LECTURA EN CASA

DISTRIBUCIÓN NORMAL ESTANDARIZADA

La distribución normal se define por la desviación estándar y media, de un conjunto de datos cuantitativos / numéricos.

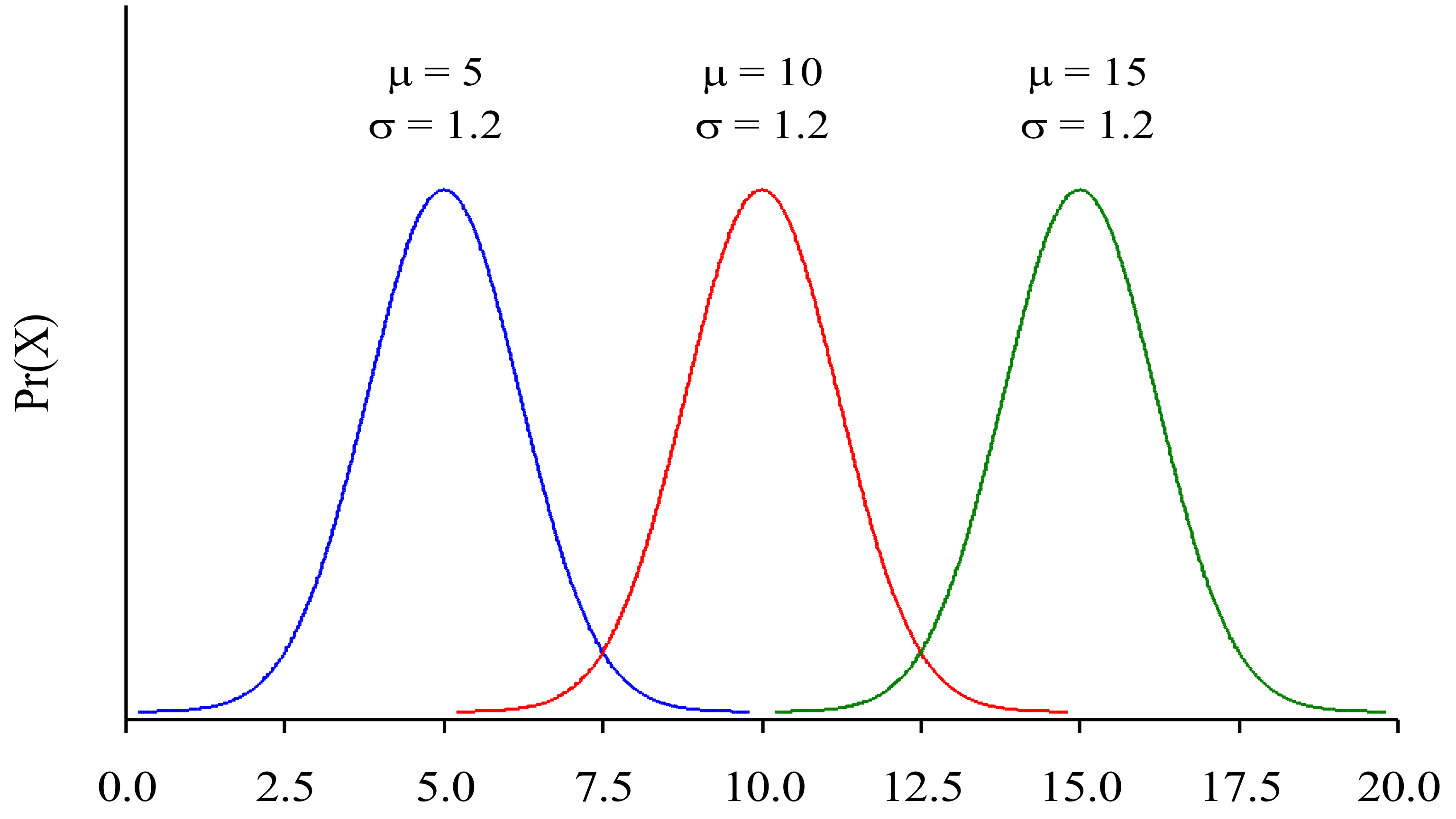
La media determina la ubicación de la curva en el eje x de un gráfico y la desviación estándar determina la altura de la curva en el eje y.

Si cambia la media, la distribución se desplaza en su eje x; si cambia la desviación estándar, cambia la extensión de la distribución.

Es importante recordar que hay un número infinito de distribuciones normales, una por cada combinación posible de media y desviación estándar.

Estos son ejemplos de tres distribuciones normales. Cada distribución tiene una media diferente, pero la misma desviación estándar.

Por lo tanto, los gráficos se desplazan a diferentes lugares en el eje x, debido a las diferentes medias, pero las formas son idénticas porque la desviación estándar es la misma.



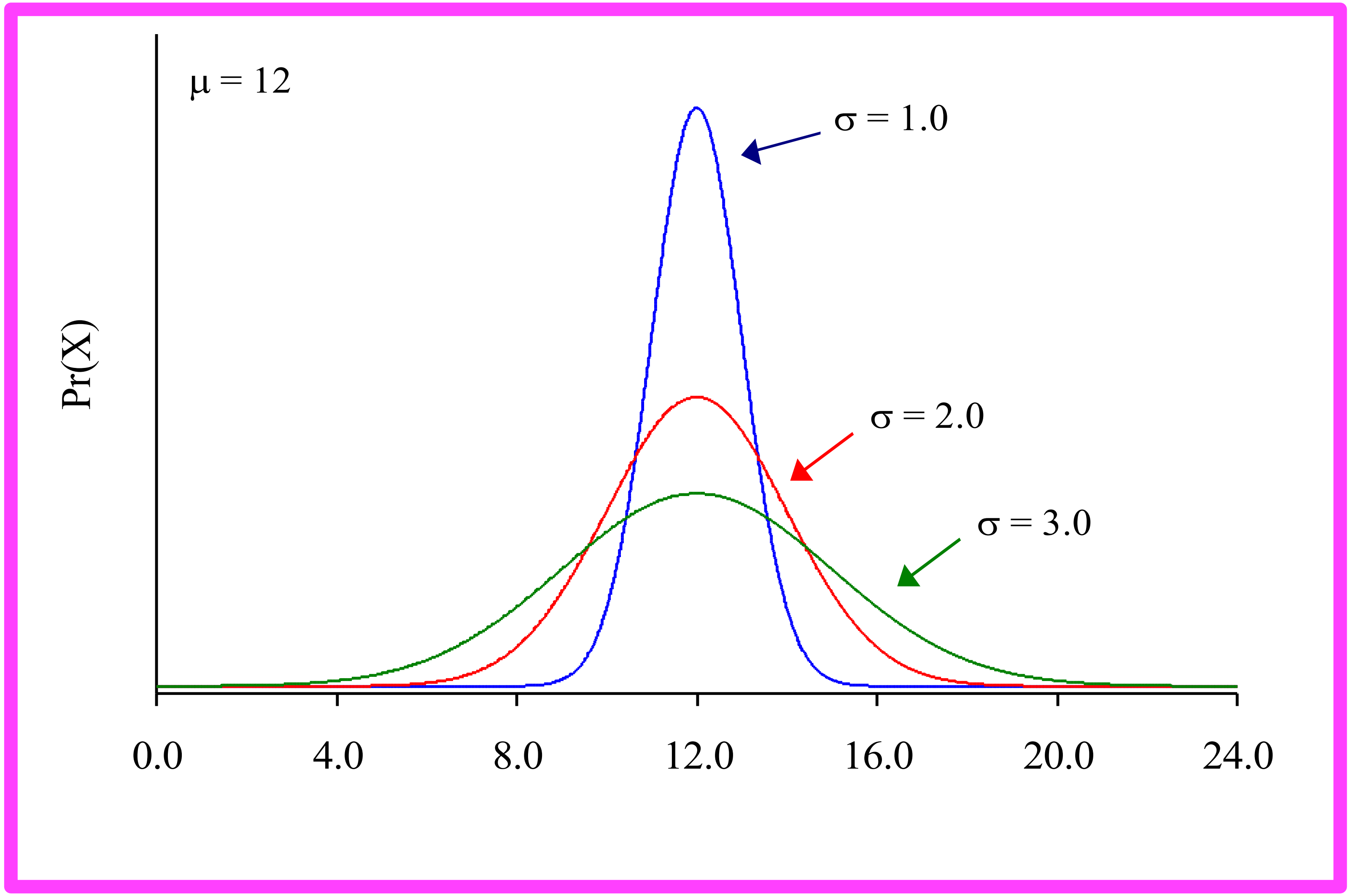
*Pr (X) en el eje y se refiere a frecuencia o probabilidad.

Estas tres distribuciones normales tienen todas la misma media, pero diferentes desviaciones estándar.

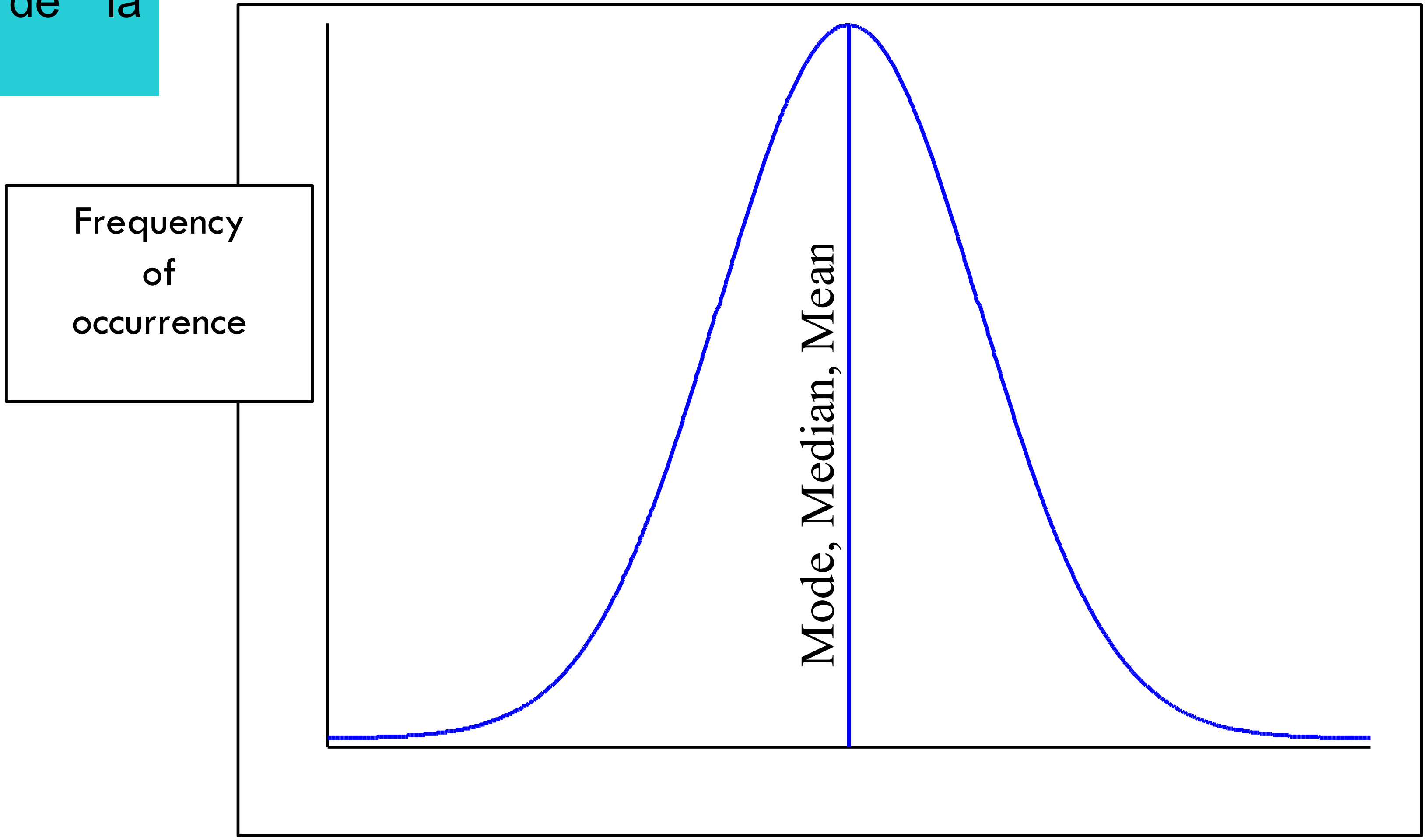
Los cambios en la desviación estándar dan como resultado cambios en la forma de la distribución, sin afectar el punto medio.

Una desviación estándar más pequeña da como resultado una curva más estrecha y puntiaguda.

Una desviación estándar mayor da como resultado una curva más ancha y plana.



Cuando los datos se distribuyen normalmente, la moda, la mediana y la media son idénticas y se encuentran en el centro de la distribución.

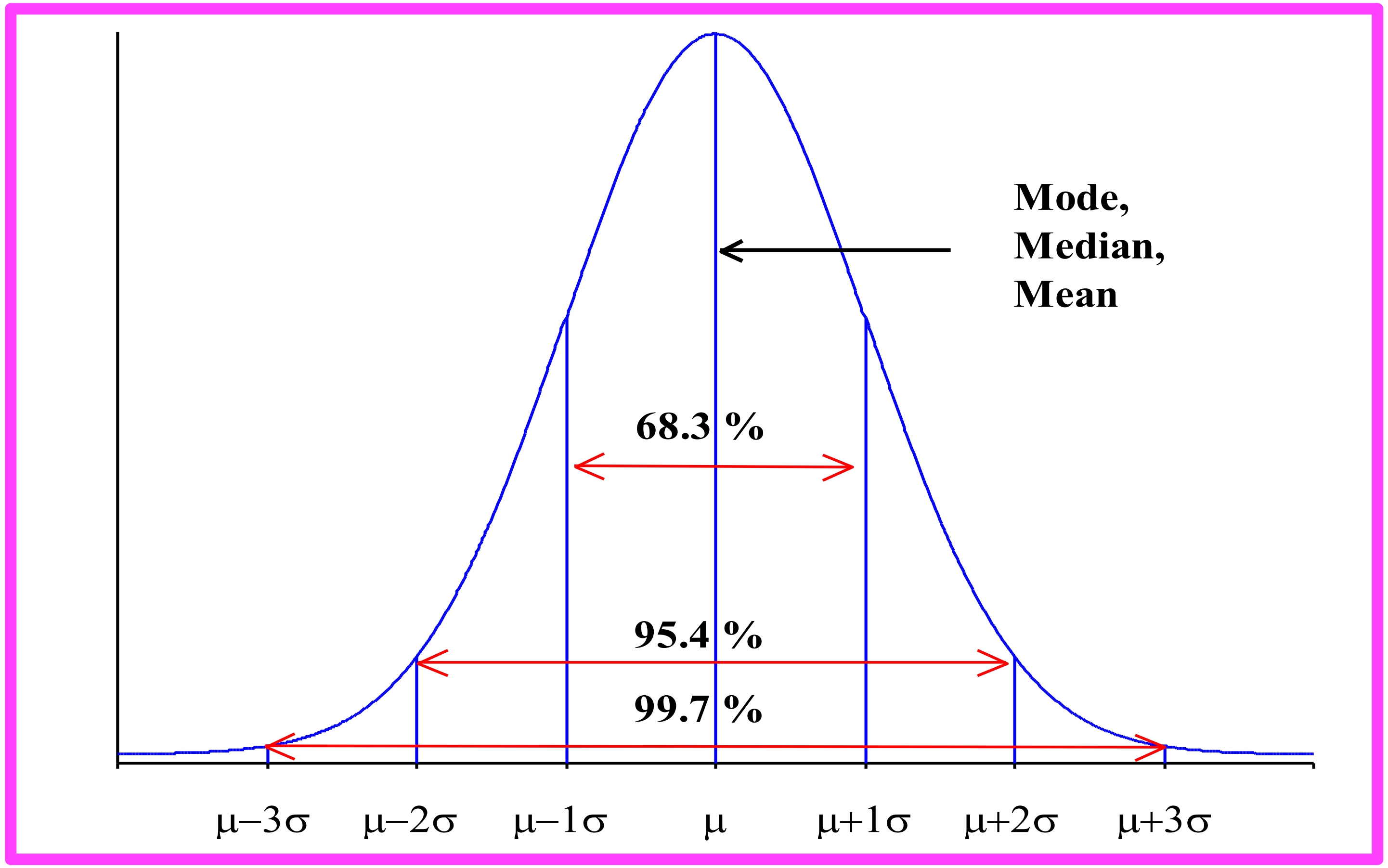


Cuando los datos se distribuyen normalmente con la moda, la mediana y la media en el centro de la curva

El 68,3% de las observaciones se encuentran entre la media y una desviación estándar a cada lado de la media, señaladas como + o - 1 desviación estándar.

El 95,4% de las observaciones se encuentran entre la media y + o - 2 desviaciones estándar

El 99,7% de las observaciones se encuentran entre la media y + o - 3 desviaciones estándar.



PERO...

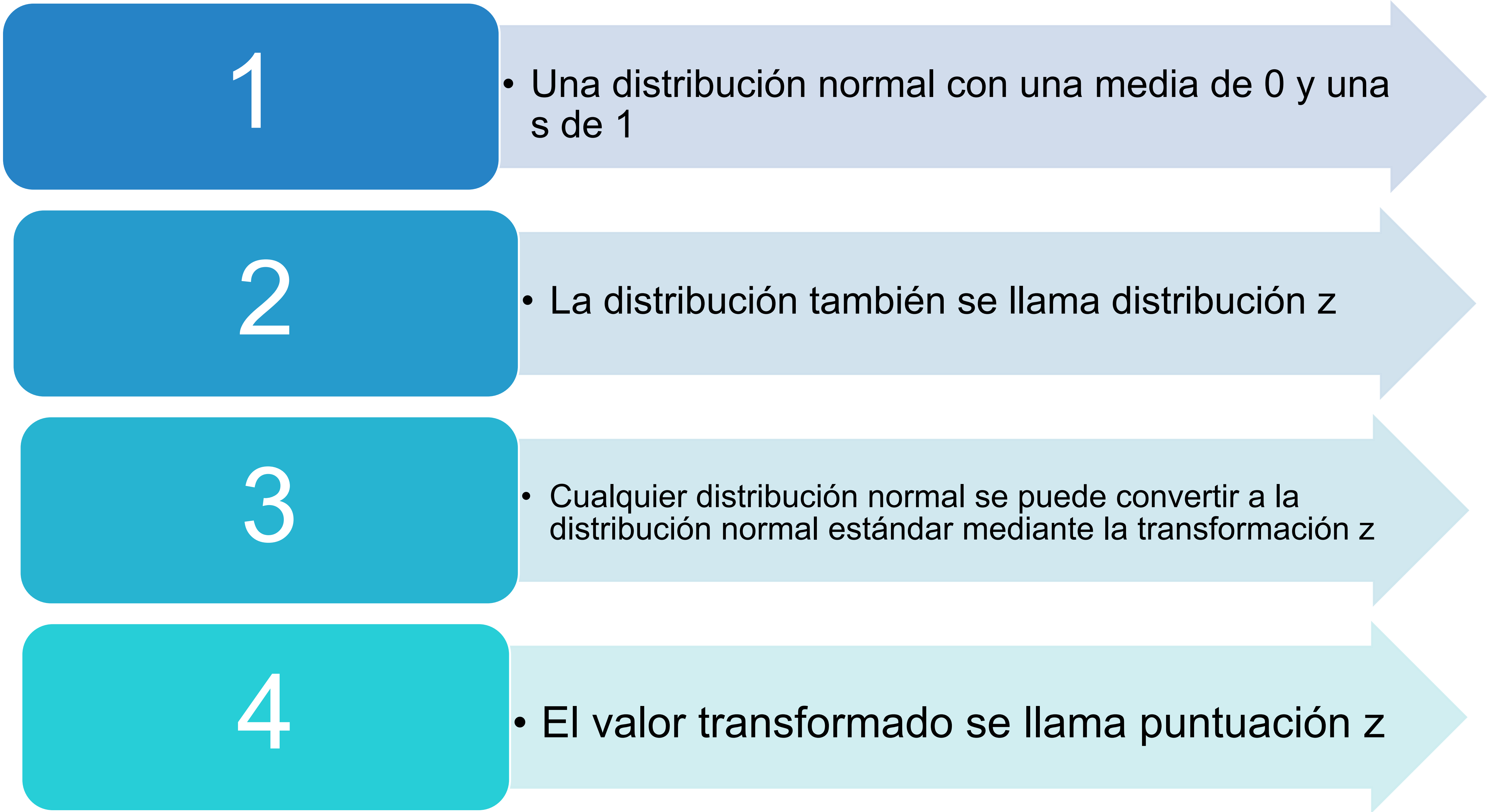
Existen miles de probabilidades para armar una distribución normal:

Media 7 – Desviación estándar de 1,5
Media 3 – Desviacón estándar de 2,8

Esto equivale a miles tablas para encontrar la probablilidad. Dado que eso no es práctico se estandariza la distribución con la tipificación de la variable aleatoria X a Z la cual sigue una distribución normal de media 0 y desviación estándar 1. Por lo tanto se utiliza una tabla estandarizada, la tabla Z.

Para encontrar las probabilidades o cantidad de datos entre determinados valores de la variable, se calcula el área bajo la curva normal, que se encuentra en la tabla z.

La prueba Z más simple es la prueba Z de una muestra, la cual evalúa la media de una población normalmente distribuida con varianza conocida.


1
• Una distribución normal con una media de 0 y una s de 1
2
• La distribución también se llama distribución z
3
• Cualquier distribución normal se puede convertir a la distribución normal estándar mediante la transformación z
4
• El valor transformado se llama puntuación z

La fórmula para transformar datos en puntuaciones z es X (el valor) - mu (la media de la población) dividida para sigma, la desviación estándar de la población. Las puntuaciones z transformadas se pueden utilizar para determinar las áreas bajo la curva para cualquier distribución normal.

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

La distribución normal surge de datos esenciales de la Media y Desviación Estándar.

En la distribución normal estándar los valores se asocian a PROBABILIDADES. Aquí el valor de la media se tranforma en 0 es decir separa en dos partes iguales a la distribución.

Las probabilidades en donde se encuentra la curva provienen de una integral.

Esos son los porcentajes de la curva para fucionarlas se usa la Tipificación = ajustar a lo que tienen en común las distribuciones a través de z. Esta ecuación vincula las dos.

EJERCICIOS DE REPASO - Z

1. Duración de un embarazo humano

La duración (en días) de un embarazo humano elegido al azar es una variable aleatoria normal con media (μ) = 266 y desviación estándar (σ) = 16 días.

(a) ¿Cuál es la probabilidad de que un embarazo elegido al azar dure menos de 246 días?

(b) ¿Cuál es la probabilidad de que un embarazo elegido al azar dure más de 240 días?

(c) ¿Cuál es la probabilidad de que un embarazo elegido al azar dure más de 500 días?

(d) Supongamos que el esposo de una mujer embarazada ha programado sus viajes de negocios para que esté en la ciudad entre los días 235 y 295. ¿Cuál es la probabilidad de que el nacimiento tenga lugar durante ese tiempo?

a) ¿Cuál es la probabilidad de que un embarazo elegido al azar dure menos de 246 días?
Z < a (debido a que el valor es < a, entonces, el valor que se obtiene de Excel es con el que nos quedamos.

$Z= (246 – 266) / 16 = -1,25$

$P(X < 246) = P(Z < −1,25) = 0,1056$ **Nota: Colocar en Excel =DISTR.NORM.ESTAND.N(-1,25;VERDADERO)**

(b) ¿Cuál es la probabilidad de que un embarazo elegido al azar dure más de 240 días?

Z> a (usamos la 1era propiedad)

$Z = (240 - 266) / 16 = -1,63$

$P(X > 240) = P(Z > −1,63) = 1 - P(-1,63) = 0,9484$ **Nota: Colocar en Excel =DISTR.NORM.ESTAND.N(-1,63;VERDADERO)**

(c) ¿Cuál es la probabilidad de que un embarazo elegido al azar dure más de 500 días?
Método 1: El sentido común nos dice que esto sería **imposible**.
Método 2: Z> a
$Z=(500 – 266) / 16 = 14.625.$
$P(X > 500) = P(Z > 14.625) = 1 – P(14,625) = 0$ **Nota: Colocar en Excel =DISTR.NORM.ESTAND.N(14,625;VERDADERO)**

(d) Supongamos que el esposo de una mujer embarazada ha programado sus viajes de negocios para que esté en la ciudad entre los días 235 y 295. ¿Cuál es la probabilidad de que el nacimiento tenga lugar durante ese tiempo?

$a<Z<b$ (usamos la 3era propiedad)

Z= (235 – 266) / 16) = -1,94
Z= (295 – 266) / 16 = 1,81.

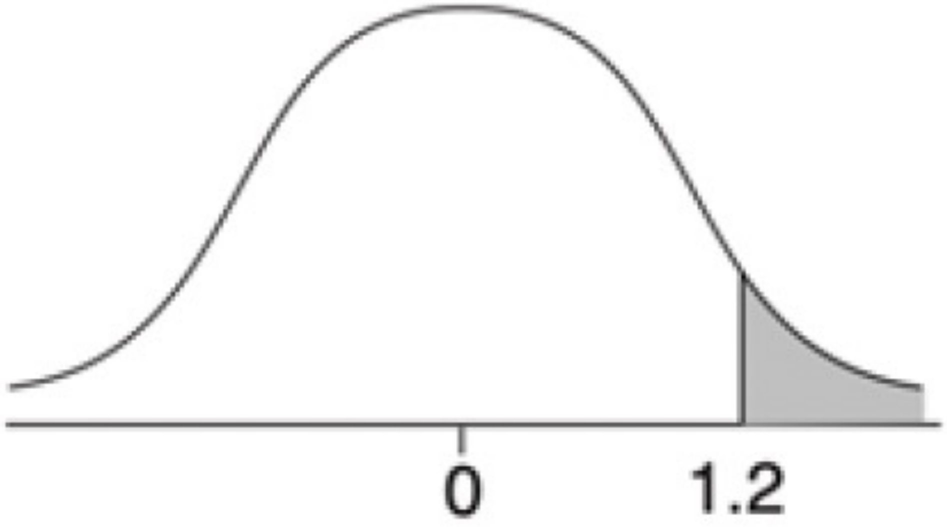
$$\mathbf{P(235 < X < 295) = P(-1{,}94 < Z < 1{,}81) = P(Z < 1{,}81) - P(Z < -1{,}94) = 0{,}9649 - 0{,}0262 = 0{,}9387}$$

Nota: Colocar en Excel **=DISTR.NORM.ESTAND.N(1,81;VERDADERO)**
=DISTR.NORM.ESTAND.N(-1,94;VERDADERO)

Hay cerca de un 94% de probabilidades de que el marido esté en la ciudad para el nacimiento.

2.

Si Z – N(0,1) encuentra:
P (Z>1,2)



$$P\ (Z \geq a) = 1 - P\ (Z \leq a)$$

$$P\ (Z > 1{,}2\ ) = 1 - P\ (Z < 1{,}2)$$

$$= 1 - P\ (0{,}8849)$$

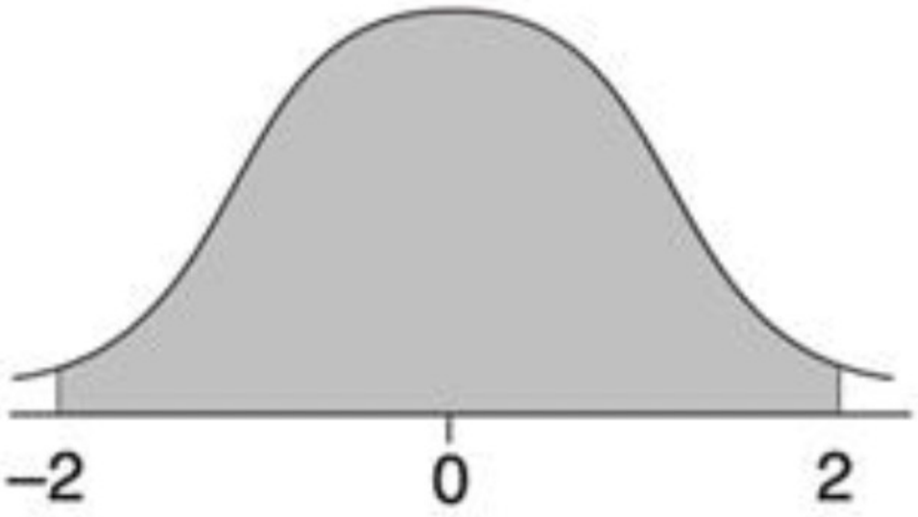
$$= 1 - 0{,}8849$$

$$= 0{,}\ 1151$$

Nota: Colocar en Excel =DISTR.NORM.ESTAND.N(1,2;VERDADERO)

3.

Si Z ~ N(0,1) encuentra:
P (-2<Z<2)



$$P\ (b \leq Z \leq a) = P\ (Z \leq a) - P\ (Z \leq b)$$

$$P\ (-2 < Z < 2) = P\ (Z < 2) - P\ (Z < -2)$$

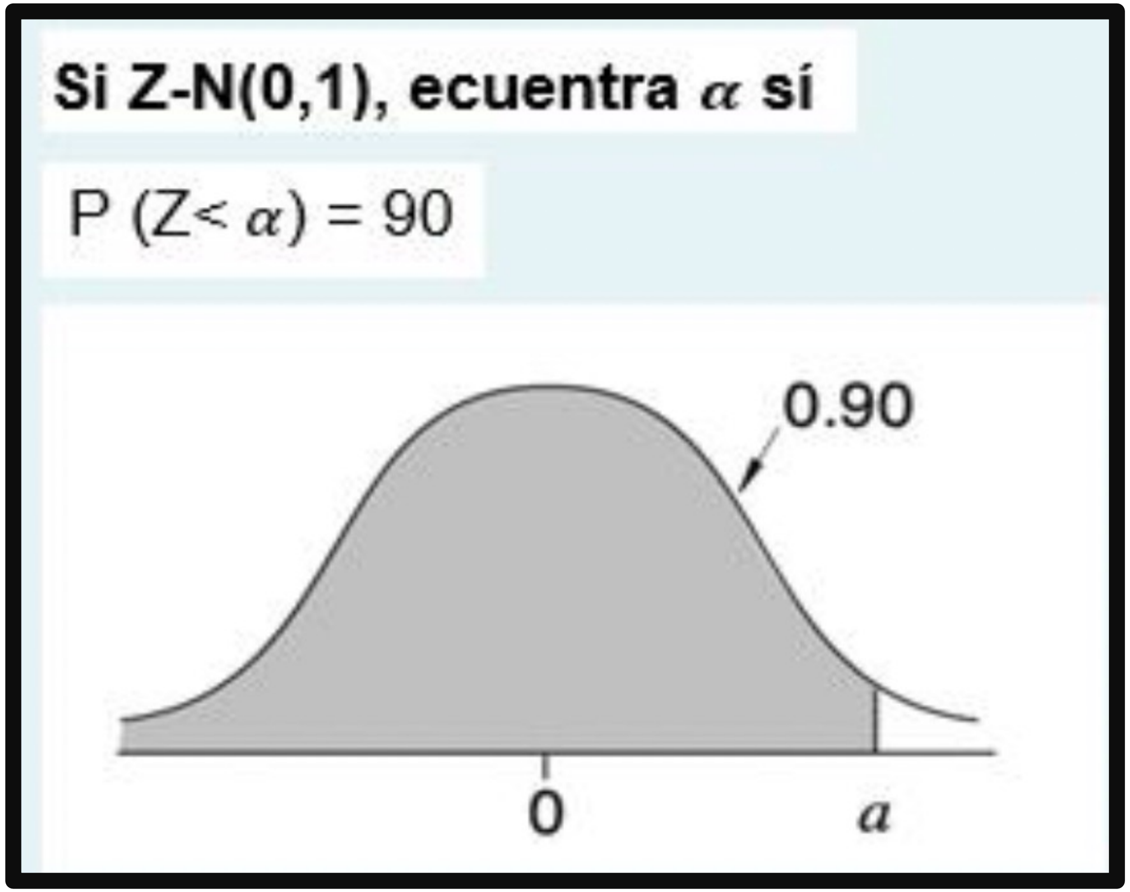
$$= P\ (Z < 0{,}9772) - P\ (Z < 0{,}0228)$$

$$= 0{,}9772 - 0{,}0228$$

$$= 0{,}9544$$

Nota: Colocar en Excel =DISTR.NORM.ESTAND.N(2;VERDADERO)
=DISTR.NORM.ESTAND.N(-2;VERDADERO)

4.



El ejercicio nos indica la probabilidad de 0,90. Es decir, vamos desde la curva a buscar a α

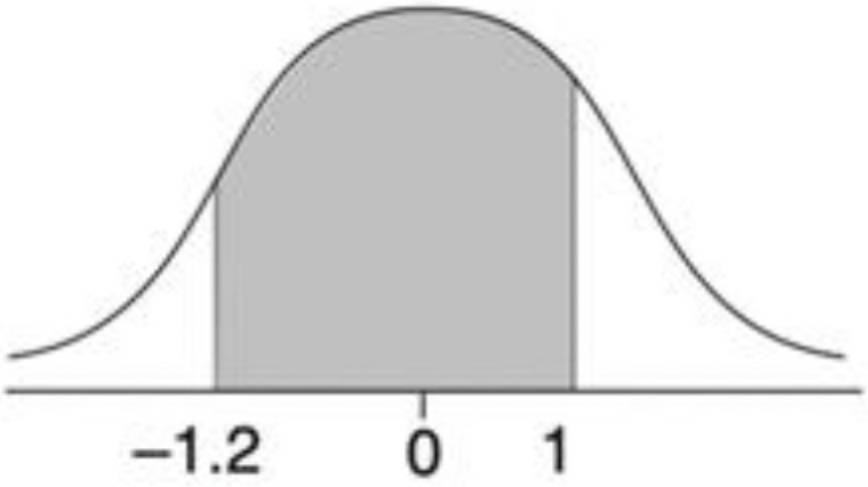
Nota: Colocar en Excel la función inversa
=DISTR.NORM.ESTAND.INV(0,9)

Por lo tanto, α equivale a 1,28

5.

Si $Z \sim N(0,1)$ encuentra:

$P(-1,2 < Z < 1)$



P (b ≤ Z ≤ a) = P (Z ≤ a) - P (Z ≤ b)

P (-1,2 < Z < 1) = P (Z < 1) – P (Z < -1,2

= P (Z <0,8413) – P (Z < 0,

= 0,8413- 0,1151

= 0,7262

Nota: Colocar en Excel **=DISTR.NORM.ESTAND.N(1;VERDADERO)**
=DISTR.NORM.ESTAND.N(-1,2;VERDADERO)

EJERCICIOS PROBABILIDADES

1.

En todos los Centros de Salud del MSP de Otavalo, como parte del Plan Estratégico Nacional Multisectorial para la Respuesta al VIH/SIDA e ITS se aplican pruebas ELISA de cuarta generación como parte del screening.

El test ELISA de cuarta generación permiten la detección simultánea de anticuerpos y antígeno p24, reduciéndose el período ventana a 13-15 días, es decir, se aproxima casi a la detección de ARN-VIH.

Los médicos del Centro eligen una población de 5014 hombres entre los 18 a 55 años, quienes se realizan la prueba de screening, la misma que se califica como positiva o negativa dependiendo del nivel de anticuerpos IgG, IgM, Agp24 en suero.

Posteriormente este mismo grupo de personas fueron sometidos a test o ensayos confirmatorios con una especificidad superior a las técnicas de ELISA.

Los resultados de las pruebas de screening se presentan a continuación:

Prueba de screening	VIH	No VIH	Total
Positivo	63	24	87
Negativo	17	4910	4927
Total	80	4934	5014

Con esta información, encuentra:

Tener VIH dado que tiene una prueba de screening positiva

Tener una prueba de screening positiva dado que tiene VIH

P (VIH / Prueba positiva) = 63 / 87 = 0,72. Es decir 72,4%
P (Prueba positiva /VIH) = 63 / 80 = 0,75. Es decir 75%

2.

En el Hospital Vicente Corral Moscoso, la prevalencia del glioma es del 0,003 %. El 25 de mayo del 2021, 80 personas ingresan al hospital quejándose de dolores de cabeza y pérdida de memoria. Los médicos sospechan de Glioma (tumor que se origina en el cerebro). El hospital cuenta con un nuevo análisis de sangre para el diagnóstico del glioma.

Resultado de la prueba	**Estado de la enfermedad**		**Total**
	Regalo de Glioma	**Glioma ausente**	
positivo	29	2	31
negativo	1	48	49
Total	30	50	80

Con esta información, encuentra:

Tener Glioma dado que tiene una prueba de sangre positiva

Tener una prueba de sangre positiva dado que tiene Glioma

P (Glioma /Prueba positiva) = 29 / 31 = 0,93. Es decir 93%
P (Prueba positiva /Glioma) = 29 / 30 = 0,96. Es decir 96%

3.

Los gerentes de una empresa de seguros de salud, cercana a asesores de alto nivel del sistema de salud pública del Ecuador, se han enterado que en 2 años se privatizará la salud pública del país. Ante la posibilidad de que en dos años exista una cepa mutante y sea más letal, se ha pedido a sus responsables de bioestadística, extraer información de una de las provincias más grandes del país. Específicamente, del grupo etario entre 30 y 75 años de edad.

Se obtiene que: durante el 2020 en el pico más alto de la pandemia, el 64% de la población era obesa (Ob), y el 36% diabética (Dm). De estas personas, el 50% de obesos y el 18% de diabéticos infectados por COVID19, ingresaron a UCI con diagnóstico de gravedad, lo cual significó un "gasto" diario de 6.000 USD para el Estado.

Con esa información se requiere determinar la probabilidad de que una persona (infectada con COVID19) tenga diabetes dado que fue ingresada a UCI.

		No UCI= 0,82		P(Dm)= PDm x PUCI
	Dm= 0,36			P(Dm)= 0,36 x 0,18
		UCI=0,18		P(Dm)= 0,065
Enfermedades				
		No UCI= 0,5		P(Ob)= POb x PUCI
	Ob= 0,64			P(Ob)= 0,64 x 0,5
		UCI=0,5		P(Ob)= 0,32

$P(Dm|UCI)= P(Dm)x(UCI|Dm)/P(UCI)$

$P(Dm|UCI)= P(0,36)x(0,18)/P(0,36x0,18+0,64x0,5)$

$P(Dm|UCI)= P(0,065)/P(0,065+0,32)$

$P(Dm|UCI)= P0,065/0,385$

$P(Dm|UCI)= 0,171 \times 100$

$P(Dm|UCI)= 17,1\%$

Esto quiere decir, que la empresa de seguros de salud, deberá presetar atención al grupo etario 30 a 75 años de una de las provincias mas grandes del Ecuador, recordándoles que el 17,1% del grupo etario con diabetes, tendrá la probabilidad de cursar con COVID19 grave, y requerir de ingreso a UCI.

4.

En el Centro de Salud de Peguche se solicita a un médico rural que examine a un adulto mayor. Según el perfil epidemiológico de las enfermedades respiratorias de la zona, el médico conoce que el 80% de los adultos mayores enfermos tiene el diagnóstico de neumonía, mientras que el otro 20% tiene tuberculosis pulmonar. Para representar el evento de un adulto mayor con neumonía se coloca N y para representar un evento de un adulto mayor con tuberculosis se coloca T.

Adicionalmente, se conoce que tener tos con expectoración por más de 15 días es un síntoma principal de tuberculosis, y se lo representa como E.

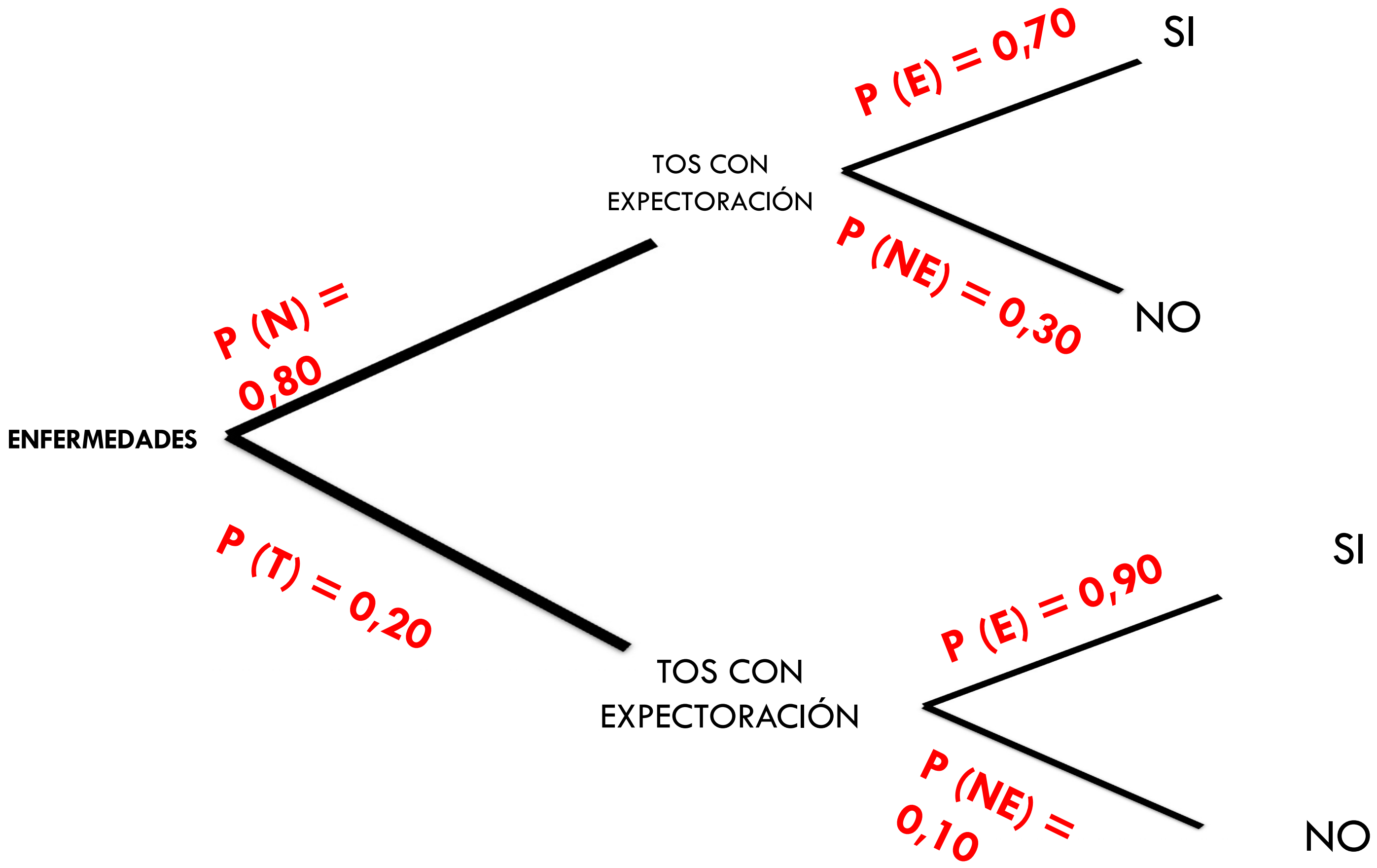
Por consiguiente, la probabilidad de tener tos con expectoración por más de 15 días si el adulto mayor tiene tuberculosis es P (E/T) = 0,90. Mientras que, la probabilidad de tener tos con expectoración por más de 15 días y tener neumonía es P (E/N) = 0,70.

Determinar la probabilidad de que el adulto mayor tenga tuberculosis dado que tiene tos con expectoración

Ordenamos los datos:

- Probabilidad de tener neumonía. **P (N) = 0,80**
- Probabilidad de tener tuberculosis pulmonar. **P (T) = 0,20**
- Probabilidad de tener tos con expectoración si tiene neumonía **P (E/N) = 0,70**
- Probabilidad de tener tos con expectoración si tiene tuberculosis **P (E/T) = 0,90**

Segundo, ante estos elementos realizamos un mapa de árbol y recordamos que las probabilidad siempre deben sumar 1



$$P\,(_{T/E}) = \frac{P(T) x\, P\,(E/T)}{P\,(E)}$$

$$P\,(_{T/E}) = \frac{P(0{,}20) x\, P\,(0{,}90)}{P\,(0{,}20\ X\ 0{,}90 + 0{,}80\ X\ 0{,}70)}$$

$$P\,(_{T/E}) = 0{,}2432$$

P (T/E)= 24,32%