1.3. Distribuciones y el TLC

1. Estadística y Probabilidad Básica

Christian E. Galarza

Programa New Dimensions

La Distribución Normal

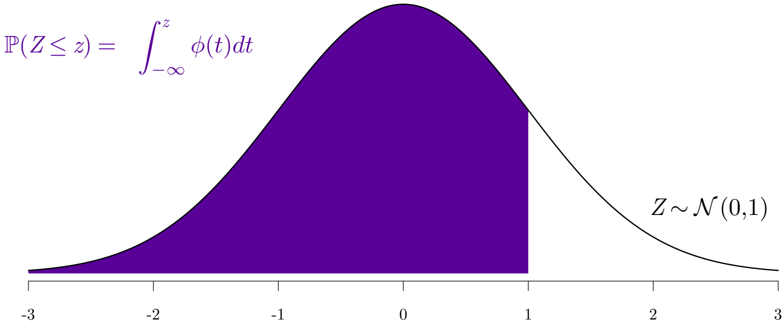
- Es una de las distribuciones más importantes en estadística.
- Es simétrica y tiene forma de campana.
- Se describe completamente por su media μ y desviación estándar σ.
- La notación: $X \sim N(\mu, \sigma^2)$.

Definición Matemática de la Distribución Normal

La función de densidad de probabilidad (PDF) de la distribución normal es:

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

donde μ es la media y σ^2 es la varianza.


$\mathbb{P}(Z \leq z) = \int_{-\infty}^{z} \phi(t)dt$
$Z \sim \mathcal{N}(0,1)$
-3
-2
-1
0
1
2
3

Probabilidades de la Distribución Normal

- Se pueden calcular probabilidades usando la tabla Z o funciones en R.
- Ejemplo: La probabilidad de que $X < x$ si $X \sim N(\mu, \sigma^2)$ se calcula con la función `pnorm` en R.

Estandarización de una Variable Aleatoria Normal

La estandarización implica transformar una variable aleatoria X con media μ y desviación estándar σ a una variable Z con media 0 y desviación estándar 1.

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

- Z es una variable aleatoria normal estándar.
- Estandarizar permite comparar diferentes variables aleatorias normales.
- Las probabilidades de Z se pueden encontrar usando tablas estándar de la distribución normal.

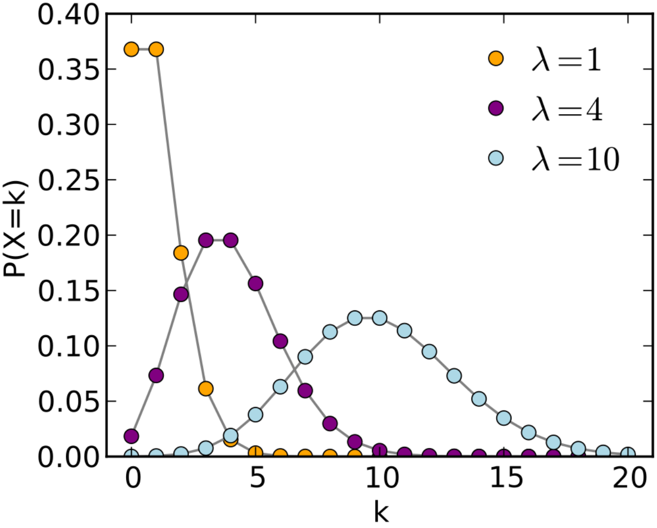
Definición Matemática de la Distribución de Poisson

La función de masa de probabilidad (PMF) de la distribución de Poisson es:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

donde λ es la tasa promedio de ocurrencia.

Distribución de probabilidades para una v.a. Poisson(λ)



Ejemplo: Número de Llamadas en una Hora

Supongamos que el número de llamadas a un centro de atención al cliente en una hora sigue una distribución de Poisson con $\lambda = 10$. Queremos calcular la probabilidad de recibir exactamente 5 llamadas en una hora.
Usando la distribución de Poisson:

$$P(X = 5) = \texttt{dpois(5, lambda = 10)}$$

Ejemplo: Clientes en una Boutique

Supongamos que el número de clientes que ingresan a una boutique en 30 minutos sigue una distribución de Poisson con $\lambda = 8$. Queremos calcular:

- La probabilidad de que exactamente 6 clientes ingresen en 30 minutos.
- La probabilidad de que al menos 10 clientes ingresen en esos 30 minutos.

Usando la distribución de Poisson:

- Probabilidad puntual: $P(X = 6) =$ `dpois(6, lambda = 8)`
- Probabilidad acumulada: $P(X \geq 10) = 1 -$ `ppois(9, lambda = 8)`

El Teorema del Límite Central

- Es tal vez el resultado más importante en la estadística.
- Afirma que la distribución de la suma (o promedio) de una gran cantidad de variables aleatorias independientes e idénticamente distribuidas se aproxima a una distribución normal.

Teorema del Límite Central

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de n observaciones tomadas de una población con media μ y varianza σ^2 (finitas).

Si n es suficientemente grande, la distribución de la suma $S_n = X_1 + X_2 + \ldots + X_n$ (o el promedio) tiende a una distribución normal, independientemente de la forma de la distribución original de la población.
Formalmente,

$$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \xrightarrow{n \to \infty} N(0, 1)$$

donde $\bar{X}$ es la media de la muestra.

!El TLC en la práctica!

Simulador del TLC

Ejercicio: Teorema del Límite Central y Poisson

Supongamos que una empresa recibe en promedio 10 correos electrónicos por hora y que la distribución de los correos sigue una distribución de Poisson. Si se observa la media de correos recibidos durante 50 horas, ¿cuál es la probabilidad de que esa media sea mayor a 11 correos por hora?

Datos:

- X_i Poisson($\lambda = 10$)
- Por el Teorema del Límite Central: $\bar{X}$ $N\left(\mu = \lambda, \sigma^2 = \frac{\lambda}{n}\right)$
- $n = 50$ (horas)

Aplicando:

$$P(\bar{X} > 11) \approx P\left(Z > \frac{11 - 10}{\sqrt{10/50}}\right)$$

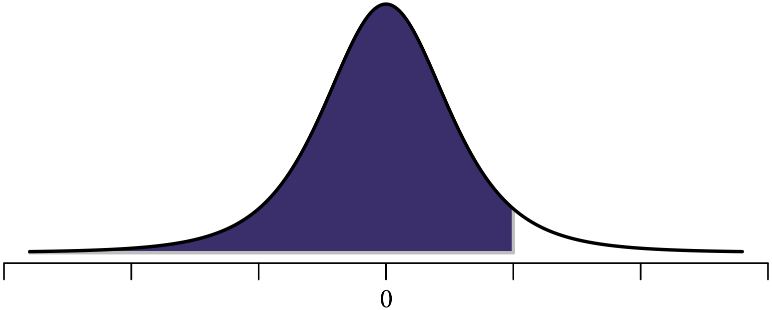
El valor resultante puede encontrarse en una tabla z o mediante software estadístico.

La Distribución t de Student

La distribución t de Student es similar a la normal, pero con colas más pesadas. Se define mediante:

$$f(t|\nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

donde ν son los grados de libertad.


0

Ventas

Supongamos que las ventas diarias de un producto siguen una distribución normal con $\mu = 100$ y $\sigma = 15$. Queremos calcular la probabilidad de que las ventas sean mayores que 120 en un día dado.
Usando la distribución normal:

$$P(X > 120) = 1 - P(X \leq 120) = 1 - \texttt{pnorm(120, mean = 100, sd = 15)}$$

Cálculo de la Probabilidad Normal en R

Podemos calcular la probabilidad en R usando la función `pnorm()`.

```
# Parámetros
mu <- 100  # Media
sigma <- 15 # Desviación estándar
x <- 120  # Valor dado

# Calcular la probabilidad
probabilidad <- 1 - pnorm(x, mean = mu, sd = sigma)
```