Università degli Studi di Cassino

Esercitazione di Statistica 2 dell'1.03.2007

Simona Balzano

Esercizio 1

La quantità di merci in transito negli aeroporti italiani si distribuisce normalmente con una media pari a 18,7 (migliaia di tonnellate) e uno scarto quadratico medio pari a 63. In un campione di 20 aeroporti viene registrato un valore medio pari a 15. Utilizzando un livello di significatività del 99%:

a) Verificare l'ipotesi che il transito medio di merci sia rimasto invariato;
b) Verificare l'ipotesi che il transito medio di merci non sia diminuito

Soluzione

$\mu = 18{,}7$
$\bar{x} = 15$
$\sigma = 63$
$n = 20$
$\alpha = 0{,}01$

a) test sulla media, bilaterale;
distribuzione normale, varianza nota

▪ Ipotesi	$H_0: \mu = 18{,}7$ $H_1: \mu \neq 18{,}7$
▪ Statistica test	$X_{test} = \dfrac{\bar{x} - \mu}{\sigma / \sqrt{n}}$
▪ Valori critici	$\pm z_{\alpha/2} = \pm 2{,}58$
▪ Regola di decisione	$- 2{,}58 \leq x_{test} \leq 2{,}58$ $\Rightarrow$ si accetta H_0 $x_{test} \leq - 2{,}58$ oppure $x_{test} \geq 2{,}58$ $\Rightarrow$ si rifiuta H_0
▪ Valore test	$x_{test} = \dfrac{15 - 18{,}7}{63/\sqrt{20}} = -0{,}26$
▪ Decisione	$- 2{,}58 \leq -0{,}26 \leq 2{,}58$ $\Rightarrow$ **si accetta H_0**

b) test sulla media, unilaterale;
distribuzione normale, varianza nota

▪ Ipotesi	H_0: μ = 18,7 H_1: μ < 18,7
▪ Statistica test	$X_{test} = \dfrac{\overline{x} - \mu}{\sigma/\sqrt{n}}$
▪ Valore critico	$-z_\alpha = -2,33$
▪ Regola di decisione	$x_{test} \geq -2,33$ $\Rightarrow$ si accetta H_0 $x_{test} \leq -2,33$ $\Rightarrow$ si rifiuta H_0
▪ Valore test	$x_{test} = \dfrac{15 - 18,7}{63/\sqrt{20}} = -0,26$
▪ Decisione	- 0,26 ≥ - 2,33 $\Rightarrow$ **si accetta H_0**

Esercizio 2

Un campione di 20 alberghi italiani presenta un numero di letti pari in media a 70, con scarto quadratico medio corretto pari a 25.
Sapendo che gli alberghi italiani contano in media 61 letti:

a) verificare l'ipotesi nulla che la media si possa considerare invariata, al livello di significatività del 99%.
b) Verificare la stessa ipotesi rispetto ad un campione di ampiezza 100.

Soluzione

$\mu = 61$
$\bar{x} = 70$
$s = 25$
$\alpha = 0,01$

a) test sulla media, bilaterale;
distribuzione non nota, varianza non nota, campione piccolo

$n = 20$

▪ Ipotesi	$H_0: \mu = 61$ $H_1: \mu \neq 61$
▪ Statistica test	$X_{test} = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$
▪ Valori critici	$\pm t_{\alpha/2;n-1} = \pm t_{0,005;19} = \pm 2,861$
▪ Regola di decisione	$-2,86 \leq x_{test} \leq 2,86$ $\Rightarrow$ si accetta H_0 $x_{test} \leq -2,86$ oppure $x_{test} \geq 2,86$ $\Rightarrow$ si rifiuta H_0
▪ Valore test	$x_{test} = \dfrac{70 - 61}{25/\sqrt{20}} = 1,97$
▪ Decisione	$-2,86 \leq 1,97 \leq 2,86$ **$\Rightarrow$ si accetta H_0**

b) <u>test sulla media, bilaterale;</u>
<u>distribuzione non nota, varianza non nota, campione grande</u>

$n = 100$

▪ <u>Ipotesi</u>	H_0: $\mu = 61$ H_1: $\mu \neq 61$
▪ <u>Statistica test</u>	$X_{test} = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}}$
▪ <u>Valore critico</u>	$\pm z_{\alpha} = \pm 2,58$
▪ <u>Regola di decisione</u>	$-2,58 \leq x_{test} \leq 2,58$ ⇒ si accetta H_0 $x_{test} < -2,58$ oppure $x_{test} > 2,58$ ⇒ si rifiuta H_0
▪ <u>Valore test</u>	$x_{test} = \dfrac{70 - 61}{25/\sqrt{100}} = 4,4$
▪ <u>Decisione</u>	$4,4 > -2,58$ **⇒ si rifiuta H_0**

Esercizio 3

La spesa in ricerca e sviluppo in due campioni di 20 università nord e 20 del sud Italia risulta pari rispettivamente a 271 (migliaia di euro) con s.q.m. corretto di 200 e 176 con s.q.m. di 150.

a) verificare l'ipotesi nulla che non vi sia in media differenza di spesa tra nord e sud, al livello di significatività del 95%;
b) verificare se le due varianze possono essere considerate uguali;
c) verificare l'uguaglianza tra le medie nell'ipotesi che le due varianze siano note e pari rispettivamente a 4000 e 3600.

Soluzione

$\bar{x}_1 = 271 \qquad \bar{x}_2 = 176$

$s_1 = 200 \qquad s_2 = 150$

$n_1 = n_2 = 20$

$\alpha = 0{,}05$

a) <u>test sulla differenza tra 2 medie, bilaterale</u>;
<u>distribuzione non nota, varianza non nota, campioni piccoli</u>

▪ <u>Ipotesi</u>	H_0: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$
▪ <u>Statistica test</u>	$X_{test} = \dfrac{\bar{x}_1 - \bar{x}_2}{\tilde{s}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}}$ in cui $\tilde{s}^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$
▪ <u>Valore critico</u>	$\pm t_{\alpha/2; n_1+n_2-2} = \pm t_{0,025;38} = \pm 2{,}02$
▪ <u>Regola di decisione</u>	$-2{,}02 \leq x_{test} \leq 2{,}02 \Rightarrow$ si accetta H_0 $x_{test} < -2{,}02$ oppure $x_{test} > 2{,}02 \Rightarrow$ si rifiuta H_0
▪ <u>Valore test</u>	$\tilde{s} = \sqrt{\dfrac{(20-1)\cdot 40.000 + (20-1)\cdot 22.500}{38}} = 176{,}78$ $X_{test} = \dfrac{271 - 176}{176{,}78 \cdot \sqrt{2 \cdot \dfrac{1}{20}}} = 1{,}7$
▪ <u>Decisione</u>	$-2{,}02 < 1{,}7 < 2{,}02$ $\Rightarrow$ **si accetta H_0**

b) <u>test sulla differenza tra 2 varianze</u>

▪ <u>Ipotesi</u>	$H_0\colon\ \sigma_1^2 = \sigma_2^2$ $H_1\colon\ \sigma_1^2 \neq \sigma_2^2$
▪ <u>Statistica test</u>	$X_{test} = \dfrac{s_1^2}{s_2^2}$
▪ <u>Valore critico</u>	$F_{\alpha;n_1-1;n_2-1} = F_{0,05;19;19} = 2,16$
▪ <u>Regola di decisione</u>	$x_{test} \leq 2,16$ $\Rightarrow$ si accetta H_0 $x_{test} < 2,16$ $\Rightarrow$ si rifiuta H_0
▪ <u>Valore test</u>	$x_{test} = \dfrac{200}{150} = 1,33$
▪ <u>Decisione</u>	$1,33 < 2,16$ **$\Rightarrow$ si accetta H_0**

c) <u>test sulla differenza tra 2 medie, bilaterale;</u>
<u>distribuzione non nota, varianze note, campioni piccoli</u>

$\sigma_1^2 = 4000 \quad \sigma_2^2 = 3600$

▪ <u>Ipotesi</u>	$H_0\colon \mu_1 - \mu_2 = 0$ $H_1\colon \mu_1 - \mu_2 \neq 0$
▪ <u>Statistica test</u>	$X_{test} = \dfrac{\overline{x}_1 - \overline{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$
▪ <u>Valori critici</u>	$\pm z_{\alpha/2} = \pm 1,96$
▪ <u>Regola di decisione</u>	$-1,96 \leq x_{test} \leq 1,96$ $\Rightarrow$ si accetta H_0 $x_{test} < -1,96$ oppure $x_{test} > 1,96$ $\Rightarrow$ si rifiuta H_0
▪ <u>Valore test</u>	$X_{test} = \dfrac{271 - 176}{\sqrt{\dfrac{4000}{20} + \dfrac{3600}{20}}} = 4,87$
▪ <u>Decisione</u>	$4,87 > 1,96$ **$\Rightarrow$ si rifiuta H_0**

Esercizio 4

Di un campione di 160 soggetti il 42% è stato colpito dall'influenza afro-cubana.

a) Al livello di significatività del 95% si può ritenere affidabile il dato ufficiale secondo cui l'incidenza è pari al 35%?
b) Qual è il livello di significatività osservato (p-value) in corrispondenza della proporzione campionaria del 42%?

Soluzione

$\pi = 0,35$
$p = 0,42$
$n = 160$

a) test sulla proporzione, unilaterale

$\alpha = 0,95$

▪ Ipotesi	$H_0: \pi = 0,35$ $H_1: \pi > 0,35$
▪ Statistica test	$X_{test} = \dfrac{p-\pi}{\sqrt{\dfrac{p\cdot(1-p)}{n}}}$
▪ Valore critico	$z_\alpha = 1,645$
▪ Regola di decisione	$x_{test} \leq 1,645 \Rightarrow$ si accetta H_0 $x_{test} > 1,645 \Rightarrow$ si rifiuta H_0
▪ Valore test	$X_{test} = \dfrac{0,42-0,35}{\sqrt{\dfrac{0,42\cdot 0,58}{160}}} = 1,79$
▪ Decisione	$1,79 > 1,96$ $\Rightarrow$ **si rifiuta H_0**

b)

Bisogna determinare il p-value associato al valore campionario della statistica test, quindi al valore osservato della proporzione.

▪ <u>Ipotesi</u>	H_0: $\pi = 0{,}35$ H_1: $\pi \neq 0{,}35$
▪ <u>Statistica test</u>	$X_{test} = \dfrac{p - \pi}{\sqrt{\dfrac{p \cdot (1-p)}{n}}}$
▪ <u>Valore test</u>	$X_{test} = \dfrac{0{,}42 - 0{,}35}{\sqrt{\dfrac{0{,}42 \cdot 0{,}58}{160}}} = 1{,}79$
▪ <u>P-value</u>	p-value = $P(Z > x_{test}) = 1 - F(1{,}79) = 1 - 0{,}9633 = 0{,}0367$
▪ <u>Decisione</u>	Il p-value è basso: l'ipotesi nulla è poco coerente con in dati (= il campione non conferma H_0) $\Rightarrow$ **si rifiuta H_0**

Esercizio 5

Il tasso migratorio registrato in un campione di 40 comuni del nord risulta 2,5% contro un valore di 2,3% registrato in un secondo campione di 50 comuni del resto d'Italia.
Verificare se al livello di significatività del 99% il tasso del nord si può considerare maggiore che nel resto del paese.

Soluzione

<u>test sulla differenza tra 2 proporzioni, bilaterale</u>

p_1 = 2,5% = 0,025
n_1 = 40
p_2 = 2,3% = 0,023
n_2 = 50

α = 0,01

▪ <u>Ipotesi</u>	$H_0: \pi_1 - \pi_2 = 0$ $H_1: \pi_1 - \pi_2 \neq 0$
▪ <u>Statistica test</u>	$X_{test} = \dfrac{p_1 - p_2}{\sqrt{\dfrac{p_1 \cdot (1-p_1)}{n_1} + \dfrac{p_2 \cdot (1-p_2)}{n_2}}} \sim N(0,1)$
▪ <u>Valore critico</u>	$\pm z_{\alpha/2} = \pm 2{,}58$
▪ <u>Regola di decisione</u>	$-2{,}58 \leq x_{test} \leq 2{,}58$ $\Rightarrow$ si accetta H_0 $x_{test} < -2{,}58$ oppure $x_{test} > 2{,}58$ $\Rightarrow$ si rifiuta H_0
▪ <u>Valore test</u>	$X_{test} = \dfrac{0{,}025 - 0{,}023}{\sqrt{\dfrac{0{,}025 \cdot (0{,}975)}{40} + \dfrac{0{,}023 \cdot (0{,}977)}{50}}} = 0{,}061$
▪ <u>Decisione</u>	$-2{,}58 \leq 0{,}061 \leq 2{,}58$ **$\Rightarrow$ si accetta H_0**

Esercizio 6

Un fabbricante di batterie afferma che la capacità media delle batterie che produce è di 140 ampere.
Per verificare la veridicità dell'affermazione si misura la capacità di un campione di 20 batterie, da cui risultano che 14 batterie hanno una capacità superiore a 140.
Utilizzando un livello di significatività del 95% a quale conclusione si può giungere?

Soluzione

Test dei segni
(1 campione)

Si ipotizza la simmetria (μ = Me) della v.c. capacità.

μ = Me = 140
k = numero di osservazioni superiori alla media = 14
n = 20 α = 0,05

▪ Ipotesi	H_0: Me = 140 H_1: Me < 140
▪ Statistica test	$X_{test} = \frac{2k-n+1}{\sqrt{n}}$
▪ Valore critico	$-z_{\alpha/2} = -1,96$
▪ Regola di decisione	$x_{test} \geq -2,58$ $\Rightarrow$ si accetta H_0 $x_{test} < -2,58$ $\Rightarrow$ si rifiuta H_0
▪ Valore test	$X_{test} = \frac{2 \cdot 14 - 20 + 1}{\sqrt{20}} = 2,01$
▪ Decisione	2,01 > - 2,58 $\Rightarrow$ **si accetta H_0**

Conclusione: al 95% il produttore afferma la verità, ossia il livello medio della capacità delle batterie che produce è di 140 ampere.

In modo equivalente l'ipotesi può essere verificata come segue:

▪ Statistica test	X_{test} = k
▪ Valore critico	$\frac{n+1}{2} - z_{\alpha/2}\frac{\sqrt{n}}{2} = \frac{20+1}{2} - 1,96\frac{\sqrt{20}}{2} = 10,5 - 4,38 = 6,12$
▪ Regola di decisione	$x_{test} \geq 6,12$ $\Rightarrow$ si accetta H_0 $x_{test} < 6,12$ $\Rightarrow$ si rifiuta H_0
▪ Valore test	x_{test} = 14
▪ Decisione	14 > 4,73,12 $\Rightarrow$ **si accetta H_0**

Esercizio 7

Per verificare l'efficacia di un trattamento chimico, vengono confrontati gli indici di crescita dei raccolti osservati settimanalmente per 15 settimane su due campi di cui uno sottoposto al trattamento e l'altro no. I risultati sono in tabella:

Trattato (x)	1	3	4	6	7	6	9	9	8	7	9	8	7	9	9
Non trattato (y)	4	4	3	4	6	7	4	4	7	8	9	4	6	7	7

Verificare l'ipotesi che il trattamento sia inutile usando il test dei ranghi con un livello di significatività del 95%.

Soluzione

$n = 15$ $\qquad\qquad$ $D_i = x_i - y_i$

Test dei ranghi (2 campioni, $n_1 = n_2$)

Se il trattamento ha effetto le differenze $D_i = x_i - y_i$ avranno tutte segno positivo, quindi una mediana positiva; se, viceversa, il trattamento è inutile differenze positive e negative si presenteranno con la stessa frequenza, quindi la loro mediana sarà prossima allo 0.
Sulla tabella che segue sono riportati i valori delle differenze Di, di cui in rosso sono quelle positive (da sommare nel calcolo di W)

Trattato	1	3	4	6	7	6	9	9	8	7	9	8	7	9	9
Non trattato	4	4	3	4	6	7	4	4	7	8	9	4	6	7	7
D_i	-3	-1	1	2	1	-1	5	5	1	-1	0	4	1	2	2

▪ Ipotesi	H_0: $Me(D_i) = 0$ H_1: $Me(D_i) \neq 0$
▪ Statistica test	1. $W = \sum_{i=1}^{n} R(\lvert D_i\rvert) \cdot I(D_i > 0)$ 2. $X_{test} = \dfrac{W - n(n+1)/4}{n(n+1)(2n+1)/24} \sim N(0,1)$
▪ Valori critici	$\pm z_{\alpha/2} = \pm 1,96$
▪ Regola di decisione	$-1,96 \leq x_{test} \leq 1,96$ $\Rightarrow$ si accetta H_0 $x_{test} < -1,96$ oppure $x_{test} > 1,96$ $\Rightarrow$ si rifiuta H_0
▪ Valore test	1. $W = \sum_{i=1}^{n} R(\lvert D_i\rvert) \cdot I(D_i > 0) = 1+2+1+5+5+1+4+1+2+2 = 24$ 2. $x_{test} = \dfrac{24 - 15(15+1)/4}{15(15+1)(30+1)/24} = \dfrac{-36}{310} = -0,116$
▪ Decisione	$-1,96 \leq -0,116 \leq 1,96$ $\Rightarrow$ **si accetta H_0**

Esercizio 8

Al fine di confrontare i rendimenti di due borse un analista estrae 2 campioni di 8 titoli quotati in una e 10 quotati nell'altra. I rendimenti osservati sono:

Borsa 1 X	**Borsa 2 Y**
5,1	2,8
1,4	7,3
1,6	9,8
5,7	7,0
9,7	9,5
9,1	5,5
11,2	5,6
8,2	10,8
	4,7
	6,5

Si verifichi se i rendimenti medi delle due borse possono essere considerati uguali, al livello di significatività del 99%, utilizzando il test delle sequenze.

Soluzione

Test delle sequenze
(2 campioni, $n_1 \neq n_2$)

Il test si basa sulle differenze tra tutte le coppie di osservazioni $x_i - y_j$, con i = 1, …, n_1 e j = 1, …, n_2, per poi dedurre il valore D_{ij} come segue:

$$D_{ij} = \begin{cases} 1 & \text{se } x_i < y_j \\ 0 & \text{se } x_i \geq y_j \end{cases}$$

Le differenze $x_i - y_j$ sono raccolte nella seguente tabella:

X \ Y	2,8	7,3	9,8	7	9,5	5,5	5,6	10,8	4,7	6,5
5,1	2,3	-2,2	-4,7	-1,9	-4,4	-0,4	-0,5	-5,7	0,4	-1,4
1,4	-1,4	-5,9	-8,4	-5,6	-8,1	-4,1	-4,2	-9,4	-3,3	-5,1
1,6	-1,2	-5,7	-8,2	-5,4	-7,9	-3,9	-4	-9,2	-3,1	-4,9
5,7	2,9	-1,6	-4,1	-1,3	-3,8	0,2	0,1	-5,1	1	-0,8
9,7	6,9	2,4	-0,1	2,7	0,2	4,2	4,1	-1,1	5	3,2
9,1	6,3	1,8	-0,7	2,1	-0,4	3,6	3,5	-1,7	4,4	2,6
11,2	8,4	3,9	1,4	4,2	1,7	5,7	5,6	0,4	6,5	4,7
8,2	5,4	0,9	-1,6	1,2	-1,3	2,7	2,6	-2,6	3,5	1,7

mentre la seguente contiene i valori D_{ij}:

X \ Y	2,8	7,3	9,8	7	9,5	5,5	5,6	10,8	4,7	6,5
5,1	1	0	0	0	0	0	0	0	1	0
1,4	0	1	0	0	0	0	0	0	0	0
1,6	0	1	0	0	0	0	0	0	0	0
5,7	1	1	0	0	0	1	1	0	1	0
9,7	1	0	0	1	1	1	1	0	1	1
9,1	1	0	0	1	0	1	1	0	1	1
11,2	1	0	1	1	1	1	1	1	1	1
8,2	1	0	0	1	0	1	1	0	1	1

Totale tabella = 37

▪ Ipotesi	H_0: $Me_1 = Me_2$ H_1: $Me_1 \neq Me_2$
▪ Statistica test	1. $U = \sum_{i=1}^{n} \sum_{j=1}^{m} D_{ij}$ 2. $X_{test} = \dfrac{U - mn/2}{mn(m+n+1)/12} \sim N(0,1)$
▪ Valori critici	$\pm z_{\alpha/2} = \pm 2{,}58$
▪ Regola di decisione	$-2{,}58 \leq x_{test} \leq 2{,}58$ $\Rightarrow$ si accetta H_0 $x_{test} < -2{,}58$ oppure $x_{test} > 2{,}58$ $\Rightarrow$ si rifiuta H_0
▪ Valore test	1. $U = \sum_{i=1}^{n} \sum_{j=1}^{m} D_{ij} = 37$ 2. $x_{test} = \dfrac{37 - 8 \cdot 10/2}{8 \cdot 10 \cdot (8+10+1)/12} = \dfrac{-3}{126{,}67} = -0{,}024$
▪ Decisione	$-2{,}58 \leq -0{,}024 \leq 2{,}58$ $\Rightarrow$ **si accetta H_0**

Esercizio 9

La seguente tabella contiene la distribuzione doppia del numero di imprese che hanno svolto attività di R&S, suddivise per settore di attività economica e tipo di programma di ricerca cui hanno partecipato.

Programma / ATT. EC.	UE	Internazionali	Nazionali	Altri	**Totale**
Industria	26	4	1	44	**75**
Costruzioni	0	0	0	50	**50**
Servizi	35	5	5	95	**140**
Totale	**61**	**9**	**6**	**189**	**265**

Verificare se esiste indipendenza al livello di significatività del 99%.

Soluzione

La verifica si basa sul calcolo dell'indice di indipendenza assoluta χ^2 attraverso la formula:

$$\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(n_{ij}-\hat{n}_{ij}\right)^2}{\hat{n}_{ij}}$$

in cui $\hat{n}_{ij} = \dfrac{n_{i.} \times n_{.j}}{n}$

La seguente tabella contiene le frequenze teoriche $\hat{n}_{ij}$:

$\hat{n}_{ij}$	UE	Internazionali	Nazionali	Altri	**Totale**
Industria	17,26	2,55	1,70	53,49	**75**
Costruzioni	11,51	1,70	1,13	35,66	**50**
Servizi	32,23	4,75	3,17	99,85	**140**
Totale	**61**	**9**	**6**	**189**	**265**

▪ <u>Ipotesi</u>	H_0: $\chi^2 = 0$ H_1: $\chi^2 > 0$
▪ <u>Statistica test</u>	$X_{test} = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(n_{ij}-\hat{n}_{ij}\right)^2}{\hat{n}_{ij}}$
▪ <u>Valore critico</u>	$\chi^2_{\alpha;(r-1)(c-1)} = \chi^2_{0,01;6} = 16,812$
▪ <u>Regola di decisione</u>	$x_{test} \leq 16,872$ $\Rightarrow$ si accetta H_0 $x_{test} > 16,872$ $\Rightarrow$ si rifiuta H_0
▪ <u>Valore test</u>	$x_{test} = 28,87$
▪ <u>Decisione</u>	28,87 > 16,872 $\Rightarrow$ **si rifiuta H_0**

Esercizio 10

Per verificare se 5 monete siano o meno truccate, vengono lanciate 1000 volte. I risultati dei 1000 lanci sono riassunti dalla seguente distribuzione di frequenze:

Numero di "testa"	Frequenza
0	38
1	144
2	342
3	287
4	164
5	25
	1000

Si può affermare che le monete sono regolari, al livello di significatività del 95%?

Soluzione

Per rispondere alla domanda bisogna effettuare un test sulla distribuzione della variabile causale:

X = "numero di teste in 1 lancio di 5 monete",

che, se le monete sono regolari, è una variabile casuale Binomiale con parametri n = 5 (le monete) e p = 0,5 (nel caso di monete regolari).
Il test si basa sul confronto tra le frequenze osservate n_i e quelle teoriche $\hat{n}_i$, che vanno determinate a partire dalle probabilità dei singoli valori che la v.c. può assumere (ossia: 0, 1, 2, 3, 4, 5).
In altre parole, la probabilità di x teste in 1 lancio di 5 monete sarà data da:

$$P(x) = \binom{5}{x} \cdot 0,5^x \cdot 0,5^{5-x} = \binom{5}{x} \cdot 0,5^5$$

La variabile casuale è definita come segue:

x_i	$P(x_i)$
0	$P(0) = \binom{5}{0} 0,5^5 = 0,03125$
1	$P(1) = \binom{5}{1} 0,5^5 = 0,15625$
2	$P(2) = \binom{5}{2} 0,5^5 = 0,3125$
3	$P(3) = \binom{5}{3} 0,5^5 = 0,3125$
4	$P(4) = \binom{5}{4} 0,5^5 = 0,15625$
5	$P(5) = \binom{5}{5} 0,5^5 = 0,03125$
	1,00

Moltiplicando le probabilità per il numero di lanci effettuati (1000) si ottengono le frequenze teoriche:

$$\hat{n}_i = 1000 \times P(x_i)$$

che sono raccolte nella seguente tabella, accanto a quelle osservate:

X = Numero di "testa"	Frequenza osservata n_i	Frequenza teorica $\hat{n}_i$
0	38	31,25
1	144	156,25
2	342	312,5
3	287	312,5
4	164	156,25
5	25	31,25
	1000	**1000**

Il test si basa sul fatto che se le monete sono regolari allora le frequenze osservate non devono essere molto diverse da quelle teoriche, quindi il χ^2 non deve essere significativamente diverso da 0:

▪ Ipotesi	H_0: $\chi^2 = 0$ (= le monete sono regolari) H_1: $\chi^2 > 0$
▪ Statistica test	$X_{test} = \sum_{i=1}^{k} \frac{(n_i - \hat{n}_i)^2}{\hat{n}_i}$
▪ Valore critico	$\chi^2_{\alpha;k-1} = \chi^2_{0,05;5} = 11,07$
▪ Regola di decisione	$x_{test} \leq 16,872$ $\Rightarrow$ si accetta H_0 $x_{test} > 16,872$ $\Rightarrow$ si rifiuta H_0
▪ Valore test	$x_{test} = \frac{(38-31,25)^2}{31,25} + \cdots + \frac{(25-31,25)^2}{31,25} = 8,9184$
▪ Decisione	8,9184 < 11,07 $\Rightarrow$ **si accetta H_0**