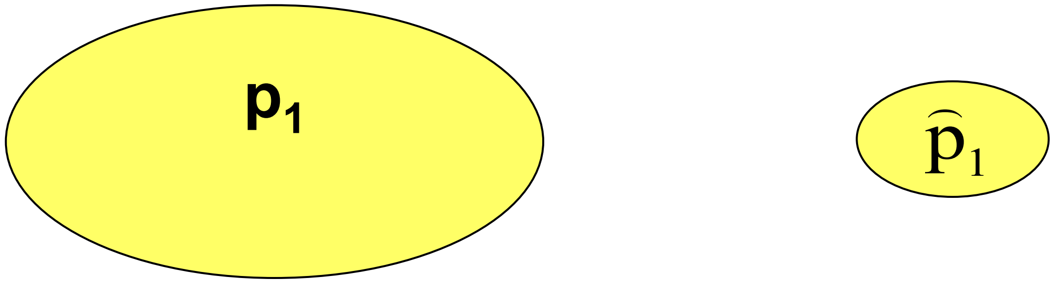
VERIFICA DI IPOTESI PER UNA PROPORZIONE



DATI

Si suppone che da una certa popolazione sia stato estratto un campione di 300 elementi. Nel campione si osservano 123 casi di una certa malattia. Si stima la proporzione di malati (evento "successo")

$\hat{p} = 123 / 300 = 0,41$

ASSUNZIONI

La variabile che conta il numero di successi sul numero delle prove segue una distribuzione binomiale.
Nel caso in cui $n \to \infty$ e $p \to 0,5$ è possibile approssimare una Distribuzione binomiale ad una distribuzione di Gauss.

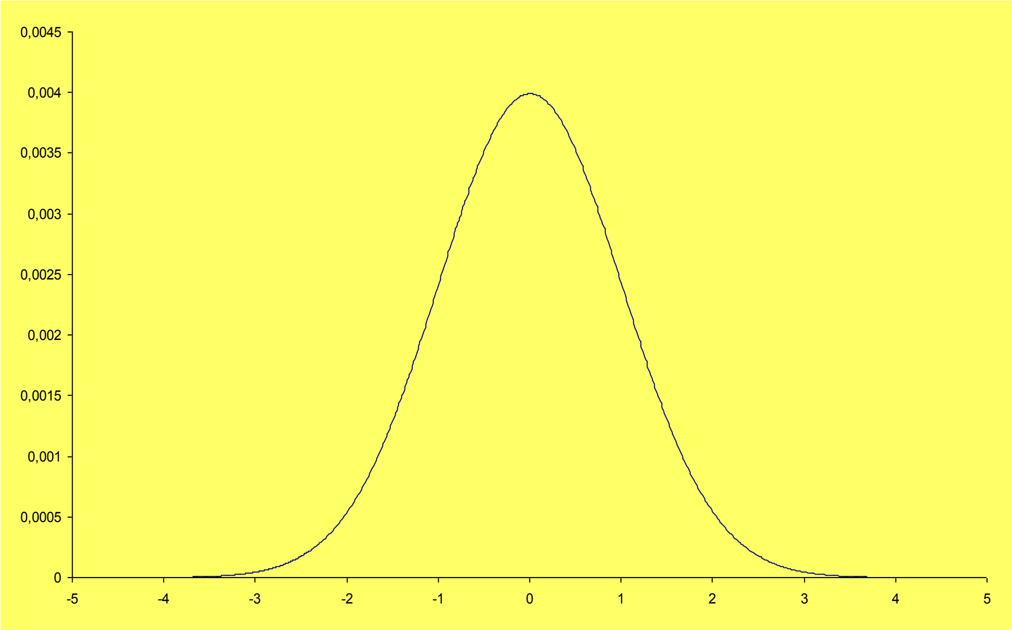
IPOTESI

$$\begin{cases} H_0: p = p_0 & \text{con } p_0 = 0,5 \\ H_1: p \neq p_0 \end{cases}$$

STATISTICA TEST

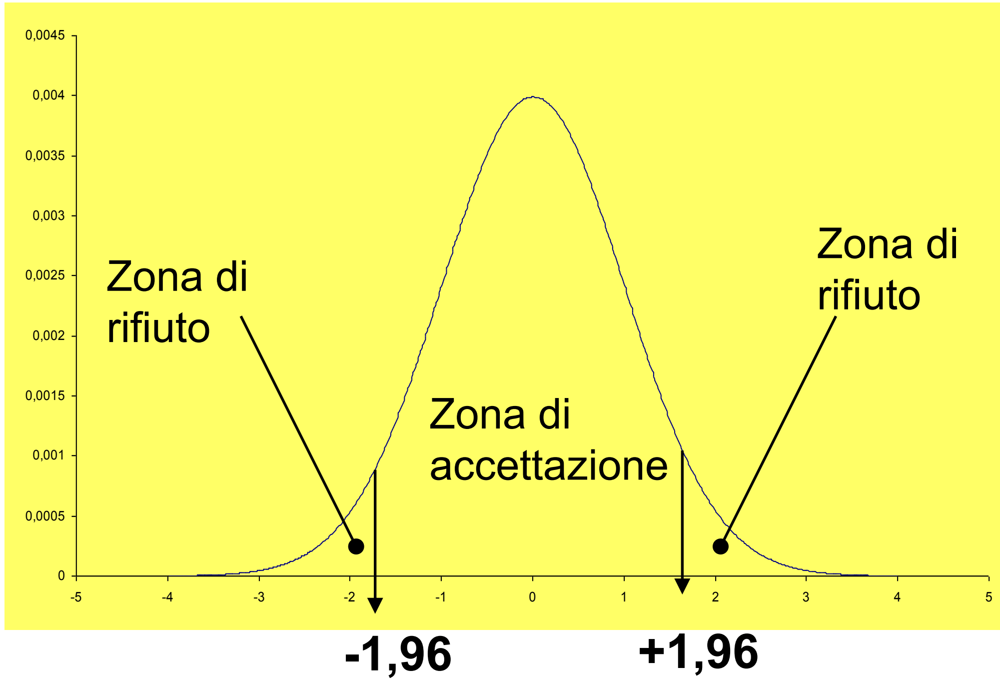
$$z = \frac{x - np_0 - 1/2}{\sqrt{np_0(1-p_0)}} = \frac{x - np_0 - 1/2}{\sqrt{np_0 q_0}}$$

DISTRIBUZIONE DELLA STATISTICA



Vera l'ipotesi nulla, applicando il teorema centrale del limite, la statistica test segue una distribuzione di Gauss standard N (0,1).

REGOLA DI DECISIONE



Fisso α, la probabilità di rifiutare l’ipotesi nulla quando è vera, accettabilmente bassa: 0,05. Per verificare l’ipotesi formulata dovremo suddividere a nei due lati della curva (α/2 nelle code). Si individueranno in questo modo i limiti della zona di rifiuto (code); sapendo che la distribuzione della statistica è di Gauss standard i limiti si ricercheranno nelle apposite tavole e saranno uguali ed opposti.

CALCOLO DELLA STATISTICA TEST

$$z = \frac{123 - (300 \times 0{,}5) - 1/2}{\sqrt{300(0{,}5)(0{,}5)}} = \frac{-27{,}5}{8{,}66} = -3{,}17$$

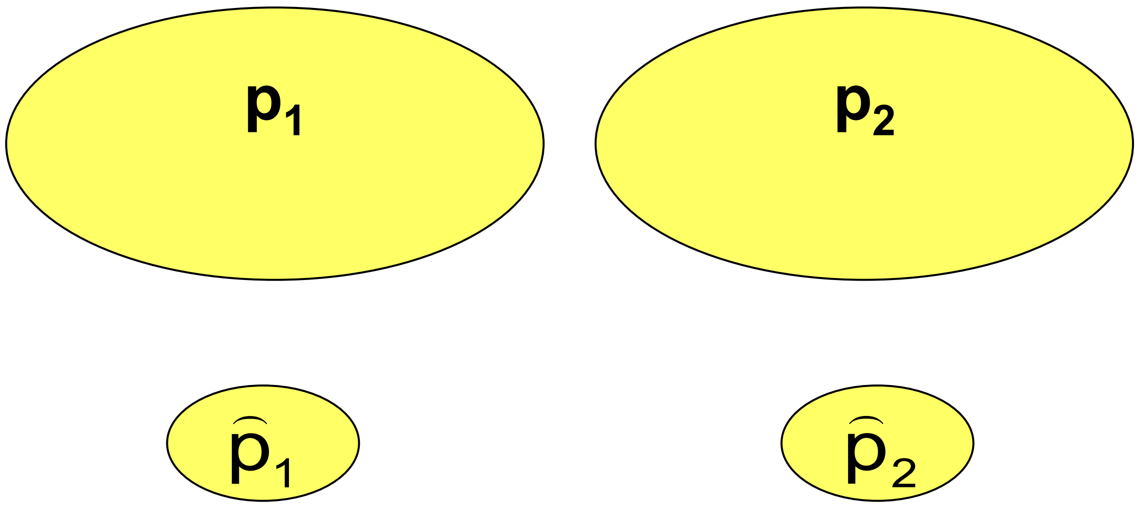
DECISIONE STATISTICA

Z=-3,17 cade nella zona di rifiuto (-3,17<-1,96) quindi rifiuterò l'ipotesi nulla.

DECISIONE DEL RICERCATORE

La frequenza con la quale la malattia in esame si presenta nella popolazione in studio è differente dalla frequenza della stessa malattia nella popolazione di "riferimento".

VERIFICA DI IPOTESI PER IL CONFRONTO TRA DUE PROPORZIONI



DATI
Si dispone del numero di pazienti trattati con due diverse terapie per una certa malattia.

$n_1 = 100$ $\widehat{p}_1$ = proporzione di guariti nel primo gruppo = 78 /100 = 0,78

$n_2 = 100$ $\widehat{p}_2$ = proporzione di guariti nel secondo gruppo = 90 /100 = 0,90

ASSUNZIONI

Campioni indipendenti.

La variabile in studio conta il numero di guarigioni ("successo") sul totale delle prove (numerosità del campione): segue una distribuzione binomiale.

Sarà possibile approssimare la distribuzione binomiale ad una distribuzione di Gauss standard se $n \to \infty$ e $p \to 0{,}5$.

IPOTESI

$$\begin{cases} H_0: p_1 = p_2 \\ H_1: p_1 \neq p_2 \end{cases} \qquad p_1 - p_2 = 0$$

STATISTICA TEST

$$Z = \frac{\dfrac{X_1}{n_1} - \dfrac{X_2}{n_2} - (p_1 - p_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}} =$$

$$= \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$$

Dove

$\hat{p} = (X_1 + X_2) / (n_1 + n_2)$

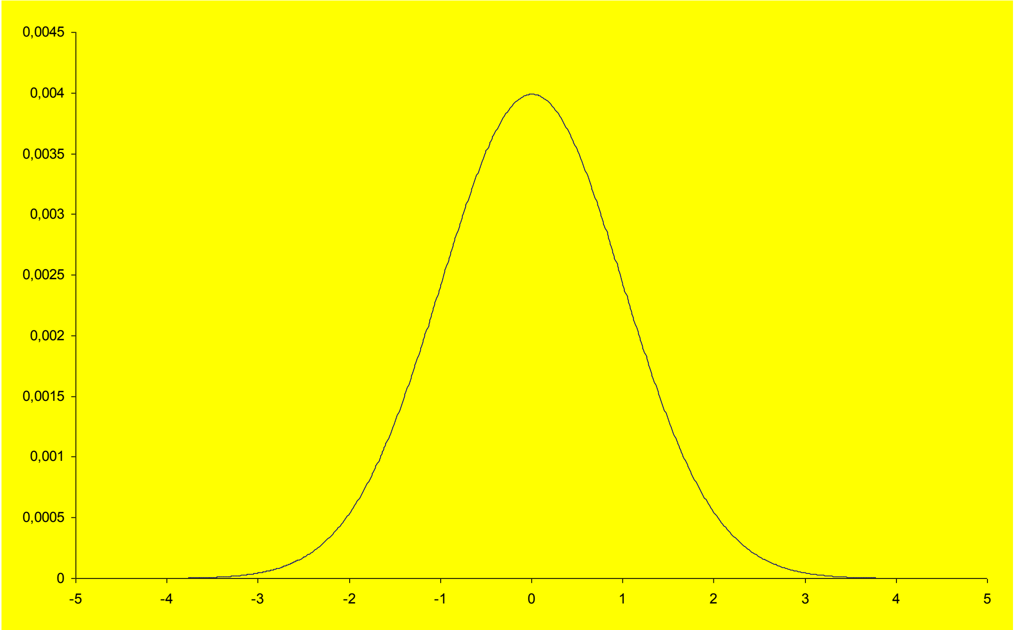
$\hat{q} = 1 - \hat{p}$

$\hat{p}_1 = X_1 / n_1$

$\hat{p}_2 = X_2 / n_2$

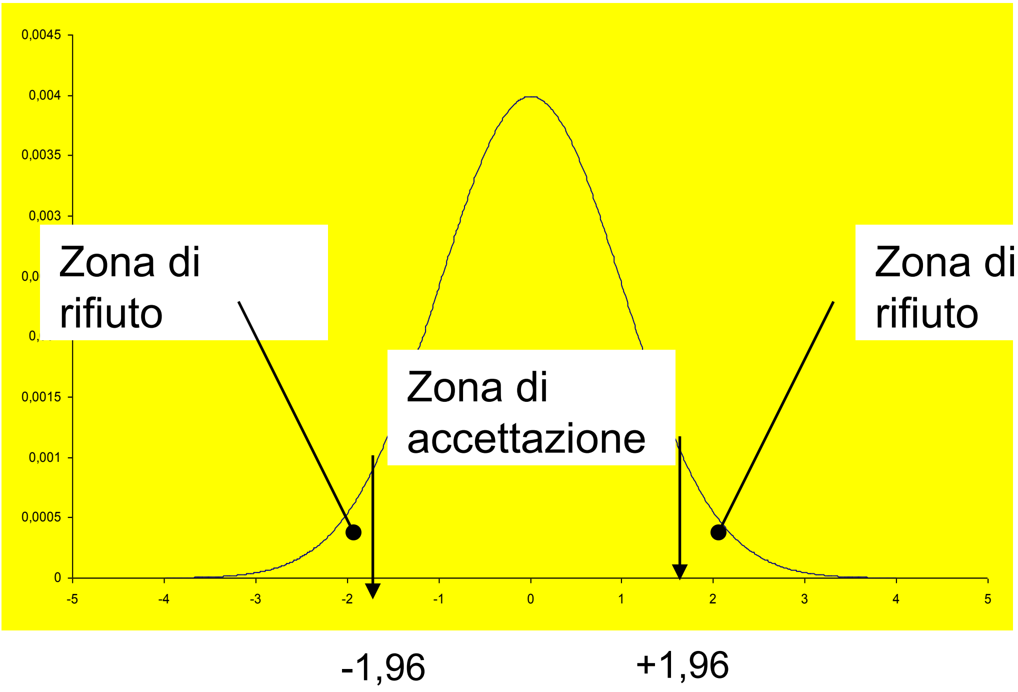
DISTRIBUZIONE DELLA STATISTICA

Con l'ipotesi nulla vera, applicando il teorema del limite centrale, la statistica test segue una distribuzione di Gauss standard N (0,1).



REGOLA DI DECISIONE

• Fisso α, la probabilità di rifiutare l'ipotesi nulla quando è vera, accettabilmente bassa: 0,05.
• Si individuano in questo modo i limiti della zona di rifiuto (code);
• Sapendo che la distribuzione della statistica è di Gauss standard i limiti si ricercheranno nelle apposite tavole e saranno uguali ed opposti.



CALCOLO DELLA STATISTICA TEST

$\hat{p}$ = (78+90) / 200 = 168 /200 =0,84
$\hat{q}$ = 1-0,84 = 0,16

$$z = \frac{\frac{78}{100} - \frac{90}{100}}{\sqrt{0{,}84 \times 0{,}16 \times \left(\frac{1}{100} + \frac{1}{100}\right)}} =$$

$$= \frac{-0{,}12}{0{,}058} = -2{,}32$$

DECISIONE STATISTICA

-2,32<-1,96 pertanto rifiuto H_0

DECISIONE DEL RICERCATORE

La differenza tra i due trattamenti non è casuale, il trattamento a cui è stato sottoposto il secondo gruppo è più efficace.

Quando non è possibile approssimare la distribuzione binomiale ad una Gauss il test X^2 permette anche di eseguire confronti tra proporzioni.

DATI

Criterio 1 – Trattamento

Criterio2	nuovo	vecchio	Totale
Guariti	90	78	168
Non guariti	10	22	32
Totale	100	100	200

Criterio 1

Criterio2	1	2	Totale
1	a	b	a+b
2	c	d	c+d
Totale	a+c	b+d	N

IPOTESI

$$\begin{cases} H_0: & p_1 = p_2 \\ H_1: & p_1 \neq p_2 \end{cases}$$

STATISTICA TEST

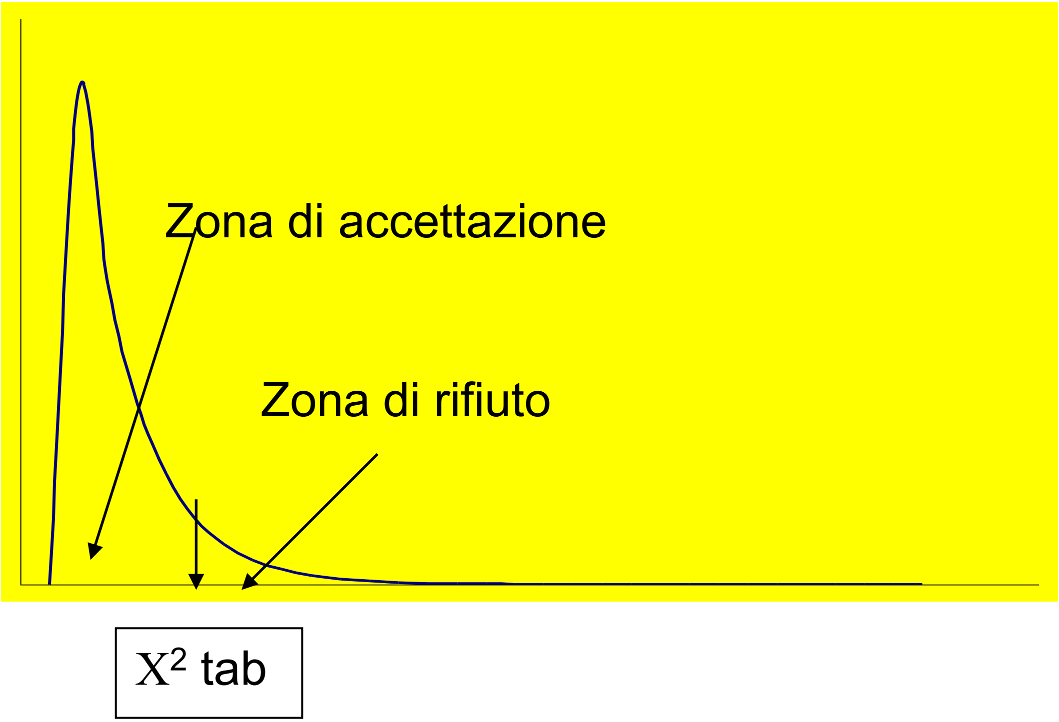
$$X^2 = \frac{N(ad - bc)^2}{(a+b)\times(c+d)\times(a+c)\times(b+d)}$$

$$X^2 = \frac{N(|ad - bc| - 0,5N)^2}{(a+b)\times(c+d)\times(a+c)\times(b+d)}$$

Nella seconda formula c'è la correzione per la continuità di Yates

DISTRIBUZIONE DELLA STATISTICA TEST

La distribuzione della statistica test è una X^2 ed è caratterizzata dai gradi di libertà.



REGOLA DI DECISIONE

Fissato α accettabilmente piccolo (0,05), troverò sulle tavole X^2 un valore in corrispondenza di α prescelto e dei gradi di libertà della statistica che nelle tabelle 2x2 sono sempre pari a 1.

Se il valore calcolato della statistica è maggiore del valore tabulato rifiuterò l'ipotesi nulla, se invece il valore calcolato è minore del tabulato accetterò l'ipotesi nulla.

Criterio 1

Criterio2	nuovo	vecchio	Totale
Guariti	90	78	168
Non guariti	10	22	32
Totale	100	100	200

$$X^2 = \frac{200(90 \times 22 - 78 \times 10)^2}{100 \times 100 \times 168 \times 32} = 5{,}357$$

DISTRIBUZIONE DELLA STATISTICA TEST

E' una distribuzione X^2 con gradi di libertà 1.

REGOLA DI DECISIONE

Con $\alpha = 0{,}05$ e 1 grado di libertà il valore X^2 tabulato è 3,841.

Rifiuto l'ipotesi nulla poiché

$X^2 calc=5.37 > X^2 tab=3{,}841$

I due trattamenti producono un effetto differente