Distribución Normal

También conocida como distribución gaussiana, es una de las distribuciones más comunmente utilizadas por físicos, químicos e ingenieros ya que

- Si repite un experimento una cierta cantidad de veces entonces la variable que representa el promedio de los resultados tiene aproximadamente una distribución normal.
- Aparece en el estudio de numerosos fenómenos físicos (por ejemplo: velocidad de moléculas (Maxwell))
- etc.

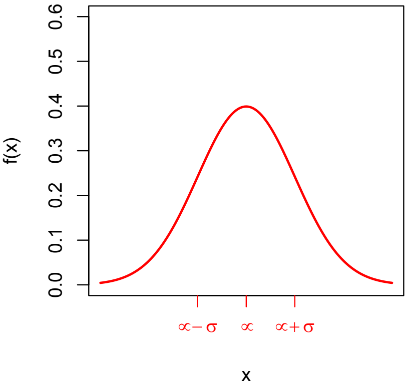
Distribución Normal

Si X es una v.a. con Distribución Normal con media μ y desviación estándar σ, entonces $X \sim N(\mu, \sigma^2)$

Su función de densidad está dada por

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty$$

- $E(X) = \mu$
- $V(X) = \sigma^2$



Distribución Normal

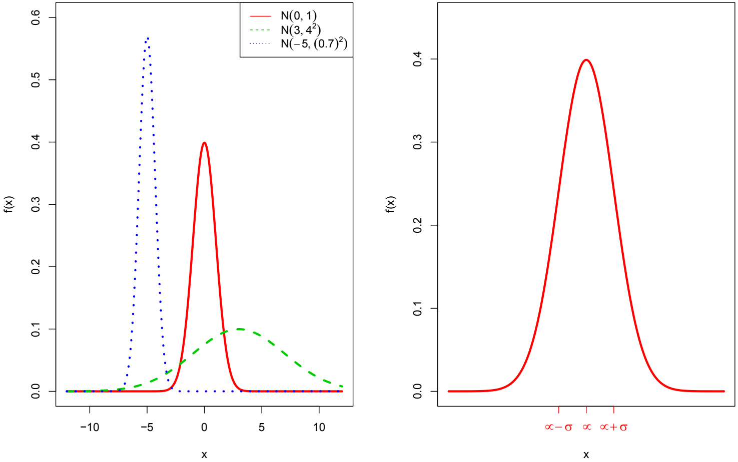
Si una v.a. con Distribución Normal tiene media $\mu = 0$ y desviación estándar $\sigma = 1$ entonces se dice que esa v.a. tiene Distribución Normal estándar, y se denota por Z

Su función de distribución acumulada se denota por:

$$\Phi(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-t^2}{2}}\, dt$$

Además, $E(X) = 0, V(X) = 1$

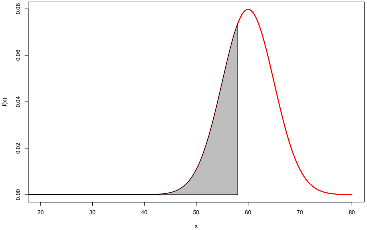
Distribución Normal



Ejemplo 4

El tiempo que tarda una célula en dividirse (mitosis) tiene distribución normal con un tiempo promedio de una hora y una desviación estándar de 5 minutos. ¿Cuál es la probabilidad que una célula se divida en menos de 58 minutos?

- $X \sim N(60, 5^2)$
- $P(X \leq 58) = \frac{1}{\sqrt{2\pi}5} \int_{-\infty}^{58} e^{\frac{-(x-60)^2}{2 \cdot 5^2}} dx$
- ¿Es posible calcular eso para cualquier valor de μ y σ?



Distribución Normal Estándar

Si, es posible. Para ello se estandariza la variable aleatoria X:

Así si $X \sim N(\mu, \sigma^2)$, entonces la v.a $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$ y :

$$P(X \leq x) = P(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}) = P(Z \leq \frac{x-\mu}{\sigma}) = \Phi(z)$$

con $z = \frac{x-\mu}{\sigma}$. La función $\Phi(z)$ es la función de distribución acumulada de una v.a con distribución normal estándar y está tabulada.

Supongamos $X \sim N(\mu, \sigma^2)$ $\quad P(X \leq x) = F_X(x)$

Defino una nueva v.a. $Z = \frac{X-\mu}{\sigma}$. Veamos cual es la distribución de Z.

Sea $y \in \mathbb{R}$

$$F_Z(y) \underset{\text{definición}}{=} P(Z \leq y) \underset{Z=\frac{X-\mu}{\sigma}}{=} P\left(\frac{X-\mu}{\sigma} \leq y\right) \underset{\text{despejo } X}{=} P(X \leq \sigma y + \mu) =$$

$$= F_X(\sigma y + \mu)$$

Por ahora tenemos que $F_Z(y) = F_X(\sigma y + \mu)$ $\quad \forall y \in \mathbb{R}$

Como Z es una v.a. continua, su función de densidad es

$$f_Z(y) = \frac{dF_Z(y)}{dy} = \frac{d}{dy} F_X(\sigma y + \mu) \underset{\frac{d}{dy}F_X = f_X}{=} \sigma . f_X(\sigma y + \mu)$$

$$\therefore f_Z(y) = \sigma . f_X(\sigma y + \mu) \underset{X \sim N(\mu, \sigma^2)}{=} \sigma . \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(\sigma y + \mu - \mu)^2}{2\sigma^2}\right)$$

$$= \frac{1}{\sqrt{2\pi}} . \exp\left(-\frac{y^2}{2}\right) \qquad \therefore f_Z(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right)$$

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

Queríamos calcular $P(X \leq 58)$, con $X \sim N(60, 5^2)$

$$P(X \leq 58) = P\left(\underbrace{\frac{X-60}{5}}_{} \leq \frac{58-60}{5}\right) \underset{Z = \frac{X-60}{5} \sim N(0,1)}{=} P\left(Z \leq -\frac{2}{5}\right) =$$

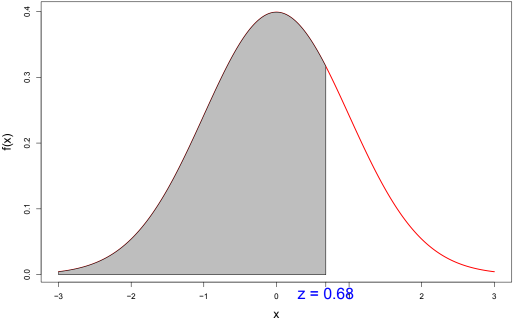
$$= P(Z \leq -0{,}4) = \Phi(-0{,}4)$$

$$= 0{,}3446$$

Distribución Normal Estándar

Si $Z \sim N(0,1)$, supongamos $z = 0.63$, entonces:

$$P(Z \leq 0.63) = \Phi(0.63)$$



Distribución Normal - Lectura de Tabla

Usando la tabla:

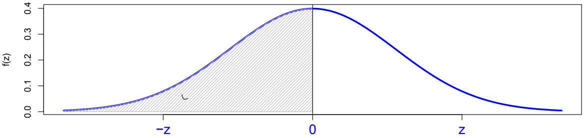
$$P(Z \leq 0.63) = \Phi(0.63)$$

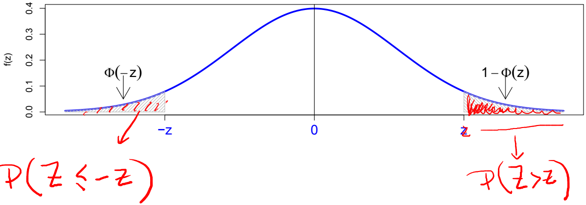
z	0.00	0.01	0.02	<u>0.03</u>	0.04	0.05 ..
0.0	0.500000	0.503989	0.507978	0.511967	0.515953	0.519939
0.1	0.539828	0.543795	0.547758	0.551717	0.555760	0.559618
0.2	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706
0.3	0.617911	0.621719	0.625516	0.629300	0.633072	0.636831
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840
<u>0.6</u>	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154
0.7	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373
⋮						

Distribución Normal Estandar

Algunas propiedades: Sean $Z \sim N(0,1)$ y $z > 0$. Como la función de densidad de esta v.a es simétrica en torno a x=0,

- $\Phi(0) = 0.5$
- $\Phi(-z) = 1 - \Phi(z)$
- $P(Z > z) = \Phi(-z)$
- $P(-z \leq Z \leq z) = 2\Phi(z) - 1$





$$P(Z > z) = \Phi(-z)$$ por simetria

$$P(Z \leq -z) = P(Z > z) \qquad z > 0$$

$$\Phi(-z) = P(Z > z) = 1 - P(Z \leq z) = 1 - \Phi(z)$$

$$P(-z < Z < z) = \Phi(z) - \Phi(-z) = \Phi(z) - [1 - \Phi(z)]$$

$$= 2\Phi(z) - 1$$

Continuación-Ejemplo 4

Si $X \sim N(60, 5^2)$ entonces $Z = \frac{X-60}{5} \sim N(0,1)$ y

$$P(X \leq 58) = P(\frac{X-60}{5} \leq \frac{58-60}{5}) = \Phi(\frac{-2}{5}) = \Phi(-0.4)$$

Buscando en la tabla Normal, se tiene que

$$\Phi(-0.4) = P(Z \leq -0.4) = 0.3445$$

Distribución Normal-Ejemplo 5

Supongamos ahora que queremos determinar el tiempo (en minutos) para el cual la probabilidad que la división de la célula suceda en menos de ese tiempo es de 0.7. Es decir, queremos encontrar el valor de t para el cual $P(X \leq t) = 0.7$.

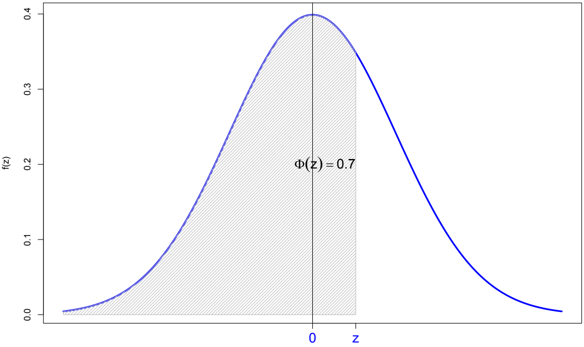
$Z \sim N(0,1)$

$$0.7 = P(X \leq t) = P\left(\frac{X-60}{5} \leq \frac{t-60}{5}\right) = \Phi\left(\frac{t-60}{5}\right)$$

Sea $z = \frac{t-60}{5}$, debemos encontrar z tal que $\Phi(z) = 0.7$

$P\left(Z \leq \frac{t-60}{5}\right)$

Continuación Ejemplo 5



•

Continuación Ejemplo 5

$z = 0{,}525$

$0{,}525 = z = \frac{t-60}{5}$

$t = 5 \cdot 0{,}525 + 60 = 62{,}62$

Buscando en la tabla vemos a que combinación de filas y columnas le corresponde la entrada 0.7

z	0.00	0.01	0.02	0.03	0.04	0.05 ..
0.0	0.500000	0.503989	0.507978	0.511967	0.515953	0.519939
0.1	0.539828	0.543795	0.547758	0.551717	0.555760	0.559618
0.2	0.579260	0.583166	0.587064	0.590954	0.594835	0.598706
0.3	0.617911	0.621719	0.625516	0.629300	0.633072	0.636831
0.4	0.655422	0.659097	0.662757	0.666402	0.670031	0.673645
0.5	0.691462	0.694974	0.698468	0.701944	0.705401	0.708840
0.6	0.725747	0.729069	0.732371	0.735653	0.738914	0.742154
0.7	0.758036	0.761148	0.764238	0.767305	0.770350	0.773373
⋮						

Así $0.52 = z = \frac{t-60}{5}$ y despejando resulta $t = 62.6$

$\frac{0{,}52 + 0{,}53}{2} = 0{,}525$

Ejercicios

1. La lectura de la temperatura medida por una termocupla que está ubicada en un medio de temperatura constante tiene distribución normal con media μ , la temperatura del medio, y desviación estándar σ. ¿Cuál debería ser el valor de σ para asegurarse que el 95% de las lecturas estén a 0.01^o de μ?
2. El peso de una zapatilla deportiva está normalmente distribuido con una media de 250 grs y una desviación estándar de 25 grs.
    2.1 ¿Cuál es la probabilidad que una de esas zapatillas pese más de 270 grs?
    2.2 ¿Cuánto debería ser la desviación estándar para que la compañía anuncie que el 99,9% de sus zapatillas pesan menos que 270 grs?
    2.3 Si la desviación estándar se mantiene en 25 grs, cuánto debería ser el peso medio para que la compañia pueda afirmar que el 95% de sus zapatillas pesa menos que 270 grs?

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

① $P(|X-\mu|<0{,}01) = 0{,}95$

$$0{,}95 = P(-0{,}01 < X-\mu < 0{,}01) = P\left(-\frac{0{,}01}{\sigma} < \frac{X-\mu}{\sigma} < \frac{0{,}01}{\sigma}\right) = (*)$$

Por propiedad, si $Z \sim N(0,1)$ $y>0$ $\Rightarrow P(-y < Z < y) = 2\Phi(y) - 1$

$$(*) = 2\Phi\left(\frac{0{,}01}{\sigma}\right) - 1$$

$$0{,}95 = 2\Phi\left(\frac{0{,}01}{\sigma}\right) - 1$$

$$\Phi\left(\underbrace{\frac{0{,}01}{\sigma}}_{z}\right) = \frac{1+0{,}95}{2} = 0{,}975$$

Buscamos en tabla el valor de z para el cual $\Phi(z) = 0{,}975$ y luego despejamos σ.

Si $z = 1{,}96 \Rightarrow \Phi(z) = 0{,}975$

$$\frac{0{,}01}{\sigma} = z = 1{,}96 \Rightarrow \sigma = \frac{0{,}01}{1{,}96} = 0{,}005$$

El valor de σ para el cual $P(|X-\mu|<0{,}01) = 0{,}95$ es aprox $\sigma \approx 0{,}005$

② X = peso de la zapatilla $\quad X \sim N(250, (25)^2)$

a) $$P(X>270) = P\left(\frac{X-250}{25} > \frac{270-250}{25}\right) =$$

$$\underset{\downarrow}{=} P(Z > 0{,}8) = 1 - \Phi(0{,}8) =$$

$z = \frac{X-250}{25}$ $\qquad = \Phi(-0{,}8)$

$$= 1 - 0{,}7881$$

$$= 0{,}2119$$

2.2 $\quad P(X \leq 270) = 0{,}999 \qquad$ Suponemos $X \sim N(250, \sigma^2)$

$\sigma = ?$

$$0{,}999 = P(X \leq 270) = P\left(\frac{X-250}{\sigma} \leq \frac{270-250}{\sigma}\right)$$

$$= P\left(Z \leq \frac{20}{\sigma}\right) = \Phi\left(\frac{20}{\sigma}\right)$$

$0{,}999 = \Phi\left(\frac{20}{\sigma}\right)$ Busco z tq $\Phi(z) = 0{,}999$

$z = 3{,}09$

$\frac{20}{\sigma} = 3{,}09 \Rightarrow \sigma = \frac{20}{3{,}09} \approx 6{,}47$

La desviación estándar debería ser $\sigma = 6{,}47$.

Generador de variables aleatorias: Método de inversión

Método de Inversión

Sea $U \sim \mathcal{U}[0,1]$ y sea F una función de distribución acumulada. Entonces la variable aleatoria $Y = F^{-1}(U)$ tiene distribución F

Ejemplo: Generación de una variable $Y \sim \mathcal{E}(\lambda)$

Si $U \sim \mathcal{U}[0,1]$ y F función de dist. acumulada (continua y estrict. creciente) $\Rightarrow$ $Y = F^{-1}(U)$ tiene función de dist. F.

Dem: Veamos cual es la func. de dist. de Y

$$F_Y(y) = P(Y \leq y) = P(F^{-1}(U) \leq y) \underset{\text{prop. de } F}{=} P(U \leq F(y))$$

$U \sim \mathcal{U}[0,1]$, además como F es una función de dist

$\Rightarrow 0 \leq F(y) \leq 1$

$$F_Y(y) = P(U \leq F(y)) = F_U(F(y)) = F(y)$$

$U \sim \mathcal{U}[0,1] \Rightarrow f_U(y) = \begin{cases} 1 & y \in [0,1] \\ 0 & cc \end{cases}$ $\quad F_U(y) = \begin{cases} 0 & y < 0 \\ y & 0 \leq y \leq 1 \\ 1 & y > 1 \end{cases}$

$\therefore F_Y(y) = F(y)$

Generemos $Y \sim \mathcal{E}(\lambda)$

Si $Y \sim \mathcal{E}(\lambda) \Rightarrow$ su función de dist. acumulada es

$F(y) = 1 - e^{-\lambda y}$. Busquemos $F^{-1}(y)$

$$x = 1 - e^{-\lambda y} \Rightarrow \begin{aligned} 1 - x &= e^{-\lambda y} \\ \ln(1-x) &= -\lambda y \\ y &= -\frac{\ln(1-x)}{\lambda} \end{aligned}$$

Método inversión:

1. generar $u \sim \mathcal{U}[0,1]$

2. Hacer $F^{-1}(u) = -\frac{\ln(1-u)}{\lambda}$