

VERIFICA DI IPOTESI STATISTICHE

La procedura della **VERIFICA DI IPOTESI STATISTICHE**, o procedura dei **test statistici**, riguarda il problema di voler avere maggiori informazioni sul valore di un certo parametro θ della popolazione, che risulta incognito, e consiste nel fissare nell'ipotesi un certo valore θ_0 del parametro. Quindi, sulla base dei risultati campionari, si stabilisce se la differenza riscontrata fra il valore ipotizzato θ_0 e il valore campionario t sia dovuta a cause accidentali o a cause sistematiche; di conseguenza si decide se accettare o rifiutare l'ipotesi prefissata.

Sulla popolazione vengono poste due ipotesi complementari e disgiunte, delle quali una deve essere vera e l'altra falsa: la procedura del test statistico ha lo scopo di far scegliere una delle due.

La prima ipotesi, che solitamente comprende la relazione di uguaglianza, viene detta *ipotesi base* o *ipotesi di nullità* o *ipotesi nulla*: H_0 .

L'altra ipotesi viene detta *ipotesi alternativa*: H_1 .

$H_0 : \theta = \theta_0$ $H_1 : \theta \neq \theta_0$ TEST BILATERALE o TEST A DUE CODE

$H_0 : \theta = \theta_0$ $H_1 : \theta > \theta_0$ TEST UNILATERALE o TEST A CODA DESTRA

$H_0 : \theta = \theta_0$ $H_1 : \theta < \theta_0$ TEST UNILATERALE o TEST A CODA SINISTRA

Al fine di verificare quale delle due ipotesi sia vera, si opera per mezzo di una **funzione TEST**: essa è una funzione di uno STIMATORE T del parametro θ in questione ed è una variabile casuale tabulata (invece i software statistici ricorrono ad altre tecniche equivalenti basate sul p -value).

Fissato un *Livello di significatività* pari ad α , lo spazio dei valori possibili della funzione TEST viene diviso in due insiemi complementari e disgiunti, detti ZONE CRITICHE (i valori che le individuano sono detti VALORI CRITICI):

- un insieme, detto ZONA o INTERVALLO DI ACCETTAZIONE, comprende i valori con densità di probabilità maggiore, pari a $1-\alpha$, ed è costruito in modo da avere una probabilità totale di verificarsi elevata, dato H_0 ;
- l'altro insieme, detto ZONA o INTERVALLO DI RIGETTO o RIFIUTO, comprende i valori con densità di probabilità minore, pari a α , ed ha una probabilità totale di verificarsi piuttosto piccola.

Sul campione estratto si calcola la funzione TEST:

a) se il valore della funzione TEST appartiene al primo insieme, allora si accetta l'*ipotesi nulla* H_0 ;

b) in caso contrario, allora si rifiuta H_0 e si accetta H_1 .

VERIFICA DI IPOTESI SULLA MEDIA DELLA POPOLAZIONE

Per fare inferenza su μ si utilizza lo stimatore Media campionaria \bar{X} . A seconda della distribuzione campionaria di \bar{X} si sceglie la funzione TEST opportuna: se \bar{X} si distribuisce normalmente allora scelgo z_c ; se \bar{X} standardizzata si distribuisce secondo una v.c. t di Student allora scelgo t_c .

$$H_0 : \mu = \mu_0 \quad H_1 : \mu \neq \mu_0 \quad \text{TEST BILATERALE o TEST A DUE CODE}$$

funzione TEST

$$z_c = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Fissato α ed essendo un test BILATERALE, si individuano i due *valori critici* $\pm z_{\frac{\alpha}{2}}$ e quindi le *zone critiche*:

$$\text{ZONA DI ACCETTAZIONE (di } H_0): \left(-z_{\frac{\alpha}{2}}, +z_{\frac{\alpha}{2}} \right)$$

$$\text{ZONE DI RIFIUTO (di } H_0): \left(-\infty, -z_{\frac{\alpha}{2}} \right) \text{ e } \left(+z_{\frac{\alpha}{2}}, +\infty \right)$$

Quindi si confronta il valore della funzione TEST calcolato sul campione (z_c) con i *valori critici*:

a) se il valore della funzione TEST appartiene alla ZONA DI ACCETTAZIONE, allora si accetta l'*ipotesi nulla* H_0 ;

b) in caso contrario, ovvero se la funzione TEST appartiene alle ZONE DI RIFIUTO, allora si rifiuta H_0 e si accetta H_1 .

OPPURE

funzione TEST

$$t_c = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Fissato α ed essendo un test BILATERALE, si individuano i due *valori critici* $\pm t_{\frac{\alpha}{2}}$ e quindi le *zone critiche*:

ZONA DI ACCETTAZIONE (di H_0): $\left(-t_{\frac{\alpha}{2}}, +t_{\frac{\alpha}{2}}\right)$

ZONE DI RIFIUTO (di H_0): $\left(-\infty, -t_{\frac{\alpha}{2}}\right)$ e $\left(+t_{\frac{\alpha}{2}}, +\infty\right)$

Quindi si confronta il valore della funzione TEST calcolato sul campione (t_c) con i *valori critici*:

a) se il valore della funzione TEST appartiene alla ZONA DI ACCETTAZIONE, allora si accetta l'*ipotesi nulla* H_0 ;

b) in caso contrario, ovvero se la funzione TEST appartiene alla ZONE DI RIFIUTO, allora si rifiuta H_0 e si accetta H_1 .

ESERCIZIO Da una certa popolazione è stato estratto un campione casuale di 1000 osservazioni ($n=1000$) di media $\bar{X}=2,065$ e varianza $S^2=2,06153$.

Verificare l'ipotesi che la media della popolazione sia uguale a **1,8** ($\alpha=5\%$).

SOLUZIONI

Le ipotesi in campo sono

$$H_0: \mu = 1,8$$

$$H_1: \mu \neq 1,8$$

Poiché il campione è di dimensione elevata ($n=1000>30$) allora la media campionaria \bar{X} si distribuisce Normalmente; quindi il TEST da applicare è il test z:

$$z_c = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Dove però è incognito σ , essendo incognita la varianza della popolazione; allora bisogna ricorrere alla varianza campionaria corretta S^2 e quindi la formula del test diventa:

$$z_c = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$$

Per determinare la varianza campionaria corretta bisogna "correggere" il valore pari a *2,06153* della varianza del campione:

$$S^2 = (2,06153) \cdot \frac{1000}{999} = 2,0636 \text{ da cui } S = \sqrt{2,0636} = 1,4365$$

Da cui:

$$z_c = \frac{2,065 - 1,8}{\frac{1,4365}{\sqrt{1000}}} = +5,83315$$

Dalle tavole della variabile normale standardizzata si ricavano i due valori critici al livello di significatività del 5% (test bilaterale):

$$\pm z_{\frac{\alpha}{2}} = \pm 1,96$$

Poiché il valore del test calcolato $z_c = +5,83315$ risulta esterno alla zona di accettazione, si rifiuta l'ipotesi nulla ovvero non si ritiene che la popolazione d'origine possa avere una media pari a 1,8.

$H_0 : \mu = \mu_0$ $H_1 : \mu > \mu_0$ TEST UNILATERALE o TEST A CODA DESTRA

funzione TEST

$$z_c = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Fissato α ed essendo un test UNILATERALE a coda destra, si individua il *valore critico* $+z_{\alpha}$ e quindi le *zone critiche*:

ZONA DI ACCETTAZIONE (di H_0): $(-\infty, +z_{\alpha})$

ZONA DI RIFIUTO (di H_0): $(+z_{\alpha}, +\infty)$

Quindi si confronta il valore della funzione TEST calcolato sul campione (z_c) con il *valore critico*:

a) se il valore della funzione TEST appartiene alla ZONA DI ACCETTAZIONE, allora si accetta l'*ipotesi nulla* H_0 ;

b) in caso contrario, ovvero se la funzione TEST appartiene alla ZONA DI RIFIUTO, allora si rifiuta H_0 e si accetta H_1 .

OPPURE

funzione TEST

$$t_c = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Fissato α ed essendo un test UNILATERALE a coda destra, si individua il *valore critico* $+t_\alpha$ e quindi le *zone critiche*:

ZONA DI ACCETTAZIONE (di H_0): $(-\infty, +t_\alpha)$

ZONA DI RIFIUTO (di H_0): $(+t_\alpha, +\infty)$

Quindi si confronta il valore della funzione TEST calcolato sul campione (t_c) con il *valore critico*:

a) se il valore della funzione TEST appartiene alla ZONA DI ACCETTAZIONE, allora si accetta l'*ipotesi nulla* H_0 ;

b) in caso contrario, ovvero se la funzione TEST appartiene alla ZONA DI RIFIUTO, allora si rifiuta H_0 e si accetta H_1 .

$H_0 : \mu = \mu_0$ $H_1 : \mu < \mu_0$ TEST UNILATERALE o TEST A CODA SINISTRA

funzione TEST

$$z_c = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Fissato α ed essendo un test UNILATERALE a coda sinistra, si individua il *valore critico* $-z_\alpha$ e quindi le *zone critiche*:

ZONA DI ACCETTAZIONE (di H_0): $(-z_\alpha, +\infty)$

ZONA DI RIFIUTO (di H_0): $(-\infty, -z_\alpha)$

Quindi si confronta il valore della funzione TEST calcolato sul campione (z_c) con il *valore critico*:

a) se il valore della funzione TEST appartiene alla ZONA DI ACCETTAZIONE, allora si accetta l'*ipotesi nulla* H_0 ;

b) in caso contrario, ovvero se la funzione TEST appartiene alla ZONA DI RIFIUTO, allora si rifiuta H_0 e si accetta H_1 .

OPPURE

funzione TEST

$$t_c = \frac{\bar{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Fissato α ed essendo un test UNILATERALE a coda sinistra, si individua il *valore critico* $-t_\alpha$ e quindi le *zone critiche*:

ZONA DI ACCETTAZIONE (di H_0): $(-t_\alpha, +\infty)$

ZONA DI RIFIUTO (di H_0): $(-\infty, -t_\alpha)$

Quindi si confronta il valore della funzione TEST calcolato sul campione (t_c) con il *valore critico*:

a) se il valore della funzione TEST appartiene alla ZONA DI ACCETTAZIONE, allora si accetta l'*ipotesi nulla* H_0 ;

b) in caso contrario, ovvero se la funzione TEST appartiene alla ZONA DI RIFIUTO, allora si rifiuta H_0 e si accetta H_1 .

VERIFICA DI IPOTESI SULLA VARIANZA DELLA POPOLAZIONE (nell'ipotesi di popolazione distribuita normalmente)

$$H_0 : \sigma^2 = \sigma_0^2 \quad H_1 : \sigma^2 \neq \sigma_0^2 \quad \text{TEST BILATERALE o TEST A DUE CODE}$$

funzione TEST

$$\chi_c^2 = \frac{S^2 \cdot \nu}{\sigma_0^2}$$

Fissato α ed essendo un test BILATERALE, si individuano i due *valori critici* $\chi_{\frac{\alpha}{2}}^2$ e $\chi_{1-\frac{\alpha}{2}}^2$, e quindi le *zone critiche*:

ZONA DI ACCETTAZIONE (di H_0): $\left(\chi_{\frac{\alpha}{2}}^2, \chi_{1-\frac{\alpha}{2}}^2 \right)$

ZONE DI RIFIUTO (di H_0): $\left(0, \chi_{\frac{\alpha}{2}}^2 \right)$ e $\left(\chi_{1-\frac{\alpha}{2}}^2, +\infty \right)$

Quindi si confronta il valore della funzione TEST calcolato sul campione (χ_c^2) con i *valori critici*:

a) se il valore della funzione TEST appartiene alla ZONA DI ACCETTAZIONE, allora si accetta l'*ipotesi nulla* H_0 ;

b) in caso contrario, ovvero se la funzione TEST appartiene alla ZONE DI RIFIUTO, allora si rifiuta H_0 e si accetta H_1 .

ESERCIZIO Da una popolazione distribuita normalmente è stato estratto un campione casuale di 11 osservazioni ($n=11$) con *varianza campionaria corretta* $S^2=1394359$.
Verificare l'ipotesi che la varianza della popolazione sia uguale a **14000** ($\alpha=1\%$).

SOLUZIONI

Le ipotesi in campo sono

$$H_0 : \sigma^2 = 14000 \quad H_1 : \sigma^2 \neq 14000$$

La *varianza campionaria corretta* si distribuisce come una v.c. chi-quadrato, a meno di un coefficiente di proporzionalità; quindi il TEST da applicare è il test χ_c^2 :

$$\chi_c^2 = \frac{S^2 \cdot \nu}{\sigma_0^2} = \frac{1394359 \cdot 10}{14000} = 9,9597$$

Con $\alpha=1\%$ e $v=n-1=10$ si individuano i due valori critici $\chi_{\frac{\alpha}{2}}^2 = \chi_{0,005}^2 = 2,16$ e

$$\chi_{1-\frac{\alpha}{2}}^2 = \chi_{0,995}^2 = 25,19.$$

Poiché il valore del test calcolato $\chi_c^2 = 9,9597$ risulta interno alla zona di accettazione, si accetta l'ipotesi nulla ovvero si ritiene che la popolazione d'origine possa avere una varianza pari a 14000.

$H_0 : \sigma^2 = \sigma_0^2$ $H_1 : \sigma^2 > \sigma_0^2$ TEST UNILATERALE o TEST A CODA DESTRA

funzione TEST

$$\chi_c^2 = \frac{S^2 \cdot v}{\sigma_0^2}$$

Fissato α ed essendo un test UNILATERALE, si individua il *valore critico* $\chi_{1-\alpha}^2$, e quindi le *zone critiche*:

ZONA DI ACCETTAZIONE (di H_0): $(0, \chi_{1-\alpha}^2)$

ZONA DI RIFIUTO (di H_0): $(\chi_{1-\alpha}^2, +\infty)$

Quindi si confronta il valore della funzione TEST calcolato sul campione (χ_c^2) con il *valore critico*:

a) se il valore della funzione TEST appartiene alla ZONA DI ACCETTAZIONE, allora si accetta l'*ipotesi nulla* H_0 ;

b) in caso contrario, ovvero se la funzione TEST appartiene alla ZONA DI RIFIUTO, allora si rifiuta H_0 e si accetta H_1 .

$H_0 : \sigma^2 = \sigma_0^2$ $H_1 : \sigma^2 < \sigma_0^2$ TEST UNILATERALE o TEST A CODA SINISTRA

funzione TEST

$$\chi_c^2 = \frac{S^2 \cdot v}{\sigma_0^2}$$

Fissato α ed essendo un test UNILATERALE, si individua il *valore critico* χ_α^2 , e quindi le *zone critiche*:

ZONA DI ACCETTAZIONE (di H_0): $(\chi_\alpha^2, +\infty)$

ZONA DI RIFIUTO (di H_0): $(0, \chi_\alpha^2)$

Quindi si confronta il valore della funzione TEST calcolato sul campione ($:\chi_c^2$) con il *valore critico*:

a) se il valore della funzione TEST appartiene alla ZONA DI ACCETTAZIONE, allora si accetta l'*ipotesi nulla* H_0 ;

b) in caso contrario, ovvero se la funzione TEST appartiene alla ZONA DI RIFIUTO, allora si rifiuta H_0 e si accetta H_1 .