Test de Hipótesis

Pablo L. De Nápoli

Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires

Probabilidades y Estadística para Computación
Primer Cuatrimestre de 2023

Un Ejemplo

Problema: (tomado del apunte de A. Bianco- E. Martinez)

Supongamos que el consumo promedio de nafta de los motores utilizados por una empresa automotriz en uno de sus modelos es de 10 litros cada 100 km. Se presenta un proyecto de mejora del motor que produciría una disminución en el consumo pero, por razones de costo, se considera viable el proyecto si la reducción lleva el consumo a un valor menor de 9 litros cada 100 km.

Para estudiar la conveniencia o no de aplicar la mejora a los motores, se aplica esta mejora a una muestra de 25 motores, los cuáles se ponen a funcionar en igualdad de condiciones durante un periodo fijo. El consumo promedio observado es de 8,9 litros cada 100 km. ¿Proveen estos datos evidencia de que vale la pena incorporar la mejora al motor o se deben simplemente al azar?

Se supone que el consumo de nafta de los motores es una v.a. con distribución normal con varianza igual a 1 y que la que la muestra es aleatoria, es decir que los 25 consumos son independientes.

Un ejemplo (2)

Si notamos por $X_1, X_2, \ldots, X_{25}$ a los consumos observados de los motores, lo que estamos suponiendo es que los (X_i) son una muestra aleatoria de la distribución normal $N(\mu, 1)$ con $\mu = 0{,}89$, entonces

$$\overline{X} = \sum_{i=1}^{25} X_i \sim N\left(\mu, \frac{1}{25}\right) \Leftrightarrow Z = \frac{\overline{X} - \mu}{\sqrt{1/25}} \sim N(0,1)$$

Si la media verdadera del consumo en el motor mejorado fuese de 9 litros cada 100 km., ¿cuál es la probabilidad de que una v.a. normal con media 9 y varianza $1/25$ tome un valor igual o menor que el observado, 8,9?

$$P\{\overline{X} \leq 8,9\} = P\left\{Z = \frac{\overline{X} - 9}{1/5} \leq \frac{8{,}9 - 9}{1/5}\right\} = \Phi(-0{,}5) \approx 0,309$$

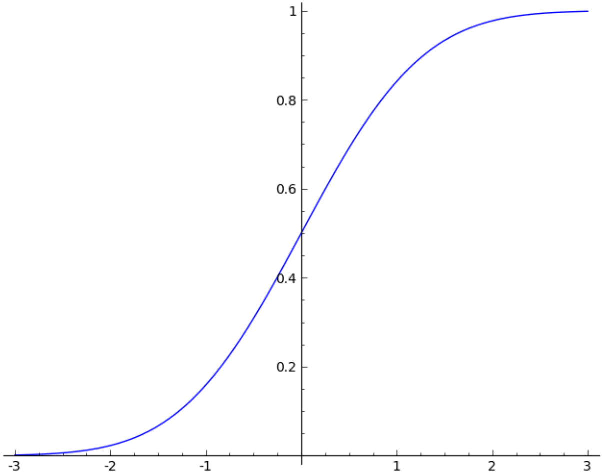
donde Φ es la función de distribución acumulada de la norma estándar. Esta probabilidad se denomina p-valor. Vemos que en este caso no es tan pequeña.

Para calcularlo: En Python: scipy.stats.norm.cdf(-0.5)

en R : pnorm(-0.5)

Recordamos...

La función Φ (distribución acumulada de la normal estándar) tiene el siguiente gráfico



Un ejemplo (3)

En cambio, si el consumo promedio observado hubiese sido $\overline{X} = 8,6$ litros cada 100 km, entonces

$$P\{\overline{X} \leq 8,6\} = P\left\{ Z = \frac{\overline{X} - 9}{1/5} \leq \frac{8{,}6 - 9}{1/5} \right\} = \Phi(-2) \approx 0,023$$

es decir que, en este último caso, hubiese sido muy poco probable que se observase un valor promedio de $8,6$ si la media verdadera es 9.

Un ejemplo (4): Hipótesis Nula e Hipótesis Alternativa

¿Qué es lo que estamos tratando de decidir? Nuestras hipótesis se refieren a la media μ, y se podrían enunciar así:

- H_0: $\mu \geq 9$ litros cada $100 km$. En este caso no se implementa la mejora a los motores
- H_1: $\mu < 9$ litros cada 100 km. En este caso conviene implementar la mejora a los motores.

A la primera hipótesis se la denomina hipótesis nula y se designa H_0. Esta hipótesis implica que no hay efecto, es la hipótesis del status quo, o sea del no cambio respecto a la situación inicial.

La segunda hipótesis se denomina hipótesis alternativa y se designa H_1. Se la suele llamar la hiótesis del investigador.

Definiciones

Definition

Un test es una regla de decisión basada en un estadístico o función de la muestra (en este caso Z) , y en una zona de rechazo, es decir un conjunto de valores para los cuáles se rechaza la hipótesis nula H_0.

¿Cómo se elige la zona de rechazo? Observemos que al tomar una decisión en base a una muestra, podemos cometer dos tipos de error.

	No se rechaza H_0	Se rechaza H_0
H_0 es cierta	OK	Error de tipo I
H_0 no es cierta	Eror de tipo II	OK

$\alpha=$ (máxima) probabilidad de error de tipo I (nivel de significancia del test).

$\beta=$ probabilidad de error de tipo II.

Volviendo al ejemplo...

Volviendo al ejemplo, sabemos que, si H_0 es cierta,

$$Z = \frac{\overline{X} - 9}{1/5} \sim N(0,1)$$

Si queremos que el test tenga nivel de significancia $\alpha = 0{,}05$ rechazaríamos H_0 si

$$Z \leq \Phi^{-1}(0,05) = -1,64$$

Para calcularlo en Ptyhon: scipy.stats.norm.ppf(0.05) en R : qnorm(0.5)

Esta es la zona de rechazo del test de nivel $0,05$. Si observamos un promedio igual a $8,9$, el valor del estadístico Z es $-0,5$ y por lo tanto no se rechaza H_0, mientras que si observamos un promedio igual a $8,6$, el valor del estadístico es -2 y se rechaza H_0.
Si queremos que el test tenga nivel de significación $\alpha = 0,10$, rechazaríamos H_0 si

$$Z \leq \Phi^{-1}(0,1) = -1,28$$

Generalizando esto: Test para la media μ con σ conocida

Suponemos que tenemos una muestra aleatoria $X_1, X_2, \ldots, X_n$ tomada de una población con distribución normal $N(\mu, \sigma^2)$.

Si queremos testear

$$H_0 : \mu = \mu_0$$

frente a

$$H_0 : \mu < \mu_0$$

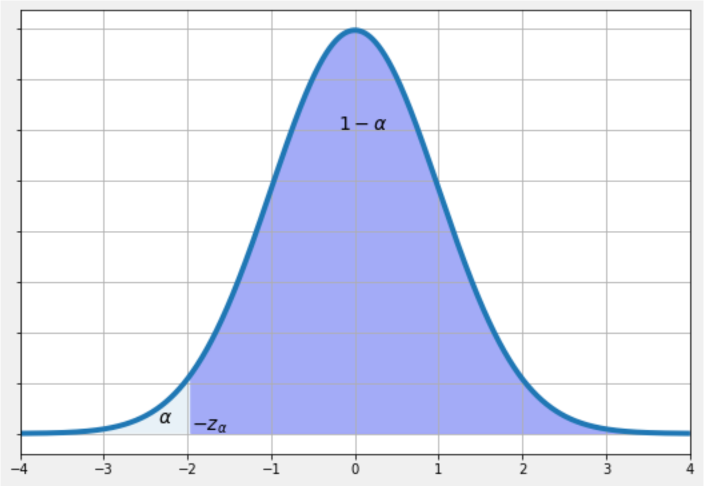
Si $\sigma = \sigma_0$ es conocida, utilizamos el estadístico

$$Z = \frac{\overline{X} - \mu_0}{\sigma_0} \sqrt{n} \sim N(0, 1)$$

Test para la media μ con σ conocida (2)

Elegido un nivel de significancia α, definimos el valor crítico del test

$$-z_\alpha = \Phi^{-1}(\alpha)$$



Test para la media μ con σ conocida (3)

Entonces podemos formular nuestro test así:

Test para $H_0 : \mu = \mu_0$ vs $H_1 : \mu < \mu_0$ con $\sigma = \sigma_0$ conocida

Calculamos el estadístico

$$Z = \frac{\overline{X} - \mu_0}{\sigma_0} \sqrt{n}$$

y rechazamos la hipótesis nula H_0 si

$$Z \leq -z_\alpha = \Phi^{-1}(\alpha)$$

Otro Ejemplo

Ejercicio de *Essential Statistics with Python and R*

Una etiqueta en cierto paquete de cereales indica que el verdadero peso medio de los paquetes es de 16 onzas (lo indicamos por $\mu = 16$). Un grupo de consumidores insiste en que la verdadera media el peso es menor que el peso indicado. Suponga que el peso total de una muestra de 100 cajas es de 1550 onzas, y que se sabe que la desviación estándar de los paquetes es 1,0 onza

(a) Pruebe la hipótesis de que $\mu = 16$ frente a la alternativa sugerida, con un nivel de significancia $\alpha = 0,05$.

Vamos a hacer la hipótesis de que la distribución de los pesos sigue una distribución normal.

Solución de (a)

```
import numpy as np
from scipy.stats import norm
n=100
mu_0= 16
alfa =0.05
sigma= 1
X_barra= 1550/100

Z= (X_barra-mu_0)/sigma * np.sqrt(n)
z_alfa = -norm.ppf (alfa)

print("X_barra=",X_barra)
print("Z=",Z)

if Z <= - z_alfa:
    print("Rechazamos H_0")
else:
    print("No hay suficiente evidencia para rechazar H_0")
```

Resultado del test

Salida del programa

```
Z= -5.0
-z_alfa= -1.6448536269514729
Rechazamos H_0
```

¡Tienen razón los consumidores!

Potencia del Test

El parámetro α controla el error de tipo I (rechazar H_0 cuaando es cierta).
¿Pero cuál es la probabilidad β de cometer un error de tipo II (no rechazar H_0 cuando no es cierta)?
Para cuantificar ambos tipos de errores, introducimos el concepto de función potencia

Definición

La función potencia de un test estadístico $\pi(\mu)$ es la probabilidad de rechazar la hipótesis nula H_0 cuando el verdadero valor del parámetro es μ.

Notamos que

$$\pi(\mu) = \begin{cases} \alpha(\mu) & \text{si} \quad \mu \in H_0 \\ 1 - \beta(\mu) & \text{si} \quad \mu \in H_1 \end{cases}$$

¿Y en nuestro test?

$$\pi(\mu) = P_\mu\{Z \leq -z_\alpha\} = P_\mu\left\{\frac{\overline{X}-\mu_0}{\sigma_0}\sqrt{n} \leq -z_\alpha\right\}$$
$$= P_\mu\left\{\overline{X} \leq \mu_0 - z_\alpha \frac{\sigma_0}{\sqrt{n}}\right\}$$

Pero $\overline{X} \sim N(\mu, \frac{\sigma_0^2}{n}) \Rightarrow X^* = \frac{\overline{X}-\mu}{\sigma_0} \sim N(0,1)$

$$\pi(\mu) = P_\mu\left\{\mu + X^*\frac{\sigma_0}{\sqrt{n}} \leq \mu_0 - z_\alpha\frac{\sigma_0}{\sqrt{n}}\right\} = P_\mu\left\{X^* \leq -z_\alpha + \frac{\mu_0-\mu}{\frac{\sigma_0}{\sqrt{n}}}\right\}$$
$$= \Phi\left(-z_\alpha + \frac{\mu_0-\mu}{\frac{\sigma_0}{\sqrt{n}}}\right)$$

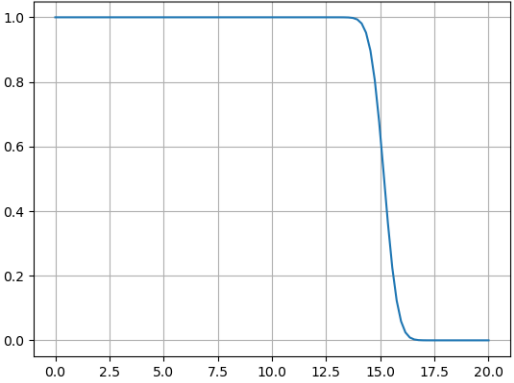
(cuenta en la página 198 del apunte de Bianco-Martínez)

¡En la computadora!

```
import numpy as np
from scipy.stats import norm
import matplotlib.pyplot as plt
n=100
mu_0= 16
sigma_0= 5
alfa= 0.05

z_alfa = -norm.ppf (alfa)

def potencia(mu):
    return norm.cdf(-z_alfa+ (mu_0-mu)/(sigma_0/np.sqrt(n)))

x=np.linspace(start=0,stop=20,num=100)
y=potencia(x)
plt.plot(x,y)
plt.grid()
plt.savefig("../graficos/potencia_del_test.png")
plt.show()
```

Con

```
mu_0= 16
sigma_0= 5
alfa= 0.05
```



Test para $H_0 : \mu = \mu_0$ vs $H_1 : \mu > \mu_0$ con σ conocida

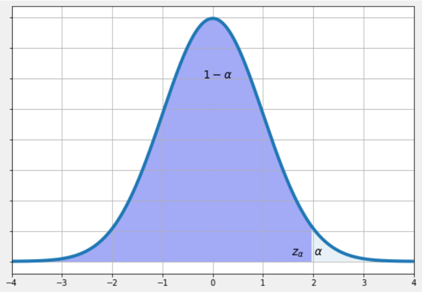
Test para $H_0 : \mu = \mu_0$ vs $H_1 : \mu > \mu_0$ con $\sigma = \sigma_0$ conocida

Calculamos el estadístico

$$Z = \frac{\overline{X} - \mu_0}{\sigma_0} \sqrt{n}$$

y rechazamos la hipótesis nula H_0 si

$$Z > z_\alpha = \Phi^{-1}(1 - \alpha)$$



Test para $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$ con σ conocida

Test para $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$ con $\sigma = \sigma_0$ conocida

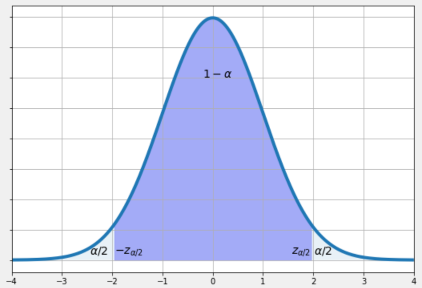
Calculamos el estadístico

$$Z = \frac{\overline{X} - \mu_0}{\sigma_0} \sqrt{n}$$

y rechazamos la hipótesis nula H_0 si

$$|Z| > z_\alpha/2 = \Phi^{-1}(1 - \alpha/2)$$

test bilátero



¿Qué pasa si no conocemos σ? (más realista)

Reemplazamos a σ por su estimador insesgado S dado por

$$S^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X})^2$$

y consideramos test basados en el estadístico

$$T = \frac{\overline{X} - \mu_0}{S} \cdot \sqrt{n} \sim t_{n-1}$$

(distribución t de Student con $n-1$ grados de libertad)

Variantes

Los test resultantes van a ser similares a los que vimos antes, pero utilizando los niveles críticos provenientes de la distribución t de Student con $n-1$ grados de libertad.

Test para $H_0 : \mu = \mu_0$ (o $\mu \leq \mu_0$) vs $H_1 : \mu > \mu_0$ con σ desconocida

Rechazamos la hipótesis nula H_0 si

$$T \geq t_{n-1,\alpha}$$

Test para $H_0 : \mu = \mu_0$ (o $\mu \geq \mu_0$) vs $H_1 : \mu < \mu_0$ con σ desconocida

Rechazamos la hipótesis nula H_0 si

$$T \leq -t_{n-1,\alpha}$$

Test para $H_0 : \mu = \mu_0$ vs $H_1 : \mu \neq \mu_0$ con σ conocida

Rechazamos la hipótesis nula H_0 si

$$|T| \geq t_{n-1,\alpha/2}$$