Indice

1. PROBABILIDAD[1]

Objetivo:

Al finalizar el tema, el alumno será capaz de aplicar la teoría de probabilidad y los conceptos relacionados con ella en la solución de problemas que tengan relación con su carrera.

Introducción:

En un principio podemos decir que la probabilidad es el estudio de fenómenos puramente aleatorios.

La teoría de la probabilidad[2] es una parte de las matemáticas que se utiliza para descubrir e investigar las **características regulares** de los **eventos aleatorios**. Aunque en realidad no es posible dar una definición precisa y simple de los que significan las palabras **aleatorio** y **regular**, se espera que la explicación y los ejemplos que se dan a continuación faciliten la comprensión de estos conceptos.

Ciertos fenómenos del mundo real pueden considerarse como **fenómenos de azar**. Estos fenómenos no siempre producen el mismo resultado observado y el resultado de cualquier observación dada no se puede predecir. Estos fenómenos tienen un **comportamiento "fluctuante"** conocido como **regularidad estadística**.

En algunos casos, se conoce bastante bien el fenómeno que se estudia como para estar convencido de hacer las **predicciones exactas** con respecto al resultado de cada observación individual. Por ejemplo, si se desea saber la hora y el lugar de un eclipse solar, no se dudará en predecir una fecha exacta si se toman como base datos astronómicos.

Sin embargo, en muchas ocasiones el conocimiento disponible no es lo bastante preciso para que se puedan elaborar predicciones exactas en situaciones particulares.

Algunos ejemplos de estos casos, conocidos como **eventos aleatorios**, son[3]:

---

[1] A Girolamo cardano (1501-1576), físico, astrólogo y matemático, se le atribuye la primera discusión sobre probabilidad en su manual para jugadores, "**Liber De Ludo Aleae**" (manual para tirar los dados). Su trabajo más notable es "**Ars Magna**", en el cual se presentan raíces negativas de una ecuación y algunos calculus con números imaginarios. (nota tomada del libro: "Probabilidad y estadística". Stephen S. Willoughby. Publicaciones cultural S.A. p.p: prologo)

[2] Mizrahi y Sullivan. "Matemáticas finitas con aplicaciones a la Administración y Economía". Editorial: Limusa wiley. 2ª. Edición. p.p 332-333.

[3] Mizrahi y Sullivan. "Matemáticas finitas con aplicaciones a la Administración y Economía". Editorial: Limusa wiley. 2ª. Edición. p.p 332-333.

a) al lanzar una moneda se obtiene sol o águila. Para cualquier lanzamiento, no se puede predecir el resultado, aunque es obvio que está determinado por causas definidas (como la velocidad inicial de la moneda, el ángulo inicial de lanzamiento y la superficie sobre la que cae la moneda). Aunque se puedan controlar algunas de estas causas, no se puede predeterminar el resultado de un lanzamiento particular. Así, el resultado de lanzar una moneda es un evento aleatorio.

b) En una serie de lanzamientos de un dado común, cada vez que éste se lance da como resultado uno de los números 1, 2, 3, 4, 5, o 6. por consiguiente el resultado de lanzar un dado es un evento aleatorio.

c) Un recién nacido es varón o hembra. Sin embargo, en ningún caso particular se puede predecir el sexo de un recién nacido. Éste es también un ejemplo de evento aleatorio.

Los ejemplos anteriores demuestran que al estudiar una secuencia de experimentos aleatorios, no es posible predecir resultados particulares. Éstos están sujetos a fluctuaciones irregulares, aleatorias, que no se pueden predecir con exactitud. Sin embargo, si el número de observaciones es grande, es decir, si se trata de un **fenómeno en masa**, aparece alguna **regularidad**.

En el ejemplo (a) no se puede predecir el resultado de un lanzamiento cualquiera. No obstante, si se efectúa una larga serie de lanzamientos, se notará que el número de veces que cae sol es aproximadamente igual al número de veces que cae águila. O sea, parece **razonable** decir que en cualquier lanzamiento de la moneda, es **igualmente probable** que resulte un solo un águila. Como resultado, se puede asignar una probabilidad de 1/2 a la obtención de un sol (o águila) en un lanzamiento particular.

Para el ejemplo (b), la aparición de cualquier cara particular del dado es un evento aleatorio. Sin embargo, si se realizan varios lanzamientos, cualesquiera de las caras tiene **igual probabilidad** de aparecer. En este caso, se puede **asignar una probabilidad** de 1/6 a la obtención de una cara particular.

En el ejemplo (c), la intuición indica que es **igualmente probable** que nazca niño o niña. Si se sigue este razonamiento, se puede **asignar una probabilidad** *de* 1/2 al suceso de tener una niña. Sin embargo, si se consultan los datos de la tabla siguiente[4]:

[4] Mizrahi y Sullivan. "Matemáticas finitas con aplicaciones a la Administración y Economía". Editorial: Limusa wiley. 2ª. Edición. p.p 332-333.

Fecha de nacimiento	Número de nacimientos		Número total de nacimientos	Proporción de nacimientos	
	Niños **b**	Niñas **g**	**b+g**	$\frac{b}{b+g}$	$\frac{g}{b+g}$
1970	1 915378	1 816008	3731386	.513	.487
1971	1 822910	1 733 060	3 555 970	.513	.487
1972	1 669927	1 588484	3258411	.512	.488
1973	1 608 326	1528639	3 136 965	.513	.487
1974	1 622114	1 537 844	3 159 958	.513	.487
1975	1 613 135	1531063	3144198	.513	.487
1976	1 624 436	1 543 352	3 167 788	.513	.487
1977	1 705 916	1 620 716	3 326 632	.513	.487
1978	1 709 394	1 623885	3 333 279	.513	.487
1979	1 791 267	1 703 131	3 494 398	.513	.487
Total	17082803	16 226 182	33308985	.513	.487

se observa que sería más preciso **asignar una probabilidad** de .487 al hecho de tener una niña.

La palabra probabilidad[5] se usa para indicar la posibilidad de que ocurra un evento o resultado. La definición clásica de probabilidad y, en cierto modo, la más simple, se usa cuando un experimento puede tener solamente ciertos resultados definidos, cada uno de los cuales es igualmente probable: si hay ***N*** elementos en el conjunto de resultados posibles, la probabilidad para cualquiera de ellos es: ***1/N.***

1.1. Probabilidad subjetiva[6].

La probabilidad subjetiva es una cuestión de opinión y se basa en el hecho de que un experimento se realiza una sola vez. Dos personas por ejemplo pueden asignar diferentes probabilidades a un mismo evento, aun cuando tengan la misma información. Tal diversidad de opiniones se puede ver en las proyecciones económicas que hacen los asesores en inversiones y los economistas para los

[5] "Probabilidad y estadística". Stephen S. Willoughby. Publicaciones cultural S.A. p.p 1 y 3
[6] Estadística aplicada a la administración y a la economía. David K. Hildebrand y R. Lyman Ott. Editorial Addison Wesley Longman. P p 75

años venideros. Aunque muchos de estos individuos trabajan con los mismos datos, ellos se forman distintas opiniones acerca de las condiciones económicas más probables. Tales proyecciones son inherentemente subjetivas.

Las leyes matemáticas de la probabilidad pueden ayudarnos a hacer estimaciones lógicamente consistentes de la misma, pero no pueden garantizar que esas estimaciones sean correctas. Así, un buen gerente correrá menos riesgos que uno malo y deberá mostrar un mejor rendimiento en su trabajo en el largo plazo. No obstante no hay ninguna manera de garantizar que la evaluación de una probabilidad subjetiva sea correcta y otra incorrecta. Esto como dicen, ¡es lo que da lugar a las carreras de caballos!.

1.2. Probabilidad como frecuencial[7]:

La interpretación de la probabilidad como frecuencia relativa en el límite se basa en la observación de un gran número de ensayos. Donde con un número finito de ensayos es posible aproximar la verdadera probabilidad de un evento y valorar la aproximación. Muy a menudo se recurre a la interpretación de la frecuencia relativa, pues se utiliza siempre que sea razonable imaginar un gran número de ensayos repetidos de un experimento.

Espacios muestrales y asignación de probabilidades[8].

Al estudiar probabilidad interesan los experimentos, reales o conceptuales, y sus resultados. En este estudio se trata de formular con precisión una teoría matemática que produzca el experimento en cuestión. La primera etapa en la formulación de una teoría matemática es la elaboración de lo que se llama un *modelo matemático.* Luego, dicho modelo se emplea para predecir los resultados del experimento. El propósito de esta sección es aprender a diseñar un **modelo probabilístico.**

espacios muestrales

Se debe empezar con la descripción del **espacio muestral** del experimento; o sea, anotar todos los posibles resultados que pueden ocurrir al realizar el experimento.

Por ejemplo, si el experimento consiste en lanzar una moneda, generalmente se estaría de acuerdo en que los únicos resultados posibles son sol, **S***,* y águila, **A***.* Por consiguiente, el espacio muestral del experimento es el conjunto **{S, A}.**

Ejemplo 1[9]

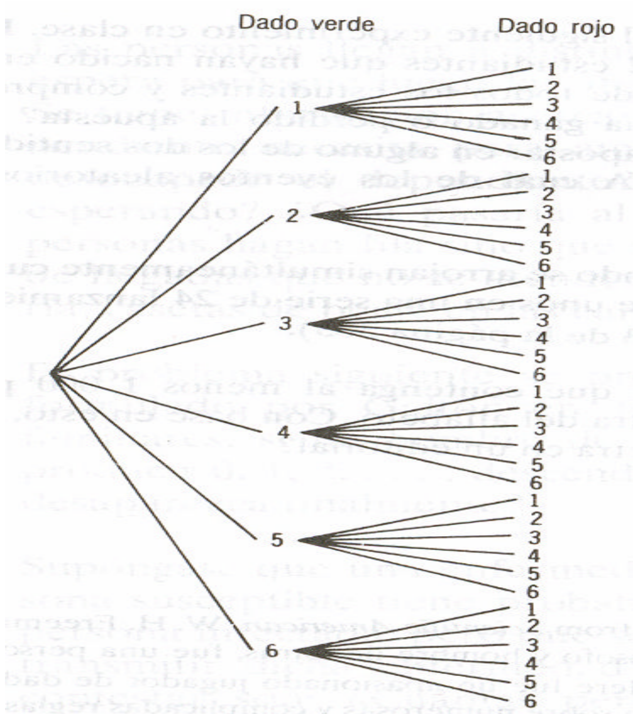
Considérese un experimento en el que, por ejemplo, un dado es verde y el otro rojo. Cuando se lanzan los dos dados, el conjunto de resultados consiste en todas las formas diferentes en que pueden caer. Esto se conoce como un conjunto de todas las **posibilidades lógicas.** Este

[7] Estadística aplicada a la administración y a la economía. David K. Hildebrand y R. Lyman Ott. Editorial Addison Wesley Longman. P. p 75

[8] Mizrahi y Sullivan. “Matemáticas finitas con aplicaciones a la Administración y Economía”. Editorial: Limusa wiley. 2ª. Edición. p.p 336-337.

[9] Mizrahi y Sullivan. “Matemáticas finitas con aplicaciones a la Administración y Economía”. Editorial: Limusa wiley. 2ª. Edición. p.p 336

experimento se puede describir en dos formas diferentes. Una forma consiste en emplear un diagrama de árbol, como se indica en la figura Siguiente:

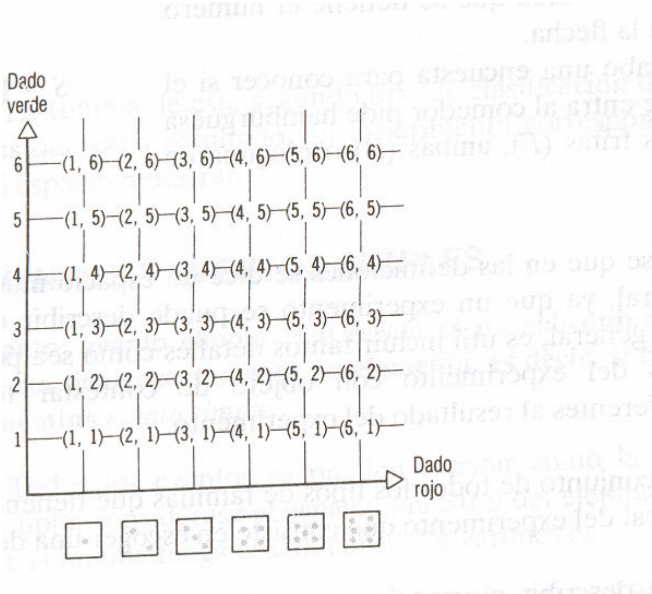


Otra forma consiste en denotar con **V** y con **R** los números que aparecen en los dados verde y rojo, respectivamente. Entonces, un **resultado** puede representarse con un par ordenado (**V**, **R**), donde **V** *y* **R** pueden tomar cualquiera de los valores 1, 2, 3, 4, 5, 6. Por lo tanto, el espacio muestral **S** para este experimento es el conjunto:

$$\mathbf{S = \{(V, R)/1 \leq V \leq 6, \quad 1 \leq R \leq 6\}}$$

Obsérvese también que el principio de multiplicación indica que el número de elementos en **S** es 36, ya que hay 6 posibilidades para ***V*** y 6 posibilidades para **R**, y que (6) (6) = 36.

La figura siguiente es la representación gráfica de los elementos de **S**[10].



Espacio muestral. Un *espacio muestral* S asociado a un experimento real o conceptual es el conjunto de todas las posibilidades lógicas que puedan ocurrir como resultado del experimento. Cada elemento de un espacio muestral S se llama resultado*.*

La teoría de la probabilidad se emplea tan pronto se especifica un espacio muestra!. El espacio muestral tiene la misma función que el conjunto universal en la teoría de los conjuntos, en lo que se refiere a todas las preguntas del experimento.

Para construir un modelo probabilístico se necesita lo siguiente[11]:

Paso 1. Hacer una lista de todos los resultados del experimento que se

investiga; o sea, dar el espacio muestral o, si esto no es fácil de realizar, determinar el número de eventos simples en el espacio muestral.

Paso 2. Asignar a cada evento simple una probabilidad $P_{(e)}$ de modo que se

satisfaga (1).

---

10 Mizrahi y Sullivan. “Matemáticas finitas con aplicaciones a la Administración y Economía”. Editorial: Limusa wiley. 2ª. Edición. p.p 337

11 Mizrahi y Sullivan. “Matemáticas finitas con aplicaciones a la Administración y Economía”. Editorial: Limusa wiley. 2ª. Edición. p.p 343

1.3. Espacio muestral[12]:

El espacio muestral "S" es el conjunto de todos los resultados posibles de un experimento.

Ejemplo[13]:

Suponga que se está llevando a cabo una auditoria de los servicios de tres nuevos distribuidores de automóviles seleccionados dentro de cierta área geográfica del país. Cada compañía auditada se señala con una "H" si todas las quejas sobre el servicio fueron resueltas en menos de dos meses, y si marcamos con una "T" en el caso contrario. Describa el espacio muestral obtenido de la forma: (resultado de la 1ª. compañía, resultado de la 2ª. Compañía, resultado de la 3ª. Compañía)

Solución:

S = {(HHH); (HHT); (HTH); (THH); (HTT); (THT); (TTH); (TTT)}

1.4. Eventos

Si un resultado es un elemento de un espacio muestral dado, entonces, Un Evento, es cualquier colección de resultados o, en lenguaje matemático, es un subconjunto del espacio muestral[14]. Definición: un **evento elemental** es un evento que contiene un solo elemento[15].

Ejemplo:

En el experimento lanzar un dado; el número 5 es un evento simple, y "cae" un número par" es un evento que contiene más de un elemento: (2, 4, 6)

1.5. Eventos mutuamente excluyentes, eventos independientes, regla de la multiplicación y regla de la adicion

Es importante hacer notar que antes de hablar de las leyes aditivas y de la multiplicación, debemos enunciar lo que son los:

Eventos mutuamente excluyentes[16]:

Se dice que dos eventos "**A**" y "**B**" son mutuamente excluyentes (disjuntos, lógicamente incompatibles) si no tienen resultados en común. Para los eventos mutuamente excluyentes, la intersección entre ellos no tiene resultados, es decir, es igual a cero; el que suceda uno de ellos significa automáticamente que el otro no puede ocurrir. Así, los eventos: **A**, **B**, **C**, **D,...** son

---

[12] Estadística aplicada a la administración y a la economía. David K. Hildebrand y R. Lyman Ott. Editorial Addison Wesley Longman. P. p. 84
[13] Estadística aplicada a la administración y a la economía. David K. Hildebrand y R. Lyman Ott. Editorial Addison Wesley Longman. P. p 85
[14] Estadística aplicada a la administración y a la economía. David K. Hildebrand y R. Lyman Ott. Editorial Addison Wesley Longman. P. p 84
[15] Probabilidad y estadística. Stephen S. Willoughby. Publicaciones cultural S.A. p.p 43
[16] Estadística aplicada a la administración y a la economía. David K. Hildebrand y R. Lyman Ott. Editorial Addison Wesley Longman. P. p 88

mutuamente excluyentes si todas las parejas posibles de ellos son mutuamente excluyentes. Lo cual significa que: si uno de tales eventos sucede, ninguno de los otros puede ocurrir.

Decisiones de inversión mutuamente excluyentes[17].

Si dos inversiones, X e Y son mutuamente excluyentes, el realizar una de ellas significa que no se puede llevar a cabo la otra. (este problema de decisión, se llega a presentar incluso si existe una sola TIR[18].)

Por ejemplo:

Si se posee un terreno en un esquina, se puede construir en él una estación de venta de gasolina o un edificio de apartamentos, pero no ambos. Estas son alternativas mutuamente excluyentes.

Ley aditiva para eventos mutuamente excluyentes[19]:

Esta ley dice que: Si dos eventos **A** y **B** son mutuamente excluyentes, entonces:

$$P_{(AoBsucede)} = P(A \cup B) = P(A) + P(B)$$

Siempre que los eventos se hayan definido como mutuamente excluyentes, el "o" lógico corresponde a la suma de probabilidades. Obviamente, esta idea no se restringe a dos eventos; es aplicable a cualquier número finito o infinito de ellos.

Ejemplo[20]:

En el experimento que consiste en lanzar dos dados no cargados, ¿cuál es la probabilidad de obtener una suma de 7 o una suma de 11 ?

Solución: Sean **E** Y **F** los eventos

E*:* la suma es 7 **F*:* la suma es 11**

Puesto que los dados no están cargados,

---

[17] Stephen A. Ross. Et al. "Fundamentos de Finanzas Corporativas". Editorial: Mc Graw Hill- Irwin. P.p. 240 (las decisions de inversión mutuamente excluyentes, son situaciones en las que el hecho de aceptar una inversión evita que se acepte otra).

[18] TIR significa: Tasa Interna de Rendimiento; que no es otra cosa que la tasa de descuento que hace que el Valor Presente Neto (VPN) de una inversión sea cero.

[19] Estadística aplicada a la administración y a la economía. David K. Hildebrand y R. Lyman Ott. Editorial Addison Wesley Longman. P. p 89

[20] Mizrahi y Sullivan. "Matemáticas finitas con aplicaciones a la Administración y Economía". Editorial: Limusa wiley. 2ª. Edición. p.p 349

$P(E) = 6/\underline{36}$

$P(F) = 2/\underline{36}$ '

Los dos eventos *E* y *F* se excluyen. Por lo tanto, de acuerdo a la fórmula para eventos mutuamente excluyentes la probabilidad de que la suma sea 7 u 11 es

$$P(E \cup F) = P(E) + P(F) = 6/\underline{36} + 2/\underline{36} = 8/\underline{36} = 2/9$$

Eventos independientes:

La probabilidad de dos o más eventos independientes que se presentan juntos o en sucesión es el producto de sus probabilidades marginales:

$$P(AB) = P(A) \times P(B)$$

Simbólicamente, la probabilidad condicional se escribe:

P(B/A)

Y se lee "la probabilidad de que se presente el evento B, dado que el evento A se ha presentado".

La probabilidad condicional es la probabilidad de que un segundo evento (B) se presente, si un primer evento (A) ya ha sucedido.

Para eventos estadísticamente independientes, la probabilidad condicional de que suceda el evento B dado que el evento A se ha presentado, es simplemente la probabilidad del evento B:

$$P(B/A) = P(B)$$

Ley aditiva general[21]:

Esta ley nos dice que: Si **A** y **B** son eventos cualesquiera, entonces:

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

Ejemplo[22]:

[21] Estadística aplicada a la administración y a la economía. David K. Hildebrand y R. Lyman Ott. Editorial Addison Wesley Longman. P. p 90

[22] Estadística aplicada a la administración y a la economía. David K. Hildebrand y R. Lyman Ott. Editorial Addison Wesley Longman. P. p 77-78

Un distribuidor minorista acepta pedidos de tres maneras distintas: por teléfono, a través de una forma que va adjunta a su catálogo o repitiendo las órdenes de compra de sus clientes. Las órdenes de compra están clasificadas como:

1. pequeñas (menos de $25.00)
2. medianas (de $25.00 a $99.00)
3. grande (de $100.00 a $299.99) y
4. mayores (de $300.00 o más)

En la siguiente tabla se muestra un análisis de las últimas 4000 órdenes de compra que recibió el distribuidor:

Tamaño

	Pequeña	Mediana	Grande	Mayor	Total
Catálogo	1021	216	109	14	1360
Repetida	86	371	308	49	814
Telefónica	1497	230	86	13	1826
total	2604	817	503	76	4000

a) Si los pedidos procedentes de las órdenes repetidas y del catálogo pasan por un proceso de entrada. ¿Cuál es la probabilidad de que un dato seleccionado al azar pertenezca a esta categoría?

b) Si las órdenes mayores y las órdenes telefónicas se retienen mientras se verifica el crédito. ¿Cuál es la probabilidad de que una orden seleccionada al azar haya sido retenida?

Solución:

a) para este caso tenemos que:

P(proceso de entrada) = P(de catálogo o repetida)

= P(de catálogo) + P(repetida)

$$= \frac{1360}{4000} + \frac{814}{4000}$$

= 0.3400 +0.2035

= 0.5435

por lo tanto, la probabilidad de que un dato seleccionado al azar pertenezca a la categoría: órdenes repetidas o del catálogo es de 0.5435.

no necesitamos preocuparnos por el doble conteo, pues "orden por catálogo" y "repetición" son categorías mutuamente excluyentes.

b) hay órdenes telefónicas que son mayores. Debemos utilizar el principio aditivo general. Por lo tanto:

P(retenida) = P(mayor o telefónica)

= P(mayor) + P(telefónica) – P(mayor y telefónica)

$$= \frac{76}{4000} + \frac{1826}{4000}$$

$$= \frac{13}{4000}$$

= 0.01900 + 0.45650 – 0.00325

= 0.47225

por lo tanto la probabilidad de que la orden sea retenida es de: 0.47225

Ejemplo 2[23]

Considerar dos eventos:

***E:* un comprador gasta al menos 40 dólares en alimentos**

***F:* un comprador gasta al menos 15 dólares en carne**

Como resultado de estudios recientes, se pueden dar las probabilidades siguientes:

$P(E) = .56$

$P(F) = .63$

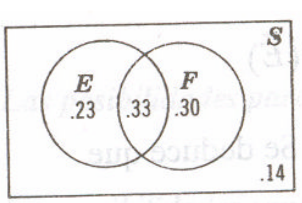
Supóngase que la probabilidad de que un comprador gaste al menos 40 dólares en alimentos y 15 en carne es .33. ¿Cuál es la probabilidad de que un comprador gaste al menos 40 dólares en alimentos o al menos 15 en carne?

Solución:

Puesto que se busca la probabilidad de $E \cup F$, se emplea la ley aditiva general y lo que se obtiene es

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$
$$= .56 + .63 - .33$$
$$= .86$$

En la solución de problemas de probabilidad, muchas veces resulta útil un diagrama de Venn. El diagrama de Venn de la figura siguiente representa la información del ejemplo 2. De esta figura se concluye que la probabilidad del evento *E*, pero no la de *F*, es 0.23 y que .14 no es la probabilidad de *E* ni de *F.*



Ejemplo 3[24]

[23] Mizrahi y Sullivan. "Matemáticas finitas con aplicaciones a la Administración y Economía". Editorial: Limusa wiley. 2ª. Edición. p. 351

En un experimento con dos dados no cargados considerar los eventos:

***E:* la suma de las caras es 8**

***F:* cae lo mismo en los dos dados**

¿Cuál es la probabilidad de obtener *E* ó *F?*

Solución:

$E = \{(2, 6), (3, 5), (4,4), (5, 3), (6, 2)\}$

$F = \{(1, 1), (2, 2), (3,3), (4, 4), (5,5), (6, 6)\}$

$E \cap F = \{(4,4)\}$

También

$P(E) = 5/36$

$P(F) = 6/36$

$P(E \cap F) = 1/36$

Por lo tanto:

$$
\begin{aligned}
P(EUF) &= 5/36 + 6/36 - 1/36 \\
&= 10/36 \\
&= 5/18
\end{aligned}
$$

Ejemplo 4[25]

Un estudio reveló que es razonable atribuir a personas mayores de 40 años que tienen grado de Maestro en Administración de Empresas, una probabilidad de 0.756, de que obtendrán ingresos anuales superiores a 30 000 dólares. Entonces, la probabilidad de que una persona de éstas gane 30 000 dólares o menos es

---

[24] Mizrahi y Sullivan. "Matemáticas finitas con aplicaciones a la Administración y Economía". Editorial: Limusa wiley. 2ª. Edición. p.p 351-352

[25] Mizrahi y Sullivan. "Matemáticas finitas con aplicaciones a la Administración y Economía". Editorial: Limusa wiley. 2ª. Edición. p.p 352

$$1 - .756 = .244$$

Ejemplo 5[26]

Para un experimento con dos dados no cargados, calcular:

(a) La probabilidad de que la suma de las caras sea menor o igual que 7

(b) La probabilidad de que la suma de las caras sea mayor que 7

Solución

(a) El número de eventos simples en el evento **E** es 21. El número de eventos simples en el espacio muestral **S** es 36. Por lo tanto, debido a que los dados no están cargados:

$$P(E) = 21/36$$
$$= 7/12$$

(b) Se necesita calcular la probabilidad del complemento del evento **E**; por lo tanto:

$$P(E) = 1 - P(E)$$
$$= 1 - 7/12$$
$$= 5/12$$

Es decir, la probabilidad de que la suma de las caras sea superior a 7 es 5/12

Ejemplo 6[27]

En un experimento con 2 dados, se presentan los siguientes eventos:

---

26 Mizrahi y Sullivan. "Matemáticas finitas con aplicaciones a la Administración y Economía". Editorial: Limusa wiley. 2ª. Edición. p.p 352-353

27 Mizrahi y Sullivan. "Matemáticas finitas con aplicaciones a la Administración y Economía". Editorial: Limusa wiley. 2ª. Edición. p.p 360

(a) La suma de las caras es 3.
(b) La suma de las caras es 7.
(c) La suma de las caras es 7 o 3.
(d) La suma de las caras es 7 y 3.

Calcular la probabilidad de estos eventos para dados no cargados.

Solución

(a) La suma de las caras es 3 si, y sólo si el resultado es el evento
A = {(1, 2), (2,l)}. Por lo tanto:

$$p(A) = 2/36$$
$$= 1/18$$

(b) La suma de las caras es 7 si y sólo si el resultado es un elemento del evento
B = {1, 6), (2,5), (3, 4), (4,3), (5, 2), (6, 1)}. por lo tanto

$$P(B) = 6/36$$
$$= 1/6$$

(c) La suma de las caras es 7 ó 3 si y sólo si el resultado es un elemento de *A* U *B,* como se definieron en (a) y (b).

A U *B* = {(2, 1), (1, 2), (1,6), (2, 5), (3,4), (4, 3), (5,2), (6, 1)}

Por lo tanto:

$$P(A \cup B) = 8/36$$
$$= 2/9$$

(d) La suma de las caras es 7 y 3 si y sólo si el resultado es un elemento de *A* n *B.* Dado que *A* n *B* = 0, el evento es imposible. Es decir, *P(A* n *B)* = 0.

Ejemplo 7[28]

¿Cuál es la probabilidad de que un número de 4 dígitos de una extensión telefónica tenga uno o más dígitos repetidos?

Solución:

Hay 10^4 = 10 000 números de 4 dígitos distintos para las extensiones telefónicas. Por lo tanto, éste es el número de eventos simples en el espacio muestra!.

Se desea calcular la probabilidad de que un número de 4 dígitos de extensión telefónica escogido al azar tenga uno o más dígitos repetidos. Dado que es difícil contar los elementos que hay en este evento, se calcula primero la probabilidad del evento.

***E:* ningún dígito repetido en un número de 4 dígitos**

Obsérvese que el evento $\overline{E}$ consiste en números con uno o más dígitos repetidos. Por el principio de multiplicación, el número de extensiones de 4 dígitos con **ninguno** repetido es

$$(10)\ (9)\ (8)\ (7) = 5\ 040$$

De aquí,

$$P(E) = 5040/10000$$
$$= 0.504$$

Por lo tanto, la probabilidad de tener uno o más dígitos repetidos es:

$$P(\overline{E}) = 1 - 0.504 = 0.496$$

[28] Mizrahi y Sullivan. "Matemáticas finitas con aplicaciones a la Administración y Economía". Editorial: Limusa wiley. 2ª. Edición. p.p 360-361

Ejemplo 8[29]

Una caja contiene 12 bombillas eléctricas, de las cuales 5 están defectuosas. Todas las bombillas se ven iguales y tienen igual probabilidad de ser escogidas. Se escogen tres bombillas al azar. ¿Cuál es la probabilidad de que las 3 estén defectuosas?

Solución

El número de elementos en el espacio muestral **S** es igual al número de combinaciones de 12 bombillas tomadas de 3 en 3, o sea

$$C_3^{12} = \frac{12!}{3!9!}$$

$$= 220$$

Se define el evento **E** como "3 bombillas están defectuosas". Entonces **E** puede ocurrir en $_5C_3$ formas, o sea, el número de formas en que se pueden escoger 3 bombillas defectuosas de un conjunto de 5 que estén defectuosas. La probabilidad *P(E)* es

$$P(E) = \frac{C_3^5}{C_3^{12}}$$

$$P(E) = \frac{\frac{5!}{3!2!}}{220}$$

$$P(E) = 0.04545$$

Ley Multiplicativa[30]:

Esta ley enuncia que: Si *A* y *B* son dos eventos cualesquiera, entonces:

$$P_{(AyB)} = P(A \cap B) = P(A)P(B/A)$$

---

[29] Mizrahi y Sullivan. "Matemáticas finitas con aplicaciones a la Administración y Economía". Editorial: Limusa wiley. 2ª. Edición. p.p 362-363

[30] Estadística aplicada a la administración y a la economía. David K. Hildebrand y R. Lyman Ott. Editorial Addison Wesley Longman. p p 94

la única diferencia entre la ley multiplicativa y la definición de probabilidad condicionada radica en cuáles probabilidades se suponen y cuáles han de ser calculadas. Así, cuando suponemos que se conocen las llamadas **Probabilidad conjunta**: $P_{(A \cap B)}$ y la **probabilidad marginal:** $P_{(A)}$, la probabilidad condicionada : $P_{(B/A)}$ se puede calcular por medio de la definición. Y cuando se conocen: $P_{(A)}$ y $P_{(B/A)}$ entonces podemos calcular la probabilidad de su intersección con la ley multiplicativa: $P_{(A \cap B)}$.

Ejemplo[31]:

Se va a seleccionar un equipo de evaluación de dos personas a partir de un grupo formado por 10 hombres y 6 mujeres. Si cada grupo de dos personas tiene la misma probabilidad de ser seleccionado, encuentre la probabilidad de que el equipo de evaluación esté integrado por dos mujeres.

Solución:

Sea el evento "A" "la primera persona seleccionada es una mujer" y sea el evento "B" "la segunda persona seleccionada es mujer". Lo que queremos calcular entonces es $P_{(A \cap B)}$ lo cual hacemos claro esta por medio de la ley multiplicativa como sigue:

$$P_{(A \cap B)} = P_{(A)} P_{(B/A)}$$

de donde sustituyendo datos obtenemos fácilmente que:

$$P_{(A \cap B)} = \left[\frac{6}{16}\right]\left[\frac{5}{15}\right] = 0.125$$

la probabilidad de que el equipo de evaluación esté integrado por dos mujeres es de: 0.125

31 Estadística aplicada a la administración y a la economía. David K. Hildebrand y R. Lyman Ott. Editorial Addison Wesley Longman. p 94

1.6. Tablas de probabilidad conjunta

Probabilidad conjunta es una probabilidad que mide la posibilidad de que dos o más eventos ocurran juntos.

1.7. Probabilidad marginal y condicional

Probabilidad marginal
Es la probabilidad particular de una de las variables dada un variable aleatoria bidimensional, y se define como:

$$P(X) = \sum_y P(x_i, y_j)$$

Probabilidad condicional
Dada la probabilidad conjunta y marginal, la probabilidad condicional se define como:

$$P(X \mid Y) = P(X, Y)/P(Y)$$

Ejemplo[32]:

En un despacho contable hay 216 cuentas, 80 de las cuales son grandes y 16 de las cuales son grandes y tienen errores. Si elegimos una cuenta al azar. ¿Cuál es la probabilidad de que la cuenta seleccionada tenga un error, si se trata de una cuenta grande?.

Solución: la fórmula a utilizar es la siguiente:

$$P_{(error/\,grande)} = \frac{P_{(error \cap grande)}}{P_{(grande)}}$$

Es decir, se trata de un problema de probabilidad condicional. Por lo tanto, al sustituir datos en la fórmula obtenemos que:

$$P_{(error/\,grande)} = \frac{\frac{16}{216}}{\frac{80}{216}} = \frac{16}{80} = 0.2$$

[32] Estadística aplicada a la administración y a la economía. David K. Hildebrand y R. Lyman Ott. Editorial Addison Wesley Longman. P. p 93

es decir, la probabilidad de que una cuenta seleccionada al azar tenga un error dado que sabemos que se trata de una cuenta grande es de: 0.2

Ejemplo 2[33]

Suponer que una población de 1 000 personas incluye a 70 contadores y 520 mujeres. Sea **E** el evento "una persona es contador". Sea **F** el evento "una persona es mujer". Entonces:

$$P(E) = \frac{70}{1000} \qquad\qquad P(F) = \frac{520}{1000}$$

$$P(E) = 0.07 \qquad\qquad P(F) = 0.52$$

En lugar de estudiar toda la población, se desea investigar a la población femenina y calcular la probabilidad de que una mujer elegida al azar sea contador. Si hay 40 mujeres contadores, la razón $\frac{40}{520}$ representa la probabilidad condicional del evento **E** (contador) siempre que ya haya ocurrido el evento **F** (la persona elegida es mujer). En símbolos, se escribiría:

$$P(E/F) = \frac{P_{(E \cap F)}}{P_{(F)}}$$

De donde al sustituir valores tenemos que:

$$P(E/F) = \frac{\frac{40}{1000}}{\frac{520}{1000}}$$

$$P(E/F) = \frac{(40)(1000)}{(520)(1000)}$$

33 Mizrahi y Sullivan. "Matemáticas finitas con aplicaciones a la Administración y Economía". Editorial: Limusa wiley. 2ª. Edición. p.p 369

$$P(E/F)=\frac{40}{520}$$

De donde finalmente:

$$P(E/F)=\frac{1}{13}$$

1.8. Independencia estadística.

La probabilidad de dos o más eventos independientes que se presentan juntos o en sucesión es el producto de sus probabilidades marginales:

P (AB) = P(A) X P(B)

Simbólicamente, la probabilidad condicional se escribe:

P(B/A)

Y se lee "la probabilidad de que se presente el evento B, dado que el evento A se ha presentado".

La probabilidad condicional es la probabilidad de que un segundo evento (B) se presente, si un primer evento (A) ya ha sucedido.

Para eventos estadísticamente independientes, la probabilidad condicional de que suceda el evento B dado que el evento A se ha presentado, es simplemente la probabilidad del evento B:

P(B/A) = P(B)

1.9. Teorema de Bayes:

Si $\mathbf{A_1,\dots,A_k}$ son estados naturales mutuamente excluyentes y si $\mathbf{B_1,\dots,B_m}$ son "***m***" eventos observables posiblemente excluyentes entre sí, entonces:

$$P_{(A_i/B_j)}=\frac{P_{(A_i)}P_{(B_j/A_i)}}{P_{(A_1)}P_{(B_j/A_1)}+P_{(A_2)}P_{(B_j/A_2)}+\ldots+P_{(A_k)}P_{(B_j/A_k)}}$$

$$P_{(A_i/B_j)}=\frac{P_{(A_i)}P_{(B_j/A_i)}}{\sum_i P_{(A_i)}P_{(B_j/A_i)}}$$

Una probabilidad a priori son aquellas que tenemos antes de obtener nueva información; y las probabilidades a son aquellas que llegamos a tener después de obtenida alguna otra información.

Ejemplo[34]:

Los anuncios para la televisión varían en su efectividad. Una agencia de publicidad produjo un anuncio para TV de un producto ya conocido (neumáticos radiales para automóvil). El gerente de marca estima subjetivamente que el anuncio tiene un 20% de posibilidades de ser efectivo (la participación en el mercado aumentará después de su exhibición), un 70% de posibilidades de ser adecuado (la participación en el mercado no cambiará) y un 10% de posibilidades de ser desastroso (la participación en el mercado se reducirá). El anuncio se puede poner a prueba con un grupo de consumidores. Experiencias anteriores con tales grupos indican que son moderadamente fiables para predecir la efectividad. El director de la marca estima la verosimilitud de las reacciones positivas, neutrales y negativas del grupo (dado el resultado eventual) como sigue:

Reacción del grupo

Resultado Del anuncio	*Positiva*	*Neutral*	*negativa*
Efectivo	0.60	0.30	0.10
Adecuado	0.40	0.30	0.30
Desastroso	0.10	0.30	0.60

¿Cómo debería cambiar la probabilidad de que el anuncio será efectivo con una reacción neutral del grupo?.

Solución:

El evento observable "reacción neutral" es estadísticamente independiente de los estados naturales: "resultado del anuncio". La probabilidad (condicionada) de una reacción neutral es la misma para todos los resultados. Por consiguiente, la probabilidad de un

[34] Estadística aplicada a la administración y a la economía. David K. Hildebrand y R. Lyman Ott. Editorial Addison Wesley Longman. p 113-114

anuncio publicitario efectivo no debería cambiar con una reacción neutral. Utilizando el teorema de Bayes tenemos que:

$$P_{(efectivo\,/\,neutral\,)} = \frac{P_{(neutral\,/\,efectivo\,)}P_{(efectivc)}}{P_{(neutral\,/\,efectivo)}P_{(efectivo)} + P_{(neutral\,/\,adecuado\,)}P_{(adecuado)} + P_{(neutral\,/\,desastroso\,)}P_{(desastroso)}}$$

por lo tanto, al sustituir valores tenemos que:

$$P_{(efectivo\,/\,neutral\,)} = \frac{(0.20)(0.30)}{(0.20)(0.30) + (0.70)(0.30) + (0.10)(0.30)} = 0.20$$

es decir, la probabilidad es la misma que la probabilidad a priori, como debe de ser cuando el evento observable es estadísticamente independiente del verdadero estado natural.

Ejercicios propuestos:

De la población masculina[35] entre los 30 y los 35 años que viven en la ciudad de México, el 25% tiene un título universitario, el 15% gana más de $25,000.00 al año y el 65% no tiene grado y gana menos de $25,000.00 al año. ¿Cuál es la probabilidad de que una persona elegida al azar de este grupo gane más de $25,000.00 al año dado que posee un título?

De la población masculina[36] entre los 30 y los 35 años que viven en la ciudad de México, el 25% tiene un título universitario, el 15% gana más de $25,000.00 al año y el 65% no tiene grado y gana menos de $25,000.00 al año. ¿Cuál es la probabilidad de que una persona elegida al azar de este grupo gane más de $25,000.00 al año dado que no tiene un título?

[35] Arya, Jagdisch C. y Lardner, Robin W. “Matemáticas aplicadas a la administración y a la economía” Editorial: Prentice-Hall. Tercera edición. p.p 313.

[36] Arya, Jagdisch C. y Lardner, Robin W. “Matemáticas aplicadas a la administración y a la economía” Editorial: Prentice-Hall. Tercera edición. p.p 313.

2. Estadística Descriptiva.

Objetivo:

Inducir al alumno a que tome gusto por la estadística al observar la aplicación profesional de ésta a diferentes áreas y ciencias tales como:

- Administración
- Contabilidad
- Mercadotecnia
- Sociología
- Control de calidad
- Ingeniería
- Finanzas
- Recursos Humanos
- Almacén (de una empresa)
- En la vida diaria, etcétera.

El crecimiento de la población[37] y con ella el surgimiento de nuevos problemas que resolver hicieron posible la ampliación de las aplicaciones de la matemática de las ciencias físicas a otras como: las ciencias del comportamiento, las ciencias biológicas y las ciencias sociales entre otras.

Teniendo esto en consideración no es de extrañar que los usos y aplicaciones de la Probabilidad y la Estadística pasaran de los juegos de azar y los fines militares y posteriormente recaudatorios a las "entrañas" de las organizaciones convirtiéndose en un bastión de las empresas exitosas.

Históricamente, el crecimiento y desarrollo de la estadística moderna puede trazarse desde dos fenómenos separados:

- La necesidad del gobierno de recabar datos sobre sus ciudadanos y
- El desarrollo en las matemáticas, de la teoría de probabilidades.

[37] Indagar de donde se obtuvo la información de estos dos párrafos.

Así por ejemplo, durante las civilizaciones egipcia, griega y romana, los datos se obtenían principalmente con propósitos de impuestos y reclutamiento militar.

En la edad media, las instituciones eclesiásticas a menudo mantenían registros de nacimientos, muertes y matrimonios. En nuestro país, organismos tales como el INEGI realizan levantamientos de censos.

por otra parte, la mayoría de los autores coinciden en que la estadística proporciona los elementos básicos para fundamentar una investigación, como son:

1. Cómo planear la obtención de los datos para que de ellos se puedan extraer conclusiones confiables.
2. Cómo analizar estos datos.
3. Qué tipo de conclusiones pueden obtenerse con los datos disponibles.
4. Cuál es la confianza que nos merecen los datos.

Como puede observarse, la estadística nos permite realizar estudios de tipo descriptivo y explicativo por medio de sus dos ramas, prácticamente en todas las áreas del conocimiento humano; Claro esta siempre y cuando apliquemos un método.

Elementos para realizar una investigación usando estadística.

1. **formulación del problema.** Consiste en la identificación y especificación adecuada de un problema de investigación. En esta etapa es muy importante establecer con precisión la o las hipótesis, el o los objetivos del estudio, su alcance y la población de datos asociada al mismo.

2. **diseño del experimento.** En esta segunda etapa el investigador deberá seleccionar la técnica de recolección de datos (observación directa, entrevistas, encuesta, investigación documental) que le permitan obtener información a un mínimo costo (dinero y tiempo) posible. En esta etapa el investigador deberá definir el tamaño de la muestra, la calidad requerida y el tipo de datos que le permitan resolver el problema planteado de la manera más eficiente.

3. **recolección de datos.** Es, tal vez la etapa de mayor importancia en la investigación, ya que la calidad de los datos obtenidos depende de una buena recolección, por lo tanto ésta deberá sujetarse a reglas estrictas de tal forma que permitan obtener la información deseada.

4. **proceso de datos y su descripción.** Consiste en la elaboración de cuadros estadísticos de trabajo, cuadros estadísticos de referencia, gráficas y cálculo de medidas estadísticas apropiadas al proceso inferencial seleccionado. Esto es, se exponen los datos muestrales mediante de representaciones tabulares, gráficas y medidas estadísticas con el objeto de hacer una descripción de los resultados.

5. **inferencia estadística y conclusiones.** Esta etapa proporciona una contribución muy importante, ya que en ella se define el nivel de confianza y significancia del proceso inferencial, lo cual sirve como orientación a quien o quienes deben tomar una decisión sobre el tema objeto de estudio. Esto último permite al investigador, establecer una conclusión sobre el problema y, en algunas ocasiones, elaborar sugerencias para la solución del mismo.

Los subtemas que forman este primer tema de INTRODUCCIÓN son:

Definición de estadística.

Cuyo objetivo es: Que el alumno se forme una idea o definición de lo es la Estadística después de haber analizado y cuestionado las diferentes definiciones que nos ofrecen diferentes autores en diferentes libros.

¿qué es la estadística?

De una manera muy sencilla, podemos decir, que:

La estadística es una ciencia relativamente nueva que tiene por objeto la colección e interpretación de datos. Y aunque sus orígenes y actividades se remontan a la época del viejo testamento, y a los registros que ya elaboraban los babilonios y los romanos, sobre la población, no fue, sino hasta mediados del siglo XVII cuando Adolph Quetelet, aplico por primera vez métodos modernos al estudio de un conjunto de datos.

Formalmente, la palabra Estadística, surge en el tiempo a partir de la interpretación de los vocablos:

Status. Palabra derivada del latín, que significa:
Situación, posición, estado
Statera. Palabra derivada del griego, que
Significa: balanza.
Staat. Palabra de origen alemán que se refiere al
Estado como expresión de unidad política Superior.

Hoy, esta ciencia moderna puede interpretarse como:

Un método que permite organizar, sintetizar, presentar, analizar, cuantificar e interpretar grandes volúmenes de datos, de tal manera que se puedan obtener conclusiones válidas (dar información), sobre los fenómenos de estudio.

Este amplio concepto ha llevado a los especialistas es estadística a llegar a un acuerdo que permite clasificar a esta materia en dos grandes ramas, **estadística descriptiva y estadística inferencial.** Donde ambas desempeñan funciones distintas pero complementarias en el análisis estadístico.

estadística descriptiva. Es aquella rama de la estadística que trata del resumen y descripción de los datos. Donde este resumen puede ser tabular, gráfico, o numérico. Su análisis y descripción se limita exclusivamente a los datos coleccionados, por lo que esta no puede inferir o generalizar acerca de la totalidad de donde provienen dichas observaciones (población).

estadística inferencial. Es la otra rama de la estadística que tiene como objetivo generalizar o inferir conclusiones útiles sobre la totalidad de las observaciones (población) a partir del análisis de los datos coleccionados (muestra). En otras palabras, la inferencia estadística constituye la base teórica del muestreo, es decir, permite conocer el todo con cierta aproximación a partir del estudio de una parte.

Usos y abusos de la Estadística.

Mostrar los usos correctos de la estadística y como ha sido mal utilizada por algunas personas.

La aparición de los juegos de azar dieron un gran impulso a la estadística inferencial, y de hecho hay quien los considera como precursores del objeto de la misma; así, en la correspondencia entre el matemático Pascal y el jugador Chevalier se encontró un gran ímpetu por la formulación de las matemáticas de la teoría de probabilidades (esto fue a mediados del siglo XVII), estos y otros desarrollos matemáticos como los de Bernoulli, DeMoivre y Gauss fueron los precursores de la Estadística Inferencial, pero no fue sino a principios del siglo XX cuando estadísticos como Pearson, Ficher, Gosset, Neyman, Wald y Tukey sentaron las bases del desarrollo de los métodos de la estadística inferencial que actualmente tiene tantas aplicaciones en otros tantos campos de la ciencia actual[38].

Poblaciones y muestras.

Dar la definiciones de Población y Muestra.

- **población.** Es el conjunto formado por un número determinado o indeterminado de unidades, (personas, objetos, etc.). que comparten características comunes a un objeto de estudio. Por ejemplo, en un estudio de las preferencias de los votantes en una elección presidencial, la población estaría formada por todas las personas registradas en un padrón electoral.

- **muestra.** Es cualquier subconjunto seleccionado de una población, siguiendo ciertos criterios establecidos en la teoría del muestreo. La muestra es el elemento básico sobre el cual se fundamenta la posterior inferencia acerca de la población de donde se ha tomado.

Datos, problemas de definición y medición.

[38] Tomado del libro: Estadística Basica en Administración del Autor: Levine, pag.3

En la definición de Estadística hemos utilizado una serie de conceptos que es conveniente definir para el correcto entendimiento de la estadística misma.

- **dato.** Es un número o medida que ha sido recopilada como resultado de una observación. Los datos pueden ser producto de un conteo, una medición o una denominación. Por ejemplo: el número de personas en una población, número de empresas en un país, el peso de una persona, el sexo de una persona, el nombre de una persona, etc.

- **variable.** Para obtener estadísticas manejamos conjuntos que poseen un determinado o indeterminado numero de unidades (personas, objetos, etc.). Las unidades de estudio tienen determinadas características, por ejemplo, para un ciudadano Mexicano, podríamos señalar: Sexo, edad, estatura, peso, lugar de nacimiento, estrato social, grado de escolaridad, religión, estado civil, etc. así pues, todas y cada una de estas características, que adquieren diferentes valores en cada persona, lugar o cosa, y que son susceptibles de una medición, reciben el nombre de Variables. De esta forma, el estudio de los habitantes de una población, requeriría tal vez del uso de variables como: sexo, edad, estatura, peso, estrato social, religión, etcétera. Como puede observarse, la variable es una construcción que el investigador genera para analizar una realidad.

- **parámetro.** Es la medida estadística que cuantifica una característica que ha sido estudiada para una población. Este valor estadístico se considera verdadero, ya que su origen parte del estudio de cada uno de los datos que constituyen a la población.

- **estadígrafo o estadística.** Es la medida que cuantifica una característica estudiada en una muestra. Por ejemplo, si se considera una muestra de 100 posibles clientes de un producto y se les entrevista para conocer las preferencias sobre el mismo, obteniéndose como resultado que 70 prefieren el producto, entonces la proporción muestral será de 0.70, constituyendo ésta una estadística.

2.1. Tabulación de datos

Una vez recogidos los datos, parece oportuno presentarlos de una forma coherente y ordenada que sea más manejable. Es decir, tabulemos.

La tabulación consiste simplemente en ordenar en una tabla los datos recogidos para ofrecer una visión conjunta.

La tabulación constará, en su forma más sencilla, de dos columnas: en la primera, se reflejarán los distintos valores observados de la variable, ordenados en sentido creciente o decreciente, y en la segunda, el número de veces que cada uno de ellos aparece (frecuencia). Por esta razón, las estadísticas así organizadas reciben el nombre de distribuciones de frecuencia.

Para el ejemplo anterior de ausencias a clase en una determinada semana, la tabla correspondiente será:

Ausencia (valores de variable)	**Número de alumnos (Frecuencia)**
0	11
1	4
2	2
3	2
4	1
5	1

La presentación de datos cualitativos suele hacerse de forma análoga a la de las variables, indicando las distintas clases o atributos observados y sus frecuencias de aparición, tal como se recoge en la tabla siguiente sobre color de pelo en un grupo de 100 turistas italianos:

Color de pelo	**Número de personas**
Negro	60
Rubio	25
Castaño	15

2.2. Distribución de frecuencias

Una distribución de frecuencia o tabla estadística es un arreglo sistemático de datos numéricos presentados en columnas e hileras para propósito de comparación.

Frecuencias absolutas y relativas[39]

La **frecuencia absoluta** no es otra cosa que el número que indica cuántas veces el valor correspondiente de una variable de medición (dato), se presenta en la muestra y también se le conoce simplemente como frecuencia de ese valor de "x" (dato) en la muestra.

Si ahora dividimos la frecuencia absoluta entre el tamaño de la muestra "n", obtenemos la **frecuencia relativa** correspondiente.

A manera de teorema podemos decir que: la frecuencia relativa es por lo menos igual a 0 y cuando más igual a 1. además, la suma de todas las frecuencias relativas en una muestra siempre es igual a 1.

2.3. Presentación gráfica de datos

"Un dibujo vale más que diez mil palabras", dice el viejo proverbio chino, siendo este principio tan cierto con respecto a números como a dibujos. Frecuentemente, es posible resumir toda la información importante, que se tiene de una gran cantidad de datos, en un dibujo sencillo. Así, uno de los métodos más ampliamente utilizados para representar datos es mediante gráficas[40].

Cuando hay problemas, confusión y en consecuencia una crisis, es difícil encontrar por donde empezar. Pero existen formas y una de ellas se atribuye a Wilfredo Pareto, nacido en Francia a mediados del siglo XIX, hijo de padres italianos y educado en Italia, por lo que se considera Italiano. El famoso ecónomo austriaco Schumpeter lo ve como uno de los 10 mejores economistas de todos los tiempos. También fue director de "Las fabricas del Hierro" en Italia y expresó el principio social que observó en su tiempo y sigue siendo válido en la actualidad[41].

Pareto vio que en la sociedad existen clases y que dentro de cada clase hay sólo una élite que toma las acciones y decisiones, por lo que dijo: **"pocos deciden la suerte de muchos".**

Este principio se aplica muy frecuentemente en la industria pero es universal. El llamado **diagrama de pareto** se construye de la siguiente forma:

1. Primero se deben ordenar los datos de mayor a menor.

---

[39] Matemáticas avanzadas para ingenieria. Erwin Kreysig Volumen 2. tercera edición, editorial: Limusa. P.p 908-909

[40] del libro: Probabilidad y Estadística. Stephen S. Willougbhy. Publicaciones cultural S.A p.p 69

[41] del libro del Ing. Carlos González

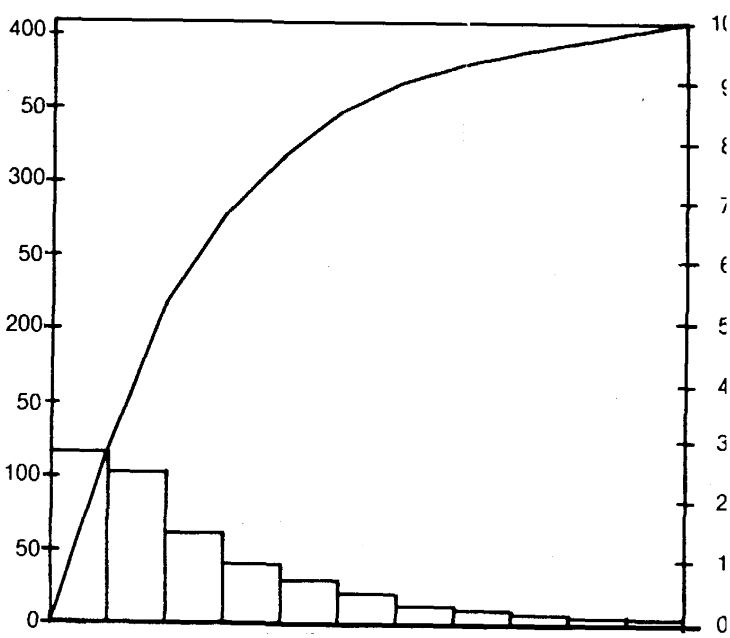
2. A continuación se hará uso de la parte de la estadística llamada descriptiva, que se auxilia del plano para mostrar gráficamente lo colectado en una tabla.

3. Para dibujar la gráfica, el primer paso será trazar los ejes X y Y; de preferencia en un plano cuadriculado o milimétrico.

4. a continuación se dibujará arriba de cada país un rectángulo que corresponda al valor numérico, según la escala del eje "Y" del lado izquierdo.

5. Ahora dibújese una diagonal en el primer rectángulo, desde el origen hasta la parte superior derecha del mismo rectángulo. Éste es el primer paso para dibujar el polígono de frecuencia acumulada; en seguida, donde termina esta primera diagonal, se unirá una segunda línea que será igual en tamaño y ángulo a la diagonal del segundo rectángulo; y así sucesivamente hasta observar que el polígono llegará a la esquina superior derecha del eje en el que se podrá apreciar el 100%. Y en efecto, la frecuencia acumulada total corresponde al 100%.

Es decir, gráficamente tendríamos la representación de estos pasos en la siguiente figura:



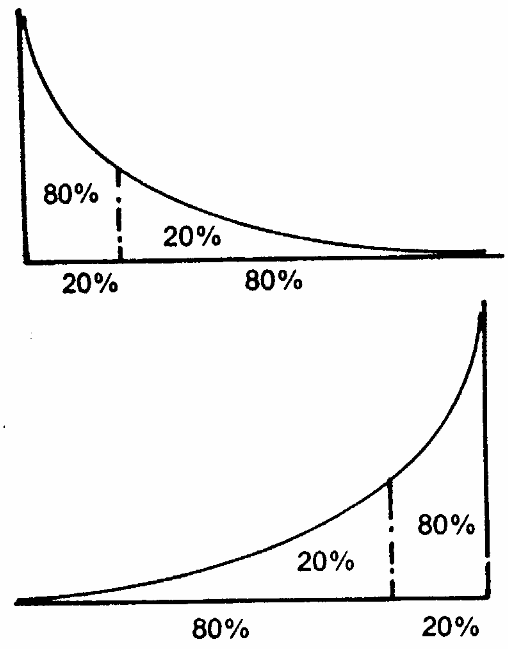
La ley de Pareto se puede extender ampliamente y podemos decir que es casi universal y general, siendo una práctica común como herramienta en el análisis de problemas y toma de decisiones, donde también se le expresa como: **"pocos vitales, muchos triviales".**

Veamos a continuación algunos ejemplos para comprender su alcance:

1. Si se trata del valor nutritivo de nuestros alimentos diarios, se podrá comprobar fácilmente que el 20 0 25% de los mismos contienen el 80% el valor nutritivo total.

2. Si proseguimos con algo cotidiano y común, imagínese un guardarropa, se podrá comprobar que el 20 o 25% del mismo es el que se usa en el 80% de las ocasiones.
3. Considérese ahora un radiorreceptor, y cuéntese el total de radiodifusoras del cuadrante. Se puede comprobar con facilidad que las personas sintonizan durante el 80% de su tiempo sólo el 20 o 25% de las estaciones.
4. Sucede lo mismo con los libros; en el 20 o 25% de sus capítulos estará contenido el 80% de su valor real.
5. En un Hospital en el que lógicamente hay pacientes, esta regla se aplica de la siguiente manera: el 80% de los recursos del hospital se usan y gastan sólo en el 20 o 25% de los pacientes.
6. Se puede decir que en una escuela, el 80% de la tecnología total de valor real para los alumnos, se concentra en el 20 o 25% de los maestros.
7. De igual modo, en un negocio, el 80% de las ventas se concentra en el 20 o 25% de sus productos o servicios y que el 80% de las ventas se lleva a cabo en el 20 o 25% de sus clientes.

Esto simplifica el análisis de los problemas cotidianos de todas las personas. Si este principio se vuelve acto reflejo, servirá para ayudar a concentrar el esfuerzo de individuos, grupos, compañías o instituciones.



Datos e histogramas:

Como ya se ha dicho, la estadística descriptiva ayuda a comprender mejor la naturaleza de los datos. La distribución de frecuencias y el histograma son dos herramientas que auxilian al analizador de problemas a visualizar mucho mejor las causas de éstos con el fin de tomar acciones[42],

Pasos a seguir para la elaboración de un diagrama de frecuencias y un histograma.

Considere el siguiente conjunto de datos:

8.9	8.3	9.2	8.4	9.1
8.6	8.9	9.1	8.8	8.8
8.8	9.1	8.9	8.7	8.8
8.9	9.0	8.6	8.7	8.4
8.6	9.0	8.8	8.9	9.1
9.4	9.0	9.2	9.1	8.8
9.1	9.3	9.0	9.2	8.8
9.7	8.9	9.7	8.3	9.3
8.9	8.8	9.3	8.5	8.9
8.3	9.2	8.2	8.9	8.7
8.9	8.8	8.5	8.4	8.0
8.5	8.7	8.7	8.8	8.8
8.3	8.6	8.7	9.0	8.7
8.4	8.8	8.4	8.6	9.0
9.3	8.8	8.5	8.7	9.6
8.5	9.1	9.0	8.8	9.1
8.6	8.6	8.4	9.1	8.5
9.1	9.2	8.8	8.5	8.3
9.3	8.6	8.7	8.7	9.1
8.8	8.7	9.0	9.0	8.5
8.5	8.8	8.9	8.2	9.0
9.0	8.7	8.7	8.9	9.4
8.3	8.6	9.2	8.7	8.7
8.7	9.7	8.9	9.2	8.8
8.3	8.6	8.5	8.6	9.7

[42] El Dr. Juran dijo que: llevar las estadísticas y la gráfica de los accidentes de determinada autopista no disminuía los mismos, sino las acciones reales que se tomaran.

9.7	9.7	9.7	9.2	9.2	máximo = 9.7
8.3	8.3	8.2	8.2	8.0	mínimo = 8.0

1º. Cuente el número de datos en la población o muestra:

en este caso son 125 lecturas, por lo tanto: n = 125

2º. Calcule el rango de los datos (R)

para determinar **el rango de los datos** lo único que se debe hacer es encontrar el número mayor y el número menor de las 125 lecturas que se tienen en la tabla. Para hacer esto, el doctor Kaouru Ishikawa[43] recomendó lo siguiente:

se toman filas o columnas, en este caso columnas y se identifica mínimo y máximo por columna, creando dos filas. En este caso, una de máximos y otra de mínimos. Estos dos números podrán ser identificados como el **máximo** y **mínimo** de las lecturas en la tabla, en este caso: MÁX = 9.7 y MÏN = 8.0, por lo que el rango: R = MÁX – MÍN. R = 9.7-8.0 = 1.7, así que: R = 1.7 unidades.

3º. Determine el número de clases, celdas o intervalos.

En la construcción de un diagrama de frecuencias o de un histograma es necesario encasillar las lecturas. Aunque existe un fórmula matemática para el cálculo del número de clases que debe tener el diagrama de frecuencias o el histograma, que depende del número de elementos en la población; hay una receta práctica a seguir que dice que el número de clases en un diagrama de frecuencias o histograma no debe ser menos de 6 ni más de 15.

Se llamará "Q" a la cantidad de clases que tendrá el histograma, y se recomienda lo siguiente:

Número de lecturas	número de clases
Menos de 50	6 – 8
50 – 100	9 – 11
100 – 250	8 – 13
más de 250	10 – 15

4º. Determinación de "c" (ancho de celda, clase o intervalo). Para este caso utilizamos la siguiente fórmula:

$$C = \frac{R}{Q} = \frac{1.7}{10} = 0.17$$

[43] Kaouru Ishikawa falleció el 16 de abril de 1989.

generalmente es necesario redondear "**c**" para trabajar con un número cómodo en las gráficas y en los cálculos. En esta ocasión daremos un valor de **c = 0.20** unidades que debe mantenerse constante a lo largo del rango, que en este caso es de **R = 1.7**

5º. Fijación de los límites de clase.

En muchos casos esto sucede automáticamente y depende de la costumbre. Por ejemplo si se le pregunta su edad a una persona, ésta contestará con el número de años que tiene. En este caso, el ancho de clase es automáticamente de un año aunque la persona haya cumplido años ayer o hace 11 meses. En otras ocasiones, la resolución en los instrumentos de medición es la que determina el ancho de clase aun cuando es necesario dar una regla general que se mantenga para lograr una normalización del histograma. En el ejemplo, la lectura menor fue de 8.0, por lo que se podría fijar el ancho de clase "c" y repetir esta operación hasta que todos los valores de la tabla queden contenidos.

6º. Construcción de la distribución de frecuencias:

Clase	Límite de clase	Marca de clase	Frecuencia	total
1	**8.00-819**	**8.1**	I	**1**
2	**8.20-839**	**8.3**	**IIIII IIII**	**9**
3	**8.40-859**	**8.5**	**IIIII IIIII IIIII I**	**16**
4	**8.60-879**	**8.7**	**IIIII IIIII IIIII IIIII IIIII II**	**27**
5	**8.80-8.99**	**8.9**	**IIIII IIIII IIIII IIIII IIIII IIIII I**	**31**
6	**9.00-9.19**	**9.1**	**IIIII IIIII IIIII IIIII III**	**23**
7	**9.20-9.39**	**9.3**	**IIIII IIIII II**	**12**
8	**9.40-9.59**	**9.5**	**II**	**2**
9	**9.60-9.79**	**9.7**	**IIII**	**4**
10	**9.80-9.99**	**9.9**		**0**
			Suma de "f" = N =	**= 125**

De donde al graficar obtenemos:

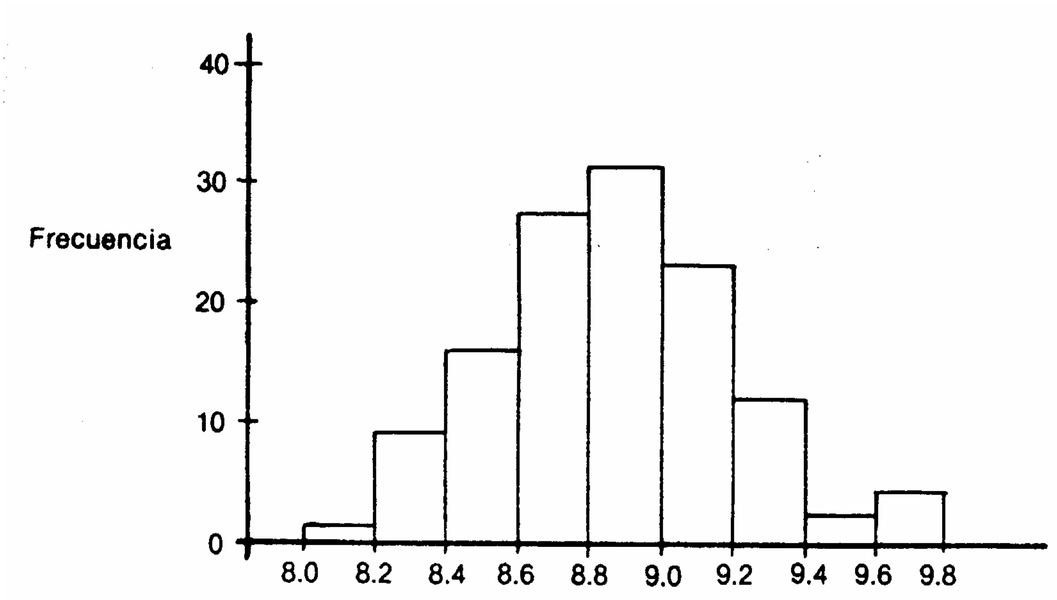
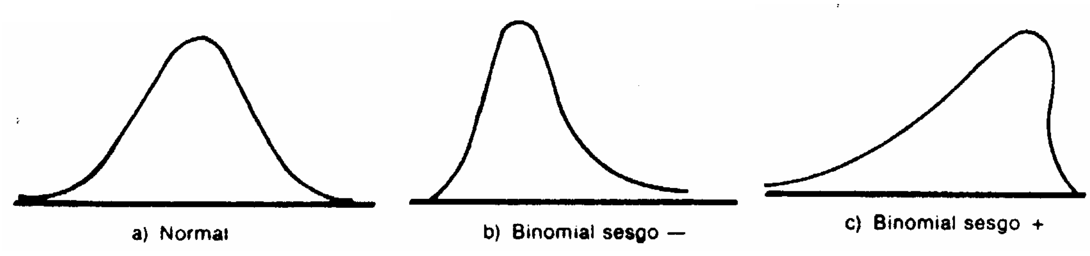
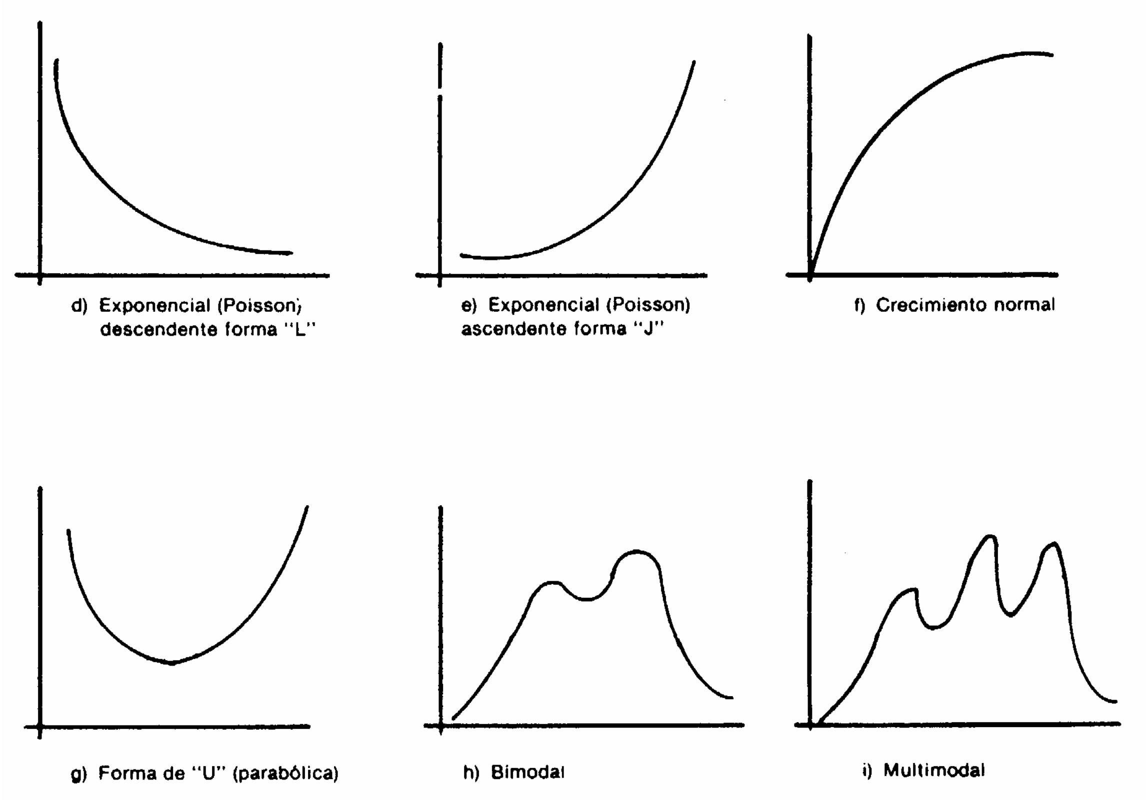


Figura 1-3

Formas de los histogramas:

Al graficar datos y construir diagramas de frecuencia e histogramas se observará que toman varias formas, las que se identificarán frecuentemente como se indica a continuación:





Representación Gráfica

Como ya lo dijimos, una gráfica es un método de presentar datos estadísticos en forma visual.

La forma más habitual de representar la información contenida en una tabla es a partir de un sistema de ejes cartesianos. Hay, no obstante, otras formas de representar datos, como posteriormente veremos, que están básicamente orientadas a caracteres no cuantitativos o atributos.

Diagrama circular

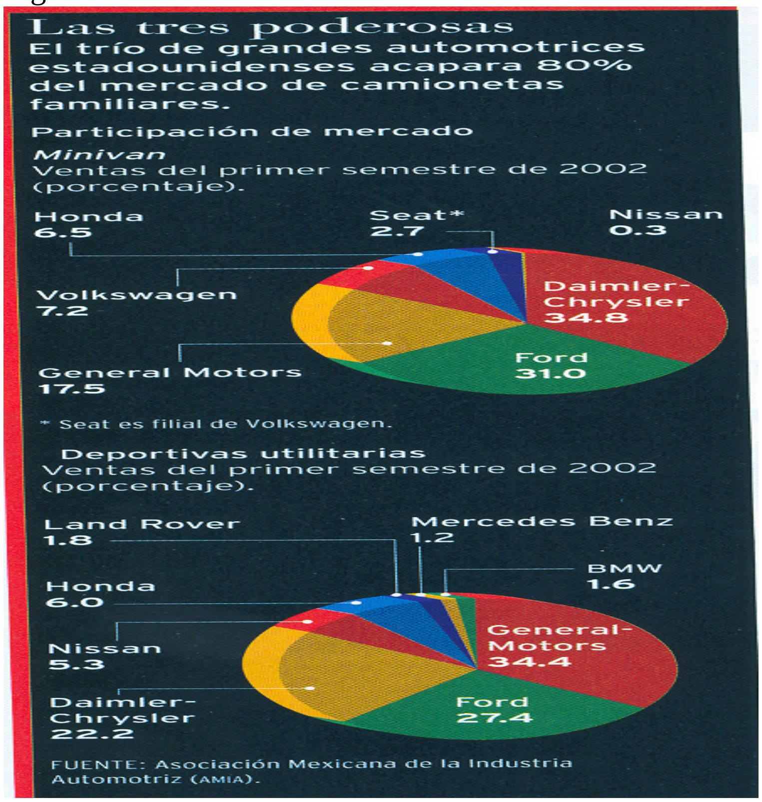
Esta representación gráfica es especialmente adecuada en aquellos casos en que se desea que los datos estadísticos lleguen a todo tipo de persona, incluso a las que no tienen por qué tener una formación científica.

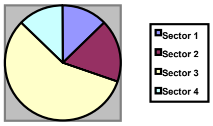
Este tipo de diagrama muestra la importancia relativa de las diferentes partes que componen un total. La forma de elaborarlo es la siguiente:

Se toma como total el circulo.
A continuación, se divide éste en tantas partes como componentes haya, siendo el tamaño de cada una de ellas proporcional a la importancia relativa de cada componente.
Más concretamente, como el círculo tiene 360ª, estos se reparten proporcionalmente a las frecuencias absolutas de cada componente.

La ventaja intrínseca de este tipo de representaciones no debe hacer olvidar que plantea ciertas desventajas, que pasamos a enumerar:

Requiere cálculos adicionales.
Es más difícil comparar segmentos de un circulo que comparar alturas de un diagrama de barras.
No da información sobre las magnitudes absolutas, a menos que las incorporemos en cada segmento.





fuente: Revista “Expansión” No. 852 (octubre 30 del 2002) p.p 52

Gráficas para distribuciones agrupadas en intervalos:

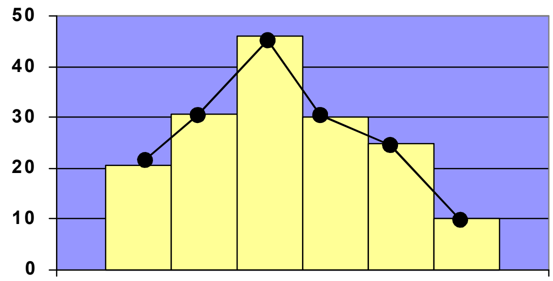
Para distribuciones agrupadas en intervalos existen básicamente tres tipos de representaciones gráficas: el histograma, el polígono de frecuencias y las ojivas.

Histograma

El histograma es similar al diagrama de barras o rectángulos, aunque con una diferencia importante: mientras que en los diagramas sólo estamos interesados en las alturas de las barras o rectángulos, en el histograma son fundamentales tanto la altura como la base de los rectángulos, haciendo el área del rectángulo proporcional a su frecuencia.

Polígono de frecuencias

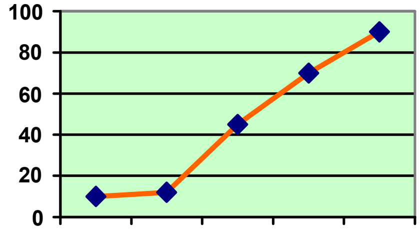
El polígono de frecuencias, que es una representación en línea, se obtiene a partir del histograma.
Su construcción se lleva a cabo uniendo los puntos medios de los lados superiores de los rectángulos del histograma



Ojivas

Si en lugar de frecuencias absolutas utilizamos sus correspondientes acumuladas, obtendremos, en vez del histograma, una representación gráfica en forma de línea creciente que se conoce con el nombre de ojiva. Estos gráficos son especialmente adecuados cuando se tiene interés en saber cuántas observaciones hay en las zona izquierda o inferior del límite superior de cualquier intervalo.

fuente: Revista Expansión No. 852 (octubre 30 del 2002) p.p 14



2.4. Medidas de tendencia central

medidas de tendencia central:

Esta serie de estadísticos, como su nombre lo indica, representa a una serie de valores con un solo valor que se localiza al centro de su distribución.
Supóngase que se tiene una serie de datos $x_1, x_2, x_3, \ldots, x_n$ y se desea encontrar la lectura que mejor los represente. *mediana, modo o moda, media geométrica, media armónica, raíz cuadrática media.*

media:
Tambien conocida como (x´barra) testada, media aritmética, bogie, línea central, valor típico, típica.

La media o "X barra", denotada por (x´barra) está representada por la suma de las observaciones, dividida entre el número total de datos que hay en la serie o conjunto de ellos.

Es decir, la media aritmética se calcula a través de la siguiente formula[44]:

$$\varpi = \frac{\sum_{i=1}^{i=n} x_i}{n} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n}$$

Por ejemplo:

sean los valores: 3, 4, 5. encuentre la media aritmética[1]:

$$\varpi = \frac{3+4+5}{3} = \frac{12}{3} = 4$$

esta fórmula se aplica a series de valores no agrupados, más adelante se verá cómo se calcula la media de valores agrupados.

mediana: (Me)
La mediana es otra medida de tendencia central que puede encontrarse arreglando una serie de medidas o datos en forma ascendente o descendente y localizándose al centro. Es decir, si el número de observaciones es non, será la del centro sin ningún cálculo posterior, pero si el número de observaciones es par, habrá que localizar las dos del centro, sumarlas y entonces dividir esa suma entre dos y esa será la mediana identificada como (Me).

Ejemplo:

sea la serie de lecturas: 3,5,4,6,2,9,8
Primero las ordenamos: 2,3,4,5,6,8,9

Entonces Me = 5 sin mayor cálculo, pues es la que se localiza en medio, exactamente a la mitad.

Pero:

[44] En el presente trabajo utilizamos "W barra" en lugar de la clásica "X barra"

Si fuera el caso ahora: 8,4,5,6,9,3
Ordenarlas: 3,4,5,6,8,9
Entonces: sumar 5 y 6 y dividir entre dos:

$$Me = \frac{5+6}{2} = 5.5$$

modo o moda (Mo)
El modo o moda es el valor que más frecuentemente ocurre en una serie de observaciones.

Ejemplos:
Por ejemplo: sean las lecturas u observaciones: 3,4,5,5,5,8,9
Mo= 5 ya que se repite 3 veces el 5 siendo una distribución *unimodal*.

Ejemplo2. sean ahora: 3,8,4,4,5,9,9,7
Tiene dos modos Mo = 4 y Mo = 9
Ya que se repiten 2 veces cada uno siendo una distribución *bimodal.*

Ejemplo3: sean ahora: 3,4,5,7,8,9,10.
Esta distribución no tiene modo por lo que es una distribución *amodal.*

media geométrica:
La media geométrica es la medida de tendencia central que nos sirve para identificar distribuciones de datos que no varían en forma aritmética, sino geométricamente y se calcula sacando la raíz enésima del producto de todas las observaciones. Se expresa matemáticamente por la fórmula:

$$G = \sqrt[n]{(x_1)(x_2)(x_3)\ldots(x_n)}$$

Ejemplo[45]:

el número de pacientes contagiados de una enfermedad en una localidad ha sido para iguales periodos: 1,2,4,8
¿cuál es el valor que mejor identifica los datos?

$$G = \sqrt[4]{(1)(2)(4)(8)}$$

[45] Del libro del Sr. Carlos González

$$G = \sqrt[4]{64}$$
$$G = 2.83$$

a diferencia de la media aritmética cuyo cálculo da x´barra=3.75

Media armónica "h"
La media armónica es otra medida de tendencia central que se define matemáticamente como la división del número de observaciones entre la suma de sus inversos. Se recomienda su uso para datos que estén en función del tiempo (velocidades físicas), como Km/h, m/s.

Se presenta matemáticamente de la siguiente manera:

$$H = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \frac{1}{x_3} + \ldots + \frac{1}{x_n}}$$

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

Ejemplo[46] :

Se quiere calcular la velocidad substituto y única que supla matemáticamente la siguiente condición: entre las ciudades A y B hay 30 Km de distancia. Se va de A a B a una velocidad de 30 Km/h y se regresa de B a A a una velocidad de 60 Km/h. ¿cuál es la velocidad única para ir y venir que cumple con el tiempo original?

Solución: aplicando la fórmula correspondiente tenemos que:

$$H = \frac{2}{\frac{1}{30} + \frac{1}{60}} = \frac{2}{\frac{2+1}{60}} = \frac{2}{\frac{3}{60}} = 40 \frac{Km}{hr.}$$

Finalmente vemos que la velocidad única requerida es de 40 Km/h

[46] Del libro del Sr. Carlos González

2.5. Medidas de dispersión

La **varianza de una población** de **N** observaciones: **$y_1, y_2, y_3, \ldots, y_N$** se define como el promedio del cuadrado de las desviaciones con respecto a su media **μ**. La varianza de la población se denota por la letra griega sigma elevada al cuadrado: δ^2 y está dada por la fórmula siguiente:

$$\delta^2 = \frac{1}{N}\sum_{i=1}^{i=N}(y_i - \mu)^2$$

nótese que se utiliza la letra mayúscula N para denotar el número de elementos en una población.

La **varianza de una muestra** de "**n**" observaciones: **$y_1, y_2, y_3, \ldots, y_n$** se define como la suma de los cuadrados de las desviaciones de las observaciones respecto de su media: ϖ^2, dividida esta suma entre (**n-1**). La varianza de la muestra se denota por: S^2 y está dada por la siguiente fórmula:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{i=n}(y_i - \varpi)$$

nótese que se utiliza la letra minúscula "n" para denotar el número de elementos de la muestra.

La **desviación estándar** de un conjunto de "**n**" observaciones: **$y_1, y_2, y_3, \ldots, y_n$** es igual a la raíz cuadrada positiva de la varianza. Así por ejemplo, la desviación estándar muestral sería igual a:

$$S = \sqrt{S^2} = \sqrt{\frac{\sum_{i=1}^{i=n}(y_i - \varpi)^2}{n-1}}$$

la desviación estándar de la población sería un caso similar y se denota por la letra griega sigma, cuyo símbolo es: δ .

Es importante hacer notar que la varianza se mide en términos del cuadrado de las unidades originales así, si las observaciones están medidas en centímetros, la varianza estaría dada en centímetros cuadrados. Al tomar la raíz cuadrada de la varianza, se obtiene la desviación estándar, con lo que se regresa a las unidades originales de las observaciones.

El coeficiente de variación[47]:

El coeficiente de variación es una medida relativa de variabilidad, porque evalúa la desviación estándar en relación con la media.

En resumidas cuentas, el coeficiente de variación es una medida estadística descriptiva que nos indica lo grande que es la desviación estándar en comparación con la media. Y se calcula como sigue:

$$C.V = \frac{desviación estándar}{media aritmética} \times 100$$

por lo tanto, si tenemos un caso en el que el promedio muestral sea de 44 y la desviación estándar de 8, el coeficiente de variación nos dará un valor de 18.2. interpretando estos datos, el coeficiente de variación nos dice que la desviación estándar de la muestra es el 18.2% del valor de la media de la muestra. En general, el coeficiente de variación es un estadístico útil para comparar la dispersión de conjuntos de datos que tienen distintas desviaciones estándar y distintos promedios.

2.6. Teorema de Chebyssheff y regla empírica[48]:

El teorema de Chebyshev nos permite inferir el porcentaje de elementos que deben quedar dentro de una cantidad específica de desviaciones estándar respecto a la media.

Teorema de Chebyshev:

Cuando menos $1 - \frac{1}{k^2}$ de los elementos en cualquier conjunto de datos debe estar a menos de "**k**" desviaciones estándar de separación respecto a la media, siendo "**k**" cualquier valor mayor que **1.**

Por ejemplo, veamos algunas implicaciones de este teorema con k= 2, 3, y 4 desviaciones estándar:

- cuando menos, el 0.75 o 75% de los elementos deben estar a menos de **z= 2** desviaciones estándar del promedio.
- cuando menos, el 0.89 u 89% de los elementos deben estar a menos de

47 Estadística para administración y economía. Anderson, Sweeney & Williams. Editorial: international Thomson. 7a. edición. P.p: 78
48 Estadística para administración y economía. Anderson, Sweeney & Williams. Editorial: international Thomson. 7a. edición. P.p 82-84

z= 3 desviaciones estándar del promedio.

- cuando menos, el 0.94 o 94% de los elementos deben estar a menos de

  z= 4 desviaciones estándar del promedio.

Ejemplo:

Como ejemplo de aplicación, supongamos que las calificaciones de 100 alumnos en un examen parcial de estadística tuvieron un promedio de 70 y una desviación estándar de 5. ¿cuántos alumnos tuvieron calificaciones entre 60 y 80? ¿cuántos entre 58 y 82?

Solución:

Para las calificaciones entre 60 y 80 vemos que el valor de 60 está a 2 desviaciones estándar abajo del promedio y que el valor de 80 está a dos desviaciones estándar arriba. Al aplicar el teorema de Chebyshev, cuando menos el 0.75 o 75% de los elementos deben tener valores a menos de dos desviaciones estándar del promedio. Así, cuando menos 75 de los 100 alumnos deben haber obtenido calificaciones entre 60 y 80.

Y para las calificaciones entre 58 y 82. (58-70)/5 = 2.4 indica que 50 está a 2.4 desviaciones estándar abajo del promedio, y que (82-70)/5 = 2.4 indica que 82 está a 2.4 desviaciones estándar arriba cdel promedio. Y aplicando el teorema de Chebyshev con **z= 2.4** tenemos que:

$$1-\frac{1}{k^{2}}=1-\frac{1}{2.4^{2}}=0.826$$

cuando menos el 82.6% de los alumnos deben tener calificaciones entre 58 y 82.

Como podemos ver, en el teorema de Chebyshev se requiere que z sea mayor que 1 pero no necesariamente debe ser un entero.

La regla empírica:

Una de las ventajas del teorema de Chebyshev es que se aplica a cualquier conjunto de datos, independientemente de la forma de la distribución de los mismos. Sin embargo, en las aplicaciones prácticas se ha encontrado que muchos conjuntos de datos tienen una distribución en forma de colina o de campana, o sea que tienen una distribución normal. Cuando se cree que los datos tienen aproximadamente esa distribución se puede aplicar la regla empírica para determinar el porcentaje de elementos que debe estar dentro de determinada cantidad de desviaciones estándar respecto del promedio.

La regla empírica dice que: "Para datos que se distribuyen de una manera normal (en forma de campana):

- aproximadamente el 68% de los elementos están a menos de una desviación estándar de la media.

- aproximadamente el 95% de los elementos están a menos de dos desviaciones estándar de la media.

- casi todos los elementos están a menos de tres desviaciones estándar de la media".

V.gr.

En una línea de producción se llenan, automáticamente, envases de plástico con detergente líquido. Con frecuencia, los pesos de llenado tienen una distribución en forma de campana. Si el peso promedio de llenado es de 16 onzas y la desviación estándar es de 0.25 onzas, se puede aplicar la regla empírica para sacar las siguientes conclusiones:

- aproximadamente el 68% de los envases llenos tienen entre 15.75 y 16.25 onzas (esto es, a menos de una desviación estándar del promedio)

- aproximadamente el 95% de los envases llenos tienen entre 15.50 y 16.50 onzas (esto es, a menos de dos desviaciones estándar del promedio)

- casi todos los envases llenos tienen entre 15.25 y 16.75 onzas (esto es, a menos de tres desviaciones estándar del promedio).

3. Estadística inferencial

3.1. Teoría del muestreo[49].

La teoría del muestreo estudia la relación entre una población y las muestras tomadas de ella.

La teoría del Muestreo es de gran utilidad en muchos campos. Por ejemplo, hoy por hoy, todos sabemos que Vivimos[50] en un mundo en el que la mayoría de los países hacen enérgicos esfuerzos para aumentar el nivel de vida de sus poblaciones, con el fin de lograr el desarrollo equilibrado se elaboran planes detallados y se ejecutan en la medida de los posible. Para elaborar esos planes de una manera científica es necesario disponer de los hechos básicos en términos numéricos para las diferentes regiones del país y para éste como un todo.

Los recursos de los países pequeños no bastan para recopilar datos, año tras año de cada persona, empresa o institución. Por fortuna como sabemos ahora, no es indispensable incluir cada una de las unidades del universo para llegar a una cifra aceptable para el total.

Así, tenemos la aplicación de los métodos de muestreo a problemas prácticos que enfrentan los países en desarrollo, lo que ha sido de igual importancia para la creación de sistemas nacionales de Estadística. Por lo general, esos países no tienen una tradición larga de censos o de recopilaciones periódicas similares que puedan usar como contexto de su muestreo. Por lo tanto, en esas condiciones la aplicación de métodos de muestreo que produzcan resultados aceptables requiere de gran ingenio.

Una **muestra** cuidadosamente diseñada puede proporcionar la información necesaria para establecer los lineamientos que requiere un país[51], a un costo que este último podría muy posiblemente absorber.

Es decir, la teoría del muestreo se utiliza para **estimar** magnitudes desconocidas de una población, tales como la media y la varianza, llamadas a menudo **parámetros de la población** o simplemente **parámetros**, a partir del conocimiento de esas magnitudes sobre muestras, que se llaman **estadísticos de la muestra** o simplemente **estadísticos**.

La teoría del muestreo es útil también para determinar si las diferencias observadas entre dos muestras son debidas a variaciones fortuitas o si son realmente significativas. Tales cuestiones aparecen, por ejemplo, al probar un nuevo suero como tratamiento de una enfermedad o al decidir si un proceso de producción es mejor que otro. Las respuestas implican el uso de los llamados

---

[49] Murray R. Spiegel. **"Estadística"** serie Schaum. Editorial McGraw-Hill. P.p 186

[50] Raj, Des. "Teoría del Muestreo". Fondo de cultura económica. P.p 9

[51] **NOTA:** La oficina de Estadística de las Naciones Unidas se dedica entre otras funciones a encontrar la forma y los medios de ayudar a los gobiernos nacionales en la obtención de los datos estadísticos tan indispensables para la planificación del desarrollo económico y social, controlar la ejecución real de los programas y evaluar los resultados.

contrastes (o tests) de hipótesis y de significación, que son importantes en la **teoría de las decisiones**.

Metodología de muestreo[52].

Es importante hacer notar que la Metodología de muestreo debe quedar plasmada en los papeles de trabajo correspondientes.

1. **definir el objetivo.**

2. **seleccionar el plan de muestreo adecuado.**
    a) Muestreo de atributos.
    b) Muestreo de suspensión o continuación.
    c) Muestreo de variables.
    d) Muestreo de descubrimiento.
    e) Muestreo dirigido.

3. **definir el nivel de confianza y precisión deseada.**
    a) Determinar el tamaño de la muestra.

4. **seleccionar la muestra.**
    a) Aleatoria.
    b) Sistemática o intervalos.
    c) Estratificada.
    d) Conglomerados.
    e) Automatizada.

5. **realizar las pruebas.**

6. **determinar estadísticos.**
7. **evaluar resultados.**

Muestreo de Atributos[53]:

---

[52] Notas tomadas durante el curso: "El muestreo estadístico aplicado a la auditoria" impartido por el Maestro en Ciencias: José Refugio Ruiz Piña. Octubre del 2001.

La función del muestreo de atributos, es determinar “cuantos elementos” existen. Y se utiliza para estimar la frecuencia probable con la cual ocurre un determinado evento.

Donde este evento puede ser una clase de error u otro atributo de la población. Es aplicable cuándo el propósito de una auditoria puede lograrse mediante una respuesta de “sí” o “no”, “bueno” o “malo”, “blanco” o “negro”, etc.

Pero, si el tamaño de la muestra lo determinamos de acuerdo con la fórmula genérica para este caso, tendremos que:

Para la determinación del tamaño de la muestra se requiere[54]:

1. tamaño del universo.
2. tasa de error esperada.
3. homogeneidad-heterogeneidad del fenómeno.
4. precisión o margen de error.
5. exactitud o nivel de confianza.
6. número de estratos.
7. etapas de muestreo.
8. conglomeración de unidades.
9. estado del marco muestral.
10. efectividad de la muestra.
11. técnica de recolección de datos. Y
12. recursos disponibles.

Dentro de la teoría del muestreo y probabilidad existen diversos procedimientos para el cálculo de los tamaños de la muestra: todos ellos consideran los elementos que hemos enumerado. A continuación se presenta una fórmula genérica para el cálculo del tamaño de muestra. Las variables que considera la fórmula son los siguientes:

Variable	Descripción
N	Tamaño de la muestra
N	Tamaño del universo
P	Probabilidad de ocurrencia (homogeneidad del fenómeno)
Q	Probabilidad de no ocurrencia (1-p)
Me	Margen de error o precisión. Expresado como probabilidad.

[53] Tomado del curso: “el muestreo estadístico aplicado a la auditoria”. Impartido por el M.C José Refugio Ruiz Piña. Octuble del 2001. pag. 25
[54] Galindo, Caceres. Jesús. “Técnicas de Investigación en sociedad, cultura y comunicación” editorial: Addison Wesley Longman (Pearson) p.p 49-62

Nc	Nivel de confianza o exactitud. Expresado como valor **z** que determina el área de probabilidad buscada.

La fórmula utilizada es la siguiente:

$$n = \frac{NPQ}{\left[\frac{Me^2}{Nc^2}(N-1)\right] + PQ}$$

por ejemplo: supongamos que queremos calcular el tamaño de una muestra para el siguiente caso:

Variable	Descripción
n	?
N	3,000,000
P	Desconocemos la probabilidad de ocurrencia. Por esta razón asumimos el mayor punto de incertidumbre, que es de 50%, que al ser expresada como probabilidad queda como: 0.5
q	1 – 0.5 = 0.5
Me	+/- 5% de margen de error. Que expresado como probabilidad queda como: 0.05
Nc	95% de nivel de confianza o exactitud. Que expresado como valor "z" que determina el área de probabilidad buscada queda como: 1.96[1]

Al sustituir estos valores en la fórmula, quedaría como sigue:

$$n = \frac{(3{,}000{,}000)(0.5)(0.5)}{\left[\frac{(0.05)^2}{(1.96)^2}(3{,}000{,}000-1)\right] + (0.5)(0.5)}$$

de donde, al realizar las operaciones indicadas, el valor de "n" obtenido es de: 384.1.

[1] El valor de "z" se busca en las tablas de distribución de la curva normal. La mayoría de los textos de probabilidad y estadística contienen esta tabla.

Tipos de muestras[55]:

En general podemos decir que existes dos tipos de muestras, a saber:

- Muestras probabilísticas y
- Muestras no probabilísticas.

De ambas, aquellas que nos interesan son las muestras probabilísticas, pues los resultados de un muestreo no probabilístico pueden estar sesgados. Por lo tanto, nos surge la pregunta: ¿Qué es una muestra probabilística?

Muestra probabilística:

Es una muestra seleccionada de tal forma que cada artículo o persona dentro de la población tiene la misma probabilidad (distinta de cero) de ser incluida en la muestra.

No existe un método "mejor" para seleccionar una muestra probabilística de una población de interés. Quizá el método que se utilizó para seleccionar una muestra de facturas de un cajón no sea el idóneo para elegir una muestra nacional de votantes. Sin embargo, **todos los métodos de muestreo probabilistico tiene similar finalidad: permitir que el azar determine los artículos o personas que incluye la muestra.**

Muestras aleatorias y números aleatorios[56].

Para que las conclusiones de la teoría del muestreo y de la inferencia estadística sean válidas, las muestras deben ser "**representativas**" de la población. El análisis de los métodos de muestreo y problemas relacionados se llama el **diseño del experimento**.

Una forma de obtener una muestra representativa es mediante un **muestreo aleatorio**, de acuerdo con el cual, cada miembro de la población tiene la misma probabilidad de ser incluido en la muestra. Existen al menos dos métodos para lograr obtener una muestra representativa; a saber

[55] Douglas A. Lind., et al. "Estadística para administración y economía" editorial: Irwin-McGraw-Hill. P.p 222
[56] Murray R. Spiegel. "Estadística" serie Schaum. Editorial: Mc Graw-Hill. P.p 186

1. El primer método consiste en asignar un número a cada elemento de la población, escribir dicho número en una papeleta, y realizar un sorteo justo con ellas en una urna.
2. Un método alternativo consiste en recurrir a una **tabla de números aleatorios**.

Error de muestreo.

Es la diferencia entre un estadístico y el parámetro correspondiente.

Introducción[57].

La acción directiva tiene en las crisis su campo natural de trabajo, pues es atributo del director el enfrentamiento de ellas, sea para prevenirlas, sea para resolverlas o bien sea para amortiguar sus consecuencias, pero sea cual fuere la situación de la alta dirección, el uso de los métodos cuantitativos es innegable; y entre ellos, la Estadística no es precisamente el de menor uso. Y claro está, que el análisis e interpretación de los estados financieros provenientes de la contabilidad es el terreno del cual parten las acciones directivas.

Los modelos estratégicos para el manejo de las empresas suelen estudiarse para su aplicación en situaciones de normalidad, siendo muy pocos los estudios que se refieren precisamente a las coyunturas[58] de turbulencia, complejidad e incluso caos indomeñable[59], siendo aquí, donde la maximización de los recursos de la empresa y la minimización de los costos, requiere con mayor fuerza de los métodos cuantitativos y en particular de la investigación de operaciones junto con la Estadística.

El estudio o tratamiento de empresas en crisis, es ahora inevitable, ya que los momentos de crisis son ahora más comunes y frecuentes que los momentos de normalidad, provocando que en nuestros tiempos la acción directiva puede resumirse sucintamente así: Acción de síntesis sobre las situaciones criticas.

El trabajo de dirección se caracteriza por no contar con reglas fijas conocidas, y cuyos resultados son inciertos (aunque implique la obligación de acertar).

---

[57] E. González et al. “Rescate de empresas en crisis”. International thomson editores.
[58] Combinación de factores y circunstancias que, para la decisión de un asunto importante, se presenta en una nación. (diccionario de la real academia española)
[59] indomable. (DRAE)

Variables aleatorias[60].

Supongamos que hay un experimento aleatorio que genera un espacio muestral con sus puntos muestrales $E_1, E_2, \ldots$ y probabilidades asociadas $P_r(E_1)$, $P_r(E_2)$,...ahora se definirá una función en este espacio muestral.

Supongamos que hay una regla por la cual un número "U" está asociado con cada punto del espacio muestral. De conformidad con dicha regla, asignamos los números reales: $U_1, U_2, \ldots$ a los puntos $E_1, E_2, \ldots,$ respectivamente. Reuniendo todos los puntos con los cuales está asociado el numero "U_i" formamos el evento $U = u_i$ que interpretamos como: "la variable aleatoria "U" toma el valor de "u_i".

El conjunto de relaciones:

$P_r(U = u_i) = g(u_i)$, $\qquad \sum g(u_i) = 1$, $\qquad (i = 1, 2, \ldots)$

Define la distribución de probabilidad de la variable aleatoria "U".

3.2. Distribución muéstrales y el teorema central del limite.

Cada muestra de tamaño n que podemos extraer de una población proporciona una media. Si consideramos cada una de estas medias como valores de una variable aleatoria podemos estudiar su distribución que llamaremos distribución muestral de medias.

- Si tenemos una población normal N(m,s) y extraemos de ella muestras de tamaño **n**, la distribución muestral de medias sigue también una distribución normal

$$N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

- Si la población no sigue una distribución normal pero n>30, aplicando el llamado Teorema central del límite la distribución muestral de medias se aproxima también a la normal anterior.

El Teorema Central del Límite dice que si tenemos un grupo numeroso de variables independientes y todas ellas siguen el mismo modelo de distribución (cualquiera que éste sea), la suma de ellas se distribuye según una distribución normal.

[60] Raj, Des. "Teoría del Muestreo". Fondo de cultura económica. P.p 16

Ejemplo : la variable "tirar una moneda al aire" sigue la distribución de Bernouilli. Si lanzamos la moneda al aire 50 veces, la suma de estas 50 variables (cada una independiente entre si) se distribuye según una distribución normal.

Este teorema se aplica tanto a suma de variables discretas como de variables continuas.
Los parámetros de la distribución normal son:

Media : n * m (media de la variable individual multiplicada por el número de variables independientes)

Varianza : n * s2 (varianza de la variable individual multiplicada por el número de variables individuales)

Veamos ahora un ejemplo:
Se lanza una moneda al aire 100 veces, si sale cara le damos el valor 1 y si sale cruz el valor 0. Cada lanzamiento es una variable independiente que se distribuye según el modelo de Bernouilli, con media 0,5 y varianza 0,25. Calcular la probabilidad de que en estos 100 lanzamientos salgan más de 60 caras.

La variable suma de estas 100 variables independientes se distribuye, por tanto, según una distribución normal.

$$\text{Media} = 100 * 0{,}5 = 50$$
$$\text{Varianza} = 100 * 0{,}25 = 25$$

Para ver la probabilidad de que salgan más de 60 caras calculamos la variable normal tipificada equivalente:

$$Y = \frac{60 - 50}{5{,}0} = 2{,}00$$

(*) 5 es la raiz cuadrada de 25, o sea la desviación típica de esta distribución
Por lo tanto:

$$P(X > 60) = P(Y > 2{,}0) = 1 - P(Y < 2{,}0) = 1 - 0{,}9772 = 0{,}0228$$

Es decir, la probabilidad de que al tirar 100 veces la moneda salgan más de 60 caras es tan sólo del 2,28%.

El teorema del límite central[61]:

El teorema del límite central afirma que, para grandes muestras aleatorias, la distribución muestral de las medias de muestras está más próxima a una distribución de probabilidad normal. La aproximación es más precisa para muestras grandes. Ésta es una de las conclusiones más útiles en

[61] Douglas A. Lind., et al. "Estadística para administración y economía" editorial: Irwin-McGraw-Hill. P.p 234

Estadística. Es posible razonar sobre la distribución muestral de las medias de muestras sin contar con información alguna sobre la forma de la distribución original de la que se toma la muestra. En otras palabras, el teorema del límite central es válido para todas las distribuciones.

El enunciado formal del teorema del límite central es el siguiente:

Si en cualquier población se seleccionan muestras de un tamaño específico, la distribución muestral de las medias de muestras es aproximadamente un distribución normal. Esta aproximación mejora con muestras de mayor tamaño.

ley de los grandes números[62]:

La ley de los grandes números sugiere que la probabilidad de una desviación significativa de un valor de probabilidad determinado empíricamente[63], a partir de una determinado teóricamente, es menor cuanto más grande sea el número de repeticiones del experimento[64].

Muestreo[65].

Es el proceso para obtener información acerca del conjunto de una población o universo examinando solo una parte del mismo.

3.3. Estimación de parámetros

En el último tema vimos como se puede emplear la teoría del muestreo para recabar información acerca de muestras aleatorias tomadas -de una población conocida[66]. Desde un punto de vista práctico, no obstante, suele resultar más importante ser capaz de inferir información sobre la población a partir de muestras suyas. Con tal situación trata la **inferencia estadística**, que usa los principios de la teoría del muestreo.

Un problema importante de la inferencia estadística es la **estimación de parámetros de la población**, o brevemente **parámetros** (tales como la media o la varianza de la población), de los correspondientes **estadísticos muestrales**, o simplemente **estadísticos** (tales como la media y la varianza de la muestra). Consideremos este problema en nuestro presente tema.

Demos una definición técnica para poder continuar con el análisis. Utilicemos "**a**" como un símbolo genérico de un parámetro poblacional y, "**â**" para indicar una estimación de "**a**" basada en datos de la muestra. Una vez acordado esto podemos decir que:

[62] Kohler, Heinz. Estadística para negocios y economía. Editorial: CECSA primera reimpresión, México 1998. p.p 158

[63] empírico, ca. Del lat. empiricus, y este del gr. mpeirikÕj, que se rige por la experiencia.
1. adj. Relativo a la experiencia o fundado en ella. (DRAE)

[64] Jacob Bernoulli (1654-1705) fue uno de los primeros que estudiaron la probabilidad matemática. En su libro "Ars Conjectandi" (1713) apareceró la primera proposición de la ley de los grandes números. En su honor, su ombre va asociado a varios conceptos matemáticos, como los experimentos de Bernoulli en probabilidad, los números de Bernoulli en la teoría de los números y la lemniscata de Bernoulli en cálculo. (nota tomada del libro: "Probabilidad y Estadística" de Stephen S. Willoughby. Publicaciones cultural, s.a. p.p 100.

[65] Notas tomadas durante el curso: El muestreo Estadístico aplicado a la Auditoria; impartido por el M.C José Refugio Ruiz Piña. (dirección general de asuntos del personal académico. Programa de actualización académica para profesores de licenciatura. Octubre del 2001)

[66] Cuando la población de interés no es muy grande y se tiene acceso a ella, se puede calcular fácilmente los parámetros de la misma; sin embargo, en la mayoría de los casos es necesario estimar la media de la población y algunos otros parámetros. (Douglas A. Lind., et al "Estadística para administración y economía" editorial: Irwin-McGraw-Hill. P.p 242)

Estimador.

Un estimador "â" de un parámetro "a" es una función de los valores muestrales aleatorios, que proporciona una estimación puntual de "a". Un estimador es en sí una variable aleatoria y por consiguiente tiene una distribución muestral teórica.

Estimador puntual:

Valor que se calcula a partir de la información de la muestra, y que se usa para estimar el parámetro de la población.

Existe una distinción técnica entre un **estimador** como una función de variables aleatorias y una **estimación** como un único número. Tal distinción se refiere al proceso en sí (estimador) y el resultado de dicho proceso (la estimación.) Lo que en realidad importa de esta definición es que: nosotros solo podemos definir buenos procesos (estimadores), mas no garantizar buenos resultados(estimaciones).

Por ejemplo:

la media muestral (**ē**)* es el mejor estimador de una población normal (μ), sin embargo no podemos garantizar que el resultado sea óptimo todas las veces. Es decir, no podemos garantizar que, para cada muestra, la media muestral esté siempre más cerca de la media poblacional, que, digamos, la mediana muestral. Así, lo más que podemos hacer es encontrar estimadores que den buenos resultados en el limite.

* en realidad debe ser una "x barra".

Como una aproximación[67] de la media μ de una población, puede tomarse la media **ē**[68] de una muestra correspondiente, lo cual da la estimación: **û = ē,** para μ, es decir:

$$\hat{u} = \bar{e} = \frac{1}{n}\sum_{i=1}^{i=n} x_i \quad \text{----------------(1)}$$

donde **n**= tamaño de la muestra.

Del mismo modo, una estimación para la varianza de una población, es la varianza de una muestra correspondiente; es decir:

$$\sigma^2 = s^2 = \frac{1}{n-1}\sum_{i=1}^{i=n}(x_i = x)^2 \quad \text{-------------(2)}$$

[67] Kreyszig. Erwin. "Matemáticas avanzadas para ingeniería" vol. 2. editorial: limusa. P.p 958

[68] considerse a este símbolo como la media aritmética de la muestra.

evidentemente estos casos 1 y 2 son estimaciones de los parámetros para distribuciones en las que μ o bien la varianza aparecen explícitamente como parámetros, tales como las distribuciones Normal y de Poisson. Aquí, podemos mencionar que (1) es un caso muy especial del llamado **Método de los momentos**. En este método, los parámetros que van a estimarse se expresan en términos de los momentos de la distribución[69], en las fórmulas resultantes, esos momentos se reemplazan por los momentos correspondientes de la muestra. Esto proporciona las estimaciones deseadas. Aquí, el k-ésimo momento de una muestra $x_1, x_2, \ldots x_n$, es:

$$m_k = \frac{1}{n}\sum_{i=1}^{i=n}(x_i)^k$$

Estimador insesgado

Un estimador **â** que es una función de datos muestrales, se conoce como: **Estimador insesgado** del parámetro poblacional **a** si su valor esperado es igual a **e**. Dicho de otra manera, **â** es un estimador insesgado del parámetro **a** sí:

E(â) = a

La condición de que el estimador **â** es insesgado supone que el valor **promedio** de **â** es exactamente correcto.

Cuando es estimador es sesgado, la magnitud del sesgo viene dada por:

Sesgo (â) = E (â) – a

Si la media de las distribuciones de muestreo de un estadístico es igual que la del correspondiente parámetro de la población, el estadístico se llama un estimador sin sesgo del parámetro; si no, se llama un estimador sesgado. Los correspondientes valores de tales estadísticos se Llaman estimaciones sin sesgo y sesgadas, respectivamente.

Por ejemplo:

La media de las distribuciones de muestreo de medias $\mu_{\bar{x}}$ y μ, la media de la población. Por tanto, la media muestral $\mu_{\bar{x}}$ es una estimación sin sesgo de la media de la población μ.

En términos de Esperanzas, podríamos decir que un estadístico es insesgado si su esperanza es igual al correspondiente parámetro de población.

[69] Para mayor información consulte la sección 19.8 del libro: “Matemáticas avanzadas para ingeniería” de Erwin Kreyszig. Editorial Limusa. Vol. 2.

Estimador Eficiente

Se dice que un estimador es **el más eficiente** para un problema particular cuando tiene el error estándar más pequeño de todos los estimadores insesgados posibles.

Se utiliza la palabra eficiente porque, en una situación dada, el estimador hace el mejor uso posible de los datos muestrales. Y de acuerdo con la teoría estadística clásica, en términos generales se debe preferir el estimador insesgado más eficiente sobre cualquier otro. De aquí, más adelante veremos que las **Hipótesis** nos dicen cual es el estimador más eficiente de un cierto parámetro en un momento dado.

Así por ejemplo

si las distribuciones de muestreo de dos estadísticos tienen la misma media (o esperanza), el de menor varianza se llama un Estimador eficiente de la media, mientras que el otro se llama un estimador ineficiente. Los valores correspondientes de los estadísticos se llaman estimación eficiente y estimación ineficiente, respectivamente.

Si consideramos todos los posibles estadísticos cuyas distribuciones de muestreo tienen la misma media, aquel de varianza mínima se llama a veces el estimador de máxima eficiencia, o sea, el mejor estimador.

Ejemplo

Las distribuciones de muestreo de media y mediana tienen ambas la misma media, a saber, la media de la población. Sin embargo, la varianza de la distribución de muestreo de medias es menor que la varianza de la distribución de muestreo de medianas. Por tanto, la media muestral da una estimación eficiente de la media de la población, mientras la mediana de la muestra da una estimación ineficiente de ella.

De todos los estadísticos que estiman la media de la población, la media muestral proporciona la mejor (la más eficiente) estimación.

En la practica, las estimaciones ineficientes se usan con frecuencia a causa de la relativa sencillez con que se obtienen algunas de ellas.

De manera Desafortunada, las declaraciones de eficiencia dependen fuertemente de algunos supuestos. Por ejemplo, cuando la distribución de la población no es normal, la media muestral no es siempre el estimador más eficiente. Con lo cual surge un tema de investigación en la teoría estadística, es el de los llamados **estimadores robustos**: estadísticos casi insesgados y casi eficientes para una gran variedad de distribuciones poblacionales. Semejantes estimadores todavía son motivo de estudio en la teoría estadística.

Estimador consistente

Un estimador es consistente si se aproxima al parámetro poblacional con

probabilidad uno a medida que el tamaño de la muestra tiende a infinito.

Por ejemplo:

la media muestral $\mu_{\bar{x}}$ de una muestra aleatoria tiene valor esperado μ y un error estándar que se aproxima a cero a medida que "n" tiende a infinito. Por lo tanto, cuando el tamaño de la muestra tiende a infinito, la media muestral $\mu_{\bar{x}}$ se aproxima a μ tanto como se quiera. Y de acuerdo con la definición, la media muestral $\mu_{\bar{x}}$ es consistente.

Un estimador inconsistente es a todas luces un mal estimador y no es aconsejable dar una estimación imprecisa basada en una infinidad de datos, cosa que puede suceder si el sesgo de un estimador se aproxima a cero a medida que "n" tiende a infinito. Por ejemplo, utilizar el 25 percentil para estimar la mediana poblacional produciría un estimador inconsistente. También habría inconsistencia si el error estándar de un estimador no tiende a cero a medida que el tamaño muestral crece.

Por lo general, los estimadores inconsistentes son el resultado de alguna equivocación o, lo que es más probable, resultan del fracaso de una hipótesis clave.

Distribución muestral de las medias de las muestras[70]:

Una vez que se descubrió la posibilidad del error de muestreo cuando se utilizan los resultados de la muestra para estimar el parámetro de una población,

- ¿Cómo es posible hacer una predicción precisa sobre el éxito de una pasta de dientes de reciente desarrollo con base sólo en los resultados de la muestra?
- ¿De qué manera puede el departamento de control de calidad de una empresa que se dedica a la producción masiva liberar un embarque de microprocesadores, con base en una muestra de sólo diez unidades?
- ¿Cómo puede Gallup o Harris hacer una predicción precisa de una votación presidencial con base en una muestra de sólo 2000 votantes registrados, de una población de casi 90 millones de votantes?

Para responder a estas preguntas, se examina la distribución muestral de las medias de la muestra.

Al organizar la medias de todas las muestras posibles de un cierto tamaño en una distribución de probabilidad, se obtiene una **distribución muestral de las medias de las muestras.**

Distribución muestral de las medias de las muestras.

70 Douglas A. Lind., et al. "Estadística para administración y economía" editorial: Irwin-McGraw-Hill. P.p 230

Es la distribución de probabilidad de todas la medias posibles de las muestras de un tamaño de muestra dado.

En este caso debemos tomar en consideración lo siguiente: Supongamos que se toman todas las posibles muestras de tamaño "**n**" sin reposición, de una población finita de tamaño $N > n$. Si denotamos la media y la desviación típica de la distribución de muestreo de medias por: $\mu_{\bar{x}}$ y $\sigma_{\bar{x}}$ y las de la población por μ y σ, respectivamente, entonces:

$$\mu_{\bar{x}} = \mu \qquad \text{y} \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\ \mu$$

si la población es infinita o si el muestreo es con reposición, los resultados anteriores se reducen a:

$$\mu_{\bar{x}} = \mu \qquad \text{y} \qquad \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

para valores grandes de "**n**" ($n \geq 30$), la distribución de muestreo de medias es aproximadamente normal con media $\mu_{\bar{x}}$ y desviación típica $\sigma_{\bar{x}}$ independientemente de la población (en tanto en cuanto la media poblacional y la varianza sean finitas y el tamaño de la población sea al menos el doble que el de la muestra). Este resultado para una población infinita es un caso especial del **teorema central del límite** de la teoría avanzada de probabilidades, que afirma que la precisión de la aproximación mejora al crecer "**n**". Esto se indica en ocasiones diciendo que la distribución de muestreo es **asintóticamente normal**.

En caso de que la población esté normalmente distribuida, la distribución de muestreo de medias también lo está, incluso para pequeños valores de "n" (o sea, n<30).

Por ejemplo[71]:

El número de unidades producidas por un obrero que trabaja de lunes a sábado en una fábrica que produce "latas" para refresco es la siguiente: 80, 80, 76, 70, 70 y 68. Suponga que estos números constituyen la población de la cual se desea tomar una muestra de tamaño 3.

[71] problema tomado con ligeros cambios del libro: "Probabilidad y Estadística" de Stephen S. Willoughby. Editorial: Publicaciones cultural s.a. p.p 126

a) determine la media aritmética de estos números.

b) Determine la desviación estándar de los números.

c) Calcule el número de muestras de tamaño 3

d) Liste cada una de las muestras

e) Calcule la media de cada una de las muestras.

f) Encuentre la media de la distribución de las medias de las muestras.

Problema tomado con ligeros cambios del libro: "Probabilidad y Estadística" de stephen S. Willoughby. Editorial: Publicaciones cultural s.a. p.p 126

Solución al problema propuesto:

a) para encontrar la media aritmética de los numero solicitada, procedemos a utilizar la fórmula correspondiente, tomando en consideración de que si se trata de una población, entonces el símbolo a utilizar es: μ , por lo tanto:

$$\mu = \frac{1}{N}\sum_{1}^{n} x_i$$

de donde sustituyendo datos tenemos que:

$$\mu = \frac{1}{6}[80 + 80 + 76 + 70 + 70 + 68]$$

$\mu = 74$ **solución al a)**

b) para el inciso b, es recomendable elaborar la tabla indicada a continuación:

# de experimento	**Datos**	**media aritmética**	**Dato - media**	**(dato - media)elevado al cuadrado**
i	x_i	μ	$(x_i - \mu)$	$(x_i - \mu)^2$
1	80	74	6	36

2	80	74	6	36
3	76	74	2	4
4	70	74	-4	16
5	70	74	-4	16
6	68	74	-6	36
Sumatoria	444		0	**144**

En esta tabla podemos observar que la sumatoria de la columna correspondiente a la diferencia del dato menos la media, es cero, por lo tanto, hasta ese punto nuestro proceso es correcto.

Finalmente para este inciso, aplicamos la fórmula correspondiente:

$$\sigma = \sqrt{\frac{1}{N}\sum_{1}^{N}(x_i - \mu)^2}$$

de donde sustituyendo valores tenemos que:

$$\sigma = \sqrt{\frac{1}{6}(144)}$$

σ = 4.9 **respuesta al b)**

c) para dar respuesta a este inciso, debemos aplicar la fórmula correspondiente a el cálculo de combinaciones. Es decir:

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

de donde sustituyendo valores tenemos que:

$$C_r^n = \frac{6!}{3!(6-3)!}$$

$$C_r^n = \frac{6x5x4x3!}{3!(3x2x1)}$$

de donde fácilmente vemos que el número de combinaciones de 6 objetos tomados de 3 en 3 es:

$C_r^n = 20$ **respuesta al c)**

d) para dar respuesta a este inciso, es necesario realizar los siguientes pasos:

1. identificar cada uno de los datos. En nuestro caso, en virtud de que algunos datos se repiten se procede a identificarlos de las siguiente manera: **80_1, 80_2, 76, 70_1, 70_2, 68.**

2. a continuación, se colocan estos datos en forma horizontal, es decir de la siguiente forma:

80_1, 80_2, 76, 70_1, 70_2, 68.

3. como siguiente punto, se elabora una tabla donde se colocaran todas las combinaciones obtenidas siguiendo el orden indicado a continuación: la primera terna o combinación se obtiene de los tres primero datos, es decir:

Si los datos son: **80_1, 80_2, 76, 70_1, 70_2, 68.**

Entonces, la primera terna es: **80_1, 80_2, 76,**

para la segunda terna, se toman los dos primeros datos junto con el cuarto dato (es decir, no saltamos el tercer dato), por lo tanto:

recordemos que los datos son: **80_1, 80_2, 76, 70_1, 70_2, 68.**

Entonces, la segunda terna sería: **80_1, 80_2, 70_1.**

Para la tercera terna se hace lo mismo, sólo que en este caso utilizamos los dos primeros datos más el quinto dato, y así sucesivamente hasta que cubrimos todos los datos que se encuentran a la derecha de los dos primeros datos. Mediante este procedimiento, obtenemos las siguientes ternas:

Para este caso también recordemos que los datos son:

80_1, 80_2, 76, 70_1, 70_2, 68.

80_1	**80_2**	**76**
80_1	**80_2**	**70_1**
80_1	**80_2**	**70_2**
80_1	**80_2**	**68**

Continuando con este procedimiento, nos "saltamos" el segundo dato, continuando con el tercero y cuarto dato; es decir, la siguiente terna tendría la forma siguiente:

Recordemos que los datos son: **80_1, 80_2, 76, 70_1, 70_2, 68.**

80_1, 76, 70_1

siguiendo este procedimiento, podemos encontrar fácilmente las siguientes ternas:

tengamos siempre presente los datos: **80_1, 80_2, 76, 70_1, 70_2, 68.**

80_1	***76***	***70_1***
80_1	**76**	**70_2**
80_1	**76**	**68**
80_1	**70_1**	**70_2**
80_1	**70_1**	**68**
80_1	**70_2**	**68**

Una vez que hemos terminado con todas las posibles combinaciones que empiezan con el primer dato, nos continuamos de la misma forma para el segundo dato; mediante este procedimiento podemos encontrar todas las restantes combinaciones, que son:

En este caso también y por comodidad recordemos los datos:

80_1, 80_2, 76, 70_1, 70_2, 68.

80_2	76	70_1
80_2	76	70_2
80_2	76	68
80_2	70_1	70_2
80_2	70_1	68
80_2	70_2	68
76	70_1	70_2
76	70_1	68
76	70_2	68
70_1	70_2	68

e) para calcular la media de cada una de las muestras, conviene elaborar una tabla donde estén incluidas todas las muestras de tamaño 3 encontradas, por lo tanto elaboramos la siguiente tabla, donde fácilmente podemos calcular la media de cada una de las muestras requerida.

	M U E S T R A S			Media
1	80_1	80_2	76	78 2/3
2	80_1	80_2	70_1	76 2/3
3	80_1	80_2	70_2	76 2/3
4	80_1	80_2	68	76

5	***80_1***	76	70_1	75 1/3
6	***80_1***	76	70_2	75 1/3
7	***80_1***	76	68	74 2/3
8	***80_1***	70_1	70_2	73 1/3
9	***80_1***	70_1	68	72 2/3
10	***80_1***	70_2	68	72 2/3
11	***80_2***	76	70_1	75 1/3
12	***80_2***	76	70_2	75 1/3
13	***80_2***	76	68	74 2/3
14	***80_2***	70_1	70_2	73 1/3
15	***80_2***	70_1	68	72 2/3
16	***80_2***	70_2	68	72 2/3
17	***76***	70_1	70_2	72
18	***76***	70_1	68	71 1/3
19	***76***	70_2	68	71 1/3
20	***70_1***	70_2	68	69 1/3

f) si ahora consideramos el conjunto de todas las medias de las muestras como un nuevo conjunto al que podemos llamar **distribución de las medias de las muestras**, fácilmente podemos calcular la media de la distribución de las medias de las muestras, para lo cual procedemos a aplicar la formula correspondiente:

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

de donde sustituyendo datos tenemos que:

$$\mu_{\bar{x}} = 69$$

En resumen se tomaron todas las muestras aleatorias posibles de una población y para cada muestra se calculó un estadístico de muestra (la media). Debido a que cada muestra posible tiene la misma posibilidad de ser seleccionada, se puede determinar la probabilidad de que la media obtenida tenga un valor comprendido en un rango. La distribución de los valores de las medias obtenidas se conoce como distribución muestral de las medias de muestras.

Aunque en la práctica sólo se ve una muestra aleatoria específica, en teoría podría surgir cualquiera de las muestras. En consecuencia, el proceso de muestreo repetido genera la distribución muestral. Luego, la distribución muestral se utiliza para medir lo probable que podría ser obtener un resultado específico.