

Elementi di Psicometria

14-Introduzione all'analisi della potenza statistica vers. 1.1 (13 dicembre 2011)

Germano Rossi¹

`germano.rossi@unimib.it`

¹Dipartimento di Psicologia, Università di Milano-Bicocca

2011-2012

Analisi della potenza

- È il capitolo 14 del libro di testo
- Studiare fino a p.265
- Leggere a p. 265 “La verifica d’ipotesi sulla media di una popolazione”
- Saltare “proporzione” a p. 270, “correlazione” a p.272
- Leggere “differenza fra le medie” a p. 273, “medie appaiate” a p.278, “come stimare” a p.279
- Leggere “interpretazione” a p. 281

Analisi della potenza (Power analysis)

- Chiamiamo **errore di I tipo** la possibilità di accettare per vera H_1 che, invece, è falsa ovvero H_0 **è vera ma noi la rifiutiamo**
- La probabilità associata a questo errore è indicata con α ed è il rischio massimo che siamo disposti a correre, per sbagliare (rifiutare qualcosa che è vero)
- Chiamiamo **errore di II tipo** o β la possibilità di rifiutare H_1 , ritenendola falsa, mentre invece è vera ovvero H_0 è falsa ma noi la accettiamo come vera
- Si chiama **potenza di un test** la capacità di una statistica di accettare H_1 quando è vera ($1-\beta$)

Analisi della potenza

- La **potenza statistica** di un test è la sua capacità di rifiutare un'ipotesi nulla falsa, perché noi, in genere, verifichiamo un'ipotesi nulla rispetto ad una “gamma” di ipotesi alternative (ad es. $H_1 : \mu_1 \neq \mu_2$)
- Come ricercatori, facciamo molti sforzi per organizzare e fare una ricerca che ci dia conoscenze “sicure” e “affidabili”. Ma i nostri sforzi sono vani se non riusciamo a trovare i risultati che ci aspettiamo, o meglio, se non riusciamo a falsificare con maggior sicurezza la nostra ipotesi.
- Per molti anni, i ricercatori si sono focalizzati sul rischio di rifiutare H_0 quando è vera (atteggiamento conservatore)
- Di recente ha acquisito importanza anche l'errore opposto.
- Riassumiamo un momento le procedure di verifica d'ipotesi

Verifica d'ipotesi

- All'inizio di una ricerca, partiamo generalmente da un'ipotesi che è espressa a parole. Ad es. “A causa delle nuove tecnologie di comunicazione veloce (e-mail, sms, chat, cellulari) gli studenti passano *meno tempo* a stabilire relazioni personali dirette fra di loro”.
- Siccome qualcuno ha raccolto dati sul tempo trascorso in relazioni personali negli anni precedenti ($M=6$ ore alla settimana; $s=2$), posso raccogliere un nuovo campione da confrontare con il precedente
- Possiamo trasformare la nostra ipotesi verbale in ipotesi statistica:

$$H_0 : \mu = 6.0$$

$$H_1 : \mu < 6.0$$

Verifica d'ipotesi

- Ricordiamo che noi verifichiamo l'ipotesi nulla confrontandola con un'ipotesi alternativa.
- L'ipotesi nulla è ciò che è noto o che si assume in base alla teoria o alle ricerche precedenti.
- Nel nostro esempio, la ricerca precedente, ci ha detto che gli studenti universitari hanno speso circa 6 ore al giorno della settimana in contatti faccia-a-faccia (più o meno 2 ore).
- Così, la nostra ipotesi è che $\mu = 6.0$.
- L'errore α ci protegge dal prendere una decisione errata basata su un campione “particolarmente anomalo” estratto dalla popolazione corretta
- La potenza di un test $(1 - \beta)$ ci dice la probabilità di aver accettato correttamente l'ipotesi alternativa

Concetti chiave della potenza

- Ricordiamo che la **potenza statistica** di un test è la sua capacità di rifiutare un'ipotesi nulla falsa e che è legata al test statistico usato.
- Ci sono 4 variabili che agiscono sulla potenza:
 - 1 Il livello di significatività cioè α : più è severo (vicino a 0), più è difficile rifiutare l'ipotesi nulla (anche quando è falsa).
All'aumentare di α aumenta la potenza del test. Tuttavia non possiamo usare α molto grandi; un buon criterio (non troppo basso, né troppo alto) è $\alpha = 0.05$
 - 2 L'ampiezza del campione cioè N ; quando un campione è grande, è meno probabile fare errori di campionamento e trovare dati che portino a stime inaffidabili dei parametri della popolazione.
L'errore standard è sempre basato su N . Quindi all'aumentare di N , aumenta la potenza

Concetti chiave della potenza

- 3 La dimensione dell'effetto nella popolazione cioè d ; d indica genericamente l'ampiezza dell'effetto, ovvero quanto grande è il risultato che abbiamo ottenuto; ricordiamo che d ha senso solo se H_0 è falsa; quindi possiamo considerare d come una misura di quanto è falsa l'ipotesi nulla; tanto più d è grande, tanto più H_0 è falsa, tanto più aumenta la potenza
- 4 La potenza statistica stessa, cioè $1 - \beta$
- Infine, questi quattro indici sono legati fra loro matematicamente; ovvero conoscendo il valore di tre di loro, si può calcolare il valore del quarto

Concetti chiave della potenza

Riassumendo:

	Potenza ($1 - \beta$)	
	aumenta	diminuisce
α	grande	piccola
N	grande	piccolo
d	grande	piccolo

- La formula che lega i quattro indici è abbastanza complessa, per cui sono state predisposte delle tavole (Tavola D a p. 476) che utilizzano α e una combinazione di d ed N chiamata δ :

$$\delta = \frac{d}{f(N)}$$

dove $f()$ indica “funzione di”

Uso dell'analisi di potenza

L'analisi di potenza viene usata, generalmente, per due obiettivi

- 1 **a posteriori** per determinare la **potenza di un test**: dal momento che la ricerca viene effettuata su un certo campione (di ampiezza N) e usando un certo livello α , e dai risultati ottenuti possiamo calcolare d , ne consegue la possibilità di stimare la potenza di un test, cioè la probabilità di aver fatto la scelta giusta con α ;
- 2 **a priori** per determinare la **numerosità del campione**: se vogliamo fare una ricerca che abbia una determinata potenza (di solito, almeno .80), una volta stabilito un determinato α e ipotizzato un determinato d , quale dev'essere l'ampiezza del campione?

Al punto 2, ci potrebbe servire anche un'ipotesi alternativa puntuale (ad es. $H_1 : \mu = 4$)

Calcolare la potenza di un test

- La potenza di un test si calcola, generalmente, a posteriori
- Abbiamo raccolto un certo campione
- Per la verifica d'ipotesi, abbiamo stabilito un livello α
- L'analisi dei dati, ci ha fornito i valori da utilizzare per il calcolo di d
- A questo punto ci serve solo di sapere qual è la funzione da applicare ad N per trovare δ , e per cercare il valore corrispondente alla potenza sulla tavola D
- La funzione $f(N)$ cambia in base alla tecnica statistica utilizzata

Numerosità del campione

- Ipotizziamo di voler fare una ricerca su un campione patologico (ad es. pazienti di un servizio mentale)
- Possiamo fare una ricerca veloce (in termini di tempo) su un piccolo campione
- oppure una ricerca che duri più tempo per poter raccogliere un campione più grande
- certamente non vogliamo fare una ricerca che non abbia abbastanza “potenza” e che possa essere criticata
- Decidiamo quindi una potenza minima che vogliamo raggiungere, un d che ci aspettiamo e calcoliamo **quanto** dev’essere ampio il campione da raccogliere.

Relazioni fra errori e ipotesi

	Ipotesi	
	H_0 - Vera H_1 - Falsa	H_0 - Falsa H_1 - Vera
Accetto H_0 ; rifiuto H_1	Corretta $1 - \alpha$	Errore II tipo β
Rifiuto H_0 ; accetto H_1	Errore I tipo α	Corretta $1 - \beta$

Interpretazione

- **Un risultato non significativo non è necessariamente vero:** se accettiamo per vera H_0 , questo non significa che sia “realmente” vera. Se la potenza di un esperimento è bassa (ad es. .50), *beta* sarà alto ($1 - .50 = .50$); quindi avremo il 95% di probabilità ($1 - .05$) che H_0 sia vera, ma il 50% che sia falsa
- **Significatività con piccoli campioni:** Nel caso del test t , t è parte di **g**; se t è molto alto, anche **g** lo sarà, se **g** è alto anche la potenza sarà alta (e *beta* sarà basso). Quindi un risultato t molto alto è molto più importante se ottenuto su piccoli campioni che su grandi campioni.
- **Grandi campioni e significatività:** con grandi campioni è più facile ottenere un t grande e quindi significativo; bisogna quindi considerare **d** (o **g**) per sapere l'effettiva sicurezza da dare al risultato raggiunto