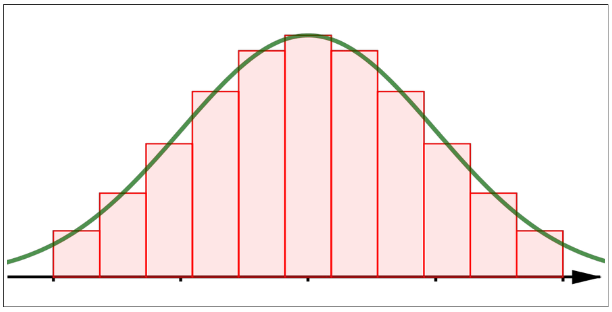
8 Distribuciones Normales

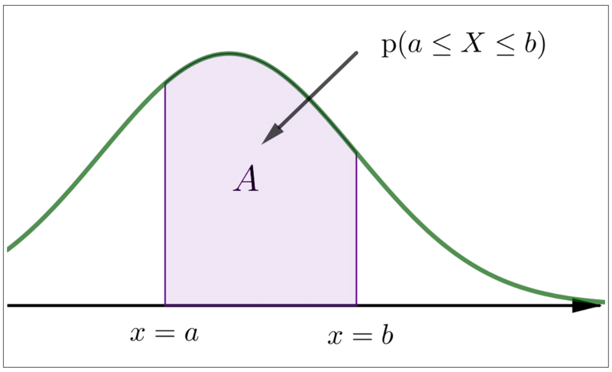
Variables Normales

Gauss descubrió que una gran cantidad de variables estadísticas tenían una características semejantes que podían estudiarse mediante unas curvas exponenciales llamadas, en su honor, campanas de Gauss.



Dichas variables se denominan normales y se dice también que sus datos siguen una distribución normal.

La característica fundamental de las variables normales es la siguiente: la probabilidad de encontrar un individuo en el que la variable normal X tome un valor comprendido entre a y b viene dada por el área bajo la curva de Gauss en dicho intervalo:



Casi toda variable que sea suma de una serie elevada de características sigue una distribución normal.

Hay muchísimas variables que se distribuyen siguiendo el modelo normal: morfológicas (como la talla o el peso de las personas, la altura de los árboles de cierta especie), fisiológicas (como la presión sanguínea o la frecuencia cardíaca), sociológicas (como el consumo de ciertos productos por personas de un mismo colectivo), físicas (como la resistencia a la rotura de piezas),...

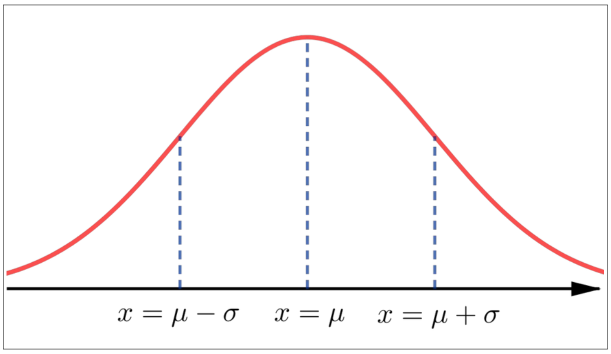
Nota: una función cuya gráfica proporciona así la probabilidad se denomina **función de densidad**

Campana de Gauss

Para cada media media μ y cada desviación típica σ hay una curva normal que se designa por $\mathrm{N}(\mu\,,\sigma)$. Su ecuación es

$$y = \frac{1}{\sigma\sqrt{2\pi}}\,\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma}}$$

Se trata de una función continua, simétrica, cuyo máximo se alcanza para $x = \mu$ y que tiene dos puntos de inflexión situados a distancia σ de dicha media:



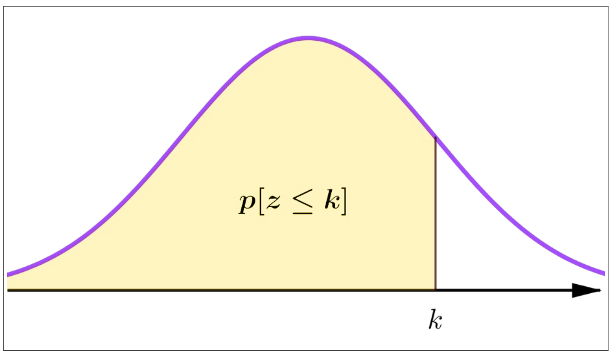
Una característica fundamental de todas ellas es que el área bajo cualquier curva normal es 1.

La Normal Típica

Todas las curvas normales son esenciales idénticas, sólo se distinguen en un cambio de eje de simetría y un cambio de escala. Esto permite calcular cualquier área a partir de una de ellas.

Se ha tomado como base de cálculo a la que tiene $\mu = 0$ y $\sigma = 1$, denominándose normal típica y se reserva para ella la letra z:

$$Z \sim \mathrm{N}(0\,,1)$$



La tabla proporciona $p\,[z \leq k]$ para $k \geq 0$.

8 Distribuciones Normales

Probabilidades Típicas

Así calculamos con la tabla $Z \sim \mathrm{N}(0,1)$ si $k \geq 0$:

$\mathrm{p}[z \leq k] =$ ver tabla

$\mathrm{p}[z > k] = 1 - \mathrm{p}[z \leq k]$

$\mathrm{p}[z > -k] = \mathrm{p}[z \leq k]$

$\mathrm{p}[z \leq -k] = \mathrm{p}[z \geq k]$

$\mathrm{p}[a \leq z \leq b] = \mathrm{p}[z \leq b] - \mathrm{p}[z \leq a]$

Tipificación

Dada una variable $\mathbf{X} \sim \mathcal{N}(\mu, \sigma)$, es:

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

Esta propiedad es conocida “fórmula de tipificación”:

Ejemplo: La estatura de una población se distribuye normalmente con media 170 cm. y desviación típica 6 cm ¿Qué probabilidad hay de que un individuo tenga una estatura inferior a los 175 cm?

Es $\mathbf{X}$ = “estatura de un individuo” una variable $\mathcal{N}(170, 6)$. Tipificando obtenemos:

Buscamos en la tabla y tenemos:

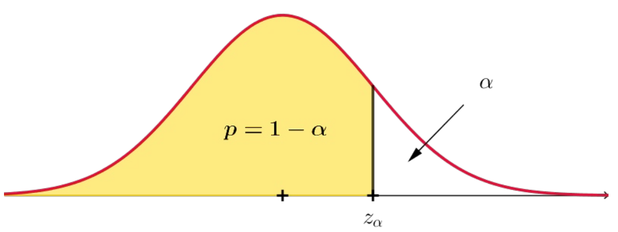
$$\mathrm{p}[X \leq 175] = \mathrm{p}\left[X \leq \frac{175 - 170}{6}\right] = \mathrm{p}[Z \leq 0.83]$$
$$= 0.7967$$

Valores Críticos

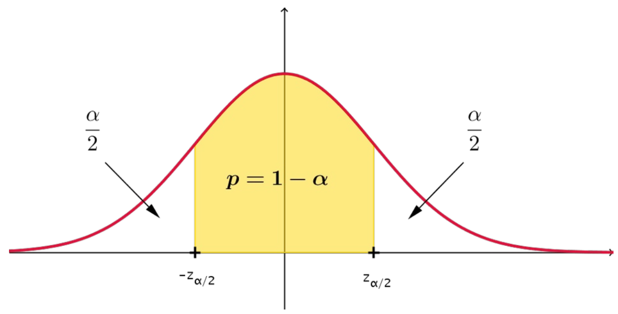
A veces es necesaria a búsqueda inversa. El vocabulario y la notación son rebuscados.

Consideremos un número α, comprendido entre 0 y 1, que denominaremos nivel de significación.

Valor unilateral asociado es el z_α con $\mathrm{p}[z > z_\alpha] = \alpha$



Valor bilateral es el $z_{\alpha/2}$ con $\mathrm{p}\left[z > z_{\alpha/2}\right] = \frac{\alpha}{2}$:



Nota: al porcentaje asociado $(1-\alpha) \cdot 100\%$ se le suele llamar nivel de confianza.

Normal y Binomial

Si conoces la fórmula de la de probabilidad binomial:

$$\mathrm{p}[X = k] = \binom{n}{k} p^k q^{n-k}$$

resulta que cuando es $np > 5$ y $nq > 5$, la distribución normal $\mathbf{N} \sim \mathcal{N}(np, \sqrt{npq})$ permite una excelente aproximación. Concretamente:

$\mathrm{p}[X = k] = \mathrm{p}[k - 0.5 \leq N \leq k + 0.5]$

$\mathrm{p}[X \leq k] = \mathrm{p}[N \leq k + 0.5]$

$\mathrm{p}[X \geq k] = \mathrm{p}[N \geq k - 0.5]$