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 γ della popolazione, che risulta incognito, e consiste nel fissare nell'ipotesi un certo valore γ_0 del parametro. Quindi, sulla base dei risultati campionari, si stabilisce se la differenza riscontrata fra il valore ipotizzato γ_0 e il valore campionario g 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.

$H_0 : \gamma = \gamma_0 \qquad H_1 : \gamma \neq \gamma_0$ TEST BILATERALE o TEST A DUE CODE

$H_0 : \gamma = \gamma_0 \qquad H_1 : \gamma > \gamma_0$ TEST UNILATERALE o TEST A CODA DESTRA

$H_0 : \gamma = \gamma_0 \qquad H_1 : \gamma < \gamma_0$ 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 ***G*** del parametro $\boldsymbol{\gamma}$ 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 $\boldsymbol{\alpha}$, 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 $\boldsymbol{1-\alpha}$, ed è costruito in modo da avere una probabilità totale di verificarsi elevata, dato $\boldsymbol{H_0}$;

- l'altro insieme, detto ZONA o INTERVALLO DI RIGETTO o RIFIUTO, comprende i valori con densità di probabilità minore, pari a $\boldsymbol{\alpha}$, 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* $\boldsymbol{H_0}$;

b) in caso contrario, allora si rifiuta $\boldsymbol{H_0}$ e si accetta $\boldsymbol{H_1}$.

VERIFICA DI IPOTESI SULLA MEDIA DELLA POPOLAZIONE

Per fare inferenza su $\boldsymbol{\mu}$ si utilizza lo stimatore Media campionaria $\boldsymbol{m}$. A seconda della distribuzione campionaria di $\boldsymbol{m}$ si sceglie la funzione TEST opportuna: se $\boldsymbol{m}$ si distribuisce normalmente allora scelgo $\boldsymbol{u_c}$; se $\boldsymbol{m}$ standardizzata si distribuisce secondo una v.c. *t* di Student allora scelgo $\boldsymbol{t_c}$.

$H_0 : \mu = \mu_0 \qquad H_1 : \mu \neq \mu_0$ TEST BILATERALE o TEST A DUE CODE

funzione TEST

$$u_c = \frac{m - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

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

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

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

Quindi si confronta il valore della funzione TEST calcolato sul campione (: $\boldsymbol{u_c}$) con i *valori critici*:

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

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

OPPURE

funzione TEST

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

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 **m**=2,065 e varianza 'S^2=2,06153.
Verificare l'ipotesi che la media della popolazione sia uguale a **1,8** (α=5%).

SOLUZIONI

Le ipotesi in campo sono

$H_0: \quad \mu = 1{,}8 \qquad\qquad H_1: \quad \mu \neq 1{,}8$

Poiché il campione è di dimensione elevata (: n=1000>30) allora la media campionaria m si distribuisce Normalmente; quindi il TEST da applicare è il test u:

$$u_c = \frac{m - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

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:

$$u_c = \frac{m - \mu_0}{\frac{S}{\sqrt{n}}}$$

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

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

Da cui:

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

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

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

Poiché il valore del test calcolato $u_c = +5{,}83315$ 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.

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

funzione TEST

$$u_c = \frac{m - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Fissato $\boldsymbol{\alpha}$ ed essendo un test UNILATERALE a coda destra, si individua il *valore critico* $+u_\alpha$ e quindi le *zone critiche*:
ZONA DI ACCETTAZIONE (di H_0): $(-\infty, +u_\alpha)$
ZONA DI RIFIUTO (di H_0): $(+u_\alpha, +\infty)$

Quindi si confronta il valore della funzione TEST calcolato sul campione (: $\boldsymbol{u_c}$) con il *valore critico*:

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

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

OPPURE
funzione TEST

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

Fissato $\boldsymbol{\alpha}$ 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 (: $\boldsymbol{t_c}$) con il *valore critico*:

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

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

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

funzione TEST

$$u_c = \frac{m - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

Fissato $\boldsymbol{\alpha}$ ed essendo un test UNILATERALE a coda sinistra, si individua il *valore critico* $-u_\alpha$ e quindi le *zone critiche*:
ZONA DI ACCETTAZIONE (di H_0): $(-u_\alpha, +\infty)$
ZONA DI RIFIUTO (di H_0): $(-\infty, -u_\alpha)$

Quindi si confronta il valore della funzione TEST calcolato sul campione (: $\boldsymbol{u_c}$) con il *valore critico*:

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

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

OPPURE

funzione TEST

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

Fissato $\boldsymbol{\alpha}$ 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 (: $\boldsymbol{t_c}$) con il *valore critico*:

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

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

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

$H_0 : \sigma^2 = \sigma^2{}_0$ $\quad H_1 : \sigma^2 \neq \sigma^2{}_0$ TEST BILATERALE o TEST A DUE CODE

<u>funzione TEST</u>

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

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

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

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

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

a) se il valore della <u>funzione TEST</u> appartiene alla ZONA DI ACCETTAZIONE, allora si accetta l'*ipotesi nulla* $\boldsymbol{H_0}$;

b) in caso contrario, ovvero se la <u>funzione TEST</u> appartiene alla ZONE DI RIFIUTO, allora si rifiuta $\boldsymbol{H_0}$ e si accetta $\boldsymbol{H_1}$.

ESERCIZIO Da una popolazione distribuita normalmente è stato estratto un campione casuale di 11 osservazioni (: n=11) con *varianza campionaria corretta* $\mathbf{S^2}$=1394359.
Verificare l'ipotesi che la varianza della popolazione sia uguale a **14000** (α=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 $\boldsymbol{\alpha}$=1% e ν=n-1=10 si individuano i due *valori critici* $\chi^2_{1-\frac{\alpha}{2}} = \chi^2_{0,995} = 2{,}16$ e $\chi^2_{\frac{\alpha}{2}} = \chi^2_{0,005} = 25{,}19$.

Poiché il valore del test calcolato $\chi^2_c = 9{,}9597$ 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.

$H_0 : \sigma^2 = \sigma^2{}_0 \qquad H_1 : \sigma^2 > \sigma^2{}_0$ TEST UNILATERALE o TEST A CODA DESTRA

funzione TEST

$$\chi^2_c = \frac{S^2 \cdot \nu}{\sigma^2_0}$$

Fissato $\boldsymbol{\alpha}$ ed essendo un test UNILATERALE, si individua il *valore critico* χ^2_α, e quindi le *zone critiche*:

ZONA DI ACCETTAZIONE (di H_0): $\left(0, \chi^2_\alpha\right)$

ZONA DI RIFIUTO (di H_0): $\left(\chi^2_\alpha, +\infty\right)$

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

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

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

$H_0 : \sigma^2 = \sigma^2{}_0 \qquad H_1 : \sigma^2 < \sigma^2{}_0$ TEST UNILATERALE o TEST A CODA SINISTRA

funzione TEST

$$\chi^2_c = \frac{S^2 \cdot \nu}{\sigma^2_0}$$

Fissato $\boldsymbol{\alpha}$ ed essendo un test UNILATERALE, si individua il *valore critico* $\chi^2_{1-\alpha}$, e quindi le *zone critiche*:
ZONA DI ACCETTAZIONE (di H_0): $(\chi^2_{1-\alpha}, +\infty)$
ZONA DI RIFIUTO (di H_0): $(0, \chi^2_{1-\alpha})$

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

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

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