VERIFICA DI IPOTESI STATISTICHE

Psicometria 1 - Lezione 11
Lucidi presentati a lezione

AA 2000/2001 dott. Corrado Caudek

Una decisione statistica riguarda la scelta tra due ipotesi contrapposte: *l'ipotesi nulla* e *l'ipotesi sostantiva*. La convenzione è di ritenere vera l'ipotesi sostantiva quando l'ipotesi nulla è falsa. L'ipotesi sostantiva è l'ipotesi complementare rispetto all'ipotesi nulla.

La notazione usata è H_0 per l'ipotesi nulla e H_1 per l'ipotesi sostantiva.

1. Si enuncia esplicitamente l'ipotesi che viene sottoposta a test (ipotesi nulla) insieme all'ipotesi sostantiva.

2. Si specifica la statistica che è maggiormente rilevante per il test dell'ipotesi nulla. Si introducono delle assunzioni che consentono di determinare la distribuzione campionaria della statistica in questione.

3. Si determina la distribuzione campionaria della statistica rilevante. La distribuzione campionaria è costruita assumendo che sia vera l'ipotesi nulla.

4. Si sceglie un livello di significatività. Questo consente di dividere la distribuzione campionaria in due parti. Una regione (regione di rifiuto) conterrà valori relativamente improbabili della statistica (se l'ipotesi nulla è vera) e una regione conterrà valori relativamente probabili della statistica (se l'ipotesi nulla è vera).

5. Si calcola la statistica rilevante per il test dell'ipotesi nulla sulla base dei dati forniti da un campione.

6. Si prende una decisione in base al valore della statistica calcolata. Se la statistica cade all'interno della *regione di rifiuto*, l'ipotesi nulla viene rigettata in favore dell'ipotesi sostantiva. Se invece la statistica calcolata cade al di fuori della regione di rifiuto, l'ipotesi nulla non viene rigettata e <u>lo sperimentatore sospende il suo giudizio</u>. Una statistica del campione che cade all'interno della regione di rifiuto si dice *statisticamente significativa*.

Esempio. Supponiamo di volere stabilire se una moneta è truccata o meno.

In questo caso, l'*ipotesi nulla* è: "la moneta non è truccata"; l'*ipotesi sostantiva* è: "la moneta è truccata".

Per sottoporre a test l'*ipotesi nulla*, consideriamo il numero di successi (esito *T*) di una serie di *n* lanci della moneta. Supponiamo di osservare **9 successi in 10 prove.**

Decidiamo di rifiutare l'ipotesi nulla se il risultato osservato è molto improbabile. Dire che un dato risultato è improbabile equivale a dire che, ripetendo l'esperimento tante volte, un dato risultato viene osservato molto raramente (diciamo in meno del 5% dei casi).

Anziché ripetere tante volte l'esperimento, facciamo ricorso alla distribuzione campionaria della statistica considerata (numero r di successi in n prove).

La distribuzione campionaria ci dice appunto quale è la probabilità di osservare tutti i possibili esiti di un esperimento, se questo venisse ripetuto infinite volte.

Nel caso in questione dobbiamo usare la distribuzione binomiale. La distribuzione binomiale, infatti, ci dice quale è la probabilità associata a ciascuno dei possibili esiti di questo esperimento (r successi in n prove bernoulliane).

La distribuzione binomiale è completamente specificata quando sappiamo il numero di prove e alla probabilità di un successo. Nel nostro caso, il numero di prove è 10. Quale è la probabilità di un successo?

L'ipotesi nulla ipotizza che la moneta sia onesta, dunque, poniamo che la probabilità di un successo sia .5.

Nel test di un ipotesi statistica la distribuzione campionaria viene costruito assumendo che sia vera l'ipotesi nulla.

Se il valore osservato (9 successi) è molto improbabile all'interno della distribuzione campionaria, allora concludiamo che l'ipotesi in base alla quale abbiamo costruito la distribuzione campionaria (ovvero, l'ipotesi nulla) è falsa.

Nel caso presente abbiamo scelto un livello di significatività del 5%. Questo ci consente di trovare i valori della statistica esaminata che porteranno a rifiutare l'ipotesi nulla.

Quali sono questi valori? Sono i valori estremi della distribuzione campionaria, ovvero quelli a cui è associata una probabilità complessiva non superiore a .05.

Quali sono i valori più improbabili nel caso di questa distribuzione campionaria? Sono valori molto alti (9 o 10) e molto bassi (0 o 1) di successi in 10 prove.

$b(n = 10, p = .5; r = 9) = .0098;$
$b(n = 10, p = .5; r = 10) = .0016.$

La probabilità complessiva di osservare 9 successi in 10 prove, o un risultato ancora più estremo, dunque, è uguale a .0114.

In conclusione, la statistica osservata è molto improbabile all'interno della distribuzione campionaria. Per questa ragione l'ipotesi nulla (*la moneta non è truccata*) viene rifiutata e l'ipotesi sostantiva (*la moneta è truccata*) viene ritenuta vera.

Esempio. Supponiamo di sapere che la popolazione ha una distribuzione normale con $s = 10$, e ipotizziamo che la media della popolazione sia uguale a 75.

Un campione casuale di 25 individui estratto da questa popolazione ha una media di 79.

Assumendo un livello di significatività del 5%, che cosa possiamo concludere?

Le ipotesi nulla e sostantiva sono:

H_0: *m = 75, s = 10, distribuzione normale.*

H_1: *m ? 75, s = 10, distribuzione normale.*

Se l'ipotesi nulla è vera, la distribuzione campionaria della media sarà normale, con media 75 ed errore standard uguale a $10 / \sqrt{25} = 2$.

(La distribuzione campionaria della media è normale per campioni estratti da una popolazione normale)

Quale è la regione di rifiuto? Quali sono i valori della media che portano a rifiutare l'ipotesi nulla in base alla quale abbiamo costruito la distribuzione campionaria?

La regione di rifiuto è quella che contiene valori che si discostano molto dalla media ipotizzata dall'ipotesi nulla. La regione di rifiuto, dunque, corrisponde all'insieme di casi contenuti nelle code della distribuzione.

L'ipotesi alternativa dice che la media è diversa da 75. Potrebbe essere più grande o più piccola. Dunque, risultati possibili che portano a rigettare l'ipotesi nulla sono i risultati estremi sia molto maggiori che molto minori di 75.

In questo caso, la regione di rifiuto corrisponde, dunque, a entrambe le code della distribuzione.

Quanto è grande la regione di rifiuto? Il livello di significatività prescelto è quello del 5%.

Questo significa che nella regione di rifiuto cade il 5% dei casi della distribuzione.

Metà di questi casi saranno nella coda della distribuzione che contiene valori molto grandi e metà nella coda che contiene valori molto piccoli.

Ciascuna delle due regioni di rifiuto contiene dunque il 2.5% dei casi possibili.

Come si calcolano i limiti delle due regioni di rifiuto?

Quello che dobbiamo fare è trovare un punto a tale per cui l'area sottesa dal grafico della funzione tra a e ∞ è uguale a .025.

Allo stesso modo, dobbiamo trovare un punto b tale per cui l'area sottesa dal grafico della funzione tra b e $-\infty$ è uguale a .025.

Nel caso presente, questi valori sono 71.08 e 78.92.

m=75; s=2;
NIntegrate[((1/Sqrt[2 Pi s^2]) E^(-(x-m)^2/(2 s^2))),{x,-Infinity,71.08}] = 0.0249979
NIntegrate[((1/Sqrt[2 Pi s^2]) E^(-(x-m)^2/(2 s^2))),{x,-Infinity,78.92}] = 0.975002

Dato che il valore osservato (ovvero, 79) cade all'interno della regione di rifiuto di destra, concludiamo rigettando l'ipotesi nulla (H_0: $\mu = 75$, $\sigma = 10$, distribuzione normale) e accettando l'ipotesi sostantiva (H_1: μ ? 75, $\sigma = 10$, distribuzione normale).

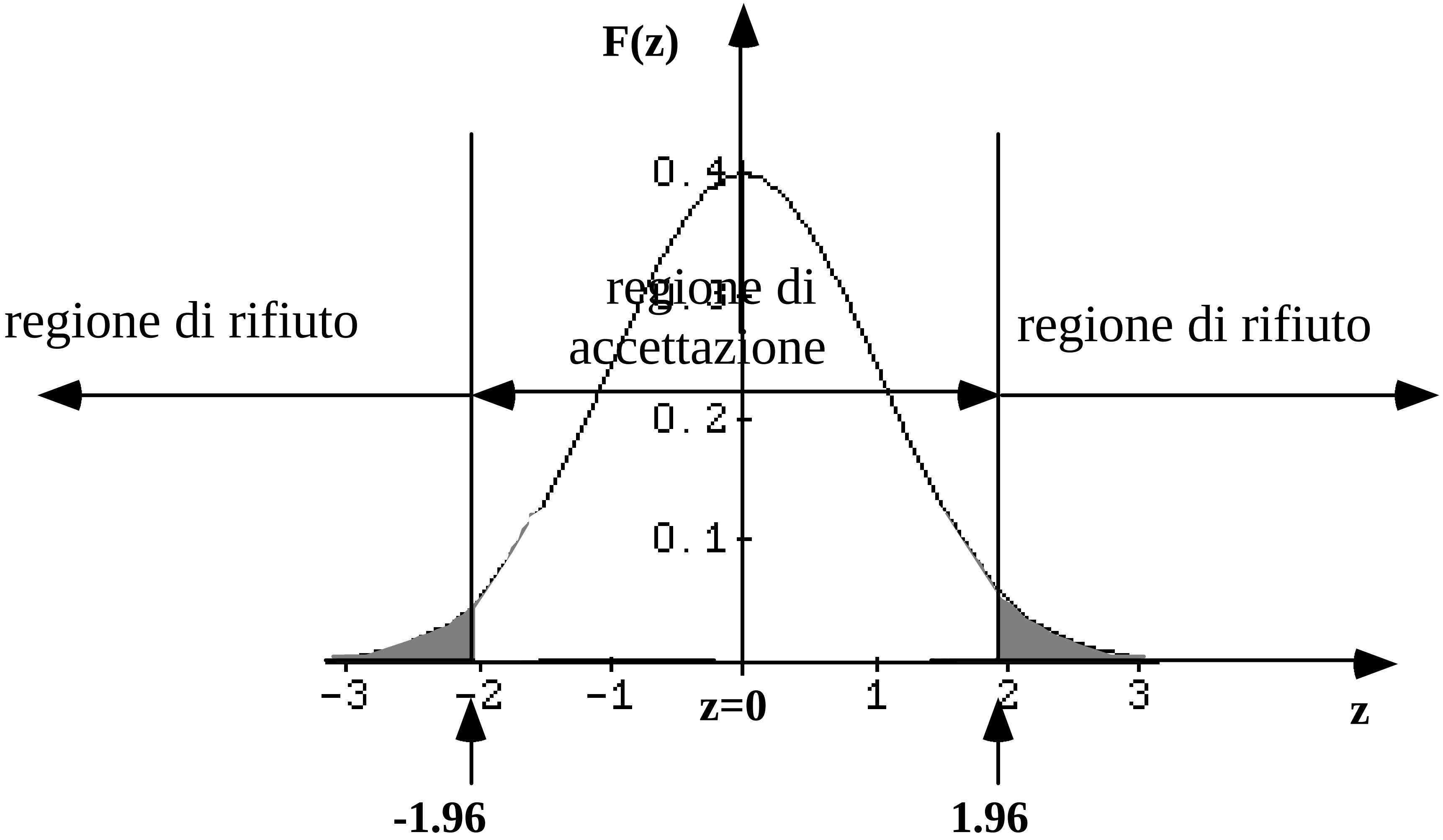
Poniamoci ora il problema di trovare i due valori critici che delimitano le regioni di rifiuto.

Se trasformiamo la media osservata del campione in un punteggio standardizzato, possiamo facilmente stabilire se cade all'interno della regione di rifiuto consultando le tavole della distribuzione normale standardizzata.

Nel caso di una distribuzione normale standardizzata, il 95% dei casi è contenuto nell'intervallo ($-1.96 < z < 1.96$).

Le due regioni di rifiuto sono dunque

$$-\infty \le \mathbf{z} \le -1.96 \qquad 1.96 \le \mathbf{z} \le \infty$$



F(z)
0.4
0.3
0.2
0.1
regione di rifiuto
regione di accettazione
regione di rifiuto
-3
-2
-1
z=0
1
2
3
z
-1.96
1.96

La media del campione espressa in termini di unità standard è

$$z = \frac{\overline{X} - m}{s_M} = \frac{\overline{X} - m}{s / \sqrt{N}} = \frac{79 - 75}{10 / \sqrt{25}} = 2$$

Questo valore è maggiore del valore critico di 1.96.

La media osservata cade quindi all'interno della regione di rifiuto.

Come si calcolano i valori che delimitano la regione di rifiuto?

Con il livello di significatività del 5%, i punteggi z che delimitano la regione di rifiuto sono -1.96 e 1.96.

Se vogliamo trasformare questi punti z nelle unità di misura originarie, scriviamo

$$1.96 = \frac{\overline{X} - m}{s_M}, \qquad -1.96 = \frac{\overline{X} - m}{s_M}$$

$$\overline{X}_{\min} = m - 1.96 s_M = 75 - 1.96 (10 / \sqrt{25}) = 71.08$$
$$\overline{X}_{\max} = m + 1.96 s_M = 75 + 1.96 (10 / \sqrt{25}) = 78.92$$

Notate che la procedura seguita è simile a quella usata per calcolare l'intervallo di fiducia della media.

Nel caso della procedura di verifica di ipotesi statistiche costruiamo un intervallo sulla base delle informazioni fornite dall'ipotesi nulla.

Se il valore osservato cade al di fuori di questo intervallo, l'ipotesi nulla viene rifiutata.

ERRORI DI PRIMO E DI SECONDO TIPO

Quando prendiamo una decisione statistica (rigettiamo l'ipotesi nulla oppure non rigettiamo l'ipotesi nulla) ci esponiamo alla possibilità di commettere un errore. Due sono gli errori possibili.

		Vera	
		H_0	**H_1**
Decisione	**H_0**	**corretta** **$1 - \alpha$**	**errata** β
	H_1	**errata** α	**corretta** **$1 - \beta$**

Se decidiamo H_0 quando è vera H_1, allora commettiamo un tipo di errore.

Un altro tipo di errore si commette quando decidiamo H_1 mentre è vera H_0.

Le altre due possibilità corrispondono a decisioni corrette (decidere che H_0 è vera quando H_0 è vera, e decidere che H_1 è vera quando H_1 è vera).

Consideriamo il caso in cui è vera l'ipotesi H_0.

Se l'ipotesi nulla è vera, commettiamo un errore ogni qualvolta decidiamo che è vera l'ipotesi H_1. Questo errore si chiama errore di I tipo e la probabilità di fare un errore di questo genere è indicato con α.

La grandezza di α è stabilita a priori e corrisponde al livello di significatività del test (ad esempio .05).

Stabilita la grandezza di α, ne deriva che la probabilità di essere corretti (se l'ipotesi H_0 è vera) è uguale a $1 - \alpha$.

Consideriamo il caso in cui è vera l'ipotesi H_1.

Se l'ipotesi sostantiva è vera, commettiamo un errore ogni qualvolta decidiamo di non rifiutare l'ipotesi nulla. Questo è un *errore di II tipo*. La probabilità di commettere un tale errore si indica con β.

Dato β, ne deriva che la probabilità di essere corretti (se l'ipotesi H_1 è vera) è uguale a $1 - \beta$. Come si calcola β?

POTENZA DELLA VERIFICA

La probabilità di rigettare l'ipotesi nulla quando l'ipotesi alternativa è effettivamente vera è chiamata *potenza della verifica.*

Questa probabilità si indica con $1 - \beta$, ovvero a 1 meno la probabilità di commettere l'errore per cui decidiamo H_0 quando in realtà è vera H_1.

Esempio. Supponiamo che un ricercatore ipotizzi che la popolazione abbia media minore o uguale a 138 e deviazione standard uguale a 20 quando in realtà la vera media della popolazione è 142 e la deviazione standard è uguale a 20.

Supponiamo inoltre che la popolazione sia distribuita normalmente.

Le ipotesi nulla e sostantiva sono:

H_0: $\mu \leq 138$; $\sigma = 20$; distribuzione normale
H_1: $\mu > 138$; $\sigma = 20$; distribuzione normale

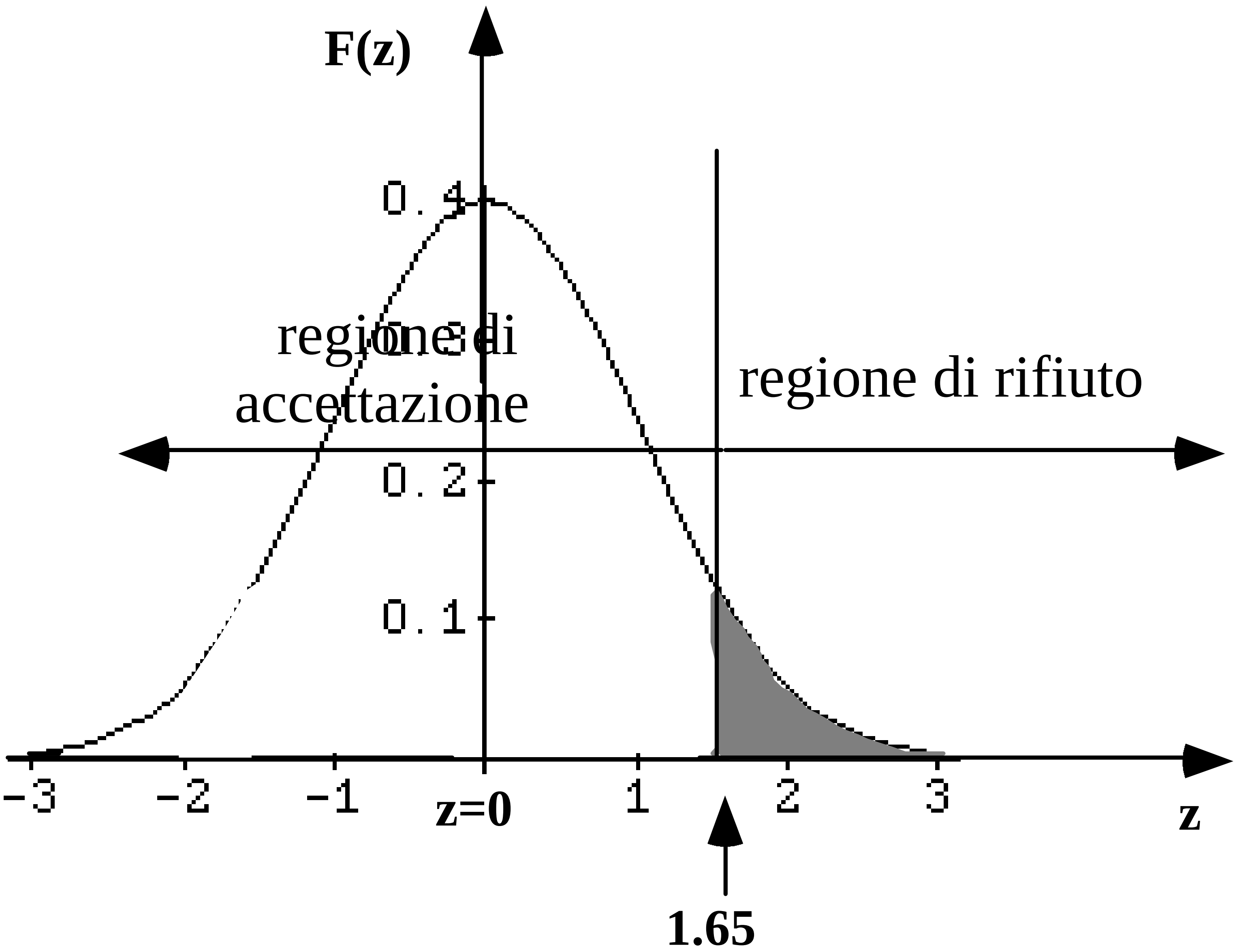
Viene raccolto un campione di grandezza n = 100 e si sceglie il livello di significatività del 5%.

Quale è la probabilità di accettare H_1 quando l'ipotesi H_1 è effettivamente vera?

Per prendere la sua decisione, il ricercatore ipotizza che sia vera l'ipotesi nulla H_0.

In base a quest'assunzione, costruisce una distribuzione campionaria con media 138 e errore standard $20/\sqrt{100}$.

Determiniamo dunque la *regione di rifiuto* all'interno di questa distribuzione campionaria.


F(z)
0.4
0.3
0.2
0.1
regione di
accettazione
regione di rifiuto
-3
-2
-1
z=0
1
2
3
z
1.65

Con un livello di significatività del 5%, la regione di rifiuto corrisponde all'area sottesa dalla curva normale per valori di z maggiori di 1.65.

Il valore critico che delimita la regione di rifiuto è:

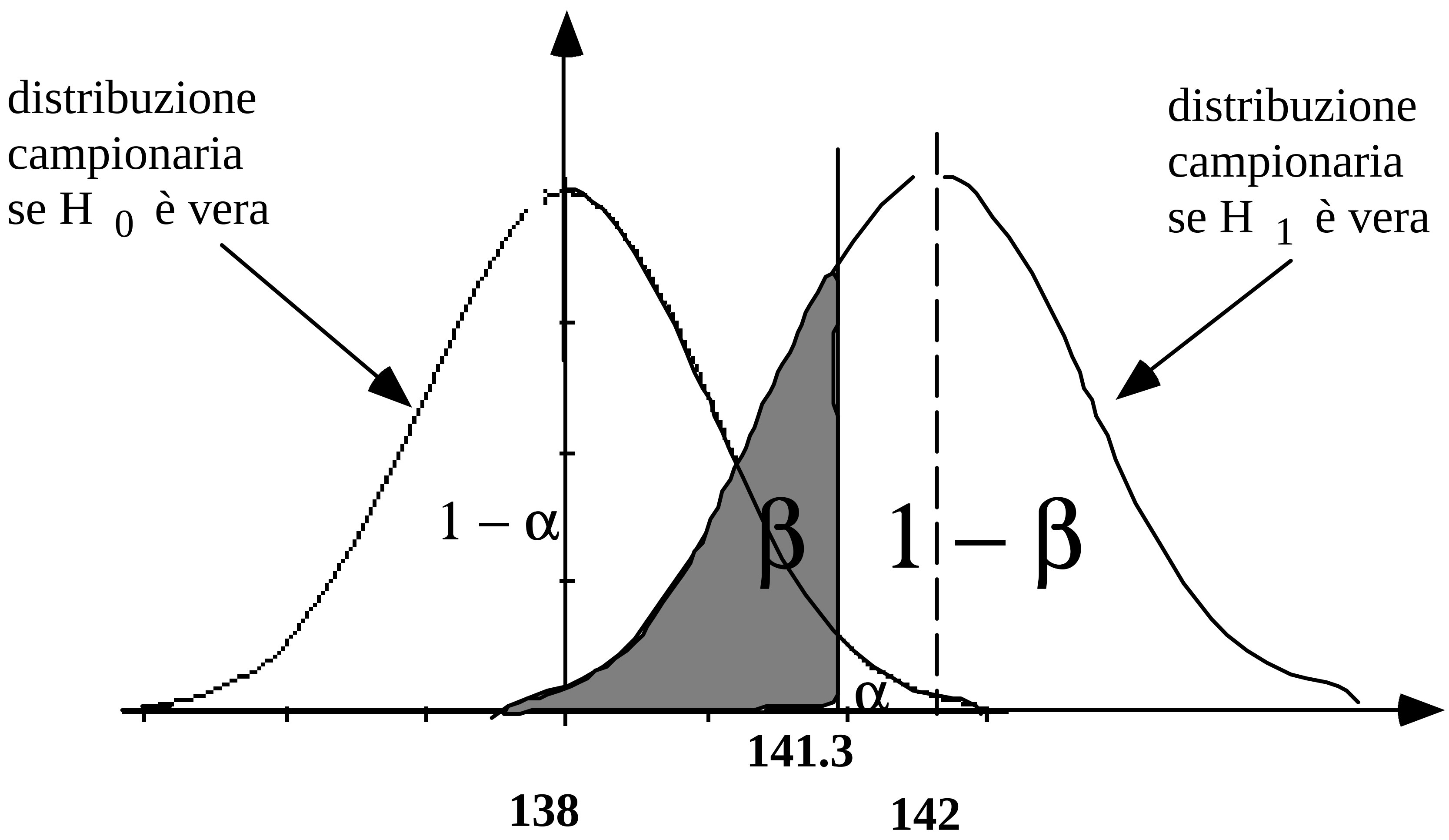
$$\bar{X} = 138 + 1.65 s_{\bar{X}} = 138 + 1.65 \frac{20}{\sqrt{100}} = 141.3$$

Il ricercatore accetta dunque l'ipotesi nulla in tutti i casi in cui il valore osservato della media del campione ha un valore *minore* di 141.3.

L'ipotesi nulla però è falsa. Ogni qualvolta il ricercatore osserva una media minore di 141.3 l'ipotesi nulla viene accettata e un errore di II tipo viene commesso.

Abbiamo chiamato β la probabilità di commettere un errore di questo genere. Come si calcola β?

Per calcolare β dobbiamo dunque prendere in considerazione la vera distribuzione campionaria della media, ovvero la distribuzione con media 142 ed errore standard $20/\sqrt{100}$.


distribuzione
campionaria
se H_0 è vera
distribuzione
campionaria
se H_1 è vera
$1-\alpha$
β
$1-\beta$
α
141.3
138
142

La probabilità di commettere un errore del secondo tipo (ovvero, β) si calcola esaminando, all'interno della vera distribuzione campionaria, l'area delimitata dal valore 141.3.

$$z_M = \frac{\bar{X} - m}{s_M} = \frac{141.3 - 142}{20 / \sqrt{100}} = -.35$$

L'area sottesa dalla normale standardizzata tra $-\infty$ e $-.35$ è uguale a .36.

Il ricercatore, dunque, ha una probabilità .36 di commettere un errore del II tipo (decidere H_0 quando in realtà è vera H_1).

La potenza della verifica (decidere H_1 quando l'ipotesi H_1 è vera) è uguale a 1 - β.

Nel nostro caso, la potenza della verifica è: 1 - .3632 = .64.

In conclusione, il ricercatore ha una probabilità di .64 di decidere l'ipotesi alternativa <u>quando in effetti l'ipotesi alternativa è vera</u>.

In altre parole, se l'esperimento venisse ripetuto molte volte, la decisione giusta (decidere H_1 dato che H_1 è effettivamente vera) verrebbe presa soltanto nel 64% dei casi.

ESERCIZIO

La variabile aleatoria X è distribuita normalmente con deviazione standard uguale a 4.2. Ventisei osservazioni vengono effettuate in maniera casuale e indipendente e generano una media campionaria di 31.

(a) Usando $\alpha = .05$, sottoponete a test l'ipotesi che la media della variabile X sia 28.6 contro l'ipotesi alternativa che la media corrisponda a qualche altro valore.

(b) Supponete che la vera media della popolazione sia 32. Calcolate la potenza della verifica del test.

IPOTESI ESATTE E INESATTE

Si possono distinguere due tipi di ipotesi alternative, le ipotesi esatte e inesatte.

Le ipotesi esatte forniscono una stima puntuale di un parametro della popolazione.

Quelle inesatte, invece, ipotizzano soltanto che il parametro della popolazione sia contenuto all'interno di un intervallo di valori.

Un esempio di ipotesi esatta è: $H: m = 100$.
Invece, $H: m \geq 100$ è un'ipotesi inesatta.

Tra le ipotesi inesatte distinguiamo tra ipotesi monodirezionali e ipotesi bidirezionali.

IPOTESI MONODIREZIONALI

Un esempio di ipotesi inesatta corrisponde alla situazione in cui ci si chiede se una media è maggiore di un altra oppure no. In questo caso, non disponiamo dei valori esatti delle due medie, ma solo di *due classi di valori*.

L'ipotesi nulla e sostantiva hanno la seguente forma:

$$H_0 : m \leq m_0$$

$$H_1 : m > m_0$$

I valori possibili della media della popolazione vengono divisi in due classi: la classe dei valori superiori e quella dei valori inferiori a un certo punteggio μ_0.

Lo sperimentatore si pone il problema di stabilire *in quale classe si colloca la media della popolazione date le evidenze fornite dal campione*.

Quando è possibile definire un solo valore critico che consente di dividere la distribuzione campionaria in due regioni contigue, allora parliamo di <u>ipotesi monodirezionali</u>.

In questo caso, tutti i valori della statistica che hanno grandezza maggiore (o minore) del punteggio critico cadono nella regione di rifiuto e portano a rigettare l'ipotesi nulla.

Nel caso di ipotesi monodirezionali, per rigettare l'ipotesi nulla non è sufficiente che la statistica calcolata a partire dai dati del campione sia molto diversa in valore assoluto dal valore critico dell'ipotesi nulla.

È necessario che il valore della statistica si differenzi dal valore critico nella *direzione ipotizzata dall'ipotesi sostantiva.*

Esempio. Un test di lingua inglese è stato messo a punto da uno psicologo americano. Questo test è stato standardizzato su una popolazione di soggetti americani adulti in modo tale che i soggetti ottengano un punteggio medio uguale a 100 con uno scarto quadratico medio uguale a 15. La distribuzione di questi punteggi può ritenersi approssimativamente normale.

Lo psicologo è interessato a applicare il test in Inghilterra e sospetta che i soggetti inglesi, per qualche ragione, ottengano in media dei punteggi superiori a quelli americani. Per verificare quest'ipotesi, lo psicologo sottopone il test a un campione di soggetti inglesi.

Lo psicologo sottopone a test 50 soggetti inglesi ottenendo una media di 105.

Si sottoponga a verifica l’ipotesi dello psicologo con α = .01.

Le ipotesi nulla e sostantiva sono ipotesi inesatte:

$$H_0 : m \leq 100$$

$$H_1 : m > 100$$

L'ipotesi sostantiva è monodirezionale.

In base all'ipotesi nulla non ci sono differenze tra le popolazioni inglese e americana rispetto al test (oppure, i punteggi della popolazione inglese sono addirittura più bassi di quelli della popolazione americana).

L'ipotesi sostantiva afferma invece che la popolazione inglese ha una media superiore a quella americana.

Lo psicologo adotta un livello α = .01 così da includere nella regione di rifiuto dell'ipotesi nulla tutti i punteggi *z* con un valore maggiore o uguale a 2.33. Dato che l'ipotesi sostantiva è μ > 100, la regione critica sarà delimitata da un valore uguale a:

$$\overline{X}_{critico} = m_0 + 2.33 s_{\overline{X}}$$

L'errore standard della distribuzione campionaria è

$$s_{\overline{X}} = \frac{s}{\sqrt{n}} = \frac{15}{\sqrt{50}} = 2.12$$

Il valore critico della media diventa:

$$\overline{X}_{critico} = m_0 + 2.33s_{\overline{X}} = 100 + 2.33(2.12) = 104.95$$

Dato che il valore ottenuto è maggiore di quello che delimita la regione di rifiuto, l'ipotesi nulla può essere rigettata al livello di significatività di .01.

Lo psicologo può dunque ritenere vera l'ipotesi sostantiva secondo la quale i soggetti inglesi ottengono punteggi superiori a quelli americani nel test considerato.

Esempio. Per un particolare compito sperimentale, lo sperimentatore ipotizza che circa il 20% della popolazione dei soggetti è in grado di completare il compito nel tempo richiesto. Per sottoporre a test quest'ipotesi, lo sperimentatore esamina 300 soggetti scelti in maniera casuale e somministra il test indipendentemente a ciascuno di essi. Il 25% di questo gruppo è in grado di compleatare il test nel tempo richiesto.

Sottoponete a test l'ipotesi che la vera proporzione della popolazione sia .20 (o minore) contro l'ipotesi che la proporzione sia maggiore di .20. Sia a = .05.

Il valore atteso e la varianza della distribuzione campionaria di una proporzione sono:

$$E(M_n) = E\left(\frac{S_n}{n}\right) = \frac{1}{n}E(S_n) = \frac{1}{n}np = p$$

$$V(M_n) = V\left(\frac{S_n}{n}\right) = \frac{1}{n^2}V(S_n) = \frac{1}{n^2}npq = \frac{pq}{n}$$

Quale è la forma di questa distribuzione campionaria? Dato che stiamo considerando un processo bernoulliano, la forma della distribuzione sarà quella binomiale.

Sappiamo però che la distribuzione binomiale si approssima alla normale con il crescere delle dimensioni del campione.

$$\boldsymbol{z = (.25 - .20)/(\mathrm{Sqrt}[(.25\ .75)/300]) = 2.0}$$

Quale è il valore critico che delimita la regione di rifiuto?

L'ipotesi sostantiva è monodirezionale ($p > .2$) e può essere accettata soltanto se la proporzione osservata nel campione è maggiore di .20.

Nel caso di una distribuzione normale standardizzata il limite critico della regione di rifiuto con $\alpha = .05$ corrisponde a un punteggio z di 1.65.

Dato che il valore della statistica osservata cade all'interno della regione critica, concludiamo rigettando l'ipotesi nulla e accettando l'ipotesi sostantiva.

ESERCIZIO

80 ratti sono scelti a caso e devono imparare a percorrere un nuovo tipo di labirinto. Il numero medio di prove necessario per l'apprendimento è 15.91. In base a una consolidata esperienza precedente con un labirinto simile, la media di prove necessarie per l'apprendimento si era attestata a 15 prove con una deviazione standard di 2. Possiamo concludere che il nuovo labirinto è più difficile del precedente? Sia $\alpha = .01$.

IPOTESI BIDIREZIONALI

Alle volte uno sperimentatore affronta un problema senza sapere che tipo di differenza aspettarsi se l'ipotesi H_0 è falsa.

Questo si verifica, per esempio, quando le domande poste sono del tipo "qualcosa è accaduto?", "c'è una differenza?", "c'è stato un cambiamento?"

Quando lo sperimentatore si pone questo tipo di domande, allora l'ipotesi alternativa è un'ipotesi bidirezionale.

Esempio. Un istituto scolastico ha messo a punto un test per valutare le prestazioni complessive degli studenti di vari livelli d'età. Per gli studenti della seconda superiore, nel corso degli anni precedenti il test ha prodotto un punteggio medio di 32 con una deviazione standard di 8. Si è stabilito, inoltre, che la distribuzione dei punteggi del test è approssimativamente normale.

Una classe di 16 studenti del secondo anno esibisce problemi di comportamento nel corso di tutto l'anno scolastico. Il preside ipotizza che questi studenti hanno capacità di apprendimento diverse dalla media. Per questi studenti, il punteggio medio nel test a fine d'anno è 37.

Quali sono le conclusioni a cui può giungere il preside sulla base di questi dati?

H_0: $m = 32$, $s = 8$, distribuzione normale
H_1: $m \neq 32$, $s = 8$, distribuzione normale

La distribuzione campionaria è normale con media 32 e errore standard di $s_M = 8 / \sqrt{16} = 2$.

Il punteggio z della media osservata è:

$$z = \frac{37 - 32}{2} = 2.5$$

L'ipotesi sostantiva è bidirezionale (accettiamo l'ipotesi nulla sia che gli studenti ottengano punteggi molto alti oppure molto bassi nel test).

Sappiamo che la probabilità di un punteggio z compreso nell'intervallo tra ±1.96 è 95.

In un test bidirezionale, dunque, l'intervallo che definisce la regione di accettazione dell'ipotesi nulla al livello di significatività del 5 per cento è compreso tra ± 1.96.

In altri termini, decidiamo di rifiutare l'ipotesi nulla se il valore osservato della statistica standardizzata ha un valore maggiore di 1.96 oppure minore di -1.96.

INTERVALLI DI FIDUCIA E IPOTESI BILATERALI

Le nozioni di <u>ipotesi bidirezionale</u> e <u>intervallo di fiducia</u> sono strettamente collegate.

Consideriamo il livello di significatività del 5 per cento e l'ipotesi nulla: $H_0\!: m = m_0$. In questo caso, l'intervallo di fiducia è:

$$\overline{X} - 1.96 s_{M} \leq m_0 \leq \overline{X} + 1.96 s_{M}$$

$$\bar{X} - 1.96 s_M \le m_0 \le \bar{X} + 1.96 s_M$$

$$-\bar{X} + 1.96 s_M \ge -m_0 \ge -\bar{X} - 1.96 s_M$$

$$1.96 s_M \ge \bar{X} - m_0 \ge -1.96 s_M$$

$$1.96 \ge \frac{\bar{X} - m_0}{s_M} \ge -1.96$$

In altre parole, troviamo i valori critici che delimitano le regioni di rifiuto di un'ipotesi bidirezionale.

L'intervallo di fiducia contiene dunque tutti i valori delle medie dei campioni che non possono essere rigettati al livello di significatività prescelto.

In conclusione, se un risultato è statisticamente significativo, la statistica calcolata in base al campione è poco probabile *se è vera l'ipotesi nulla*, mentre è relativamente più probabile se è vera l'ipotesi sostantiva.

La significatività statistica ha a che fare con la verosimiglianza di un certo risultato date alcune assunzioni e non ha altre implicazioni.

A proposito dei risultati statisticamente significativi dobbiamo notare che:

(1) un risultato statisticamente significativo non mette al riparo da errori metodologici che sono stati compiuti nel raccogliere i dati;

(2) è un grave errore valutare la "bontà" di un esperimento nei termini del livello di significatività del risultato ottenuto. È preferibile un risultato significativo ottenuto con un livello di significatività generoso (diciamo .05) nel caso di un esperimento ben controllato, piuttosto che un risultato significativo con un livello di significatività molto severo (diciamo .001) nel caso di un esperimento poco controllato.

Un risultato significativo, inoltre, non fornisce una risposta alle seguenti domande: "A quale popolazione si riferisce il risultato ottenuto?", "Il risultato significativo è stato ottenuto facendo una *selezione* dei soggetti che limita la generalizzabilità?", "Per ottenere un campione numeroso, lo sperimentatore ha soddisfatto i criteri del campionamento casuale?" Queste sono domande molto importanti e vanno al di là di quello che la significatività statistica di un risultato può dire.

Varianza della popolazione non conosciuta

Negli esempi precedenti abbiamo assunto che la varianza della popolazione sia conosciuta. Sfortunatamente, questo raramente accade in pratica.

Dobbiamo dunque *stimare la varianza della popolazione per mezzo della varianza del campione*.

In questo caso, l'errore standard stimato della distribuzione campionaria della media diventa:

$$\hat{s}_{\overline{X}} = \frac{s}{\sqrt{n}} = \frac{S}{\sqrt{n-1}}$$

Se il campione è grande abbastanza, una volta stimato l'errore standard, possiamo calcolare il punteggio standardizzato della media assumendo che la distribuzione campionaria sia distribuita normalmente.

Esempio. In un campione di 175 soggetti, le prestazioni medie dei soggetti in un dato compito corrispondono a un punteggio di 8.8 con una deviazione standard di 2.3.

Ricerche precedenti hanno rivelato che la prestazione media dei soggetti nel compito considerato dà luogo a un punteggio uguale a 9.

Il campione esaminato può essere considerato come un campione casuale della popolazione in esame, oppure rivela caratteristiche diverse? Si usi a = .01.

$$H_0 : m_0 = 9$$
$$H_1 : m_0 \neq 9$$

Con α = .01 e un'ipotesi bidirezionale, i limiti delle regioni di rifiuto corrispondono a un punteggio z di ± 2.58 .

Se l'ipotesi nulla è vera, la distribuzione campionaria è normale con media uguale a 9. L'errore standard <u>stimato</u> della distribuzione campionaria è:

$$\hat{s}_{\overline{X}} = \frac{S}{\sqrt{n-1}} = \frac{2.3}{\sqrt{175-1}} = .174$$

Il valore standardizzato della media del campione all'interno della distribuzione campionaria diventa:

$$z_{\overline{X}} = \frac{\overline{Y} - m}{\hat{s}_{\overline{X}}} = \frac{8.8 - 9}{.174} = -1.15$$

Dato che il valore standardizzato della media è minore in valore assoluto di 2.58, l'ipotesi nulla non può essere rigettata.

Anche gli intervalli di fiducia possono essere calcolati usando la stima dell'errore standard della distribuzione campionaria. In questo caso, l'intervallo di fiducia del 99% è:

$$\bar{Y} + 2.58\hat{s}_{\bar{X}} = 9.25$$

$$\bar{Y} - 2.58\hat{s}_{\bar{X}} = 8.35$$

FORMA DELLA DISTRIBUZIONE CAMPIONARIA DELLA MEDIA

La distribuzione campionaria di $\frac{\overline{X} - m}{s / \sqrt{n}}$ (σ conosciuto) è

	n grande	*n* piccolo
Popolazione normale	**Normale**	**Normale**
Popolazione NON normale	**Normale** (TLC)	**?**

La distribuzione campionaria di $\frac{\overline{X} - m}{s / \sqrt{n}}$ (σ non conosciuto) è

Popolazione normale	**Normale** s^2 fornisce una stima sufficientemente accurata di σ^2. Ci comportiamo come se σ fosse conosciuto	***t***
Popolazione NON normale	**Normale** (TLC) s^2 fornisce una stima sufficientemente accurata di σ^2. Ci comportiamo come se σ fosse conosciuto	**?**

Esempio. Un campione casuale di 15 osservazioni viene estratto da una popolazione normale con media e varianza non conosciute.

La media del campione è uguale a 100 e la varianza del campione è:

$$S^2 = \frac{\sum \left(X_i - \overline{X}\right)^2}{n} = 35$$

Lo sperimentatore ipotizza che la media della popolazione è uguale a 120.

Sia $\alpha = .05$.

a) Formulate le ipotesi nulla e sostantiva.

$$H_0: \mu = 120$$

$$H_1: \mu \neq 120.$$

b) L'ipotesi sostantiva è monodirezionale o bidirezionale?

Bidirezionale

c) Specificate la distribuzione campionaria per la statistica rilevante.

La statistica rilevante è la media del campione.

La distribuzione campionaria della media è normale con media uguale a 120.

Per calcolare l'errore standard dobbiamo prima "correggere" la varianza del campione affinché ci fornisca una stima priva di errore sistematico della varianza della popolazione:

$$s^2=(n\ /(n-1))\ S^2 = (15/14)\ 35 = 37.5$$

$$SE = Sqrt[37.5\ /\ 15] = 1.58114$$

d) In base a quale principio siete stati in grado di determinare la forma della distribuzione campionaria della statistica rilevante?

Quando la popolazione è normale e la varianza della popolazione viene stimata allora, per piccoli campioni, la statistica

$$\frac{\overline{Y} - E\left(\overline{Y}\right)}{s/\sqrt{n}}$$

segue la distribuzione *t*.

e) La distribuzione campionaria della statistica rilevante è stata costruita assumendo che sia vera l'ipotesi nulla oppure che l'ipotesi nulla sia falsa?

La distribuzione campionaria si costruisce assumendo che l'ipotesi nulla sia VERA.

f) In base al livello di significatività prescelto, quali sono i valori critici che delimitano le regioni di rifiuto?

Con a = .05 i limiti critici della regione di rifiuto per la distribuzione $t(14)$ sono: +2.145 e -2.145.

g) Calcolate il valore che assume la statistica rilevante all'interno della distribuzione campionaria standardizzata.

$$t_{14} = \frac{\bar{Y} - E\left(\bar{Y}\right)}{s \big/ \sqrt{n}} = \frac{100 - 120}{\sqrt{37.5} \big/ \sqrt{15}} = -12.6491$$

h) Che cosa potete concludere a partire dal risultato ottenuto in riferimento all'ipotesi nulla e sostantiva?

Il valore della media del campione cade all'interno della regione di rifiuto della distribuzione campionaria. Concludiamo rigettando l'ipotesi nulla in base alla quale distribuzione campionaria è stata costruita.

i) Costruite l'intervallo di confidenza al livello di fiducia del 95%.

$$100 + 2.145 \times (\ 37.5\ /\ \mathrm{sqrt}(15)\ ) = 103.392$$

$$100 - 2.145 \times (\ 37.5\ /\ \mathrm{sqrt}(15)\ ) = 96.6085$$

l) Quali sono i limiti che delimitano la regione di rifiuto dell'ipotesi nulla con $\alpha = .05$?

Con un ipotesi bidirezionale, sono gli stessi.

Esempio. Ci si chiede se la psicoterapia venga usata più spesso della terapia farmacologica nel caso di una particolare patologia. Uno psicologo si occupa di questo problema e raccoglie un campione casuale di 132 cartelle cliniche relative a pazienti affetti da questa patologia. Relativamente a questo campione, lo psicologo scopre che la terapia farmacologica è stata usata nel 62% dei casi.

(a) Si costruisca l'intervallo di fiducia del 95%.

(b) Sulle basi dell'intervallo calcolato si può concludere che, nella popolazione da cui il campione è stato tratto, la psicoterapia e la terapia farmacologica vengono usate con eguale frequenza?

$$.62 + 1.96\ \mathrm{sqrt}[(.62\ .38)\,/\,132] = 0.703$$

$$.62 - 1.96\ \mathrm{sqrt}[(.62\ .38)\,/\,132] = 0.537$$

Dato che l'intervallo di confidenza non copre il valore di .5, l'ipotesi che entrambe le terapie vengano usate con eguale frequenza può essere rigettata con $\alpha = .05$.

Esempio. Un campione di 15 osservazioni indipendenti è estratto a caso da una popolazione normale con media conosciuta uguale a 33. Il campione ha una media uguale a 46 con una varianza (priva di errore sistematico) di 24. Quale è la probabilità di osservare una media minore o uguale a quella del campione osservato all'interno della distribuzione campionaria della media per campioni di 15 osservazioni?

(a) Quale statistica è necessario calcolare per risolvere il problema?

(b) Specificate la forma, la media e l'errore standard della distribuzione campionaria della statistica rilevante.

(c) Calcolate il valore che assume la statistica rilevante e, a parole, descrivete la procedura necessaria per trovare la probabilità richiesta dal problema.

Quale statistica è necessario calcolare per risolvere il problema?

$$t = \frac{\overline{Y} - E\left(\overline{Y}\right)}{\hat{s}_{\overline{Y}}}$$

Specificate la forma, la media e l'errore standard della distribuzione campionaria della statistica rilevante.

Distribuzione *t* con 14 gradi di libertà.

In base all'ipotesi nulla, $E(\overline{Y})$ = 33.

Errore standard stimato = 24 / Sqrt[15] = 6.19677

Calcolate il valore che assume la statistica rilevante e, a parole, descrivete la procedura necessaria per trovare la probabilità richiesta dal problema.

$$t = \frac{\bar{Y} - E(\bar{Y})}{\hat{s}_{\bar{Y}}} = \frac{46 - 33}{6.19677} = 2.09787$$

E' necessario trovare l'area sottesa dalla distribuzione *t* con 14 gradi di libertà tra -∞ e 2.09787.