Inferenza statistica parametrica: verifica d'ipotesi

20 maggio 2019

Verifica di un'ipotesi statistica

Supponiamo, come in precedenza di disporre di un campione aleatorio proveniente da una distribuzione nota a meno di uno o più parametri incogniti.

Problema: non vogliamo più stimare direttamente questi parametri, ma piuttosto utilizzare il campione per verificare qualche ipotesi sui parametri.

Esempio. Un'impresa edile acquista una grossa partita di cavi elettrici con una resistenza media alla rottura di 48.26 Megapascal. La ditta vuole verificare se quei cavi hanno effettivamente quella resistenza. A questo scopo prende un campione di 10 esemplari e li testa. I dati così ottenuti possono essere utilizzati per stabilire se accettare o meno l'ipotesi del produttore che la resistenza media sia almeno 48,26 Megapascal.

Definizione: ipotesi statistica.

Una ipotesi statistica è una affermazione su uno o più parametri della distribuzione della popolazione.

N.B. Si parla di ipotesi perchè a priori non sappiamo se è vera o no.

Definizione: test per la verifica di un'ipotesi.

Data una ipotesi statistica, un test per la verifica di questa ipotesi è una procedura per determinare se i valori di un campione aleatorio e l'ipotesi sono compatibili oppure no.

Esempio iniziale. Assumiamo che la resistenza dei cavi (misurata in Megapascal) sia un v.a. con una densità di media incognita, che indichiamo con μ. Osserviamo 10 campioni e misuriamo la loro resistenza alla rottura. Vogliamo sulla base di questo campione verificare l'ipotesi statistica:

$$\mu \geq 48.26.$$

Se secondo una procedura (test) il campione sarà giudicato compatibile con l'ipotesi considerata diremo che quest'ultima è "accettata", altrimenti che è "rifiutata".

ATTENZIONE: Quando accettiamo un'ipotesi non stiamo affermando che essa sia necessariamente vera, ma solo che i dati raccolti sono "accettabilmente " in accordo con essa, cioè che non la escludono.

Regione critica e livello di significativià di un test

Consideriamo una popolazione avente distribuzione F_θ che dipende da un parametro incognito θ e supponiamo di voler verificare una qualche ipotesi su θ che chiameremo ipotesi nulla e che indicheremo con $\mathbb{H}_0$.

Esempio: $F_\theta = \mathcal{N}(\theta, 4)$. Due possibili ipotesi sono:

1. $\mathbb{H}_0 : \theta = 1$;
2. $\mathbb{H}_0 : \theta \leq 1$.

Definizione: ipotesi semplice e composta.

Una ipotesi statistica si dice semplice se caratterizza completamente la distribuzione della popolazione, altrimenti è detta composta.

Nell'esempio 1 l'ipotesi $\mathbb{H}_0$ è semplice in 2 l'ipotesi $\mathbb{H}_0$ è composta.

Sia $X_1, \ldots, X_n$ un campione aleatorio estratto da F_θ.

Supponiamo di voler usare il campione per eseguire una verifica o test di una certa ipotesi $\mathbb{H}_0$.

Poichè dobbiamo decidere se accettare o meno $\mathbb{H}_0$ basandosi esclusivamente sui valori assunti dal campione (i dati), il test sarà definito da quella che viene chiamata la sua regione critica.

Definizione: regione critica di un test.

Un test per la verifica di un'ipotesi $\mathbb{H}_0$ ha regione critica $C \subseteq \mathbb{R}^n$ se, avendo osservato $X_1 = x_1, \ldots, X_n = x_n$,

- accetto $\mathbb{H}_0$ se $(x_1, \ldots, x_n) \notin C$;
- rifiuto $\mathbb{H}_0$ se $(x_1, \ldots, x_n) \in C$.

Esempio: sia $X_1, \ldots, X_n$ un campione aleatorio estratto da $\mathcal{N}(\theta, 4)$. Vogliamo testare:

$$\mathbb{H}_0 : \ \theta = 1.$$

Una regione critica è

$$C = \left\{(x_1, \ldots, x_n) : \ \left|\frac{1}{n}\sum_{i=1}^{n} x_i - 1\right| \geq k\right\},$$

per k opportuno (vedremo come determinarlo).

È importante notare che in un qualunque test per verificare un'ipotesi nulla $\mathbb{H}_0$ si può sbagliare in due modi differenti:

- commettiamo un errore di $\mathbb{I}$ specie (o tipo) quando i dati ci portano a rifiutare $\mathbb{H}_0$ e $\mathbb{H}_0$ è corretta;
- commettiamo un errore di $\mathbb{II}$ specie (o tipo) quando i dati ci portano a accettare $\mathbb{H}_0$ e $\mathbb{H}_0$ è falsa.

Non vi è simmetria tra i due tipi di errore.

L'obbiettivo nella verifica d'ipotesi non è quello di dire se l'ipotesi $\mathbb{H}_0$ sia vera o falsa, ma piuttosto di dire se l'ipotesi fatta sia compatibile con i dati raccolti. Quindi vi è un'ampia tolleranza nell'accettare $\mathbb{H}_0$, mentre per rifiutarla occorre che i dati campionari siano molto improbabili quando $\mathbb{H}_0$ è soddisfatta.

Questo bilanciamento si ottiene specificando una soglia α per la probabilità dell'errore di $\mathbb{I}$ specie:

Livello di significatività di un test

Si dice che un test ha livello di significatività (almeno) α se

$$\mathbb{P}(\text{errore di } \mathbb{I} \text{ specie}) = \mathbb{P}_{\mathbb{H}_0}(\text{rifiutare } \mathbb{H}_0\ ) \leq \alpha$$

Valori tipici del livello di significatività sono 0.1, 0.05 e 0.01.

Regione critica di livello α

La regione critica C di un test è detta di livello di significatività α se

$$\sup_{\theta \in \mathbb{H}_0} \mathbb{P}_\theta((X_1, \dots, X_n) \in C) \leq \alpha.$$

Quindi il livello di significatività di un test è una soglia per la probabilità di rifiutare l'ipotesi $\mathbb{H}_0$, calcolata supponendo $\mathbb{H}_0$ vera.

Come si costruisce la regione critica di un test?

Supponiamo di voler verificare l'ipotesi nulla

$$\mathbb{H}_0 : \theta \in \Psi$$

dove Ψ è un insieme di valori possibili per il parametro.

$\Rightarrow$ Un approccio naturale: individuare uno stimatore puntuale $d(X_1, \ldots, X_n)$ di θ e quindi

rifiutare l'ipotesi quando $d(X_1, \ldots, X_n)$ è "lontano" dall'insieme Ψ.

Quanto deve essere è "lontano" $d(X_1, \ldots, X_n)$ da Ψ per giustificare un rifiuto di $\mathbb{H}_0$ ad un livello di significatività α?.

$\Rightarrow$ Occorre conoscere la distribuzione di $d(X_1, \ldots, X_n)$, quando $\mathbb{H}_0$ è vera. Questa conoscenza permette di determinare la regione critica del test, cioè la regione di lontananza in modo tale che la probabilità dell'errore di $\mathbb{I}$ specie sia minore o uguale ad α.

Verifica d'ipotesi per la media di una popolazione gaussiana quando la varianza è nota: test z

Sia $X_1, \ldots, X_n$ campione estratto da una popolazione $\mathcal{N}(\mu, \sigma_0^2)$ con σ_0^2 varianza nota.

1. Test bilatero.

Fissata una costante μ_0, vogliamo verificare l'ipotesi nulla

$\mathbb{H}_0 : \mu = \mu_0$

contro l'ipotesi alternativa

$\mathbb{H}_1 : \mu \neq \mu_0$.

Essendo $\overline{X}_n$ lo stimatore M.L. di μ appare ragionevole rifiutare $\mathbb{H}_0$ quando $\overline{X}_n$ è lontana da μ_0, cioè considerare una regione critica del test del tipo

$$C := \left\{ (x_1, \ldots, x_n) \ : |\bar{x}_n - \mu_0| \geq k \right\}$$

dove $\bar{x}_n = \frac{x_1 + \cdots + x_n}{n}$.

Vogliamo che il test abbia livello di significatività α, quindi fissiamo k in modo tale che l'errore di $\mathbb{I}$ specie sia α.

$$\begin{aligned}\alpha &= \mathbb{P}(\text{errore di } \mathbb{I} \text{ specie})\\ &= \mathbb{P}_{\mathbb{H}_0}((X_1, \ldots, X_n) \in C)\\ &= \mathbb{P}_{\mu_0}(|\overline{X}_n - \mu_0| \geq k)\end{aligned}$$

Se $\mu = \mu_0 \implies \overline{X}_n \sim \mathcal{N}(\mu_0, \sigma_0^2/n) \implies \frac{(\overline{X}_n - \mu_0)\sqrt{n}}{\sigma_0} \sim \mathcal{N}(0,1)$.
Quindi

$$\begin{aligned}\alpha &= \mathbb{P}_{\mu_0}\left(\frac{|\overline{X}_n - \mu_0|}{\sigma_0/\sqrt{n}} \geq \frac{k\sqrt{n}}{\sigma_0}\right)\\ &= \mathbb{P}\left(|Z| \geq \frac{k\sqrt{n}}{\sigma_0}\right) = 2P\left(Z \geq \frac{k\sqrt{n}}{\sigma_0}\right)\end{aligned}$$

dove $Z \sim \mathcal{N}(0,1)$.

$$\Longrightarrow \mathbb{P}\left(Z \geq \frac{k\sqrt{n}}{\sigma_0}\right) = \frac{\alpha}{2} \iff \frac{k\sqrt{n}}{\sigma_0} = z_{\alpha/2}$$

$$\Longrightarrow C := \left\{(x_1, \ldots, x_n) \ : \ \frac{|\bar{x}_n - \mu_0|\sqrt{n}}{\sigma_0} \geq z_{\alpha/2}\right\}.$$

Possiamo concludere che un test di livello di significatività α è quello che, avendo osservato il valore della media campionaria $\overline{X}_n = \bar{x}_n$,

$$\text{rifiuta } \mathbb{H}_0 \text{ se } \frac{|\bar{x}_n - \mu_0|\sqrt{n}}{\sigma_0} \geq z_{\alpha/2}$$

$$\text{accetta } \mathbb{H}_0 \text{ se } \frac{|\bar{x}_n - \mu_0|\sqrt{n}}{\sigma_0} < z_{\alpha/2}.$$

$\dfrac{(\overline{X}_n - \mu_0)\sqrt{n}}{\sigma_0}$ è detta statistica test $\Longrightarrow$ rifiuto $\mathbb{H}_0$ se il modulo del valore osservato della statistica test è maggiore o uguale al quantile $z_{\alpha/2}$.

Esercizio.

Un segnale di valore μ trasmesso da una sorgente A, viene raccolto dal ricevente B con un rumore gaussiano di media nulla e varianza 4, il segnale ricevuto da B è quindi una v.a. gaussiana $\mathcal{N}(\mu, 4)$. Per ridurre il rumore, viene inviato 5 volte e la media campionaria dei valori ricevuti è $\overline{X}_5 = 9.5$. Si sa inoltre che B aveva motivo di supporre che il valore inviato deve essere 8. Si verifichi questa ipotesi.

soluzione

Si tratta di verificare:

$$\mathbb{H}_0 : \mu = 8 \quad \text{contro} \quad \mathbb{H}_1 : \mu \neq 8.$$

La regione critica di un test di livello $\alpha = 0.05$ è

$$C := \left\{ (x_1, \dots, x_5) \; : \; \frac{|\bar{x}_5 - 8|\sqrt{5}}{2} \geq z_{0.025} \simeq 1.96 \right\}$$

$\Longrightarrow$ rifiuto $\mathbb{H}_0$ se il valore osservato $\overline{X}_5 = 9.5$ sta in C:

$$\frac{|\bar{x}_5 - 8|\sqrt{5}}{2} = \frac{|9.5 - 8|\sqrt{5}}{2} \simeq 1.68 < 1.96 \simeq z_{\alpha/2}$$

$\Longrightarrow$ accetto $\mathbb{H}_0$ al livello di significatività $\alpha = 0.05$.

Se $\alpha = 0.1$, $z_{\alpha/2} = z_{0.05} \simeq 1.645$

$$\frac{|\bar{x}_5 - 8|\sqrt{5}}{2} = \frac{|9.5 - 8|\sqrt{5}}{2} \simeq 1.68 > 1.645 \simeq z_{\alpha/2}$$

$\Longrightarrow$ rifiuto $\mathbb{H}_0$ al livello di significatività $\alpha = 0.1$.

N.B. Più α è grande più facilmente rifiuto.

Qual è il livello di significatività corretto?
Dipende.

$p-value$

La regione critica del test in funzione del livello α è

$$C_\alpha := \left\{(x_1, \ldots, x_5) \ : \ \frac{|\bar{x}_5 - 8|\sqrt{5}}{2} \geq z_{\alpha/2}\right\}$$

i.e. $\mathbb{P}_{\mu_0=8}\left\{(X_1, \ldots, X_5) \in C_\alpha\right\} = \mathbb{P}_{\mu_0=8}\left\{\frac{|\overline{X}_5-8|\sqrt{5}}{2} \geq z_{\alpha/2}\right\} = \alpha.$
Avendo osservato $\overline{X}_n = 9.5$ (i dati) è fissato il valore del modulo della statistica test

$$\frac{|\overline{X}_n - \mu_0|\sqrt{n}}{\sigma_0} = \frac{|9.5 - 8|\sqrt{5}}{2} \simeq 1.68$$

quindi

$\forall\alpha$ t.c. $z_{\alpha/2} \leq 1.68 \Rightarrow$ rifiuto $\mathbb{H}_0$ con i dati che ho a disposizione;

$\forall\alpha$ t.c. $z_{\alpha/2} > 1.68 \Rightarrow$ accetto $\mathbb{H}_0$ con i dati che ho a disposizione.

Inoltre z_α decresce al crescere di α. Quindi esiste un valore critico del livello di significatività α, detto $p-value$, al di sopra del quale rifiuto $\mathbb{H}_0$ con i dati a disposizione. In formule:

$$\text{rifiuto } \mathbb{H}_0 \iff z_{\alpha/2} \leq 1.68 \iff \Phi(z_{\alpha/2}) \leq \Phi(1.68)$$

perchè Φ è una funzione crescente. Ma

$$\begin{aligned} & 1-\alpha/2 = \Phi(z_{\alpha/2}) \leq \Phi(1.68) \\ \iff & \alpha/2 \geq 1-\Phi(1.68) \\ \iff & \quad \alpha \geq 2(1-\Phi(1.68)) = p-value. \end{aligned}$$

N.B. $p-value = 2(1-\Phi(1.68)) = \mathbb{P}_{\mu_0=8}\left\{\frac{|\overline{X}_5-8|\sqrt{5}}{2} \geq 1.68\right\}$.
O anche il $p-value$ è quel valore di p tale che $z_{p/2} = 1.68$.

Definizione: $p-value$ (o $p-dei\ dati$)

Data una famiglia di test al variare del livello di significatività α, si definisce $p-value$ è il più piccolo valore della significatività per cui rifiuto $\mathbb{H}_0$ con i dati a disposizione.

Quindi

$$\text{se } \alpha \geq p-value \Longrightarrow \text{rifiuto } \mathbb{H}_0 \text{ con i dati a disposizione;}$$

$$\text{se } \alpha < p-value \Longrightarrow \text{accetto } \mathbb{H}_0 \text{ con i dati a disposizione.}$$

Nell'esercizio:

$$p-value = 2(1-\Phi(1.68)) = \mathbb{P}_{\mu_0=8}\left\{\frac{|\overline{X}_5-8|\sqrt{5}}{2} \geq 1.68\right\} \simeq 0.087$$

$\alpha = 0.1 > p-value \Longrightarrow$ rifiuto $\mathbb{H}_0$ con i dati a disposizione;

$\alpha = 0.05 < p-value \Longrightarrow$ accetto $\mathbb{H}_0$ con i dati a disposizione.

Probabilità dell'errore di $\mathbb{II}$ specie e potenza del test

Definizione: Curva OC (Operating Characteristic curve)

Si definisce curva OC del test la funzione del parametro incognito μ

$$\beta(\mu) := \mathbb{P}_\mu(\text{accettare } \mathbb{H}_0).$$

Quindi, se μ soddisfa l'ipotesi alternativa $\mathbb{H}_1$, $\beta(\mu)$ rappresenta una probabilità di un errore di $\mathbb{II}$ specie.

Definizione: Funzione di potenza

Si definisce potenza del test la funzione del parametro incognito μ, per μ che soddisfa $\mathbb{H}_1$

$$\pi(\mu) := \mathbb{P}_\mu(\text{rifiutare } \mathbb{H}_0) = 1 - \beta(\mu).$$

Per il test di livello α considerato prima

$$\begin{aligned}
\beta(\mu) := \mathbb{P}_\mu(\text{accettare } \mathbb{H}_0) &= \mathbb{P}_\mu\Big(-z_{\alpha/2} < \frac{(\overline{X}_n - \mu_0)\sqrt{n}}{\sigma_0} < z_{\alpha/2}\Big) \\
&= \mathbb{P}_\mu\Big(-z_{\alpha/2} < \frac{(\overline{X}_n - \mu)\sqrt{n}}{\sigma_0} + \frac{(\mu - \mu_0)\sqrt{n}}{\sigma_0} < z_{\alpha/2}\Big) \\
&= \mathbb{P}_\mu\Big(-z_{\alpha/2} + \frac{(\mu_0 - \mu)\sqrt{n}}{\sigma_0} < \frac{(\overline{X}_n - \mu)\sqrt{n}}{\sigma_0} < z_{\alpha/2} + \frac{(\mu_0 - \mu)\sqrt{n}}{\sigma_0}\Big) \\
&= \Phi\Big(z_{\alpha/2} + \frac{(\mu_0 - \mu)\sqrt{n}}{\sigma_0}\Big) - \Phi\Big(-z_{\alpha/2} + \frac{(\mu_0 - \mu)\sqrt{n}}{\sigma_0}\Big).
\end{aligned}$$

Esercizio. $\mu_0 = 8$, $\sigma_0^2 = 4$, $n = 5$, $\alpha = 0.05$ e $z_{\alpha/2} = z_{0.025} = 1.96$ e, per esempio, il valore della potenza del test $\mu = 10$

$$1 - \beta(10) = 1 - \left[\Phi\left(z_{\alpha/2} + \frac{(\mu_0 - \mu)\sqrt{n}}{\sigma_0}\right) - \Phi\left(-z_{\alpha/2} + \frac{(\mu_0 - \mu)\sqrt{n}}{\sigma_0}\right)\right]$$
$$= 1 - [\Phi(1.96 - \sqrt{5}) - \Phi(-1.96 - \sqrt{5})]$$
$$= 1 - \Phi(-0.276) + \Phi(-4.196) \simeq 1 - 0.391 = 0.609.$$

La curva OC può essere usata per dimensionare il campione in modo tale che l'errore di $\mathbb{II}$ specie soddisfi delle condizioni specifiche.

Esempio. Determinare n in modo tale che la probabilità di accettare $\mathbb{H}_0 : \mu = \mu_0$ quando il vero valore è $\mu = \mu_1$ sia minore o uguale a β cioè:

$$\beta(\mu_1) = \Phi\left(\frac{\mu_0 - \mu_1}{\sigma_0/\sqrt{n}} + z_{\alpha/2}\right) - \Phi\left(\frac{\mu_0 - \mu_1}{\sigma_0/\sqrt{n}} - z_{\alpha/2}\right) \leq \beta.$$

Supponiamo $\mu_1 > \mu_0$. Osserviamo che

$$\beta(\mu_1) \leq \Phi\left(\frac{\mu_0 - \mu_1}{\sigma_0/\sqrt{n}} + z_{\alpha/2}\right)$$

e

$$\Phi\left(\frac{\mu_0 - \mu_1}{\sigma_0/\sqrt{n}} + z_{\alpha/2}\right) \leq \beta \iff \frac{\mu_0 - \mu_1}{\sigma_0/\sqrt{n}} + z_{\alpha/2} \leq -z_\beta$$
$$\iff n \geq \sigma_0^2 \left(\frac{z_{\alpha/2} + z_\beta}{\mu_0 - \mu_1}\right)^2 .$$

Anche nel caso $\mu_1 < \mu_0$ si ottiene la stessa formula.

2. Test unilaterale.

Sia $X_1, \ldots, X_n$ campione aleatorio estratto da $\mathcal{N}(\mu, \sigma_0^2)$ con σ_0^2 nota.

$\mathbb{H}_0 : \mu = \mu_0$
contro l'ipotesi alternativa
$\mathbb{H}_1 : \mu > \mu_0$.

Chiaramente in questo caso valori bassi di $\overline{X}_n$ non ci dovrebbero far rifiutare l'ipotesi nulla. Infatti è più probabile ottenere valori piccoli di $\overline{X}_n$ quando $\mathbb{H}_0$ è vera che non quando è vera $\mathbb{H}_1$.

Perciò in questo caso dovremmo rifiutare l'ipotesi nulla quando $\overline{X}_n$, lo stimatore di μ, è molto più grande di μ_0 e quindi una regione critica ragionevole per il test è:

$$C := \left\{(x_1, \dots, x_n) \ : \ \frac{x_1 + \cdots + x_n}{n} - \mu_0 \geq k\right\}$$

dove la costante k viene scelta in modo tale che il test abbia significatività α:

$$\begin{aligned}\alpha &= P_{\mathbb{H}_0}((X_1, \dots, X_n) \in C) \\ &= P_{\mu_0}(\overline{X}_n - \mu_0 > k) = P_{\mu_0}\left(\frac{(\overline{X}_n - \mu_0)\sqrt{n}}{\sigma_0} \geq \frac{k}{\sigma_0}\sqrt{n}\right).\end{aligned}$$

$$\Longrightarrow \quad \frac{k\sqrt{n}}{\sigma_0} = z_\alpha \text{ cioè } k = \frac{\sigma_0 z_\alpha}{\sqrt{n}}.$$

Quindi un test di livello α per il problema di verifica d'ipotesi

$\mathbb{H}_0 : \mu = \mu_0$ contro $\mathbb{H}_1 : \mu > \mu_0$ ha regione critica

$$C := \left\{ (x_1, \dots, x_n) \ : \ \frac{\frac{(x_1+\cdots+x_n)}{n} - \mu_0}{\sigma_0} \sqrt{n} \geq z_\alpha \right\}.$$

Quindi, osservato il valore $\overline{X}_n = \frac{X_1+\cdots+X_n}{n} = \bar{x}_n$ per la media campionaria,

$$\text{rifiuta } \mathbb{H}_0 \text{ se } \frac{(\bar{x}_n - \mu_0)}{\sigma_0}\sqrt{n} \geq z_\alpha$$

$$\text{accetta } \mathbb{H}_0 \text{ se } \frac{(\bar{x}_n - \mu_0)}{\sigma_0}\sqrt{n} < z_\alpha.$$

$\frac{(\overline{X}_n - \mu_0)\sqrt{n}}{\sigma_0}$ statistica test $\Longrightarrow$ rifiuto $\mathbb{H}_0$ se il valore osservato della statistica test è maggiore o uguale al quantile z_α.

Operativamente il *p−value* si ottiene nel seguente modo: si calcola $\dfrac{(\bar{x}_n - \mu_0)}{\sigma_0}\sqrt{n} = \bar{y}$, allora

$$
\begin{aligned}
p\text{−}value &= P_{\mu_0}\left(\frac{(\overline{X}_n - \mu_0)}{\sigma_0}\sqrt{n} \geq \bar{y}\right) \\
&= P(Z \geq \bar{y}) = 1 - \Phi(\bar{y}).
\end{aligned}
$$

Infatti

$$
\begin{aligned}
\alpha < 1 - \Phi(\bar{y}) \quad &\Rightarrow \quad \bar{y} < z_\alpha \quad \Rightarrow \quad \text{accetto } \mathbb{H}_0 \\
\alpha \geq 1 - \Phi(\bar{y}) \quad &\Rightarrow \quad \bar{y} \geq z_\alpha \quad \Rightarrow \quad \text{rifiuto } \mathbb{H}_0
\end{aligned}
$$

La curva OC del test unilaterale sopra presentato:

(ricorda X_i sono i.i.d., $X_i \sim \mathcal{N}(\mu, \sigma_0^2) \Rightarrow \overline{X}_n \sim \mathcal{N}(\mu, \sigma_0^2/n)$.)

$$\begin{aligned}
\beta(\mu) :&= P_\mu(\text{accettare } \mathbb{H}_0) = P_\mu\Big(\frac{\overline{X}_n - \mu_0}{\sigma_0/\sqrt{n}} < z_\alpha\Big) \\
&= P_\mu\Big(\frac{\overline{X}_n - \mu}{\sigma_0/\sqrt{n}} + \frac{\mu - \mu_0}{\sigma_0/\sqrt{n}} < z_\alpha\Big) = P_\mu\Big(\frac{\overline{X}_n - \mu}{\sigma_0/\sqrt{n}} < z_\alpha + \frac{\mu_0 - \mu}{\sigma_0/\sqrt{n}}\Big) \\
&= \Phi\Big(z_\alpha + \frac{\mu_0 - \mu}{\sigma_0/\sqrt{n}}\Big).
\end{aligned}$$

N.B. Poichè Φ è una funzione crescente, $\beta(\mu)$ è una funzione decrescente di μ e $\beta(\mu_0) = \Phi(z_\alpha) = 1 - \alpha$.

La funzione di potenza $1 - \beta(\mu)$ è quindi crescente.

Analogamente per testare

$\mathbb{H}_0 : \mu \leq \mu_0$

contro l'ipotesi alternativa

$\mathbb{H}_1 : \mu > \mu_0$.

a livello di significatività α usiamo ancora il test che, avendo osservato $\overline{X}_n = \bar{x}_n$,

$$\text{rifiuta } \mathbb{H}_0 \text{ se } \frac{(\bar{x}_n - \mu_0)\sqrt{n}}{\sigma_0} \geq z_\alpha$$

$$\text{accetta } \mathbb{H}_0 \text{ se } \frac{(\bar{x}_n - \mu_0)\sqrt{n}}{\sigma_0} < z_\alpha.$$

$\dfrac{(\overline{X}_n - \mu_0)\sqrt{n}}{\sigma_0}$ statistica test $\Longrightarrow$ rifiuto $\mathbb{H}_0$ se il valore osservato della statistica test è maggiore o uguale al quantile z_α.

In questo caso l'ipotesi nulla $\mathbb{H}_0$ non è semplice. Mostriamo che il livello di significatività è ancora α. Si tratta di mostrare che la probabilità dell'errore di $\mathbb{I}$ specie è $\leq \alpha$. Infatti:

$$\begin{aligned}\sup_{\mu \leq \mu_0} \mathbb{P}_\mu(\text{rifiutare } \mathbb{H}_0) &= \sup_{\mu \leq \mu_0} \mathbb{P}_\mu\Big(\frac{\overline{X}_n - \mu_0}{\sigma_0/\sqrt{n}} \geq z_\alpha\Big) \\ &= \sup_{\mu \leq \mu_0} (1 - \beta(\mu)) = 1 - \beta(\mu_0) = \alpha\end{aligned}$$

poichè, come abbiamo visto, $\beta(\mu)$ è decrescente e $\beta(\mu_0) = 1 - \alpha$. (N.B. $\mu \leq \mu_0 \Leftrightarrow \mathbb{H}_0$ è vera)

Con analoghi ragionamenti possiamo dire che un test di livello α per

$$\mathbb{H}_0 : \mu = \mu_0 \quad (o\ \mu \geq \mu_0)$$

contro l'ipotesi alternativa

$$\mathbb{H}_1 : \mu < \mu_0.$$

è quello che, avendo osservato $\overline{X}_n = \bar{x}_n$,

$$\text{rifiuta } \mathbb{H}_0 \text{ se } \frac{(\bar{x}_n - \mu_0)\sqrt{n}}{\sigma_0} \leq -z_\alpha$$

$$\text{accetta } \mathbb{H}_0 \text{ se } \frac{(\bar{x}_n - \mu_0)\sqrt{n}}{\sigma_0} > -z_\alpha.$$

$\dfrac{(\overline{X}_n - \mu_0)\sqrt{n}}{\sigma_0}$ statistica test $\Longrightarrow$ rifiuto $\mathbb{H}_0$ se il valore osservato della statistica test è minore o uguale a $-z_{\alpha/2}$.

Esercizio. Mostrare che è un test di livello α per il problema di verifica d'ipotesi

$\mathbb{H}_0 : \mu = \mu_0 \;\; (o \; \mu \geq \mu_0)$

contro l'ipotesi alternativa

$\mathbb{H}_1 : \mu < \mu_0$.

Mostrare che il $p-value$ per questo test, in corrispondenza dei dati a disposizione, è

$$p-value = \Phi\left(\frac{(\bar{x}_n - \mu_0)\sqrt{n}}{\sigma_0}\right)$$

dove $\bar{x}_n$ è il valore osservato della media campionaria. La funzione di potenza del test è

$$\pi(\mu) = \Phi\left(\frac{(\mu_0 - \mu)\sqrt{n}}{\sigma_0} - z_\alpha\right) \qquad \text{per} \;\; \mu < \mu_0$$

Verifica d'ipotesi per la media di una popolazione gaussiana quando la varianza è incognita: test t

Sia $X_1, \ldots, X_n$ campione estratto da una popolazione $\mathcal{N}(\mu, \sigma^2)$ con (μ, σ^2) incognito.

1. Test bilatero.

Fissata una costante μ_0, vogliamo verificare l'ipotesi nulla

$\mathbb{H}_0 : \mu = \mu_0$
contro l'ipotesi alternativa

$\mathbb{H}_1 : \mu \neq \mu_0$.

N.B. In questo caso $\mathbb{H}_0$ non è un'ipotesi semplice. Infatti non specifica completamente la distribuzione del campione in quanto non fornisce il valore della varianza.

idea per costruire il test

Come in precedenza, sembra ragionevole rifiutare l'ipotesi nulla quando $\overline{X}_n$ cade lontano da μ_0. Tuttavia la distanza a cui deve essere da μ_0 per giustificare un rifiuto dipende dalla deviazione standard σ che in questo caso non è nota. Quindi stimiamo σ con un suo stimatore, la deviazione standard campionaria

$$S_n := \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2}$$

e rifiutiamo l'ipotesi nulla quando

$$\left|\frac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}}\right|$$

è "troppo grande". Questa quantità, come prima, viene fissata in base al livello di significatività.

Se α è il livello di significatività dobbiamo fissare c in modo tale che

$$\sup_{(\mu,\sigma^2)\in\mathbb{H}_0} \mathbb{P}_{(\mu,\sigma^2)}\left(\left|\frac{\overline{X}_n-\mu_0}{S_n/\sqrt{n}}\right|\geq c\right)=\sup_{\sigma^2>0}\mathbb{P}_{\mu_0,\sigma^2}\left(\left|\frac{\overline{X}_n-\mu_0}{S_n/\sqrt{n}}\right|\geq c\right)=\alpha$$

Ma quando $\mathbb{H}_0$ è vera, cioè $\mu=\mu_0$ e σ^2 qualsiasi, la statistica

$$\frac{\overline{X}_n-\mu_0}{S_n/\sqrt{n}} \sim t(n-1)$$

quindi

$$\begin{aligned}\alpha =&\mathbb{P}_{(\mu_0,\sigma^2)}\left(\left|\frac{\overline{X}_n-\mu_0}{S_n/\sqrt{n}}\right|\geq c\right)\\ =&\mathbb{P}(|T_{n-1}|\geq c)=2\mathbb{P}(T_{n-1}\geq c)\end{aligned}$$

dove $T_{n-1}\sim t(n-1)$ e quindi $c=t_{\alpha/2,n-1}$.

Quindi la regola del test è:

rifiuta $\mathbb{H}_0$ se il valore osservato di $\left|\frac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}}\right| \geq t_{\alpha/2,n-1}$

accetta $\mathbb{H}_0$ se il valore osservato di $\left|\frac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}}\right| < t_{\alpha/2,n-1}$.

(t-test bilatero di livello α).

Quindi la statistica test è $\frac{(\overline{X}_n - \mu_0)\sqrt{n}}{S_n}$ e rifiuto $\mathbb{H}_0$ se il modulo del valore osservato della statistica test è maggiore o uguale al quantile $t_{\alpha/2,n-1}$.

Sia $\bar{t}$ il valore assunto dalla statistica test $\dfrac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}}$ in corrispondenza dei dati (cioè in corrispondenza delle osservazioni campionarie). Allora il $p-value$ è

$$p-value = \mathbb{P}_{\mathbb{H}_0}\left(\left|\frac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}}\right| \geq |\bar{t}|\right) = \mathbb{P}_{(\mu_0,\sigma^2)}\left(\left|\frac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}}\right| \geq |\bar{t}|\right).$$

e $\dfrac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}} \sim t(n-1)$ sotto $\mathbb{H}_0$.

Quindi si rifiuta $\mathbb{H}_0$ a tutti i livelli di significatività $\alpha \geq p-value$ altrimenti si accetta.

$X_1, \ldots, X_n$ campione da $\mathcal{N}(\mu, \sigma^2)$, (μ, σ^2) incognito

2. Test unilatero: t-test unilaterale di livello α.

Fissata una costante μ_0, vogliamo verificare l'ipotesi nulla

$\mathbb{H}_0 : \mu = \mu_0$ (oppure $\mathbb{H}_0 : \mu \leq \mu_0$)

contro l'ipotesi alternativa

$\mathbb{H}_1 : \mu > \mu_0$.

Ad un livello di significatività α:

si rifiuta $\mathbb{H}_0$ se il valore osservato di $\dfrac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}} \geq t_{\alpha, n-1}$

si accetta $\mathbb{H}_0$ se il valore osservato di $\dfrac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}} < t_{\alpha, n-1}$.

Sia $\bar{t}$ il valore assunto dalla statistica test $T := \frac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}}$ in corrispondenza dei dati, allora

$$p\!-\!value = P_{(\mu_0, \sigma^2)}(T \geq \bar{t}) = P_{(\mu_0, \sigma^2)}\left(\frac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}} \geq \bar{t}\right).$$

Ricorda che, se $\mu = \mu_0$. allora $\frac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}} \sim t(n-1)$ per ogni σ^2.

Analogamente, vogliamo verificare l'ipotesi nulla

$\mathbb{H}_0 : \mu = \mu_0$ (oppure $\mathbb{H}_0 : \mu \geq \mu_0$)

contro l'ipotesi alternativa

$\mathbb{H}_1 : \mu < \mu_0$.

Ad un livello di significatività α:

si rifiuta $\mathbb{H}_0$ se il valore osservato di $\dfrac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}} \leq -t_{\alpha,n-1}$

si accetta $\mathbb{H}_0$ se il valore osservato di $\dfrac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}} > -t_{\alpha,n-1}$.

(t-test unilaterale di livello α)

Sia $\bar{t}$ il valore assunto dalla statistica test $\frac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}}$ in corrispondenza dei dati, allora il $p-value$ è

$$p-value = \mathbb{P}_{(\mu_0,\sigma^2)}\left(T \leq \bar{t}\right) = \mathbb{P}_{(\mu_0,\sigma^2)}\left(\frac{\overline{X}_n - \mu_0}{S_n/\sqrt{n}} \leq \bar{t}\right).$$

Verifica d'ipotesi per la varianza di una popolazione gaussiana

Sia $X_1, \dots, X_n$ campione estratto da una popolazione $\mathcal{N}(\mu, \sigma^2)$ con (μ, σ^2) incognito. Fissata una costante positiva σ_0^2, vogliamo

verificare l'ipotesi nulla

$\mathbb{H}_0 : \sigma^2 = \sigma_0^2$

contro l'ipotesi alternativa

$\mathbb{H}_1 : \sigma^2 \neq \sigma_0^2$.

Idea per costruire il test.

Ricordiamo che, se $\sigma^2 = \sigma_0^2$ e per qualunque valore della media μ, (cioè se $\mathbb{H}_0$ è vera),

$$\frac{(n-1)S_n^2}{\sigma_0^2} \sim \chi^2(n-1).$$

Quindi

$$\sup_{\mu} \mathbb{P}_{(\mu,\sigma_0^2)}\left(\chi^2_{1-\alpha/2,n-1} < \frac{(n-1)S_n^2}{\sigma_0^2} < \chi^2_{\alpha/2,n-1}\right) = 1-\alpha$$

o equivalentemente

$$\sup_{\mu} \mathbb{P}_{(\mu,\sigma_0^2)}\left(\chi^2_{1-\alpha/2,n-1} \geq \frac{(n-1)S_n^2}{\sigma_0^2} \cup \frac{(n-1)S_n^2}{\sigma_0^2} \geq \chi^2_{\alpha/2,n-1}\right) = \alpha$$

Perciò si usa la statistica test $\frac{(n-1)S_n^2}{\sigma_0^2}$ e un test di livello di significatività α è quello che, avendo osservato $S_n^2 = s_n^2$,

si rifiuta $\mathbb{H}_0$ se $\frac{(n-1)s_n^2}{\sigma_0^2} \leq \chi^2_{1-\alpha/2,n-1}$ oppure

se $\frac{(n-1)s_n^2}{\sigma_0^2} \geq \chi^2_{\alpha/2,n-1}$

si accetta $\mathbb{H}_0$ se $\chi^2_{1-\alpha/2,n-1} < \frac{(n-1)s_n^2}{\sigma_0^2} < \chi^2_{\alpha/2,n-1}$.

Esercizio. Mostrare che, se $\bar{c}$ è il valore assunto dalla statistica test $\dfrac{(n-1)S_n^2}{\sigma_0^2}$ in corrispondenza dei dati, allora

$$p\!-\!value = 2\min\{\mathbb{P}(W \leq \bar{c}), \mathbb{P}(W > \bar{c}) = 1 - \mathbb{P}(W \leq \bar{c})\} := \tilde{p}$$

dove $W \sim \chi^2(n-1)$, cioè è una v.a. che ha la stessa distribuzione della statistica test $\dfrac{(n-1)S_n^2}{\sigma_0^2}$ sotto $\mathbb{H}_0$. Infatti:

$$\alpha < \tilde{p} \quad \Rightarrow \quad \frac{\alpha}{2} < \mathbb{P}(W > \bar{c}) \quad \text{e} \quad 1 - \frac{\alpha}{2} > \mathbb{P}(W > \bar{c}) \quad \Rightarrow$$

$$\chi^2_{1-\alpha/2,n-1} < \bar{c} < \chi^2_{\alpha/2,n-1} \quad \Rightarrow \quad \text{accetto}$$

$$\alpha \geq \tilde{p} \quad \Rightarrow \quad \frac{\alpha}{2} \geq \mathbb{P}(W > \bar{c}) \quad \text{o} \quad 1 - \frac{\alpha}{2} \leq \mathbb{P}(W > \bar{c}) \quad \Rightarrow$$

$$\bar{c} \leq \chi^2_{1-\alpha/2,n-1} \quad \text{o} \quad \bar{c} \geq \chi^2_{\alpha/2,n-1} \quad \Rightarrow \quad \text{rifiuto}$$