Distribuciones variables aleatorias continuas

Jessica Nathaly Pulzara Mora
jessica.pulzara@udea.edu.co

Departamento de ingeniería de sistemas

UNIVERSIDAD DE ANTIOQUIA

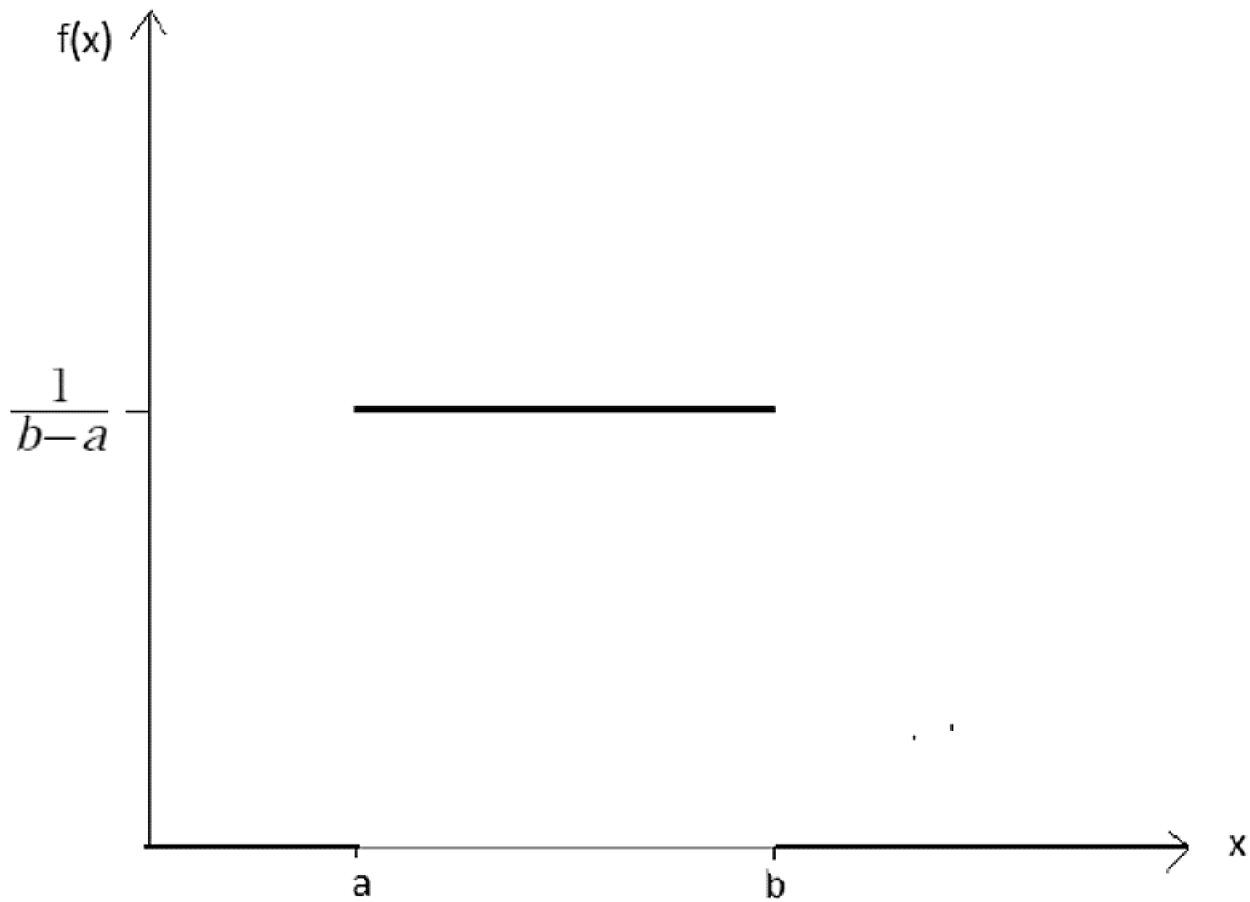
Distribución uniforme continua

Distribución uniforme continua

La variable aleatoria X tiene distribución uniforme continua sobre el intervalo $[a, b]$ si su función de densidad de probabilidad está dada por:

$$f(x; a, b) = \begin{cases} \frac{1}{b-a}, & \text{si } a \leq x \leq b \\ 0, & \text{en otro caso.} \end{cases}$$

- Se dice entonces que $X \sim U_{[a,b]}$.
- Todos los valores de X dentro de $[a, b]$ son equiprobables.


f(x)
$\frac{1}{b-a}$
a
b
x

Entonces, si $X \sim U_{[a,b]}$, entonces la f.d.a es:

$$F_X(x) = \begin{cases} \frac{x-a}{b-a}, & \text{si } a \leq x \leq b \\ 0, & \text{en otro caso.} \end{cases}$$

Sea $X \sim U_{[a,b]}$:

$$E[X] = \frac{a+b}{2} \qquad V[X] = \frac{(b-a)^2}{12}$$

Ejemplo

El tiempo de viaje de cierto camión de entrega no supera los 50 minutos ni se da en menos de 20 minutos, pero se da de forma equiprobable en dicho intervalo.

1. ¿Qué proporción de entrega se hace en menos de media hora?
2. Si el conductor tardó más de 25 minutos, ¿qué tan probable es que entregue en menos de 40 minutos?
3. ¿Cuál es el tiempo medio de entrega y su varianza?

Ejercicio

Se escoge un número al azar en el intervalo $[1, 3]$. ¿Cuál es la probabilidad de que el primer dígito al lado derecho del punto decimal sea cinco?

Distribución normal

Esta distribución juega un papel clave en el desarrollo de la inferencia estadística, pues muchas de las herramientas usadas en la toma de decisiones o en el modelamiento estadístico, tienen su fundamento en ésta distribución.

Un gran número de estudios pueden ser aproximados usando una distribución normal. Algunas variables físicas, datos meteorológicos (temperatura, precipitaciones, presión atmosférica, etc), mediciones en organismos vivos, notas o puntajes en pruebas de admisión o de aptitud, errores en instrumentación, proporciones de errores en diversos procesos, etc.

Distribución Normal

Una variable aleatoria continua X tiene una distribución normal si su función de densidad de probabilidad es:

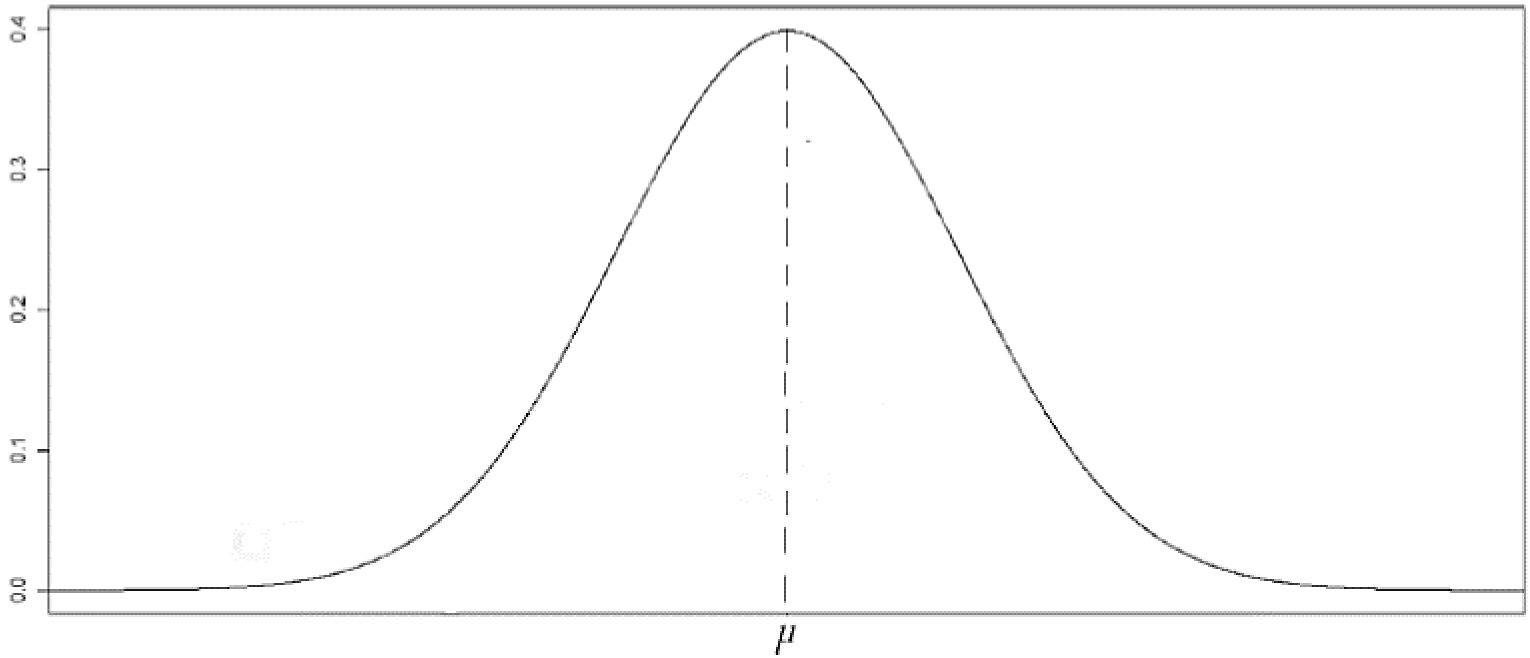
$$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \text{ si } -\infty \leq x \leq \infty$$

$\mu \in \mathbb{R}$ y $\sigma > 0$ son los parámetros de localización y escala.

Se dice entonces que $X \sim N(\mu, \sigma^2)$

Características distribución Normal

- El área bajo la curva Gaussiana entre $(-\infty, \infty)$ es igual a 1.
- La curva es simétrica respecto a μ. El área bajo $[\mu, \infty)$ es 0.5.
- La curva queda completamenta caracterizada al conocer los parámetros de localización y escala.
- Cuanto más grande es σ^2, más *achatada* (amplia) es la gráfica de la función.


0.4
0.3
0.2
0.1
0.0
μ

Media y varianza

Sea $X \sim N(\mu, \sigma^2)$:

Media

$$E[X] = \mu$$

Varianza

$$Var[X] = \sigma^2$$

Función de distribución acumulada

- Como estudiamos anteriormente, la obtenemos por medio de una integral.

- Entonces, si $X \sim N(\mu, \sigma^2)$, entonces la f.d.a es:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dy$$

Suponga que $X \sim N(\mu, \sigma^2)$ y que se quiere hallar $P(x_1 < X < x_2)$. Observe que para hallar esta probabilidad debe resolverse la siguiente integral:

$$P(x_1 < X < x_2) = \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$

Esta integral no puede ser resuelta de manera explícita; se requiere de métodos numéricos para aproximarla. Además, cada vez que se modifique alguno de los parámetros μ o σ^2, se debe calcular de nuevo dicha integral.

Para evitar este problema, se utiliza el cambio de variable $Z = \frac{X-\mu}{\sigma}$. Del cual se obtiene que $dz = \frac{1}{\sigma} dx$.

Con esto se tiene que:

$$
\begin{aligned}
P(x_1 < X < x_2) &= \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&= \int_{\frac{x_1-\mu}{\sigma}}^{\frac{x_2-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz \\
&= P(z_1 < Z < z_2)
\end{aligned}
$$

donde $z_1 = \frac{x_1-\mu}{\sigma}$ y $z_2 = \frac{x_2-\mu}{\sigma}$. Esto indica que cualquier cálculo de probabilidades para una variable aleatoria $N(\mu, \sigma^2)$ puede reducirse al cálculo de probabilidades con una variable aleatoria $N(0, 1)$. El proceso de transformar una v.a normal en una normal estándar se conoce como **Estandarización** o **Normalización**.

Observe que

$$E[Z] = E\left[\frac{X-\mu}{\sigma}\right] = \frac{1}{\sigma}(E[X]-\mu) = \frac{1}{\sigma}(\mu-\mu) = 0$$

$$V[Z] = V\left[\frac{X-\mu}{\sigma}\right] = \frac{1}{\sigma^2}(V[X]-0) = \frac{1}{\sigma^2}(\sigma^2) = 1$$

En la práctica la estanzarización se haría como:

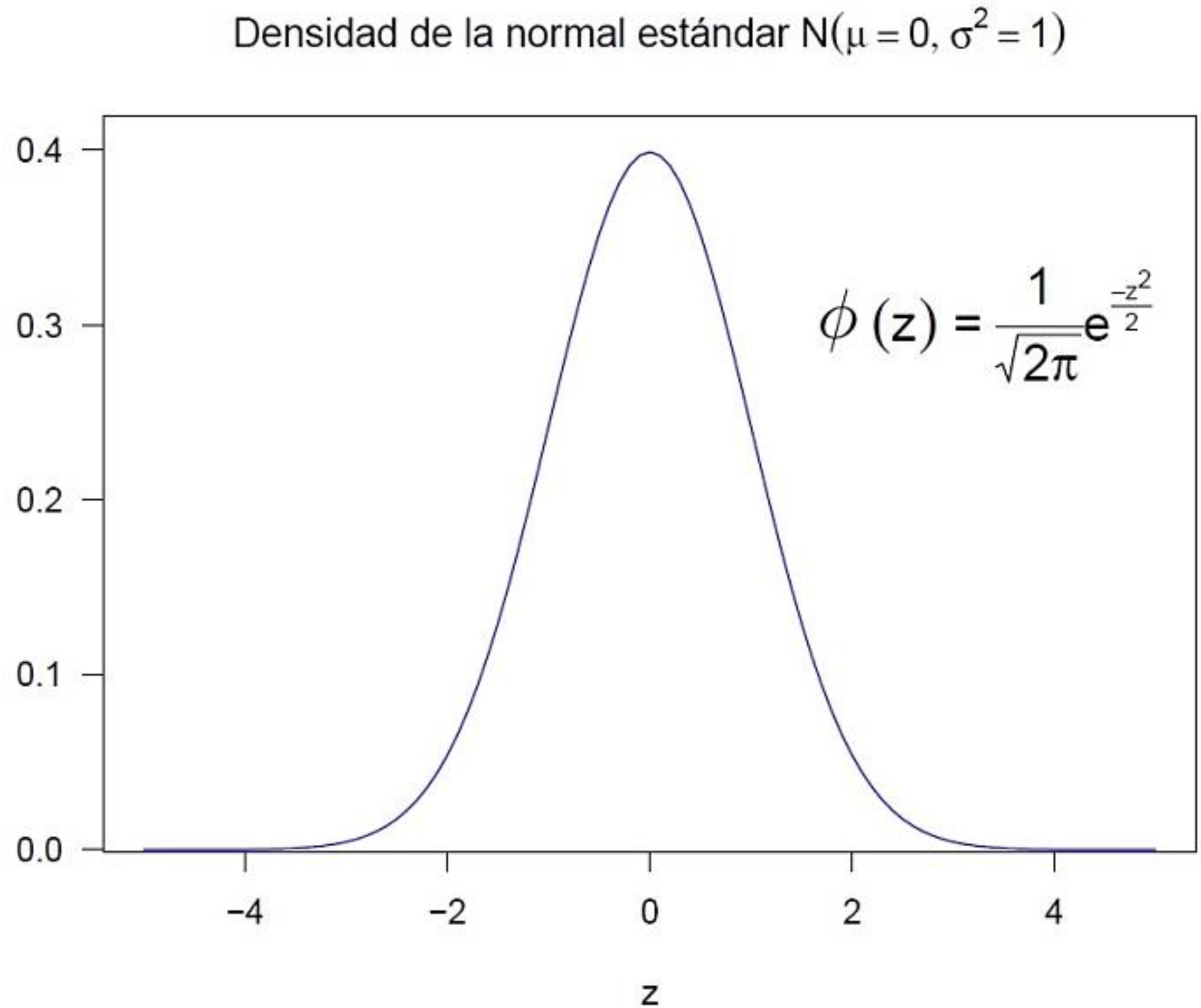
$$\begin{aligned}
P(x_1 < X < x_2) =& P(x_1 - \mu < X - \mu < x_2 - \mu) \\
=& P\left(\frac{x_1 - \mu}{\sigma} < \frac{X - \mu}{\sigma} < \frac{x_2 - \mu}{\sigma}\right) \\
=& P(z_1 < Z < z_2)
\end{aligned}$$

Para el cálculo de probabilidades con una variable $N(0,1)$, la f.d.a. estará dada por:

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

De esta manera se tiene que:

$$P(x_1 < X < x_2) = P(z_1 < Z < z_2) = \Phi(z_2) - \Phi(z_1)$$



Densidad de la normal estándar $N(\mu = 0, \sigma^2 = 1)$
0.4
0.3
0.2
0.1
0.0
−4
−2
0
2
4
z
$\phi(z) = \frac{1}{\sqrt{2\pi}} e^{\frac{-z^2}{2}}$

Sea Z una variable aleatoria tal que $Z \sim N(0\,,\,1)$:

- $P(Z < -z) = P(Z > z)$.
- $P(Z \geq -z) = P(Z \leq z)$.
- $P(-z < Z < 0) = P(0 < Z < z)$.
- $P(Z < 0) = P(Z > 0) = \frac{1}{2}$.
- $P\left(-z_1 < Z < -z_2\right) = P\left(z_2 < Z < z_1\right)$.
- $P(-z \leq Z \leq z) = 2P(Z \leq z) - 1$

Ejemplo

Sea $X \sim N(1, 4)$. Calcule:

1. $P(0 \leq X < 4)$

2. $P(X^2 > 1)$

Ejemplo

Suponga que Z es una variable aleatoria tal que $Z \sim N(0,1)$.
Determine el valor de a para el cuál se cumple:

a. $P(Z < a) = 0.9$
b. $P(Z > a) = 0.05$
c. $P(-1.24 < Z < a) = 0.8$

Ejemplo

El diámetro de un cable electrónico está distribuido normalmente con un diámetro medio de 0.8 pulgadas y una desviación estándar de 0,02. Un cable es defectuoso si su diámetro difiere del promedio en más de 0,025 pulgadas. ¿Qué proporción de cables son defectuosos?

Ejemplo

Se utilizan medidores para rechazar todos los componentes en los que cierta dimensión no esté dentro de la especificación $1,50 \pm d$. Se sabe que esta medida se distribuye normalmente con una media de 1.50 y una desviación estándar de 0,2. Determine el valor d tal que las especificaciones "cubran" 95 % de las mediciones.