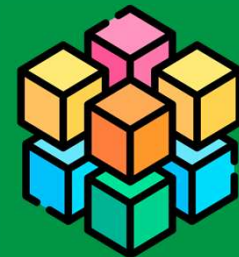


# Estadística II

Variables aleatorias continuas

Distribución NORMAL

Matemáticas Aplicadas II





# V.A. Normales

Las variables aleatorias **CONTINUAS** que siguen una distribución **NORMAL** aparecen en muchísimos fenómenos naturales y reales, siendo las más habituales.

## Ejemplos:

- Altura y peso de una población
- Ingresos de una empresa en muchos días diferentes
- Puntuaciones de exámenes
- Errores de medición.

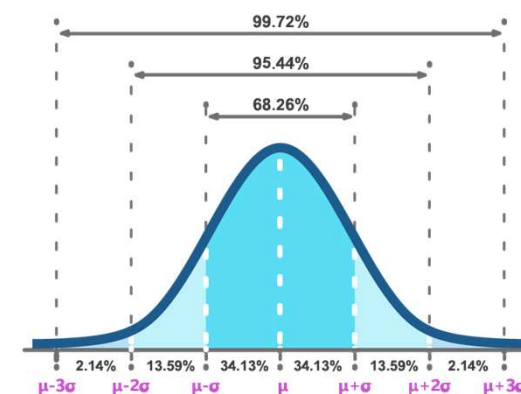


# Distribución NORMAL

La gráfica de su función de distribución tiene forma de campana, conocida como **campana de Gauss**.

## Características:

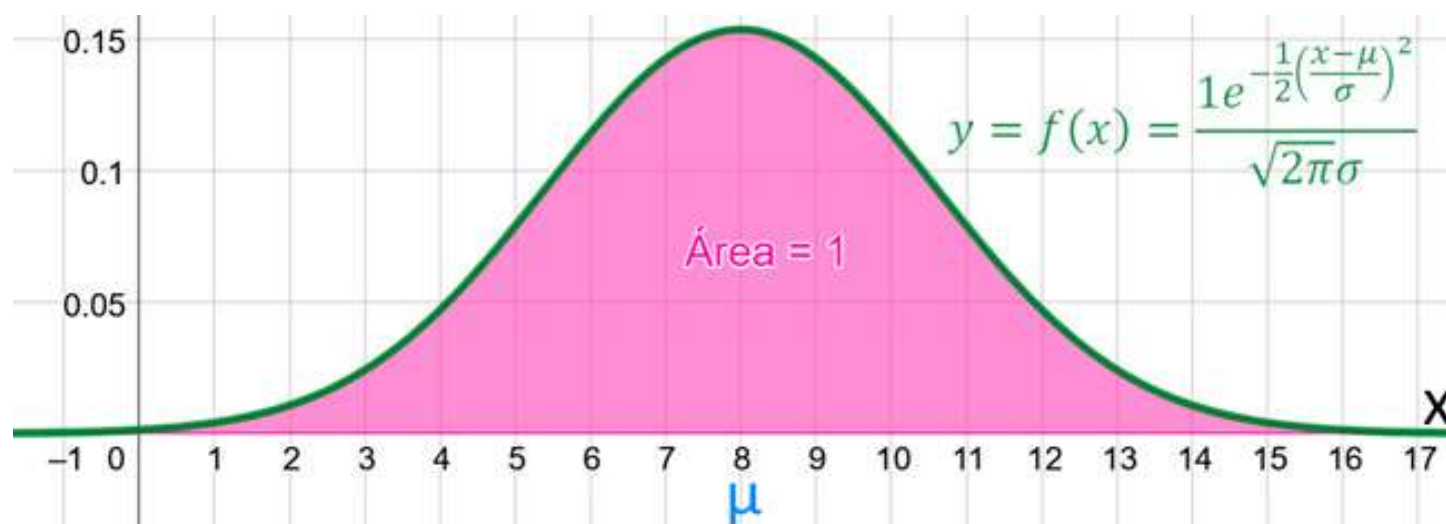
- Forma de campana
- Simétrica y centrada en la media  $\mu$
- los valores más cercanos al centro son los más frecuentes
- Las “colas”, valores alejados, marcados por la desviación típica  $\sigma$ , son los casos raros o menos frecuentes.





# De dónde viene

La gráfica de la función de distribución de una V.A. normal se llama **campana de Gauss**, es muy conocida y útil



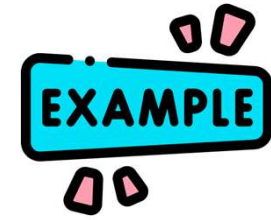
La probabilidad  $P(Z < a)$  es el área bajo la curva hasta el punto  $a$

# Para qué sirve



- 👉 Modelar fenómenos naturales y sociales donde los datos se agrupan en torno a un valor central, la media.
- 👉 Permite determinar la probabilidad de que un valor caiga en un intervalo específico (área bajo la curva)
- 👉 Son fundamentales en inferencia estadística, control de calidad, análisis de riesgos y aproximación de otras distribuciones.

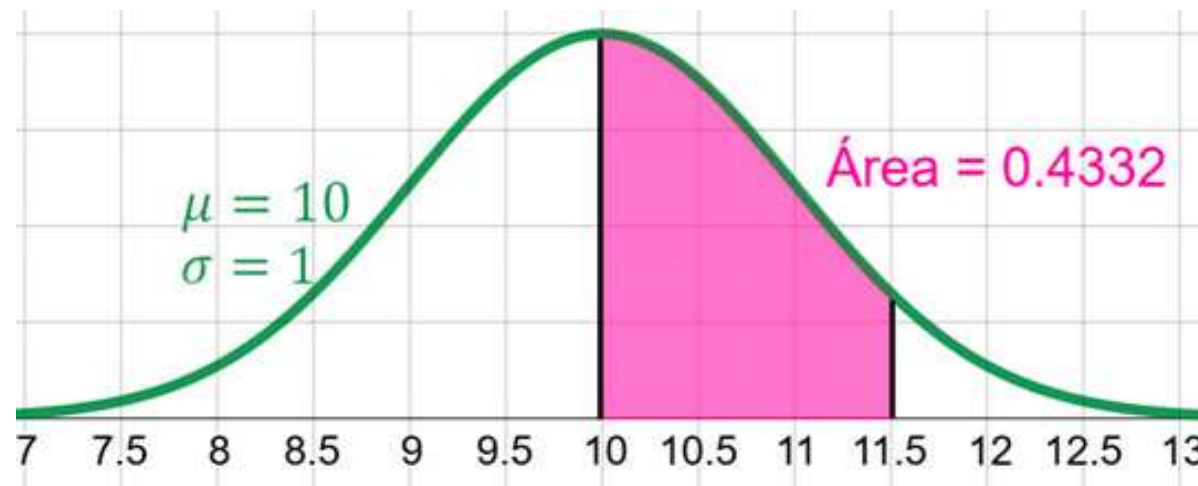
# Ejemplo



En una oficina de correos el tiempo medio de espera es de 10 minutos, con una desviación típica de 1 minuto. Queremos saber qué porcentaje de clientes tendrán que esperar un minuto y medio más de lo habitual

$X = \text{"tiempo que espera un cliente"} \sim N(10, 1)$

lo que buscamos es  $P(10 < X < 11.5) =$  al área comprendida entre 10 y 11.5 =  $0.4332 = 43,32\%$



# Media $\mu$



Es el **valor central o promedio**. Es el valor más esperado o probable, también se le llama esperanza. Es el centro de la campana.



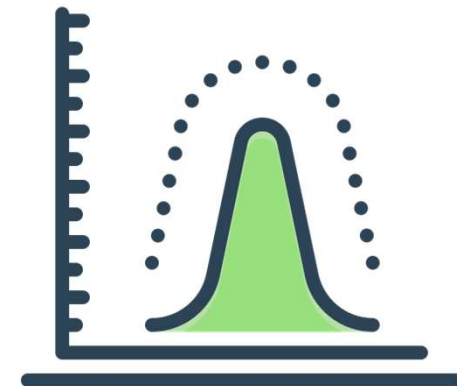
# Desviación típica $\sigma$



Mide **cuánto se alejan los datos de la media**. Ahí decimos que empiezan las “colas” o valores “atípicos”

Interpretación intuitiva:

- pequeña  $\rightarrow$  datos muy concentrados
- grande  $\rightarrow$  datos muy dispersos



**Ejemplo:**

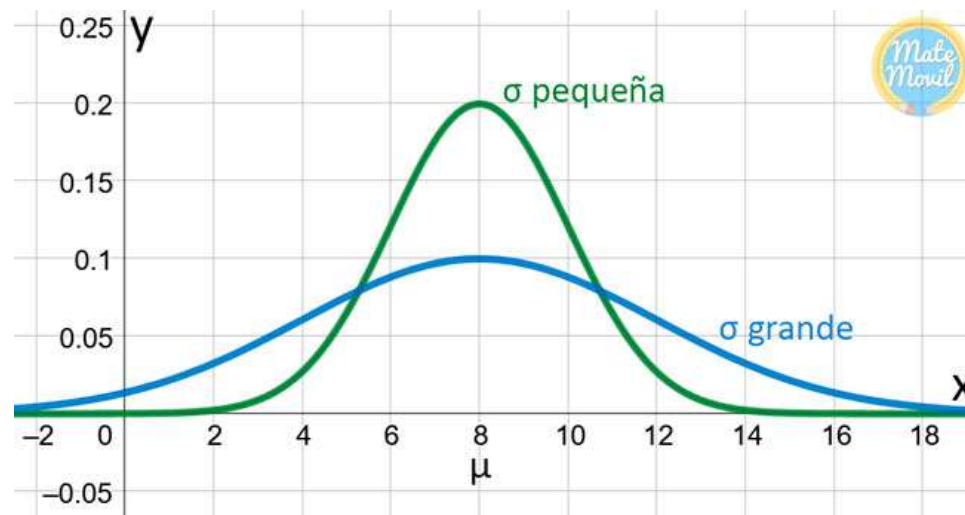
Si estudiamos el salario de dos países que tienen el mismo salario medio, la desviación típica nos informará en cuál de ellos hay **más desigualdad**.

# Desviación típica $\sigma$



Interpretación intuitiva:

- pequeña  $\rightarrow$  datos muy concentrados
- grande  $\rightarrow$  datos muy dispersos



# Normal estándar

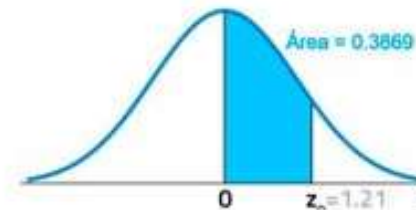


Es la más utilizada, con  $\mu = 0$  y  $\sigma = 1$ .

Si  $Z$  es una variable aleatoria normal estándar se indica así:  $Z \sim N(0,1)$



Todos sus valores están calculados y organizados en una tabla de acceso universal. Así que no hay que aprenderse fórmulas ni calcular áreas con integrales



| z   | 0.00   | 0.01   | 0.02   | 0.03   | 0.04   | 0.05   | 0.06   | 0.07   | 0.08   | 0.09   |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 1.3 | 0.4032 | 0.4049 | 0.4066 | 0.4082 | 0.4099 | 0.4115 | 0.4131 | 0.4147 | 0.4162 | 0.4177 |
| 1.4 | 0.4192 | 0.4207 | 0.4222 | 0.4236 | 0.4251 | 0.4265 | 0.4279 | 0.4292 | 0.4306 | 0.4319 |
| 1.5 | 0.4332 | 0.4345 | 0.4357 | 0.4370 | 0.4382 | 0.4394 | 0.4406 | 0.4418 | 0.4429 | 0.4441 |
| 1.6 | 0.4452 | 0.4463 | 0.4474 | 0.4484 | 0.4495 | 0.4505 | 0.4515 | 0.4525 | 0.4535 | 0.4545 |
| 1.7 | 0.4554 | 0.4564 | 0.4573 | 0.4582 | 0.4591 | 0.4599 | 0.4608 | 0.4616 | 0.4625 | 0.4633 |

# Cómo se usa



Uso de tablas de la normal

$$Z \sim N(0,1)$$

Las tablas nos dan el porcentaje o probabilidad de que los valores estén **por debajo** de un cierto valor establecido  $P(Z < a)$

$$\text{Ejemplo: } P(Z < 1.25) = 0.8944$$

| z   | 0,00   | 0,01   | 0,02   | 0,03   | 0,04   | 0,05   | 0,06   | 0,07   | 0,08   | 0,09   |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0,0 | 0,5000 | 0,5040 | 0,5080 | 0,5120 | 0,5160 | 0,5199 | 0,5239 | 0,5279 | 0,5319 | 0,5359 |
| 0,1 | 0,5398 | 0,5438 | 0,5478 | 0,5517 | 0,5557 | 0,5596 | 0,5636 | 0,5675 | 0,5714 | 0,5753 |
| 0,2 | 0,5793 | 0,5832 | 0,5871 | 0,5910 | 0,5948 | 0,5987 | 0,6026 | 0,6064 | 0,6103 | 0,6141 |
| 0,3 | 0,6179 | 0,6217 | 0,6255 | 0,6293 | 0,6331 | 0,6368 | 0,6406 | 0,6443 | 0,6480 | 0,6517 |
| 0,4 | 0,6554 | 0,6591 | 0,6628 | 0,6664 | 0,6700 | 0,6736 | 0,6772 | 0,6808 | 0,6844 | 0,6879 |
| 0,5 | 0,6915 | 0,6950 | 0,6985 | 0,7019 | 0,7054 | 0,7088 | 0,7123 | 0,7157 | 0,7190 | 0,7224 |
| 0,6 | 0,7257 | 0,7291 | 0,7324 | 0,7357 | 0,7389 | 0,7422 | 0,7454 | 0,7486 | 0,7517 | 0,7549 |
| 0,7 | 0,7580 | 0,7611 | 0,7642 | 0,7673 | 0,7704 | 0,7734 | 0,7764 | 0,7794 | 0,7823 | 0,7852 |
| 0,8 | 0,7881 | 0,7910 | 0,7939 | 0,7967 | 0,7995 | 0,8023 | 0,8051 | 0,8078 | 0,8106 | 0,8133 |
| 0,9 | 0,8159 | 0,8186 | 0,8212 | 0,8238 | 0,8264 | 0,8289 | 0,8315 | 0,8340 | 0,8365 | 0,8389 |
| 1,0 | 0,8413 | 0,8438 | 0,8461 | 0,8485 | 0,8508 | 0,8531 | 0,8554 | 0,8577 | 0,8599 | 0,8621 |
| 1,1 | 0,8643 | 0,8665 | 0,8686 | 0,8708 | 0,8729 | 0,8749 | 0,8770 | 0,8790 | 0,8810 | 0,8830 |
| 1,2 | 0,8849 | 0,8869 | 0,8888 | 0,8907 | 0,8925 | 0,8944 | 0,8962 | 0,8980 | 0,8997 | 0,9015 |

👉 interpretación: el **89,44%** de los valores están por debajo de 1.25



# Propiedades 1

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

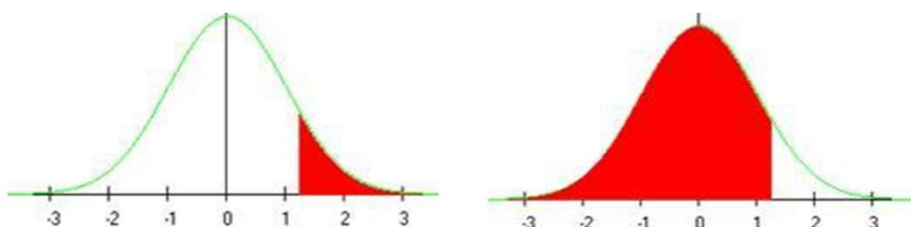
- $P(Z = a) = 0$  como es una variable continua, su probabilidad se mide como el área bajo la curva y **un punto... no tiene área.**
- $P(Z < a) = P(Z \leq a)$
- El **área total** bajo la curva es **1**. Ese área representa la probabilidad de que ocurra **cualquier valor posible**, lo que equivale a la certeza total, el  $100\% = 1$



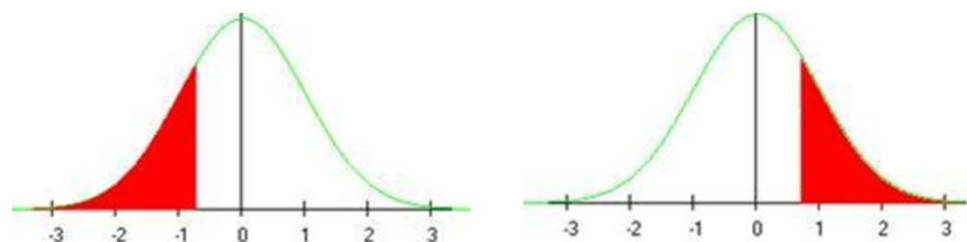
# Propiedades 2

Las probabilidades que NO están en las tablas se deducen fácilmente por simetría, no hay que aprenderlas, es más intuitivo dibujar:

- $P(Z > a) = 1 - P(Z < a)$



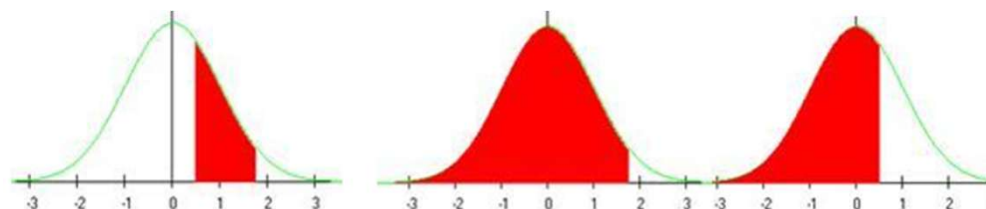
- $P(Z > -a) = P(Z < a)$





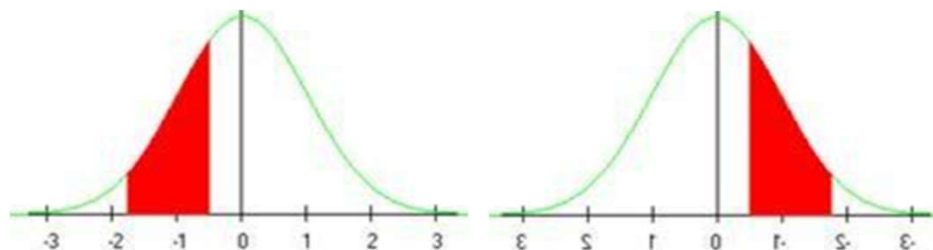
# Propiedades 3

- Entre dos valores positivos:  $P(a < Z < b) = P(Z < b) - P(Z < a)$



- Entre dos negativos:

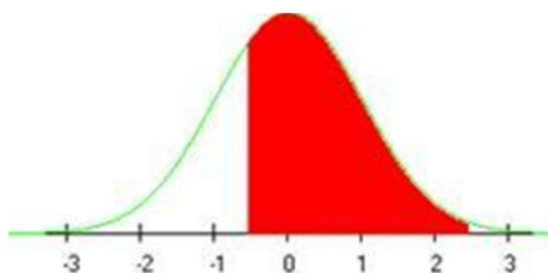
$$P(-b < Z < -a) = P(a < Z < b)$$





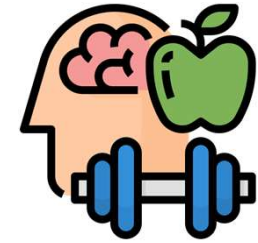
# Propiedades 4

- Entre uno negativo y uno positivo:  $P(-a < Z < b) =$   
 $P(Z > -a) - P(Z > b) = P(Z < a) - [1 - P(Z < b)]$   
 $= P(Z < a) + P(Z < b) - 1$



- Caso particular  $P(-a < Z < a) = 2P(Z < a) - 1$

# Ejercicio 1



Sea  $Z \sim N(0,1)$  sabiendo que  $P(Z < 0.72) = 0.7642$  y  $P(Z < 0.5) = 0.6915$  calcula:

- a)  $P(Z > 0.72)$
- b)  $P(Z < -0.5)$
- c)  $P(Z > -0.72)$
- d)  $P(-0.72 < Z < 0.72)$
- e)  $P(-0.5 < Z < 0.72)$

# Tabla “inversa”



En este caso, en lugar de calcular la probabilidad de un valor, estamos buscando justo lo contrario.

Partimos de que conocemos la probabilidad y buscamos en la tabla el valor que le corresponde.

## Ejemplo:

buscamos un valor  $k$  tal que el 67,38% estén por debajo de él

$$P(Z < k) = 0,6738$$

Mirando la tabla:  $k \approx 0,45$

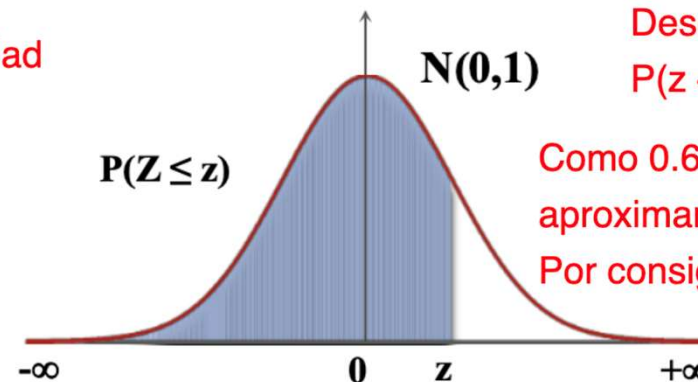
👉 Interpretación: el **67,36% de los valores están por debajo de 0,45.**

# Tabla “inversa”



## FUNCIÓN DE DISTRIBUCIÓN NORMAL $N(0,1)$

Si el valor de la probabilidad  
no aparece en la tabla,  
lo aproximamos



Deseamos hallar  $k$  con  
 $P(z < k) = 0.6738$

Como 0.6768 no aparece en la tabla, lo  
aproximamos a 0.6736  
Por consiguiente,  $k \approx 0.45$

| z   | 0,00   | 0,01   | 0,02   | 0,03   | 0,04   | 0,05   | 0,06   | 0,07   | 0,08   | 0,09   |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| 0,0 | 0,5000 | 0,5040 | 0,5080 | 0,5120 | 0,5160 | 0,5199 | 0,5239 | 0,5279 | 0,5319 | 0,5359 |
| 0,1 | 0,5398 | 0,5438 | 0,5478 | 0,5517 | 0,5557 | 0,5596 | 0,5636 | 0,5675 | 0,5714 | 0,5753 |
| 0,2 | 0,5793 | 0,5832 | 0,5871 | 0,5910 | 0,5948 | 0,5987 | 0,6026 | 0,6064 | 0,6103 | 0,6141 |
| 0,3 | 0,6179 | 0,6217 | 0,6255 | 0,6293 | 0,6331 | 0,6368 | 0,6406 | 0,6443 | 0,6480 | 0,6517 |
| 0,4 | 0,6554 | 0,6591 | 0,6628 | 0,6664 | 0,6700 | 0,6736 | 0,6772 | 0,6808 | 0,6844 | 0,6879 |
| 0,5 | 0,6915 | 0,6950 | 0,6985 | 0,7019 | 0,7054 | 0,7088 | 0,7123 | 0,7157 | 0,7190 | 0,7224 |
| 0,6 | 0,7257 | 0,7291 | 0,7324 | 0,7357 | 0,7389 | 0,7422 | 0,7454 | 0,7486 | 0,7517 | 0,7549 |
| 0,7 | 0,7580 | 0,7611 | 0,7642 | 0,7673 | 0,7704 | 0,7734 | 0,7764 | 0,7794 | 0,7823 | 0,7852 |

# Tipificación



La tipificación nos permite usar la **misma tabla** para cualquier variable **normal**  $X \sim N(\mu, \sigma)$

Transformamos cualquier variable normal en la normal estándar mediante:

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

Esto se llama **tipificación**.

# Tipificación



$$X \sim N(\mu, \sigma)$$



tipificación



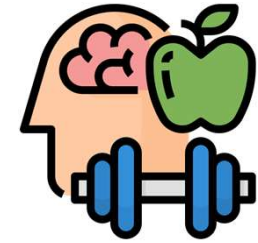
$$Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$



$$P(X < a) = P\left(Z < \frac{a - \mu}{\sigma}\right)$$

Y ahora ya puedo usar la tabla estándar  $N(0,1)$

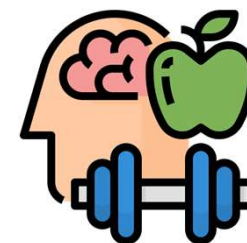
# Ejercicio



En una empresa el **tiempo de entrega** sigue una distribución normal  $N(30,5)$  minutos

- a) Calcula la probabilidad de que una entrega tarde **menos de 25 minutos**.
- b) Calcula la probabilidad de que tarde **entre 25 y 35 minutos**.

# Ejercicio



Tipificamos:

$$Z = \frac{X - 30}{5} \sim N(0,1)$$

$$\begin{aligned} a) P(X < 25) &= P\left(Z < \frac{25 - 30}{5}\right) = P(Z < -1) = P(Z > 1) \\ &= 1 - P(Z < 1) \approx 1 - 0.8413 = \mathbf{0.1587} \end{aligned}$$

$$\begin{aligned} b) P(25 < X < 35) &= P\left(\frac{25 - 30}{5} < Z < \frac{35 - 30}{5}\right) = P(-1 < Z < 1) = \\ &P(Z < 1) - P(Z < -1) \approx 0.8413 - 0.1587 = \mathbf{0.6826} \end{aligned}$$

# Transición de Binomial a Normal



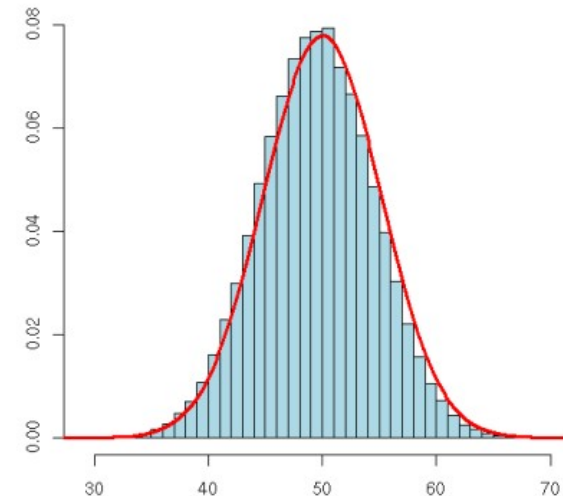
¿Qué pasa si  $n$  es muy grande? De Binomial  $\rightarrow$  Normal

El diagrama de la distribución de la binomial empieza a parecerse a una campana.

Si  $n$  es muy grande podremos aproximar la variable Binomial a una Normal.

Para ello su diagrama debiera ser:

- simétrico
- centrado en la media
- la mayoría de valores cerca del centro



# Transición de Binomial a Normal



La aproximación de una distribución binomial a una normal es posible cuando el tamaño de la **muestra es grande** y la probabilidad de éxito no es extrema.

Las condiciones clave son que el número de éxitos y fracasos esperados

$$n \cdot p \geq 5 \qquad n \cdot q \geq 5$$

En ese caso  $X \sim B(n, p) \rightarrow Z \sim N(\mu, \sigma)$  recordando que

$$\mu = np \qquad \sigma = \sqrt{np(1-p)}$$

# Errores típicos en las PAU



los alumnos fallan porque:

1. no entienden **qué variable es**
2. no distinguen **binomial vs normal**
3. no interpretan resultados.

# Errores típicos en las PAU



Por eso conviene usar este esquema mental:

