

Tecniche di indagine statistica

10 gennaio 2018

1. Sui dati dell'indagine PISA Italia 2003 è stato stimato il seguente modello logit (fonte: D. Contini, 2007) per l'analisi della probabilità di iscrizione ad un istituto professionale ($Y = 1$ se istituto professionale; $Y = 0$ se liceo/ist. tecnico/ist. magistrale) al variare delle caratteristiche socio-demografiche dello studente, controllando per la zona geografica di residenza.

	β	<i>s.e.</i>	Wald	<i>df</i>	<i>p-value</i>	$exp(\beta)$
NordEst	.366	.243	2.264	1	.132	1.442
Centro	.235	.235	1.001	1	.317	1.265
SudIsole	-.196	.202	.945	1	.331	.822
Sex	.512	.148	11.926	1	.001	1.669
Immigd	.343	.522	.432	1	.511	1.410
Escs	-.850	.085	100.266	1	.000	.427
Constant	-1.621	.185	76.500	1	.000	.198

Con:

NordEst: residenza nord-est = 1, altro = 0

Centro: residenza centro = 1, altro = 0

SudIsole: residenza sud-isole = 1, altro = 0

Sex: genere F = 0, M = 1

Immigd = nazionalità nativo in Italia = 0, *nonnativo* = 1

Escs: status socio-economico (v. continua con valori approssimativamente compresi tra $[-3, 2.5]$)

- a. descrivere che cosa rappresentano e come si interpretano i valori nelle colonne da 2 a 7;
- b. calcolare la probabilità di iscrizione a scuole non professionali per una ragazza con Escs = 0.5 (status vicino alla media) di nazionalità italiana e non e confrontare i valori ottenuti per residenza;
- c. che cosa si può dire dei risultati ottenuti al punto b) in relazione alla significatività dei coefficienti ?
- d. indicare se e come potrebbe essere migliorato il modello.

2. La tabella seguente riporta i risultati di uno studio (Schuman e Converse, 1971) sulle risposte di capifamiglia neri ad alcuni quesiti su questioni razziali, distinte secondo la razza dell'intervistatore.

		Risposta nella categoria con intervistatore	
<i>Quesito</i>	<i>Categoria di risposta</i>	Bianco	Nero
Si può avere fiducia nei bianchi ?	Per la maggior parte	35%	7%
I genitori neri vanno più d'accordo con insegnanti neri?	Sì	14%	29%
Con chi si intrattiene più facilmente?	Nominate solo persone di razza nera	16%	43%

Sapendo che l'assegnazione degli intervistatori agli intervistati è stata effettuata in modo casuale:

- a. indicare gli obiettivi dello studio e interpretare i risultati;
 - b. che considerazioni si possono trarre ai fini di progettazione di indagini con contenuti simili?
 - c. i risultati riportati potrebbero essere 'spiegati' da qualche altra caratteristica dell'intervistatore? Motivare la risposta.
3. Da una lista di 2000 scuole superiori con 1000 studenti ciascuna, è estratto un campione *epsem* di 3000 studenti secondo una procedura a due stadi. Al primo stadio sono selezionate 100 scuole e al secondo sono estratti 30 studenti in ogni scuola selezionata. Il 30% degli studenti intervistati dichiara di possedere un computer personale e lo standard error di tale stima è pari a 1.4%. Ignorando il fattore di correzione per popolazione finita e con l'approssimazione $n - 1 = n$, stimare:
- a. il *design effect* per la percentuale campionaria;
 - b. il coefficiente di correlazione intraclasse ρ entro le scuole per il possesso di computer da parte degli studenti;
 - c. lo standard error della percentuale campionaria nel caso di un campione che prevede la selezione di 300 scuole e di 10 studenti per scuola.
4. In tre città diverse, Roma, Campobasso e Bologna, si estrae un campione casuale semplice senza reinserimento di maschi adulti pari all'1% della popolazione di maschi adulti delle tre città per stimare la proporzione di favorevoli alla pena di morte. In quale città la varianza campionaria della stima è più bassa ? (motivare la risposta).