Domande per verificare la comprensione del significato di errori del II tipo e di potenza.

N.B. Le domande potrebbero contenere informazioni irrilevanti.

Quesito 1. Consideriamo una popolazione normale con $\sigma = 6$ e media $\mu \in \{2, 5, 9\}$ ignota.

Vogliamo testare $H_0 : \mu = 5$ contro $H_A : \mu \in \{2, 9\}$. Preleviamo un campione di rango $n = 16$ e decidiamo di rifiutare H_0 se la media campionaria $\bar{x}$ non appartiene all'intervallo $[4,\ 6]$.

1. Qual è la significatività del test?
2. Qual è la potenza del test?

Esprimere il risutato numerico tramite (solo) le funzioni elencate in calce.

Risposta

$$\alpha \;=\; 1 - \Pr\left(4 \leq \bar{X} \leq 6\right) \text{ dove } \bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \quad \text{con} \quad \mu = 5 \quad \text{e} \quad \frac{\sigma^2}{n} = \frac{9}{4}$$

$$= 1 - \Pr\left(-\frac{2}{3} \leq Z \leq \frac{2}{3}\right) \;=\; \texttt{2 * norm.cdf(-2/3)} \;\texttt{= 0.505}$$ Risposta 1

Per calcolare la potenza assumo H_A valga nel caso più sfaforevole $\mu = 2$

$$1 - \beta \;=\; 1 - \Pr\left(4 \leq \bar{X} \leq 6\right) \text{ dove } \bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \quad \text{con} \quad \mu = 2 \quad \text{e} \quad \frac{\sigma^2}{n} = \frac{9}{4}$$

$$= 1 - \Pr\left(\frac{4}{3} \leq Z \leq \frac{8}{3}\right) \;=\; \texttt{norm.cdf(4/3) + norm.cdf(-8/3)} \;\texttt{= 0.9126}$$ Risposta 2

Quesito 2. Preleviamo un campione di rango $n = 9$ da una popolazione con distribuzione $N(\mu, \sigma^2)$. Sappiamo che la deviazione standard è $\sigma = 3$. La media μ invece potrebbe avere uno qualsiasi valori dei nell'intervallo $[1,\ 4]$.

Vogliamo testare $H_0 : \mu = 1$ contro $H_A : \mu \in (1, 4]$. Fissiamo come significatività $\alpha = 0.1$ otteniamo che per uno z-test a coda superiore la zona di rifiuto per la media campionaria è $[2.282,\ +\infty)$.

1. Nel caso $H_A : \mu \in [3,\ 4]$ qual'è la massima probabilità β di non rigettare H_0 (errore II tipo)?
2. Calcolare la potenza del test con l'effect-size suggerito nel punto precedente.

Esprimere il risutato numerico tramite (solo) le funzioni elencate in calce.

Risposta

Il caso più sfavorevole si ottiene quando $\mu = 3$. Sia $\bar{X} \sim N(3,\ \sigma^2/n)$

$$\beta \;=\; \Pr\left(\bar{X} < 2.282\right) \;=\; \Pr\left(\frac{\bar{X} - 3}{\sigma/\sqrt{n}} \;<\; \frac{-0.718}{\sigma/\sqrt{n}}\right) \;=\; \Pr\left(Z < -0.718\right)$$

β = `norm.cdf(-0.718)` = `0.2364` Risposta 1

Con un effect size $\delta = 2$ la potenza del test è $1 - \beta =$ `1 - norm.cdf(-0.718)` = `0.7636` Risposta 2

Quesito 3. Vogliamo testare $H_0 : \mu = \mu_0$ contro $H_A : \mu > \mu_0$ per una popolazione distribuita normalmente con deviazione standard nota σ. Fissiamo una significatività α e una potenza $1 - \beta$. L'effect-size che ci interessa è δ. Esprimere, in funzione dei parametri che assumiamo noti, le condizioni cui deve soddisfare il rango n del campione.

Risposta

Il rango necessario è il minimo n tale che $\Pr\left(Z < \dfrac{x_\alpha - \mu_0 - \delta}{\sigma/\sqrt{n}}\right) \leq \beta$

dove x_α tale che $\Pr\left(Z \geq \dfrac{x_\alpha - \mu_0}{\sigma/\sqrt{n}}\right) = \alpha$ Risposta

---

Si assuma noto il valore delle seguenti funzioni della libreria `scipy.stats`

`norm.cdf(z)` $= \Pr\big(Z < \mathtt{z}\big)$ per $Z \sim N(0, 1)$

`norm.ppf(`α`)` $= z_\alpha$ dove z_α è tale che $\Pr\big(Z < z_\alpha\big) = \alpha$ per $Z \sim N(0, 1)$