

TEST χ^2 PER PIU' DI DUE GRUPPI (O RISULTATI)

giovedì 29 maggio 2014 10.10

Supponiamo di avere r campioni di v. al.

$$X_1^j, \dots, X_{n_j}^j \sim \text{Be}(p_j), \quad j=1, \dots, r$$

(nota: $c=2$). Raccogliamo ogni campione in una riga della tabella di contingenza osservata:

$$O_{j1} = \sum_{i=1}^{n_j} X_i^j \quad O_{j2} = n_j - O_{j1}$$

	Tot	
	Tot	
camp. 1	O_{11}	O_{12}
camp. 2	1	
$\vdots$	1	
camp. r	O_{r1}	O_{r2}
TOT.		

Vogliamo testare $H_0: p_1 = p_2 = \dots = p_r$ contro

$H_1: \exists i, j \text{ t.c. } p_i \neq p_j$ - Sotto $H_0, p_1 = \dots = p_r = p$,

per cui una stimatore è

$$\hat{p} = \frac{O_{11} + O_{21} + \dots + O_{r1}}{n_1 + \dots + n_r}$$

Calcoliamo le tabelle attese:

$$E_{j1} = n_j \cdot \hat{p}, \quad E_{j2} = n_j(1 - \hat{p}) = n_j - E_{j1}$$

Se n_j sono "grandi" (cioè $E_{ij} > 5 \quad \forall i, j$), allora possiamo applicare l'approssimazione normale e costruire

$$\chi^2 = \sum_{j=1}^r \sum_{i=1}^{(c=2)} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

sotto H_0 , $\chi^2 \sim \chi^2(\nu)$, con $\nu = (r-1)(c-1)$

regole:

- se $\chi^2 < \chi^2_{1-\alpha}(\nu)$, accettiamo H_0
- se $\chi^2 > \chi^2_{1-\alpha}(\nu)$, rifiutiamo H_0 e accettiamo H_1

In quest'ultimo caso, quali sono i campioni diversi dagli altri?

esempio: controllo → joggers → athlete

sono amate o no dal ginecologo per mestruazioni saltate?

tabella osservata:

	gin. SÌ	NO	TOT.
C	14 (22,6)	40 (31,4)	54
J	9 (9,6)	14 (13,4)	23
A	46 (36,8)	42 (51,2)	88

	46	(36,8)	42	(51,2)	88
A					
	69		96		165

test: $H_0: p_C = p_T = p_A$, $H_1: \exists i, j \text{ t.c. } p_i \neq p_j$

$$\hat{p} = \frac{69}{165} = 0,418$$

$$0,418 \times 54 = 22,6$$

$$0,418 \times 23 = 9,6$$

$$0,418 \times 88 = 36,8$$

$$E_{ij} > 5 \quad \forall i, j \Rightarrow$$

$$\text{test } \chi^2$$

$$\chi^2 = \frac{(14 - 22,6)^2}{22,6} + \frac{(40 - 31,4)^2}{31,4} + \frac{(9 - 9,6)^2}{9,6} + \frac{(14 - 13,4)^2}{13,4} + \frac{(46 - 36,8)^2}{36,8} + \frac{(42 - 51,2)^2}{51,2} = 9,65$$

$$\alpha = 0,05 : \text{val. critica } \chi^2_{1-\alpha}(v) = \chi^2_{0,95}(2) = 5,991$$

$$\chi^2 > \chi^2_{1-\alpha}(v) \Rightarrow \text{rif. } H_0 \text{ e acc. } H_1$$

che hanno i tassi uguali/diversi?

Individuare le tabelle di contingenza

Yaccino su diversi test sulle possibili coppie di campi

con ipotesi $H_0: p_i = p_j$ contro $H_1: p_i \neq p_j$.

Vogliamo tenere la probabilità di commettere un'errore di 1^a specie in almeno uno dei test bassa. In particolare:

$A_{ij} = \{ \text{un. 1^a specie su test } p_i = p_j \}$
 per attribuire il livello α' a questo test, vogliamo

$$P_0(A_{ij}) \leq \alpha'$$

Qual è la prob. di commettere almeno un un. di 1^a specie?

$$A = \bigcup_{i \neq j} A_{ij}$$

$$P_0(A) = P_0\left(\bigcup_{i \neq j} A_{ij}\right) \leq \sum_{i \neq j} P_0(A_{ij})$$

Vogliamo fissare un livello α , in modo che facendo n test multipli di livello α' , abbiamo $\alpha \geq P_0(A)$: come fissare α' ?

Disuguaglianza di Bonferroni: poniamo $\alpha' = \frac{\alpha}{n}$; allora

$$P_0(A) \leq \sum_{i=1}^n P_0(A_i) \leq \sum_{i=1}^n \alpha' = n \alpha' = n \frac{\alpha}{n} = \alpha$$

esempio: test multipli su C-J-A

$$H_0: p_J = p_A, \quad H_1: p_J \neq p_A$$

	SÌ	NO	TOT
T	0 (1,1)	1,1 (1,1)	2,2

	SÌ	NO	TOT
J	9 (11)	14 (12)	23
A	46 (44)	42 (44)	88
	55	56	111

$$\hat{p} = \frac{55}{111} = 0,495$$

$$23 \times 0,495 = 11,4 \simeq 11$$

$$\chi^2 = \frac{(9-11)^2}{11} + \frac{(14-12)^2}{12} + \frac{(46-44)^2}{44} + \frac{(42-44)^2}{44} =$$

$$= 0,88$$

$$0,88 < \chi^2_{0,95}(1) = 3,84 \Rightarrow P > 0,05$$

$$\alpha' = \begin{cases} (\text{tutto, test}) \frac{\alpha}{3} = 0,0166 \\ (\text{solo 2}) \frac{\alpha}{2} = 0,025 \end{cases}$$

in ogni caso, $\alpha' < P \Rightarrow$ accetto H_0

Posso allora considerare joggens = atlete e unificare i gruppi!

	SÌ	NO	TOT
C	14 (22,6)	40 (31,4)	54
J + A	55 (46,4)	56 (64,6)	111
	69	96	165

$$H_0: p_c = p_{JA}$$

$$H_1: p_c \neq p_{JA}$$

$$\hat{p} = \frac{69}{165} = 0,418$$

$$\chi^2 = \frac{(14 - 22,6)^2}{22,6} + \frac{(40 - 31,4)^2}{31,4} + \frac{(55 - 46,4)^2}{46,4} + \frac{(56 - 64,6)^2}{96} = 8,36$$

abbiamo fatto 2 test multipli, di livello $\alpha' = \frac{\alpha}{2} =$

$$= 0,025$$

val. critica: $\chi^2_{1-\alpha'}(1) = \chi^2_{0,975}(1) = 5,024 < \chi^2$

ripiamo H_0 accettiamo H_1

conclusione globale: $p_c \neq p_J = p_A$

controllo di "robustezza": cambiare l'ordine dei test multipli influisce sul risultato?

Proviamo a fare prima il test $H_0: p_c = p_J$, $H_1: p_c \neq p_J$

C	14 (16)	40 (38)	54
J	9 (7)	14 (16)	23
	23	54	77

$$\hat{p} = \frac{23}{77} = 0,298$$

$$E_{11} = 54 \times 0,298 = 16,0$$

... .. χ^2

127

$E_{ij} > 5 \quad \forall i, j \Rightarrow$ test χ^2

$$\begin{aligned}\chi^2 &= \frac{(14-16)^2}{16} + \frac{(40-38)^2}{38} + \frac{(9-7)^2}{7} + \frac{(14-16)^2}{16} = \\ &= 4 \left(\frac{1}{16} + \frac{1}{38} + \frac{1}{7} + \frac{1}{16} \right) = 1,176\end{aligned}$$

$$\chi'_{0,95}(1) = 3,84 > \chi^2 \Rightarrow p > 0,05 \text{ e acc. } H_0: p_c = p_j$$