Educaguía.com

ESTADÍSTICA

DISTRIBUCIÓN NORMAL

DISTRIBUCIÓN NORMAL

La distribución normal se llama así porque precisamente durante mucho tiempo se pensó que ése era el comportamiento normal de todos los fenómenos.
Para el caso de la distribución normal teórica, se dice que una variable aleatoria continua X sigue una distribución normal de media μ y desviación típica σ, y se designa por N (μ,σ), si se cumplen las siguientes condiciones:

- La variable recorre toda la recta real, es decir, ($+\infty$,$-\infty$)
- La función de densidad, que es la expresión en términos de ecuación

$$f(x) = \frac{1}{\sigma.\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

  matemática de la curva de Gauss, es:

A los valores μ y σ se les denomina parámetros de la distribución normal.

Función de densidad

La función de densidad tiene una expresión muy complicada, su representación gráfica tiene las siguientes características:

- La curva tiene forma campaniforme y es simétrica respecto a la recta vertical x = μ.
- La ordenada máxima se obtiene para x = μ.
- El área determinada por la campana y el eje X es igual a 1, ya que es una función de densidad, y como es simétrica respecto al eje vertical implica que el área es igual a 0,5 a cada lado de la campana.
- Las ordenadas de la curva decrecen, primero lentamente y luego con mayor rapidez, hasta hacerse nula.
- Los puntos x = $\mu-\sigma$ y x = $\mu+\sigma$ son los puntos de inflexión de la curva

De las infinitas distribuciones N(μ,σ), tiene especial interés la distribución N(0,1). Esta distribución se llama ley normal estándar, o bien distribución normal reducida.
La función de distribución estándar proporciona según sus distintos valores el área de la campana; para facilitar el cálculo de esta área se utilizan tablas.

Tipificación de la variable

Tipificar la variable consiste en transformar la variable X de una distribución N(μ, σ) en la variable Z de una distribución N(0, 1) que como se dijo anteriormente está tabulada. Para llevar a cabo esta transformación utilizamos el siguiente cambio de variable:

$$Z = \frac{X-\mu}{\sigma}$$

Manejo de tablas

Sea Z una variable que sigue una distribución normal N(0, 1). Los casos más frecuentes que se suelen presentar a la hora de buscar en las tablas son los siguientes:

- $p(Z \le 1,45)$. La probabilidad pedida se encuentra directamente en la tabla, sin más que buscar 1,4 en la columna y 0,05 en la fila; su intersección nos da la probabilidad, es decir:
  $p(Z \le 1,45) = 0,9265$.
  Esto quiere decir que el 92,65% de las observaciones son menores que 1,45.
- $p(Z \le -1,45)$. La tabla sólo proporciona probabilidades para valores de Z positivos. Pero teniendo en cuenta la simetría de la función de densidad, y que el área encerrada por toda la curva es igual a la unidad, resulta:
  $p(Z \le -1,45) = p(Z> 1,45) = 1- p(Z \le 1,45) = 1-0,9265 = 0,0735$
  Este resultado indica que sólo el 7,35% de las observaciones son menores que –1,45
- $p(1,25<Z \le 2,57)$. Su cálculo lo realizaremos restando al área mayor la menor:
  $p(1,25<Z \le 2,57) = p(Z \le 2,57) – p(Z \le 1,25) = 0,9949 – 0,8944 = 0,1005$
- $p(-2,57<Z \le -1,25)$. Como consecuencia de la simetría de la función se tiene:
  $p(-2,57<Z \le -1,25) = p(1,25<Z \le 2,57) = p(Z \le 2,57) – p(Z \le 1,25) = 0,9949 – 0,8944 = 0,1005$; Por lo tanto sólo cumplen la condición pedida el 10,05% de las observaciones.
- $p(-0,53<Z \le 2,46)$. Teniendo en cuenta todos los casos anteriores tendremos:
  $p(-0,53<Z \le 2,46) = p(Z \le 2,46) – p(Z \le -0,53) = p(Z \le 2,46) – p(Z>0,53) =$
  $= p(Z \le 2,46) – [1 – p(Z \le 0,53)] = 0,9931 – (1 – 0,7019) = 0,695.$
  Esto quiere decir que el 69,5% de las observaciones cumplen la condición pedida.

Caso particular de áreas bajo la curva:

- El 68,26 % de las observaciones se encuentran en el intervalo $(\mu-\sigma,\mu+\sigma)$.
- El 95,4 % de las observaciones se encuentran en el intervalo $(\mu-2\sigma,\mu+2\sigma)$
- El 99,7 % de las observaciones se encuentran en el intervalo $(\mu-3\sigma,\mu+3\sigma)$

Aproximación de la binomial a la normal

Teniendo en cuenta que, como se dijo anteriormente, las expresiones de la media y la desviación típica de una distribución binomial B(n, p) son:

$$\mu = n.p$$

$$\sigma = \sqrt{n.p.q}$$

De Moivre demostró que bajo determinadas condiciones, la distribución binomial B(n, p) se puede aproximar, mediante la distribución normal N(n.p, √n.p.q)

$$X \text{ es } B(n, p) \rightarrow \text{por aproximación de Moivre} \rightarrow X' \text{ es } N(n.p, \sqrt{n.p.q})$$

$$\downarrow \text{por tipificación}$$

$$Z = \frac{X - n.p}{\sqrt{n.p.q}} \text{ es } N(0,1)$$

Las condiciones de aplicabilidad que mencionamos anteriormente son:
$n.p \geq 5$ y $n.q \geq 5$.
Gracias a esta aproximación es fácil de hallar probabilidades binomiales, para valores de n muy grandes, que por el método tradicional sería muy complicada de calcular y por lo tanto prácticamente inutilizable