

## LM-88 SOCIOLOGIA E RICERCA SOCIALE

### Metodi Statistici per la Ricerca Sociale

#### Confronto tra due gruppi

1. In una indagine sul lavoro femminile sono state rilevate le ore lavorate giornalmente di un campione di 37 lavoratrici residenti in Toscana, e di un campione di 25 lavoratrici residenti in Lombardia:

| Regione   | Media campionaria | Varianza campionaria (corretta) | Dimensione campionaria |
|-----------|-------------------|---------------------------------|------------------------|
| Toscana   | 5.5               | 8                               | 37                     |
| Lombardia | 6.5               | 9                               | 25                     |

Verificare se le osservazioni del campione possono avvalorare l'ipotesi che in Lombardia il numero di ore lavorate giornalmente da una donna è una maggiore che in Toscana, al livello di significatività  $\alpha = 0.05$ . (Assumere che le ore lavorate giornalmente nelle due popolazioni abbiano distribuzione normali con uguali varianze  $\sigma_1^2 = \sigma_2^2 = \sigma^2$ )

**Soluzione:** Ipotesi:  $H_0 : \mu_L = \mu_T$  versus  $H_a : \mu_L > \mu_T$

Regione critica:  $RC_{0.05} = T > 1.6706$ .

Valore osservato della statistica test:  $T^{oss} = 1.333$

Decisione: Non si rifiuta  $H_0$  al livello di significatività del 5%

2. Si supponga di voler comparare la durata media delle lampadine prodotte da 2 fabbriche, e di disporre delle seguenti informazioni campionarie:

| Regione    | Media campionaria | Varianza campionaria (corretta) | Dimensione campionaria |
|------------|-------------------|---------------------------------|------------------------|
| Fabbrica A | 100               | 48400                           | 200                    |
| Fabbrica B | 150               | 8000                            | 160                    |

- (a) Si sottoponga a test l'ipotesi di uguaglianza tra le medie, al livello di significatività  $\alpha = 0.01$ .  
(b) Calcolare e interpretare il  $p$ -valore per il test al punto (a)  
(c) Costruire un intervallo di confidenza al livello di confidenza del 99% per la differenza tra medie e confrontare il risultato con il risultato del test.

**Soluzione:**

- (a) Ipotesi:  $H_0 : \mu_A = \mu_B$  versus  $H_a : \mu_A \neq \mu_B$

Regione critica:  $RC_{0.01} = Z < -2.58$  oppure  $Z > 2.58$ .

Valore osservato della statistica test:  $Z^{oss} = -2.93$

Decisione: Si rifiuta  $H_0$  al livello di significatività del 1%.

- (b)  $p$ -valore = 0.00339

- (c)  $IC_{0.99} = (-94.087; -5.913)$ .

3. Un medico sportivo utilizza un test per misurare la condizione fisica degli atleti. Si considerino due campioni casuali di giocatori di calcio, formati rispettivamente da  $n_1 = 25$  titolari  $n_2 = 26$  riserve. Su tali campioni risulta un punteggio medio a questo test pari a  $\bar{y}_1 = 47.5$  con deviazione standard campionaria  $s_1 = 4.8$  e  $\bar{y}_2 = 37.5$  con deviazione standard campionaria  $s_2 = 5.1$ .
- (a) Verificare l'ipotesi di uguaglianza tra le medie dei giocatori titolari e delle riserve, a un livello  $\alpha = 0.01$ . Si supponga che il punteggio del test sulla condizione fisica si distribuisca secondo una Normale e che nelle due popolazioni il punteggio al test abbia uguale varianza.
  - (b) Interpretare il  $p$ -valore per il test svolto al punto (a) il quale è uguale a  $p = 0.000$
  - (c) Costruire un intervallo di confidenza al livello di confidenza del 99% per la differenza tra medie.

**Soluzione:**

- (a) Ipotesi:  $H_0 : \mu_1 = \mu_2$  versus  $H_a : \mu_1 \neq \mu_2$ .

Regione critica:  $RC_{0.01} = T < -2.68$  oppure  $T > 2.68$ .

Valore osservato della statistica test:  $T^{oss} = 7.204$

Decisione: Si rifiuta  $H_0$  al livello di significatività del 1%.

- (b) Forte evidenza contro  $H_0$

- (c)  $IC_{0.99} = (6.28; 13.72)$ .

4. Tra i visitatori di una libreria, la proporzione di chi acquista almeno un libro è risultata essere, sulla base di un campione di 500 clienti rilevati nell'arco di una settimana, il 30%. Per promuovere le vendite, viene lanciata una campagna promozionale, e durante la settimana della campagna si osserva un campione casuale di 400 visitatori. Tra di essi la proporzione di coloro che acquista almeno un libro è pari al 41%. Sottoporre a test l'ipotesi che la proporzione di visitatori che acquista libri è aumentata nella settimana della campagna ( $\alpha = 0.05$ ). Calcolare e interpretare il  $p$ -valore.

**Soluzione:** Ipotesi:  $H_0 : \pi_1 = \pi_2$  versus  $H_a : \pi_2 > \pi_1$  ( $\pi_1$  = proporzione di libri venduti nella settimana senza campagna;  $\pi_2$  = proporzione di libri venduti nella settimana della campagna).

Regione critica:  $RC_{0.05} = Z > 1.645$ .

Valore osservato della statistica test:  $Z^{oss} = 3.44$

Decisione: Si rifiuta  $H_0$  al livello di significatività del 5%.

$p$ -valore = 0.00029. Forte evidenza contro  $H_0$

5. In due regioni, A e B, sono estratti due campioni indipendenti di famiglie, rispettivamente di 850 e 400 unità. Nel primo campione si osservano 220 famiglie monoreddito, e nel secondo 70. Sottoporre a test l'ipotesi (a un livello  $\alpha = 0.05$ ) che le proporzioni di famiglie monoreddito nelle due regioni siano uguali, contro quella che nella regione A tale proporzione sia maggiore che nella regione B. Calcolare e interpretare il  $p$ -valore.

**Soluzione:** Ipotesi:  $H_0 : \pi_A = \pi_B$  versus  $H_a : \pi_A > \pi_B$ .

Regione critica:  $RC_{0.05} = Z > 1.645$ .

Valore osservato della statistica test:  $Z^{oss} = 3.275$

Decisione: Si rifiuta  $H_0$  al livello di significatività del 5%.

$p$ -valore = 0.00052. Forte evidenza contro  $H_0$

6. A due campioni di donne viene chiesta l'opinione circa un certo metodo contraccettivo. La numerosità dei due campioni e la proporzione di donne che ha espresso opinione favorevole su quel metodo contraccettivo è indicata nella seguente tabella:

|                                     | Proporzione di<br>donne favorevoli | Dimensione<br>campionaria |
|-------------------------------------|------------------------------------|---------------------------|
| Campione A (Donne tra 20 e 30 anni) | 0.75                               | 100                       |
| Campione B (Donne tra 40 e 50 anni) | 0.65                               | 100                       |

Verificare l'ipotesi nulla che la proporzione di donne favorevoli sia la stessa tra le ventenni e le quarantenni, contro l'ipotesi alternativa che la proporzione sia inferiore tra le quarantenni. Si consideri un livello di significatività  $\alpha = 0.05$ . Calcolare e interpretare il  $p$ -valore.

**Soluzione:** Ipotesi:  $H_0 : \pi_B - \pi_A = 0$  versus  $H_a : \pi_B - \pi_A < 0$ .

Regione critica:  $RC_{0.05} = Z < -1.645$ .

Valore osservato della statistica test:  $Z^{oss} = -1.54$

Decisione: Non si rifiuta  $H_0$  al livello di significatività del 5%.

$p$ -valore = 0.06178. Non c'è una forte evidenza contro  $H_0$

7. Si vuole rispondere alla seguente domanda: "Fare jogging regolarmente, porta a ridurre il battito cardiaco?" A tale fine, si organizza un esperimento con 8 volontari che non hanno mai praticato jogging e che ora lo faranno per un mese. I dati raccolti sono rappresentati in tabella:

| Soggetto | 1  | 2  | 3  | 4   | 5  | 6  | 7  | 8  |
|----------|----|----|----|-----|----|----|----|----|
| Prima    | 74 | 86 | 98 | 102 | 78 | 84 | 79 | 70 |
| Dopo     | 70 | 85 | 90 | 110 | 71 | 80 | 69 | 74 |

Fare il test di ipotesi per verificare se fare jogging porta ad una riduzione del battito cardiaco (Fissare  $\alpha = 0.05$ ).

**Soluzione:** Ipotesi:  $H_0 : \mu_D = 0$  versus  $H_a : \mu_D < 0$  ( $\mu_D = \mu_{dopo} - \mu_{prima}$ ).

Regione critica:  $RC_{0.05} = T < -1.8946$ .

Valore osservato della statistica test:  $T^{oss} = -1.263$

Decisione: Non si rifiuta  $H_0$  al livello di significatività del 5%.

8. Una ditta che pubblicizza un prodotto dietetico assicura che dopo 2 settimane di trattamento si ottiene una riduzione media di peso superiore a 5 Kg. Ad un campione casuale di 25 persone viene somministrato il prodotto e dopo due settimane si osserva una riduzione media di peso di 5.2 Kg, con  $s = 12.5$ .
- (a) Supponendo che la riduzione di peso sia una variabile aleatoria Normale, verificare l'ipotesi nulla che la riduzione sia di almeno 5 kg come pubblicizzato, con un livello di significatività  $\alpha = 0.05$ .
- (b) Si potrebbe accettare l'ipotesi nulla a un livello  $\alpha = 0.01$ ?

**Soluzione:**

(a) Ipotesi:  $H_0 : \mu_D = 0$  versus  $H_a : \mu_D < 0$  ( $\mu_D = \mu_{dopo} - \mu_{prima}$ ).

Regione critica:  $RC_{0.05} = T < -1.7109$ .

Valore osservato della statistica test:  $Z^{oss} = -2.08$

Decisione: Si rifiuta  $H_0$  al livello di significatività del 5%.

(b) Regione critica:  $RC_{0.01} = T < -2.4922$ .

Decisione: Non si rifiuta  $H_0$  al livello di significatività del 1%.

9. Si considerino due campioni di due diversi tipi di fusibili elettrici che sono sottoposti a un voltaggio crescente fino a quando fondono. Il voltaggio massimo sopportato prima di fondere e le relative varianze (che si suppongono note) sono riportate nella tabella seguente:

| Fusibili   | Voltaggio medio sopportato<br>prima della fusione | Varianza nella<br>popolazione | Dimensione<br>campionaria |
|------------|---------------------------------------------------|-------------------------------|---------------------------|
| Fusibili A | 42.786                                            | 50                            | 14                        |
| Fusibili B | 55.833                                            | 100                           | 12                        |

Quanto è plausibile l'affermazione che il tipo B di fusibili possa sopportare un voltaggio superiore a quello sopportato dai fusibili di tipo A? Effettuare un opportuno test di significatività (al livello  $\alpha = 0.05$ ). Calcolare e interpretare il  $p$ -valore

**Soluzione:** Ipotesi:  $H_0 : \mu_B - \mu_A = 0$  versus  $H_a : \mu_B - \mu_A > 0$

Regione critica:  $RC_{0.05} = Z > 1.96$ .

Valore osservato della statistica test:  $Z^{oss} = 3.78$

Decisione: Si rifiuta  $H_0$  al livello di significatività del 5%.

$p$ -valore = 0.00008. Forte evidenza contro  $H_0$ .

10. In due sondaggi pre-elettorali effettuati in due città si osservano i seguenti risultati. Città A: su un campione di 1000 unità il 29% ha dichiarato di votare il partito  $D$ . Città B: su un campione di 800 unità il 33% ha dichiarato di votare il partito  $D$ . Costruire un intervallo di confidenza al 99% per la differenza tra le proporzioni dei votanti per il partito  $D$  nelle due città.

**Soluzione:**  $IC_{0.95}(\pi_B - \pi_A) (-0.0167; 0.0967)$

**Esercizi dal libro di testo: Capitolo 7**

7.3   7.4   7.5   7.7   7.8   7.9   7.10   7.14   7.15   7.19   7.20   7.21   7.22   7.23   7.24  
 7.25   7.27   7.32   7.34   7.35(a)   7.59   7.60   7.61   7.62   7.63