ESERCIZIO 7.1

In una banca i libretti di risparmio sono caratterizzati da un importo medio di 5.000 Euro, con varianza pari a 1225. Al fine di verificare se la situazione è variata, il direttore seleziona un campione di 30 depositanti, da cui risulta una media pari a 4.993 Euro. A quale conclusione deve giungere se utilizza un livello di significatività del 99%?

Soluzione

Ipotesi:

$H_0 : \mu = 5000$

$H_1 : \mu \neq 5000$

Dati:

$\sigma^2 = 1225$

$n = 30$

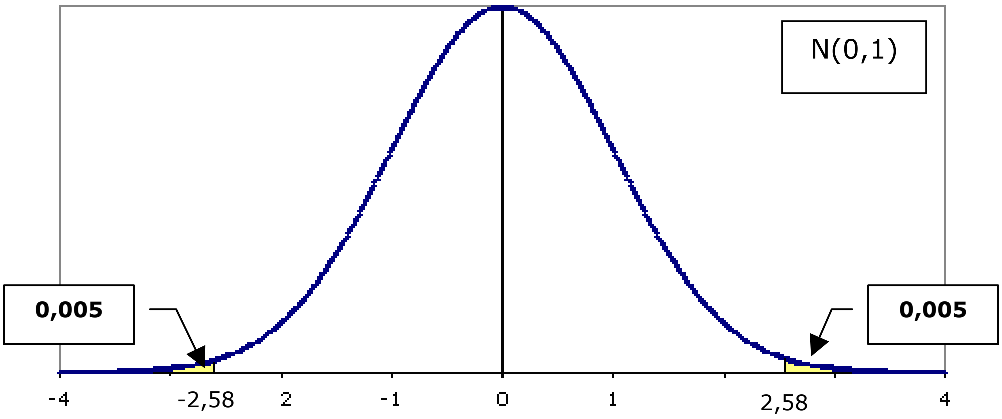
$\alpha = 0{,}01;\ \alpha/2 = 0{,}005$

Statistica test:

$$X_{test} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Valori critici:

$\pm z_{0,005} = \pm 2{,}58$



Regola di decisione:

$$-2{,}58 \leq X_{test} \leq +2{,}58 \quad \Rightarrow \text{si accetta } H_0$$

$$X_{test} < -2{,}58; X_{test} > +2{,}58 \quad \Rightarrow \text{si rifiuta } H_0$$

Valore test:

$$x_{test} = \frac{4993 - 5000}{35/\sqrt{30}} = \frac{-7}{6{,}39} = -1{,}09$$

Decisione:

$$-2{,}58 \leq x_{test} \leq +2{,}58 \quad \Rightarrow \quad \boxed{\textbf{Accettiamo } H_0}$$

ESERCIZIO 7.2

In una fabbrica di lampadine il processo di produzione è definito "sotto controllo" se la durata media delle lampadine è pari a 375 ore. Un campione casuale di 49 lampadine ha una durata media di 350 ore. Sapendo che la varianza del processo produttivo è pari a 100, si verifichi al livello del 95% se tale rilevazione può essere considerata un "fuori controllo".

Soluzione

Un processo di produzione deve rispettare le specifiche di prodotto, quindi va considerato fuori controllo ogni volta che le specifiche non vengono rispettate. In teoria, quindi, in questo caso, non conoscendo le specifiche, dovremmo formulare un'ipotesi alternativa bilaterale, ma, dato il tipo di prodotto, è presumibile che eventuali differenze in eccesso rispetto allo standard costituiscano un risultato "auspicabile" piuttosto che un "fuori controllo" da evitare. In casi come questo sembra, dunque, più ragionevole ricorrere ad un test unilaterale.

Ipotesi:

$H_0 : \mu = 375$

$H_1 : \mu < 375$

Dati:

$\sigma^2 = 100$

$n = 49$

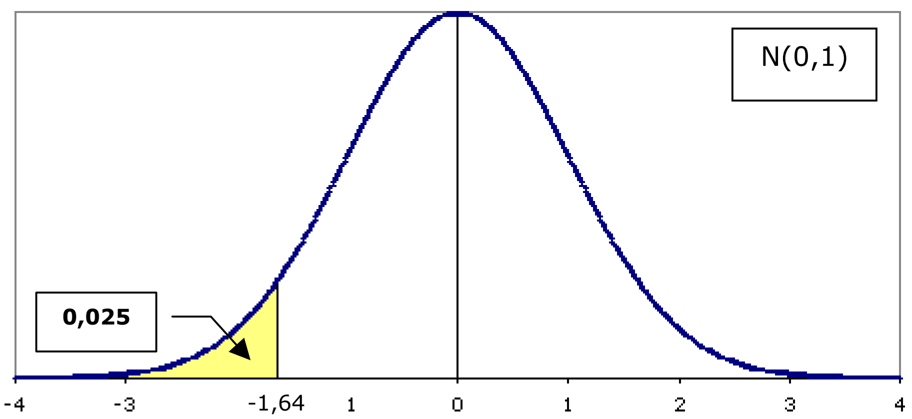
$\alpha = 0,05$

Statistica test:

$$X_{test} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$

Valore critico:

$- z_{0,025} = - 1,65$



Regola di decisione:

$$X_{test} \geq -1,65 \quad \Rightarrow \text{si accetta } H_0$$

$$X_{test} < -1,65 \quad \Rightarrow \text{si rifiuta } H_0$$

Valore test:

$$x_{test} = \frac{350-375}{10/7} = \frac{-25}{1,429} = -17,5$$

Decisione:

$x_{test} < -1,65 \Rightarrow$ $\boxed{\textbf{Rifiutiamo } H_0}$

ESERCIZIO 7.3

Una società di gestione di carte di credito sa che il traffico medio mensile dei suoi clienti è pari a 75 Euro. Estraendo un campione di 100 conti si rileva che il debito medio è di 83,4 Euro, con uno scarto quadratico medio campionario corretto di 23,65 Euro. È lecito concludere, al livello di significatività del 95%, che il bilancio medio è diverso da 75 Euro?

Soluzione

Ipotesi:

$H_0 : \mu = 75$

$H_1 : \mu \neq 75$

Dati:

$\hat{s} = 23,65$

$n = 100$

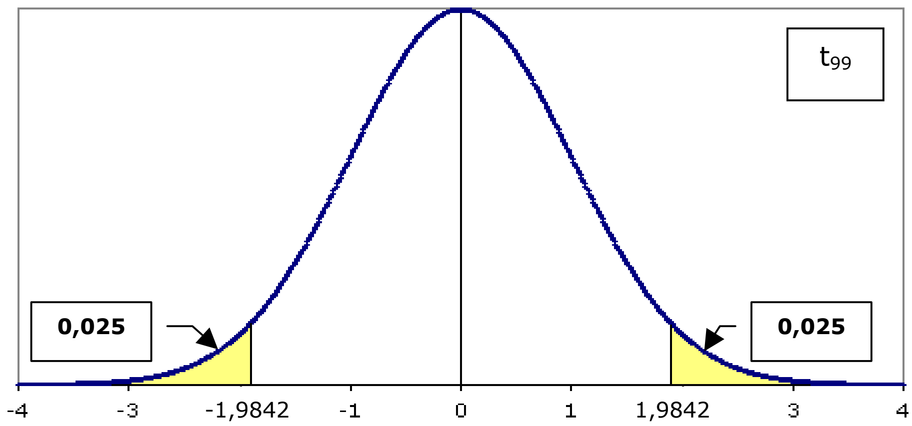
$\alpha = 0,05;\ \alpha/2 = 0.025$

Statistica test:

$$x_{test} = \frac{\bar{x} - \mu}{\hat{s}/\sqrt{n}} \sim t_{n-1}$$

Valori critici:

$\pm\, t_{0,025;\ 99} = \pm\, 1,9842$



Regola di decisione:

$$-1,9842 \le x_{test} \le +1,9842 \quad \Rightarrow \text{si accetta } H_0$$

$$x_{test} < -1,9842;\ x_{test} > +1,9842 \quad \Rightarrow \text{si rifiuta } H_0$$

Valore test:

$$x_{test} = \frac{83,40 - 75}{23,65/10} = \frac{83 - 75}{\sqrt{0,64}} = \frac{8,4}{2,365} = 3,55$$

Decisione:

$x_{test} > 1,9842 \quad \Rightarrow$ **Rifiutiamo H_0**

ESERCIZIO 7.4

Il filo metallico prodotto dalla ditta A è caratterizzato da un punto di rottura di 80 Kg. Un cliente seleziona a caso 9 matasse sulle quali rileva un punto di rottura medio di 79 Kg. e una varianza pari a 16: può procedere all'acquisto della partita, al livello di significatività del 95%?

Soluzione

Ipotesi:

$H_0 : \mu = 80$
$H_1 : \mu < 80$

Dati:

$\overline{x} = 79$ Kg.
$\hat{s}^2 = 16$
$n = 9$

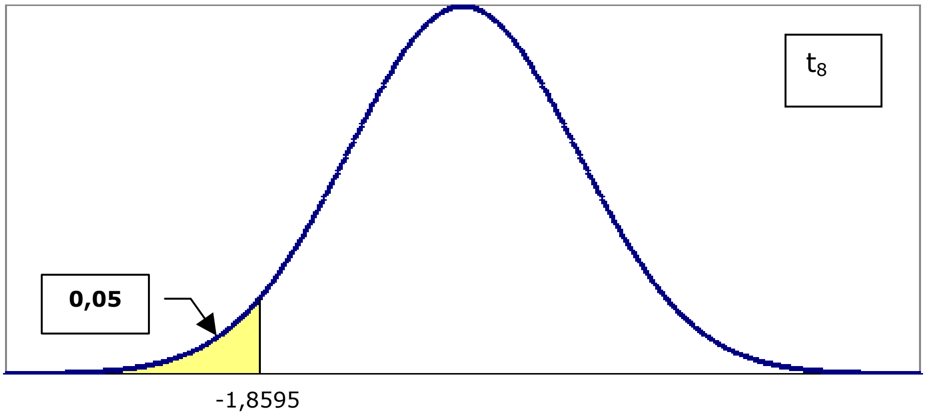
$\alpha = 0,05$

Statistica test:

$$x_{test} = \frac{\overline{x} - \mu}{\hat{s}/\sqrt{n}} \sim t_{n-1}$$

Valore critico:

$$-t_{0,05;8} = -\ 1,8595$$



Regola di decisione:

$$-1,8595 \le x_T \le +1,8595 \quad \Rightarrow \text{ si accetta } H_0$$
$$x_{test} < -1,8595; x_{test} > +1,8595 \quad \Rightarrow \text{ si rifiuta } H_0$$

Valore test:

$$x_t = \frac{79 - 80}{4/3} = \frac{3}{4} = 0,75$$

Decisione:

$$-1,8595 \le x_{test} \le +1,8595 \quad \Rightarrow \quad \boxed{\textbf{Accettiamo } H_0}$$

ESERCIZIO 7.5

Per verificare se gli studenti di due scuole A e B hanno lo stesso rendimento due campioni di 100 studenti vengono sottoposti ad un test. I punteggi medi riportati dai due campioni sono $\overline{x}_A = 83$ e $\overline{x}_B = 75$. Si ha motivo di ritenere che la varianza sia identica per le due scuole e pari a 32.

a) Si verifichi se vi sia o meno una differenza significativa, al livello del 95%, tra le due scuole;

b) si costruisca un intervallo di confidenza per μ_A - μ_B al livello del 95%

Soluzione

Ipotesi:

$H_0 : \mu_A = \mu_B$

$H_1 : \mu_A \neq \mu_B$

Dati:

$\sigma_A^2 = \sigma_B^2 = \sigma^2 = 32$

$n_A = n_B = 100$

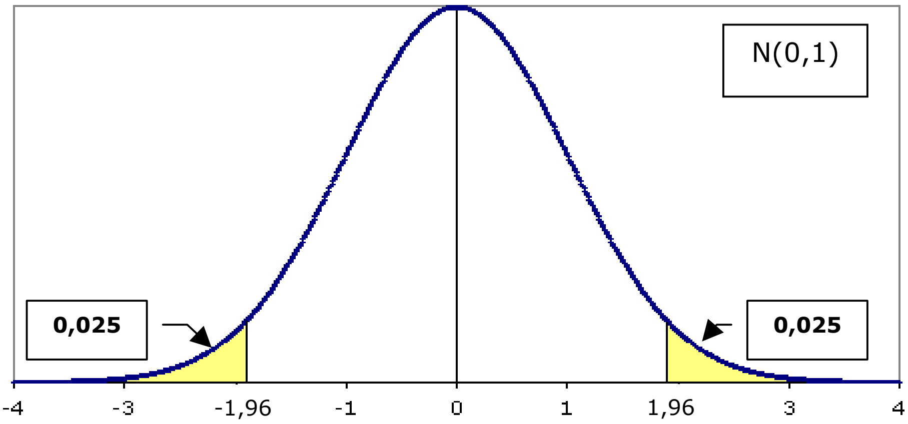
$\alpha = 0,05; \alpha/2 = 0.025$

Statistica test:

$$x_{test} = \frac{\overline{x}_A - \overline{x}_B}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} \sim N(0,1)$$

Valori critici:

$\pm z_{0,025} = \pm 1,96$



Regola di decisione:

$-1,96 \le x_{test} \le +1,96 \quad \Rightarrow$ si accetta H_0

$x_{test} < -1,96; x_{test} > +1,96 \quad \Rightarrow$ si rifiuta H_0

Valore test:

$$x_{test} = \frac{\overline{x}_A - \overline{x}_B}{\sqrt{2\frac{\sigma^2}{n}}} = \frac{83-75}{\sqrt{0,64}} = \frac{8}{0,8} = 10$$

Decisione:

$\boxed{\textbf{Rifiutiamo } H_0}$

$x_{test} < -1{,}96 \quad \Rightarrow$

b)

La differenza tra le due medie può essere stimata tramite l'intervallo:

$$P\left(\left(\overline{x}_A - \overline{x}_B\right) - z_{\alpha/2}\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} \leq \mu_A - \mu_B \leq \left(\overline{x}_A - \overline{x}_B\right) + z_{\alpha/2}\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}} \right) = 1 - \alpha$$

ovvero, essendo $\sigma_A^2 = \sigma_B^2 = \sigma^2$ ed $n_A = n_B = n$:

$$P\left(\left(\overline{x}_A - \overline{x}_B\right) - z_{\alpha/2}\sqrt{\frac{2\sigma^2}{n}} \leq \mu_A - \mu_B \leq \left(\overline{x}_A - \overline{x}_B\right) + z_{\alpha/2}\sqrt{\frac{2\sigma^2}{n}} \right) = 1 - \alpha$$

Il valore campionario della differenza tra le due medie è:

$$\overline{x}_A - \overline{x}_B = 83 - 75 = 8$$

Inoltre:

$\alpha = 0{,}05$
$\alpha/2 = 0.025$
$z_{0{,}025} = 1{,}96$,

quindi, sostituendo:

$$P\left(8 - 1{,}96\sqrt{\frac{64}{100}} \leq \mu_A - \mu_B \leq 8 + 1{,}96\sqrt{\frac{64}{100}} \right) = 1 - \alpha$$

da cui l'intervallo di confidenza al 95%:

$$IC_{0{,}95} = \left[8 - 1{,}568; 8 + 1{,}568\right] = \left[6{,}432; 9{,}568\right]$$

P.S.:
l'intervallo di confidenza può essere indirettamente utilizzato per verificare l'ipotesi di uguaglianza tra le medie. La regola di decisione è la seguente: si accetta H_0 (ipotesi di uguaglianza tra le medie) se l'intervallo di confidenza contiene il valore 0 (in quanto questo significa che una differenza pari a 0 è "plausibile").
Infatti, l'intervallo di confidenza calcolato non contiene lo 0.

ESERCIZIO 7.6

Si estraggano due campioni x ed y di numerosità $n_x = 12$ ed $n_y = 14$, da cui risulta $\overline{x} = 84,2$ e $\overline{y} = 89,3$ con varianze campionarie $s_x = 101,3$ ed $s_y = 81,4$. Supposto che le varianze σ_X^2 e σ_Y^2 siano uguali, si verifichi al livello del 98% se è possibile affermare che due campioni provengano dalla medesima popolazione o meno.

Soluzione

Ipotesi:

$H_0 : \mu_X = \mu_Y$

$H_1 : \mu_X \neq \mu_Y$

Dati:

$\sigma_X^2 = \sigma_Y^2$

$s_x = 101,3$ ed $s_y = 81,4$

$n_X = 12$; $n_Y = 14$

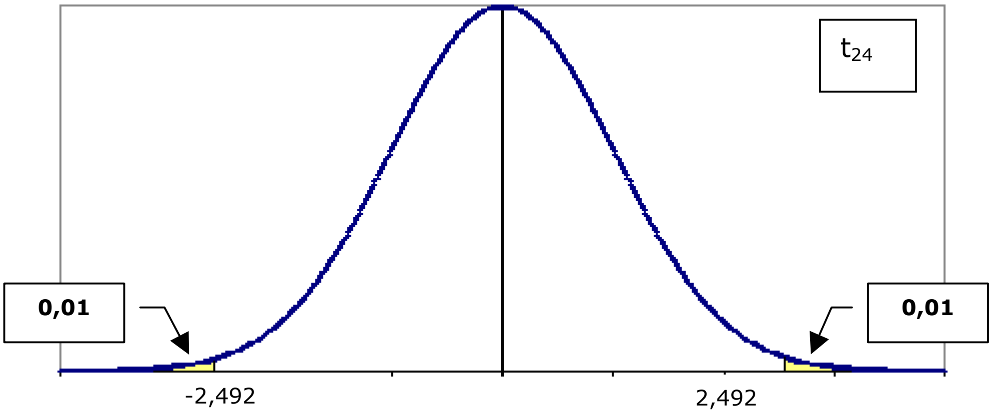
$\alpha = 0,02$; $\alpha/2 = 0.01$

Statistica test:

$$x_{test} = \frac{\overline{x} - \overline{y}}{\sqrt{s_x^2 (n_x - 1) + s_y^2 (n_y - 1)}} \sqrt{\frac{n_x n_y (n_x + n_y - 2)}{n_x + n_y}} \sim t_{n_X + n_Y - 2}$$

Valori critici:

$\pm t_{0,01;24} = \pm 2,492$



Regola di decisione:

$-2,492 \leq x_{test} \leq +2,492 \quad \Rightarrow$ si accetta H_0

$x_{test} < -2,492; x_{test} > +2,492 \quad \Rightarrow$ si rifiuta H_0

Valore test:

$$x_{test} = \frac{84,2-89,3}{\sqrt{101,3(12-1)+81,4(14-1)}}\sqrt{\frac{12\times14\times(12+14-2)}{12+14}} = \frac{-5,1}{46,6}\times12,45 = -1,36$$

Decisione:

$-2,492 \le x_{test} \le +2,492 \quad \Rightarrow$ **Accettiamo H_0**

b)

Anche in questo caso l'ipotesi può essere verificata attraverso l'intervallo di confidenza ottenuto da:

$$P\left(\left(\bar{x}_A-\bar{x}_B\right)-t_{\alpha/2;n_x+n_y-2}\hat{\sigma}\sqrt{\frac{n_x+n_y}{n_xn_y}} \le \mu_A-\mu_B \le \left(\bar{x}_A-\bar{x}_B\right)+t_{\alpha/2;n_x+n_y-2}\hat{\sigma}\sqrt{\frac{n_x+n_y}{n_xn_y}}\right) = 1-\alpha$$

in cui:

$$\hat{\sigma} = \sqrt{\frac{(n_x-1)s_x^2+(n_y-1)s_y^2}{(n_x+n_y-2)}}$$

è la stima non distorta dello scarto quadratico medio σ.
A questo punto, sostituendo le quantità note:

$$\hat{\sigma} = \sqrt{\frac{(12-1)101,3+(14-1)81,4}{(12+14-2)}} = 9,51$$

e

$$P\left(-5,1-2,492\times9,51\sqrt{\frac{12+14}{12\cdot14}} \le \mu_A-\mu_B \le \left(\bar{x}_A-\bar{x}_B\right)+2,492\times9,51\sqrt{\frac{12+14}{12\cdot14}}\right) = 0,98$$

da cui l'intervallo di confidenza al 98%:

$$IC_{0,98} = \left[-5,1-9,31;-5,1+9,31\right] = \left[-14,41;4,21\right]$$

$IC_{0,98}$ contiene il valore 0 $\Rightarrow$ **Accettiamo H_0**

ESERCIZIO 7.7

In due città è stata rilevata la spesa in generi alimentari impiegando due campioni di numerosità $n_1 = 15$ ed $n_2 = 20$, da cui risultano le varianze $s_1^2 = 186$ ed $s_2^2 = 81$. Si verifichi al livello di significatività del 95% se le varianze σ_1^2 e σ_2^2 delle spese delle due città possono essere considerate uguali.

Soluzione

Ipotesi:

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 > \sigma_2^2$

Dati:

$s_1^2 = 186$, $s_2^2 = 81$

$n_1 = 15$, $n_2 = 20$

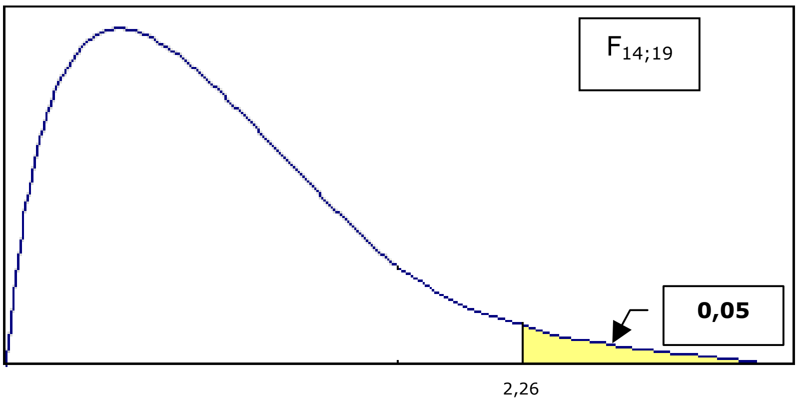
$\alpha = 0,05$

Statistica test:

$$X_{test} = \frac{\hat{s}_1^2}{\hat{s}_2^2} \sim F_{n_1-1;n_2-1}$$

Valore critico:

$F_{0,05;14;19} = 2,26$



Regola di decisione:

$x_{test} \leq 2,26 \quad \Rightarrow$ si accetta H_0

$x_{test} > 2,26 \quad \Rightarrow$ si rifiuta H_0

Valore test:

$$x_{test} = \frac{s_1^2}{s_2^2} = \frac{186}{81} = 2,30$$

Decisione:

$x_{test} > 2,26 \quad \Rightarrow$ **Rifiutiamo H_0**

ESERCIZIO 7.8

La produzione di pneumatici in una fabbrica avviene in tre turni. Il controllo della conformità dei pneumatici prodotti si basa su un campione di 400 pezzi, che ha rivelato quanto segue:

frequenze osservate n_{ij}		**TURNO DI PRODUZIONE**			
		Giorno	*Sera*	*Notte*	totale
ESITO	*Conformità*	194	108	66	**368**
	Non conformità	6	12	14	**32**
	totale	**200**	**120**	**80**	**400**

Verificare l'ipotesi che l'esito non dipenda dal turno di produzione, al livello di significatività del 95%.

Soluzione

Ipotesi:

$$H_0 : n_{ij} = \frac{n_{i.} \times n_{.j}}{N} \quad \text{(indipendenza)}$$

$$H_1 : n_{ij} \neq \frac{n_{i.} \times n_{.j}}{N}$$

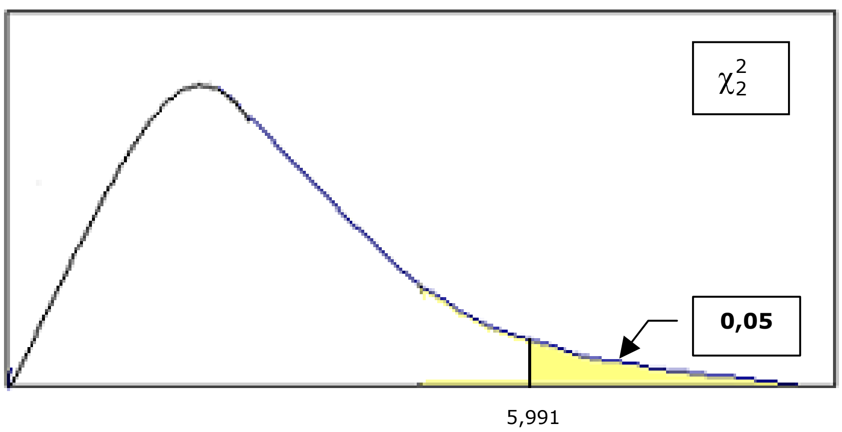
Statistica test:

$$X_{test} = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\left(n_{ij} - n_{ij}^*\right)^2}{n_{ij}^*} \sim \chi^2_{(r-1)\bullet(c-1)}$$

in cui, appunto, $n_{ij}^* = \frac{n_{i.} \times n_{.j}}{N}$.

Valore critico:

$\chi^2_{0,05;2}$ = 5,991



Regola di decisione:

$$x_{test} \le 5{,}991 \quad \Rightarrow \quad \text{si accetta } H_0$$
$$x_{test} > 5{,}991 \quad \Rightarrow \quad \text{si rifiuta } H_0$$

Valore test:

Il calcolo del valore test presuppone la conoscenza delle frequenze teoriche n^*_{ij}, che procediamo a determinare:

$$n^*_{11} = \frac{n_{1.} \times n_{.1}}{N} = \frac{184 \times 100}{200} = 92\,;$$
$$n^*_{12} = \frac{n_{1.} \times n_{.2}}{N} = \frac{184 \times 60}{200} = 55{,}2\,;$$

...

e così via per tutte le altre n_{ij}:

frequenze teoriche n^*_{ij}

		TURNO DI PRODUZIONE			
		Giorno	*Sera*	*Notte*	totale
ESITO	*Conformità*	184	110,4	73,6	**368**
	Non conformità	16	9,6	6,4	**32**
	totale	**200**	**120**	**80**	**400**

$$x_{test} = \sum_{i=1}^{2} \sum_{j=1}^{3} \frac{\left(n_{ij} - n^*_{ij}\right)^2}{n^*_{ij}} = \frac{(194 - 184)^2}{184} + \frac{(108 - 110{,}4)^2}{110{,}4} + \ldots + \frac{(14 - 6{,}4)^2}{6{,}4} = 17{,}255$$

Decisione:

$x_{test} > 5{,}991 \quad \Rightarrow$ **Rifiutiamo H_0**

ESERCIZIO 7.9

Si verifichi l'indipendenza tra rendimento scolastico e numero di figli della famiglia di appartenenza tra gli alunni di una scuola elementare attraverso un campione di 228 alunni, da cui risulta la seguente distribuzione congiunta:

frequenze osservate n_{ij}		**RENDIMENTO SCOLASTICO**			
		scarso	*sufficiente*	*buono*	totale
NUMERO DI FIGLI	*1*	7	13	17	**37**
	2	19	30	22	**71**
	3	13	11	13	**37**
	4	10	9	12	**31**
	5	10	9	12	**31**
	più di 5	8	6	7	**21**
	totale	**67**	**78**	**83**	**228**

Si consideri un livello di significatività del 99%.

<u>Soluzione</u>

Ipotesi:

$$H_0 : n_{ij} = \frac{n_{i.} \times n_{.j}}{N}$$

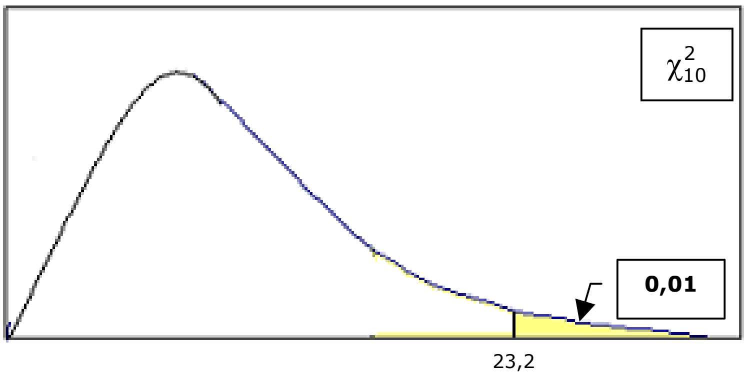
$$H_1 : n_{ij} \neq \frac{n_{i.} \times n_{.j}}{N}$$

Statistica test:

$$X_{test} = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\left(n_{ij} - n_{ij}^*\right)^2}{n_{ij}^*} \sim \chi^2_{(r-1)\cdot(c-1)}$$

Valore critico:

$\chi^2_{0,01;10} = 23,2$



Regola di decisione:

$x_{test} \leq 23,2 \quad \Rightarrow$ si accetta H_0

$x_{test} > 23,2 \quad \Rightarrow$ si rifiuta H_0

Valore test:

Le frequenze teoriche n^*_{ij} sono:

$$n^*_{11} = \frac{n_{1.} \times n_{.1}}{N} = \frac{37 \times 67}{228} = 10,87\,;$$

$$n^*_{12} = \frac{n_{1.} \times n_{.2}}{N} = \frac{37 \times 78}{228} = 12,66\,;$$

…

e così via per tutte le altre n_{ij}:

frequenze teoriche n^*_{ij}

		RENDIMENTO SCOLASTICO			
		scarso	*sufficiente*	*buono*	totale
NUMERO DI FIGLI	*1*	10,87	12,66	13,47	**37**
	2	20,86	24,29	25,85	**71**
	3	10,87	12,66	13,47	**37**
	4	9,11	10,61	11,29	**31**
	5	9,11	10,61	11,29	**31**
	più di 5	6,17	7,18	7,64	**21**
	totale	**67**	**78**	**83**	**228**

$$x_{test} = \sum_{i=1}^{2}\sum_{j=1}^{3}\frac{\left(n_{ij} - n^*_{ij}\right)^2}{n^*_{ij}} = \frac{(7-10,87)^2}{10,87} + \frac{(13-12,66)^2}{12,66} + \ldots + \frac{(7-7,64)^2}{7,64} = 6,588$$

Decisione:

$x_{test} < 23,2 \quad \Rightarrow$ **Accettiamo H_0**