

Scuola di Scienze Politiche
Compito di Metodi Statistici per la Ricerca Sociale
12 Settembre 2018

Esercizio 1

A un campione di 31 soggetti in pensione viene somministrato un test per valutare le loro capacità cognitive. La seguente tabella mostra alcune statistiche descrittive del punteggio al test (su scala 0 – 100) in gruppi di soggetti classificati rispetto al titolo di studio:

Titolo di studio	Statistiche descrittive		
	Numero di osservazioni	Media	Varianza
Media inferiore o meno	6	70	20.25
Media superiore	16	76	25
Diploma universitario o più	9	80	25

- (a) Calcolare la devianza tra gruppi e la devianza entro gruppi
- (b) Verificare l'ipotesi che non ci sia differenza tra il punteggio medio al test nelle tre popolazioni definite dal titolo di studio al livello di significatività del 1%: (i) Scrivere l'ipotesi nulla e l'ipotesi alternativa; (ii) specificare la regione critica; (iii) calcolare il valore della statistica test; (iv) prendere una decisione (giustificandola).
- (c) Il p -valore per tale test è 0.0025. Interpretare tale valore anche in base al risultato ottenuto al punto (b)
- (d) Scrivere il modello di analisi della varianza che pone il punteggio al test in funzione della titolo di studio usando Medie inferiori o meno come livello di riferimento. Stimare quindi i parametri del modello e interpretarli.

Esercizio 2

In uno studio sulla relazione tra Y = tasso di criminalità (numero di crimini ogni 1000 abitanti) e X = tasso di povertà (percentuale di popolazione con reddito sotto il livello di povertà) nelle città di un certo paese si è considerato un campione di 20 città. Per tale campione si sono osservate le seguenti statistiche:

$$\bar{x} = 14 \quad \bar{y} = 8.5 \quad s_x^2 = 49 \quad s_y^2 = 144 \quad s_{xy} = 73.5$$

- (a) Stimare il coefficiente di correlazione
- (b) Stimare i coefficienti del modello di regressione che pone il tasso di criminalità in funzione del tasso di povertà
- (c) Stimare la varianza degli errori
- (d) Calcolare il tasso di criminalità medio per città con tasso di povertà uguale a 12 utilizzando il modello di regressione. Costruire quindi il relativo intervallo di confidenza al livello di confidenza del 95%.

Esercizio 3

Su un campione di $n = 32$ soggetti anziani si sono osservate le seguenti variabili: Y = punteggio a un test per valutare le capacità cognitive; X_1 = Titolo di studio (1 = Media inferiore o meno; 2 = Media superiore; 3 = Diploma universitario o più); X_2 = Stato civile (1 = Single; 2 = Sposato/a; 3 = Separato/Divorziato; 4 = Vedovo/a). Si supponga di adattare un modello di regressione che pone il punteggio al test, Y , in funzione delle variabili esplicative X_1 e X_2 . Si consideri tale modello come modello esteso. Per tale modello la somma dei quadrati dei residui è $SSE_e = 495.8$. Si supponga di voler confrontare tale modello con un modello ridotto in cui si considera come variabile esplicativa solo il titolo di studio, sapendo che per il modello ridotto la somma dei quadrati dei residui è $SSE_r = 680.5$.

- (c) Completare la seguente tavola di analisi della varianza per il modello esteso e per il modello ridotto sapendo che la varianza del punteggio al test è $s_Y^2 = 40$

Fonte di variabilità	Somma dei Quadrati	Gradi di libertà
Regressione		
Residua		
Totale		

- (b) Scrivere l'ipotesi nulla e l'ipotesi alternativa relativa al confronto tra il modello esteso e il modello ridotto. Scrivere quindi la regione critica al livello di significatività $\alpha = 0.10$
- (c) Calcolare il valore osservato della statistica test e prendere una decisione.

Domanda 1

Su un campione di soggetti in pensione di età compresa tra i 50 e i 70 anni si sono rilevate le seguenti variabili: una variabile binaria Y uguale a 1 se il soggetto presenta sintomi di senilità e uguale a zero altrimenti; e una variabile X che indica il numero di anni che il soggetto è in pensione

Variabile	Coefficiente	Errore standard
Costante	-1.8	0.53
Anni in pensione	0.25	0.03

- (a) Scrivere l'equazione che definisce la provabilità stimata $\hat{\pi}_i = \hat{Pr}(Y_i = 1|X_i = x_i)$.
- (b) Interpretare il coefficiente relativo agli anni in pensione in termini di odds.
- (c) Calcolare la probabilità stimata di presentare sintomi di senilità per un soggetto che è in pensione da 5 anni

Domanda 2

In uno studio finalizzato ad analizzare le capacità cognitive di soggetti in pensione di età compresa tra i 65 e i 75 anni viene stimato un modello di regressione in cui il punteggio a un test volto a misurare le capacità cognitive di un soggetto viene posto in relazione a $X_1 =$ numero di anni che il soggetto è in pensione e $X_2 =$ età (entrambe scartate dalla propria media). Di seguito i risultati ottenuti:

```
> mean(X1)
[1] 2.5
> mean(X2)
[1] 70
>
> mX1<- X1-mean(X1)
> mX2<- X2-mean(X2)

> summary(lm(Y~mX1+mX2))

Call:
lm(formula = Y ~ mX1 + mX2)

Residuals:
Min      1Q  Median      3Q      Max
-4.6979 -1.4957  0.3079  1.6041  2.8337

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept)  79.9096     0.2045 390.840 < 2e-16 ***
mX1          -0.4929     0.1765  -2.792  0.00644 **
mX2          -0.6787     0.1058  -6.414  7.14e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.94 on 87 degrees of freedom
Multiple R-squared:  0.4052, Adjusted R-squared:  0.3915
F-statistic: 29.63 on 2 and 87 DF,  p-value: 1.533e-10
```

- (a) Scrivere l'equazione di regressione stimata
- (b) Interpretare i coefficienti di regressione
- (c) Si consideri il valore della statistica test **F-statistic**: (i) Scrivere l'ipotesi nulla e l'ipotesi alternativa a cui è relativo tale valore; (ii) Interpretare il relativo p-valore