Facoltà di Psicologia
Università di Padova
Anno Accademico 2010-2011

Elementi di inferenza statistica

L4, Corso Integrato di Psicometria - Modulo B

Dr. Marco Vicentini
marco.vicentini@unipd.it

Rev. 18/04/2011

Sommario

- Inferenza statistica
  - Formulazione di ipotesi
  - Errore I e II tipo
  - Regole decisionali e Test statistici
    - Monodirezionali e Bidirezionali
    - Potenza di un test

Formulazione di ipotesi

Inferenza statistica

- L'**inferenza statistica** può essere definita come un insieme di metodi, basati sulla teoria della probabilità, atti a condurre lo sperimentatore a conclusioni plausibili sulla variabile casuale associata ad una particolare caratteristica di una popolazione, rilevata mediante l'osservazione di un campione di elementi della popolazione stessa;
- Secondo l'impostazione classica, la teoria dell'inferenza statistica si sviluppa lungo due direttrici:
  - teoria della ***stima di parametri***;
  - teoria della ***verifica di ipotesi***.

Stima dei parametri

- Uno **stimatore (t)** è una funzione dei valori osservati sul campione $t(X_1,X_2,\ldots,X_n)$, utilizzato per la stima di un parametro θ (generalmente incognito) di una popolazione e che (possibilmente) rispetta una serie di caratteristiche;
  - Quando il metodo di stima conduce ad un unico valore per il parametro in oggetto, si parla di *stima puntuale*;
  - Quando il metodo di stima conduce ad un intervallo di valori, all'interno dei quali si suppone che si trovi con una certa probabilità il valore vero del parametro, si parla di *stima intervallare*.

Verifica di ipotesi

"Scientific theories… cannot be verified or confirmed by experience"

"Le teorie non sono mai verificabili empiricamente"

K. Popper (1935)

- Un'**ipotesi statistica** è una asserzione o supposizione sulla distribuzione di una o più variabili casuali, si indica con la lettera ***H***.

- **Esempio:**
  - Sia $X_1, X_2, \ldots, X_n$ un campione casuale estratto da una popolazione normale con media incognita θ, possiamo formulare l'ipotesi che la media di tale popolazione sia uguale ad un qualsiasi valore, quale, per esempio, 12.
  - Formalmente ciò viene espresso nella notazione:

$$H : \theta = 12$$

Ipotesi nulla

- Generalmente si mettono a confronto due ipotesi, espresse in modo tale da essere contrarie tra loro, mutualmente esclusive.
- **Ipotesi H_0**
- **È l'ipotesi relativa ad un parametro della popolazione**
  - Generalmente, è l'ipotesi che costituisce l'oggetto della verifica; specifica i valori dei parametri della popolazione da cui si suppone provenga il campione (o i campioni) in esame.
  - *Stabilisce che i valori ottenuti nel campione sono puramente casuali o che non vi siano relazioni oltre l'indipendenza.*
  - Formalmente si esprime con:

$$H_0 : \theta = \theta_0$$

  - in cui θ indica il parametro oggetto di indagine della popolazione e θ_o il valore che ci si attende.

Ipotesi alternativa

- **Ipotesi H_1**
- **È l'ipotesi che lo sperimentatore intende mettere alla prova.**
  - È l'ipotesi alternativa, solitamente l'ipotesi di ricerca secondo la quale i risultati ottenuti non sono casuali ma sono spiegabili secondo una qualche regola, relazione o teoria.
  - L'ipotesi alternativa può essere *bidirezionale*, ed in tal caso si esprime semplicemente come

$$H_1 : \theta \neq \theta_0$$

  - oppure può essere *monodirezionale (destra o sinistra):*

$$H_1 : \theta > \theta_0$$

$$H_1 : \theta < \theta_0$$

Esempio 1

- Si vuole stabilire, attraverso un esperimento, se una moneta sia truccata o meno. Indicheremo con *p* la probabilità di ottenere testa in un lancio.
- Prima di eseguire l'esperimento dobbiamo formulare le due ipotesi alternative relativamente al parametro *p*:
  - $H_0 : p = 0,5$ la moneta non è truccata
  - $H_1 : p \neq 0,5$ la moneta è truccata

Errore I e II tipo

Errore I e II tipo

- Dal punto di vista decisionale, la suddivisione dello spazio campionario in una regione di accettazione ed una di rifiuto porta a quattro soluzioni possibili.

Ipotesi

Decisione	*Ipotesi nulla H_0 vera*	*Ipotesi alternativa H_1 vera*
accetto H_0	corretto	errore II tipo falso negativo
rigetto H_0	errore I tipo falso positivo	potenza

Errore I e II tipo

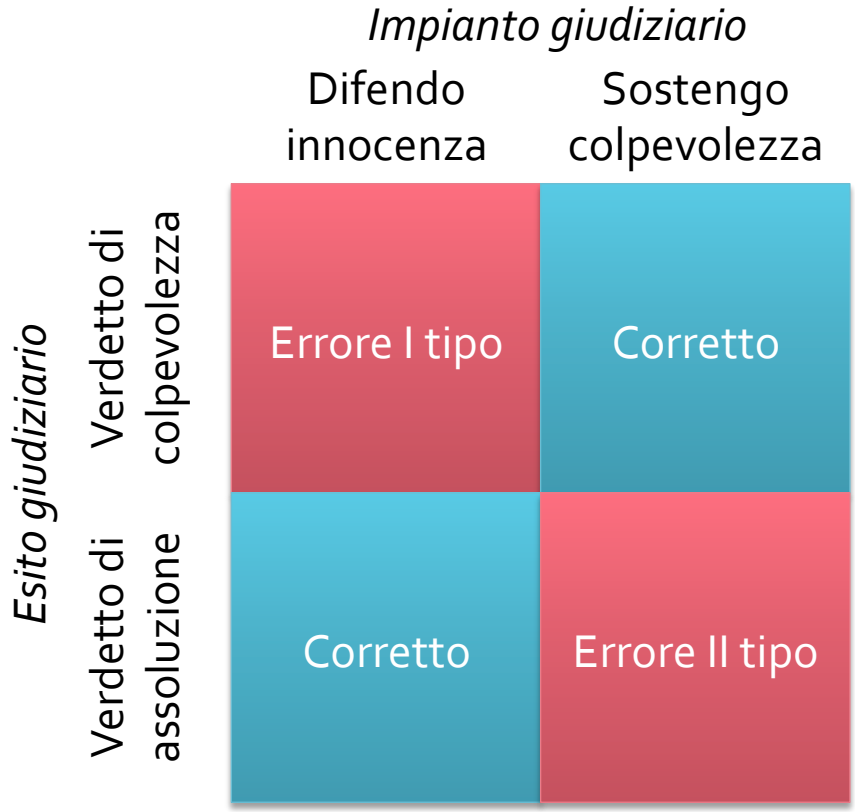
- Errore I tipo:
  - È conosciuto come il "falso positivo", quando si rigetta l'ipotesi nulla H_0 quando essa è vera.
    - Ad esempio il parametro nella popolazione è davvero nullo, una donna è incinta quando non lo è, …
    - È un errore di eccessiva crudeltà che porta a falsi allarmi
- Errore II tipo:
  - È conosciuto come il "falso negativo", quando si accetta l'ipotesi nulla quando essa è falsa, e si sarebbe dovuta accettare l'ipotesi alternativa
    - Ad esempio quando non si da una invalidità ad una persona davvero malata …
    - È un errore di eccessivo scetticismo che porta a non vedere le cose palesi.

Errore I e II tipo: un esempio "civile"

- "La polizia arresta una persona per un crimine"
- Formulazione delle ipotesi
  - H_0: la persona è innocente
  - H_1: la persona è colpevole
    - (il motivo per cui è stato arrestato)
- Errore
  - **Errore I tipo**: falso positivo.
    - Rifiuto H_0 quando questa è vera.
    - La persona viene condannata ingiustamente
  - **Errore II tipo**: falso negativo.
    - Accetto H_0 quando questa è falsa (e avrei dovuto accettare H_1)
    - Un criminale è in libertà

Errore I e II tipo: un esempio "civile"

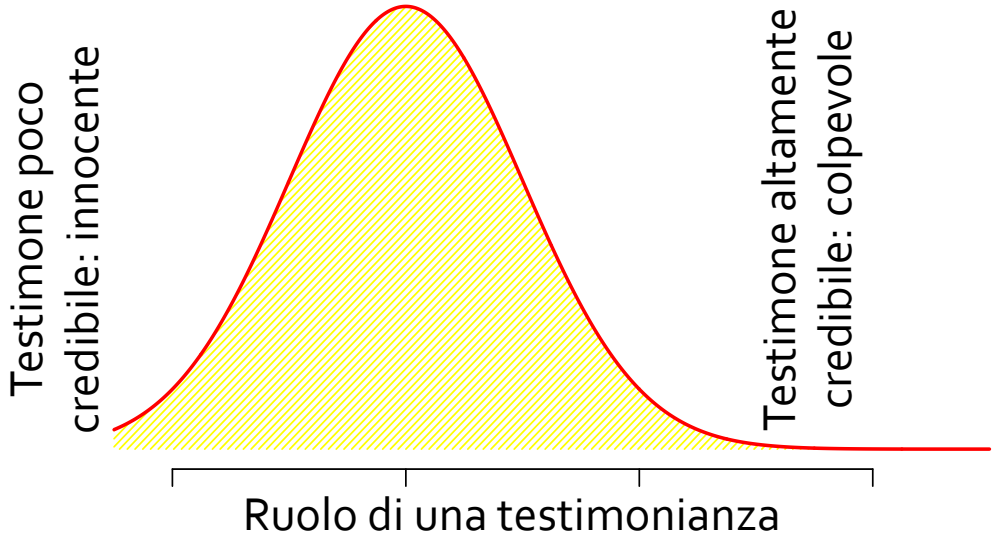
- H_0: la persona è innocente
- H_1: la persona non è innocente



Errore I e II tipo: un esempio "civile"

- Distribuzione d'errore.

Una persona innocente è arrestata



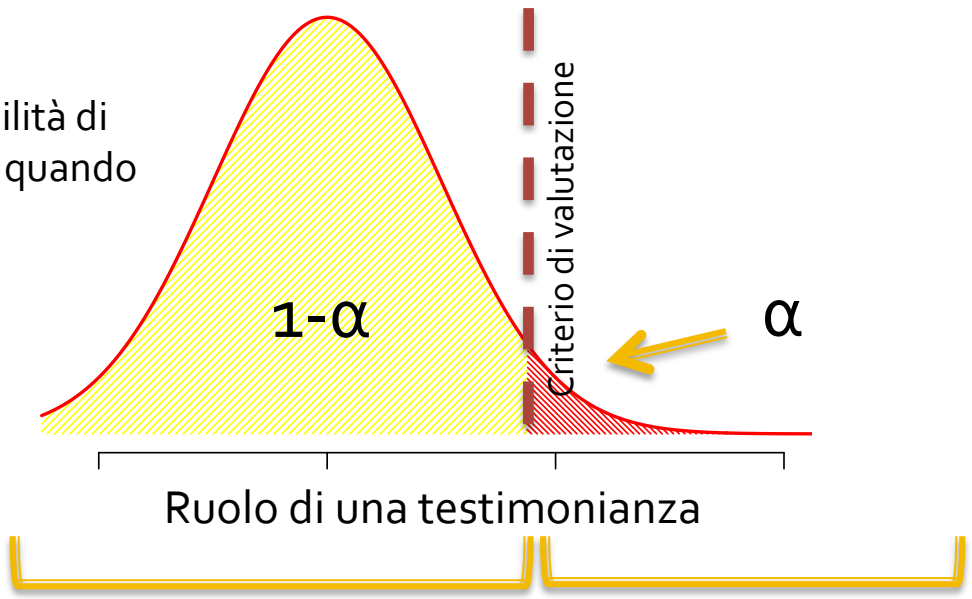
Errore I e II tipo: un esempio "civile"

- Errori di primo tipo per ipotesi monodirezionale destra.

Una persona innocente è arrestata

α = falso positivo

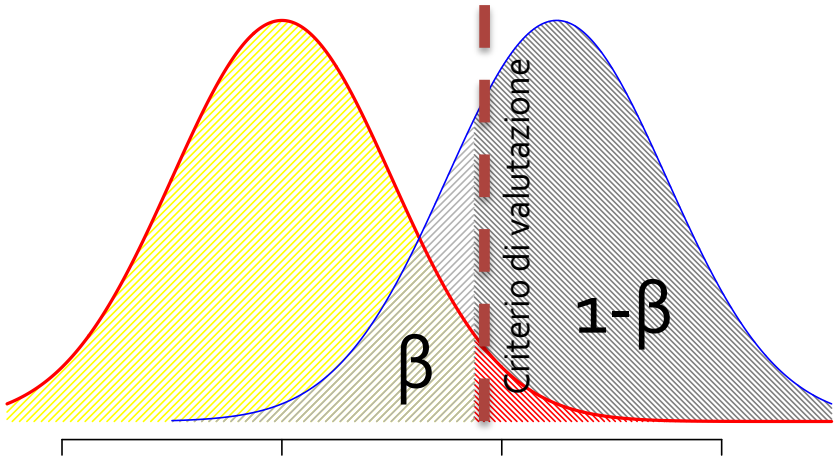
1-α = probabilità di accettare H_0 quando H_0 è vera



Errore I e II tipo: un esempio "civile"

- Errori di primo e secondo tipo, e potenza del test per ipotesi monodirezionale destra.

Una persona innocente è arrestata Una persona colpevole è arrestata



α = falso positivo

β = falso negativo

1-β = probabilità di accettare H_1 quando H_1 è vera

Ruolo di una testimonianza

H_0

Giudizio di non colpevolezza

H_1

Giudizio di non innocenza (colpevolezza)

Errore I e II tipo

- Una diminuzione dell'errore di I tipo porta necessariamente all'aumentare dell'errore di II tipo.
  - Al diminuire dei falsi positivi si avrà un conseguente aumento dei falsi negativi.
- L'unica possibilità per variare la distanza tra le due distribuzioni è aumentare la numerosità campionaria
  - Per arrivare alla diminuzione della varianza della media campionaria.

Errore a-priori

- All'*errore di primo tipo* si attribuisce un valore **α** piccolo fissato a priori, generalmente può essere .05, .01 oppure .001;
  - Tale valore rappresenta la *probabilità di respingere l'ipotesi nulla H_0 quando essa è vera*.
  - Il corrispondente 1-α corrisponde dunque alla probabilità di accettare H_0 quando H_0 è vera. Ovviamente si vuole che tale probabilità sia alta.
- La probabilità dell'*errore di secondo tipo* non viene fissata a priori ma dipende da α e la numerosità campionaria.
  - il valore β rappresenta la *probabilità di non respingere (quindi accettare) l'ipotesi nulla H_0 quando è falsa*.

  - **Cohen J. (1994). *The earth is round (p < 0.05)***

Potenza di un test

- la quantità **1-β** , cioè la probabilità di rifiutare H_0 quando essa è falsa o equivalentemente la probabilità di accettare H_1 quando essa è vera, si definisce "**potenza di un test statistico**".
- nell'ambito della verifica d'ipotesi, oltre a fissare opportunamente la probabilità di errore di primo tipo α, è importante mantenere sotto controllo la potenza del test che si utilizza;
- Nell'ambito delle scienze sociali, c'è un generale accordo sul fatto che la potenza di un test non dovrebbe scendere al di sotto di 0.80 (Cohen, 1977)

Accettere l'ipotesi nulla ?

- Alcuni ricercatori dicono che un test di ipotesi può avere uno dei due risultati: **si accetta l'ipotesi nulla** o si rifiuta l'ipotesi nulla.
- Molti statistici, tuttavia, sono in disaccordo con il concetto di "accettare l'ipotesi nulla". Affermano che sia preferibile dire:
  - si rifiuta l'ipotesi nulla
  - non si riesce a rifiutare l'ipotesi nulla
- Perché la distinzione tra "accettazione" e di "fallimento di rifiutare?"
  - L'accettazione implica che l'ipotesi nulla è vera.
  - Il mancato rifiuto implica che i dati non sono sufficientemente convincenti per far preferire l'ipotesi alternativa all'ipotesi nulla.

Test d'ipotesi e test statistico

Valutazione di ipotesi

- In statistica si è soliti seguire un processo formale per decidere se rifiutare una ipotesi nulla, sulla base di dati di un campione.
- Questo processo, chiamato **test di ipotesi**, si compone di quattro fasi.

1. **Formulare le ipotesi**. Ciò comporta indicare l'ipotesi nulla e l'ipotesi alternativa. Le ipotesi sono indicati in modo tale che si escludono a vicenda. Cioè, se una è vera, l'altra deve essere falsa.
2. **Formulare un piano di analisi**. Il piano di analisi descrive come utilizzare i dati sperimentali per valutare l'ipotesi nulla. La valutazione si focalizza sull'utilizzo di un singolo test statistico.
3. **Analizzare i dati del campione**. Trovare il valore della statistica test (punteggio medio, proporzione, T-score, Z-score, ecc) previsto nel piano di analisi.
4. **Interpretare i risultati**. Applicare la regola di decisione descritta nel piano di analisi. Se il valore di probabilità associato al test statistico è sfavorevole ($p(t) > \alpha_{crit}$) rifiutare l'ipotesi nulla.

Test statistico e regole decisionali

- Un **test statistico** è una variabile casuale che può assumere valori compresi in un particolare insieme, che costituisce lo spazio campionario del test, secondo una determinata distribuzione di probabilità;

- Ogni test statistico divide lo spazio campionario in due sottoinsiemi complementari:
  1. la *regione di accettazione di* H_o
  2. la *regione di rifiuto di* H_o

- In base alla regione in cui andrà a cadere il valore campionario *t*, si prenderà una decisione sull'ipotesi H_o.

Errore I e II tipo

- Dal punto di vista decisionale, la suddivisione dello spazio campionario in una regione di accettazione ed una di rifiuto porta a quattro soluzioni possibili.

	Ipotesi	
Decisione	H_0 *vera*	H_0 *falsa*
accetto H_0	$1 - \alpha$	β
rigetto H_0	α	$1 - \beta$

Regole decisionali

- In statistica vi sono due fondamentali regole decisionali: in riferimento al valore p di probabilità o in riferimento ad una regione di accettazione.
  - **P-value**. La forza di prova a sostegno di una ipotesi nulla è misurata dal valore-p associato ad una statistica S.
    - Il valore di probabilità p associato è pari alla probabilità di osservare secondo la distribuzione del test il valore S, supponendo che l'ipotesi nulla sia vera.
    - **Se il valore è è inferiore al livello di significatività prescelto, possiamo rifiutare l'ipotesi nulla**.
  - **Regione di accettazione**. La regione di accettazione è un intervallo di valori.
  - Se la statistica del test cade nella regione di accettazione, l'ipotesi nulla non viene respinta. La regione di accettazione è definita in modo che la possibilità di fare un errore di tipo I è uguale al livello di significatività. L'insieme dei valori al di fuori della regione di accettazione è chiamata la regione di rifiuto.
  - Se la statistica test rientra nella regione di rifiuto, l'ipotesi nulla è respinta. In tali casi, diciamo che l'ipotesi è stata respinta a livello α di significatività.

- Questi approcci sono equivalenti. Alcuni manuali utilizzano l'approccio basato sul valore p, altri usano la regione di accettazione.
  - Nell'ambito internazionale (norme APA) la modalità più utilizzata è basata sul valore p.

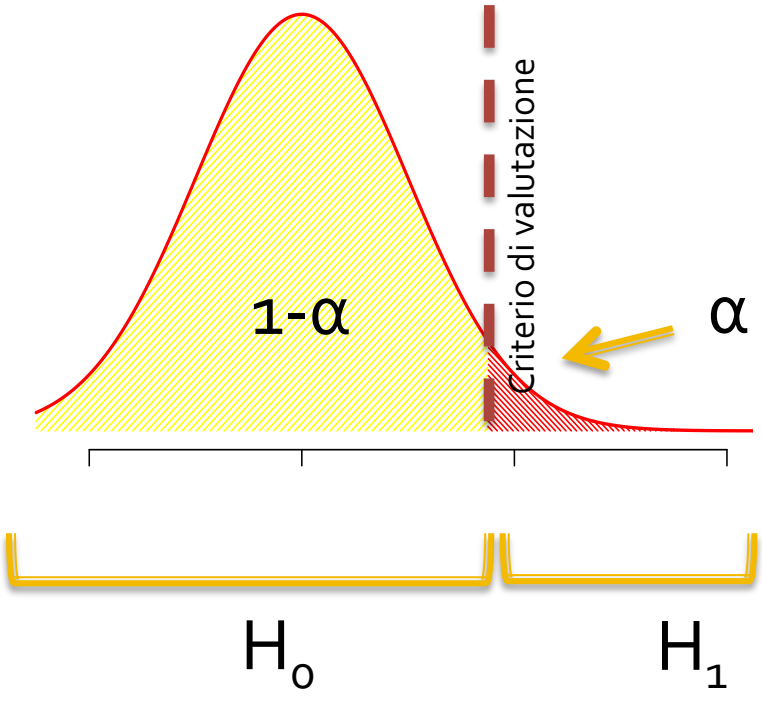
Test a una e due code

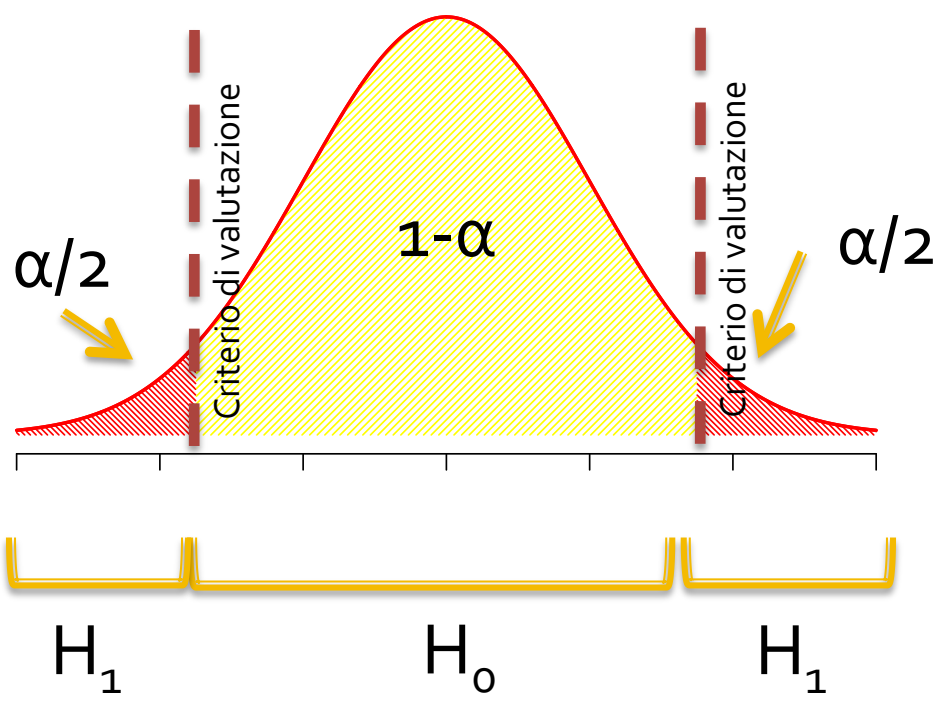
- Un test statistico di ipotesi in cui la regione di rifiuto prevede solo un lato della distribuzione di campionamento, viene chiamato un **test ad una coda, o monodirezionale**.
  - Ad esempio, supponiamo che l'ipotesi nulla preveda che la media sia inferiore o uguale a 10. La regione di rifiuto consiste in una serie di valori che si trovano situati sul lato destro della distribuzione di campionamento, cioè un insieme di numeri maggiori di 10.

- Un test statistico di ipotesi in cui la regione di rifiuto è su entrambi i lati della distribuzione di campionamento, è chiamato un **test a due code**, o bidirezionale.
  - Ad esempio, supponiamo che l'ipotesi nulla preveda che la media è pari a 10. L'ipotesi alternativa sarebbe che la media è inferiore a 10 o superiore a 10. La regione di rifiuto consisterebbe di una serie di numeri situati su entrambi i lati del campionamento di distribuzione, cioè la regione di rifiuto sarebbe composto in parte da valori che sono inferiori 10 e in parte di numeri che sono superiori a 10.

Test a una e due code

- Test monodirezionale
- Test bidirezionale





Tipi di test

1. **Test ad una coda.** Il valore del parametro θ è minore o maggiore di un valore θ_0

$$H_1 : \theta > \theta_0$$

2. **Test a due code.** Il valore del parametro θ è interno o esterno ad un intervallo delimitato da due valori specificati θ_1 e θ_2 o è diverso (maggiore e minore) da un valore θ_2

$$H_1 : \theta \neq \theta_0$$

$$H_1 : \theta_1 < \theta < \theta_2$$

Esemplificazioni

Esempio 1

- Ipotizziamo che la popolazione degli studenti di psicologia abbia una disponibilità mensile media μ di euro 100, con deviazione standard σ pari a 100.
  - Preso un gruppo di n=25 studenti, si osserva come questi abbiano una disponibilità media X di euro 150.
  - L'ipotesi nulla è che la media osservata nel campione sia non diversa dalla media della popolazione (ipotesi bidirezionale)
  - Tramite uno opportuno test (che approfondiremo successivamente) è possibile calcolare la statistica associata a tale differenza

$$z = \frac{\bar{X} - \mu}{\sigma_{\bar{X}}} = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

  - Emerge un valore della statistica z = 2.5

- A questo punto si possono prevedere due strade:
  1. p(z) = 0.9938 > α/2 = 0.975
  2. z > z(α/2) = 1.96

- Quindi si rifiuta H_0

Esempio 2

"Calcolare la potenza di un test significa calcolare la probabilità di cadere nella regione di rifiuto del test dato H_1 vera"

- La durata dei tradizionali pneumatici per auto è distribuita normalmente con media 50,000 Km e deviazione standard pari a 8,000.
- La ditta Y vuole introdurre nel mercato un nuovo tipo di pneumatico affermando che il nuovo prodotto ha una durata maggiore rispetto a quelli tradizionali.
- A tal proposito è stato condotto un test su 100 pneumatici della ditta Y, che ha riscontrato una durata media dei nuovi pneumatici pari a 51,500 Km.

Esempio 2

1. Ad un livello di significatività del 5% (α = 0.05), si può affermare che i pneumatici della ditta Y abbiano una durata maggiore rispetto a quelli tradizionali ?
2. In riferimento al valore osservato di 51,500 Km, calcolare la potenza del test utilizzato per la verifica d'ipotesi.

- formulazione delle ipotesi

$$H_0 : \mu = 50{,}000$$

$$H_1 : \mu > 50{,}000$$

Esempio 2 (soluzione I)

- calcolo del valore z
  - Visto che la varianza della popolazione è nota, utilizziamo la distribuzione normale e calcoliamo il valore di z osservato:

$$z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} = \frac{51{,}500 - 50{,}000}{8{,}000 / \sqrt{100}} = 1.875$$

Esempio 2 (soluzione I)

- **determinazione dei valori critici**
  - Fissato un livello di significatività $\alpha = 0{,}05$, e dato che il test è monodirezionale, bisogna trovare sulla tavola della normale il punto *z* che lascia a destra una probabilità di 0,05.
  - Tale valore è pari a 1.645.

- **decisione e conclusione**
  - Poiché $z_{cal} = 1.875$ è superiore al valore critico $z_c = 1.645$, dobbiamo rigettare l'ipotesi H_0.
  - Ad un livello di significatività del 5%, si può quindi affermare che i pneumatici della ditta Y hanno una durata superiore rispetto a quelli tradizionali.

Esempio 2 (soluzione II)

- Determinazione della regione di rifiuto del test
  - Ad un livello di significatività del 5% la regione di rifiuto del test è costituita dall'insieme delle medie campionarie che soddisfano la seguente disequazione:

$$\frac{\bar{X} - \mu}{\sigma / \sqrt{n}} > z_{\text{critico}}$$

  - sostituendo:

$$\frac{\bar{X} - 50000}{8000 / \sqrt{100}} > 1.645$$

Esempio 2 (soluzione II)

- risolvendo la disequazione:

$$\bar{X} > \left(1.645 \times \frac{8000}{\sqrt{100}}\right) + 50000$$

- la regione di rifiuto del test è quindi composta da tutte le medie campionarie tali che:

$$\bar{X} > 51316$$

Esempio 2 (potenza del test)

- Calcolo della potenza del test
  - Per calcolare la potenza del test in corrispondenza del valore osservato di 51,500, dobbiamo calcolare la probabilità di cadere nella regione di rifiuto del test, ipotizzando che sia vera H_1:

$$Potenza(51500) = P(\bar{X} > 51316 / H_1 \text{ è vera})$$

  - cioè, standardizzando entrambi i membri della disequazione:

$$Potenza(51500) = P\left(z > \frac{51316 - 51500}{8000/\sqrt{100}}\right) = P(z > -0.23)$$

Esempio 2 (potenza del test)

- Utilizzando le tavole della distribuzione normale osserviamo che:

$$Potenza(51{,}500) = P(z > -0.23) = 0.59$$

- Possiamo quindi concludere che la potenza del test in corrispondenza del valore osservato di 51,500 è pari al 59%
- Conseguentemente che la probabilità di errore del secondo tipo β è pari al 41% (0.41 = 1 - 0.59).