X}+\bar{Y}} = \frac{(5-9)^2+(7-9)^2+(9-9)^2+(10-9)^2+(11-9)^2+(12-9)^2}{6} = \frac{34}{6} = \frac{17}{3}$$

b) Diferencias. Todas las posibles diferencias de promedios aritméticos de X con los de Y serían

$$\begin{matrix} 3-2 & 7-2 & 8-2 \\ 3-4 & 7-4 & 8-4 \end{matrix} \Rightarrow \begin{matrix} 1 & 5 & 6 \\ -1 & 3 & 4 \end{matrix}$$

Es decir,

$$\mu_{\bar{X}-\bar{Y}} = \frac{-1+1+3+4+5+6}{6} = \frac{18}{6} = 3$$

$$\sigma^2_{\bar{X}-\bar{Y}} = \frac{(-1-3)^2+(1-3)^2+(3-3)^2+(4-3)^2+(5-3)^2+(6-3)^2}{6} = \frac{34}{6} = -$$

Segundo procedimiento.–

Sabemos que

$$\mu_{\bar{x}+\bar{y}} = \mu_{\bar{x}} + \mu_{\bar{y}} \; ; \; \sigma^2_{\bar{x}+\bar{y}} = \sigma^2_{\bar{x}} + \sigma^2_{\bar{y}}$$

$$\mu_{\bar{x}-\bar{y}} = \mu_{\bar{x}} - \mu_{\bar{y}} \; ; \; \sigma^2_{\bar{x}-\bar{y}} = \sigma^2_{\bar{x}} + \sigma^2_{\bar{y}}$$

Por ello,

$$\mu_{\bar{x}} = \frac{3+7+8}{3} = \frac{18}{3} = 6$$

$$\mu_{\bar{y}} = \frac{2+4}{2} = \frac{6}{2} = 3$$

$$\sigma^2_{\bar{x}} = \frac{(3-6)^2+(7-6)^2+(8-6)^2}{3} = \frac{14}{3}$$

$$\sigma^2_{\bar{y}} = \frac{(2-3)^2+(4-3)^2}{2} = \frac{2}{2} = 1$$

a) Sumas.

$$\mu_{\bar{x}+\bar{y}} = 6+3 = 9$$

$$\sigma^2_{\bar{x}+\bar{y}} = \frac{14}{3}+1 = \frac{17}{3}$$

b) Diferencias.

$$\mu_{\bar{x}-\bar{y}} = 6-3 = 3$$

$$\sigma^2_{\bar{x}-\bar{y}} = \frac{14}{3}+1 = \frac{17}{3}$$

Se debe observar que ambos procedimientos conducen a los mismos resultados.

Table 8.1

Standard Errors for Some Sampling Distributions

Sampling Distribution	Standard Error	Special Remarks
Means	$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{N}}$	This is true for large or small samples. The sampling distribution of means is very nearly normal for $N \geq 30$ even when the population is non-normal. $\mu_{\bar{X}} = \mu$, the population mean in all cases.
Proportions	$\sigma_p = \sqrt{\frac{p(1-p)}{N}} = \sqrt{\frac{pq}{N}}$	The remarks made for means apply here as well. $\mu_p = p$ in all cases.
Standard Deviations	(1) $\sigma_s = \frac{\sigma}{\sqrt{2N}}$ (2) $\sigma_s = \sqrt{\frac{\mu_4 - \mu_2^2}{4N\mu_2}}$	For $N \geq 100$, the sampling distribution of s is very nearly normal. σ_s is given by (1) only if the population is normal (or approximately normal). If the population is non-normal, (2) can be used. Note that (2) reduces to (1) when $\mu_2 = \sigma^2$ and $\mu_4 = 3\sigma^4$, which is true for normal populations. For $N \geq 100$, $\mu_s = \sigma$ very nearly.
Medians	$\sigma_{\text{med}} = \sigma\sqrt{\frac{\pi}{2N}} = \frac{1.2533\sigma}{\sqrt{N}}$	For $N \geq 30$, the sampling distribution of the median is very nearly normal. The given result holds only if the population is normal (or approximately normal). $\mu_{\text{med}} = \mu$.
First and Third Quartiles	$\sigma_{Q_1} = \sigma_{Q_3} = \frac{1.3626\sigma}{\sqrt{N}}$	The remarks made for medians apply here as well. μ_{Q_1} and μ_{Q_3} are very nearly equal to the first and third quartiles of the population. Note that $\sigma_{Q_2} = \sigma_{\text{med}}$.
Deciles	$\sigma_{D_1} = \sigma_{D_9} = \frac{1.7094\sigma}{\sqrt{N}}$, $\sigma_{D_2} = \sigma_{D_8} = \frac{1.4288\sigma}{\sqrt{N}}$ $\sigma_{D_3} = \sigma_{D_7} = \frac{1.3180\sigma}{\sqrt{N}}$ $\sigma_{D_4} = \sigma_{D_6} = \frac{1.2680\sigma}{\sqrt{N}}$	The remarks made for medians apply here as well. $\mu_{D_1}, \mu_{D_2}, \ldots$ are very nearly equal to the first, second, ... deciles of the population. Note that $\sigma_{D_5} = \sigma_{\text{med}}$.
Semi-interquartile Ranges	$\sigma_Q = \frac{0.7867\sigma}{\sqrt{N}}$	The remarks made for medians apply here as well. μ_Q is very nearly equal to the population semi-interquartile range.
Variances	(1) $\sigma_{s^2} = \sigma^2\sqrt{\frac{2}{N}}$ (2) $\sigma_{s^2} = \sqrt{\frac{\mu_4 - \mu_2^2}{N}}$	The remarks made for standard deviation apply here as well. Note that (2) yields (1) in case the population is normal. $\mu_{s^2} = \sigma^2(N-1)/N$, which is very nearly σ^2 for large N.
Coefficients of Variation	$\sigma_V = \frac{v}{\sqrt{2N}}\sqrt{1+2v^2}$	Here $v = \sigma/\mu$ is the population coefficient of variation. The given result holds for normal (or nearly normal) populations and $N \geq 100$.

Ejemplo.– Cierta población consiste de cuatro elementos, uno de los cuales está defectuoso. Calcular

a) la proporción de elementos defectuosos, p
b) la desviación estándar, σ

Si de esa población se extraen todas las muestras posibles de tamaño 2 con remplazo, calcular

c) la media de la distr. muestral de proporciones, μ_p
d) la desv. est. 〃 〃 〃 〃 〃 〃, σ_p

Solución

a) $p = \dfrac{\text{número de elementos defectuosos}}{\text{número de elementos de la población}} = \dfrac{1}{4} = 0.25$

b) Denotando con $\underline{1}$ al elemento defectuoso y con $\underline{0}$ a los que están en buen estado, se tiene

$$\sigma = \sqrt{\frac{(0-1/4)^2+(0-1/4)^2+(0-1/4)^2+(1-1/4)^2}{4}} = \sqrt{\frac{1/16+1/16+1/16+9/16}{4}}$$

$$= \sqrt{12/16} = \sqrt{3/16} \quad \left(\text{ó } \sigma = \sqrt{pq} = \sqrt{1/4 \times 3/4} = \sqrt{3/16} \right)$$

c) Denotando con b_1, b_2, b_3 a los elementos en buen estado y con d al defectuoso, las muestras posibl son

(b_1, b_1)	(b_1, b_2)	(b_1, b_3)	(b_1, d)
(b_2, b_1)	(b_2, b_2)	(b_2, b_3)	(b_2, d)
(b_3, b_1)	(b_3, b_2)	(b_3, b_3)	(b_3, d)
(d, b_1)	(d, b_2)	(d, b_3)	(d, d)

ó sea

(0, 0)	(0, 0)	(0, 0)	(0, 1)
(0, 0)	(0, 0)	(0, 0)	(0, 1)
(0, 0)	(0, 0)	(0, 0)	(0, 1)
(1, 0)	(1, 0)	(1, 0)	(1, 1)

La distribución muestral de proporcions es, entonces

0	0	0	0.5
0	0	0	0.5
0	0	0	0.5
0.5	0.5	0.5	1.0

Por lo que

$$\mu_p = \frac{9(0)+6(0.5)+1}{16} = \frac{3+1}{16} = \frac{4}{16} = 1/4 = p$$

d) $$\sigma_p = \sqrt{\frac{9(0-1/4)^2+6(1/2-1/4)^2+(1-1/4)^2}{16}}$$

$$= \sqrt{\frac{9(1/16)+6(1/16)+9/16}{16}} = \sqrt{\frac{24/16}{16}}$$

$$= \sqrt{3/32}$$

o $$\sigma_p = \sqrt{\frac{pq}{n}} = \sqrt{\frac{1/4 \times 3/4}{2}} = \sqrt{\frac{3/16}{2}} = \sqrt{3/32}$$

$$= \frac{\sigma}{\sqrt{n}} = \frac{\sqrt{3/16}}{\sqrt{2}} = \sqrt{\frac{3/16}{2}} = \sqrt{3/32}$$

Intervalo de confianza para determinar tamaño de muestra

Supongamos que quisiéramos determinar con un ±1% de precisión en la estimación a la media de la población de tubos de albañal tratada en el ejemplo de la página 66 del libro. Para hacerlo debemos utilizar los valores de $\bar{X}$ (32 cm) y σ (2 cm) obtenidos de la muestra de 100 tubos. Hay que recordar que la expresión $\pm Z_c \sigma/\sqrt{n}$ nos va a proporcionar los dos números que hay que sumar y restar al promedio aritmético de la muestra para definir el intervalo de confianza de μ para un nivel de confianza deseado. Por lo tanto, si deseamos un ±1% de precisión en la estimación de μ escribimos

$$0.01\,\bar{X} = Z_c \frac{\sigma}{\sqrt{n}}$$

O, lo que es lo mismo

$$(0.01)(32) = 1.96\,(2/\sqrt{n})$$

$$0.32 = 3.92/\sqrt{n}$$

$$\Rightarrow \quad \sqrt{n} = \frac{3.92}{0.32} = 12.3$$

$$\therefore \; n = (12.3)^2 \doteq 150$$

O sea que se requeriría una muestra de 150 tubos para que la precisión en la estimación de μ sea de ±1%, a un nivel de confianza de 95%.

El mismo procedimiento anterior se puede emplear si se desea hacer el cálculo del valor de n para otros niveles de confianza.

Ejemplo.— a) Supóngase que los clientes arriban a un sistema a razón de uno cada 12 minutos, y que cada persona es atendida en ocho minutos. Se desea conocer el número medio de personas en el sistema, así como el tiempo medio que pasa cada persona en el mismo.

Solución.— Considerando que $\lambda = 1/12$ y $\mu = 1/8$, se obtiene

$$L = \frac{\lambda}{\mu - \lambda} = \frac{1/12}{1/8 - 1/12} = \frac{2/24}{3/24 - 2/24} = \frac{2/24}{1/24} = 2 \text{ personas}$$

$$W = \frac{1}{\mu - \lambda} = \frac{1}{1/8 - 1/12} = \frac{1}{3/24 - 2/24} = \frac{1}{1/24} = 24 \text{ minutos}$$

Por lo tanto, el número medio de clientes que se encuentran en el sistema es 2, siendo el tiempo que pasa cada uno en el mismo igual a 24 minutos.

b) Supóngase que en el ejemplo anterior se incrementan los arribos en cerca del 20%. ¿Cuánto valen L y W?

Solución.— Ahora $\lambda \doteq 1/10$. Por lo tanto,

$$L = \frac{1/10}{1/8 - 1/10} = \frac{8/80}{10/80 - 8/80} = \frac{8/80}{2/80} = 4 \text{ personas}$$

$$W = \frac{1}{1/8 - 1/10} = \frac{1}{10/80 - 8/80} = \frac{1}{2/80} = 40 \text{ minutos}$$

Lo anterior indica que un aumento de cerca del 20% en la frecuencia de arribos duplica el número medio de personas que se encuentran en el sistema.

El aumento en el tamaño de la muestra no es directamente proporcional al aumento en la precisión. La siguiente comparación lo demuestra. Aquí $\bar{X}=100$ y $\sigma=10$

Precisión deseada	Mejoría en la precisión	tamaño de la muestra, n	Aumento en el tamaño de n
± 5%		16	
± 4%	± 1%	25	9
± 3%	± 1%	45	20
± 2%	± 1%	100	55
± 1%	± 1%	400	300
± 0.5%	± 0.5%	1,600	1,200
± 0.3%	± 0.2%	3,600	2,000
± 0.2%	± 0.1%	10,000	6,400
± 0.1%	± 0.1%	40,000	30,000

Ejemplo.—

a) Supóngase que los clientes arriban a un sistema a razón de uno cada 12 minutos, y que cada persona es atendida en ocho minutos. Se desea conocer el número medio de personas en el sistema, así como el tiempo medio que pasa cada persona en el mismo.

Solución.— Considerando que $\lambda = 1/12$ y $\mu = 1/8$, se obtiene

$$L = \frac{\lambda}{\mu - \lambda} = \frac{1/12}{1/8 - 1/12} = \frac{2/24}{3/24 - 2/24} = \frac{2/24}{1/24} = 2 \text{ personas}$$

$$W = \frac{1}{\mu - \lambda} = \frac{1}{1/8 - 1/12} = \frac{1}{3/24 - 2/24} = \frac{1}{1/24} = 24 \text{ minutos}$$

Por lo tanto, el número medio de clientes que se encuentran en el sistema es 2, siendo el tiempo que pasa cada uno en el mismo igual a 24 minutos.

b) Supóngase que en el ejemplo anterior se incrementan los arribos en cerca del 20%. ¿Cuánto valen L y W?

Solución.— Ahora $\lambda \doteq 1/10$. Por lo tanto,

$$L = \frac{1/10}{1/8 - 1/10} = \frac{8/80}{10/80 - 8/80} = \frac{8/80}{2/80} = 4 \text{ personas}$$

$$W = \frac{1}{1/8 - 1/10} = \frac{1}{10/80 - 8/80} = \frac{1}{2/80} = 40 \text{ minutos}$$

Lo anterior indica que un aumento de cerca del 20% en la frecuencia de arribos duplica el número medio de personas que se encuentran en el sistema.

DISTRIBUCION EXPONENCIAL

Se dice que una variable aleatoria X tiene distribución exponencial con parámetro λ ($\lambda>0$) si su función de densidad de probabilidades se encuentra dada por

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & ; \quad x \geq 0 \\ 0 & ; \quad x < 0 \end{cases}$$

Se puede demostrar que la esperanza y la varianza de una variable de este tipo están dadas por

$$E[X] = 1/\lambda \qquad ; \quad Var(X) = 1/\lambda^2$$

Una propiedad importante de la variable aleatoria distribuida exponencialmente es que no tiene <u>memoria</u>. Es decir, si X es la variable aleatoria "tiempo de duración de una guitarra", dicha propiedad indica que la probabilidad de que sirva una guitarra por lo menos t+s horas sabiendo que ya ha dado servicio durante t horas es igual a la probabilidad de que dure al menos s horas. Matemáticamente esto se escribe

$$P\{X > s+t \mid X > t\} = P\{X > s\} \quad \forall\, s,t \geq 0$$

En otras palabras, lo que quiere decir lo anterior es que si el instrumento se encuentra en servicio en el tiempo t, entonces la distribución del tiempo de servicio restante es igual a la distribución del tiempo que lleva la guitarra en uso, es decir, el instrumento no recuerda que ya ha "tocado" durante un tiempo t!.

1º) Sistema de cola (línea de espera) con una estación de servicio.

1. La gente arriba a una sola estación de acuerdo con un proceso de Poisson con parámetro λ. Es decir

$$E[\text{ t. de arribo}] = 1/\lambda \quad (\text{variables aleatorias indep de dist exponencial})$$

2. Los tiempos de servicio sucesivos se suponen variables aleatorias exponenciales con media

$$E[\text{t. de servicio}] = \frac{1}{\mu}$$

3. Nos interesa conocer lo siguiente:

L = número medio de personas en el sistema
L_Q = " " " " esperando en la cola
W = tiempo promedio que la persona está en el sistema
W_Q = " " " " " espera en la cola

4. Para calcular lo anterior, definimos

$$X(t) = \text{número de personas en el sistema en el tiempo } \underline{t}$$

y $$P_n(t) = P\{X(t) = n\}$$

5. $P_n(t)$ se podría obtener mediante ecuaciones diferenciales de difícil solución; por tanto, nos ocuparemos en encontrar

$$P_n = \lim_{t\to\infty} P_n(t) = \lim_{t\to\infty} P\{X(t)=n\}$$

(Probabilidad de que a la larga haya en el sist. exactamente n pers

6. El número medio de personas en el sistema es, de acuerdo con la definición de P_n,

$$L = \sum^{\infty} n P_n$$

7. Siempre que hay n personas en el sistema $(n \geq 1)$, a una de ellas se le da servicio y $n-1$ esperan en la cola; por ello,

$$L_Q = \sum_{n=1}^{\infty} (n-1) P_n$$

8. Para obtener el tiempo medio que está una persona en el sistema condicionemos al número de personas que ya están cuando esa persona llega; es decir

$$W = \sum_{n=0}^{\infty} E\left[\begin{array}{l} \text{tiempo que está la} \\ \text{persona en el sist} \end{array} \middle| \begin{array}{l} \text{había } n \text{ personas} \\ \text{en el sist. cuando} \\ \text{arribó} \end{array}\right] \times P\{n \text{ personas en el sist cuando arriba}\}$$

8.1 Si hay $n=0$ personas en el sistema cuando arriba un individuo, el tiempo que pasa en el mismo es el tiempo medio en el que le darán servicio, es decir, $1/\mu$.

8.2 Si hay $n \geq 1$ personas en el sistema al arribar el individuo, una de ellas está recibiendo servicio y otras $n-1$ están en la cola. Puesto que el tiempo de servicio es exponencial sin memoria, el tiempo medio que estará en el sistema el individuo es

$$W = \sum_{n=0}^{\infty} \frac{1}{\mu} \Big(\underbrace{n-1}_{\text{en la cola}} + \underbrace{1}_{\text{recibiendo servicio}} + \underbrace{1}_{\text{nuestro individuo}}\Big) P_n$$

$$= \sum_{n=0}^{\infty} \frac{n+1}{\mu} P_n$$

9. El tiempo medio que un individuo está en la cola es igual al tiempo medio que pasa en el sistema menos el tiempo medio en el que recibe servicio es decir,

$$W_Q = \sum_{n=0}^{\infty} \frac{n+1}{\mu} P_n - \frac{1}{\mu}$$

$$= \sum_{n=0}^{\infty} \frac{n}{\mu} P_n$$

10. Para determinar las probabilidades límite P_n, para $n=0,1,2,\ldots$ supongamos que tenemos un número infinito de cuartos numerados $0,1,2,\ldots$, y que pedimos a un individuo que entre al cuarto <u>n</u> cuando estén <u>n</u> personas en el sistema. Por ejemplo:

10.1 El individuo se meterá al cuarto nº <u>2</u> siempre que en el sistema estén dos personas.

10.2 Suponiendo que esté en el cuarto <u>2</u> y otra persona llegue al sistema, dejará ese cuarto y se irá al <u>3</u>.

10.3 Si está en el cuarto <u>2</u> y se termina de proporcionar un servicio, dejará este cuarto y se irá al <u>1</u>.

11. Si nuestro individuo a la larga entra al cuarto <u>1</u> a razón de 10 veces por hora, también deberá salir de ese cuarto el mismo número de veces por hora (o tal vez una ocasión menos). De lo anterior se desprende que para cada $n \geq 0$, <u>la proporción de veces en que el sistema entra a un estado n es igual a la proporción de veces en que sale de él</u>.

12. Consideremos el estado <u>0</u> (ninguna persona en el sistema). El proceso puede dejar ese estado únicamente si arriba una persona al sistema. Siendo P_0 la proporción de tiempo en que el proceso se encuentra en el estado <u>0</u> y λ la tasa de arribos, nuestro proceso <u>abandona</u> ese estado con una proporción λP_0.

13. Al estado <u>0</u> se puede llegar del estado <u>1</u> (una persona en el sistema) si deja el sistema una persona al habérsele proporcionado servicio. Siendo μ la tasa de servicio y P_1 la proporción de tiempo que en el sistema se encuentra una persona, el proceso <u>llega</u> al estado <u>0</u> con una proporción μP_1. De acuerdo con el punto 11, se tiene que

$$\lambda P_0 = \mu P_1$$

14. Consideremos ahora el estado $\underline{1}$ (una persona en el sistema). Puesto que este estado puede ser dejado por el arribo de una persona al sistema o porque a otra se le haya dado servicio, el proceso dejará el estado $\underline{1}$ con una tasa $\lambda + \mu$. Por tanto, el proceso <u>abandona</u> dicho estado con una proporción $(\lambda + \mu)P_1$.

15. El proceso puede llegar al estado 1 desde el estado $\underline{0}$ (por el arribo de una persona) o desde el estado $\underline{2}$ (por haberle sido dado el servicio a otra persona). Por tanto, el proceso <u>llega</u> a dicho estado con una proporción $\lambda P_0 + \mu P_2$. De acuerdo con el punto 11, se tiene que

$$(\lambda + \mu)P_1 = \lambda P_0 + \mu P_2 .$$

16. Generalizando, se tiene que

	Estado	Proporción en que el proceso <u>llega</u>		Proporción en que el proceso <u>abandona</u>
ecuaciones de Balanceo	0	λP_0	$=$	μP_1
	1	$(\lambda + \mu)P_1$	$=$	$\lambda P_0 + \mu P_2$
	$\vdots$			
	n	$(\lambda + \mu)P_n$	$=$	$\lambda P_{n-1} + \mu P_{n+1}$

17. Despejando de las ecuaciones anteriores los valores P_n $(n \geq 1)$ se tiene

$$P_1 = \frac{\lambda}{\mu} P_0$$

$$\vdots$$

$$P_{n+1} = \frac{\lambda}{\mu} P_n + \left(P_n - \frac{\lambda}{\mu} P_{n-1}\right)$$

18. Poniendo las ecuaciones anteriores en términos de P_0 se llega a

$$P_0 = P_0$$

$$P_1 = \frac{\lambda}{\mu} P_0$$

$$P_2 = \frac{\lambda}{\mu} P_1 + \left(P_1 - \frac{\lambda}{\mu} P_0\right) = \frac{\lambda}{\mu} P_1 = \left(\frac{\lambda}{\mu}\right)^2 P_0$$

$$\vdots$$

$$P_{n+1} = \frac{\lambda}{\mu} P_n + \left(P_n - \frac{\lambda}{\mu} P_{n-1}\right) \equiv \frac{\lambda}{\mu} P_n = \left(\frac{\lambda}{\mu}\right)^{n+1} P_0$$

19. Para determinar P_0 consideramos que $\sum_{n=0}^{\infty} P_n = 1$, es decir,

$$\sum_{n=0}^{\infty} P_n = \sum_{n=0}^{\infty} \left(\frac{\lambda}{\mu}\right)^n P_0 = \frac{P_0}{1 - \frac{\lambda}{\mu}} = 1$$

por lo que

$$P_0 = 1 - \frac{\lambda}{\mu}$$

$$P_1 = \left(\frac{\lambda}{\mu}\right)\left(1 - \frac{\lambda}{\mu}\right)$$

$$P_2 = \left(\frac{\lambda}{\mu}\right)^2 \left(1 - \frac{\lambda}{\mu}\right)$$

$$\vdots$$

$$P_n = \left(\frac{\lambda}{\mu}\right)^n \left(1 - \frac{\lambda}{\mu}\right)$$

20. Para las ecuaciones anteriores, el cociente λ/μ debe ser tal que

$$\frac{\lambda}{\mu} < 1$$

para que se satisfaga $0 \leq P_n \leq 1$. Ello implica también que el tiempo medio de servicio $\left(\frac{1}{\mu}\right)$ sea menor que el tiempo medio de arribos $\left(\frac{1}{\lambda}\right)$, siendo esta condición necesaria para que el sistema sea óptimo

21. Sustituyendo los valores P_n obtenidos anteriormente en las expresiones para L, L_Q, W y W_Q ya calculados, y empleando la fórmula $\sum_{n=0}^{\infty} n x^n = x(1-x)$, se llega a

$$L = \frac{\lambda}{\mu - \lambda} \qquad W = \frac{1}{\mu - \lambda}$$

$$L_Q = \frac{\lambda^2}{\mu(\mu - \lambda)} \qquad W_Q = \frac{\lambda}{\mu(\mu - \lambda)}$$

Ejemplo:

Supóngase que se tiene una población finita formada por los datos 1, 2, 3, 4, 5. Se desea conocer la media y la desviación estándar de la distribución muestral del promedio aritmético, considerando las muestras de tamaño 3 obtenidas sin remplazo.

Primer procedimiento.-

Siendo la población finita y el muestreo sin remplazo, es posible obtener la distribución muestral correspondiente, para calcular después sus parámetros considerando que el número total de muestras distintas de tamaño 3 que pueden obtenerse a partir de una población de 5 elementos es

$$\frac{5!}{3!\,(5-3)!} = 10$$

Dichas muestras son las siguientes, junto con sus promedios aritméticos correspondientes.

	$\bar{x}_i$		$\bar{x}_i$
1, 2, 3	6/3	3, 4, 5	12/3
1, 2, 4	7/3	3, 4, 1	8/3
1, 2, 5	8/3	4, 5, 1	10/3
2, 3, 4	9/3	4, 5, 2	11/3
2, 3, 5	10/3	5, 1, 3	9/3

Para calcular la media y la desviación estándar, se emplea la siguiente tabla

$\bar{x}_i$	6/3	7/3	8/3	8/3	9/3	9/3	10/3	10/3	11/3	12/3
$\bar{x}_i^2$	36/9	49/9	64/9	64/9	81/9	81/9	100/9	100/9	121/9	144/9

$$\sum_{i=1}^{10} \bar{x}_i = 90/3 \qquad \sum_{i=1}^{10} \bar{x}_i^2 = 840/9$$

$$\mu_{\bar{x}} = \bar{\bar{x}} = \frac{1}{10}\sum_{i=1}^{10} \bar{x}_i = \frac{1}{10}\cdot\frac{90}{3} = 3$$

$$\sigma^2_{\bar{x}} = \frac{1}{10}\sum_{i=1}^{10} \bar{x}_i^2 - \bar{\bar{x}}^2 = \frac{1}{10}\cdot\frac{840}{9} - (3)^2 =$$

$$= 9.333 - 9.000 = 0.333 \quad \Rightarrow \sigma_{\bar{x}} = \sqrt{0.333} = 0.577$$

Es decir, $\mu_{\bar{x}} = 3$ y $\sigma_{\bar{x}} = 0.577$

Segundo Procedimiento.

Por tratarse de una población finita, se verifica que

$$\mu_{\bar{x}} = \mu \qquad y \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N_p - n}{N_p - 1}}$$

en donde $N_p = 5$, $n = 3$ y $\mu = 3$.

El valor de σ de la población es

$$\sigma^2 = \frac{1+4+9+16+25}{5} - (3)^2 = \frac{55}{5} - 9 = 11 - 9 = 2$$

Por lo tanto, $\sigma = \sqrt{2} = 1.4145$ y

$$\sigma_{\bar{x}} = \frac{1.4145}{\sqrt{3}}\sqrt{\frac{5-3}{5-1}} = (0.8164)(0.7071) = 0.577$$

Es decir, $\mu_{\bar{x}} = 3$ y $\sigma_{\bar{x}} = 0.577$

Comparando los resultados, se puede observar que ambos procedimientos conducen a la obtención de los mismos valores de $\mu_{\bar{x}}$ y $\sigma_{\bar{x}}$ para la distribución muestral del promedio aritmético.

Ejemplo. Cierta compañía fabrica cables de acero, cuyas tensiones de ruptura tienen una media de 300 lb, con desviación estándar igual a 24 lb. Se piensa que mediante el empleo de un nuevo proceso de fabricación la tensión media de ruptura se puede incrementar.

a. Obténgase el tamaño de muestra necesario para rechazar el proceso antiguo con $\alpha = 0.01$ y cuando el promedio aritmético de la muestra sea > 307

b. Considerando la regla de decisión adoptada en el inciso anterior, encuéntrese la probabilidad de aceptar el proceso antiguo cuando en realidad el nuevo proceso ha incrementado la tensión media de ruptura hasta 310 lb. Considérese que σ es igual a 24 lb

Solución.- a. Si μ es la media de las tensiones de ruptura, entonces se tiene que decidir entre las hipótesis

$$H_0 : \mu = 300 \text{ lb}$$
$$H_1 : \mu > 300 \text{ lb}$$

Para una prueba de hipótesis de una cola, se rechazará la hipótesis nula con $\alpha = 0.01$ si el valor estandarizado de la muestra (307 lb) es mayor de 2.33 (de tabla 7), es decir

$$z_c = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \Rightarrow 2.33 = \frac{307 - 300}{24/\sqrt{n}}$$

$$= \frac{7}{24/\sqrt{n}} = \frac{7\sqrt{n}}{24}$$

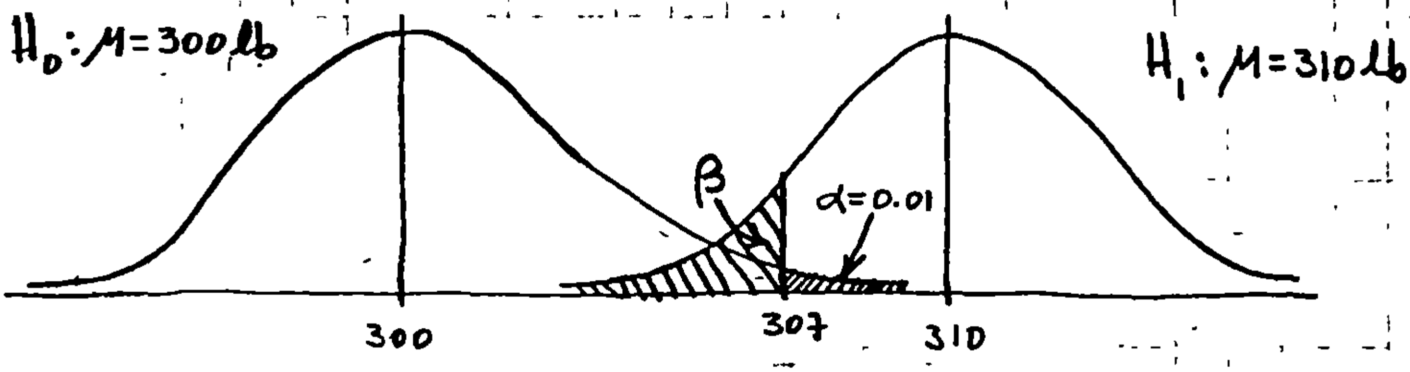
O sea que

$$n = \left(\frac{24 \times 2.33}{7}\right)^2 \doteq 8^2 = 64$$

b. Considérense las hipótesis

$$H_0: \mu = 300 \text{ lb}$$
$$H_1: \mu = 310 \text{ lb}$$

Las distribuciones muestrales correspondientes serían



La probabilidad de aceptar el proceso antiguo con $\mu = 300$ lb cuando en realidad el proceso nuevo proporciona una $\mu = 310$ lb se representa por la región β en la figura anterior. Dicha probabilidad es

$\beta =$ área bajo la curva normal para H_1 a la izquierda de $z = \dfrac{307-310}{24/\sqrt{64}} = -1.00$

$$= 0.5000 - 0.3413 = 0.1587$$

Entonces, $\beta = 0.1587$ es la probabilidad de aceptar H_0 siendo H_1 verdadera, es decir, es la probabilidad de cometer un error de tipo II.

Comentario.– Conviene observar en la figura de arriba que si se aumenta el valor del promedio aritmético de la muestra (307 lb) se puede disminuir el valor de α, pero a la vez se aumenta el de β. Por otro lado, si se disminuye el valor del promedio aritmético, el valor de β disminuye, pero el de α aumenta. La forma en la cual

se podría disminuir el valor de β para un valor dado de α sería aumentando el tamaño de la muestra, lo cual no siempre es posible. Para nuestro ejemplo se tendría lo siguiente:

1. Con $n=100$ y $\alpha = 0.01$

$$2.33 = \frac{\bar{x}-300}{24/\sqrt{100}} = \frac{\bar{x}-300}{2.4}$$

$$\Rightarrow \bar{x} = 305.6$$

Por lo cual

β = área bajo la curva normal para H_1 a la izquierda de $z = \frac{305.6-310}{2.4} = -1.833$

$$= 0.5000 - 0.4666 = 0.0334$$

2. Con $n = 225$ y $\alpha = 0.01$

$$2.33 = \frac{\bar{x}-300}{24/\sqrt{225}} = \frac{\bar{x}-300}{1.6}$$

$$\Rightarrow \bar{x} = 303.73$$

Por lo cual

β = área bajo la curva normal para H_1 a la izquierda de $z = \frac{303.73-310}{1.6} = -3.92$

$= 0.5000 - 0.5000 = 0$ (hasta cuatro decimales)

DIRECTORIO DE ASISTENTES AL CURSO DE PROBABILIDAD Y ESTADISTICA, FUNDAMENTOS Y APLICACIONES (DEL 17 DE FEBRERO AL 25 DE MARZO DE 1976)

NOMBRE Y DIRECCION	EMPRESA Y DIRECCION
1. GUSTAVO H. AGUILAR RUIBAL Cholula No. 109-1 Col. Condesa México 11, D. F. Tel· 5-15-06-18	SERVICIOS INDUSTRIALES PEÑOLES,S.C. Paseo de la Reforma 383 Col. Cuauhtémoc México 5, D. F. Tel: 5-25-92-20
2. ING. HECTOR ARMAND Cerro Compostela No. 12 Campestre Churubusco México 21, D. F. Tel: 5-49-10-38	PETROLEOS MEXICANOS Av. Marina Nacional No. 329 Col. Anáhuac México 13, D. F. Tel: 5-31-63-20
3. ING. PEDRO C. CABRERA CAVAZOS Las Noche Buenas No. 52 Las Margaritas Edo. de México Tel· 3-97-84-54	SECRETARIA DEL PATRIMONIO NACIONAL Insurgentes Sur No. 552-1o.Piso Col. Roma Sur México 7, D. F. Tel: 5-64-97-23
4. ING. ARMANDO CALDERON CAULLIERS Carlota No. 132-6 Col. Guadalupe Tepeyac México 14, D. F. Tel 5-37-78-41	SECRETARIA DEL PATRIMONIO NACIONAL Insurgentes Sur No. 552 Col. Roma Sur México 7, D. F. Tel: 5-64-97-23
5. LIC. ALBERTO CARREÑO ZEPEDA Donceles 90-502 México 1, D. F. Tel: 5-12-85-72	PRODUCTOS FORESTALES MEXICANOS Av. Hidalgo No. 5-8o. Piso México 1, D. F. Tel: 5-12-17-16
6. LIC. ANTONIO CARRILLO ARRONTE Av. Contreras No. 401 San Jerónimo México 20, D. F. Tel: 5-95-35-49	JOHNSON & JOHNSON DE MEXICO, S. A. Av. Ermita Iztapalapa No. 557 Col. Esmeralda México 13, D. F. Tel: 5-82-07-11
7. MARIO A. COELLO LEMARROY Ret. 11 de Ignacio Zaragoza U-4 Depto. 1008 Jardín Balbuena México 9, D. F. Tel: 7-62-74-49	DEPARTAMENTO DELDISTRITO FEDERAL Presidente Carranza No. 51 Coyoacán México 21, D. F. Tel: 5-54-60-75

DIRECTORIO DE ASISTENTES AL CURSO DE PROBABILIDAD Y ESTADISTICA, FUNDAMENTOS Y APLICACIONES (DEL 17 DE FEBRERO AL 25 DE MARZO DE 1976)

NOMBRE Y DIRECCION	EMPRESA Y DIRECCION
8. ING. SALVADOR DIAZ DIAZ Luis Kuhne No. 32 Col. Las Aguilas México 20, D. F. Tel: 5-93-14-93	INGENIEROS Y ARQUITECTOS, S. A. Minería No. 145 Edificio "A" ler. Piso Col. Escandón México 18, D. F. Tel: 5-36-49-58
9. RAUL DURAND ARIAS Oriente 63 No. 243 Col. Villa de Cortés México 13, D. F. Tel: 5-79-12-45	DIRECCION GENERAL DE PLANEACION EDUCATIVA-SUBDIRECCION DE EVALUACION Av. Instituto Politécnico Nacional No. 3600 Zacatenco México 14, D. F. Tel: 5-86-82-77 Ext. 142
10. ING. JOSE J. ENRIQUEZ FELIX San Juan de Letran 466-1403 Nonoalco México 3, D. F. Tel: 5-97-12-66	UNIVERSIDAD TECNOLOGICA DE MEXICO Marina Nacional No. 162 México, D. F. Tel: 5-27-03-16
11. SR. ALFREDO ESCALANTE MUNGUIA Andrinos No. 46 Manz. 114 Valle de las Flores Edo. de México Tel: 4-07-34	S.C.T.-D.G.T. CIDET Niño Perdido y Cumbres de Acultzingo Col. Narvarte México, D. F. Tel: 5-30-15-06
12. ARMANDO FRIAS VALDEZ Alfonso Reyes 123-9 Col. Condesa México 11, D. F. Tel: 5-16-84-87	SECRETARIA DE MARINA Insurgentes Sur No. 345 Col. Roma Sur México 7, D. F. Tel: 5-64-54-27
13. SR. GREGORIO FUENTES JUAREZ Av. Juárez y Plaza Asunción Xochimilco, D. F. Tel: 6-76-05-64	I.P.N. ESIA EDIF. 4 Zacatenco México 14, D. F. Tel: 5-86-28-05

...DE ASISTENTES AL CURSO DE PROBABILIDAD Y ESTADISTICA, ... Y APLICACIONES (DEL 17 DE FEBRERO AL 23 DE MARZO DE

...E Y DIRECCION	EMPRESA Y DIRECCION
... ROBERTO GALLEGOS MENESES Sta. Ma. La Redonda 209-6 Col. Peralvillo México 2, D. F. Tel: 5-26-70-62	COMISION FEDERAL DE ELECTRICIDAD Ródano No. 14 Col. Cuauhtémoc México 5, D. F. Tel: 5-33-71-33
15. RAMON GAMEZ ARISTA Nonoalco 81-4 Depto. 319 Unidad Nonoalco México 3, D. F. Tel. 5-83-62-54	CIA. DE LUZ Y FUERZA DEL CENTRO, S. A. Melchor Ocampo No. 171 Col. Anáhuac México, D. F. Tel: 5-73-66-68
16. LIC. ROBERTO M. GARCIA ESCOTO Newton 105-13 Col. Polanco México 5, D. F. Tel. 5-31-70-98	COMISION FEDERAL DE ELECTRICIDAD Ródano No. 14 Col. Cuauhtémoc México, D. F. Tel: 5-53-71-33
17. ING. ARTURO GARCIA GALINDO Ramón Prida No. 30 Col. Jardín Balbuena México 9, D. F. Tel· 5-52-11-93	COMISION FEDERAL DE ELECTRICIDAD Paseo de la Reforma No. 107-4o.P. Col. San Rafael México, D. F. Tel: 5-91-09-93
18. JOAS GOMEZ GARCIA Alvaro Obregón No. 40-2 Col. Roma México 7, D. F. Tel: 5-84-85-40	INSTITUTO MEXICANO PARA LA INFANCIA Y LA FAMILIA Emiliano Zapata No. 340 Col. Portales México 13, D. F. Tel: 5-75-37-11
19. ING. IGNACIO GONZALEZ GOMEZ Turquesa No. 71 Col. Estrella México 15, D. F. Tel: 5-17-89-43	CIA. DE LUZ Y FUERZA DEL CENTRO, S. A. Melchor Ocampo No. 171 Col. Anáhuac México 17, D. F. Tel: 5-92-04-13

DIRECTORIO DE ASISTENTES AL CURSO DE PROBABILIDAD Y ES
FUNDAMENTOS Y APLICACIONES (DEL 17 DE FEBRERO AL 23 D
1976)

NOMBRE Y DIRECCION	EMPRESA Y DIRECCI
20. ANTHAR GERMAN LOPEZ TIRADO Isla Cedros No. 13 Col. Prado Vallejo México 14, D. F. Tel 5-67-32-52	SECRETARIA DEL TRABAJO Y PREV SOCIAL, INSTITUTO NACIONAL DE DIOS DEL TRABAJO Dr. Vertiz No. 96 México 7, D. F.
21. MARIA DE LA CRUZ MALDONADO PEREZ 5 de febrero No. 1100 Col. Américas Unidas México 13, D. F. Tel 5-32-18-32	
22. JUAN J. MARTINEZ DE LA BARREDA Retorno 506 No. Unidad Modelo México 13, D. F. Tel. 5-82-40-73	COMISION FEDERAL DE ELECTRICIDAD Paseo de la Reforma No. 107-5o.Pi México, D. F. Tel: 5-35-27-34
23. GUILLERMO AMANDO MARTINEZ OMAÑA Rio Tigris No. 118 Col. Cuauhtémoc México 5, D. F. Tel: 5-11-68-80	CONDUMEX, S. A. Poniente 140 No. 720 Col. Industrial Vallejo México 16, D. F. Tel: 5-67-88-33 Ext. 552
24. ING. EDUARDO RAMIREZ S. Cruz Azul No. 161 Col. Industrial México 14, D. F.	FACULTAD DE INGENIERIA, UNAM Ciudad Universitaria México 20, D. F.
25. ING. NORBERTO REYNOSA RAMIREZ Calle 4 No. 82 Col. Espartaco México 22, D. F. Tel: 5-44-54-20	CONSULTOR EN CELULOSA Y PAPEL
26. ING. JESUS RIOS ALVARADO Av. Río de Churubusco 363-A Unidad Modelo México 13, D. F. Tel· 5-82-90-87	SECRETARIA DE COMUNICACIONES Y TRANSPORTES Cumbres de Acultzingo y Niño Perdido México 12, D. F. Tel: 5-19-02-73

RIO DE ASISTENTES AL CURSO DE PROBABILIDAD Y ESTADISTICA,
NTOS Y APLICACIONES (DEL 17 DE FEBRERO AL 23 DE MARZO DE

MBRE Y DIRECCION	EMPRESA Y DIRECCION
JOSE ROBLES TAMAYO panilla No. 45 Jardines de Coyoacán México 22, D. F. Tel: 5-52-07-60	SECRETARIA DE OBRAS PUBLICAS Xola 1755-4o. Piso Col. Narvarte México 13, D. F. Tel: 5-19-80-46
8. ING. SERGIO RAMON ROMERO Tecualiapan 36-X-7 Coyoacán México 21, D. F. Tel 5-54-35-69	COMISION FEDERAL DE ELECTRICIDAD Ródano No. 14 Col. Cuauhtémoc México 5, D. F. Tel: 5-53-71-33 Ext. 2053
29. ING. CARLOS TERCERO BONIFAZ Chiclayo No. 874 Col. Lindavista México 14, D. F. Tel: 5-86-84-26	COMISION FEDERAL DE ELECTRICIDAD Ródano No. 14-6o. Piso Col. Cuauhtémoc México 12, D. F. Tel· 5-53-71-33 Ext. 2665
30. ING. FERNANDO TREVIÑO SOJO Arcos de Alba Pirules Oriente No. 37 Cuautitlán Izcalli	DIRECCION GENERAL DE PLANEACION EDUCATIVA Argentina No. 28-1er. Piso México, D. F. Tel: 5-12-15-27
31. ARMANDO VALENZUELA NOVELO Benito Juárez No. 101 Fraccionamiento Los Robles Col. Espartaca México 22, D. F. Tel: 5-39-62-30	COMISION FEDERAL DE ELECTRICIDAD Río Ródano No. 14-6o. Piso Col. Cuauhtémoc México, D. F. Tel: 5-53-71-33
32. LIC. SILVIA VAZQUEZ R. Prol. Ayuntamiento No. 136 Col. Romero de Terreros México 21, D. F. Tel: 5-54-00-51	BANCO DE MEXICO, S. A. Condesa No. 6-6o. Piso México 1, D. F. Tel: 5-85-42-99 Ext. 181

DIRECTORIO DE ASISTENTES AL CURSO DE PROBABILIDAD Y ESTAD
FUNDAMENTOS Y APLICACIONES (DEL 17 DE FEBRERO AL 23 DE M
1976)

NOMBRE Y DIRECCION	EMPRESA Y DIRECCION
33. ING. ALFONSO VILLEGAS JASSO Ret. 2 No. 9 Fray Servando T. M. Col. Jardín Balbuena México 9, D. F. Tel 5-52-92-61	SECRETARIA DE RECURSO COS Viena No. 20-2o. Piso Col. Juárez México 2, D. F. Tel: 5-92-35-68
34. BENITO ZYCHLINSKA ZYCHLINSKA Jojutla No. 65 Col. Condesa México 11, D. F. Tel: 5-53-04-98	LABORATORIOS LIOMONT, S. A. José Ma. Olloqui No. 28 Col. del Valle México 12, D. F. Tel: 5-34-60-85

DIRECTORIO DE PROFESORES

PROBABILIDAD Y ESTADISTICA: FUNDAMENTOS Y APLICACIONES

1. DR. OCTAVIO A. RASCON CHAVEZ
   Jefe de la División Estudios Superiores
   Facultad de Ingeniería
   Investigador del Instituto de Ingeniería
   U. N. A. M.
   Ciudad Universitaria
   México 20, D. F.
   548-09-50

2. M. en I. AUGUSTO VILLAREAL ARANDA
   Secretario Académico
   División de Estudios Superiores
   Facultad de Ingeniería
   Investigador del Instituto de Ingeniería
   U. N. A. M.
   México 20, D. F.
   548-09-50