# PROBABILIDAD Y ESTADÍSTICA
# UNIDAD 4: PROBABILIDAD

## Definición de Probabilidad

Cuando un experimento aleatorio se repite un gran número de veces, los posibles resultados tienden a presentarse un número muy parecido de veces, lo cual indica que la frecuencia de aparición de cada resultado tiende a estabilizarse. La probabilidad se establece como la frecuencia relativa ($fr$)

*Ejemplo: La siguiente tabla muestra el número de visitantes, clasificados según su edad, que visitaron una muestra de pintura.*

| ***Edad*** | ***Nº de visitantes ($f_i$)*** | $x_i = \frac{L_i + L_s}{2}$ | ***fa*** | ***fr*** | ***fra*** |
|---|---|---|---|---|---|
| *1-14* | *175* | *7,5* | *175* | *0,172* | *0,17241379* |
| *15-28* | *215* | *21,5* | *390* | *0,212* | *0,38423645* |
| *29-42* | *340* | *35,5* | *730* | *0,335* | *0,71921182* |
| *43-56* | *200* | *49,5* | *930* | *0,197* | *0,91625616* |
| *57-70* | *70* | *63,5* | *1000* | *0,069* | *0,98522167* |
| *71-84* | *15* | *77,5* | *1015* | *0,015* | *1* |
| | *1015* | | | *1* | |

*Hay una probabilidad del 33.5% de que las personas que visitan la muestra de pintura tengan entre 29 y 42 años*

Nos interesa ahora la medida numérica de la posibilidad de que ocurra un suceso *E* cuando se realiza un experimento aleatorio. A esta medida la llamaremos probabilidad del suceso ***E*** y la representaremos por ***P(E)***.

El concepto de probabilidad no es único, pues se puede considerar desde distintos puntos de vista:

## *Definición clásica o a priori*

La probabilidad es una medida sobre la escala *0* a *1* de tal forma que:

- Al suceso imposible le corresponde el valor *0*
- Al suceso seguro le corresponde el valor *1*
- El resto de sucesos tendrán una probabilidad comprendida entre *0* y *1*

La probabilidad a priori sucede cuando se determina la posibilidad de ocurrencia de un evento dentro de un conjunto sin llevar a cabo un experimento previo.

$$P = \frac{m}{n}$$

donde:

*m= eventos favorables n= eventos totales*
*y se cumple m< n*

Esta definición clásica de probabilidad fue una de las primeras que se dieron y se atribuye a Laplace; para calcularla, es necesario conocer, antes de realizar el experimento aleatorio, el espacio muestral y el número de resultados o sucesos elementales que entran a formar parte del suceso.

La aplicación de la definición clásica de probabilidad puede presentar dificultades de aplicación cuando el espacio muestral es infinito o cuando los posibles resultados de un experimento no son equiprobables.

*Ejemplo: En un proceso de fabricación de piezas puede haber algunas defectuosas y si queremos determinar la probabilidad de que una pieza sea defectuosa no podemos utilizar la definición clásica pues necesitaríamos conocer previamente el resultado del proceso de fabricación*.

*Definición frecuentista o a posteriori*

Cuando se determina la posibilidad numérica de un evento dentro de un conjunto mediante un experimento previo:

$$P = \frac{e}{E}$$

donde:
*e= eventos favorables*
*E= total de experimentos*
*y se cumple E< e*

Esta definición de probabilidad se denomina a posteriori ya que sólo podemos dar la probabilidad de un suceso después de repetir y observar un gran número de veces el experimento aleatorio correspondiente.

*Definición Subjetiva de Probabilidad*

Tanto la definición a priori como la a posteriori se basan en las repeticiones del experimento aleatorio; pero existen muchos experimentos que no se pueden repetir bajo las mismas condiciones y por tanto no puede aplicarse la interpretación objetiva de la probabilidad.

En esos casos es necesario acudir a un punto de vista alternativo, que no dependa de las repeticiones, sino que considere la probabilidad como un concepto subjetivo que exprese el grado de creencia o confianza individual sobre la posibilidad de que el suceso ocurra.

Se trata por tanto de un juicio personal o individual y es posible por tanto que, diferentes observadores tengan distintos grados de creencia sobre los posibles resultados, igualmente válidos.

*Definición Axiomática de la Probabilidad*

La definición axiomática de la probabilidad es quizás la más simple de todas las definiciones y la menos controvertida ya que está basada en un conjunto de axiomas que establecen los requisitos mínimos para dar una definición de probabilidad.

La ventaja de esta definición es que permite un desarrollo riguroso y matemático de la probabilidad. Fue introducida por A. N. Kolmogorov y aceptada por estadísticos y matemáticos en general:

1. La probabilidad del espacio muestral es del 100%. El espacio muestral contiene todos los eventos posibles que pueden ocurrir, por lo tanto, la probabilidad de este es de 1:

$$P(S) = 1$$

2. La probabilidad sólo puede tomar valores comprendidos entre 0 y 1. Si los eventos ($E_i$) de un espacio muestral son partes mutuamente excluyentes de este, generalmente la probabilidad de ocurrencia de cada uno debe ser menor a 1 y mayor a 0, de manera que la suma de las probabilidades de todos los eventos sea igual a la probabilidad del espacio muestral.

*Sin tenemos:*

*$E_1$, $E_2$, $E_3$,……………$E_n$ eventos*

*Sus probabilidades son:*

*$P(E_1)$, $P(E_2)$, $P(E_3)$,……………….$P(E_n)$*

*donde:*

*$0<P(E_1)<100$; $0<P(E_2)<100$; $0<P(E_3)<100$;…………………… $0<P(E_n)<100$*

*Por lo tanto:*
*$P(E_1)+P(E_2)+P(E_3)+$……………….$P(E)=1$*

3. La probabilidad del suceso seguro es 1
4. La probabilidad del suceso imposible debe ser 0.
5. La probabilidad de la intersección de dos sucesos debe ser menor o igual que la probabilidad de cada uno de los sucesos por separado, es decir:

$$P(E_1 \cap E_2) \leq P(E_1)$$
$$P(E_1 \cap E_2) \leq P(E_2)$$

6. La probabilidad de la unión de sucesos debe ser mayor que la de cada uno de los sucesos por separado:

$$P(E_1 \cup E_2) \geq P(E_1)$$
$$P(E_1 \cup E_2) \geq P(E_2)$$

7. Si los sucesos son disjuntos (incompatibles) debe ocurrir que :

$$E_1 \cap E_2 = \varphi \Rightarrow P(E_1 \cup E_2) = P(E_1) + P(E_2)$$

8. La probabilidad del suceso contrario de $E_1$, es igual al complemento de ese suceso:

$$P(E_1^C) = 1 - P(E_1)$$

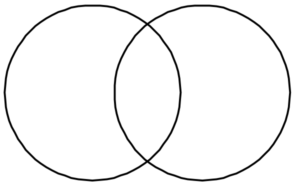
*esto en realidad puede deducirse del siguiente razonamiento:*

$$E_1 \cap E_1^C = \varphi$$
$$\Rightarrow si\ P(S) = 1 = P(E_1 \cup E^c_1) = P(E_1) + P(E^c_1)$$
$$\Rightarrow P(E^c_1) = 1 - P(E_1)$$

9. Si los eventos son independientes, la probabilidad de ocurrencia de uno u otro es igual a la multiplicación de sus respectivas probabilidades individuales:

$P(E_1 \cap E_2) = P(E_1) * P(E_2)$

Si los elementos del espacio muestral tienen elementos en común, se dicen que son dependientes, es decir, la probabilidad de ocurrencia de uno afecta a la probabilidad de ocurrencia del otro:




$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$
$$P(E_1 \cap E_2) = P(E_1) + P(E_2) - P(E_1 \cup E_2)$$

Resumiendo, a la hora de definir la probabilidad de un suceso dentro de un espacio muestral debemos tener en cuenta que:

- La función de probabilidad debe calcularse sobre el espacio muestral $S$.
- Entre las leyes que debe cumplir una función de probabilidad podemos observar que algunas son redundantes, ya que se pueden deducir de las demás. Con la definición axiomática de la probabilidad pretendemos dar el menor conjunto posible de estas reglas, para que las demás se deduzcan como una simple consecuencia de ellas.

Probabilidad Condicional

Si la probabilidad de ocurrencia de un evento depende completamente de la ocurrencia de otros eventos dentro del mismo espacio muestra, entonces el segundo evento está condicionado a la ocurrencia del primero.

$$P(E_2 \backslash E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)}$$

donde:

$P(E_2 \backslash E_1)$ = *probabilidad de ocurrencia del segundo evento dada la ocurrencia del primero*

$P(E_1 \cap E_2)$ = *probabilidad que se den tanto el primero como el segundo evento*

$P(E_1)$ = *probabilidad del primer evento*

*Ejemplo: La cancelación de un boleto de avión está sujeta a que antes ese boleto se haya vendido, entonces la cancelación (segundo evento) depende completamente de la previa venta del boleto (primer evento)*

$$P(cancelación \setminus venta) = \frac{P(venta \cap cancelación)}{P(venta)}$$

Regla de Eliminación y Teorema de Bayes

Una aplicación común de la probabilidad condicional se presenta cuando un espacio muestral formado por eventos mutuamente excluyentes se encuentra traslapado por un evento cuya ocurrencia está supeditada a la ocurrencia de al menos uno de los eventos del espacio muestral

| | | | | |
|---|---|---|---|---|
| $E_1$ | $E_2$ | $E_3$ | …………………… | $E_n$ |
| $e_1$ | $e_2$ | $e_3$ | ……….……… | $e_n$ |
| | | | | |

S

*Regla de eliminación*

Establece la posibilidad de que ocurra el evento que se traslapa parcialmente a los eventos del espacio muestral en condiciones normales de operación.

$$P(e) = P(E_1)P(e_1 \setminus E_1) + P(E_2)P(e_2 \setminus E_2) + P(E_3)P(e_3 \setminus E_3) + \dots\dots\dots + P(E_n)P(e_n \setminus E_n)$$

*donde*:

$P(e)$ =*probabilidad del evento traslapado*

$P(E_i)$ = *probabilidad del evento*

$P(e_i \setminus E_i)$ = *probabilidad de ocurrencia del evento traslapado una vez que ya ocurrió el evento principal i=1,2,3,…………….n*

*Ejemplo: Una agencia de viajes vende en promedio 80 boletos de avión, 120 excursiones, 10 visitas a bodegas y 90 reservas hoteleras. Sus estadísticas muestran que por lo general se cancelan 10 boletos de avión, 12 excursiones, 2 visitas a bodegas y 9 reservas hoteleras. Se desea conocer la probabilidad de tener al menos una cancelación el próximo mes:*

| Boletos De Avión $E1=80$ | Excursiones $E2=120$ | Visitas a Bodegas $E3=10$ | Reservas Hoteleras $E4=90$ |
|---|---|---|---|
| cancelación $e1=10$ | cancelación $e2=12$ | cancelación $e3=2$ | cancelación $e4=9$ |
| | | | |

S=300

$$P(E_1) = \frac{80}{300} = 0{,}2667$$

$$P(E_2) = \frac{120}{300} = 0{,}40$$

$$P(e_1\backslash E_1) = \frac{10}{80} = 0{,}125$$

$$P(e_2\backslash E_2) = \frac{12}{120} = 0{,}10$$

$$P(E_3) = \frac{10}{300} = 0{,}0333$$

$$P(E_4) = \frac{90}{300} = 0{,}30$$

$$P(e_3\backslash E_3) = \frac{2}{10} = 0{,}20$$

$$P(e_4\backslash E_4) = \frac{9}{90} = 0{,}10$$

$$P(e) = \left(\frac{80}{300} * \frac{10}{80}\right) + \left(\frac{120}{300} * \frac{12}{120}\right) + \left(\frac{10}{300} * \frac{2}{10}\right) + \left(\frac{90}{300} * \frac{9}{90}\right) = \frac{33}{300} = 0{,}11$$

*Teorema de Bayes*

En este caso se supone que el evento que se traslapa parcialmente sobre los eventos mutuamente excluyentes del espacio muestral ya ocurrió. Lo que se desea es conocer la probabilidad de que este evento pertenezca a una de estas partes:

$$P(E_i \setminus e_i) = \frac{P(E_i)P(e_i \setminus E_i)}{P(e)}$$

*donde:*

$P(E_i \setminus e_i)$ =*probabilidad de conocer el origen del evento traslapado*

$P(e)$ =*probabilidad del evento traslapado*

$P(E_i)$ = *probabilidad del evento*

$P(e_i \setminus E_i)$ = *probabilidad de ocurrencia del evento traslapado una vez que ya ocurrió el evento principal* $i=1,2,3,\ldots\ldots\ldots\ldots n$

*Ejemplo: si se sabe que hay una cancelación en la agencia de viajes. ¿Cuál es la probabilidad de que la misma pertenezca a una excursión?*

$$P(E_3) = \frac{10}{300}$$

$$P(e_3 \setminus E_3) = \frac{2}{10} = 0.2$$

$$P(e) = \frac{33}{100} = 0.11$$

$$P(E_3 \setminus e_3) = \frac{P(E_3)P(e_3 \setminus E_3)}{P(e)} = \frac{\frac{10}{300} * \frac{2}{10}}{\frac{33}{100}} = \frac{\frac{2}{300}}{\frac{33}{300}} = \frac{2}{33} = 0.0606$$

Distribuciones de Probabilidad

Una distribución de probabilidad la podemos concebir como una distribución teórica de frecuencia, es decir, es una distribución que describe como se espera que varíen los resultados. Dado que esta clase de distribuciones se ocupan de las expectativas son modelos de gran utilidad para hacer inferencias y tomar decisiones en condiciones de incertidumbre.

*Conceptos fundamentales: Variable aleatoria y Valor Esperado.*

- Variable aleatoria

Una variable aleatoria es una función que asocia un número real con cada elemento del espacio muestral.

- Espacio Muestral

Si un espacio muestral contiene un número finito de posibilidades o una serie interminable con tantos elementos como números enteros existen, se llama Espacio Muestral Discreto.

Si un espacio muestral contiene un número infinito de posibilidades igual al número de puntos en un segmento de línea, se llama espacio muestral continuo.

- Valor Esperado.

El valor esperado es un concepto fundamental en el estudio de las distribuciones de probabilidad. Desde hace muchos años este concepto ha sido aplicado ampliamente en el negocio de seguros y en los últimos veinte años ha sido aplicado por otros profesionales que casi siempre toman decisiones en condiciones de incertidumbre.

El valor esperado es un promedio y no es necesariamente un resultado posible del experimento. El valor esperado de una variable aleatoria discreta resulta de sumar el producto de la probabilidad asociada a cada valor que toma la variable aleatoria por el valor de la variable aleatoria, es decir

$E(X)= x1f(x1)+ x2.f(x2)+\ldots+ xn.f(xn)$

Ejemplo: suponga que se lanzan dos monedas, y se quiere averiguar el valor esperado de caras, o el promedio de caras que saldrán en dicho lanzamiento:

El espacio muestral será:

S:{(c,c); (c,s);(s,c);(s,s)}

Donde: c: el lado de la moneda que sale es cara.
s: el lado de la moneda que sale es sello.

Como los 4 puntos muestrales son igualmente probables entonces

P(X=0)=P(s,s)= ¼, es decir que la probabilidad de que no salga ninguna cara es ¼ lo que equivale a un 25%.

P(X=1)=P(c,s)+P(s,c)= ½. Es decir que la probabilidad de obtener una sola cara en un lanzamiento es 0,50 lo que equivale al 50%.

P(X=2)= ¼. Es decir que la probabilidad de obtener dos caras en el lanzamiento es ¼, lo que equivale al 50%.

E(X)=0* 1/4 + 1* 1/2 + 2 * 1/4 =1

Este resultado significa que una persona que lance dos monedas, en promedio, obtendrá una cara por lanzamiento.

*Distribución de probabilidad*

Una distribución de probabilidad es la relación que se da entre los diferentes eventos de un espacio muestral y sus respectivas probabilidades de ocurrencia.

Una forma de identificar una distribución de probabilidad es el relacionar los elementos de un arreglo de frecuencia o de una distribución de frecuencia con su respectiva frecuencia relativa, la cual se considera como la probabilidad de ocurrencia de los elementos o las clases, del arreglo de frecuencia o la distribución de frecuencias, respectivamente. De esta manera se puede trabajar de manera inductiva con la información proveniente de la estadística descriptiva.

*Distribución de probabilidad por probabilidades*

Este tipo de distribución puede ser discreta o continua. Su principal característica radica en que se tiene una probabilidad de base para la ocurrencia p y esta se aplica a una serie N de eventos sucesivos donde el orden único no cambia (por lo tanto los eventos que están afuera del tiempo de ocurrencia no cambian).
Cada evento tiene la posibilidad de ser favorable o desfavorable.

> *Ejemplo: si deseamos conocer la probabilidad de ganar la lotería, vemos que si tengo la posibilidad de participar en N sorteos, estos corresponden a eventos sucesivos, ya que los sorteos ocurren uno después de otro. Dado que existe una probabilidad p de ganar la lotería que corresponde a la probabilidad de base, se establece la probabilidad de ganar la lotería (evento favorable) o de perderla (evento desfavorable)en cada uno de los sorteos.*

*Nota: Combinaciones posibles*

*Si hay un conjunto de m elementos, una forma de encontrar las combinaciones posibles en* *dicho conjunto está dada por:*

$$\frac{n(n-1)(n-2)(n-3)\ldots(n-m+1)}{m!}$$

*Por lo tanto:*

$$\binom{n}{m}=\frac{n!}{(n-m)!m!}$$

*Ejemplo: Si tenemos 8 estudiantes. ¿Cuántos grupos de 6 integrantes podemos formar?. Si tomamos 6 de los 8 estudiantes y tenemos un primer grupo, a continuación sacamos el primer integrante y ponemos alguno de los que no habían entrado primero, de esta manera tenemos el segundo grupo; luego sacamos el segundo integrante y hacemos entrar al otro que no había estado, luego sacamos al tercero y volvemos a meter al primer alumno que sacamos la primera vez, y así sucesivamente. Por lo tanto las combinaciones posibles las obtenemos:*

$$\binom{8}{6}=\frac{8!}{(8-6)!*6!}=\frac{8!}{2!*6!}=\frac{40.320}{2*720}=28$$ Podemos formar 28 grupos de 6 estudiantes de un grupos de 8

- Distribución binomial

Un experimento a menudo consiste en pruebas repetidas, cada una con dos posibles resultados que se pueden etiquetar como éxito o fracaso. La aplicación más obvia tiene que ver con la prueba de artículos a medida que salen de una línea de producción, donde cada prueba o experimento puede indicar si un artículo está defectuoso o no. Podemos elegir definir cualquiera de los resultados como éxito.

Proceso de Bernoulli

El proceso de bernoulli tiene las siguientes propiedades

1. El experimento consiste en n pruebas que se repiten
2. Cada prueba produce un resultado que se puede clasificar como éxito o fracaso
3. La probabilidad de un éxito, se denota con p, permanece constante en cada prueba.
4. Las pruebas que se repiten son independientes.

Un experimento de bernoulli puede tener como resultado un éxito con probabilidad p o un fracaso con probabilidad q = 1 – p. Entonces la distribución de probabilidad de la variable aleatoria binomial X, el número de éxitos en n pruebas independientes, es b(x;n,p) = $\binom{n}{x} p^x q^{n-x}$ $\qquad x = 0, 1, 2, \ldots n$

*Ejemplo: el gerente de un banco estima que la probabilidad de que una empresa de primera línea requiera una autorización para girar en descubierto es de 25%. Si tienen en cartera 8 empresas. ¿Cuál es la probabilidad de que 3 empresas giren en descubierto?*

$p=0.25 \Rightarrow q=0.75$

$n=8$

$x=3$

b(3;8;0,25) = $\binom{8}{3} 0{,}25^3 0{,}75^{(8-3)} = 0{,}2076$ $\qquad x = 0, 1, 2, \ldots 8$

*Hay un 20.76% de probabilidad de que tres empresas de primera línea soliciten una autorización para girar en descubierto.*

Media, Varianza y Desviación estándar

La media de la distribución binomial es µ= *n.p*
La Varianza de la distribución binomial es $s^2$= *n.p.q*
La desviación estándar es *s* = Ö*n.p.q*

- Distribución de Poisson

Es otra distribución de probabilidad discreta muy útil cuando la variable aleatoria representa el número de eventos independientes que ocurren a una velocidad constante. Muchos eventos ocurren de manera independiente a una velocidad constante en el tiempo o espacio. Algunos ejemplos típicos son el número de personas que llegan a un hotel en un tiempo determinando, o el número de reservas realizadas en un hotel en un período de tiempo específico, etc. De hecho, la distribución de Poisson es el principal modelo de probabilidad empleado para analizar los problemas de líneas de espera. Además, ofrece una aproximación excelente a la función de probabilidad binomial cuando p es pequeño y n es grande.

Definición. La distribución de probabilidad de la variable aleatoria de Poisson X, que representa el número de resultados que ocurren en un intervalo dado o región específica que se denota con t, es

$$P(x;\mu) = \frac{e^{-\mu}\mu^x}{x!} \qquad x = 0,1,2,3 \ldots.$$

donde: e = 2,71828

μ: número promedio de resultados por unidad de tiempo o región.

***Ejemplo:***

El gerente de una pequeña empresa de turismo ha determinado que el número promedio de personas que se alojan en la posada de la empresa es 15 por fin de semana, la posada cuenta con una capacidad máxima de alojamiento de 20 personas, si para este fin de semana se han realizado 22 reservas ¿Cuál es la probabilidad de que se produzca un overbooking (es decir que la cantidad de huéspedes que asistan a la posada sea mayor a su capacidad máxima de alojamiento)?

**Solución:**

Sea X la cantidad de huéspedes que se alojan en la posada durante el fin de semana. Debemos calcular la probabilidad de que en un fin de semana cualquiera exista un overbooking, es decir que la cantidad de huéspedes que llegan a alojarse sea mayor que 20.
La variable aleatoria tiene una distribución de Poisson, por lo cual debemos calcular:

$$P(X > 20) = 1 - P(X \leq 20) = 1 - \sum_{x=0}^{20} \frac{e^{-15} 15^{x}}{x!}$$

$$= 1 - F(20)$$

$$= 1 - 0{,}9170$$

$$= 0{,}083$$

- Distribución Normal

Si una variable discreta toma los valores $x_1, ..., x_n$, existe la probabilidad igual a 1 de que al hacer un experimento, $x$ tome uno de esos valores, de modo que cada posible valor $x_i$ contribuye con una cantidad $f(x_i)$ al total:

$$\sum_{i=1}^{n} f(x_i) = \sum_{i=1}^{n} P[x = x_i] = 1$$

Aun cuando la variable tomase un número infinito de valores, $x_1, x_2, \ldots$, no hay ningún problema en comprobar que cada $x_i$ contribuye con una cantidad $f(x_i)$

Cuando la variable es continua, no tiene sentido hacer una suma de las probabilidades de cada uno de los términos en el sentido anterior, ya que el conjunto de valores que puede tomar la variable es no numerable.

La distribución Normal es frecuentemente utilizada en las aplicaciones estadísticas. Su propio nombre indica su extendida utilización, justificada por la frecuencia o normalidad con la que ciertos fenómenos tienden a parecerseensucomportamientoa estadistribución.

Muchas variables aleatorias continuas presentan una función de densidad cuya gráfica tiene forma de campana.

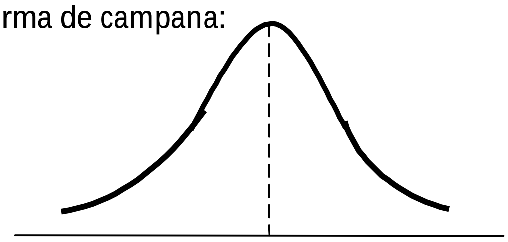
El uso extendido de la distribución normal en las aplicaciones estadísticas puede explicarse, además, por otras razones. Muchos de los procedimientos estadísticos habitualmente utilizados asumen la normalidad de los datos observados. Aunque muchas de estas técnicas no son demasiado sensibles a desviaciones de la normal y, en general, esta hipótesis puede obviarse cuando se dispone de un número suficiente de datos, resulta recomendable contrastar siempre si se puede asumir o no una distribución normal. La simple exploración visual de los datos puede sugerir la forma de su distribución. No obstante, existen otras medidas, gráficos de normalidad y contrastes de hipótesis que pueden ayudarnos a decidir, de un modo más riguroso, si la muestra de la que se dispone procede o no de una distribución normal. Cuando los datos no sean normales, podremos o bien transformarlos o emplear otros métodos estadísticos que no exijan este tipo de restricciones (los llamados métodos no paramétricos).

La distribución de una variable normal está completamente determinada por dos parámetros, su media(μ) y su desviación estándar(σ).

Con esta notación, la densidad de la normal viene dada por:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{ \frac{-1}{2} \left( \frac{x-\mu}{\sigma} \right)^2 \right\}, \quad -\infty < x < \infty$$

que determina la curva en forma de campana:



El área bajo la curva indica la probabilidad de que la variable de interés, *x*, tome un valor cualquiera en ese intervalo. Puesto que la curva alcanza su mayor altura en torno a la media, mientras que sus "ramas" se extienden asintóticamente hacia los ejes, cuando una variable siga una distribución normal, será mucho más probable observar un dato cercano al valor medio que uno que se encuentre muy alejado de éste.

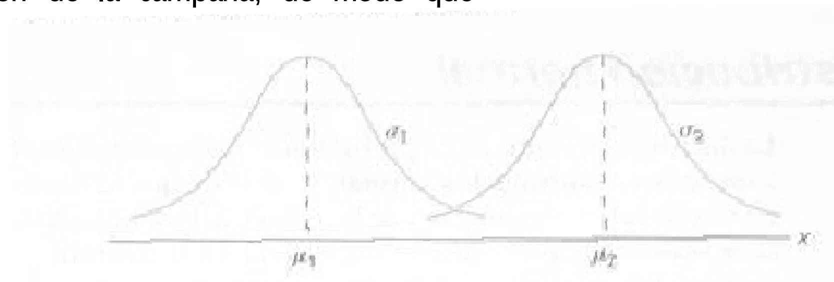
Así, se dice que una variable sigue una distribución normal de media ($\mu$) y varianza ($\sigma^2$), y se denota como:

$$X \sim N(\mu;\sigma\ )$$

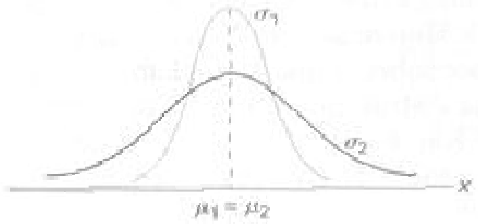
o *Propiedades de la distribución normal*:

1. Tiene una única moda, que coincide con su media y su mediana
2. La curva normal es asintótica al eje de abscisas. Por ello, cualquier valor entre $-\infty$ y $+\infty$ es teóricamente posible. El área total bajo la curva es, por tanto, igual a 1.
3. Es simétrica con respecto a su media. Por lo tanto, para este tipo de variables existe una probabilidad de un 50% de observar un dato mayor que la media, y un 50% de observar un dato menor.
4. La distancia entre la media y el punto de inflexión de la curva es igual a una desviación estándar ($\sigma$). Cuanto mayor sea la desviación estándar, más aplanada será la curva.
5. El área bajo la curva comprendida entre los valores situados aproximadamente a cuatro desviaciones estándar de la media es igual a 0.9999. En concreto, existe un 99.99% de posibilidades de observar un valor comprendido en el intervalo ($\mu \pm 4\sigma$).

6. La media indica la posición de la campana, de modo que para diferentes valores de μ, la gráfica se desplazada a lo largo del eje horizontal. Por otra parte, la desviación estándar determina el grado de apuntamiento de la curva. Cuanto mayor sea el valor de σ, más se dispersarán los datos en torno a la media y la curva será más plana. Un valor pequeño de este parámetro indica, por tanto, una gran probabilidad de obtener datos cercanos al valor medio de la distribución.




.3 Curvas normales con $\mu_1 < \mu_2$ y $\sigma_1 = \sigma_2$.




.4 Curvas normales con $\mu_1 = \mu_2$ y $\sigma_1 < \sigma_2$.

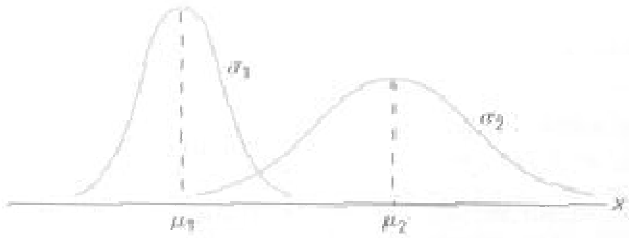



**Figura 6.5** Curvas normales con $\mu_1 < \mu_2$ y $\sigma_1 < \sigma_2$.

No existe una única distribución normal, sino una familia de distribuciones con una forma común, diferenciadas por los valores de su media y su varianza. De entre todas ellas, la más utilizada es la distribución normal estándar, que corresponde a una distribución de media $\mu = 0$ y varianza $\sigma^2 = 1$. Por lo tanto, a partir de cualquier variable *X* que siga una distribución normal, se puede obtener otra variable Z transformando la variable X de la siguiente manera:

$$z = \frac{x - \mu}{\sigma}$$

Esta propiedad resulta especialmente interesante en la práctica, ya que para una distribución Normal estándar se utiliza una tabla a partir de las que se puede obtener de modo sencillo la probabilidad de observar un dato menor o igual a un cierto valor *z*, y que permitirán resolver preguntas de probabilidad acerca del comportamiento de variables de las que se sabe o se asume que siguen una distribución aproximadamente normal.

Ejemplo: Dada un distribución normal con $\mu$=50 y $\sigma$=10, encuentra la probabilidad de que X tome un valor entre 45 y 62.

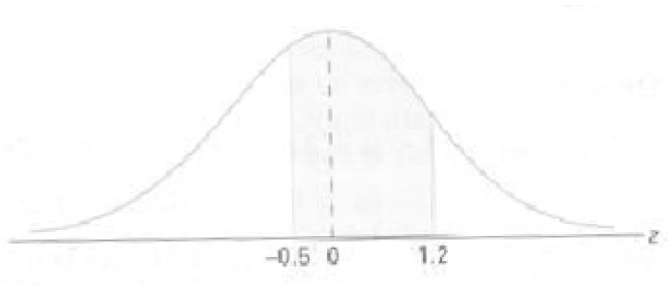
Solución: lo primero que debemos hacer es transformar los valores de X en Z, es decir, normalizar la variable X, los valores $z_i$ que corresponden a cada $x_i$ son

$$z_1 = \frac{45 - 50}{10} = -0{,}5$$

$$z_2 = \frac{62 - 50}{10} = 1{,}2$$

Por lo tanto;

P(45 < X < 62) = P(-0,5 < Z < 1,2)




Utilizando la tabla de la normal estándar buscamos el valor del área entre 0 y 1,2 y le adicionamos el valor del área entre 0 y -0,5:
P(-0,5 < Z < 1,2) = 0,57693.

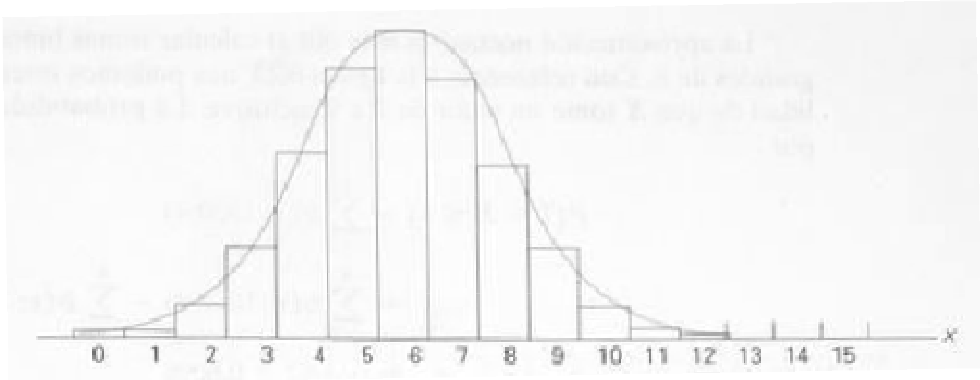
La probabilidad de que una variable aleatoria normal con media 50 y desviación estándar de 10 tome valores entre 45 y 62 es igual a 0,5769, es decir un 57,69%.

- Aproximación Normal a Variables Discretas

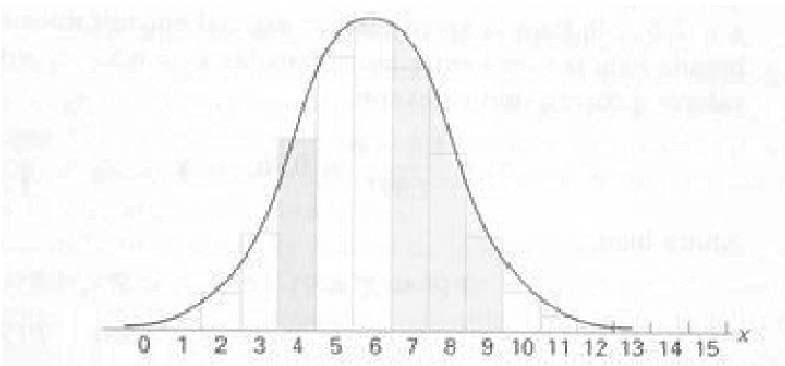
La distribución normal a menudo es una buena aproximación a una distribución discreta cuando la última adquiere una forma de campana. La distribución normal es una distribución de aproximación conveniente pues la función de distribución acumulada se tabula muy fácil.

1. Aproximación Normal a la *Binomial*: La distribución binomial se aproxima bien por la normal en problemas prácticos cuando se trabaja con la función de distribución acumulada. La distribución normal con $\mu$=n.p y $\sigma=\sqrt{n.p.q}$ proporciona una aproximación muy precisa a la distribución binomial cuando n es grande y p no está extremadamente cercana a 0 o 1, y también cuando n es pequeña y p está razonablemente cercana a ½.
Utilizaremos la aproximación normal para evaluar probabilidades binomiales siempre que n sea grande. Si n es pequeña debemos evaluar los valores de n.p y n.q, si n.p y n.q son mayores o iguales a 5, la aproximación normal a la binomial será buena.

Para ilustrar la aproximación normal a la distribución binomial, primero dibujaremos el histograma para una binomial con $n = 15$ y $p = 0,4$ y luego superponemos la curva normal particular que tenga la misma media y varianza que la variable binomial X. Por lo tnato dibujamos una curva normal con
$\mu=np = 15*0,4=6$ y $\sigma^2= npq = 15*0,4*0,6 = 3,6$ è $\sigma=\sqrt{3,6} = 1,897$




La probabilidad exacta de que la variable aleatoria binomial X tome un valor dado x es igual al ára cuya base se centra en x. Por ejemplo la probablidad de que X tome el valor 4 es igual al área del rectángulo con base centrada en x=4.




P(X=4) = $\binom{15}{4} 0,4^4 0,6^{11} = 0,1268$

Esto es aproximadamente igual al área de la región sombreada bajo la curva normal entre las dos ordenadas $x_1= 3,5$ y $x_2= 4,5$ de la figura anterior. Convirtiendo estos valores a z, tenemos

$$z_1 = \frac{3,5-6}{1,897} = -1,32$$

$$z_2 = \frac{4,5-6}{1,897} = -0,79$$

Si X es una variable aleatoria binomial y Z una variable normal estándar, entonces
$P(X=4) = P(-1,32<Z<-0,79)=0,1214$

Este valor está bastante de acurdo con el valor exacto de 0,1268.

2. Aproximación Normal a la *Poisson*: Aunque para n finito las distribuciones Poisson y normal no coinciden, es posible aproximar la Poisson por la normal de acuerdo a la siguiente regla:

| | |
|---|---|
| $\mu \geq 10$ | Aproximar a la Normal de media $\mu$, varianza $\mu$ y desv. estándar $\sqrt{\mu}$ |
| $\mu < 10$ | no aproximar, calcular con la variable original |

Referencias Bibliográficas

Canavos, George; Probabilidad y Estadística, aplicaciones y métodos; Mc Graw Hill; México; 1.988.

Ibarra Martinez, Oscar Mario; Estadística para la administración turística; Editorial Diana; México; 1998.

Walpole, Ronald E. y otros; Probabilidad y Estadística para Ingenieros, 6ta edición; Prentice Hall Hispanoamericana S.A; México; 1999.