Graphical user interface, application

Description automatically generated

Apuntes Unidad 3

Distribución Normal



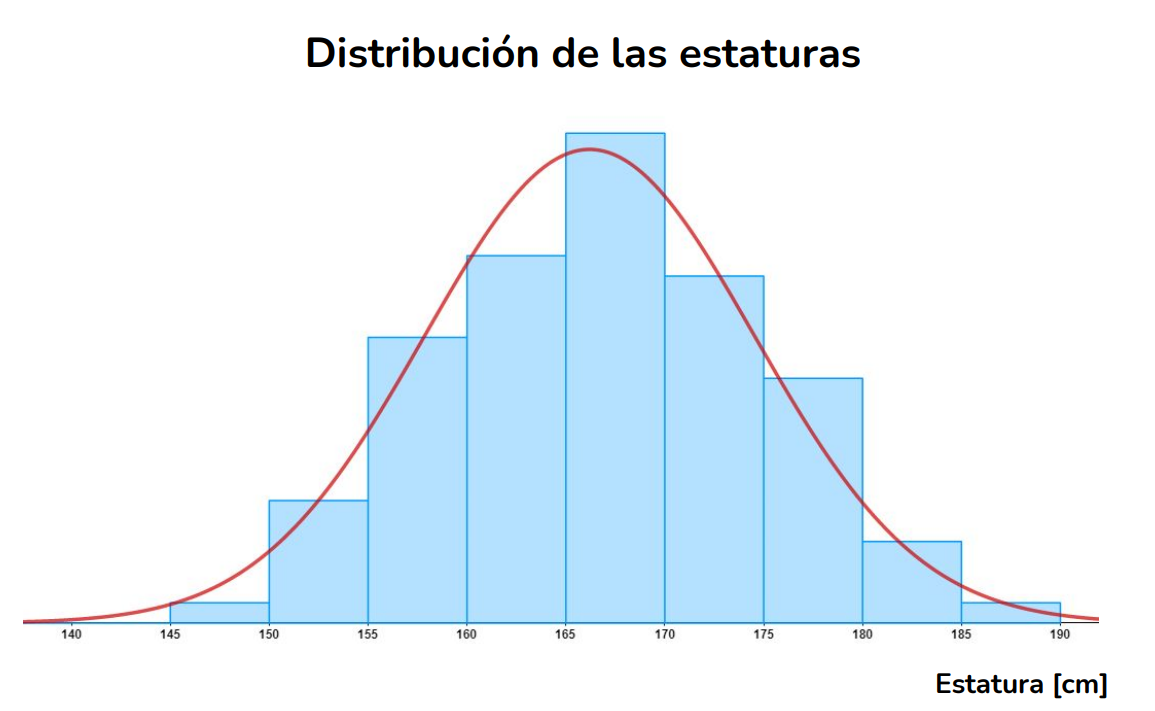
Shape, arrow

Description automatically generated

**FUNCIÓN DE DENSIDAD DE UNA DISTRIBUCIÓN NORMAL**

Muchos fenómenos que ocurren en la naturaleza o que son estudiados por las ciencias sociales presentan un histograma simétrico con forma de campana. Las variables aleatorias continuas usadas para modelar estos fenómenos se dice que siguen una distribución aproximadamente “normal”, esto en referencia a un distribución teórica que se conoce como **distribución normal.**

La gráfica de una variable aleatoria continua que sigue una distribución normal tiene una curva perfectamente simétrica con respecto a la media y su forma se asemeja a una campana. Esta gráfica se presenta en color rojo en el siguiente histograma:

Figura 1: Ejemplo de histograma que representa el fenómeno   
junto a la curva que gráfica la distribución normal.  


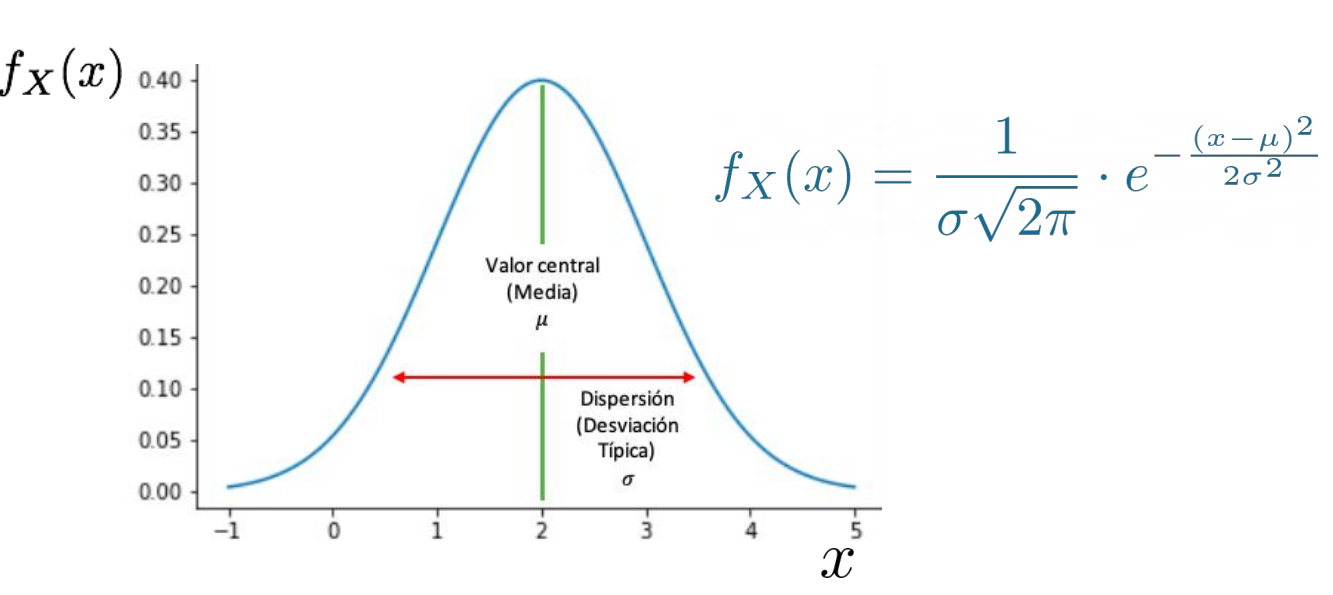
El valor esperado y la desviación estándar de una variable aleatoria que tiene una distribución normal se designan como y , respectivamente. Estos valores son los parámetros de una distribución normal.

| **Parámetros de la distribución normal** | |
| --- | --- |
|  |  |

De esta manera, una variable aleatoria continua con distribución normal de parámetros y toma valores en los números reales, y se anota:



La letra nos indica que la distribución es normal. La función de densidad de este tipo de variable tiene la siguiente forma:

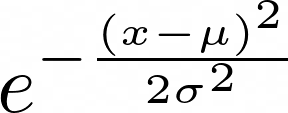
Figura 2: Forma de la función de densidad de   
una distribución normal.  


Las constantes y en la expresión para la función de densidad corresponden al número de Euler y a la razón entre el perímetro y el diámetro de un círculo, respectivamente:



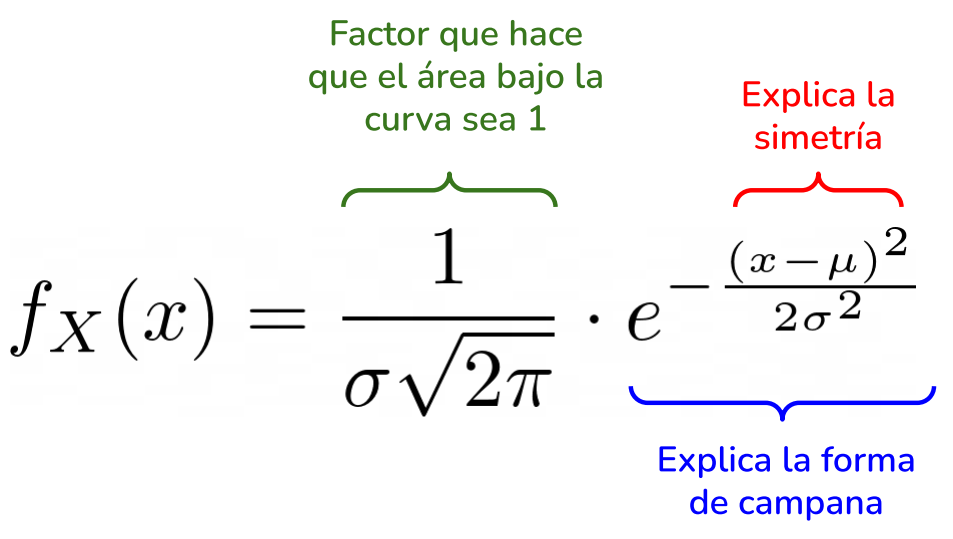
Si bien no vamos a entrar en detalles en la fórmula para esta función de densidad, podemos notar varias características de la curva de una distribución normal:

1. **Es simétrica respecto del valor esperado** . Esto se debe a que la fórmula depende del término , el cual es simétrico respecto de . Por ejemplo, si , se tiene que , ya que . Debido a esta simetría, el valor esperado se encuentra en el centro y además está exactamente en el punto más alto de la gráfica.
2. **Tiene forma de campana.** Esto se debe a la simetría la función de densidad y que la curva está dada por una función exponencial con coeficiente que es siempre negativo:



1. **El área bajo la curva es igual a 1.** Esto se explica por el factor , el cual se elige apropiadamente para que el área bajo la gráfica de la función de densidad sea exactamente 1.

En resumen, tenemos lo siguiente:

**INTERPRETACIÓN GRÁFICA DE LOS PARÁMETROS DE UNA DISTRIBUCIÓN NORMAL**

Como hemos mencionado anteriormente, las distribuciones que se asemejan a una normal aparecen habitualmente en muchos fenómenos. Sin embargo, estas distribuciones difieren en cuanto a los parámetros y . El valor esperado determina el eje de simetría de la función de densidad. Por tanto, al aumentar o disminuir el valor de , manteniendo fijo el valor de , la curva mantiene su forma y solo se desplaza en el eje horizontal. Por ejemplo, las siguientes tres variables aleatorias se distribuyen de forma normal, tienen la misma desviación estándar , pero distintos valores esperados:

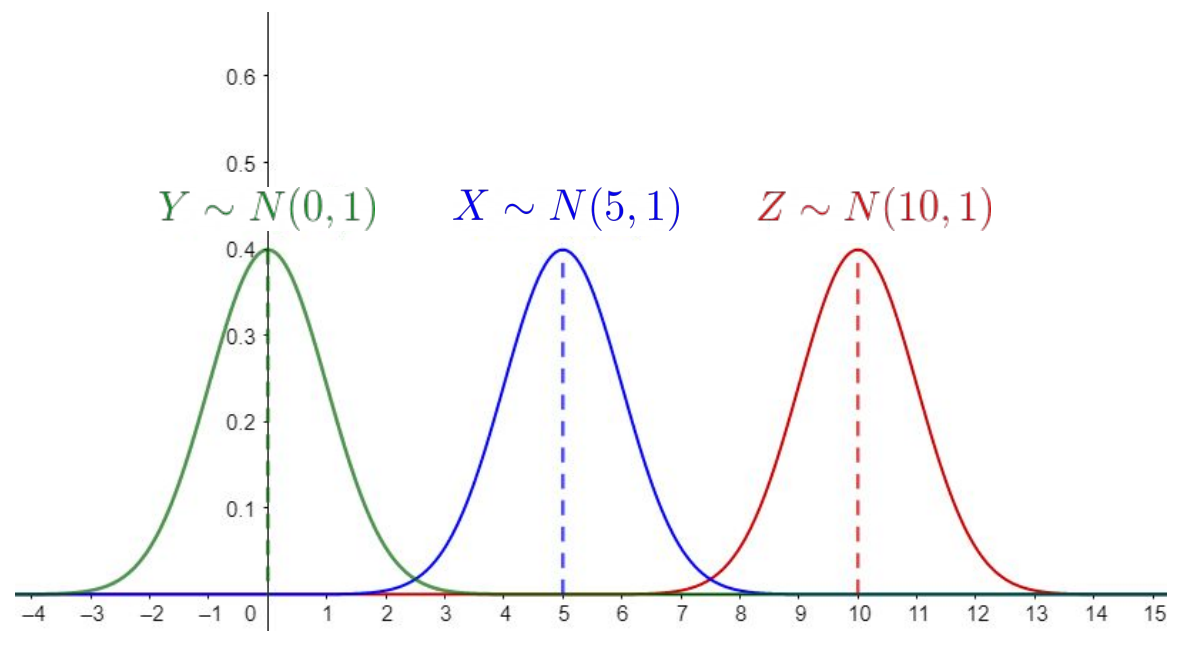
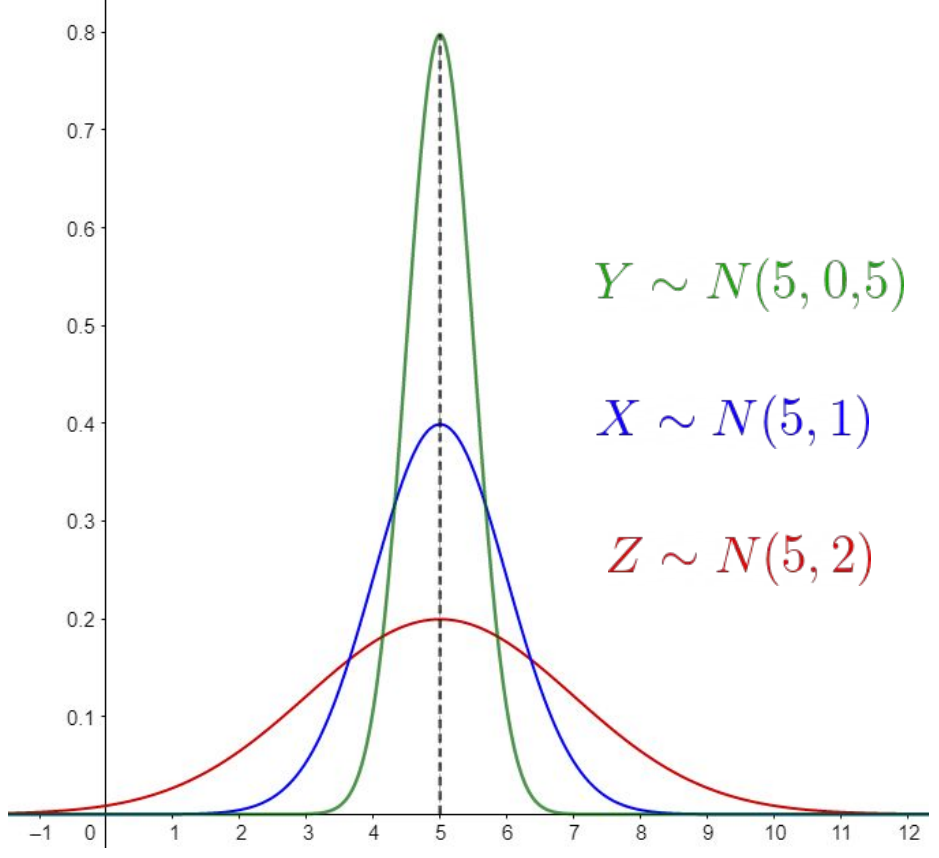


Figura 3: Distribuciones normales de distinta media   
pero con misma desviación estándar.

Por otro lado, la desviación estándar determina qué tan alta o plana es la curva de la función de densidad. Debido a la simetría de la distribución normal, mientras más grande sea el valor de , más se dispersan los datos en torno a la media y la curva se hace más plana. En cambio, cuando este parámetro se hace pequeño, los datos se concentran cerca del valor esperado de la distribución y la curva alcanza una mayor altura.

Por ejemplo, las siguientes tres variables aleatorias se distribuyen de forma normal y tienen el mismo valor esperado , por lo que poseen el mismo eje de simetría dado por la recta . Sin embargo, tienen distintas desviaciones estándar, lo que varía la amplitud de las curvas:

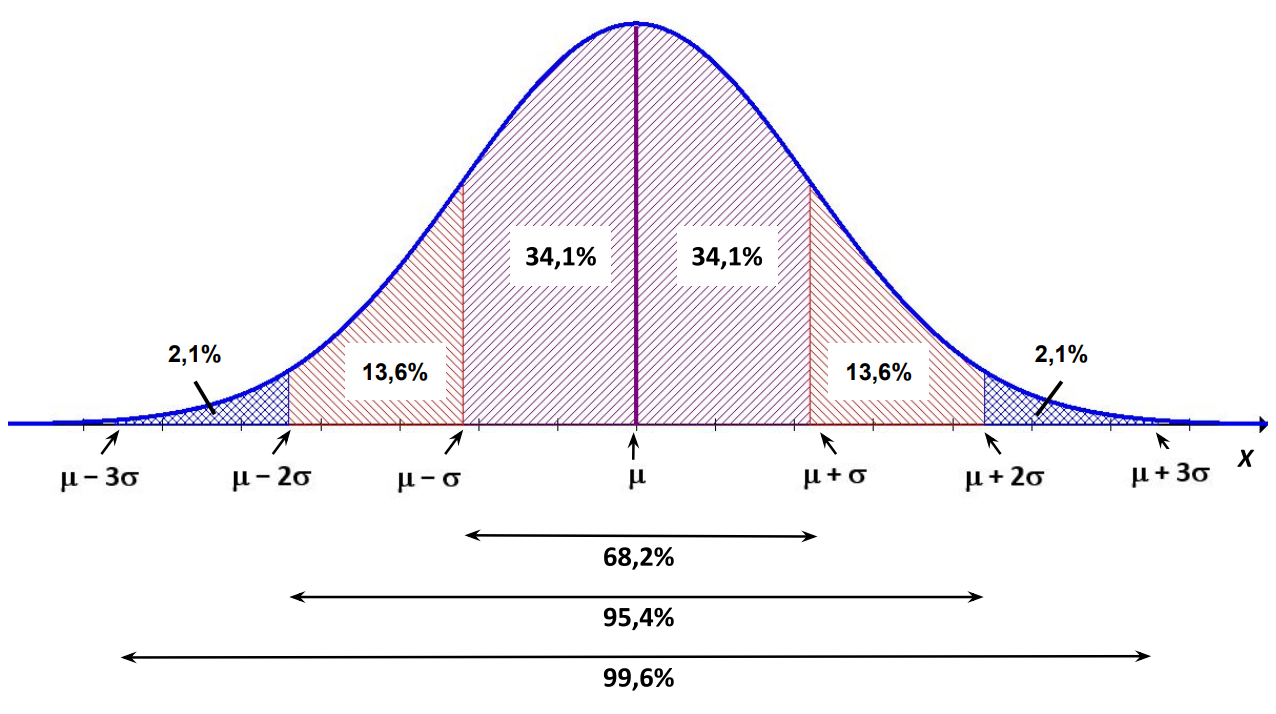
Figura 4: Distribuciones normales de misma media pero distinta desviación estándar.  


Las variables aleatorias continuas de muchos fenómenos naturales o sociales tienen una forma de campana simétrica y se pueden modelar por una distribución que se conoce como normal. Consiguientemente, una variable aleatoria continua que sigue una distribución normal con valor esperado y desviación estándar se anota:



Los valores y se conocen como parámetros de la distribución normal y determinan completamente su forma. Teniendo esto en cuenta, exploraremos una nueva propiedad, la cual utiliza estos parámetros.

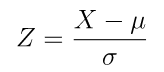
**REGLA EMPÍRICA**

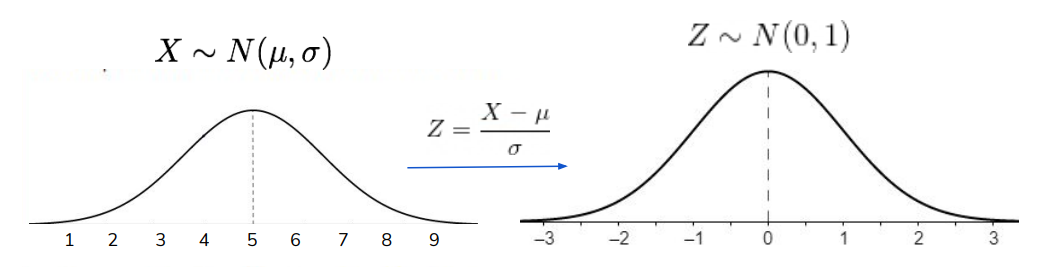
Considera una variable aleatoria continua que sigue una distribución normal con valor esperado y desviación estándar . La distribución de cumple la siguiente propiedad**:**

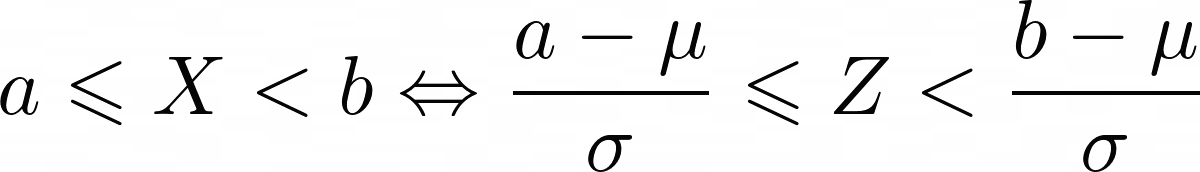
Esta propiedad, que se denomina usualmente “la regla empírica”, puede interpretarse de la siguiente manera:

* La probabilidad o el 68,2% de los valores de se encuentran en el intervalo .
* La probabilidad o el 95,4% de los valores de se encuentran en el intervalo .
* La probabilidad o el 99,6% de los valores de se encuentran en el intervalo .

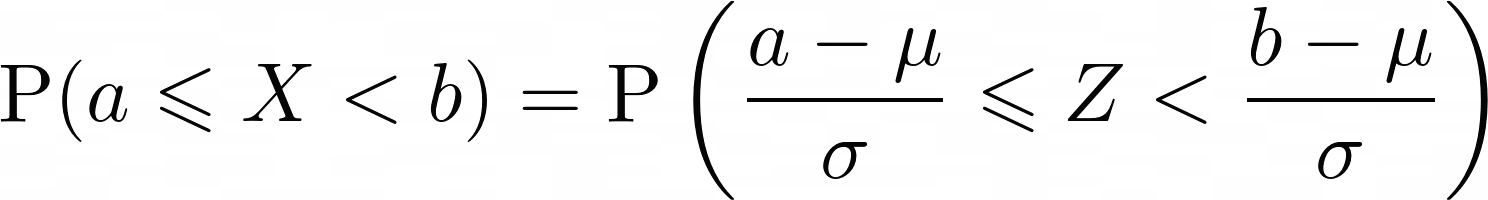
**DISTRIBUCIÓN NORMAL ESTÁNDAR**Para cada variable aleatoria continua que se distribuye como una normal con parámetros y , es decir, es posible definir una nueva variable aleatoria:



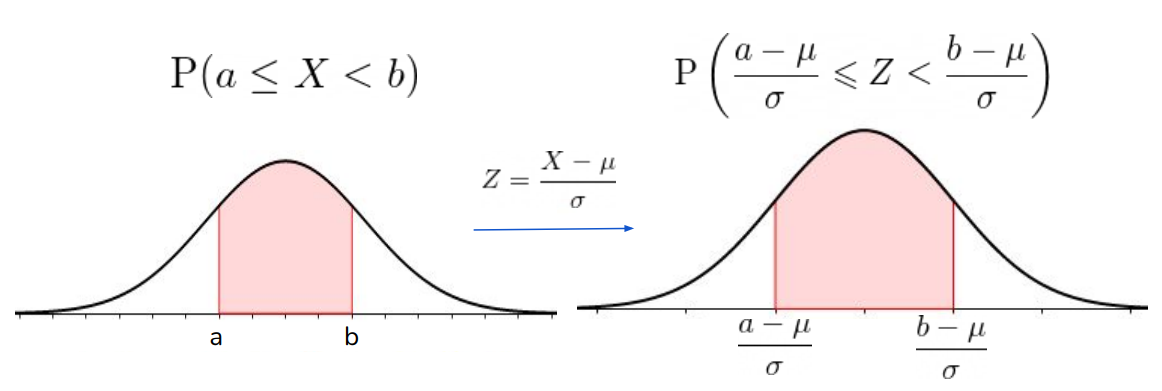
Esta nueva variable se denomina **distribución normal estándar** y tiene la propiedad que su media es y su desviación estándar es . Es decir, :  
  
La transformación que produce la nueva variable a partir de se conoce como **estandarización** y, como veremos, resulta útil para calcular probabilidades y además comparar distribuciones normales que tienen distintos parámetros.

Notemos que para cualesquiera valores y se tiene que: 

Y por tanto se tiene que:



Ilustrativamente, esto se aprecia de la siguiente manera:

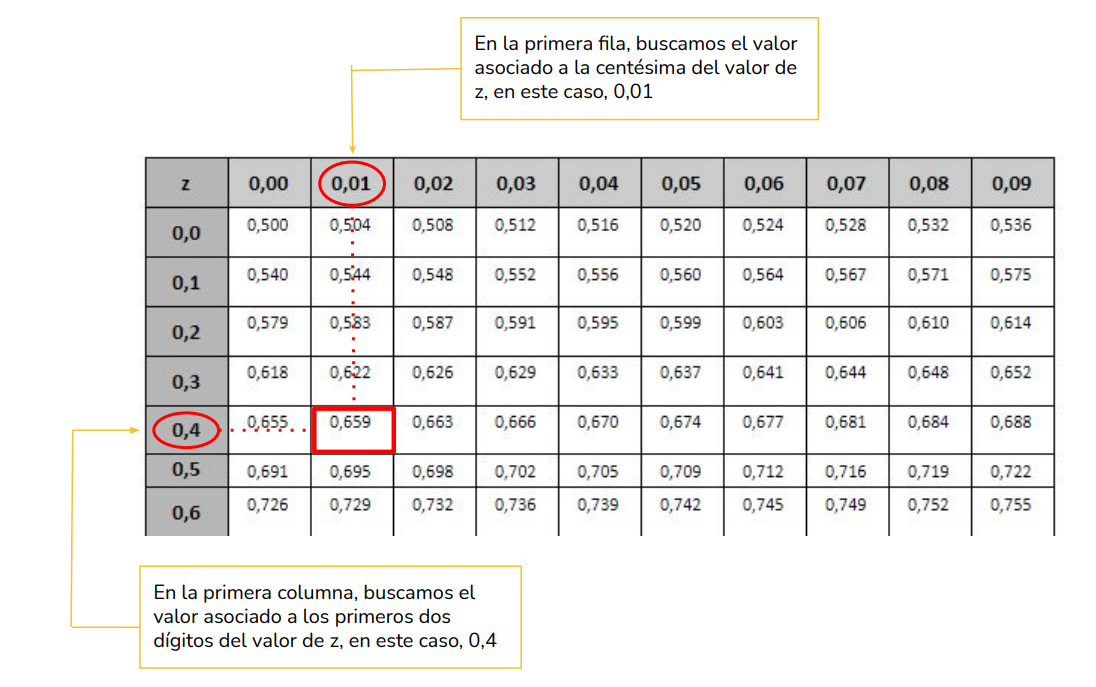


Lo anterior nos muestra que para calcular las probabilidades asociadas a la variable podemos usar la variable aleatoria estandarizada . Como lo que se obtiene es una variable aleatoria normal estándar, es común encontrarse con una tabla de probabilidades que permite encontrar la probabilidad . Esta tabla se encuentra al final de este documento.

Para aprender a utilizarla, estudiemos el siguiente ejemplo. Supongamos que la población chilena se caracteriza por tener una estatura media de 173 centímetros y desviación estándar de 7,3 centímetros. Si queremos calcular qué porcentaje de la población mide como máximo 180 centímetros, podemos estandarizar esa variable y encontrar el valor de la probabilidad asociada.  
Los valores obtenidos en estos cálculos se muestran en la siguiente tabla:

|  |  |
| --- | --- |
| 0,41 | 0,659 |

Dado que conocemos el valor de , lo buscamos en la tabla para obtener el valor de la probabilidad :

De esta forma, encontramos la probabilidad 0,659 asociada al valor , la cual coincide con la que calculamos antes.

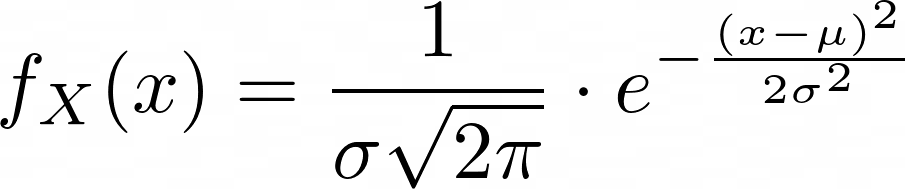
**SÍNTESIS**

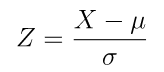
* Las variables aleatorias continuas de muchos fenómenos naturales o sociales tienen una forma de campana y se pueden modelar por una distribución que se conoce como normal.
* Una variable aleatoria continua que sigue una distribución normal con valor esperado y desviación estándar se anota:



Los valores y se conocen como parámetros de la distribución y determinan completamente su forma.

* La función de densidad de una variable aleatoria que sigue una distribución normal se puede expresar como:



* Consideremos la curva de la función de densidad de una variable aleatoria continua que sigue una distribución normal:
* La curva normal, al ser simétrica y con forma de campana, se encuentra centrada en su valor esperado . Para diferentes valores de la gráfica de la función se desplaza a lo largo del eje horizontal..
* La desviación estándar determina que tan alta o plana es la curva. Mientras más grande sea el valor de , más se dispersan los datos en torno a la media y la curva será más plana. En cambio, cuando este parámetro se hace pequeño, los datos se concentran cerca del valor esperado de la distribución y la curva alcanza una mayor altura.
* La distribución normal con parámetros y , cumple la denominada regla empírica.
* A partir de esta propiedad, una variable aleatoria que sigue una distribución normal con parámetros y cumple que:
  + La probabilidad de que corresponde al 68,2% de los valores de .
  + La probabilidad de que corresponde al 95,4% de los valores de .
  + El 99,6% de los valores de la variable se encuentran a menos dedel valor esperado.
* Para cada variable aleatoria continua que se distribuye como una normal con parámetros y , es posible definir una nueva variable aleatoria:   
  

Esta nueva variable se denomina **distribución normal estándar** y tiene la propiedad que su media es y su desviación estándar es . Es decir, .

* La transformación que produce la nueva variable \(Z\) a partir de \(X\) se conoce como **estandarización.**
* Para la variable aleatoria existe una **tabla de probabilidades** que puede ser usada para calcular la probabilidad :   
  