TEST DI IPOTESI

VARIANZA NOTA E IGNOTA

Dr. R. Lo Franco – Lezione 11

Riferimenti: [1] Sheldon M. Ross, *Introduzione alla statistica*, Apogeo Editore;
[2] Maria Garetto, *Statistica*, Università di Torino

8. *Test di ipotesi*

8.1 Introduzione

Come è già stato messo in evidenza, uno degli scopi più importanti di un'analisi statistica è quello di utilizzare dei dati provenienti da un campione per fare inferenza sulla popolazione da cui è stato tratto il campione.
Nel Cap. 7 si è visto ad esempio come, utilizzando la media campionaria, si può stimare il valore del corrispondente parametro della popolazione.
Ci sono altri problemi in cui si sottopone a test un'ipotesi su un parametro di una popolazione, con lo scopo di decidere, esaminando un campione tratto dalla popolazione, se l'affermazione riguardante il parametro è vera o falsa.
Ad esempio il responsabile della produzione in un'azienda può ipotizzare che le confezioni prodotte abbiano un peso medio di 250g; un medico può ipotizzare che un certo farmaco sia efficace nel 90% dei casi in cui viene usato. Con la verifica delle ipotesi si può determinare se tali congetture sono compatibili con i dati disponibili dal campione.

Definizioni 1
Un'ipotesi formulata in termini di parametri di una popolazione, come media o varianza, è detta **ipotesi statistica**.
Il procedimento che consente di rifiutare o accettare un'ipotesi statistica utilizzando i dati di un campione, viene chiamato **test di ipotesi**.

8.2 Ipotesi statistiche

Per illustrare i concetti generali riguardanti la verifica delle ipotesi, consideriamo i seguenti esempi.

Esempio 1
Si vuole sottoporre a test l'affermazione di un produttore di vernici secondo cui il tempo medio di asciugatura di una nuova vernice è non superiore a $\mu = 20$ minuti.
A questo scopo si prende un campione di 35 lattine di vernice, si effettuano 35 prove di verniciatura con la vernice delle diverse confezioni e si calcola il tempo medio di asciugatura, con l'intenzione di rifiutare l'affermazione del produttore se la media osservata supera il valore di 20 minuti, o di accettarla in caso contrario.

Esempio 2
Si vuole verificare se le lattine di caffè confezionate automaticamente da una ditta contengono in media il peso dichiarato $\mu = 250$ g. A tale scopo si estrae un campione di 30 lattine, se ne pesa il contenuto e si calcola il peso medio, per stabilire se il peso medio differisce da 250g.

La verifica delle ipotesi statistiche inizia con la definizione del problema in termini di ipotesi sul parametro oggetto di studio.
Per prima cosa si stabilisce l'ipotesi da sottoporre a test, detta **ipotesi nulla**, indicata con $\mathbf{H_0}$.
Oltre all'ipotesi nulla occorre specificare anche un'adeguata **ipotesi alternativa**, indicata con $\mathbf{H_1}$, ossia un'affermazione che contraddice l'ipotesi nulla.
Nell'esempio 1 l'ipotesi nulla e l'ipotesi alternativa sono

$$H_0: \quad \mu \leq 20 \text{ minuti}$$
$$H_1: \quad \mu > 20 \text{ minuti.}$$

Nell'esempio 2 l'ipotesi nulla e l'ipotesi alternativa sono

$$H_0: \quad \mu = 250 \text{ g}$$
$$H_1: \quad \mu \neq 250 \text{ g.}$$

In generale attualmente il termine “ipotesi nulla” viene usato per ogni ipotesi sottoposta a test con lo scopo di decidere se deve essere rifiutata in favore dell’ipotesi alternativa.

Regole per la scelta delle ipotesi

Un problema importante nel predisporre un test di ipotesi è la scelta delle ipotesi: come si può decidere qual è l’ipotesi nulla e quale deve essere l’ipotesi alternativa?

Non c’è purtroppo una risposta semplice alla domanda, perché la scelta dipende anche da fattori soggettivi: chi effettua il test ha in genere convinzioni e idee personali su quanto intende mostrare.

Tuttavia si possono indicare alcune linee guida; facciamo riferimento per comodità a un test sulla media di una singola popolazione, ma gli stessi principi si possono applicare a ogni test riguardante uno o più parametri.

Se ad esempio il test ha come scopo di decidere se il valore della media μ della popolazione è diverso dal valore 100 , l’ipotesi nulla e l’ipotesi alternativa saranno della forma

$$H_0: \quad \mu = 100$$
$$H_1: \quad \mu \neq 100$$

Se invece il test ha lo scopo di decidere se il valore di μ è minore di 100, allora l’ipotesi nulla e l’ipotesi alternativa saranno

$$H_0: \quad \mu \geq 100$$
$$H_1: \quad \mu < 100$$

Se infine il test ha lo scopo di decidere se il valore di μ è maggiore di 100, allora l’ipotesi nulla e l’ipotesi alternativa saranno

$$H_0: \quad \mu \leq 100$$
$$H_1: \quad \mu > 100$$

In conclusione:

1 – nell'ipotesi alternativa viene messo ciò che si spera o ci si aspetta di poter concludere come risultato del test;

2 – l'ipotesi nulla è posta con lo scopo di essere screditata, quindi ciò che si oppone alla conclusione che il ricercatore cerca di raggiungere rappresenta l'ipotesi nulla;

3 – nell'ipotesi nulla deve sempre comparire un segno di uguaglianza ($=$, $\leq$ o $\geq$);

4 – le due ipotesi sono complementari, ossia considerate insieme esauriscono tutte le possibilità riguardanti il valore che può assumere il parametro in esame.

Esempio 3

Si supponga di voler stabilire se possiamo concludere che il tempo medio richiesto per svolgere una certa operazione è minore di 30 minuti. In tal caso si scelgono le ipotesi

$$H_0: \quad \mu \geq 30 \text{ minuti}$$
$$H_1: \quad \mu < 30 \text{ minuti.}$$

Esempio 4

Il contenuto dichiarato dal produttore delle bottiglie di acqua minerale di una certa marca è 920ml. Un'associazione di consumatori sostiene che in realtà le bottiglie contengono in media una quantità inferiore di acqua. In questo caso le ipotesi sono

$$H_0: \quad \mu \geq 920 \text{ ml}$$
$$H_1: \quad \mu < 920 \text{ ml.}$$

Esempio 5

Un ingegnere suggerisce alcune modifiche che si potrebbero apportare a una linea produttiva per aumentare il numero di pezzi prodotti giornalmente.

Per decidere se applicare queste modifiche occorre che i dati sperimentali indichino con forte evidenza che la macchina modificata è più produttiva di quella originaria.

Se μ_0 è il numero medio di pezzi prodotti prima della modifica, si scelgono le ipotesi

$$H_0: \quad \mu \leq \mu_0$$
$$H_1: \quad \mu > \mu_0.$$

Osservazione

E' importante sottolineare che con la verifica delle ipotesi, e in generale con l'inferenza statistica, non si arriva alla dimostrazione di un'ipotesi; si ha solo un'indicazione del fatto che l'ipotesi sia o meno avvalorata dai dati disponibili: quando non si rifiuta un'ipotesi nulla, non si dice che essa è vera, ma che può essere vera; in altre parole se non rifiutiamo l'ipotesi nulla, possiamo solo concludere che il campione non fornisce prove sufficienti a garantirne il rifiuto, ma ciò non implica alcuna dimostrazione.

Riassumendo, le **possibili conclusioni per un test di ipotesi** sono:

1 – se l'ipotesi nulla H_0 è rifiutata, si conclude che l'ipotesi alternativa H_1 è probabilmente vera;

2 – se l'ipotesi nulla non è rifiutata si conclude che i dati non forniscono una sufficiente evidenza per sostenere l'ipotesi alternativa.

8.3 Tipi di errore e livello di significatività

Dopo aver formulato le ipotesi, occorre specificare quale risultato del campione porterà al rifiuto dell'ipotesi nulla.
Ricordiamo che le statistiche campionarie media e varianza sono stimatori corretti del corrispondente parametro della popolazione. Poiché il valore della statistica è calcolato da un campione, anche se l'ipotesi nulla è vera, è però molto probabile che la statistica differisca di una certa quantità dal valore del parametro della popolazione, per effetto del caso; ciò nonostante, se l'ipotesi nulla è vera, ci aspettiamo che la statistica campionaria sia vicina al parametro della popolazione.

Tuttavia il processo decisionale non può certo essere basato sui termini "abbastanza vicino" o "troppo grande" usati negli esempi; la teoria della verifica dei test di ipotesi si basa sullo studio della distribuzione campionaria di una statistica, detta statistica test.
La **statistica test** è una statistica che viene calcolata dai dati del campione e può assumere tanti valori quanti sono i possibili campioni estraibili dalla popolazione, quindi il particolare valore calcolato dipende dal campione estratto.
Un esempio di statistica test è la quantità

$$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

La distribuzione di campionamento della statistica test è, di solito, una distribuzione nota, come la distribuzione normale o la distribuzione *t*, e ricorriamo a queste distribuzioni per sottoporre a verifica un'ipotesi nulla; ad esempio la statistica test *Z* ha la distribuzione normale standardizzata.
Utilizzando le proprietà della distribuzione di campionamento della statistica soggetta a test, si può identificare un intervallo di valori di quella statistica che verosimilmente non si presentano se l'ipotesi nulla è vera.
La distribuzione di campionamento della statistica test è divisa in due regioni, una **regione di rifiuto** e una **regione di accettazione**, delimitate da uno o più valori, detti **valori critici**.

Definizioni 2
La **regione di rifiuto** corrisponde all'insieme dei valori di una statistica test che conducono al rifiuto dell'ipotesi nulla.
L'insieme dei valori che portano invece all'accettazione dell'ipotesi nulla si chiama **regione di accettazione**.
I **valori critici** sono i valori della statistica test che separano le regioni di rifiuto e di accettazione.

Se la statistica test, in base ai dati del campione, assume un valore che cade nella regione di rifiuto, l'ipotesi nulla deve essere rifiutata; se al contrario il valore cade nella regione di accettazione, l'ipotesi nulla non può essere rifiutata.
La regione di rifiuto può essere vista come l'insieme dei valori della statistica test che non è probabile che si verifichino quando l'ipotesi nulla è vera, mentre è probabile che si verifichino quando l'ipotesi nulla è falsa. Pertanto, se il campione porta a un valore della statistica test che cade nella regione di rifiuto, rifiutiamo l'ipotesi nulla perché non è probabile che sia vera.

I test di ipotesi possono essere classificati in due gruppi: **test a una coda** (o **test unilaterale**) e **test a due code** (o **test bilaterale**).
Quando la regione di rifiuto è costituita da un intervallo, il **test** si dice **a una coda**: questo caso è illustrato dalle figure 3 e 4, pag. 226-227; quando invece la regione di rifiuto è costituita da due intervalli, ossia da due code della distribuzione, il **test** si dice **a due code**: questo caso è illustrato dalla figura 5, pag. 227.
Un semplice modo per stabilire di che tipo è un test, senza dover conoscere la/le regioni di rifiuto è il seguente: per un test a due code nell'ipotesi alternativa compare il segno $\neq$, mentre per un test a una coda compare uno dei segni $>$ oppure $<$.

Quando si usa una statistica campionaria per prendere una decisione sul parametro della popolazione si corre sempre il rischio di giungere a una conclusione sbagliata. Questo dipende dal fatto che un'informazione parziale, ottenuta da un campione, è usata per trarre conclusioni sull'intera popolazione. Nella verifica di ipotesi si individuano due **tipi di errore**.
Per illustrare questo problema riprendiamo in esame l'esempio 1. Supponiamo di aver scelto la regione di accettazione, stabilendo di accettare l'ipotesi nulla se la media del campione non supera i 20.50 minuti.
C'è una prima possibilità che la media del campione superi i 20.50 minuti stabiliti, mentre la media effettiva della popolazione è $\mu = 20$ minuti; c'è anche una seconda possibilità che la media del campione possa essere minore o uguale a 20.50 minuti, ma la media effettiva non sia $\mu = 20$ minuti, ma sia ad esempio $\mu = 21$ minuti.
La situazione appena descritta in questo esempio è tipica dei test di ipotesi: anche se si fa il test in modo corretto, si possono commettere questi due tipi di errore, che possono portare a conseguenze dannose.

Definizioni 3
Se l'ipotesi H_0 è vera, ma viene erroneamente rifiutata, si commette un **errore del I tipo**; la probabilità di commettere tale errore è indicata con α.
Se l'ipotesi H_0 è falsa, ma erroneamente non viene rifiutata, si commette un **errore del II tipo**; la probabilità di commettere questo tipo di errore è indicata con β.

	H_0 vera	H_0 falsa
Rifiutiamo H_0	Errore del I tipo Probabilità (errore I tipo) = α	Decisione corretta
Accettiamo H_0	Decisione corretta	Errore del II tipo Probabilità (errore II tipo) = β

Un'analogia che può chiarire le idee precedenti è quella del processo a un imputato. In tribunale una persona sottoposta a processo viene ritenuta innocente fino a prova contraria. L'ipotesi nulla H_0 è quindi "l'imputato è innocente"; l'ipotesi alternativa H_1 è "l'imputato è colpevole".
L'errore del I tipo è condannare un innocente, l'errore del II tipo è assolvere un colpevole.
Riassumiamo questi concetti con lo schema seguente.

	Imputato innocente	Imputato colpevole
Imputato condannato	Errore del I tipo	Decisione corretta
Imputato assolto	Decisione corretta	Errore del II tipo

Scegliere come ipotesi nulla H_0 "l'imputato è innocente" significa ritenere che condannare un innocente sia un errore più grave che assolvere un colpevole.

In generale l'errore di I tipo è quello considerato più grave: questo significa che l'ipotesi nulla H_0 va formulata in modo che quello che si ritiene sia l'errore più grave coincida con l'errore di I tipo.
Servendoci ancora degli esempi 1 e 2, calcoliamo la probabilità α di commettere un errore del I tipo; usiamo a tale scopo le proprietà della distribuzione della media campionaria.

Esempio 1 – parte 2

Assumiamo che sia noto dall'esperienza che lo scarto quadratico medio del tempo di asciugatura della vernice è $\sigma = 2$ minuti e studiamo la probabilità di commettere un errore del I tipo, ossia la probabilità α che la media del campione superi 20.5 minuti, anche se la media effettiva della popolazione è $\mu \leq 20$ minuti[2].

Come è noto dal Cap. 6, la distribuzione della media campionaria per grandi campioni ($n \geq 30$) è approssimativamente normale, quindi la probabilità suddetta è data dall'area della regione rappresentata nella figura 1.

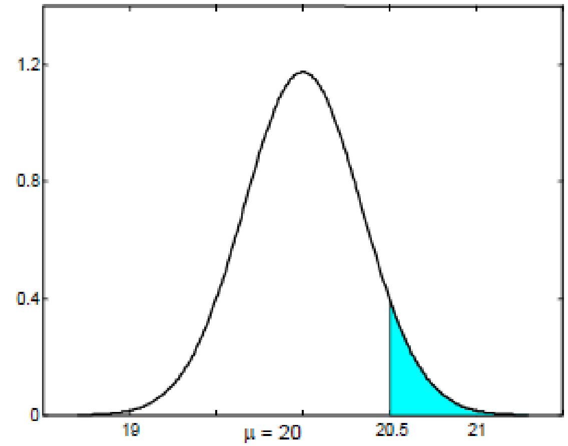

Figura 1

La regione a destra del valore 20.5 è la regione di rifiuto, quella a sinistra è la regione di accettazione: se il valore della media campionaria cade a destra di 20.5 l'ipotesi nulla viene rifiutata, altrimenti non viene rifiutata.

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{2}{\sqrt{35}} = 0.34$$

Standardizzando il valore $\overline{x} = 20.5$ si ha

$$Z = \frac{20.5 - 20}{0.34} = 1.47 .$$

Utilizzando le tavole della distribuzione normale, si trova che l'area della regione a destra di 20.5 è

$$P(Z > 1.47) = 1 - P(Z < 1.47) = 1 - 0.9292 = 0.0708$$

quindi la probabilità di rifiutare erroneamente l'ipotesi nulla è

$$\alpha = 0.0708$$

Esempio 2 – parte 2

Assumiamo che lo scarto quadratico medio della popolazione sia $\sigma = 15$g e studiamo la probabilità α che la media del campione non sia compresa fra 245g e 255g, anche se la media effettiva della popolazione è $\mu = 250$g [4].

La probabilità che si vuole calcolare è data dalla somma delle due aree rappresentate nella figura 2

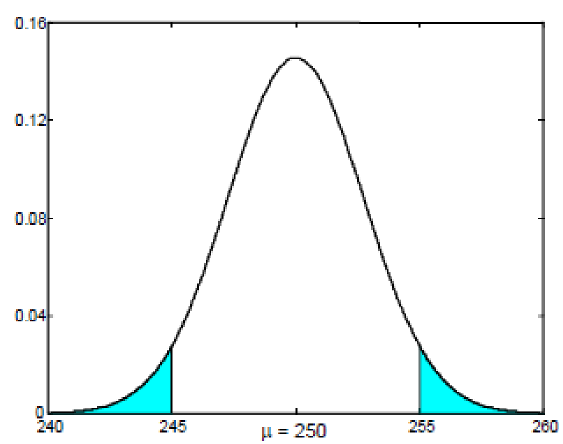

Figura 2

La regione di rifiuto in questo caso è costituita dai valori a sinistra di 245g e dai valori a destra di 255g; se il valore della media campionaria cade nell'intervallo (245, 255), che è la regione di accettazione, l'ipotesi nulla viene accettata, altrimenti viene rifiutata.
Seguendo il procedimento già descritto nell'esempio precedente si trova

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{30}} \cong 2.74\,.$$

La regione di accettazione è un intervallo simmetrico rispetto a $\mu = 250$; standardizzando il valore $\overline{x} = 255$ si ha

$$Z = \frac{255 - 250}{2.74} = 1.82\,.$$

Utilizzando le tavole della distribuzione normale, si trova che l'area della regione colorata è

$$P(|Z| > 1.82) = 2 \cdot \left[1 - P(Z < 1.82)\right] = 2 \cdot (1 - 0.9656) = 0.0688$$

quindi la probabilità di rifiutare erroneamente l'ipotesi nulla è $\alpha = 0.0688$.

Definizione 4
La probabilità α di commettere un errore del I tipo, ossia di rifiutare un'ipotesi nulla vera, è detta **livello di significatività**.

Negli esempi 1 e 2 (parte 2) si è mostrato come calcolare la probabilità α di commettere un errore del I tipo, per regioni di rifiuto scelte arbitrariamente; questo non è però il procedimento seguito di solito nelle applicazioni.
Il metodo usato più frequentemente consiste invece nello specificare un valore per il livello di significatività α e poi identificare la regione di rifiuto che soddisfa tale valore.
Poiché l'errore di I tipo è quello considerato più grave, si scelgono per α valori piccoli; i valori più usati sono $\alpha = 0.01$ e $\alpha = 0.05$.
In corrispondenza al livello di significatività α, il valore $(1-\alpha)\cdot 100\%$ coincide con il **grado di fiducia** già introdotto per gli intervalli di confidenza.
Se si sceglie ad esempio un livello di significatività $\alpha = 0.05$, ossia del 5%, ci sarà una probabilità del 5% di rifiutare un'ipotesi che non avrebbe dovuto essere rifiutata; in altre parole siamo fiduciosi al 95% di aver preso la decisione giusta.

Definizione 5
La probabilità di commettere un errore del II tipo, indicata con β, viene anche chiamata **rischio del consumatore**.

Schema riassuntivo – Test di ipotesi

1 – Si scelgono l'ipotesi nulla e l'ipotesi alternativa.

2 – Si sceglie il livello di significatività α a cui si vuole eseguire il test.

3 – In funzione del tipo di test (una coda o due code) e del valore di α scelto, si determinano i valori critici e la regione di rifiuto.

4 – Si sceglie l'ampiezza campionaria, si raccoglie il campione, si calcola dai dati del campione il valore della statistica test e si vede se appartiene o no alla regione di rifiuto.

5 – Si prende la decisione: rifiutare o non rifiutare l'ipotesi nulla al livello di significatività stabilito.

E' opportuno sottolineare che, quando l'ipotesi nulla non è rifiutata, non si dovrebbe dire che tale ipotesi viene accettata, bensì che l'ipotesi nulla non viene rifiutata: questo perché è possibile che si commetta un errore del II tipo; poiché spesso la probabilità di commettere un errore del II tipo è abbastanza elevata, non ci si dovrebbe impegnare troppo dicendo che si accetta l'ipotesi nulla.
Tuttavia, anche se impropriamente, spesso si usa il termine "si accetta l'ipotesi nulla".

8.4 Test di ipotesi sulla media (varianza nota)

Descriviamo il procedimento per eseguire un test di ipotesi sulla media di una popolazione avente varianza σ^2 nota.
Il test si basa sulla statistica

$$Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$$

dove n è l'ampiezza del campione e μ_0 è il valore della media assunto nell'ipotesi nulla
Il test qui illustrato è essenzialmente un **test per grandi campioni** ($n \geq 30$); in tal caso la distribuzione della media campionaria può essere approssimata dalla distribuzione normale e la variabile aleatoria Z ha approssimativamente la distribuzione normale standardizzata.
Nel caso particolare in cui il campione è estratto da una popolazione con distribuzione normale, la variabile Z ha distribuzione normale standardizzata, qualunque sia l'ampiezza del campione (vedere esempi 13 e 14).
Sia, come al solito, z_α il valore di Z per cui l'area a destra di z_α al di sotto della curva normale standardizzata è uguale a α.
Nelle figure seguenti si illustrano le regioni di rifiuto per un dato livello di significatività α, a seconda delle ipotesi nulla e alternativa stabilite.
Nei primi due casi si fa un test a una coda, nel terzo caso un test a due code.

1° caso – Test a una coda (figura 3)

Ipotesi nulla H_0: $\mu \leq \mu_0$.

Ipotesi alternativa H_1: $\mu > \mu_0$.

Regione di rifiuto[5] $Z > z_\alpha$

Regione di accettazione $Z < z_\alpha$

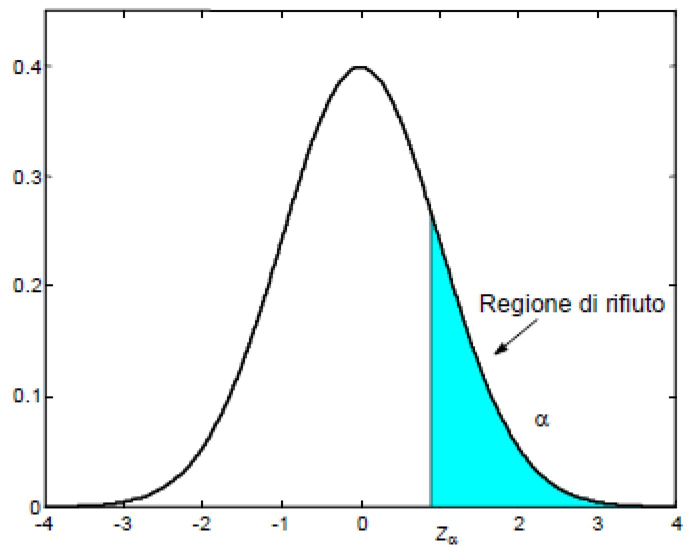

Figura 3

2° caso – Test a una coda (figura 4)

Ipotesi nulla H$_0$: $\mu \geq \mu_0$.

Ipotesi alternativa H$_1$: $\mu < \mu_0$.

Regione di rifiuto $Z < -z_\alpha$

Regione di accettazione $Z > -z_\alpha$

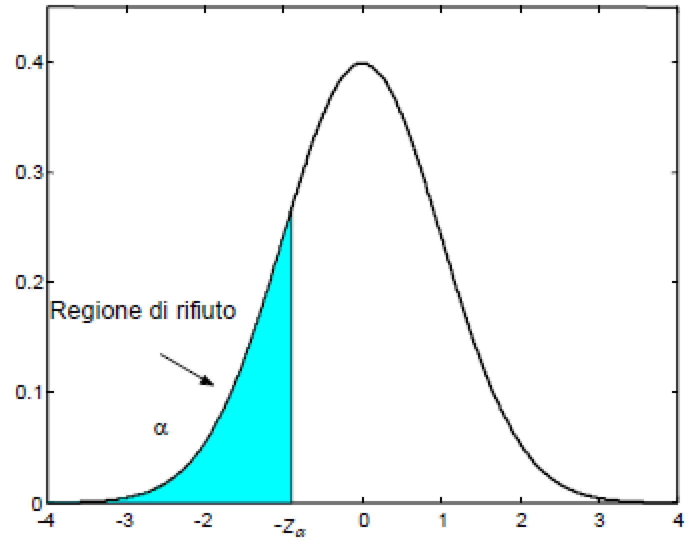

Figura 4

3° caso – Test a due code (figura 5)

Ipotesi nulla H_0: $\mu = \mu_0$.

Ipotesi alternativa H_1: $\mu \neq \mu_0$.

Regione di rifiuto $Z < -z_{\frac{\alpha}{2}}$ oppure $Z > z_{\frac{\alpha}{2}}$

Regione di accettazione $-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}$

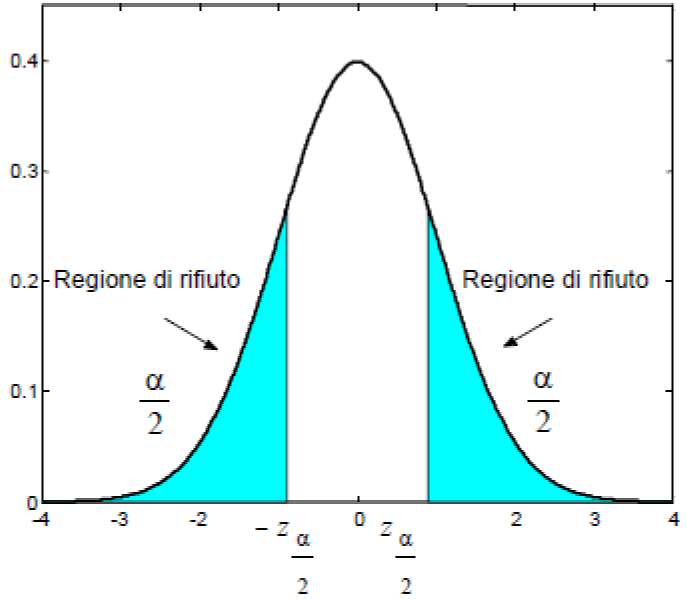

Figura 5

I valori z_α e $z_{\frac{\alpha}{2}}$ sono i valori critici del test nei tre casi; tali valori possono essere letti sulla tavola dei percentili della distribuzione normale standardizzata.

Nella tabella 1 riassumiamo i valori comunemente usati per il livello di significatività α e i corrispondenti valori critici z_α e $z_{\frac{\alpha}{2}}$ per i test a una e a due code.

Test	Ipot. nulla H_0	Ipot. altern. H_1	Liv. signif. α	Valori critici	Reg. rifiuto
una coda	$\mu \leq \mu_0$	$\mu > \mu_0$	0.01	2.326	$Z > 2.326$
			0.05	1.645	$Z > 1.645$
una coda	$\mu \geq \mu_0$	$\mu < \mu_0$	0.01	–2.326	$Z < -2.326$
			0.05	–1.645	$Z < -1.645$
due code	$\mu = \mu_0$	$\mu \neq \mu_0$	0.01	–2.576 e 2.576	$Z < -2.576$ $Z > 2.576$
			0.05	–1.96 e 1.96	$Z < -1.96$ $Z > 1.96$

Tabella 1

Esempio 6

Una ditta produttrice di lampadine sostiene che la durata media delle lampadine prodotte è di 1600 ore, con uno scarto quadratico medio $\sigma = 120$ ore.

Estraendo un campione di 100 lampadine si è calcolata una durata media di 1570 ore.

Stabilire se l'affermazione del produttore è corretta, usando come ipotesi alternativa che la durata media sia

a – inferiore a quella dichiarata;

b – diversa da quella dichiarata.

Usare in entrambi i casi il livello di significatività $\alpha = 0.05$ e il livello di significatività $\alpha = 0.01$.

a –

Ipotesi nulla	H_0:	$\mu \geq 1600$
Ipotesi alternativa	H_1:	$\mu < 1600$
Livello di significatività	$\alpha = 0.05$	

Il test è a una coda; il valore critico per questo livello di significatività è $z_\alpha = -1.645$.

La regola di decisione consiste nel rifiutare l'ipotesi se il valore della statistica Z ottenuto dai dati del campione è minore di -1.645.

Il campione ha le seguenti caratteristiche

$$n = 100 \qquad \overline{x} = 1570$$

Il valore della statistica test è

$$Z = \frac{1570 - 1600}{\frac{120}{\sqrt{100}}} = -2.50\,.$$

Dato che il valore trovato $Z = -2.50$ è minore del valore critico $z_\alpha = -1.645$, si rifiuta l'ipotesi nulla al livello di significatività $\alpha = 0.05$, ossia del 5%.

Livello di significatività $\alpha = 0.01$

Il test è a una coda; il valore critico per questo livello di significatività è $z_\alpha = -2.326$.

Anche in questo caso il valore $Z = -2.50$ è minore del valore critico $z_\alpha = -2.326$, perciò si rifiuta l'ipotesi nulla al livello di significatività $\alpha = 0.01$, ossia dell'1%.

b –

Ipotesi nulla H_0: $\mu = 1600$

Ipotesi alternativa H_1: $\mu \neq 1600$

Livello di significatività $\alpha = 0.05$

Il test è a due code; i valori critici per questo livello di significatività sono

$$z_{\frac{\alpha}{2}} = -1.96 \text{ e } z_{\frac{\alpha}{2}} = 1.96.$$

Il valore $Z = -2.50$ cade al di fuori dell'intervallo avente come estremi i valori critici, cioè appartiene alla regione di rifiuto, perciò si rifiuta l'ipotesi nulla al livello di significatività $\alpha = 0.05$, ossia del 5%.

Livello di significatività $\alpha = 0.01$.

I valori critici per questo livello di significatività sono $z_{\frac{\alpha}{2}} = -2.576$ e $z_{\frac{\alpha}{2}} = 2.576$.

Il valore $Z = -2.50$ cade fra questi estremi, perciò non si rifiuta l'ipotesi nulla al livello di significatività $\alpha = 0.01$, ossia dell'1%.

Esempio 7

La lunghezza della corda contenuta nei rotoli prodotti da una macchina ha una distribuzione avente varianza $\sigma^2 = 27.4\,\mathrm{m}^2$. La ditta produttrice afferma che la lunghezza media è $\mu = 300$m.

Viene prelevato un campione di 100 rotoli e calcolata la lunghezza media, pari a $\bar{x} = 299.2$.

Verificare se il produttore afferma il vero, oppure se la lunghezza è inferiore, al livello di significatività dell'1%.

Ipotesi nulla	H_0:	$\mu \geq 300$
Ipotesi alternativa	H_1:	$\mu < 300$
Livello di significatività	$\alpha = 0.01$.	

Il test è a una coda; il valore critico per questo livello di significatività è $z_\alpha = -2.326$. La regione di rifiuto è $z < -2.326$.

Si ha

$$n = 100 \qquad \bar{x} = 299.2$$

$$\sigma^2 = 27.4$$

Il valore della statistica test è

$$Z = \frac{299.2 - 300}{\frac{\sqrt{27.4}}{\sqrt{100}}} = -1.53.$$

Il valore $Z = -1.53$ non appartiene alla regione di rifiuto, quindi l'ipotesi nulla non viene rifiutata al livello di significatività dell'1%.

Esempio 8

La precisione di una macchina che produce componenti di dimensioni specificate viene controllata con periodiche verifiche a campione: la dimensione media richiesta è $\mu = 3.5$mm, con una varianza $\sigma^2 = 0.22\ \text{mm}^2$.

a – Valutare se il processo è da ritenersi sotto controllo oppure no, quando la media riscontrata su un campione di 60 pezzi è $\bar{x} = 3.42$ mm.

b – Ripetere la valutazione nel caso che il campione sia di 150 pezzi, con media $\bar{x} = 3.41$ mm.

Si sceglie come ipotesi nulla di ritenere che il processo sia sotto controllo e non sia quindi necessario alcun intervento

$$H_0: \qquad \mu = 3.5$$

L'ipotesi alternativa è che il processo sia fuori controllo

$$H_1: \qquad \mu \neq 3.5$$

e in questo caso occorre attuare qualche intervento per riportarlo sotto controllo. Si effettua quindi un test a due code.

Se il processo è sotto controllo, cioè H_0 è vera, ma erroneamente lo riteniamo fuori controllo, cioè rifiutiamo H_0, commettiamo un errore del I tipo; la probabilità di compiere tale errore è pari al livello di significatività α.

a – $\qquad n = 60 \qquad \bar{x} = 3.42 \qquad \sigma^2 = 0.22$

Il valore della statistica test è

$$Z = \frac{3.42 - 3.5}{\frac{\sqrt{0.22}}{\sqrt{60}}} = -1.32\,.$$

Livello di significatività $\alpha = 0.05$
Regione di rifiuto $Z < -1.96$ e $Z > 1.96$
Il valore $Z = -1.32$ non appartiene alla regione di rifiuto, quindi l'ipotesi nulla non viene rifiutata al livello di significatività del 5%; il processo si ritiene sotto controllo.

Livello di significatività $\alpha = 0.01$
Regione di rifiuto $Z < -2.576$ e $Z > 2.576$
Il valore $Z = -1.32$ non appartiene alla regione di rifiuto, quindi l'ipotesi nulla non viene rifiutata al livello di significatività dell'1%; anche in questo caso il processo si ritiene sotto controllo.

b – $n = 150$ $\bar{x} = 3.42$ $\sigma^2 = 0.2209$
Il valore della statistica test è

$$Z = \frac{3.41 - 3.5}{\frac{\sqrt{0.22}}{\sqrt{150}}} = -2.35 .$$

Livello di significatività $\alpha = 0.05$
Regione di rifiuto $Z < -1.96$ e $Z > 1.96$
Il valore $Z = -2.35$ appartiene alla regione di rifiuto, quindi l'ipotesi nulla viene rifiutata al livello di significatività del 5%; il processo si ritiene fuori controllo e si devono intraprendere delle modifiche al processo produttivo.

Livello di significatività $\alpha = 0.01$
Regione di rifiuto $Z < -2.576$ e $Z > 2.576$
Il valore $Z = -2.35$ appartiene alla regione di accettazione, quindi l'ipotesi nulla viene accettata al livello di significatività dell'1%; il processo si ritiene sotto controllo e non si intraprendono modifiche al processo produttivo.
Il rischio più basso per il consumatore si ha nel caso in cui $n = 150$ e $\alpha = 0.05$. Il punto di vista del produttore è ovviamente diverso.

Esempio 9

I carichi di rottura dei cavi prodotti da un'azienda hanno una media pari a 1800kg e uno scarto quadratico medio σ =100kg. Si afferma che mediante una nuova tecnica di costruzione il carico di rottura può essere reso maggiore. Per sottoporre a test questa affermazione si provano 50 cavi e si trova che il carico di rottura medio è di 1850kg.
E' possibile accettare l'affermazione ad un livello di significatività dell'1%?

Si assume come ipotesi nulla che non ci sia nessun cambiamento

$$H_0: \quad \mu \leq 1800$$

e come ipotesi alternativa che ci sia un aumento nel carico di rottura, ossia

$$H_1: \quad \mu > 1800.$$

Si effettua un test ad una coda; per il livello di significatività α = 0.01 il valore critico è $z_\alpha = 2.326$ e la regione di rifiuto è costituita dai valori $Z > 2.326$.
Il valore della statistica test è

$$Z = \frac{1850 - 1800}{\frac{100}{\sqrt{50}}} = 3.54 .$$

Dato che il valore trovato Z = 3.54 è maggiore del valore critico $z_\alpha = 2.326$, appartiene alla regione di rifiuto, perciò l'ipotesi nulla può essere rifiutata al livello di significatività $\alpha = 0.01$.

Definizione 6
In un test di ipotesi, dopo aver effettuato il campionamento e aver calcolato il valore della statistica test necessario per eseguire il test, si dice ***p*-value** il più piccolo valore del livello di significatività α per cui i dati campionari consentono di rifiutare l'ipotesi nulla.

Un *p*-value quasi uguale a zero significa che siamo praticamente certi di non sbagliare rifiutando l'ipotesi nulla; un *p*-value dell'ordine dei soliti livelli di significatività indica che la decisione se rifiutare o no l'ipotesi nulla è critica e dipende in modo cruciale dalla scelta del livello di significatività; un *p*-value maggiore indica invece che, a qualsiasi livello ragionevole di significatività, sbagliamo rifiutando l'ipotesi nulla; in questo caso si può anche dire che il test ci porta ad accettare l'ipotesi.

Una regola generale utile da ricordare è la seguente: se il *p*-value è minore o uguale ad α, rifiutiamo l'ipotesi nulla; se il *p*-value è maggiore di α, non rifiutiamo l'ipotesi nulla.

$$p-\text{value}=\begin{cases}1-P(Z<Z_0) & \text{per il test a una coda con} \quad \mathrm{H}_0:\mu\le\mu_0 \quad \mathrm{H}_1:\mu>\mu_0\\ P(Z<Z_0) & \text{per il test a una coda con} \quad \mathrm{H}_0:\mu\ge\mu_0 \quad \mathrm{H}_1:\mu<\mu_0\\ 2\left[1-P(Z<|Z_0|)\right] & \text{per il test a due code con} \quad \mathrm{H}_0:\mu=\mu_0 \quad \mathrm{H}_1:\mu\ne\mu_0\end{cases}$$

Esempio 15

Riprendiamo in esame i risultati ottenuti nell'esempio 8.

Le conclusioni tratte nel caso b sono piuttosto critiche e questo viene evidenziato dal p-value; si ha infatti

$$Z_0 = -2.35$$

$$p - \text{value} = 2[1 - P(Z < 2.35)] = 2(1 - 0.9906) = 0.0188$$

Il livello minimo che consente di rifiutare l'ipotesi nulla è del 1.88%.

Nel caso a invece le conclusioni non sono critiche; si ha infatti

$$Z_0 = -1.32$$

$$p - \text{value} = 2[1 - P(Z < 1.32)] = 2(1 - 0.9066) = 0.1868$$

In questo caso a ogni ragionevole livello di significatività possiamo accettare l'ipotesi nulla.

8.5 Test di ipotesi sulla media (varianza incognita)

Esaminiamo ora il caso in cui il campione usato per effettuare il test proviene da una popolazione di cui non è nota la varianza σ^2.
Come già osservato nel paragrafo precedente, se σ non è noto, ma il campione è grande, si può sostituire σ con il valore s dello scarto quadratico medio del campione.
Se invece il campione è piccolo, e la popolazione da cui proviene il campione ha distribuzione normale, si può usare il teorema 3, Cap. 6; sulla base di tale teorema la statistica test

$$T = \frac{\overline{X} - \mu}{\frac{S}{\sqrt{n}}}$$

è una variabile aleatoria avente la distribuzione t con grado di libertà $\nu = n - 1$.
I criteri per i test a una e a due code basati sull'uso di questa distribuzione sono analoghi a quelli già descritti nel paragrafo precedente, con z_α e $z_{\frac{\alpha}{2}}$ sostituiti da t_α e $t_{\frac{\alpha}{2}}$; questi valori critici per un dato livello di significatività α dipendono dal grado di libertà e devono essere letti di volta in volta sulle tavole della distribuzione t.

Nella tabella 2 riassumiamo i valori comunemente usati per il livello di significatività α e i corrispondenti valori critici t_α e $t_{\frac{\alpha}{2}}$ per i test a una e a due code.

Test	Ipot. nulla H_0	Ipot. altern. H_1	Liv. signif. α	Valori critici	Reg. rifiuto
una coda	$\mu \leq \mu_0$	$\mu > \mu_0$	0.01	$t_\alpha = t_{0.01}$	$T > t_{0.01}$
			0.05	$t_\alpha = t_{0.05}$	$T > t_{0.05}$
una coda	$\mu \geq \mu_0$	$\mu < \mu_0$	0.01	$t_\alpha = -t_{0.01}$	$T < -t_{0.01}$
			0.05	$t_\alpha = -t_{0.05}$	$T < -t_{0.05}$
due code	$\mu = \mu_0$	$\mu \neq \mu_0$	0.01	$t_{\frac{\alpha}{2}} = t_{0.005}$ $t_{\frac{\alpha}{2}} = -t_{0.005}$	$T > t_{0.005}$ $T < -t_{0.005}$
			0.05	$t_{\frac{\alpha}{2}} = t_{0.025}$ $t_{\frac{\alpha}{2}} = -t_{0.025}$	$T > t_{0.025}$ $T < -t_{0.025}$

Tabella 2

Esempio 16

Le bottiglie di vino poste in vendita contengono usualmente 750ml di vino.
Si effettua un controllo su un campione di 6 bottiglie e si misurano i seguenti valori in ml

747.0	751.5	752.0	747.5	747.0	749.0

Stabilire se questi dati confermano con un livello di significatività del 5% l'affermazione che le bottiglie hanno un contenuto medio pari a quanto dichiarato.

Se il test è effettuato per tutelare l'interesse del consumatore, l'ipotesi nulla e l'ipotesi alternativa sono

$$H_0: \quad \mu \geq 750$$
$$H_1: \quad \mu < 750.$$

Calcolando la media e la varianza del campione si ottengono i seguenti valori

$$\bar{x} = \frac{747.0 + 751.5 + 752.0 + 747.5 + 747.0 + 749.0}{6} = 749$$

$$s^2 = \frac{1}{5} \cdot \left(747.0^2 + 751.5^2 + 752.0^2 + 747.5^2 + \right.$$
$$\left. + 747.0^2 + 749.0^2 - 6 \cdot 749^2 \right) = 5.1$$

Il valore della statistica test è

$$T = \frac{749 - 750}{\frac{\sqrt{5.1}}{\sqrt{6}}} = -1.08 .$$

Il test è a una coda, e per il livello di significatività del 5% e il grado di libertà $\nu = 5$ il valore critico è

$$t_\alpha = -t_{0.05} = -2.015$$

La regione di rifiuto è data dai valori $T < -2.015$.

Il valore $T = -1.08$ appartiene alla regione di accettazione, perciò non si rifiuta l'ipotesi nulla: non c'è un'evidenza significativa, al livello del 5%, che le bottiglie contengano meno di 750ml di vino.

Esempio 19

Il contenuto dichiarato delle bottiglie di una certa bibita è 330ml.

Scegliendo un campione di 20 bottiglie, si riscontra un contenuto medio $\bar{x} = 328$ ml, con uno scarto quadratico medio $s = 3.2$ ml.

In base a questi dati si può ritenere che la ditta produttrice inganni il consumatore? Si assuma che la quantità di liquido contenuta nelle bottiglie segua approssimativamente la distribuzione normale e si scelga il livello di significatività dell'1%.

L'ipotesi nulla e l'ipotesi alternativa sono

$$H_0: \quad \mu \geq 330$$

$$H_1: \quad \mu < 330.$$

Il valore della statistica test è

$$T = \frac{328 - 330}{\frac{3.2}{\sqrt{20}}} = -2.8 .$$

Il test è a una coda, e per il livello di significatività dell'1% e il grado di libertà $\nu = 19$, il valore critico è

$$t_\alpha = -t_{0.05} = -2.539$$

La regione di rifiuto è data dai valori $T < -2.539$. Il valore $T = -2.8$ appartiene alla regione di rifiuto, perciò rifiutiamo l'ipotesi nulla e concludiamo che c'è una significativa evidenza, al livello dell'1%, che ci sia una frode da parte del produttore.

Per il livello di significatività del 5% il valore critico è

$$t_\alpha = -1.729$$

La regione di rifiuto è data dai valori $T < -1.729$. Il valore $T = -2.8$ appartiene alla regione di rifiuto, perciò anche al livello di significatività del 5% rifiutiamo l'ipotesi nulla, concludendo ancora che c'è una significativa evidenza di frode.

Riassumiamo nella tabella 3 i vari procedimenti da seguire per effettuare un test di ipotesi sulla media μ di una popolazione.

Procedimento	Ipotesi	Statistica test	Distribuzione della statistica test
1	$n \geq 30$ varianza σ^2 nota	$Z = \dfrac{\overline{X} - \mu_0}{\dfrac{\sigma}{\sqrt{n}}}$	Distribuzione normale
2	$n \geq 30$ varianza σ^2 incognita	$Z = \dfrac{\overline{X} - \mu_0}{\dfrac{s}{\sqrt{n}}}$	Distribuzione normale
3	$n < 30$ popolaz. normale varianza σ^2 nota	$Z = \dfrac{\overline{X} - \mu_0}{\dfrac{\sigma}{\sqrt{n}}}$	Distribuzione normale
4	$n < 30$ popolaz. normale varianza σ^2 incognita	$T = \dfrac{\overline{X} - \mu_0}{\dfrac{s}{\sqrt{n}}}$	Distribuzione t di Student (gradi libertà = n–1)

Tabella 3