



DOTTORATO DI RICERCA IN METODOLOGIA DELLE SCIENZE SOCIALI
Introduzione all'inferenza statistica
a.a. 2008-2009

Esercizi del 23 ottobre 2008

Esercizio 1

Su 432 soggetti intervistati 234 dichiarano di aver visto su una rivista la pubblicità di Radio GaGa.

- Sottoporre a verifica l'ipotesi nulla "La pubblicità è stata vista da non più della metà della popolazione obiettivo": determinare il p -value e l'esito del test ai livelli 1%, 5% e 20%.
- Calcolare la potenza del test assumendo $\alpha=0.07$ e una proporzione pari a 0.55.

Esercizio 2

Un'azienda deve decidere se modificare la limonata che produce. Su un campione di 300 consumatori che hanno assaggiato sia la versione attuale che quella modificata, 162 dichiarano di preferire la versione modificata.

- Sottoporre a verifica l'ipotesi che i consumatori che preferiscono la versione modificata siano non più del 50%: calcolare il p -value e determinare l'esito del test ai livelli di significatività 1%, 5% e 10%.
- Fissando la probabilità di errore di prima specie al 10%, qual è la probabilità di errore di seconda specie nel caso in cui la vera percentuale di consumatori che preferiscono la versione modificata sia 55%?

Esercizio 3

Dall'esame emocromocitometrico su un gruppo di 4 pazienti risultano i seguenti valori di emoglobina in g/dL: (14.5, 17.3, 13.9, 15.2). Lo stesso esame su un secondo gruppo di 7 pazienti ha fatto registrare una media di 14.3 con deviazione standard campionaria 0.89.

- Stabilire con un test al livello 10% se le varianze nei due gruppi si debbano ritenere uguali
- Costruire un intervallo di confidenza al 99% per la differenza tra le medie

Esercizio 4

In una indagine relativa alla mensa universitaria, 75 studenti su 130 intervistati si dicono soddisfatti.

- Sottoporre a verifica l'ipotesi che gli studenti soddisfatti siano non più del 50%: calcolare il p -value e determinare l'esito del test ai livelli di significatività 1%, 5% e 10%.
- Fissando la probabilità di errore di prima specie al 5%, qual è la probabilità di errore di seconda specie nel caso in cui la vera percentuale di studenti soddisfatti sia 60%?

Esercizio 5

Il modello di regressione lineare $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$ viene utilizzato con un campione di 20 unità in cui la variabile X ha media -8 e varianza (non corretta) pari a 25, la variabile Y ha media 15 e varianza (non corretta) pari a 121 e il coefficiente di correlazione lineare è -0.7 . Calcolare

- la stima di minimi quadrati di β_1 e il corrispondente intervallo di confidenza al livello 95%
- l'intervallo di previsione di Y al livello 95% in corrispondenza di $X = -2$

Soluzione

Esercizio 1

proporzione campionaria = 0.5417

- statistica test $z = 1.7395$, $p\text{-value} = 0.0410$
- $\gamma = P(Z > z_{0.93} [\pi_0(1-\pi_0)/\pi_1(1-\pi_1)]^{1/2} + (\pi_0 - \pi_1) / [\pi_1(1-\pi_1)/n]^{0.5} = P(Z > 1.48 [0.5(1-0.5)/0.55(1-0.55)]^{1/2} + (0.50-0.55) / [0.55(1-0.55)/432]^{0.5} = P(Z > -0.6015) = 0.7262$

Esercizio 2

- $H_0: \pi \leq 0.50$ vs. $H_1: \pi > 0.50$, proporzione campionaria=0.54, statistica test=1.3856, $p\text{value}=0.0829$, per cui si rifiuta H_0 al livello 10%, ma non si rifiuta al livello 5% e 1%.
- $\beta = P(Z \leq z_{0.10} [\pi_0(1-\pi_0)/\pi_1(1-\pi_1)]^{1/2} + (\pi_0 - \pi_1) / [\pi_1(1-\pi_1)/n]^{0.5} = P(Z \leq 1.2816 [0.5(1-0.5)/0.55(1-0.55)]^{1/2} + (0.50-0.55) / [0.55(1-0.55)/300]^{0.5} = P(Z > -0.6015) = 0.7262$
- $P(Z \leq -0.4528) = 0.3254$

Esercizio 3

Per il primo gruppo di pazienti: media 15.225, varianza 2.1958

- Rapporto tra varianze 2.7721, valore critico inf 0.1118, valore critico sup 4.7571 \Rightarrow le varianze non sono significativamente diverse
- varianza ponderata 1.2600, s.e. 0.7036, valore critico 3.2498, semintervallo 2.2866, differenza tra medie 0.9250, intervallo [-1.3616 3.2116]

Esercizio 4

- $H_0: \pi \leq 0.50$ vs. $H_1: \pi > 0.50$, proporzione campionaria=0.5769, statistica test=1.7541, $p\text{value}=0.0397$, per cui si rifiuta H_0 al livello 10% e 5%, ma non si rifiuta al livello 1%.
- $\beta = P(Z \leq z_{0.05} [\pi_0(1-\pi_0)/\pi_1(1-\pi_1)]^{1/2} + (\pi_0 - \pi_1) / [\pi_1(1-\pi_1)/n]^{0.5} = P(Z \leq 1.645 [0.5(1-0.5)/0.60(1-0.60)]^{1/2} + (0.50-0.60) / [0.60(1-0.60)/130]^{0.5} = P(Z \leq -0.6486) = 0.2583$

Esercizio 5

$$\hat{\beta}_1 = -0.7 \times (11/5) = -1.54$$

- $s_{Y, \text{res}}^2 = s_Y^2 (1 - r_{XY}^2) = 121 \times (1 - 0.49) = 61.71$

$$\hat{\beta}_1 \pm t_{18,025} \cdot \sqrt{s_{Y, \text{res}}^2 / (n \cdot s_X^2)} = 1.54 \pm 2.101 \sqrt{61.71 / (20 \times 25)} = [-2.2781; -0.8019]$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} = 2.68 \text{ e } X_* = -2 \Rightarrow Y_* = \hat{\beta}_0 + \hat{\beta}_1(-2) = -5.02$$

- $Y_* \pm t_{18,025} \cdot s_{Y, \text{res}}^2 \sqrt{1 + (1/n) + (x_* - \bar{x})^2 / (n \cdot s_X^2)} = [-22.5024; 12.4624]$