Domande per verificare la comprensione del significato di distribuzione continua (solo caso distribuzione normale). Richiede anche le nozioni di standardizzazione e di media campionaria.

N.B. Alcune domande potrebbero contenere informazioni irrilevanti.

Quesito 1. La variabile aleatoria X ha distribuzione normale con media $\mu = 8$ e deviazione standard $\sigma = 5$

1. Calcolare la probabilità dell'evento $X \in [2, 9]$

2. Calcolare la probabilità che da un campione di rango $n = 16$ si ottenga una media in $[2, 9]$.

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

Risposta

$$\Pr(2 \le X \le 9) = \Pr\left(\frac{2-\mu}{\sigma} \le Z \le \frac{9-\mu}{\sigma}\right) = \Pr\left(Z \le \frac{1}{5}\right) - \Pr\left(Z \le -\frac{6}{5}\right)$$

= `norm.cdf( 1/5 ) - norm.cdf( -6/5 )` Risposta 1

= `norm.cdf( 0.2 ) - norm.cdf( -1.2 ) = 0.464`

$\bar{X}$ v.a. media campionaria

$$\Pr(2 \le \bar{X} \le 9) = \Pr\left(\frac{2-\mu}{\sigma/\sqrt{n}} \le Z \le \frac{9-\mu}{\sigma/\sqrt{n}}\right) = \Pr\left(Z \le \frac{4}{5}\right) - \Pr\left(Z \le -\frac{24}{5}\right)$$

= `norm.cdf( -24/5 ) - norm.cdf( -24/5 )` Risposta 2

= `norm.cdf( 0.8 ) - norm.cdf( -4.8 ) = 0.788`

Quesito 2. La variabile aleatoria X ha distribuzione normale con media $\mu = 9$ e deviazione standard $\sigma = 3$. Per quale ε abbiamo $\Pr\big(5 \le X \le 5+\varepsilon\big) = 0.5$.

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

Risposta

$\Pr\big(5 \le X \le 5+\varepsilon\big)$ cresce al variare di ε, quindi il minimo è quando $\Pr\big(5 \le X \le 5+\varepsilon\big) = 0.5$

$$\Pr\big(5 \le X \le 5+\varepsilon\big) = \Pr\left(\frac{5-\mu}{\sigma} \le Z \le \frac{5-\mu+\varepsilon}{\sigma}\right) = \Pr\left(-\frac{4}{3} \le Z \le \frac{-4+\varepsilon}{3}\right) = 0.5$$

$$\Pr\left(Z \le \frac{-4+\varepsilon}{3}\right) = 0.5 + \Pr\left(Z \le -\frac{4}{3}\right)$$

`(-4 + ε) / 3 = norm.ppf( 0.5 + norm.cdf(-4/3) )`

`ε = 3 * norm.ppf( 0.5 + norm.cdf(-4/3) ) + 4  = 4.692` Risposta

Quesito 3. Abbiamo prelevato vari campioni di una data cultura. Ci interessa selezionare quei campioni che hanno una concentrazione ≤ 3 di una data sostanza. La misura produce risultati che differiscono dal valore corretto per un errore distribuito normalmente con media 0 e deviazione standard 5. Consideriamo la seguente procedura: se la media di 4 misure è ≤ 1 concludiamo che il campione è come desiderato altrimenti lo scartiamo.

Calcolare (nel caso più sfavorevole) la probabilità di scartare erroneamente un campione.

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

Risposta

Il caso più sfavorevole occorre quando la concentrazione vera nel campione è $\mu = 3$

$\bar{X} \sim N(\mu, \sigma^2/n)$ media di 4 misure

$$\Pr\left(\bar{X} \geq 1\right) = \Pr\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{4}} \geq \frac{1-\mu}{\sigma/\sqrt{4}}\right) = \Pr\left(Z \geq -4/5\right)$$

$=$ `1 - norm.cdf(-4/5)` = 0.788 Risposta

Quesito 4. Abbiamo prelevato 5 campioni di una data cultura. Ci interessa selezionare quei campioni che hanno una concentrazione ≤ 3 di una data sostanza. La misura produce risultati che differiscono dal valore corretto per un errore distribuito normalmente con media 0 e deviazione standard 5. Consideriamo la seguente procedura: se la media di 4 misure è ≤ 1 concludiamo che il campione è come desiderato altrimenti lo scartiamo.

Calcolare (nel caso più sfavorevole) la probabilità che nessun campione venga scartato erroreamente.

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

Risposta

Il caso più sfavorevole occorre quando tutti 5 campioni hanno concentrazione $\mu = 3$

$\bar{X} \sim N(\mu, \sigma^2/n)$ media di 4 misure

$$\Pr\left(\bar{X} \geq 1\right) = \Pr\left(\frac{\bar{X}-\mu}{\sigma/\sqrt{4}} \geq \frac{1-\mu}{\sigma/\sqrt{4}}\right) = \Pr\left(Z \geq -4/5\right)$$

$=$ `1 - norm.cdf(-4/5)` probabilità (per un singolo campione) di essere scartato erroneamente

$=$ `( norm.cdf(-4/5) )**5` = 0.00043 Risposta

Quesito 5. Da una popolazione con distribuzione normale con media μ ignota e deviazione standard 57 estraiamo un campione di 25 individui. Qual è la probabilità che la media campionaria risulti $> \mu + 5$?

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

Risposta

$n = 25$ dimensione del campione

$\varepsilon = 5$

$\sigma = 57$ deviazione standard della popolazione

$\sigma/\sqrt{n}$ errore standard (deviazione standard della media)

$P\left(Z > \dfrac{\varepsilon}{\sigma/\sqrt{n}}\right) \ = \ P\left(Z > \dfrac{25}{57}\right) \ =$ `1 - norm.cdf(25/57) = 0.330` Risposta

---

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

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