

Test chi quadrato



SAPIENZA
UNIVERSITÀ DI ROMA

annarita.vestri@uniroma1.it

Tabelle a doppia entrata

Disponendo di caratteri espressi con dati nominali (qualitativi) e/o categorici (livello di salute, classe di rischio...), è interesse del ricercatore valutare se fra tali caratteri esista un legame di dipendenza o di associazione

La dipendenza presuppone l'identificazione di una variabile temporalmente antecedente e quindi oggettivamente in grado di influenzare la seconda variabile detta dipendente (rapporto di causa ed effetto)

Esempio: esposizione ad un fattore di rischio (variabile indipendente) e sviluppo di una determinata patologia (variabile dipendente)

Per associazione si intende la tendenza di due variabili ad abbinarsi (associarsi) in un certo modo, senza che una di queste si possa univocamente identificare come variabile indipendente e l'altra come dipendente

Esempi: la preferenza di alcune categorie professionali (operai, impiegati, imprenditori...) per determinati schieramenti politici, la preferenza del sesso femminile per il corso in Medicina e del maschile per Ingegneria, ecc.

In entrambi le situazioni, si rende necessaria la costruzione di tabelle dette Tabelle doppie (Tabelle a doppia entrata, Tabelle di contingenza) nelle quali i due caratteri studiati vengono “incrociati” ovvero le frequenze delle modalità con cui i due caratteri si manifestano vengono riportate congiuntamente nella stessa tabella

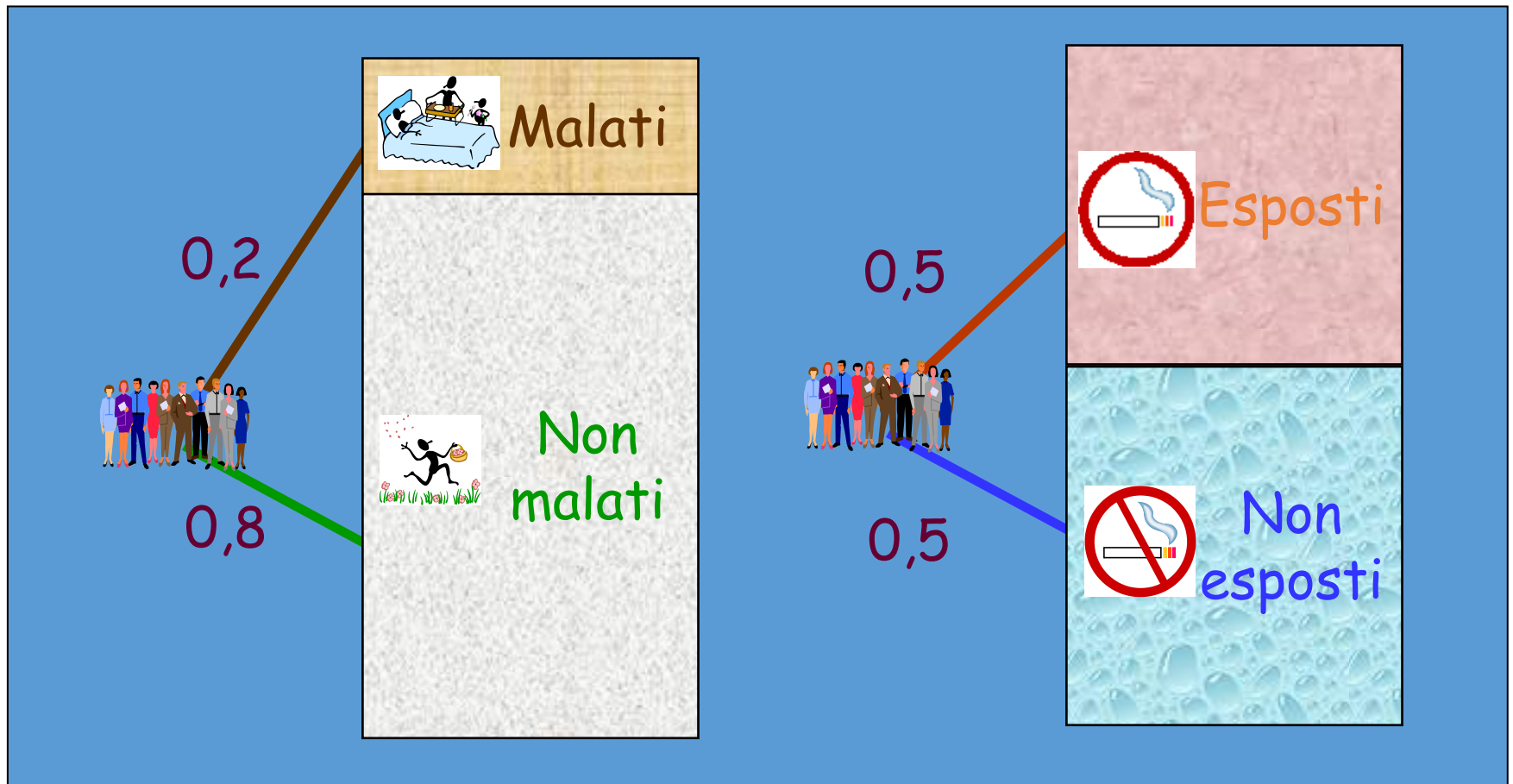
TABELLE DI CONTINGENZA

Frequenze di dati **nominali raggruppati in categorie** sono spesso organizzate in forma di **tabella di contingenza**.

Il caso più semplice implica due variabili casuali dicotomiche; le righe della tabella rappresentano i risultati di una variabile e le colonne i risultati dell'altra, i numeri all'interno della tabella sono le frequenze di una particolare combinazione di categorie.

Prendiamo in esame l'esposizione al fumo ed le bronchiti croniche

E' noto che l'esposizione e la malattia sono distribuite nella popolazione come nel seguente schema



Se rappresentiamo esposizione e patologia in una tabella a doppia entrata, considerando una ipotetica popolazione di 100 soggetti, avremo una tabella 2x2 della quale saranno noti i totali marginali

Patologia	Esposizione		Totale
	esposti	non esposti	
malati			20
sani			80
Totale	50	50	100

Ipotizzando che ci sia indipendenza fra i 2 caratteri, ci aspetteremmo di trovare frequenze (teoriche) perfettamente proporzionali ai totali marginali

Patologia	Esposizione		Totale
	esposti	non esposti	
malati	10	10	20
sani	40	40	80
Totale	50	50	100

Ipotizzando che ci sia indipendenza fra i 2 caratteri, ci aspetteremmo di trovare frequenze del tutto proporzionali ai totali marginali

...infatti...

Patologia	Esposizione		Totale
	esposti	non esposti	
malati	10 (20%)	10 (20%)	20 (20%)
sani	40 (80%)	40 (80%)	80 (80%)
Totale	50 (100%)	50 (100%)	100 (100%)

Pertanto, nel caso di perfetta indipendenza fra le due variabili, le probabilità degli eventi intersezione saranno

Patologia	Esposizione		Totale
	esposti	non esposti	
malati	$p(M/E)=0,20$	$p(M/NE)=0,20$	$p(M)=0,20$
sani	$p(S/E)=0,80$	$p(S/NE)=0,80$	$p(S)=0,80$
Totale			

Supponiamo ora, per verificare la nostra teoria, di estrarre realmente dalla popolazione un campione di 100 soggetti

Naturalmente le frequenze marginali saranno le stesse

Patologia	Esposizione		Totale
	esposti	non esposti	
malati			20
sani			80
Totale	50	50	100

Supponiamo ora, per verificare la nostra teoria, di estrarre realmente dalla popolazione un campione di 100 soggetti

Naturalmente le frequenze marginali saranno le stesse... Ma potrebbero cambiare le frequenze osservate interne alle celle

Patologia	Esposizione		Totale
	esposti	non esposti	
malati	19	1	20
sani	31	49	80
Totale	50	50	100

In questo caso, le proporzioni trovate sono diverse da quelle date dai totali marginali

E' facile vedere come i malati si "associno" agli esposti (38%), mentre i non malati siano più presenti fra i non esposti (98%)

Patologia	Esposizione		Totale
	esposti	non esposti	
malati	19 (38%)	1 (2%)	20 (20%)
sani	31 (62%)	49 (98%)	80 (80%)
Totale	50 (100%)	50 (100%)	100 (100%)

Lo squilibrio percentuale degli ammalati, sia fra esposti e non esposti, che fra questi ed i totali marginali, è evidente

Trattandosi di un risultato campionario, c'è però la possibilità che tale squilibrio sia dovuto al caso

Il risultato ottenuto dovrà quindi essere saggiato statisticamente con un test opportuno, rispettando le quattro fasi già note

- Definizione delle ipotesi
- Definizione dei rischi di errore
- Scelta del test statistico
- Decisione finale

Potremmo formulare una ipotesi nulla ed una ipotesi alternativa come segue:

$$H_0: P(M|E) = P(M|NE) = P(M)$$

$$H_1: P(M|E) \neq P(M|NE) \neq P(M)$$

In sintesi si può scrivere

H_0 : indipendenza (fra i caratteri considerati)

H_1 : dipendenza (fra i caratteri considerati)

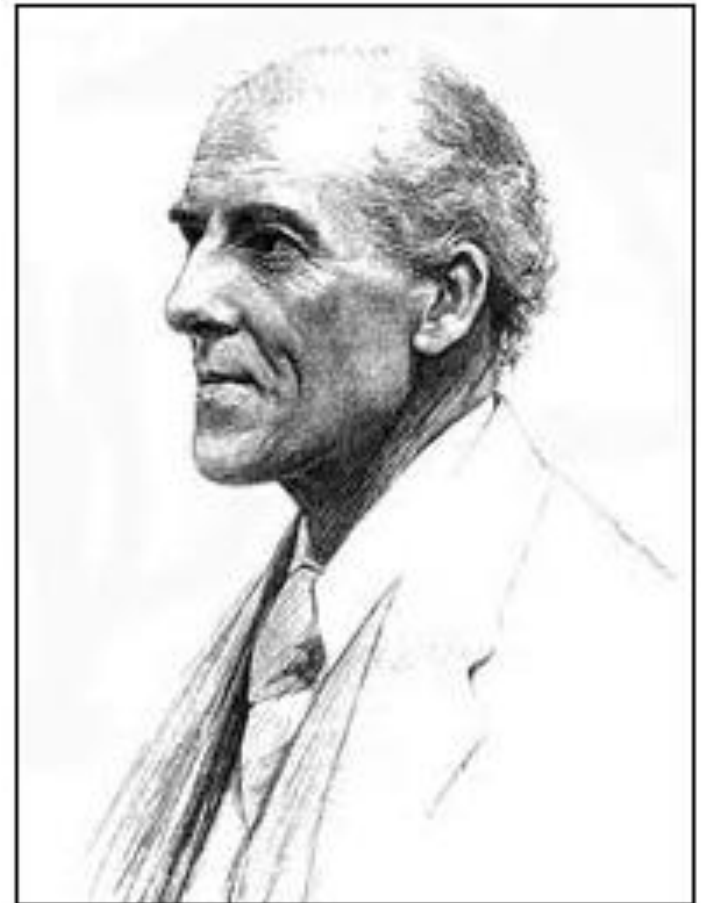
Per quanto riguarda i rischi di errore avremo, come per gli altri test

α : probabilità di respingere H_0 quando è vera (o errore di I Tipo) = 0.05

β : probabilità di accettare H_0 quando è falsa (o errore di II Tipo) = 0.10/0.20

Il test Chi quadrato

Il test comunemente utilizzato per la valutazione di dipendenza o associazione fra 2 caratteri nominali (o categorici) è il test Chi quadrato di Pearson



La formula del Chi quadrato (chisquare) è la seguente

$$\chi^2 = \sum_{i=1}^{rc} \frac{(O_i - E_i)^2}{E_i}$$

dove

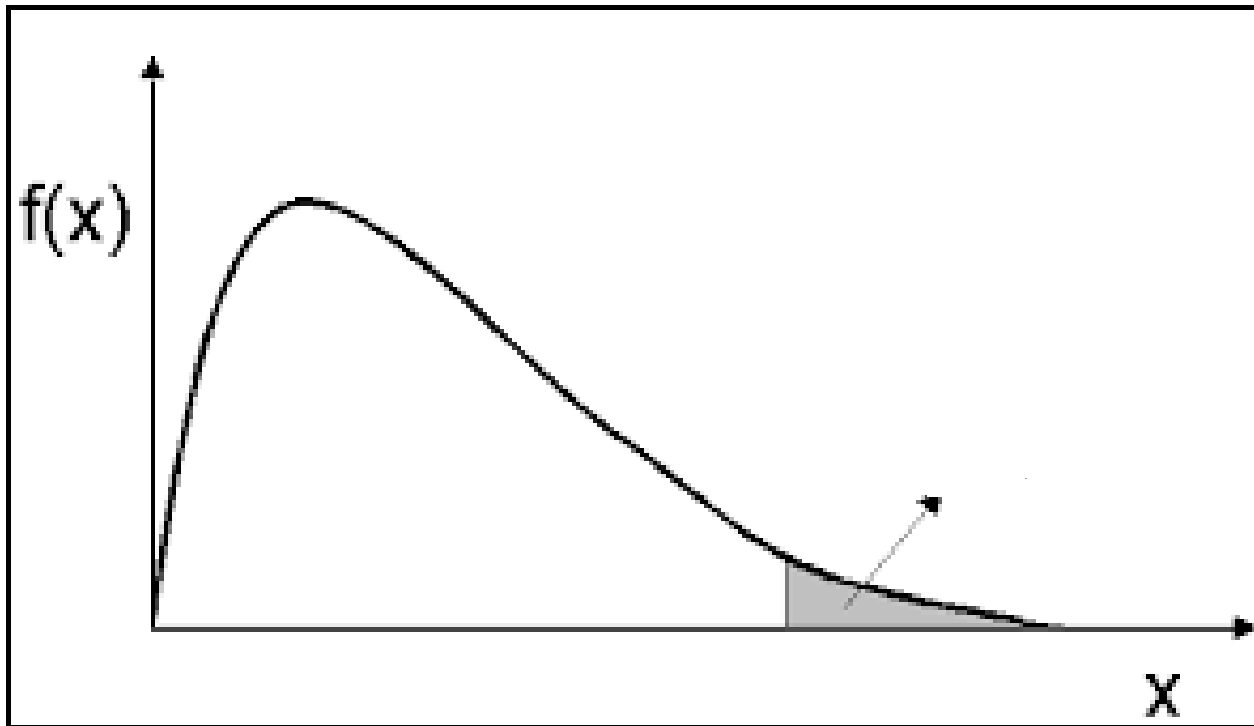
O_i = frequenze osservate (o empiriche)

E_i = frequenze teoriche (o attese), ovvero quelle che avremmo se fosse vera l'ipotesi nulla

Il test Chi quadrato si propone di stabilire se la differenza tra le frequenze osservate e le frequenze attese (detta contingenza) possa essere attribuibile al caso, o meno.

Il valore ottenuto dal test dovrà essere confrontato, utilizzando l'apposita distribuzione di probabilità, con il valore del χ^2 che definisce la probabilità α

Distribuzione di probabilità del Chi quadrato

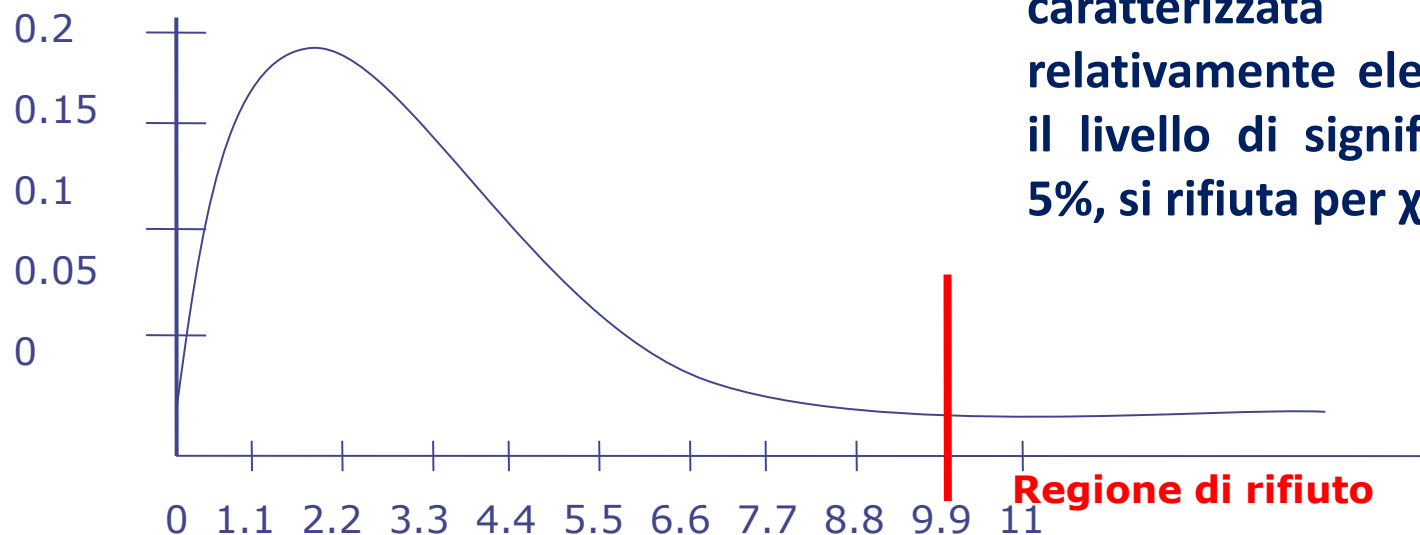


La distribuzione del χ^2

Si considera la distribuzione χ^2 , con un numero di gradi di libertà pari a $(r-1)(c-1)$, dove r è il numero di righe e c il numero di colonne della tabella di contingenza. Qui:

- H_0 : indipendenza statistica tra X e Y
- H_1 : dipendenza statistica tra X e Y

La regione di rifiuto cade nella coda di destra della distribuzione



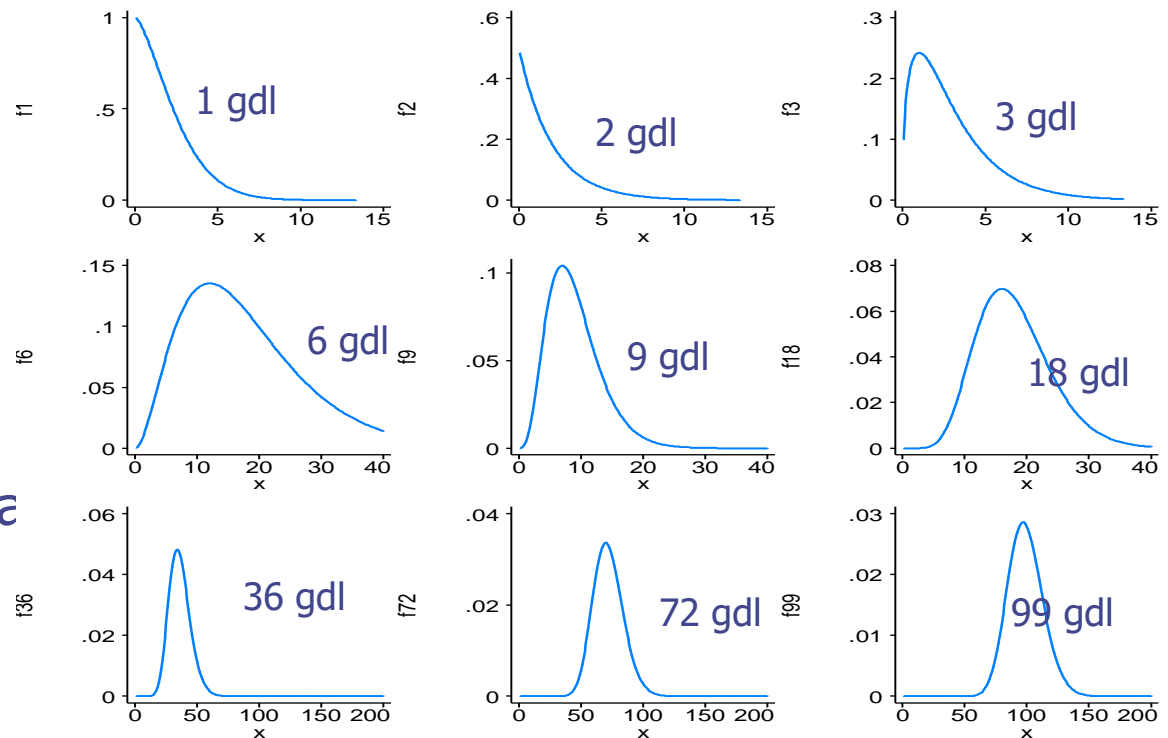
La regione di rifiuto è caratterizzata da valori relativamente elevati di χ^2 ; se il livello di significatività è al 5%, si rifiuta per $\chi^2 > \chi^2_{0.95}$

LA DISTRIBUZIONE CHI-QUADRATO

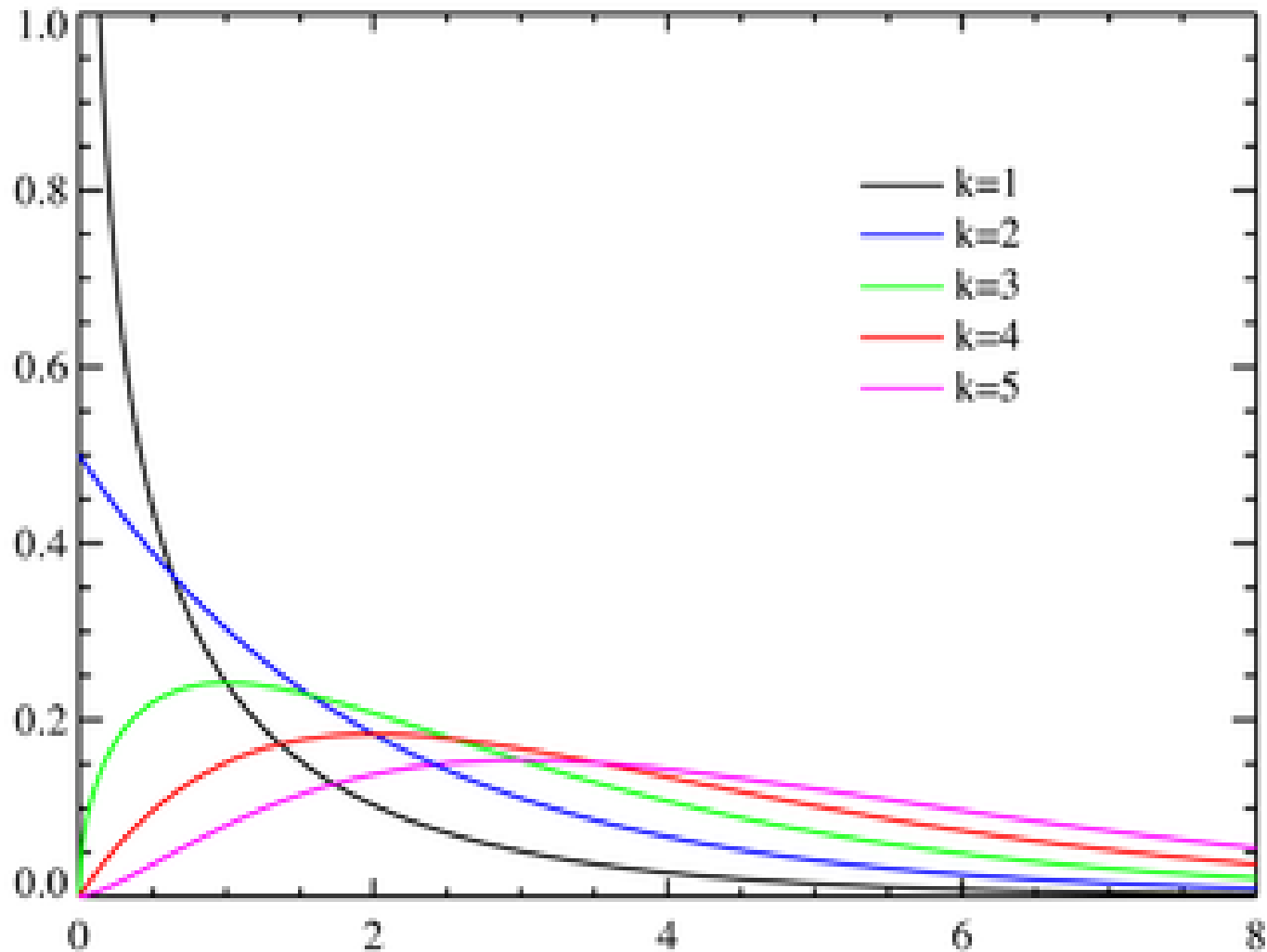
La distribuzione chi-quadrato è diversa per ogni possibile valore di gdl . Le distribuzioni con pochi gdl sono fortemente asimmetriche e tendono alla simmetria all'aumentare del numero di gdl, come illustrato in figura.

La tabella mostra le aree della distribuzione chi-quadrato a destra di valori assegnati di χ^2 e per diversi gdl .

Esempio: $\chi^2=3,84$ delimita il 5% superiore dell'area sotto la curva di una distribuzione con 1 gdl.



Modificazioni della distribuzione Chi quadro al variare dei gradi di libertà



La distribuzione di probabilità del Chi quadrato (χ^2)
varia al variare dei gradi di libertà

Sia: r = numero delle righe

c = numero delle colonne

Gradi di libertà = $(r-1) \times (c-1)$

Nel nostro esempio avremo

Gradi di libertà: $GdL = (2-1) \times (2-1) = 1$

Numero delle celle: $rc = 2 \times 2 = 4$

$$\chi^2 = \sum_{i=1}^{rc} \frac{(O_i - E_i)^2}{E_i}$$

$$\chi^2 = \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} + \frac{(O_3 - E_3)^2}{E_3} + \frac{(O_4 - E_4)^2}{E_4}$$

$$\chi^2 = \frac{(19-10)^2}{10} + \frac{(1-10)^2}{10} + \frac{(31-40)^2}{40} + \frac{(49-40)^2}{40}$$

$$\chi^2 = 8.1 + 8.1 + 2.025 + 2.025 = \mathbf{20.25}$$

Tavola del Chi quadrato, per vari livelli di α

Tavola distribuzione CHI-QUADRATO

Gradi di libertà	Livello di Probabilità α									
	1.00	0.99	0.95	0.90	0.25	0.10	0.05	0.025	0.01	0.005
1				0.02	1.32	2.71	3.84	5.02	6.64	7.88
2	0.01	0.02	0.10	0.21	2.77	4.61	5.99	7.38	9.21	10.60
3	0.07	0.12	0.35	0.58	4.11	6.25	7.82	9.35	11.35	12.84
4	0.21	0.30	0.71	1.06	5.39	7.78	9.49	11.14	13.28	14.86
5	0.41	0.55	1.15	1.61	6.63	9.24	11.07	12.83	15.09	16.75
6	0.68	0.87	1.64	2.20	7.84	10.65	12.59	14.45	16.81	18.55
7	0.99	1.24	2.17	2.83	9.04	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.73	3.49	10.22	13.36	15.51	17.54	20.09	21.96
9	1.74	2.09	3.33	4.17	11.39	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.94	4.87	12.55	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	4.58	5.58	13.70	17.28	19.68	21.92	24.73	26.76
12	3.07	3.57	5.23	6.30	14.85	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.89	7.04	15.98	19.81	22.36	24.74	27.69	29.82
14	4.08	4.66	6.57	7.79	17.12	21.06	23.69	26.12	29.14	31.32
15	4.60	5.23	7.26	8.55	18.25	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	7.96	9.31	19.37	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	8.67	10.09	20.49	24.77	27.59	30.19	33.41	35.72
18	6.27	7.02	9.39	10.87	21.61	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	10.12	11.65	22.72	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	10.85	12.44	23.83	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	11.59	13.24	24.94	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	12.34	14.04	26.04	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	13.09	14.85	27.14	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	13.85	15.66	28.24	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	14.61	16.47	29.34	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	15.38	17.29	30.44	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	16.15	18.11	31.53	36.74	40.11	43.19	46.96	49.65
28	12.46	13.57	16.93	18.94	32.62	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	17.71	19.77	33.71	39.09	42.56	45.72	49.50	52.34
30	13.79	14.95	18.49	20.60	34.80	40.26	43.77	46.98	50.89	53.67
31	14.46	15.66	19.28	21.43	35.89	41.42	44.99	48.23	52.19	55.00
32	15.13	16.36	20.07	22.27	36.97	42.59	46.19	49.48	53.49	56.33
33	15.82	17.07	20.87	23.11	38.06	43.75	47.40	50.73	54.78	57.65
34	16.50	17.79	21.66	23.95	39.14	44.90	48.60	51.97	56.06	58.96
35	17.19	18.51	22.47	24.80	40.22	46.06	49.80	53.20	57.34	60.28
36	17.89	19.23	23.27	25.64	41.30	47.21	51.00	54.44	58.62	61.58
37	18.59	19.96	24.08	26.49	42.38	48.36	52.19	55.67	59.89	62.88
38	19.29	20.69	24.88	27.34	43.46	49.51	53.38	56.90	61.16	64.18
39	20.00	21.43	25.70	28.20	44.54	50.66	54.57	58.12	62.43	65.48
40	20.71	22.16	26.51	29.05	45.62	51.81	55.76	59.34	63.69	66.77
41	21.42	22.91	27.33	29.91	46.69	52.95	56.94	60.56	64.95	68.05
42	22.14	23.65	28.14	30.77	47.77	54.09	58.12	61.78	66.21	69.34
43	22.86	24.40	28.97	31.63	48.84	55.23	59.30	62.99	67.46	70.62
44	23.58	25.15	29.79	32.49	49.91	56.37	60.48	64.20	68.71	71.89
45	24.31	25.90	30.61	33.35	50.99	57.51	61.66	65.41	69.96	73.17
46	25.04	26.66	31.44	34.22	52.06	58.64	62.83	66.62	71.20	74.44
47	25.78	27.42	32.27	35.08	53.13	59.77	64.00	67.82	72.44	75.70
48	26.51	28.18	33.10	35.95	54.20	60.91	65.17	69.02	73.68	76.97
49	27.25	28.94	33.93	36.82	55.27	62.04	66.34	70.22	74.92	78.23
50	27.99	29.71	34.76	37.69	56.33	63.17	67.51	71.42	76.15	79.49
55	31.74	33.57	38.96	42.06	61.67	68.80	73.31	77.38	82.29	85.75
60	35.53	37.49	43.19	46.46	66.98	74.40	79.08	83.30	88.38	91.95
70	43.28	45.44	51.74	55.33	77.58	85.53	90.53	95.02	100.43	104.22
80	51.17	53.54	60.39	64.28	88.13	96.58	101.88	106.63	112.33	116.32

Nella tavola del Chi quadrato, per 1 grado di libertà ed $\alpha = 0.05$, il valore soglia è pari a 3.84. Ma nel nostro caso il chi quadro empirico è

$$\chi^2 = 20.25$$

Questo valore supera ampiamente il valore soglia, cade nell'area di rifiuto, per cui rifiutiamo H_0 : risultato significativo

Si può affermare che l'esposizione al fattore di rischio favorisce l'insorgenza della patologia

Alla luce di quanto visto, appare evidente come il test Chi quadrato sia adatto per confrontare proporzioni.

Ma occorre ricordare che tale test richiede che i dati da inserire in tabella siano esclusivamente frequenze (osservate ed attese).

Utilizzare in loro vece proporzioni o qualunque altro indice, lo rende inapplicabile.

Ma come ottenere le frequenze attese?

Se avessimo solo i dati relativi alle frequenze osservate O_i come potremmo ottenere le frequenze attese E_i ?

Sarà sufficiente riproporzionare le frequenze ai totali marginali, ovvero moltiplicare il totale di riga per il totale di colonna corrispondenti alla cella considerata e dividere per il totale generale

$$\text{Quindi } E_i = (C_i \times R_i) / T$$

Frequenze osservate

		Carattere 1			
		Si	No		
Carattere 2	Si	a	b	a+b	R1
	No	c	d	c+d	R2
		a+c	b+d	a+b+c+d	T
		C1	C2		

Frequenze attese

		Carattere 1			
		Si	No		
Carattere 2	Si	$(C1 \cdot R1) / T$	$(C2 \cdot R1) / T$	a+b	R1
	No	$(C1 \cdot R2) / T$	$(C2 \cdot R2) / T$	c+d	R2
		a+c	b+d	a+b+c+d	T
		C1	C2		

Naturalmente in una tabella 2x2, una volta calcolata la prima frequenza teorica, non sarà necessario procedere con la riproporzione per le altre 3 frequenze, in quanto sarà possibile (ed anche più semplice) calcolarle per differenza

Questo dimostra che in una tabella 2x2 i gradi di libertà del Chi quadro sono realmente

$$\text{GdL} = (2-1) \times (2-1) = 1$$

ovvero una sola delle 4 frequenze teoriche è indipendente e le altre 3 sono condizionate dalla prima calcolata

Esempio traumi cranici e protezione dovuta al casco

Frequenze osservate

		Casco protettivo		
		Si	No	
Trauma Cranico	Si	17	218	235
	No	130	428	558
		147	646	793

Frequenze attese

		Casco protettivo		
		Si	No	
Trauma Cranico	Si	43,56	191,44	235
	No	103,44	454,56	558
		147	646	793

chi-quadrato 28,2555

P<0.00001

Esempio : gradimento di due colluttori

Frequenze osservate

		Risposta			
		Scarso	Suffic.	Buono	
Gruppo	A	35	78	140	253
	B	23	95	120	238
		58	173	260	491

Frequenze attese

		Risposta			
		Scarso	Suffic.	Buono	
Gruppo	A	29,89	89,14	133,97	253
	B	28,11	83,86	126,03	238
		58	173	260	491

chi-quadrato 5,238

p = 0,073

Gradi di libertà = $(2-1)*(3-1) = 2$

Test Chi-quadrato per Tabelle 2x2

Si consideri la tabella, che illustra i risultati di uno studio di validità del test da sforzo positivo (ST-2) nel predire eventi cardiovascolari CV nei 5 anni seguenti in soggetti non sintomatici

I dati si riferiscono ad un campione casuale di **514** soggetti sani arruolati in uno studio

	SFORZO		
Esito	Negativo	ST-2	Totale
Eventi CV	13	21	34
Nessuno	366	114	480
Totale	379	135	514

CV: (morte improvvisa, infarto mortale, infarto semplice, angina)

ST-2: Presenza di anomalie nel tratto ST dell'ECG eseguito sotto sforzo

514 soggetti sono stati arruolati nello studio

- ★ **135** soggetti hanno avuto **esito positivo**;
- ★ **379** soggetti hanno avuto **esito negativo**;
 - * Dei **379** con esito negativo, **13** svilupparono eventi **CV** e **366** no;
 - * Dei **135** con esito positivo, **21** svilupparono eventi **CV** e **114** no.

SFORZO			
Esito	Negativo	ST-2	Totale
Eventi CV	13	21	34
Nessuno	366	114	480
Totale	379	135	514

I numeri all'interno della tabella **13 21 366 e 114** sono le frequenze osservate

IL TEST D'IPOTESI: **Scopo:** valutare capacità predittiva della risposta ST-2 al test da sforzo

H0:

la frazione di soggetti **ST-2** all'ECG da sforzo che **sviluppano eventi CV**

è uguale

alla frazione di soggetti che **sviluppano eventi CV** essendo **NEGATIVI** all'ECG da sforzo

H1:

la frazione di soggetti **ST-2** all'ECG da sforzo che **sviluppano eventi CV**

è diversa

dalla frazione di soggetti che **sviluppano eventi CV** essendo **NEGATIVI** all'ECG da sforzo

Il livello di significatività è posto $\alpha=0,05$

IL TEST CHI-QUADRATO

Il test chi-quadrato confronta le frequenze osservate in ciascuna categoria, rappresentate da **O**, con le frequenze attese sotto l'ipotesi nulla, indicate con **E**.

Il test saggia se le differenze **(O-E)** tra frequenze osservate e attese siano troppo grandi perché siano attribuite al caso. Il test controlla le differenze in tutte le celle.

Per eseguire il test per una tabella di contingenza con (r) righe e (c) colonne, si calcola la somma:

$$\chi^2 = \sum_{i=1}^{r \times c} \frac{(O_i - E_i)^2}{E_i}$$

dove rc è il numero di celle nella tabella.

La distribuzione di probabilità di questa somma è approssimata una distribuzione chi-quadrato con gradi di libertà (gdl) = $(r-1)(c-1)$.

Il **chi quadro** indica la misura in cui le frequenze osservate in ogni casella della tabella differiscono dalle frequenze che ci aspetteremmo se non ci fosse associazione fra i due caratteri.

In altre parole verifica se esiste una relazione fra la variabile che distingue le righe e la variabile che distingue le colonne

LE FREQUENZE ATTESE

La frequenza attesa è calcolata dal *prodotto dei totali di riga e di colonna* diviso dal *totale di tabella*.

SFORZO			
Esito	Negativo	ST-2	Totale
Eventi CV	$(379 * 34) / 514 = \mathbf{25.07}$ 13	$(135 * 34) / 514 = \mathbf{8.93}$ 21	34
Nessuno	$(379 * 480) / 514 = \mathbf{353.93}$ 366	$(135 * 480) / 514 = \mathbf{126.07}$ 114	480
Totale	379	135	514

I conteggi osservati erano **13 21 366 e 114**

Calcolo della statistica Chi-quadrato

Poiché le frequenze attese sono tutte $E_i > 5$, possiamo procedere con il test chi-quadrato come sommatoria:

SFORZO		
Esito	Negativo	ST-2
Eventi CV	$(13-25.05)^2 / 25.05$	$(21-8.93)^2 / 8.93$
Nessuno	$(366-353.95)^2 / 353.95$	$(114-126.07)^2 / 126.07$

$$\chi^2 = 5.8 + 16.31 + 0.41 + 1.16 = 23.68$$

Il test ha una distribuzione chi-quadrato con $(2-1)(2-1)=1$ gdl.

Il valore tabulato vale 3.84.

Conclusione.....

$g \backslash \alpha$	0.005	0.01	0.025	0.05	0.10	0.90	0.95	0.975	0.99	0.995
1	0.00	0.00	0.00	0.00	0.02	2.71	3.84	5.02	6.63	7.88
2	0.01	0.02	0.05	0.10	0.21	4.61	5.99	7.38	9.21	10.60
3	0.07	0.11	0.22	0.35	0.58	6.25	7.81	9.35	11.34	12.84
4	0.21	0.30	0.48	0.71	1.06	7.78	9.49	11.14	13.28	14.86
5	0.41	0.55	0.83	1.15	1.61	9.24	11.07	12.83	15.09	16.75
6	0.68	0.87	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55
7	0.99	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.95
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.72	26.76
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00	34.27
17	5.70	6.41	7.56	8.67	10.09	24.77	27.59	30.19	33.41	35.72
18	6.26	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81	37.16
19	6.84	7.63	8.91	10.12	11.65	27.20	30.14	32.85	36.19	38.58
20	7.43	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57	40.00
21	8.03	8.90	10.28	11.59	13.24	29.62	32.67	35.48	38.93	41.40
22	8.64	9.54	10.98	12.34	14.04	30.81	33.92	36.78	40.29	42.80
23	9.26	10.20	11.69	13.09	14.85	32.01	35.17	38.08	41.64	44.18
24	9.89	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
25	10.52	11.52	13.12	14.61	16.47	34.38	37.65	40.65	44.31	46.93
26	11.16	12.20	13.84	15.38	17.29	35.56	38.89	41.92	45.64	48.29
27	11.81	12.88	14.57	16.15	18.11	36.74	40.11	43.19	46.96	49.64
28	12.46	13.56	15.31	16.93	18.94	37.92	41.34	44.46	48.28	50.99
29	13.12	14.26	16.05	17.71	19.77	39.09	42.56	45.72	49.59	52.34
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67

Tabella dei valori critici:
 La probabilità α che *non* venga superato il valore critico $\chi^2_{\alpha, g}$, nel caso di g gradi di libertà

Limitazioni

Per garantire la validità di questa approssimazione,

- la dimensione del campione deve essere abbastanza grande,
- nessuna cella deve avere frequenza attesa minore di 1,
- e non più del 20% delle celle deve avere frequenza attesa minore di 5.

Una variabile casuale che segue una distribuzione χ^2 quadrato deve essere positiva e assumere valori tra zero ed infinito.

L'area totale sotto la curva della distribuzione è uguale ad uno.

Condizione di applicabilità

Le frequenze attese (teoriche) non devono essere troppo basse

In tal caso infatti si rischierebbe di ottenere un Chi quadro significativo anche quando non dovrebbe esserlo

Infatti, trovandosi al denominatore, le frequenze attese troppo piccole farebbero crescere il valore del test in modo eccessivo falsando così il risultato del test

In particolare

Non oltre il 20% delle frequenze attese può risultare inferiore a 5 (e nessuna inferiore ad 1)

Questo limite è vincolante se il Chi quadro ottenuto ha una $p < \alpha$ e quindi il risultato risulti significativo

Nel caso che invece sia non significativo, il problema non si pone perché il temuto errore per eccesso non si è comunque verificato

Nel caso il Chi quadro non sia applicabile, si possono adottare altri test:

Per tabelle 2x2, si può utilizzare il test esatto di Fisher

Per tabelle superiori (NxM) si può adottare il test G di Woolf, a condizione che nessuna frequenza teorica sia inferiore ad 1

Nel caso il Chi quadro non sia applicabile, una ulteriore soluzione consiste, per tabelle con più righe o colonne, nell'aggregare tra loro, in modo logico, righe o colonne adiacenti

Unendo le celle le frequenze teoriche si sommeranno e sarà, in genere, agevole raggiungere frequenze che soddisfino la condizione di applicabilità del Chi quadro

Inoltre, aggregando righe o colonne, si avranno GdL inferiori e quindi, a parità di altre condizioni, si raggiungerà più facilmente la soglia della significatività statistica

Esempio *terapia farmacologica nell'IMA*

Questi dati si riferiscono ad una sperimentazione clinica del propranololo su pazienti affetti da infarto miocardico. I due gruppi di pazienti sono quelli trattati con propranololo e un gruppo di controllo che non riceve il farmaco.

La risposta dicotomica consisteva nell'essere ciascun paziente ancora vivo al ventottesimo giorno dopo la sua ammissione allo studio. o nell'essere egli venuto a mancare in un certo momento compreso entro questo periodo di 28 giorni.

Risposta	Trattato propranololo	trattato placebo	Totale
Sopravvissuti 28 giorni	38	29	67
Non sopravvissuti	7	17	24
Totale	45	46	91

Esempio *terapia farmacologica nell'IMA*

- Questi risultati campionari forniscono sufficiente evidenza che il propranololo aumenta il tasso di sopravvivenza nei 28 giorni se confrontato con un controllo?
- Se di fatto non ci fosse differenza nei tassi di sopravvivenza tra la sottostante popolazione dei trattati con propranololo e la popolazione di controllo, è verosimile che con i campioni di cui sopra si possa osservare uno scarto uguale o maggiore a quello osservato, puramente per caso?
- Poniamo il livello di significatività del 5%

Tabella frequenza osservate

Risposta	Trattato propranololo	trattato placebo	Totale
Sopravvissuti 28 giorni	38	29	67
Non sopravvissuti	7	17	24
Totale	45	46	91

Tabella frequenza attese

Risposta	Trattato propranololo	trattato placebo	Totale
Sopravvissuti 28 giorni	33.13	33.87	67
Non sopravvissuti	11.87	12.13	24
Totale	45	46	91

- Calcolato il valore atteso di una cella, poiché i totali marginali di riga e di colonna sono fissi, le altre quantità attese si possono derivare per sottrazione dai totali marginali: esiste quindi una sola quantità attesa indipendente e per questo si dice che per una tabella di contingenza 2x2 vi è un grado di libertà (GdL) per il calcolo del chi-quadrato.
- Il calcolo del chi-quadrato:

$$\chi^2 = 0.715 + 0.7 + 1.997 + 1.953 = 5.365$$

- Nella tabella per la distribuzione chi-quadrato con 1 GdL si nota che il valore calcolato è compreso fra i valori 5.02 e 6.63, corrispondenti a $p=0.025$ e $p=0.01$.
- Quindi la differenza fra le due mortalità è significativa perchè $p<0.05$.
- La conclusione è che i pazienti trattati con propranololo hanno un tasso di sopravvivenza a 28 giorni significativamente maggiore di quello del gruppo di controllo non trattato col propranololo.
- In altre parole la differenza osservata nei tassi di sopravvivenza non può spiegarsi come effetto del caso o della fluttuazione di campionamento.
- La stessa procedura di calcolo si può estendere a tabelle $2 * k$ o addirittura $r * k$ per confrontare più campioni.
- I GdL di una tabella $r * k$ saranno $(r-1) * (k-1)$.

Esempio: test Chi-Quadrato

I seguenti dati sono relativi all'utilizzo di antibiotici registrato in un anno in due gruppi di uomini e donne. Si vuole verificare se esiste una differenza significativa nei due gruppi.

	antibiotici		
sex	No	Si	tot
M	35	20	55
F	54	8	74
tot	89	28	129

Ipotesi e tipo di Test: H_0 : sesso e antibiotici sono indipendenti, H_1 : sono associati; Test Chi-Quadrato

freq.	37.9	11.9
attese	51	16

Ipotesi del test: tutte le frequenze attese sono >5 : il test è valido.

Calcolo della statistica test: (...) $\chi^2 = 9.8$

Metodo delle regioni di rifiuto: Fissiamo $\alpha=0.05$; il test è sempre a una coda, a destra; il valore soglia in corrispondenza di 1 grado di libertà (letto sulle tavole) è 3.841. La statistica test cade nella zona di rifiuto.

Conclusione: rifiutiamo l'ipotesi di indipendenza.

TABELLE $r \times c$

In questa situazione, i dati vengono organizzati in una tabella di contingenza $r \times c$ dove r è il numero di righe e c è il numero di colonne.

Esempio: Uno studio valuta l'accuratezza dei certificati di morte.

Scopo: Verificare se la compilazione dei certificati di morte in due Ospedali, indicati con A e B, è parimenti affidabile.

Metodi: Le cause di morte, riportate nel certificato di decesso, in due Ospedali sono confrontate con il referto di 575 autopsie. La tabella di contingenza 2×3 mostra l'esito del confronto.

	Certificato di morte			
Ospedale	Accurato	Carente	Inesatto	Totale
A	157	18	54	229
B	268	44	34	346
Totale	425	62	88	575

TABELLE r x c

Saggiamo l'ipotesi H_0 : le proporzioni delle categorie di certificati di morte nell'ospedale A sono uguali a quelle nell'ospedale B, l'ipotesi H_1 : le proporzioni sono diverse.

Utilizziamo il test chi quadrato con un livello di significatività $\alpha = 0,05$.

Calcoliamo le frequenze attese; ad esempio, ci aspettiamo che:

$229 \times 425 / 575 = 169,3$ si riscontrino nell'ospedale A,

$346 \times 425 / 575 = 255,7$ si riscontrino nell'ospedale B.

Attesi	Certificato di morte			
Ospedale	Accurato	Carente	Inesatto	Totale
A	169,3	24,7	35,0	229
B	255,7	37,3	53,0	346
Totale	425,0	62,0	88,0	575

TABELLE r x c

$$\frac{(157 - 169,3)^2}{169,3} + \frac{(18 - 24,7)^2}{24,7} + \frac{(54 - 35,0)^2}{35,0} + \frac{(268 - 255,7)^2}{255,7} + \frac{(44 - 37,3)^2}{37,3} + \frac{(34 - 53,0)^2}{53,0} =$$

$$\chi^2 = 0,89 + 1,82 + 10,31 + 0,59 + 6,81 + 1,20 = 21,62.$$

Per una distribuzione χ^2 con $(2-1)(3-1) = 2$ gdl , $p < 0.001$.

Poiché p è minore di α , rifiutiamo l'ipotesi nulla e concludiamo che le proporzioni dei certificati di morte nell'ospedale A non sono uguali per le tre categorie.

Conclusioni: L'Ospedale A presenta una proporzione maggiore di certificati che richiedono la riformulazione della causa di morte.

ODDS RATIO

Cerchiamo di stimare l'intensità dell'associazione tra due variabili nominali. In una tabella 2x2, tale stima è **l'odds ratio**.

Se un evento si verifica con probabilità p , l'odds in favore dell'evento è: $p/(1-p)$

$\text{odds} = \text{probabilità} / (1 - \text{probabilità})$

$\text{probabilità} = \text{odds} / (1 + \text{odds})$

Definizione di Odds e di P(e)

- Se un team ha la probabilità di vincere =80% il suo odds di vittoria è 4:1
- La prob di estrarre una carta rossa =1/2 l'odds è 1:1

Esempio: Se due variabili casuali dicotomiche associate indicano *malattia* (m) ed *esposizione (E) a fattore di rischio*, allora l'**odds ratio** è il rapporto tra l'odds di malattia tra i soggetti esposti e l'odds di malattia tra i soggetti non esposti, ovvero:

$$OR = \left[\frac{P(M|E)}{1 - P(M|E)} \right] / \left[\frac{P(M|\bar{E})}{1 - P(M|\bar{E})} \right]$$

ODDS RATIO

Un campione generico di n soggetti (malati, non malati) (esposti, non esposti) può dare origine ad una tabella di contingenza 2x2 di seguito riportata.

		esposizione		
malattia	Si	No	Totale	
Si	a	b	a+b	
No	c	d	c+d	
Totale	a+c	b+d	n=a+b+c+d	

Con questi rapporti esprimiamo lo stimatore dell'odds ratio:

$$OR = \left(\frac{a}{a+c} / \frac{c}{a+c} \right) / \left(\frac{b}{b+d} / \frac{d}{b+d} \right) = [a/c] / [b/d] = ad / bc$$

esempio

Scopo: determinare se il monitoraggio elettronico fetale durante il parto facilita la decisione di parto cesareo,

Metodi: uno studio ha incluso 5.824 neonati e di questi ne ha sottoposti 2.850 a monitoraggio e 2.974 no.

Monitoraggio elettronico fetale			
Parto cesareo	Si	No	Totale
Si	358	229	587
No	2492	2745	5237
Totale	2850	2974	5824

L'odds ratio di parto cesareo nel gruppo sottoposto a monitoraggio fetale verso il gruppo non sottoposto è:

$$OR = [(358) (2.745)] / [(229) (2.492)] = 1,72.$$

intervallo di confidenza di OR

Sembra esistere una moderata associazione tra l'utilizzo del monitoraggio ed il tipo di parto.

Nota: ciò non implica, tuttavia, che il monitoraggio elettronico causi un parto cesareo; è possibile che i feti a maggior rischio di parto cesareo siano quelli sottoposti a monitoraggio.

L'incertezza di questa stima è riflessa dall'ampiezza del intervallo di confidenza (IC) di OR;

Si ricordi che l'espressione per l'I.C. al 95% per una media

$$(\bar{x} - 1,96 \times es(\bar{x}) , \bar{x} + 1,96 \times es(\bar{x}))$$

si basa sulla assunzione che i valori della popolazione originaria siano normalmente distribuiti.

La distribuzione di probabilità dell'odds ratio è asimmetrica a destra; l'odds ratio assume solo valori positivi tra 0 ed infinito. Al contrario, la distribuzione di probabilità del logaritmo naturale dell'odds ratio è più simmetrica ed approssimativamente normale.

$$[\ln(OR) - 1.96 \text{ es}\{\ln(OR)\}, \ln(OR) + 1.96 \text{ es}\{\ln(OR)\}]$$

Pertanto, per calcolare un intervallo di confidenza per $\ln(OR)$ prima di tutto dobbiamo conoscere l'errore standard (es) di questa quantità. Per una tabella 2x2 rappresentata nel modo seguente :

L'errore standard di $\log(OR)$ è stimato da

$$\text{es}(\log(OR)) = [(1/a + 1/b + 1/c + 1/d)]^{1/2}$$

intervallo di confidenza di OR

Se uno dei valori della tabella è uguale a zero , l'errore standard non è definito. In questo caso, aggiungendo 0,5 ad ogni valore correggeremo la situazione. Pertanto, la stima modificata dell'errore standard è:

$$\sqrt{\frac{1}{a+0.5} + \frac{1}{b+0.5} + \frac{1}{c+0.5} + \frac{1}{d+0.5}}$$

La stima appropriata può essere sostituita nell'espressione precedente.

Per trovare L'IC al 95% per l'odds ratio, calcoliamo l'anti-logaritmo dei limiti inferiore e superiore per ottenere:

$$(e^{\ln(\text{OR}) - 1,96 \text{ es } [\ln(\text{OR})]} ; e^{\ln(\text{OR}) + 1,96 \text{ es } [\ln(\text{OR})]}).$$

Per la relazione tra monitoraggio elettronico fetale e tipo di parto, il logaritmo dell'odds ratio stimato è: $\ln(\text{OR}) = \ln(1,72) = 0,542$.

L'espressione dell'I.C. al 95% per il logaritmo naturale dell'odds ratio

$$\ln(\text{OR}) - 1,96 \times es[\ln(\text{OR})] ; \ln(\text{OR}) + 1,96 \times es[\ln(\text{OR})]$$

La stima dell'errore standard di $\ln(\text{OR})$ è:

$$\begin{aligned} es[\ln(\text{OR})] &= [(1/a + 1/b + 1/c + 1/d)]^{1/2} \\ &= (1/358 + 1/229 + 1/2492 + 1/2745)^{1/2} = 0,089 \end{aligned}$$

L'IC al 95% per il logaritmo dell'odds ratio è:

$$(0,542 - 1,96 \times (0,089) , 0,542 + 1,96 \times (0,089)) = (0.368 , 0.716).$$

L'IC al 95% per l'odds ratio è: $[(\exp(0,368), \exp(0,716))] = (1.44 , 2.05)$.

Siamo confidenti al 95% che l'odds di parto cesareo tra i feti sottoposti a monitoraggio durante il travaglio è da 1,44 a 2,05 volte maggiore dell'odds dei feti non sottoposti a monitoraggio. **Si noti che questo intervallo non include il valore 1.**

Am J Epidemiol. 2019 Nov 29.

Observational Study of the Association between Participation in High School Football and Self-Rated Health, Obesity, and Pain in Adulthood.

Abstract

American football is the most popular high school sport yet its association with health in adulthood has not been widely studied. We investigated the association between high school football and self-rated health, obesity, and pain in adulthood using a retrospective cohort study of the Wisconsin Longitudinal Study from 1957 to 2004. We matched 925 high school males who played varsity football with 1,521 males who did not play football. After matching, playing football was not associated with poor or fair self-rated health (odds ratio [OR] 0.88, 95% confidence interval [CI]: 0.63, 1.24; $P = 0.48$) or pain that limited activities (OR 0.86, 95% CI: 0.59, 1.25; $P = 0.43$) at age 65. Football was associated with an obese body mass index in adulthood (OR 1.32, 95% CI: 1.06, 1.64; $P = 0.01$). In conclusion, our findings suggest that playing football in high school was not significantly associated with poor or fair self-related health at the age of 65 but was associated with the risk of being obese as an adult compared to not playing football in high school. Our findings provide needed information about the risk of playing football to broader set of health outcomes.

Br J Anaesth. 2019 Nov 23

Association between postoperative complications and lingering post-surgical pain: an observational cohort study.

Abstract

BACKGROUND:

Post-surgical pain that lingers beyond the initial few-week period of tissue healing is a major predictor of pain chronification, which leads to substantial disability and new persistent opioid analgesic use. We investigated whether postoperative medical complications increase the risk of lingering post-surgical pain.

METHODS:

The study population consisted of patients undergoing diverse elective surgical procedures in an academic referral centre in the USA, between September 2013 and May 2017. Multivariable logistic regression, adjusting for confounding variables and patient-specific risk factors, was used to test for an independent association between any major postoperative complication and functionally limiting lingering pain 1-3 months after surgery, as obtained from patient self-reports.

RESULTS:

The cohort included 11 986 adult surgical patients; 10 562 with complete data. At least one complication (cardiovascular, respiratory, renal/gastrointestinal, wound, thrombotic, or neural) was reported by 13.3% (95% confidence interval: 12.7-14.0) of patients, and 19.7% (19.0-20.5%) reported functionally limiting lingering post-surgical pain. After adjusting for known risk factors, the patients were twice as likely (odds ratio: 2.04; 1.78-2.35) to report lingering post-surgical pain if they also self-reported a postoperative complication. Experiencing a complication was also independently predictive of lingering post-surgical pain (odds ratio: 1.95; 1.26-3.04) when complication data were extracted from the National Surgical Quality Improvement Program registry, instead of being obtained from patient self-report.

CONCLUSIONS:

Medical complications were associated with a two-fold increase in functionally limiting pain 1-3 months after surgery. Understanding the mechanisms that link complications to pathological persistence of pain could help develop future approaches to prevent persistent post-surgical pain.