

Test Chi-quadrato per l'indipendenza

- Se si considera una tabella di contingenza con r righe e c colonne il procedimento del test χ^2 può essere generalizzato per verificare l'indipendenza tra due variabili categoriali X e Y
- In questo contesto le ipotesi nulla e alternativa sono
 H_0 : le due variabili categoriali sono indipendenti ($X \perp Y$)
(non sussistono relazioni tra le due variabili)
 H_1 : le due variabili categoriali sono dipendenti ($X \not\perp Y$)
(sussiste una relazione tra le due variabili)
- Il test si basa ancora una volta sull'equazione (11.1)

$$\chi^2 = \sum_{\text{tutte le celle}} \frac{(f_o - f_e)^2}{f_e}$$

- La regola decisionale consiste nel rifiutare H_0 se il valore osservato della statistica χ^2 è maggiore del valore critico χ^2_U della distribuzione χ^2 con $(r-1) \times (c-1)$ gdl

Test Chi-quadrato per l'indipendenza

- Pure se presentano delle analogie, la differenza fondamentale tra il test chi-quadrato per le proporzioni e per l'indipendenza riguarda lo schema di campionamento:
 - Nel confronto tra proporzioni siamo di fronte a campioni estratti da popolazioni indipendenti, il cui unico fattore di interesse è rappresentato da una variabile “successo/insuccesso”
 - Nel test di indipendenza abbiamo un solo campione su cui rileviamo due variabili qualitative che possono assumere r e c modalità distinte
 - Nel caso di test chi-quadrato per l'indipendenza è possibile semplificare il calcolo delle frequenze attese applicando la seguente regola:

$$f_e = \frac{\text{totale di riga} \times \text{totale di colonna}}{n}$$

Test Chi-quadrato per l'indipendenza

Esempio: tabella della frequenze osservate con riferimento al principale motivo di insoddisfazione e all'albergo

Motivo di insoddisfazione	Albergo			Totale
	Golden Palm	Palm Royale	Palm Princess	
Prezzo	23	7	37	67
Posizione dell'albergo	39	13	8	60
Sistemazione nelle camere	13	5	13	31
Altro	<u>13</u>	<u>8</u>	<u>8</u>	<u>29</u>
Totale	88	33	66	187

Frequenze attese

Motivo di insoddisfazione	Albergo			Totale
	Golden Palm	Palm Royale	Palm Princess	
Prezzo	31.53	11.82	23.65	67
Posizione dell'albergo	28.24	10.59	21.18	60
Sistemazione nelle camere	14.59	5.47	10.94	31
Altro	13.65	5.12	10.24	29
Totale	88.00	33.00	66.00	187

Test Chi-quadrato per l'indipendenza

Calcolo della statistica χ^2 per il test di indipendenza

Cella	f_0	f_e	$(f_0 - f_e)$	$(f_0 - f_e)^2$	$(f_0 - f_e)^2/f_e$
Prezzo/Golden Palm	23	31.53	-8.53	72.7609	2.308
Prezzo/Palm Royale	7	11.82	-4.82	23.2324	1.966
Prezzo/Palm Princess	37	23.65	13.35	178.2225	7.536
Posizione/Golden Palm	39	28.24	10.76	115.7776	4.100
Posizione/Palm Royale	13	10.59	2.41	5.8081	0.548
Posizione/Palm Princess	8	21.18	-13.18	173.7124	8.202
Stanza/Golden Palm	13	14.59	-1.59	2.5281	0.173
Stanza/Palm Royale	5	5.47	-0.47	0.2209	0.040
Stanza/Palm Princess	13	10.94	2.06	4.2436	0.388
Altro/Golden Palm	13	13.65	-0.65	0.4225	0.031
Altro/Palm Royale	8	5.12	2.88	8.2944	1.620
Altro/Palm Princess	8	10.24	-2.24	5.0176	0.490
					<u>27.402</u>

Test Chi-quadrato per l'indipendenza

Regione di rifiuto e di accettazione del test χ^2 per l'indipendenza nell'esempio sulla soddisfazione dei clienti (al livello di significatività 0.05 con 6 gradi di libertà)

