UNIDAD 4: DISTRIBUCIONES DE PROBABILIDAD PARTE 3

DISTRIBUCIONES CONTINUAS DE PROBABILIDAD

Están asociadas a variables aleatorias continuas. <u>Ejemplo</u>: Alturas, pesos, tiempo de vida útil, etc.

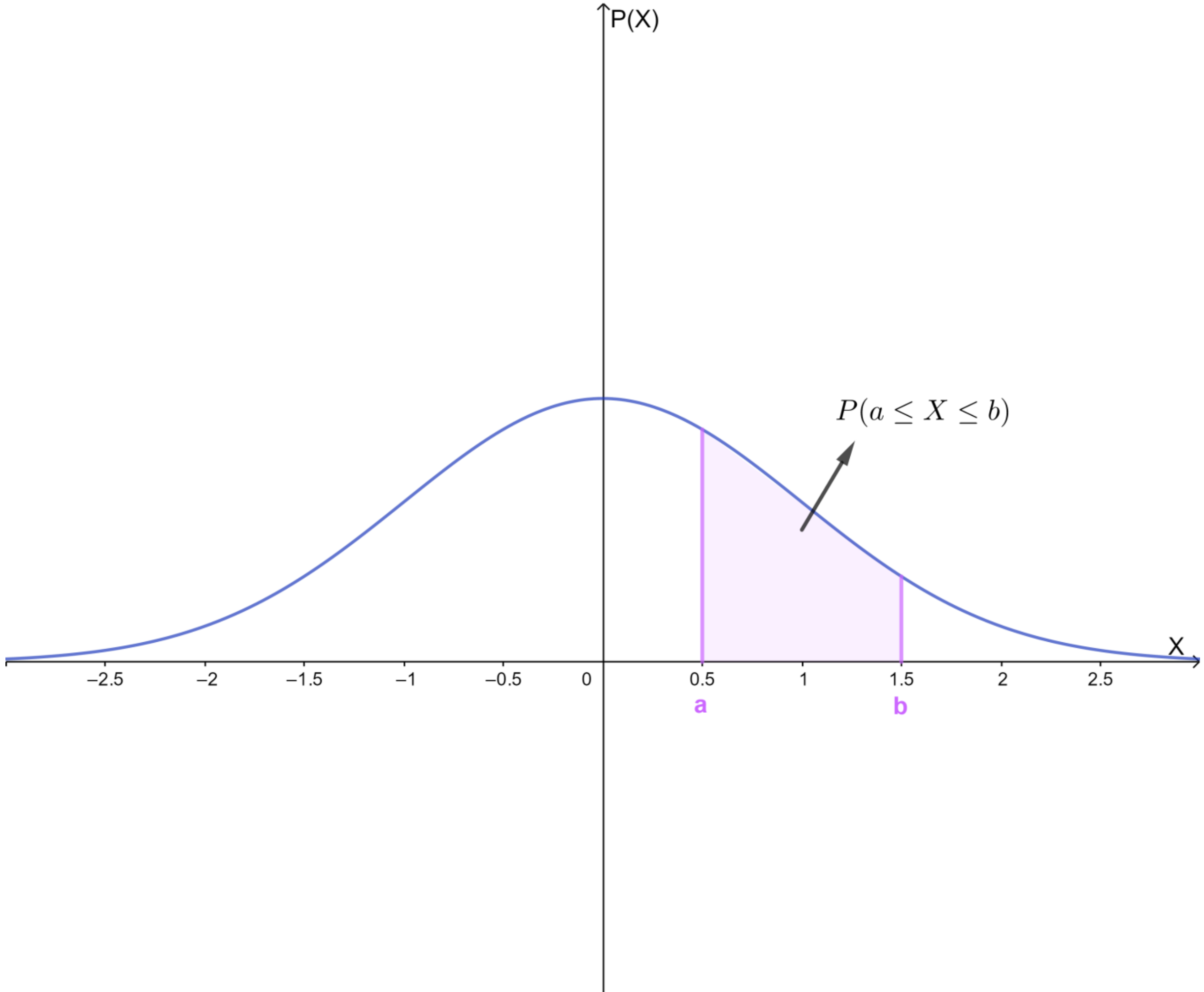
Una **curva de densidad** describe la distribución de probabilidad de una variable continua. Es la aproximación "suavizada" de un polígono de frecuencia.

La probabilidad se puede describir por una fórmula **f(X) o P(X)** llamada **distribución de probabilidad** o **función de densidad** de la variable aleatoria X.

Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

PROPIEDADES DE LAS DISTRIBUCIONES CONTINUAS DE PROBABILIDAD

- El área bajo la curva es igual a 1, ya que representa la suma de todas las probabilidades.
- $P(a \leq X \leq b)$ = **área bajo la curva entre a y b**.



Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

PROPIEDADES DE LAS DISTRIBUCIONES CONTINUAS DE PROBABILIDAD

- La probabilidad para cualquier valor particular **a** de las variables es P(X = a) = 0.
- Una de las variables aleatorias continuas más importante es la **variable aleatoria normal**.

Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

LA DISTRIBUCIÓN NORMAL

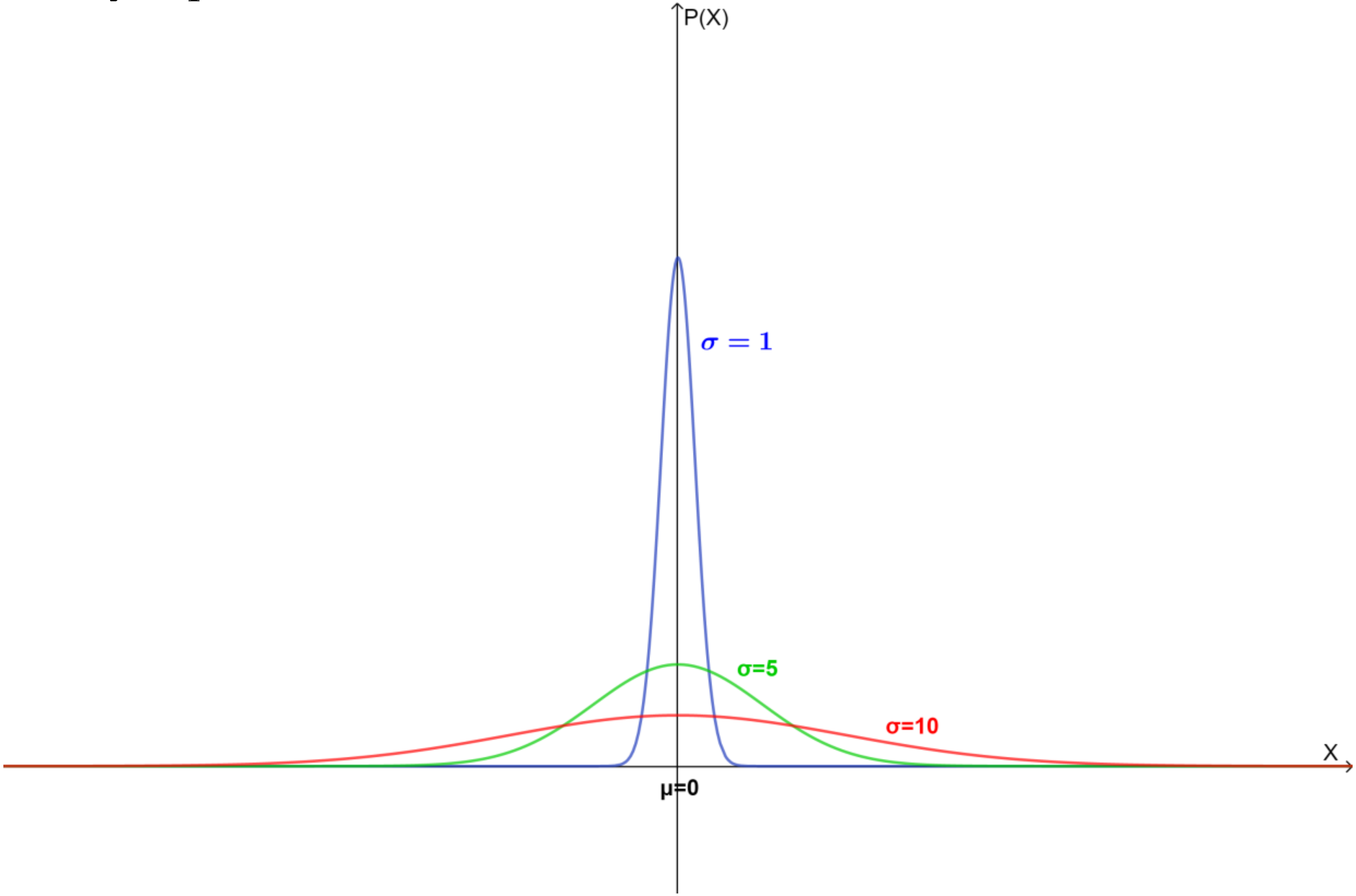
La fórmula que genera la distribución de probabilidad normal es:

$$f(x) = \frac{1}{\sigma.\sqrt{2.\pi}}.e^{-\frac{1}{2}.\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty, \; e \cong 2{,}7183 \text{ y } \pi \cong 3{,}1416$$

μ y σ son la media y desviación poblacional.

LA DISTRIBUCIÓN NORMAL

La forma y centro de la curva de densidad cambia de acuerdo a los valores de la media y de la desviación estándar. Por ejemplo:

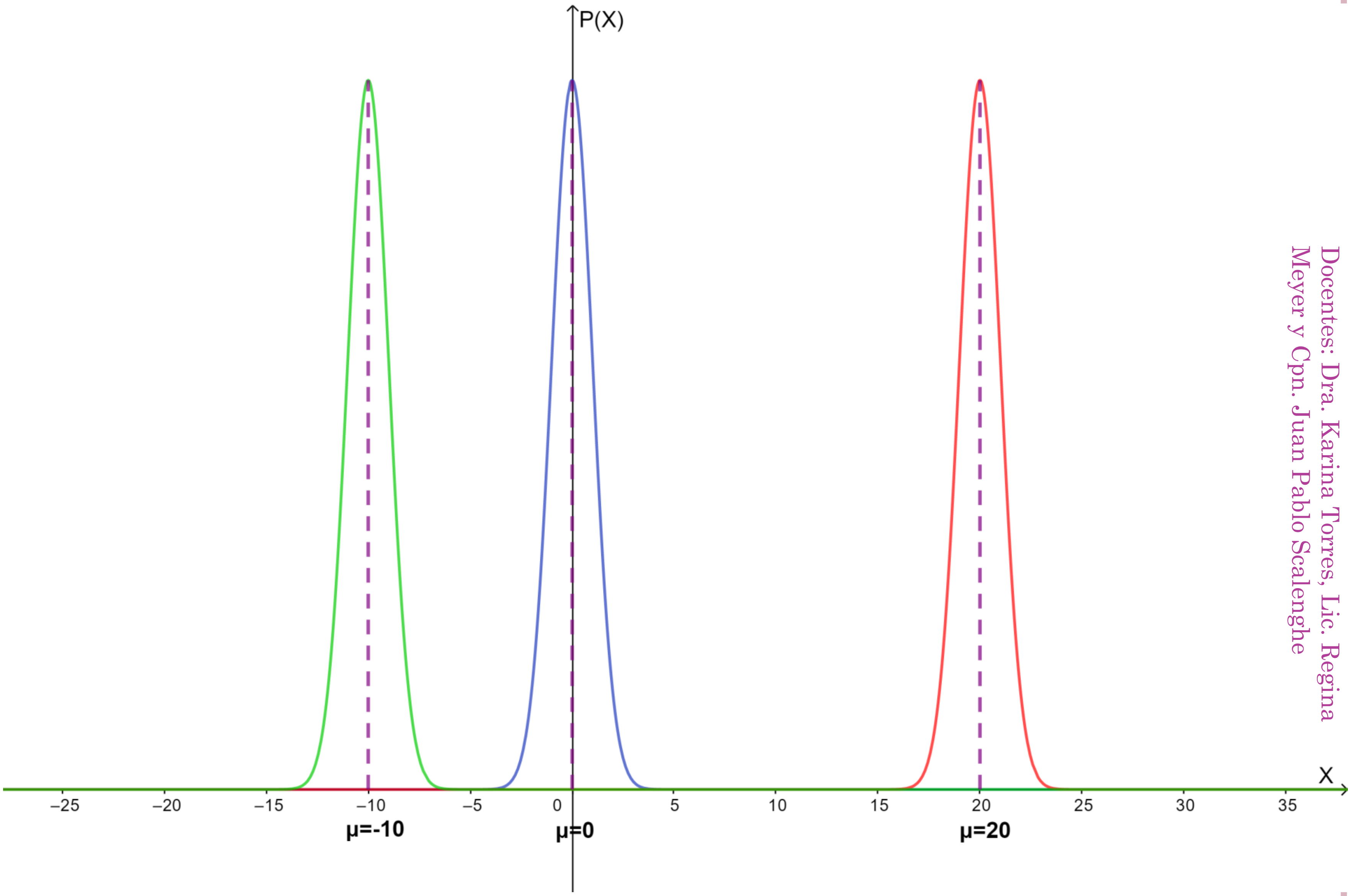


Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

LA DISTRIBUCIÓN NORMAL

Los valores distintos de μ trasladan la gráfica de la distribución a lo largo del eje X y los diferentes valores de σ determinan el grado de aplanamiento o levantamiento de la gráfica.

LA DISTRIBUCIÓN NORMAL



Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

LA DISTRIBUCIÓN NORMAL

La media de la distribución puede ser cualquier valor numérico negativo, cero o positivo.

La gráfica muestra 3 curvas normales con la misma desviación, pero con tres medias distintas.

PROPIEDADES DE LA DISTRIBUCIÓN NORMAL

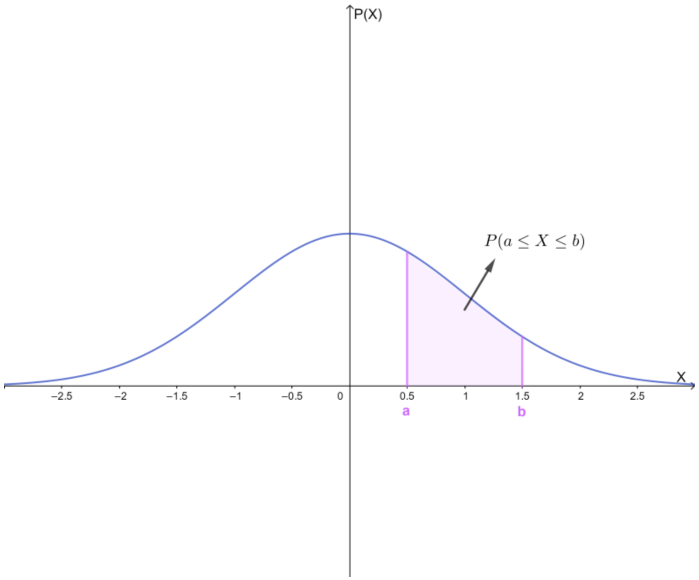
- La curva normal es **unimodal** y **simétrica** con respecto a su media μ, por lo tanto, $\overline{X} = Me = Mo$.
- La desviación estándar, representada por σ, especifica el grado de dispersión respecto de la media.
- Existe toda una familia de distribuciones normales, cada una con su correspondiente μ y σ.
- En la **distribución normal estándar μ = 0 y σ = 1**.

Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

CÁLCULO DE PROBABILIDADES

Habíamos dicho que:

$$P(a \leq X \leq b) \rightarrow$$



Este valor debe buscarse en la tabla de la distribución normal, pero para esto se debe estandarizar cada valor de X a través de la fórmula:

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

Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

CARACTERÍSTICAS DE LA DISTRIBUCIÓN NORMAL ESTÁNDAR

- $\mu = 0$ y $\sigma = 1$.

- Es simétrica alrededor de $z = 0$.

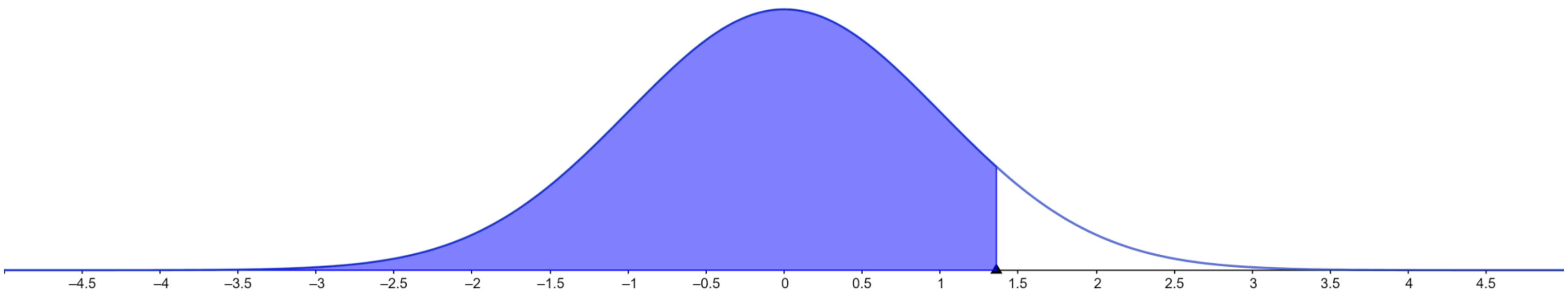
- El área total bajo la curva es 1.

Uso de la tabla: Los cuatro dígitos en una fila y columna particulares de la tabla de la distribución normal estándar proporciona el área bajo la curva en un determinado intervalo mayor o menor que un valor particular z.

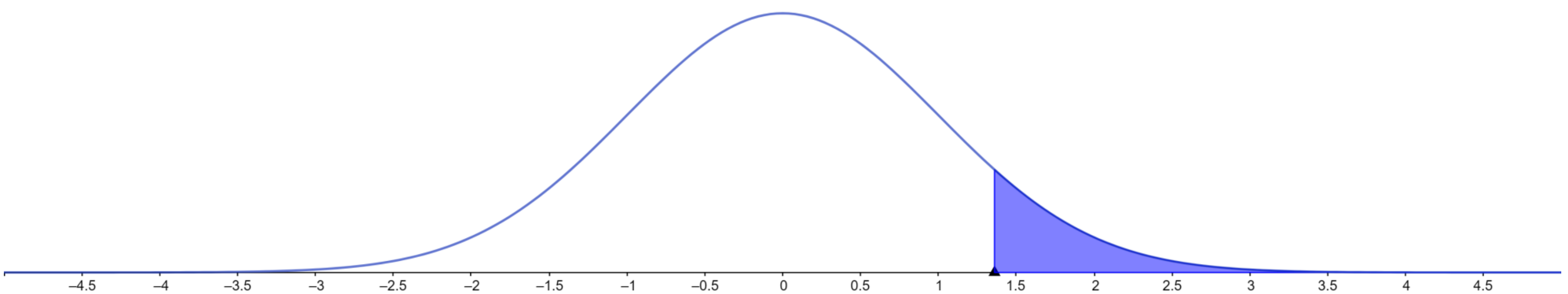
Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

EJEMPLOS DE DISTRIBUCIÓN NORMAL ESTÁNDAR

- $P(Z < 1{,}36) = 0{,}9131$



- $P(Z > 1{,}36) = 1 – P(Z \leq 1{,}36) = 1 – 0{,}9131 = 0{,}0869$

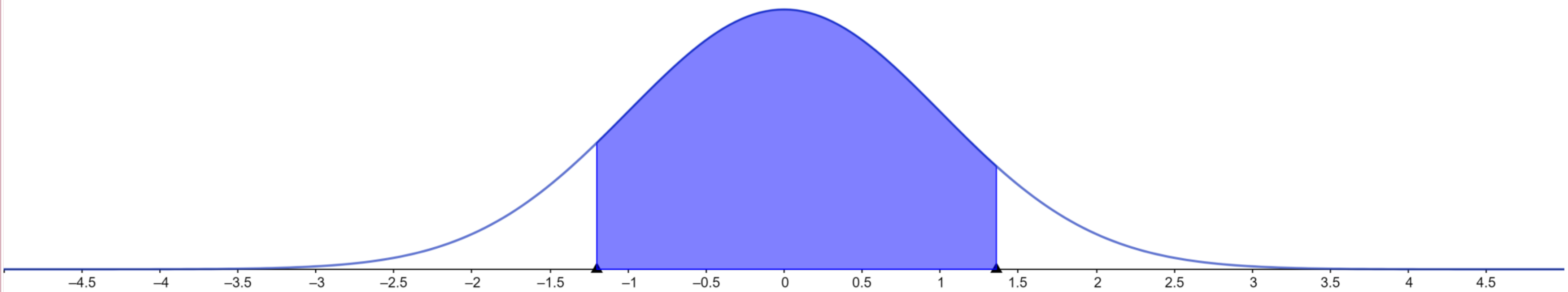


Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

EJEMPLOS DE DISTRIBUCIÓN NORMAL ESTÁNDAR

- $P(-1{,}20 \leq Z \leq 1{,}36) = P(Z \leq 1{,}36) - P(Z \leq -1{,}20)$

$= 0{,}9131 - 0{,}1151 = 0{,}7950.$

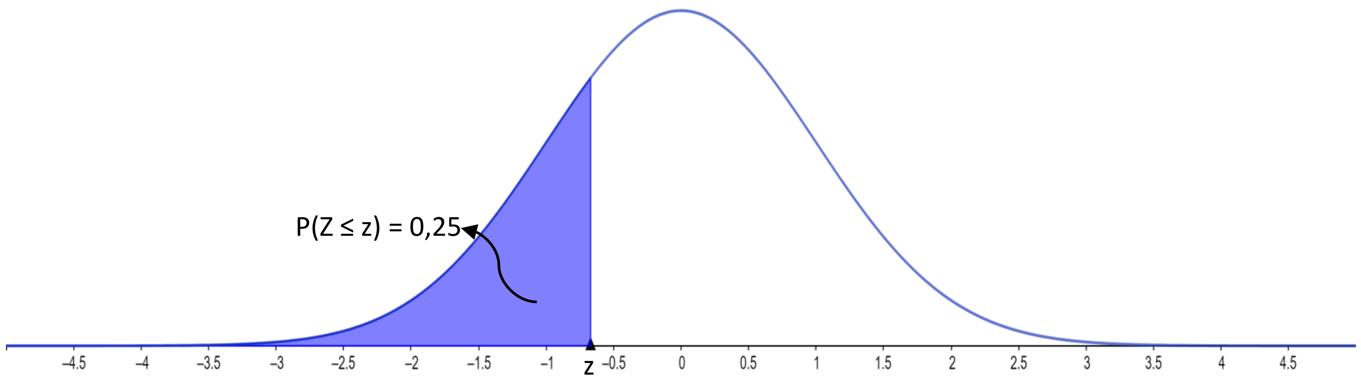


Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

EJEMPLOS DE DISTRIBUCIÓN NORMAL ESTÁNDAR

- Se sabe que la probabilidad de encontrar un valor menor o igual que Z es de 0,25. ¿Cuánto vale z?

$P(Z \leq z) = 0,25$



Primero, deberíamos buscar el valor de la tabla que más se aproxima a 0,25 (dentro de la tabla).

Luego, vemos el valor de z correspondiente.

<u>Solución</u>: **$z = -0,67$**

Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

CÁLCULO DE PROBABILIDAD PARA UNA DISTRIBUCIÓN NORMAL NO ESTÁNDAR

Si μ ≠ 0 y σ ≠ 1 se debe proceder estandarizando de la siguiente manera:

EJEMPLO: X ~ N(5,2), μ = 5 y σ = 2.

Calcular P(X > 7):

$$P(X > 7) = P\left(\frac{X-\mu}{\sigma} > \frac{7-5}{2}\right) = P\left(Z > \frac{2}{2}\right) = P(Z > 1) = 1 - P(Z \leq 1)$$

$$= 1 - 0{,}8413 = 0{,}1587$$

Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

PROBLEMAS CON DISTRIBUCIÓN NORMAL

1) En los estudios dactiloscópicos, una importante característica cuantitativa es el número total de los surcos en los dedos de un individuo. Supóngase que los números totales de surcos de los individuos de cierta población están distribuidos en forma normal con $\mu = 140$ y $\sigma = 50$.

   Encontrar la probabilidad de que un individuo seleccionado al azar dentro de la población tenga un número de surcos:

   a) de 200 o más.
   b) menor de 100.
   c) entre 100 y 200.
   d) entre 200 y 250.

Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

PROBLEMAS CON DISTRIBUCIÓN NORMAL

RESPUESTAS:

Variable: X: "Número total de surcos."

a) Planteamos lo siguiente:

$$P(X \geq 200) = P\left(Z \geq \frac{200-140}{50} \right) = P\left(Z \geq \frac{60}{50} \right) = P\left(Z \geq \frac{6}{5} \right) = P(Z \geq 1{,}2) =$$

$$= 1 - P(Z < 1{,}2) = 1 - 0{,}8849 = 0{,}1151$$

b) Tenemos lo siguiente:

$$P(X < 100) = P\left(Z < \frac{100-140}{50} \right) = P(Z < -0{,}8) = 0{,}2119$$

c) Planteamos:

$$P(100 < X < 200) = P(-0{,}8 < Z < 1{,}2) = 0{,}8849 - 0{,}2119 = 0{,}673$$

d) Por último, tenemos que:

$$P(200 < X < 250) = P(1{,}2 < Z < 2{,}2) = 0{,}9861 - 0{,}8849 = 0{,}1012$$

Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

PROBLEMAS CON DISTRIBUCIÓN NORMAL

2) El nivel de estrés en personas de 30 0 40 años de edad está normalmente distribuido con $\mu = 1$ punto y $\sigma = 0,10$ puntos.

a) ¿Qué probabilidad hay de seleccionar al azar una persona de la población que tenga un nivel de estrés comprendido entre 0,80 y 0,85?

$$P(0,80 \leq X \leq 0,85) = P(0,80 < X < 0,85) = P\left(\frac{0,80-1}{0,10} < Z < \frac{0,85-1}{0,10}\right) =$$

$$= P(-2 < Z < -1,5) = P(Z < -1,5) - P(Z < -2) = 0,0668 - 0,0228 = 0,044$$

b) ¿Cuál es el nivel mínimo de estrés que puede tener una persona que esté dentro del 1% de individuos con niveles muy altos de estrés?

$$P(X > x) = 0,01 \Rightarrow P\left(Z > \frac{x-1}{0,1}\right) = 0,01 \Rightarrow 1 - P\left(Z \leq \frac{x-1}{0,1}\right) = 0,01$$

$$\Rightarrow P\left(Z \leq \frac{x-1}{0,1}\right) = 0,99 \Rightarrow \frac{x-1}{0,1} = 2,33 \Rightarrow x = 2,33.0,1 + 1 \Rightarrow x = 1,233$$

Por lo tanto, el nivel mínimo de estrés correspondiente al grupo que tiene niveles altos de estrés es 1,233 puntos (dicho valor corresponde al P_{99})

Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

APROXIMACIÓN DE DISTRIBUCIONES EMPÍRICAS A LA NORMAL

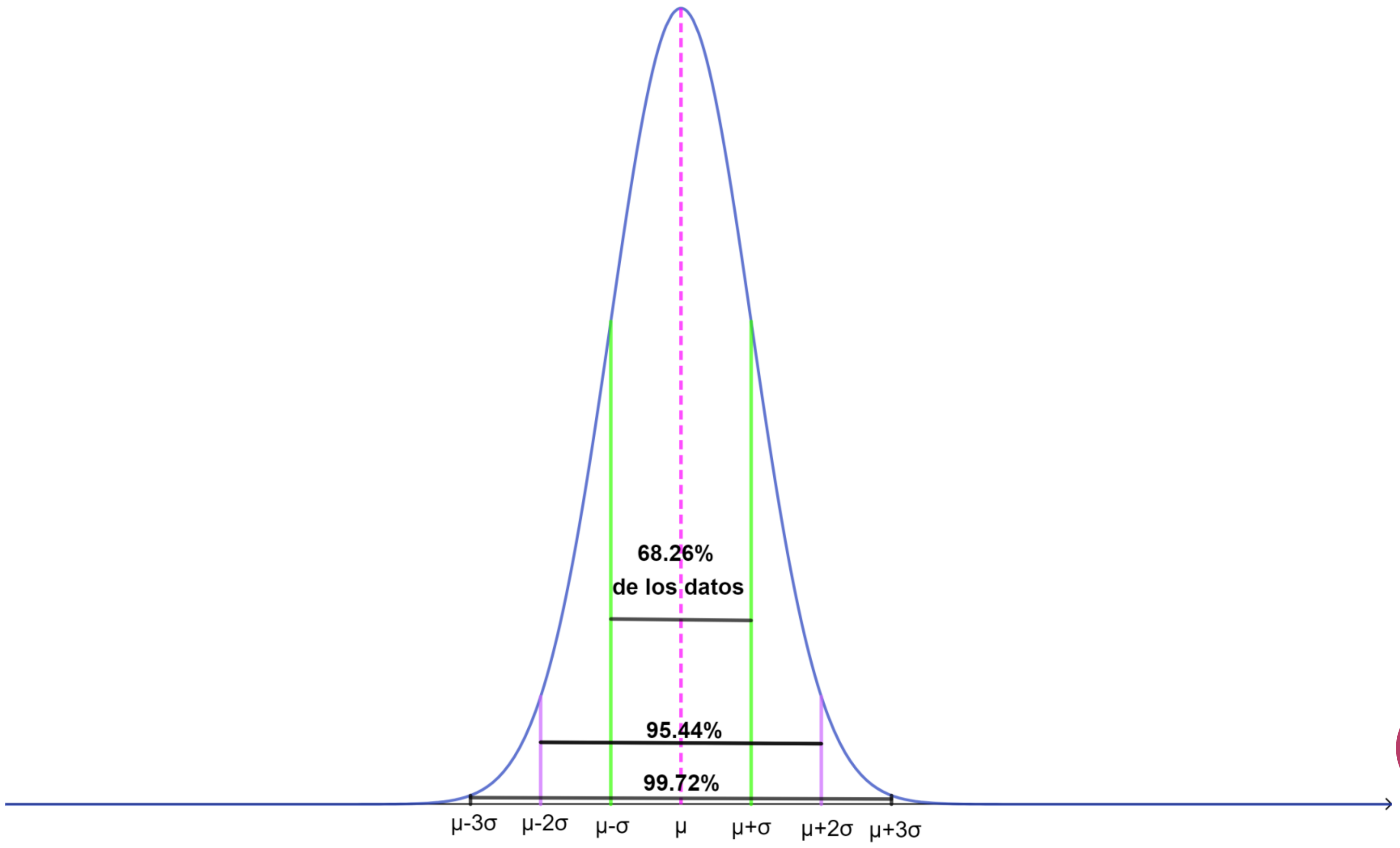
En muchas situaciones, no podemos determinar a priori si la variable que se observa tiene o no una forma normal. En esta situación, podemos buscar relaciones entre las distribuciones normal y la distribución de frecuencias empíricas y observar si la distribución empírica se aproxima a la normal.

Una forma de hacer esto (para variables cuantitativas continuas o discretas) es aplicando la regla empírica y observando si la distribución de frecuencia es unimodal y simétrica.

Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

APROXIMACIÓN DE DISTRIBUCIONES EMPÍRICAS A LA NORMAL

REGLA EMPÍRICA



Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe

APROXIMACIÓN DE DISTRIBUCIONES EMPÍRICAS A LA NORMAL

Esta regla establece que para datos con distribución en forma de campana:

1) El 68% de los datos están a menos de una desviación estándar de la media. Este es el valor del área en esa región.
2) El 95% de los datos están a menos de dos desviaciones estándar de la media.
3) Casi todos los elementos están a menos de tres desviaciones estándar de la media.

Docentes: Dra. Karina Torres, Lic. Regina Meyer y Cpn. Juan Pablo Scalenghe