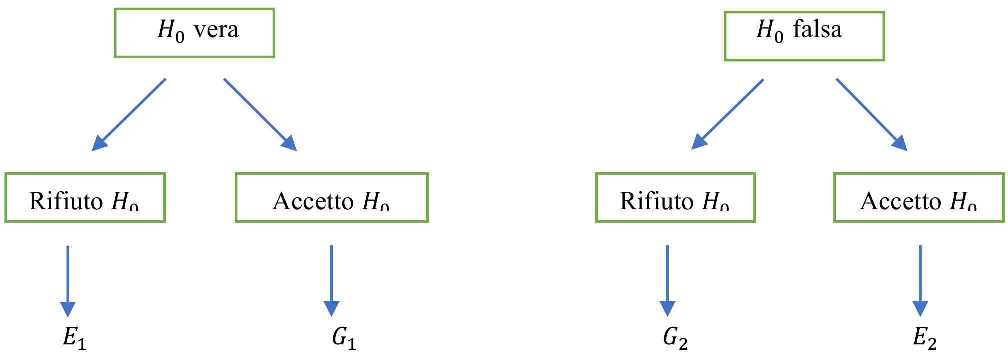
# Test Statistici

Un test statistico è una regola che permette di stabilire se un' ipotesi ($H_0$) può essere accettata (non rifiutata) o meno (rifiutata). In particolare, $H_0$ può essere vera o falsa e la sua accettazione/rifiuto porta alla definizione di alcune casistiche:




dove con $E_i$ è stato indicato l'evento "scelta errata" e con $G_i$ l'evento "scelta giusta", i=1,2.

In termini statistici, agli eventi appena indicati viene attribuita una probabilità ed una specifica definizione:

| | $H_0$ vera | $H_0$ falsa |
|---|---|---|
| Rifiuto | - $E_1$<br>- Errore di I tipo<br>- Livello di significatività<br>- $\alpha$ | - $G_2$<br>- $\gamma = 1 - \beta$ |
| Non rifiuto (accetto) | - $G_1$<br>- $1 - \alpha$ | - $E_2$<br>- Errore di II tipo<br>- $\beta$ |

Più precisamente:

1) $\alpha = P\ (E_1)$ = P (Rifiutare $H_0$ | $H_0$ vera) = Livello di significatività
2) $1 - \alpha = P\ (G_1)$ = P (Non rifiutare $H_0$ | $H_0$ vera)
3) $\beta = \ P\ (E_2)$ = P (Non rifiutare $H_0$ | $H_0$ falsa)
4) $\gamma = 1 - \beta = \ P\ (G_2)$ = P (Rifiutare $H_0$ | $H_0$ falsa) = Potenza del test

Oltre a queste quattro probabilità si è soliti calcolare il p-value che verrà definito tra breve.

In generale, i test statistici possono essere bilaterali o unilaterali: quando all'ipotesi statistica $H_0$ relativa al valore di un parametro $\theta = \theta_0$ (solitamente $\mu, \sigma, o\ p$) si contrappone un ipotesi statistica alternativa ($H_1$) del tipo $\theta \neq \theta_0$ si configura un test bilaterale; quando all'ipotesi statistica $H_0$ relativa al valore di un parametro $\theta = \theta_0$ si contrappone un ipotesi statistica alternativa ($H_1$) del tipo $\theta < \theta_0$ o del tipo $\theta > \theta_0$ si configura un test unilaterale (sinistro nel primo caso, destro nel secondo).

La logica con cui si "operativizza" un test statistico è molto simile a quella utilizzata per la costruzione di intervalli di confidenza. Ad esempio, è noto che, fissato un livello di confidenza al 95% per l'ignoto parametro $\mu$ di una v.c. Normale con varianza nota, si ha:

$$P\left(z_{\alpha/2} \leq \frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}} \leq z_{1-\alpha/2}\right) = 0.95$$

con $\bar{x}$ che indica la media (campionaria) valutata su un campione di ampiezza $n$ e $z_{\alpha/2}$ e $z_{1-\alpha/2}$ quantili della v.c. N(0,1) essendo $\frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}}$ una v.c. N(0,1).

Un test statistico bilaterale ha la seguente struttura:

$$P\left(\psi_{\alpha/2} \leq \Psi \leq \psi_{1-\alpha/2}\right) = 1-\alpha$$

con $\Psi$ che indica una generica statistica e $\psi_{\cdot}$ che indicano i quantili della distribuzione relativa a $\Psi$.
In particolare, il valore assunto da detti quantili rappresentano una soglia che divide un insieme di valori coerenti con $H_0$ detto "regione di accettazione" ed un insieme di valori non coerenti con $H_0$ detto "regione critica".

Nella pratica, una volta calcolato il valore assunto dalla statistica $\Psi$, si valuta se tale valore è contenuto nella regione di accettazione o in quella critica.

Esempi di statistiche $\Psi$ sono:

$$Z = \frac{\bar{x}-\mu_0}{\sqrt{\frac{\sigma^2}{n}}} \sim N(0,1)$$

$$T = \frac{\bar{x}-\mu_0}{\sqrt{\frac{s^2}{n}}} \sim T\ di\ Student\ (n-1) \quad con \quad s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$Z = \frac{\bar{X}-\bar{Y}}{\sqrt{\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}}} \sim N(0,1)$$

$$T = \frac{\bar{X}-\bar{Y}}{S_*\sqrt{\frac{1}{n}+\frac{1}{m}}} \sim T\ di\ Student(n+m-2) \ \text{ con } \ S_*^2 = \frac{s_X^2(n-1)+s_Y^2(m-1)}{n+m-2}$$

$$\chi^2 = \frac{(n-1)\cdot s^2}{\sigma^2} \sim \chi^2(n-1)$$

$$Z = \frac{p-p_0}{\sqrt{\frac{p_0\cdot(1-p_0)}{n}}} \sim N(0,1)$$

Un altro modo per stabilire il rifiuto o meno di $H_0$ è quello di calcolare il p-value, che è definito come la probabilità che la statistica considerata assuma un valore più "estremo" rispetto a quello osservato.
Come anticipato, i test statistici possono essere bilaterali o unilaterali in base al tipo di ipotesi alternativa. Nel caso di test unilaterali, per valori più"estremi" si intendono valori della statistica minori o maggiori, in base all'ipotesi alternativa, rispetto a quelli osservati.

Detta S una generica statistica:

$H_0 : \theta = \theta_0$, vs $H_1 : \theta > \theta_0$, p-value=P(S>s)

$H_0 : \theta = \theta_0$, vs $H_1 : \theta < \theta_0$, p-value=P(S<s)

$H_0 : \theta = \theta_0$, vs $H_1 : \theta \neq \theta_0$, p-value=2P(S>|s|)

Se p-value < $\alpha$ ciò indica che siamo in presenza di un valore "estremo" (molto lontano rispetto a quelli coerenti con $H_0$) e pertanto rifiuto $H_0$.

È quindi fondamentale conoscere la distribuzione della statistica che si sta utilizzando e il/i quantili di riferimento.
Infine, ad ogni test è associata una misura della "bontà" del test. Detta misura è la "potenza del test":

$\gamma = 1 - \beta = P(G_2)$ = P (Rifiutare $H_0$ | $H_0$ falsa) = P (Rifiutare $H_0$ | $H_1$ vera)

dove per P (Rifiutare $H_0$ | $H_0$ falsa) si intende la probabilità che il valore s della statistica S ricada nella regione/i di rifiuto definita/e dal/i quantile/e di riferimento.

ESERCIZI

1) Una ditta produttrice di barre d'acciaio sostiene che i propri prodotti abbiano una lunghezza media di 3000 cm con varianza pari a 180 cm^2. Supposto che la lunghezza delle barre si distribuisca normalmente, per verificare questa affermazione si estrae un campione casuale di 10 barre. La media campionaria risulta essere pari a 2980 cm. A livello di significatività del 5% verificare la veridicità della suddetta affermazione, contro l'alternativa che la lunghezza dichiarata sia minore. Calcolare il p-value.

In questi termini il test che si viene a configurare è:

$$H_0 : \mu = 3000 \quad \text{vs} \quad H_1 : \mu < 3000,$$

Essendo nota la varianza, la statistica di riferimento è $Z = \frac{\bar{x}-\mu_0}{\sqrt{\frac{\sigma^2}{n}}}$ = -4.7140.

- Trattandosi di un test unilaterale sinistro la regione critica è formata dai valori che lasciano a sinistra una probabilità pari a $\alpha = 0.05$ e pertanto, il quantile di riferimento è $z_\alpha = -1.64$.
  Essendo $\mathrm{Z} = -4.7140 < -1.64 = z_\alpha$, rifiuto $H_0$.

- p-value=P(Z<z)= P(Z< $-4.7140$)= 0.0000012142 e anche in questo caso il test porta a rifiutare $H_0$ (test significativo).

3) Un allenatore di una squadra di 11 corridori, per valutare l'efficacia di una nuova tipologia di allenamento, sottopone 5 atleti al nuovo tipo di allenamento e ai rimanenti l'allenamento tradizionale. Alla fine del periodo di prova, registra il tempo di percorrenza di un circuito, che si assume essere distribuito come una v.c. Normale, degli 11 atleti:

| | | | | | | |
|---|---|---|---|---|---|---|
| Allenamento tradizionale | 5.5 | 6.0 | 7.0 | 6.0 | 7.5 | 6.2 |
| Nuovo allenamento: | 5.5 | 6.0 | 8.5 | 7.0 | 6.5 | |

Ad un livello $\alpha$ =0.05 dire se esiste differenza fra i due allenamenti. Calcolare pertanto la statistica test e dire se si accetta o meno l'ipotesi nulla.

Essendo il tempo distribuito come una v.c. Normale di parametri ignoti e dovendo effettuare un confronto tra due campioni, la statistica di riferimento è $T = \frac{\bar{x}-\bar{y}}{S_*\sqrt{\frac{1}{n}+\frac{1}{m}}}$ che ha distribuzione T di Student (n+m-2), con $S_* = \frac{S_X^2(n-1)+S_Y^2(m-1)}{n+m-2}$.

Il test da effettuare è quindi $H_0 : \mu_X = \mu_Y$ vs $H_1 : \mu_X \neq \mu_Y$ o, in altri termini, $H_0 : \mu = 0$ vs $H_1 : \mu \neq 0$, con $\mu = \mu_X - \mu_Y$

Dai dati ottenuti si ricava:

$$n = 6;\ m = 5;\ \bar{x} = 6.3667;\ \bar{y} = 6.7;\ S_X^2 = 0.5467;\ S_Y^2 = 1.3250;$$
$$S_* = 0.9448;\ T = -0.5827;\ t_{\frac{\alpha}{2},m+n-2} = -2.2622\ t_{1-\frac{\alpha}{2},m+n-2} = 2.2622$$

Dato -2.2622<-0.5827<2.2622 si accetta $H_0$ e si conclude che tra i due tipi di allenamento non vi è differenza.

4) Una nota azienda di pneumatici dichiara che il 55% dei propri pneumatici percorre più di 100000 km. Al fine di verificare tale affermazione, viene condotta un'indagine basata su un campione casuale di 80 pneumatici dei quali 50 percorrono più di 100000 km. Dato un livello di significatività $\alpha$ =0.05, si richiede di effettuare un test sulle proporzioni, verificando l'affermazione fatta dall'azienda produttrice contro l'alternativa che tale percentuale sia minore.

Dato che si tratta di un test sulle proporzioni, la statistica di riferimento è $Z = \frac{p-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$ che ha distribuzione N(0,1).

Il test che si viene a configurare è unilaterale (sinistro) $H_0 : p = 0.55$ vs $H_1 : p < 0.55$ e pertanto il valore di riferimento è $z_\alpha = -1.6449$.

Dato che dall'indagine campionaria si ricava p=50/80=0.6250, si ha che:

$$Z = \frac{p - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0.6250 - 0.55}{\sqrt{\frac{0.55(1-0.55)}{80}}} = 1.3484$$

Dato che 1.3484>-1.6449 si accetta $H_0$.