

Soluzione HW 1

Modelli Statistici a.a. 2019/20

October 14, 2019

Le istruzioni di STATA per fare i calcoli si trovano nel file `HW1.do`

Esercizio 1

Prendete i dati sui punteggi al test di ingresso ed effettuate il test di uguaglianza delle medie dei punteggi dei maschi e delle femmine. Fornite il numeratore e il denominatore della statistica t di Student e la decisione (accetto/ rifiuto/livello).

I dati sono stati raccolti su due campioni indipendenti di $n_1 = 615$ maschi e $n_2 = 513$ femmine. La statistica test è $t_{oss} = \frac{\hat{\Delta}}{SE(\hat{\Delta})} = -7.0480$.

- Il NUMERATORE di t è la differenza tra medie campionarie $\hat{\Delta} = \bar{y}_1 - \bar{y}_2 = 12.43177 - 14.06179 = -1.630015$.
- Il DENOMINATORE di t e l'errore standard della differenza $SE(\hat{\Delta}) = s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 0.2312729$, dove $s = \sqrt{S^2}$.
- La varianza combinata è $S^2 = \frac{1}{n_1+n_2-2} ((n_1-1)s_1^2 + (n_2-1)s_2^2) = ((512 \times 3.847207^2 + 614 \times 3.884929^2)/(513 + 615 - 2)) = 14.96$, da cui si ottiene $s = \sqrt{S^2} = 3.8678222$

Per decidere se accettare o rifiutare $H_0 : \Delta = 0$ contro l'alternativa $H_0 : \Delta \neq 0$ dobbiamo calcolare il p -value $p_{oss} = \times P(T > |t_{oss}|) = 3.157e - 12$, con $T \sim t_{1126}$ (in alternativa possiamo utilizzare l'approssimazione normale). Poiché il p -value è minore di 0.01 e il test è ALTAMENTE significativo: possiamo dire che c'è evidenza empirica contro l'ipotesi nulla, oppure che l'ipotesi nulla è rifiutata ad un livello del test $\alpha = 0.01$.

Esercizio 2

Prendete i dati del file `occ.txt` (allegato: dati sui salari). I dati sono i salari (in sterline annue) di 72 offerte di lavoro sul *Guardian* scelti casualmente (Aprile del

1992). Ci sono due tipi di lavoro: (1) creativo, media, e marketing e (2) istruzione. I salari medi sono significativamente diversi? Fornite il numeratore e il denominatore della statistica t di Student e la decisione (accetto/ rifiuto/livello).

Si procede in modo analogo all'esercizio precedente.

- Statistica $t_{oss} = -2.13$

– NUMERATORE: differenza tra le medie $\hat{\Delta} = 17410 - 19817.74 = -2407.736$

– DENOMINATORE: $SE(\hat{\Delta}) = s\sqrt{(1/n_1 + 1/n_2)} = 1130.408$ dove $S = \sqrt{((n_1 - 1) * S_1^2 + (n_2 - 1) * S_2^2)/(n - 2)} = 6782.4504$

Per decidere se accettare o rifiutare $H_0 : \Delta = 0$ dobbiamo calcolare il p -value: $2 \times Pr(T > |t|) = 0.0349$ considerando $T \sim t_{142}$. Il p -value < 0.05 e quindi il test è significativo: c'è evidenza empirica contro l'ipotesi nulla, oppure l'ipotesi nulla è rifiutata ad un livello del test $\alpha = 0.05$.