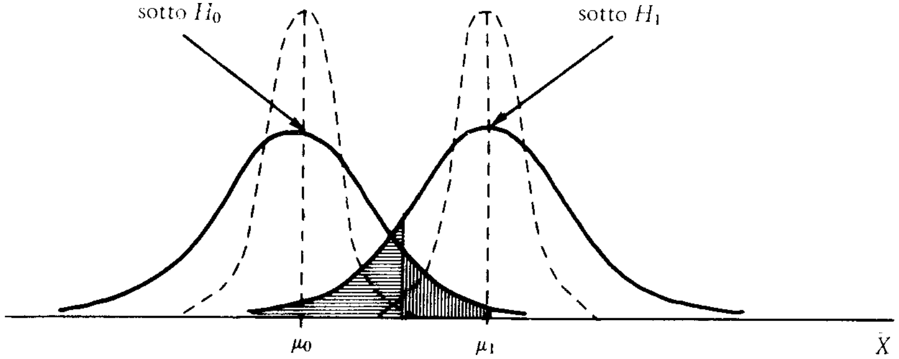


Modificazione della distribuzione campionaria del test all'aumentare della numerosità del campione.

TEST SULLA MEDIA

Si verifica che la media assuma un determinato valore

$\Rightarrow \mu_x = \mu_0$

$$x \approx N\left(\mu_x, \sigma_x^2\right)$$

$$\sigma_x \begin{cases} \text{nota} : \sigma_x = \sigma \rightarrow test\, z \\ \text{non nota} \rightarrow \begin{cases} \text{piccoli campioni} \rightarrow test\, t \\ \text{grandi campioni} \rightarrow test\, z \end{cases} \end{cases}$$

$$\boxed{z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}} \qquad \boxed{t = \frac{\overline{X} - \mu_0}{S_x / \sqrt{n}}}$$

dove : $$S_x^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2$$

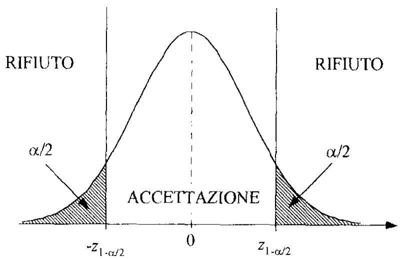
TEST SULLA MEDIA

Uso:

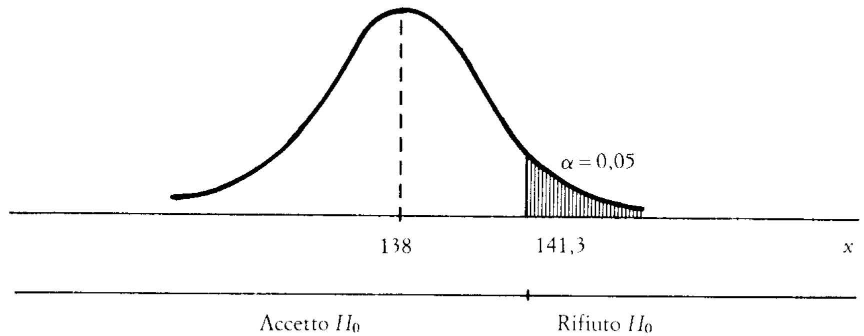
- confronto tra media campionaria e media supposta
- test sulla validità di un parametro di un modello calibrato

$$\left.\begin{aligned} H_0 &: \mu_x = \mu_0 \\ H_1 &: \mu_x \neq \mu_0 \end{aligned}\right\} \text{bilaterale; rifiuto } H_0 \text{ se } |t| > t_{\alpha/2, n-1}$$

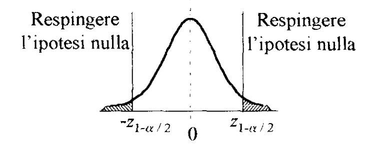
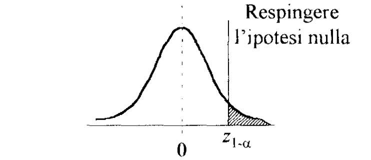
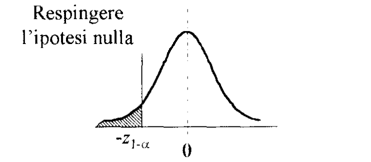
Regione di accettazione e regione di rifiuto



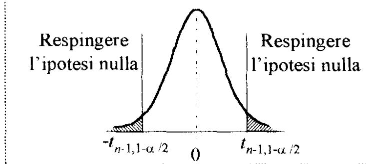
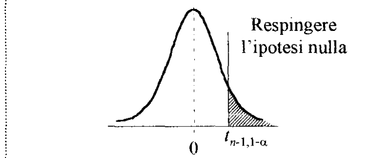
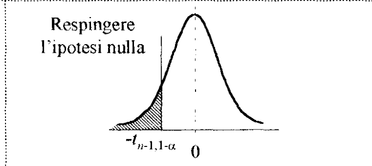
$$\left.\begin{aligned} H_0 &: \mu_x > \mu_0 \\ H_1 &: \mu_x \leq \mu_0 \end{aligned}\right\} \text{unilaterale; rifiuto } H_0 \text{ se } t < t_{\alpha, n-1}$$



Rappresentazione grafica della regione di rifiuto per alfa=0,05

Statistica $Z = (\bar{X}_n - \mu_0)\sqrt{n}/\sigma$			
Ipotesi nulla	Ipotesi alternativa	Regione di rifiuto di H_0	Schema grafico
$H_0 : \mu = \mu_0$	$H_A : \mu \neq \mu_0$	$Z > z_{1-\alpha/2}$ oppure $Z < -z_{1-\alpha/2}$	
$H_0 : \mu = \mu_0$	$H_A : \mu > \mu_0$	$Z > z_{1-\alpha}$	
$H_0 : \mu = \mu_0$	$H_A : \mu < \mu_0$	$Z < -z_{1-\alpha}$	

Test di ipotesi riguardanti la media, essendo nota la varianza.

Statistica $T = (\bar{X}_n - \mu_0)\sqrt{n}/S$ con $n-1$ gradi di libertà.			
Ipotesi nulla	Ipotesi alternativa	Regione di rifiuto di H_0	Schema grafico
$H_0 : \mu = \mu_0$	$H_A : \mu \neq \mu_0$	$T > t_{n-1,1-\alpha/2}$ oppure $T < -t_{n-1,1-\alpha/2}$	
$H_0 : \mu = \mu_0$	$H_A : \mu > \mu_0$	$T > t_{n-1,1-\alpha}$	
$H_0 : \mu = \mu_0$	$H_A : \mu < \mu_0$	$T < -t_{n-1,1-\alpha}$	

Test di ipotesi riguardanti la media con varianza incognita.

TEST SULLA DIFFERENZA TRA DUE MEDIE

Si verifica che due medie relative a due campioni differenti si possano considerare uguali

$$x \approx N(\mu_x, \sigma_x^2) \qquad y \approx N(\mu_y, \sigma_y^2)$$

$$\bar{X} \approx N\left(\mu_x, \frac{\sigma_x^2}{n_x}\right) \qquad \bar{Y} \approx N\left(\mu_y, \frac{\sigma_y^2}{n_y}\right)$$

Se σ_x^2 e σ_y^2 sono noti:

$$\boxed{z_{\bar{X}-\bar{Y}} = \frac{(\bar{X}-\bar{Y})-(\mu_x-\mu_y)}{\sqrt{\sigma_x^2/n_x+\sigma_y^2/n_y}}}$$

TEST SULLA DIFFERENZA TRA DUE MEDIE

Se σ_x^2 e σ_y^2 non sono noti, si suppone $\sigma_x^2 = \sigma_y^2 = \sigma^2$:

$$\frac{s_x^2}{\sigma^2/(n_x-1)} \approx \chi^2_{n_x-1}; \frac{s_y^2}{\sigma^2/(n_y-1)} \approx \chi^2_{n_y-1}$$

$$\Updownarrow$$

$$\frac{s_x^2}{\sigma^2/(n_x-1)} + \frac{s_y^2}{\sigma^2/(n_y-1)} \approx \chi^2_{n_x+n_y-2}$$

$$t_{\bar{X}-\bar{Y}} = \frac{(\bar{X}-\bar{Y})-(\mu_x-\mu_y)}{\sqrt{(n_x-1)s_x^2+(n_y-1)s_y^2}}\sqrt{\frac{n_x n_y(n_x+n_y-2)}{n_x+n_y}}$$

Uso: verifica di omogeneità tra coppie di gruppi di un campione

$$\left.\begin{array}{l} H_0: \mu_x = \mu_y \\ H_1: \mu_x \neq \mu_y \end{array}\right\} \text{bilaterale; rifiuto } H_0 \text{ se } |t| > t_{\alpha/2, n_x+n_y-2} \text{ o } |z| > z_{\alpha/2}$$

$$\left.\begin{array}{l} H_0: \mu_x > \mu_y \\ H_1: \mu_x \leq \mu_y \end{array}\right\} \text{unilaterale; rifiuto } H_0 \text{ se } t < t_{\alpha, n_x+n_y-2} \text{ o } z < z_{\alpha}$$

Notare che se per l'ipotesi 0, $\mu_x - \mu_y = 0$ per cui il test viene condotto sulla seguente espressione:

$$\boxed{t_{\bar{X}-\bar{Y}} = \frac{\bar{X}-\bar{Y}}{\sqrt{(n_x-1)s_x^2+(n_y-1)s_y^2}} \sqrt{\frac{n_x n_y (n_x+n_y-2)}{n_x+n_y}}}$$

Test sulla Varianza

Si verifica che la varianza campionaria assuma un valore noto.

$$x \approx N(\mu_x, \sigma_x^2)$$

si ricorda che:

$$\left.\begin{array}{l} H_0: \sigma_x = \sigma_0 \\ H_1: \sigma_x \neq \sigma_0 \end{array}\right\} \text{bilaterale}$$

$$\boxed{\frac{(n_x-1)s_x^2}{\sigma_0^2} \approx \chi^2_{n_x-1}}$$

essendo la distribuzione χ^2 asimmetrica, il test bilaterale deve essere effettuato sui due valori (max e min) di χ^2

$$\text{non rifiuto } H_0 \text{ se } \quad \chi^2_{1-\frac{\alpha}{2}, n_x-1} \leq \chi^2_{n_x-1} < \chi^2_{\frac{\alpha}{2}, n_x-1}$$

Test sulla Varianza

$$\left.\begin{array}{l} H_0: \sigma_x = \sigma_0 \\ H_1: \sigma_x > \sigma_0 \end{array}\right\} \text{unilaterale; rifiuto } H_0 \text{ se } \chi^2_{n_x-1} > \chi^2_{\alpha, n_x-1}$$

Uso:
- verifica di un peggioramento nella qualità di produzione
- verifica di appartenenza di una classe ad una popolazione con media diversa da quella della classe

TEST SULLA DIFFERENZA TRA DUE VARIANZE

Si verifica che le varianze di due campioni siano eguali.

$$\frac{(n_x-1)s_x^2}{\sigma_x^2} \approx \chi^2_{n_x-1} \quad \frac{(n_y-1)s_y^2}{\sigma_y^2} \approx \chi^2_{n_y-1}$$

$$\boxed{\frac{s_x^2}{\sigma_x^2}\frac{\sigma_y^2}{s_y^2} \approx F_{n_x-1, n_y-1}}$$

$$\left.\begin{array}{l} H_0: \sigma_x = \sigma_y \\ H_1: \sigma_x \neq \sigma_y \end{array}\right\} \text{bilaterale}$$

essendo la distribuzione F asimmetrica, il test bilaterale deve essere effettuato sui due valori (max e min) di F.

$$\text{non rifiuto } H_0 \text{ se } \quad F_{1-\frac{\alpha}{2}, n_x-1, n_y-1} \leq F_{n_x-1} < F_{\frac{\alpha}{2}, n_x-1, n_y-1}$$

Test sulla Differenza tra Due Varianze

$$\left.\begin{array}{l} H_0 : \sigma_x = \sigma_y \\ H_1 : \sigma_x > \sigma_y \end{array}\right\} \text{unilaterale; rifiuto } H_0 \text{ se } F_{n_x - 1} > F_{\alpha, n_x - 1, n_y - 1}$$

Uso:
- verifica di appartenenza ad una stessa popolazione (la cui stima della varianza è quella testata).
- nell'analisi della varianza, per verificare l'influenza di raggruppamenti o variabli

Nota: per l'ipotesi nulla il test viene effettuato sul rapporto delle s^2

TEST della MEDIA con l'USO della DISTRIBUZIONE NORMALE

Trovaree la probabilità che, in 100 lanci di una moneta buona, testa si presenti tra 40 e 60 volte comprese.

Dato che Np = 100(1/2) e Nq = 100(1/2) sono entrambi maggiori di 5, nel calcolo della somma si può ricorrere all'approssimazione normale alla distribuzione binomiale.

La media e lo scarto quadratico medio del numero di volte in cui si presenta testa in 100 lanci sono dati da

$$\mu = Np = 100(\tfrac{1}{2}) = 50 \quad \text{e} \quad \sigma = \sqrt{Npq} = \sqrt{(100)(\tfrac{1}{2})(\tfrac{1}{2})} = 5$$

Nel continuo, invece di 40 è meglio usare 39,5 e invece di 60 è meglio usare 60,5.

39,5 in unità standard vale $(39,5 - 50)/5 = -2,10$; 60,5 in unità standard vale $(60,5 - 50)/5 = 2,10$.

Probabilità richiesta = area sotto la curva normale compresa tra +/- 2,1 :

$$z = -2,10 \text{ e } z = 2,10$$

$$= 2\,(\text{area compresa tra } z = 0 \text{ e } z = 2,10) = 2(0,4821) = 0,9642$$

TEST della MEDIA con l'USO della DISTRIBUZIONE NORMALE

Per provare l'ipotesi che una moneta è buona, viene adottata la seguente regola di decisione:

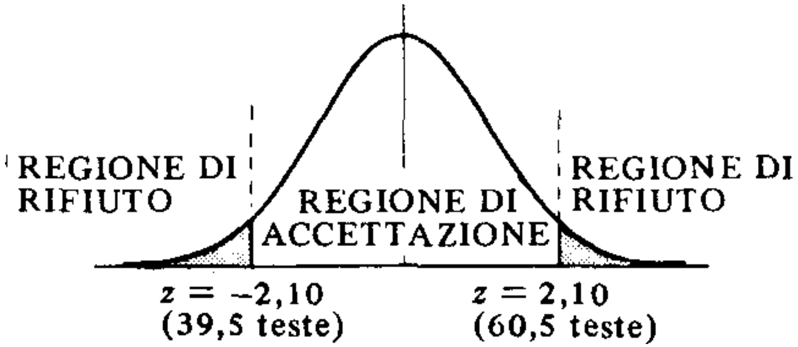
(1) si accetta l'ipotesi se il numero di teste in un campione di 100 lanci è compreso tra 40 e 60 inclusi;

(2) in caso contrario, si rifiuta l'ipotesi.

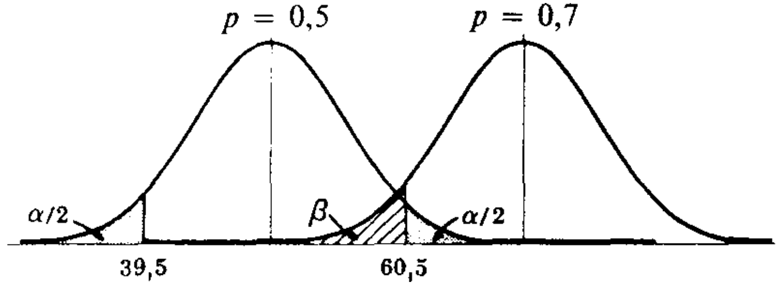
Trovare la probabilità di rifiutare l'ipotesi quando, in realtà, è corretta.

La probabilità che testa non si presenti tra 40 e 60 volte inclusi (se la moneta è buona) è 1 - 0,9642 = 0,0358.

Quindi la probabilità di rifiutare l'ipotesi quando è corretta è pari a 0,0358.



Qua] è la probabilità di accettare l'ipotesi che la moneta è buona quando la reale probabilità che si presenti testa è p = 0,7?



Dal diagramma è chiaro che la probabilità di accettare Ho quando si ha in realtà p = 0,7 (cioè la probabilità dell'errore di Il tipo) è data dall'area tratteggiata ß della figura.

Per calcolare tale area, osserviamo che la distribuzione sotto l'ipotesi p = 0,7 ha media e scarto quadratico medio dati da:

$$\mu = Np = (100)(0.7) = 70 \quad \text{e} \quad \sigma = \sqrt{Npq} = \sqrt{(100)(0,7)(0,3)} = 4,58$$

60,5 in unità standard vale $(60,5 - 70)/4,58 = -2,07$
39,5 in unità standard vale $(39,5 - 70)/4,58 = -6,66$

Quindi β = (area sotto la curva normale compresa tra $z = -6,66$ e $z = -2,07$) = 0,0192