# Esercitazione: Statistica
# 11 Maggio 2017
# Esercizi e Soluzioni

Università Mediterranea di Reggio Calabria

**A)** La polizia arresta Tizio perché sospettato di aver commesso un furto. Un giudice deve decidere se condannarlo o meno, deve quindi valutare la probabilità di colpevolezza di Tizio. Si indichino con $C$ ed $I$, rispettivamente colpevole e innocente.

*a*) Prima di aver esaminato le prove, il magistrato non sa nulla circa la colpevolezza o innocenza di Tizio. Si traduca il "non sapere nulla" in termini probabilistici. Quali sono i valori delle probabilità a priori $P(C)$ e $P(I)$?

*b*) Vi è un testimone che dichiara di aver visto Tizio commettere il furto; denotiamo con $R$ l'evento "Tizio è stato riconosciuto". Il magistrato sa che con situazioni similari, in cui vi è la stessa evidenza empirica, la probabilità che il colpevole sia riconosciuto correttamente è del 75%. Sa anche che la probabilità di falsi riconoscimenti è del 10%. Denotando con $\bar{R}$ l'evento "Tizio non è stato riconosciuto", quali sono le due probabilità $P(\bar{R} \mid C)$ e $P(R \mid I)$?

*c*) Utilizzando il teorema di Bayes, valutare la probabilità che, data la testimonianza, Tizio sia colpevole o innocente.

**B)** Si supponga che $(X_1, X_2)$ sia un campione casuale di osservazioni da una popolazione con media $\mu$ e varianza $\sigma$. Si considerino i seguenti stimatori:

$$X = \frac{1}{2}X_1 + \frac{1}{2}X_2; \quad Y = \frac{1}{4}X_1 + \frac{3}{4}X_2; \quad Z = \frac{1}{3}X_1 + \frac{2}{3}X_2.$$

*a*) Dimostrare che tutti e tre gli stimatori sono non distorti.

*b*) Qual è lo stimatore più efficiente?

**C)** Una azienda produttrice di cereali vuole monitorare il peso di ogni scatola riempita. Il peso riportato in etichetta è di 500g. I pesi relativi alle scatole, di un campione casuale di 100, hanno una media di 520.948g e varianza $89,3025$.

*a*) Determinare l'intervallo di confidenza, a livello 90%, per la media dei pesi delle scatole.

*b*) Senza svolgere i calcoli, si stabilisca se un intervallo di confidenza per la media, a livello 99%, sia più ampio, meno ampio o della stessa ampiezza di quello trovato al punto *a*).

*c*) Se e solo se si riscontra una differenza di ampiezza al punto *b*), dovuta al cambiamento dell'intervallo di confidenza, si calcoli come dovrebbe essere modificata questa ampiezza per mantenerla uguale nei due casi.

**D** Una azienda produttrice di barrette energetiche produce barrette in media lunghe 2cm e deviazione standard 0.06cm. Un campione casuale di 9 misurazioni ha determinato una media campionaria di 1.95cm. Con un livello di significatività $\alpha = 0.05$, verificare se la media campionaria sia semplicemente un valore anomalo o se suggerisca una modifica produttiva. Si calcoli il $p$-value, sapendo che per un test bilaterale il valore deve essere raddoppiato.

## Soluzioni

**A)** *a*) L'opinione iniziale del Magistrato richiede che $P(I) = P(C)$. Si noti che $P(I) + P(C) = 1$ e $P(I) = P(C) = \frac{1}{2}$

*b*) Le probabilità sono

$$P(\bar{R} \mid C) = 1 - P(R \mid C) = 1 - 0.75 = 0.25$$

$$P(R \mid I) = 0.10$$

*c*) Usando il teorema di Bayes

$$P(C \mid R) = P(C) \cdot \frac{P(R \mid C)}{P(R)} = \frac{1}{2}\frac{0.75}{P(R)}$$

$$P(I \mid R) = P(I) \cdot \frac{P(R \mid I)}{P(R)} = \frac{1}{2}\frac{0.10}{P(R)}$$

Dunque

$$P(C \mid R) = \frac{75}{85} = 0.8824$$

$$P(I \mid R) = \frac{10}{85} = 0.1176$$

**B)** *a*)

$$\mathbb{E}(\bar{X}) = \frac{\mu}{2} + \frac{\mu}{2} = \mu$$

$$\mathbb{E}(Y) = \frac{\mu}{4} + \frac{3\mu}{4} = \mu$$

$$\mathbb{E}(Z) = \frac{\mu}{3} + \frac{2\mu}{3} = \mu$$

*b*)

$$\mathrm{Var}(\bar{X}) = \frac{1}{4}\mathrm{Var}(X_1) + \frac{1}{4}\mathrm{Var}(X_2) = \frac{\sigma^2}{2}$$
$$\mathrm{Var}(Y) = \frac{5\sigma^2}{8}$$
$$\mathrm{Var}(Z) = \frac{5\sigma^2}{9}$$

Dunque $\mathrm{Var}(\bar{X}) < \mathrm{Var}(Y)$, $\mathrm{Var}(\bar{X}) < \mathrm{Var}(Z)$. Il più efficiente è $\bar{X}$.

**C)** *a*)

$$520.948 \pm 1.645\frac{9.45}{\sqrt{100}} \begin{cases} (+)522.502525 \\ (-)519.393475 \end{cases}$$

*b*) Più ampio.

*c*)

$$2 \times 1.645\frac{9.45}{10} = 2 \times 2.58\frac{9.45}{\sqrt{n}}$$

$n \approx 246$

**D)** $H_0 : \mu = 2; H_1 : \mu \neq 2$. Si ha che

$$Z = \frac{1.95 - 2.0}{0.06/\sqrt{9}} = -2.50$$

Dato che $z_{\frac{0.05}{2}} = 1.96$, $-2.50 < 1.96$ per cui si rifiuta $H_0$ concludendo che il processo produttivo deve essere modificato.

Per trovare il $p$-value, si ha che

$$F(-2.50) = 1 - P(Z < 2.50) = 0.00621$$

da cui

$$p\text{-value} = 2 \times 0.00621 = 0.0124.$$

Un livello di significatività superiore all'1.24% avrebbe significato il rifiuto della ipotesi nulla.