Documento educativo

Basic concepts of the Normal Distribution

Conceptos básicos de la Distribución Normal

Soto Espinosa Juan Luis [1] http://orcid.org/0000-0003-2600-9292, **Herrera Márquez Alma Xóchitl** [1 2] http://orcid.org/0000-0001-5039-5862 .

[1] *FES Zaragoza, UNAM*
[2] *Universidad Rosario Castellanos*

Dirección (autor principal): Batalla de 5 de Mayo esq. Fuerte de Loreto, Col Ejército de Oriente, Alcaldía Iztapalapa, C.P.09230, Ciudad de México.
Correo electrónico de contacto: *soej@unam.mx*

Fecha de envío: 01/04/2024
Fecha de aprobación: 08/05/2024

Abstract

The normal distribution is the most used distribution of continuous variables in the study of natural and social phenomena. The proposed model is applicable to different disciplines such as biology, medicine, nursing, astronomy, demography, psychology, engineering, economics, and others. disciplines.

This distribution has a characteristic graph, known as the Gaussian bell, which presents a series of characteristics that make it a fundamental tool for the study and determination of the probabilities of the occurrence of a given result, based on a set of previous historical data.

In this document we will review the basic concepts, properties, the probability density function of the distribution and the data standardization process, which gives rise to the standard normal distribution, a fundamental tool that allows the study of the different normal distributions to be standardized. existing and apply it in the study of processes and phenomena of the disciplines.

Keywords: Normal Distribution, Standard Normal Distribution, Z Value, Properties

Resumen

La distribución normal es la distribución de variables continuas más utilizada en el estudio de fenómenos naturales y sociales, el modelo propuesto es aplicable a diferentes disciplinas tales como biología, medicina, enfermería, astronomía, demografía, psicología, ingeniería, economía y muchas, muchas otras disciplinas.
Esta distribución tiene una gráfica característica, conocida como campana de Gauss, que presenta una serie de características que la convierten en una herramienta fundamental para el estudio y determinación de probabilidades de ocurrencia de un resultado determinado, partiendo de un conjunto de datos históricos previos.
En este documento revisaremos los conceptos básicos, las propiedades, la función de densidad de probabilidad de la distribución y el proceso de estandarización de datos, mismo que da origen a la distribución normal estándar, herramienta fundamental que permite homologar el estudio de las diferentes distribuciones normales existentes y aplicarla en el estudio de procesos y fenómenos de las mencionadas disciplinas.

Palabras clave: Distribución Normal, Distribución Normal Estándar, Valor Z, Propiedades

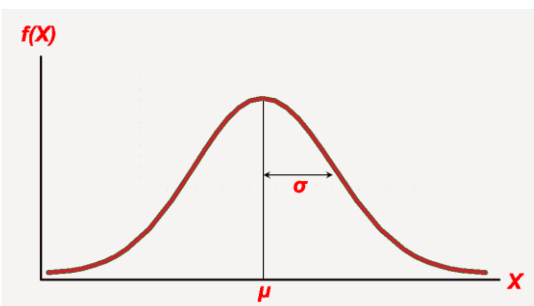
La distribución normal es considerada como la más importante de todas las distribuciones de probabilidad. Por principio de cuentas, se trata de una distribución de probabilidad de una variable continua, la cual puede adquirir valores que van de -∞ a ∞. Esta distribución es mencionada por primera vez en los trabajos en el área de juegos de azar del matemático francés Abraham de Moivre y posteriormente fue retomada por Carl Friedrich Gauss, quien profundizó en el estudio de sus propiedades y formuló por primera vez la ecuación de la curva de la distribución normal, motivo por el cual actualmente se conoce a esta curva como campana de Gauss.

Dada la amplia aplicabilidad de la distribución debido al hecho de que el comportamiento de un gran número de variables continuas en la naturaleza puede ser descrito utilizando esta forma de modelación matemática, la distribución normal es utilizada ampliamente en disciplinas como la biología, medicina, psicología, economía, astronomía, nutrición, ciencias sociales y muchas disciplinas más.

Cuando se habla de la distribución normal, hay que señalar que no existe una sola distribución normal sino toda una familia de curvas que difieren entre si, ya que corresponden a variables con diferentes escalas de medición, variabilidad y tamaño lo que ´permite que el número de distribuciones normales que se encuentran en la naturaleza sea ilimitado.

La distribución de probabilidad normal, al ser graficada, presenta una curva que tiene forma de campana y un solo pico al centro de la distribución, como se aprecia en la gráfica 1:

Gráfica 1 Curva de distribución normal



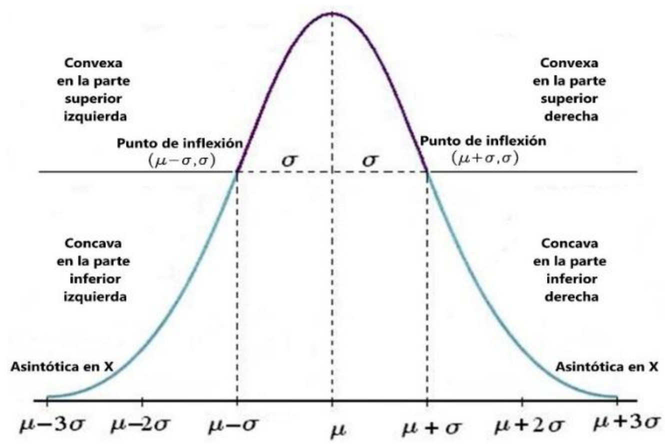
Una característica relevante es que las tres medidas de tendencia central (media aritmética, moda y mediana) tienen el mismo valor y coinciden con el pico de la curva.

Dado que la mediana se encuentra al centro de la gráfica, tenemos que el 50% de los datos se encuentran por debajo del valor de la mediana y el 50 % restante se encuentran por encima del valor de la mediana, además, el comportamiento de la probabilidad de ocurrencia de los valores se comporta de forma inversa a ambos lados de la mediana, trayendo como consecuencia que cada mitad sea una imagen perfecta de la otra, con lo que se tiene una curva simétrica respecto a la media que, como se ha mencionado, tiene el mismo valor que la mediana.

Partiendo del punto más alto de la curva, la gráfica desciende suavemente a ambos lados a partir del punto central, pasando de ser convexa en la parte superior a cóncava en la parte inferior. Cabe señalar que la curva normal nunca toca el eje de x, siendo este su límite. Debido a esta razón, se dice que la curva normal es una curva asintótica al eje de las abscisas.

La estructura de la curva se presenta en la gráfica 2:

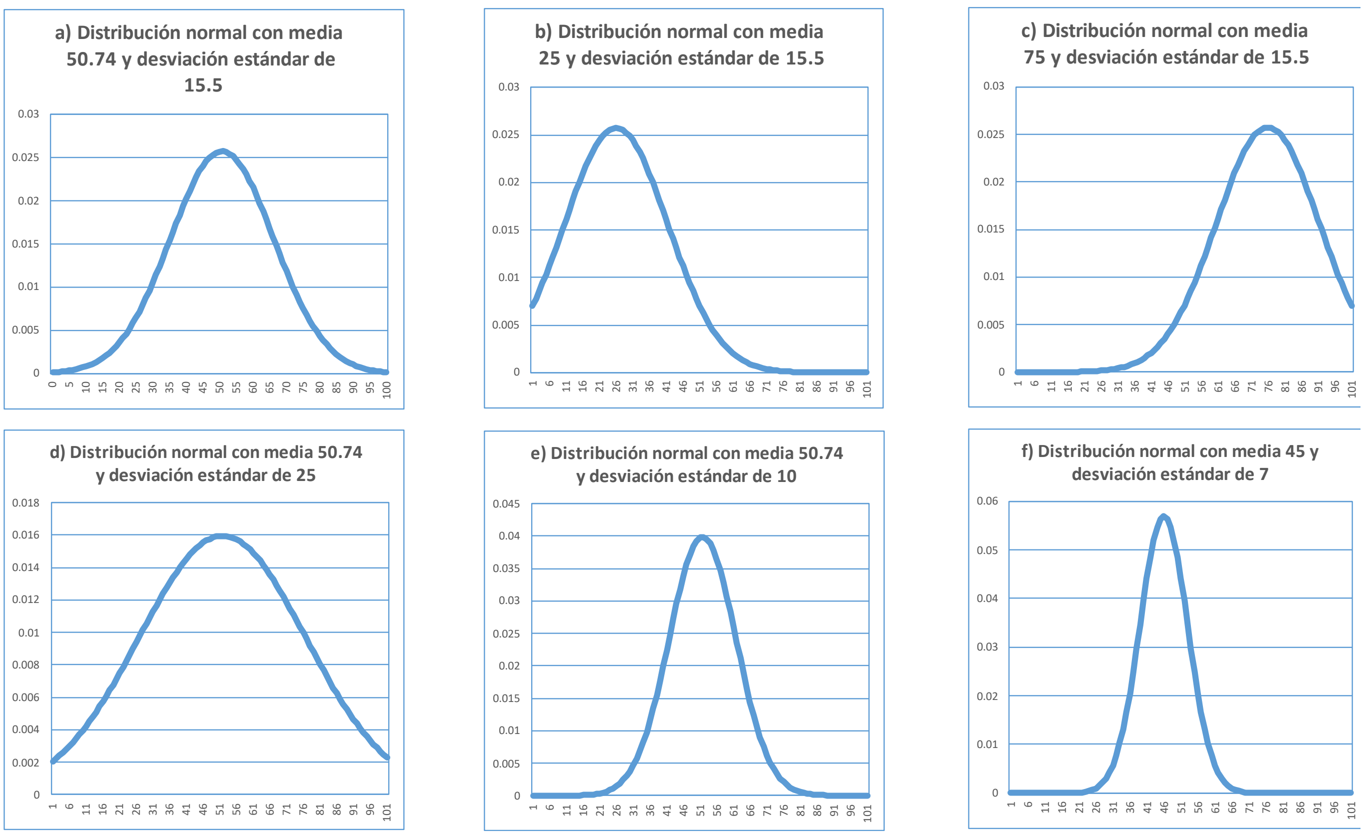
Gráfica 2 Estructura de la curva normal



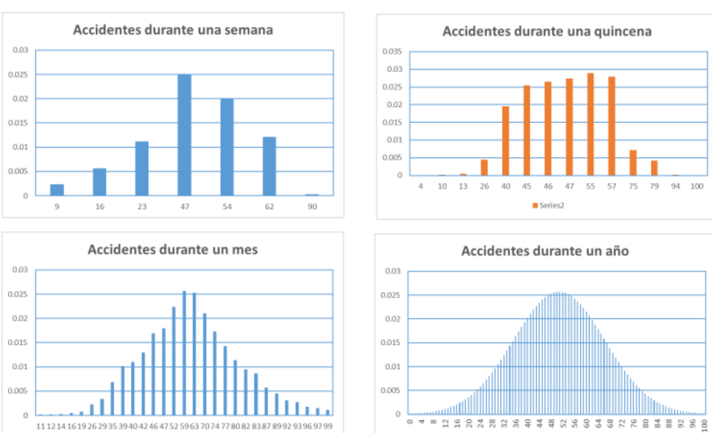
La primera curva normal, se obtuvo midiendo la distancia entre la tierra y un astro, cada una de las mediciones fue registrada. Como las mediciones fueron realizadas en intervalos de tiempo regulares y dado que tanto la tierra como el astro se encuentran en continuo movimiento, cada distancia resulta ser diferente a la anteriormente registrada, debido a la variación de las distancias y a los errores de medición cometidos durante el proceso, tanto por parte del investigador como por parte del instrumento.

Cada una de las mediciones realizadas se registra y se incluye en una gráfica, de forma que, después de un número suficientemente grande de observaciones, se va presentando un patrón en la gráfica (gráfica 3). Si se determina la expresión matemática que define el comportamiento de la curva, se obtiene una función de densidad de probabilidades. En el caso de la distribución normal, esta función de densidad depende de dos parámetros fundamentales: el valor central de la curva y la distancia que existe entre ese valor y los puntos de inflexión de la curva. Un punto de inflexión es el punto donde la curva cambia de dirección, pasando de ser cóncava a convexa y viceversa.

Gráfica 4 Ejemplos de distribución normal con media(μ) y desviación estándar(σ) diferentes



Gráfica 3 Registro de accidentes durante diferentes periodos de tiempo



Una variable aleatoria X tiene una distribución normal si los valores que adquiere, al ser graficados, forman una curva continúa acampanada. Cada distribución normal tiene una media y una desviación estándar que las define, pero con independencia de cuál sea la media y la desviación estándar, todas las distribuciones normales mantienen una forma básica de campana. Las características de cada una de las gráficas que componen a esta familia de curvas tienen una media (representada por la letra griega μ) y desviación estándar (representada por la letra griega σ) diferentes.

Al variar el valor de la media, la curva de la distribución se desplaza en el eje horizontal (abscisas), a la izquierda si el valor disminuye y a la derecha si este aumenta. Por otra parte, si la desviación estándar se incrementa, es decir aumenta la variabilidad, la campana se vuelve más ancha y su altura se reduca, mientras que, si el valor de la desviación disminuye, la curva se vuelve más angosta y la altura se incrementa. Algunos ejemplos de este comportamiento se presentan en la gráfica siguiente.

En la gráfica 4 se puede observar que las gráficas a), d) y e) tienen la misma media (50.74), pero diferente desviación estándar, de forma que la curva d tiene la desviación más alta, por lo que presenta la curva más ancha, mientras que la curva y posee la desviación más baja, por lo que presenta la curva más compacta y alargada (apuntada). Por otra parte, las gráficas a), b) y c) tienen la misma desviación estándar (15.5), pero diferente media. Nótese que el valor intermedio (50.74) ubica la curva aproximadamente al centro del plano, mientras que el valor más bajo (25) desplaza la curva hacia la izquierda y un valor más alto (75) la desplaza hacia la derecha. Si cambian al mismo tiempo la media y la desviación estándar, la gráfica se desplazará horizontalmente y adquirirá el ancho correspondiente al valor de σ, como se aprecia en la curva f.

La distribución normal es una distribución de probabilidad que se deriva del comportamiento de una variable continua, que, dicho sea de paso, es aquella característica de la población de estudio que puede ser medida y que puede adquirir un número infinito de valores dentro de un rango de medición específico. La importancia de la distribución normal se debe, principalmente a tres razones:

1. Es posible modelar un gran número de fenómenos y procesos naturales utilizando esta distribución de probabilidad. Muchas de las características que se presentan en los individuos de una población se distribuyen con este modelo.
2. Modelos matemáticos de distintas distribuciones de probabilidad después de cierto número de eventos y bajo ciertas condiciones tienden a aproximarse al comportamiento de una distribución normal.
3. El Teorema Central del Límite donde se plantea que "la suma de un gran número de variables aleatorias independientes tiende a seguir de manera asintótica una distribución normal, siempre que determinadas condiciones queden satisfechas" (Wisniewski yVelasco, 2001, pp. 211).

El parámetro de posición más utilizado en distribuciones normales es la media (μ), que se ubica en la zona de la curva con mayor densidad.

El indicador de dispersión más adecuado para definir la variabilidad de los datos es la varianza (σ2), sin embargo, el valor más utilizado es su raíz cuadrada, es decir, la desviación estándar (σ), La ventaja de este estadístico sobre la varianza es que se expresa en las mismas unidades que la variable de estudio y que la media, lo que permite comparaciones e interpretaciones más entendibles y directas.

Todas las distribuciones normales existentes comparten una serie de características, estas propiedades permiten determinar la posición relativa de un valor concreto dentro de la distribución y, como consecuencia, conocer su probabilidad de ocurrencia. Las propiedades son las siguientes:

Su distribución es simétrica, es decir, cada una de las mitades de la distribución es el reflejo exacto de la otra.

Las curvas presentan una elevación (cresta) en el centro, con colas que descienden a ambos lados de la gráfica.

La moda, la media y la mediana, como se ha mencionado, poseen el mismo valor y se encuentran al centro de la distribución.

La desviación estándar es la distancia entre la media y el punto de inflexión de la curva (el punto donde la curva pasa de ser de convexa a cóncava)

Por su forma acampanada, las probabilidades de la distribución normal cumplen la regla empírica que establece que:

1. Aproximadamente el 68% de los valores se encuentran a no más de una desviación estándar de la media, para encontrar ese rango se resta a la media la desviación estándar para el límite inferior y se suma una deviación estándar para el límite superior.
2. Cerca del 95% de los valores se encuentran a no más de dos desviaciones estándares de la media. Para obtener el rango se suma dos veces el valor de la desviación estándar y el resultado se resta a

la media para el límite inferior y se suma para el límite superior.

3. Casi todos los valores (aproximadamente el 99.7% de los valores se encuentran a no más de tres desviaciones estándar de la media. Para determinar este rango, se suma tres veces la desviación estándar y el resultado se resta a la media para el límite inferior y se suma para el límite superior.

Lo anterior puede representarse matemáticamente con las siguientes expresiones:

Si X ~ N(μ, σ), entonces

(se lee "Si X pertenece a una distribución normal de media m y desviación estándar s , entonces")

$P(\mu - \sigma < X < \mu + \sigma) \approx 0.683$

(Se lee "la probabilidad de ocurrencia de un valor de X entre la media menos la desviación estándar y la media más la desviación estándar es aproximadamente igual a 0.683)

$P(\mu - 2\sigma < X < \mu + 2\sigma) \approx 0.955$

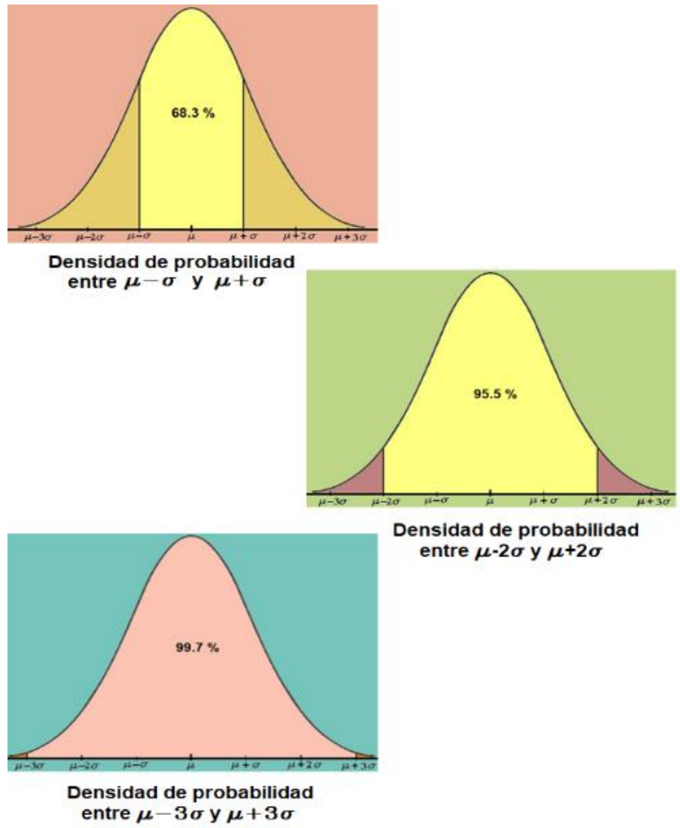
(Se lee "la probabilidad de ocurrencia de un valor de X entre la media menos dos veces la desviación estándar y la media más dos veces la desviación estándar es aproximadamente igual a 0.955)

$P(\mu - 3\sigma < X < \mu + 3\sigma) \approx 0.997$

(Se lee "la probabilidad de ocurrencia de un valor de X entre la media menos tres veces la desviación estándar y la media más tres veces la desviación estándar es aproximadamente igual a 0.997)

Gráficamente:

Gráfica 5 Probabilidad de ocurrencia de un valor entre una, dos y tres desviaciones estándar (s)



La regla empírica anterior no se aplica a todas las distribuciones y conjuntos de datos, sin embargo, de ella se deriva el Teorema de Chevyshev, que establece que "Para cualquier conjunto de datos numéricos:

- Al menos $\frac{3}{4}$ de los datos se encuentran dentro de dos desviaciones estándar de la media, es decir, en el intervalo con puntos finales $\bar{x} \pm 2s$ para muestras y con puntos finales $\mu \pm 2\sigma$ para poblaciones;
- Al menos $\frac{8}{9}$ de los datos se encuentran dentro de tres desviaciones estándar de la media, es decir, en el intervalo con puntos finales $\bar{x} \pm 3s$ para muestras y con puntos finales $\mu \pm 3\sigma$ para poblaciones;
- Al menos $1-\frac{1}{k^2}$ de los datos se encuentran dentro de las desviaciones k estándar de la media, es decir, en el intervalo con puntos finales $\bar{x} \pm ks$ para muestras y con puntos finales $\mu \pm k\sigma$ para poblaciones, donde k

representa cualquier número entero positivo que sea mayor que 1.”

Recuerde que, en notación estadística, la desviación estándar poblacional se representa con la letra griega s (sigma) mientras que la desviación estándar muestral se representa con la letra s (ese minúscula). De manera análoga, la media poblacional se representa con la letra griega m (mu) mientras que la media muestral se representa con el símbolo $\bar{x}$ (equis barra).

Es posible encontrar las probabilidades de ocurrencia para todos los intervalos de valores posibles en la distribución normal resolviendo las integrales respectivas; sin embargo, es posible obtener el valor utilizando tablas matemáticas preconstruidas, lo que facilita su aplicación y uso.

El hecho de que una variable continua (representada como X) presente una distribución normal se representa con cualquiera de las siguientes expresiones matemáticas:

$$\boldsymbol{X \rightarrow N[\mu; \sigma]}$$

O bien:

$$\boldsymbol{F(X) \rightarrow N[\mu; \sigma]}$$

Para toda variable que cumpla con esta expresión, se dice que “la variable x sigue una distribución normal con media μ y desviación estándar σ.

La función de densidad de probabilidad de la distribución está dada por:

$$\boldsymbol{f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}} \text{ para } -\infty < x < \infty$$

Donde μ y σ son la media y la desviación estándar, respectivamente, e es el número de Euler[1] y el valor de la constante p[2].

Expresado matemáticamente, la función de probabilidad correspondiente está dada por:

$$\boldsymbol{F(x) = P(X \leq x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}(\frac{v-\mu}{\sigma})^2} dv}$$

Dónde los símbolos e (constante de Euler) y π (pi) tienen valores aproximadamente 2.7183 y 3,1416 respectivamente. Como puede notarse en la función de densidad de probabilidad, la curva está definida por los valores de la media(μ) y la desviación estándar (σ), junto con las constantes e y p.

Dado el número infinito de distribuciones normales existentes (todas ellas con valores propios de media y desviación estándar), se requeriría un número infinito de tablas para representar todos los valores posibles. Para evitar este problema, se procede a normalizar los datos de cada distribución que se estudie, transformando el valor de la variable X en el valor de una variable normalizada, conocida como Z. Esta transformación de variable permite generalizar el comportamiento de todas las distribuciones existentes y abordarlas a través de un modelo de aplicación general: la Distribución Normal Estándar.

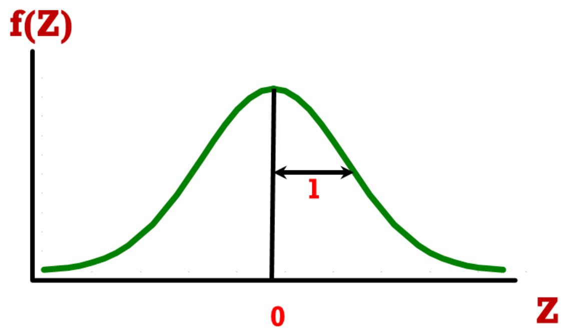
[1] La constante de Napier fue introducida por el matemático escocés John Napier, quien fue el primero en utilizar el concepto de logaritmo. Es un número irracional trascendente con infinitas cifras decimales, cuyo valor equivale a 2,71828…. Puede describir acontecimientos físicos regidos por leyes sencillas, como el vaciado de un sumidero o una veleta movida por el viento, pero también el valor del interés compuesto continuo que se utiliza en préstamos e inversiones. El número se llama así en honor de Leonhard Euler (1707-83) uno de los matemáticos más importantes de la historia. La definición formal del número es la siguiente:

$$e = \lim_{x\to\infty}(1 + \frac{1}{n})^n$$

[2] Pi es el valor obtenido de dividir la longitud de una circunferencia entre su diámetro; tiene un valor con un número infinito de cifras decimales, el cual se aproxima la mayoría de los casos a 3.1416.

Documento educativo

Gráfica 6 Distribución Normal Estándar



La **DISTRIBUCIÓN NORMAL ESTÁNDAR** (gráfica 6), conocida también como **DISTRIBUCIÓN NORMAL TÍPICA** o **DISTRIBUCIÓN NORMAL REDUCIDA**, es una distribución modelo que tiene una media con valor cero (μ =0), y una desviación estándar igual a uno (σ =1).

Esta distribución modelo, también conocida como distribución Z permite a determinar probabilidades y percentiles para el resto de las distribuciones normales. Las principales características de la distribución normal estándar.

Para obtener esta distribución modelo, se procede a transformar los valores de la variable X, en términos de la media y la desviación estándar, para obtener una variable normalizada identificada como Z. Para obtener el valor Z, se procede a restar a la variable X el valor de la media y el resultado se divide entre la desviación estándar. Matemáticamente:

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

Dónde:

Z = Variable estandarizada

X = Variable origen

μ = Media

σ = Desviación estándar

Cualquier distribución normal puede convertirse en una distribución normal estándar, transformando los valores de X en valores Z, de forma que la gráfica exprese probabilidades en término de la distancia en desviaciones estándar a la que se encuentre de la media. Cabe señalar que cuando X = μ, el valor obtenido siempre será cero dado que:

Considerando una media de 23.5, una desviación estándar de 10 y un valor de X igual a la media (23.5), tenemos que:

$$\boldsymbol{Z = \frac{X - \mu}{\sigma} = \frac{23.5 - 23.5}{10} = \frac{0}{10} = 0}$$

Esta característica de la distribución normal estándar es de suma importancia, ya que los valores de X que sean menores que la media se ubicarán a la izquierda del cero, por lo que tendrán un valor negativo; de manera análoga, los valores que sean mayores que la media se localizarán a la derecha del cero, por lo que tendrán un valor positivo.

Cuando X adquiere in valor igual a la suma de la media y la desviación estándar, tenemos que:

Considerando una media de 23.5, una desviación estándar de 10 y un valor de X igual a 23.5+10 = 33.5, tenemos:

$$\boldsymbol{Z = \frac{X - \mu}{\sigma} = \frac{33.5 - 23.5}{10} = \frac{10}{10} = 1}$$

Considerando el caso donde X es igual a la media menos la desviación estándar, tenemos:

$$\boldsymbol{Z = \frac{X - \mu}{\sigma} = \frac{13.5 - 23.5}{10} = \frac{-10}{10} = -1}$$

Nótese que cuando el valor de X a evaluar es mayor que la media y se encuentra alejado de ella una desviación estándar, el resultado de la estandarización es uno. El signo indica si el valor de X es una desviación estándar menor que la media (en el ejemplo 13.5), en cuyo caso tendrá signo negativo. Por otro lado, si el valor de la variable X es una deviación estándar mayor que la media (33.5 en el ejemplo), el signo del resultado obtenido será positivo. En estos dos casos podemos decir que el valor de X está alejado una desviación estándar de la media, pero en sentidos opuestos.

Si consideramos cualquier valor posible en la distribución, por ejemplo 45, con los mismos valores de μ y σ, obtenemos:

$$\boldsymbol{Z = \frac{X - \mu}{\sigma} = \frac{45 - 23.5}{10} = \frac{21.5}{10} = 2.15}$$

En este caso, tenemos que el valor 45 es mayo que la media, por lo que tenemos un resultado positive que se encuentra alejado de la media 2.15 veces la desviación estándar.

Consideremos ahora un valor de X igual a 15, con los mismos valores μ y σ:

$$\boldsymbol{Z = \frac{X - \mu}{\sigma} = \frac{15 - 23.5}{10} = \frac{-8.5}{10} = -0.85}$$

Para este ejemplo, el valor de X es menor que la media, por lo que esperamos un resultado con signo negative; el valor obtenido de -0.85 indica que el valor de X (15 en este caso), se encuentra alejado de la media 0.85 veces la desviación estándar.

Derivado de lo anterior, cualquier valor de X puede expresarse a través de un valor estandarizado Z, para toda X que se distribuya normalmente, si se conocen la media y la desviación estándar.

Al poder expresar cualquier valor en términos de Z, es posible generar tablas de probabilidad que puedan ser utilizadas para determinar la probabilidad de ocurrencia de un valor sin necesidad de realizar cálculos complejos. Estas tablas se conoces como tablas de distribución normal, tablas unitarias o tablas Z.

Existen varios tipos de tablas Z de uso común en estadística:

Tabla de valores acumulados desde la media: Estas tablas permiten Se indica la probabilidad de que una estadística esté entre 0 (media) y Z. Esta table permite determinar probabilidades en puntos que se encuentren de un solo lado de la media. Su principal característica es que sus valores de probabilidad van de 0.00000 a 0.50000.

Tabla de probabilidad acumulada: Esta table permite determinar la probabilidad de que un valor estadístico sea menor que Z. Esto equivale al área de la distribución normal de los valores menores que Z, partiendo de la probabilidad de la media de la distribución. Se caracteriza porque sus valores de probabilidad van de 0.5000 (probabilidad acumulada hasta la media) hasta 0.9998. Recordemos que las curvas normales son asintóticas, por lo que por mucho que se aproximen, nunca adquirirán el valor de uno.

Tabla completa de probabilidad acumulada: Esta tabla permite determinar la probabilidad de un valor en cualquier punto de la curva, los valores de probabilidad en esta tabla van de 0.0013 a 0.99999 (prácticamente 1). Esta table se obtiene uniendo las dos tablas anteriores.

Tabla de Valores complementarios de Z: en esta table se presentan probabilidades de aquellos valores que sean mayores que Z, es la inversa de la tabla de valores acumulados desde la media y determinan la probabilidad de ocurrencia de resultados de la ecuación: 1-p(Z) (uno menos la probabilidad de Z). Sus valores van desde 0.5000 hasta 0.00002 (prácticamente cero), de manera análoga, al ser asíntota la curva, por mucho que se aproxime, el valor de probabilidad nunca será cero.

Las tablas de Z Están organizadas de la siguiente manera: La primera columna contiene la parte entera y el primer decimal de la variable estandarizada Z. La primera fila de la tabla contiene el segundo decimal de Z.

Los valores dentro de la tabla son las probabilidades correspondientes al tipo de tabla. Estas probabilidades se corresponden con el área bajo la curva normal desde el origen hasta el punto definido por el valor de Z, estos valores dependen del tipo de tabla que se utilice considerando el 0 para el acumulado desde la media; el infinito negativo para el acumulado de Z; y el infinito positivo para el acumulado complementario de Z.

Ejemplo: para encontrar 1.13, se busca en la primera columna hacia abajo hasta encontrar la fila que corresponde al valor 1.1, a continuación, se recorren las columnas hacia la derecha hasta encontrar la columna cuyo encabezado es 0.03, finalmente se lee el valor que se encuentra en la celda donde se cruzan los dos valores, que en este caso representa una probabilidad de 0.3708 para una tabla acumulativa desde la media; o 0.75490 para una tabla acumulativa.

Gráfica 7 Uso de la Tabla acumulativa desde la media

z	0.00	0.01	0.02	0.03	0.04	0.05
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.019
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.059
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.098
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.136
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.173
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.208
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.242
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.273
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.302
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.328
z	0.00	0.01	0.02	0.03	0.04	0.05
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.353
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.374
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.394
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.411

Gráfica 8 Uso de la tabla acumulativa de probabilidad

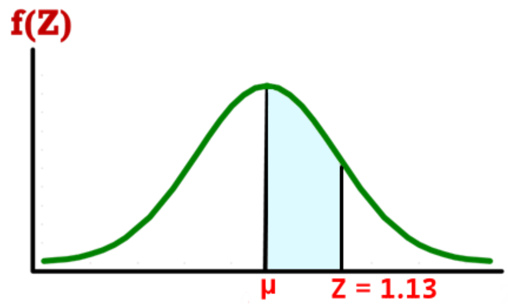
z	0.00	0.01	0.02	0.03	0.04	0.05
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.519
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.559
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.598
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.636
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.673
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.708
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.742
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.773
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.802
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.828
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.853
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.874
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.894
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.911
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.926

Nótese que el valor que se obtiene en una tabla acumulativa desde la media es de 0.3708, lo que representa el área bajo la curva normal definida entre la media y el valor Z = 1.13 (gráfica 7).

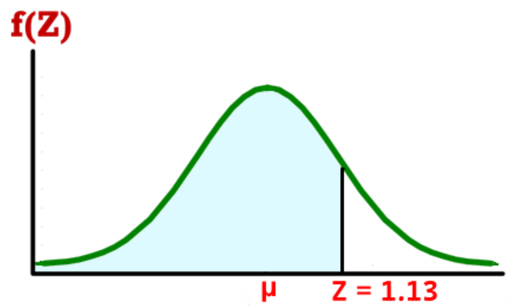
En una tabla acumulada, se obtiene un valor de 0.8708 para el mismo valor, solo que este define el área comprendida desde el valor - ∞ hasta el valor Z = 1.13 (gráfica 8)

Nótese que la diferencia entre una y otra equivale a 0.5000, que representa la mitad izquierda de la distribución (el 50 % de las probabilidades se localiza en esta zona), como puede apreciarse en la gráfica 9 y 10:

Gráfica 9 Área bajo la curva entre la media (μ) y Z=1.13

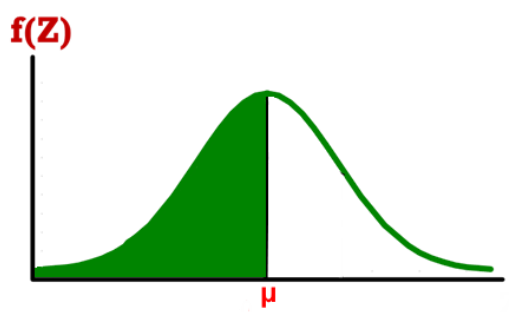


Gráfica 10 Área bajo la curva entre menos infinito (-∞) y el valor Z=1.13



Si restamos el área de la gráfica 9, que considera el valor de densidad de probabilidad desde la media hasta Z=1.13, al área de la gráfica 10, que considera la probabilidad acumulada desde menos infinito hasta Z=1.13, obtendremos la gráfica 11, que representa el área a la izquierda de la media, que equivale al 50% de las probabilidades acumuladas.

Gráfica 11 Gráfica de probabilidad acumulada desde -∞ hasta μ (distribución normal estándar)



Esta propiedad nos permite realizar todos los cálculos necesarios independientemente del tipo de tabla de valores Z con que contemos, sin embargo, ese será el tema de nuestra próxima entrega

Próxima entrega: Distribuciones de probabilidad II: Distribución normal y cálculo de probabilidades con valores Z.

Referencias:

Anderson, D. R., Sweeney, D. y Williams, T. A. (1999). Estadística para la Administración y Economía. México DF, México: International Thompson Editores.

Departamento de Didáctica de la Matemática. (2011). Estadística con proyectos. (C. Batanero, & C. Díaz, Eds.) Granada, España: Facultad de Ciencias de la Educación, Universidad de Granada.

García Pérez, A. (2008). Estadística aplicada: conceptos básicos (2a edición ed.). Madrid, España: Educación permanente / Universidad Nacional de Educación a Distancia.

Pértegas Díaz, S., & Pita Fernández, S. (2001). La distribución normal. Cad Aten Primaria, 8, 268-274.

Vargas Velasco, C. (2018). Distribución normal estándar.

Baird, D. (1991). Propiedades matemáticas de la distribución normal o de Gauss.

Rev. chil. anest. Vol. 43 Número 2 pp. 116-121|https://doi.org/10.25237/revchilanestv43n02.08

Declaración de conflicto de intereses

La autora de este artículo expresa que no tuvo ningún conflicto de intereses durante la preparación de este documento ni para su publicación.

Obra protegida con una licencia Creative Commons

Atribución - No comercial
No derivadas