Psicometria 1 (023-PS)

Michele Grassi
mgrassi@units.it

Università di Trieste

Lezione 14

Piano della presentazione

Verifica di ipotesi statistiche

Verifica di ipotesi statistiche

- Con una stima intervallare viene specificato un intervallo che contiene il parametro (per esempio, la media μ della popolazione) al livello di confidenza $1 - \alpha$.
- Il test di ipotesi statistiche consente invece di stabilire la forza delle evidenze empiriche a sostengono dell'asserzione secondo cui il parametro possiede un determinato valore.
  - Come nel caso degli intervalli di confidenza, le procedure per il test di ipotesi statistiche sono piuttosto semplici, ma le motivazioni teoriche che stanno alla base di queste procedure sono più complesse.

Verifica di ipotesi statistiche

La logica della verifica di ipotesi statistiche verrà introdotta discutendo un esempio. Si consideri il seguente esperimento fittizio.

- Uno studio di pedagogia sperimentale intende stabilire se l'utilizzo del computer in un corso introduttivo di statistica migliori l'apprendimento.
- Vengono contattati dieci docenti che insegnano un corso introduttivo di statistica presso la Facoltà di Psicologia. Ciascun docente divide in maniera casuale i suoi studenti in due gruppi: un gruppo userà il computer, mentre l'altro gruppo non lo userà.
- A fine corso, lo stesso esame viene somministrato a tutti gli studenti.

Verifica di ipotesi statistiche

Il punteggio medio all'esame per ciascun gruppo è fornito nella tabella seguente, insieme alla differenza tra i punteggi dei gruppi che hanno utilizzato i due metodi d'insegnamento.

Verifica di ipotesi statistiche

Docente	Metodo computer	Metodo tradizionale	Differenza
1	94	71	23
2	75	70	5
3	75	58	17
4	84	80	4
5	79	70	9
6	85	67	18
7	73	66	7
8	75	80	-5
9	75	72	3
10	83	70	13
		|media	$\bar{x} = 9.4$
		|deviazione standard	$s = 8.38$

Verifica di ipotesi statistiche

La differenza media degli $n = 10$ dati è $\bar{x} = 9.4$ con una deviazione standard $s = 8.38$. Si noti che anche uno studio così semplice solleva dei seri problemi interpretativi.

Si noti che anche uno studio così semplice solleva dei seri problemi interpretativi.

- Se ci fosse un vantaggio per il nuovo metodo d'insegnamento, questo potrebbe essere dovuto al fatto che i docenti si impegnano maggiormente nell'utilizzarlo, anche se in sé tale metodo non è più efficace del metodo tradizionale.

Verifica di ipotesi statistiche

Lo scopo dello studio è stabilire se l'uso del computer migliori l'apprendimento in un corso introduttivo di statistica della Facoltà di Psicologia.

- Sia μ la differenza media tra i due metodi d'insegnamento nella popolazione.
- Se l'uso del computer migliorasse il punteggio degli studenti all'esame finale, allora $\mu > 0$.
- Se i due metodi d'insegnamento fossero egualmente efficaci, allora $\mu = 0$. (Per il momento tralasciamo l'ultima possibilità, ovvero che il metodo tradizionale sia migliore, ovvero $\mu < 0$.)

Ipotesi nulla

Ipotesi nulla

L'ipotesi nulla è, in generale, l'ipotesi che si vorrebbe rifiutare.

Nel caso presente, essa afferma che i due metodi di insegnamento sono egualmente efficaci:

$$H_0 : \mu = 0$$

- Si noti che l'ipotesi nulla specifica un particolare valore del parametro μ, ovvero $\mu = 0$.
- L'ipotesi nulla specifica sempre un particolare valore per il parametro μ, ma questo valore dipende dal contesto.

Ipotesi alternativa

Ipotesi alternativa

Contrapposta all'ipotesi nulla H_0 si ha l'ipotesi alternativa.

- Nel caso presente, l'ipotesi alternativa afferma che l'uso del computer produce risultati migliori del metodo d'insegnamento tradizionale

$$H_a : \mu > 0$$

- Diversamente dall'ipotesi nulla, l'ipotesi alternativa non specifica un valore particolare per il parametro μ.

Ipotesi alternativa

Dato che l'ipotesi alternativa non specifica un valore particolare per il parametro μ, l'ipotesi alternativa non può essere direttamente verificata.

- Ci poniamo invece il problema di stabilire quanto siano forti le evidenze empiriche a sostegno dell'ipotesi nulla.
- Se i dati del campione non forniscono una forte evidenza a favore dell'ipotesi nulla, allora possiamo rifiutare l'ipotesi nulla e accettare l'ipotesi alternativa.

Ipotesi alternativa

E' questa logica inversa che rende controintuitiva la procedura di verifica di ipotesi statistiche.

- Solitamente, il nostro interesse riguarda l'ipotesi alternativa.
- L'esame delle evidenze a favore dell'ipotesi nulla ci fornisce un metodo indiretto di valutazione delle evidenze a favore dell'ipotesi alternativa.

Ipotesi alternativa

- Inoltre, dell'ipotesi nulla NON stimeremo la probabilità che sia vera, a partire dai dati campionari.
- Per questo bisogna muoversi verso al statistica di Bayes (o Bayesiana)
- Calcoleremo invece la probabilità dei nostri dati campionari, assumendo come vera l'ipotesi nulla —che sebbene suoni come la proposizione sopra, non è assolutamente la stessa cosa.

Ipotesi alternativa

- Che non si calcolino probabilità che H_0 o H_a siano vere (o false), sarà una costante rispetto a tutte le procedure che vedremo.
- Invece, la terminologia e la notazione possono variare. Talvolta l'ipotesi alternativa è detta ipotesi sostantiva, oppure ipotesi sperimentale, e può essere denotata da H_1.

Test statistico

Test statistico

Si può definire test statistico una procedura che, sulla base di dati campionari e con un certo grado di fiducia, consente di decidere se è ragionevole respingere l'ipotesi nulla H_0 (ed accettare implicitamente l'ipotesi alternativa H_a) oppure se non esistono elementi sufficienti per respingerla.

Test statistico

La scelta sulle due ipotesi (H0 e Ha) è fondata sulla probabilità di ottenere per caso il risultato osservato nel campione o un risultato ancor più distante da quanto atteso, nella condizione che l'ipotesi nulla H0 sia vera
Se questa probabilità stimata è molto piccola, allora decideremo che è ragionevole respingere H_0.

Test statistico

- In un test statistico sono poste a confronto due ipotesi :

  H_0 : formulata in maniera specifica; ad es. $\mu = 0$ e

  H_a : formulata in modo generico, complementare ad H_0.

  sulla base di una generica statistica test T;
- il rifiuto dell'una implicitamente ci porta ad accettare l'altra.
- Nel caso presente, la statistica test è la media del campione.

Test statistico

Si conviene di rifiutare H_0, quando l'evento che si è verificato (*ed eventi più estremi di esso*) ha, *sotto quella ipotesi*, una probabilità di accadimento inferiore a un livello prefissato α (detto livello di significatività).

Nel caso presente, la probabilità della statistica test è fornita dalla distribuzione campionaria della media del campione $\bar{x}$.

Se fosse vera l'ipotesi nulla ($\mu = 0$), le medie dei campioni di dimensioni $n = 10$ si distribuirebbero in maniera approssimativamente normale con media $\mu = 0$ e varianza $\sigma/\sqrt{n} = \sigma/\sqrt{10}$.

Test statistico

Ci sono due problemi che per il momento ignoriamo.

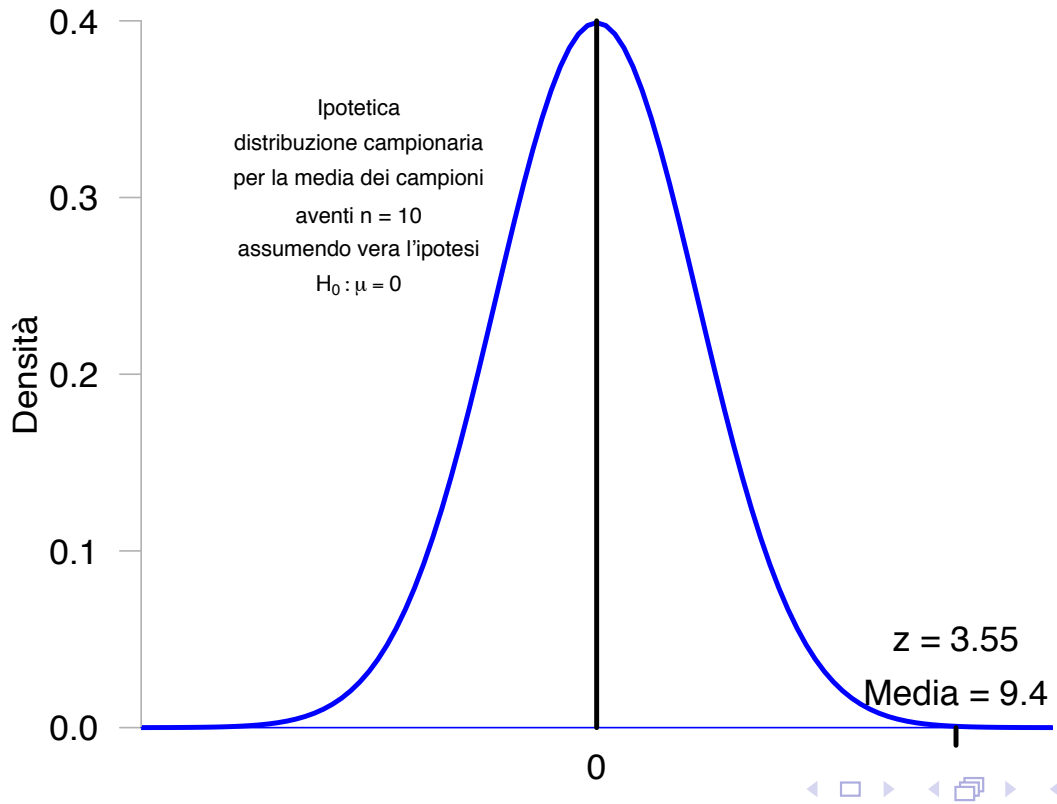
1. La deviazione standard σ della popolazione non è conosciuta. Per il momento utilizzeremo $s = 8.38$ per stimare σ, ma questa non è una buona idea quando n è piccolo.

2. Dato che il campione è piccolo, la distribuzione delle medie $\bar{x}$ potrebbe non essere sufficientemente simile alla distribuzione normale (nel caso in cui la distribuzione della popolazione sia molto diversa dalla normale).

Test statistico

Trascurando per ora questi due problemi, sotto l'ipotesi nulla la statistica test $\bar{x}$ si distribuisce come una variabile aleatoria N con $\mu = 0$ e $\widehat{\sigma} = 8.38/\sqrt{10}$:

$$\bar{x} \overset{H_0}{\sim} N(0,\ 8.38/\sqrt{10})$$



Test statistico

Nella condizione che l'ipotesi nulla H_0 sia vera, dunque, la probabilità di osservare un campione con media uguale a $\bar{x} = 9.4$, o più grande, corrisponde all'area sottesa alla distribuzione normale nell'intervallo $[9.4, \infty]$:

$$z = \frac{\bar{x} - \mu_0}{\widehat{\sigma}/\sqrt{n}} = \frac{9.4 - 0}{8.38/\sqrt{n}} = 3.55$$

$p-$valore

- Nella formula precedente,$\mu = 0$ non è il valore reale del parametro sconosciuto μ, ma il valore specificato da H_0.
- Dunque, nell'ipotesi che H_0 sia vera

$$P(\bar{x} \geq 9.4) = P(z \geq 3:55) = 1 - 0.9998 = 0.0002$$

- Questa probabilità è chiamata $p-$valore della statistica test.

Interpretazione

- La probabilità di ottenere una media campionaria pari a $\bar{x} = 9.4$ o maggiore di quella sperimentalmente osservata, nella condizione che l'ipotesi nulla H_0 sia vera, è molto piccola – circa due possibilità in 10000.
- E' dunque improbabile che l'ipotesi nulla sia vera.
- Pertanto, rifiutiamo l'ipotesi nulla e *implicitamente* accettiamo l'ipotesi alternativa concludendo che *l'uso del computer migliora ($\bar{x} = 9.4 > 0$) l'apprendimento nei corsi introduttivi di statistica.*

Procedura di verifica delle ipotesi

Procedura di verifica delle ipotesi

Contrapposta all'ipotesi nulla H_0 si ha l'ipotesi alternativa H_a. L'ipotesi alternativa può essere di tre tipi, tra loro mutuamente esclusivi:

bilaterale:

$$H_0 : \mu = \mu_0 \quad H_a : \mu \neq \mu_0$$

unilaterale destra:

$$H_0 : \mu \leq \mu_0 \quad H_a : \mu > \mu_0$$

unilaterale sinistra:

$$H_0 : \mu \geq \mu_0 \quad H_a : \mu < \mu_0$$

Procedura di verifica delle ipotesi

- La prima delle tre alternative precedenti è non-direzionale; le ultime due specificano invece degli scostamenti direzionali dall'ipotesi nulla.
- In qualsiasi applicazione concreta della procedura di verifica di ipotesi statistiche viene usata solo una delle possibili versioni dell'ipotesi alternativa.
- Si noti che le ipotesi nulla e alternativa sono specificate nei termini del parametro d'interesse (nel caso presente, μ).
  - Sarebbe errato, per esempio, scrivere

$$H_0 : \bar{x} = 0.$$

Procedura di verifica delle ipotesi

Usando la distribuzione campionaria della statistica test sotto l'ipotesi nulla

$$\bar{x} \overset{H_0}{\sim} N\left(\mu_0, \frac{\sigma}{\sqrt{n}}\right)$$

si trova la probabilità di ottenere per caso il risultato osservato nel campione o un risultato ancor più distante da quanto atteso.

Procedura di verifica delle ipotesi

Il p-valore della statistica test

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}$$

può essere trovato utilizzando la tabella della distribuzione normale standardizzata oppure un software.

```
 z<-  ( 9.4 - 0 ) / ( 8.38/sqrt(10) );  z
## [1] 3.547185
 1-pnorm(9.4,  mean=0,  sd=8.38/sqrt(10) )
## [1] 0.0001946856
 1-pnorm(z,  mean=0,  sd=1 )
## [1] 0.0001946856
```

Distribuzione dell'indice statistico

L'insieme di valori ottenibili con il test formano la distribuzione campionaria dell'indice statistico. Essa può essere divisa in due zone:

1. la regione di rifiuto dell'ipotesi nulla, detta anche regione critica, che corrisponde ai valori collocati agli estremi della distribuzione secondo la direzione dell'ipotesi alternativa H_a; sono quei valori che hanno una probabilità piccola di verificarsi per caso, quando l'ipotesi nulla H_0 è vera;

2. la regione di non rifiuto di H_0, che comprende i restanti valori, quelli che si possono trovare abitualmente per effetto della variabilità casuale.

Distribuzione dell'indice statistico

Se il valore dell'indice statistico calcolato cade nella zona di rifiuto, si respinge l'ipotesi nulla H_0.

Distribuzione dell'indice statistico

Si noti che il $p-$valore non rappresenta la probabilità che l'ipotesi nulla sia vera.

- L'ipotesi nulla può essere vera oppure no – μ è uguale a μ_0 oppure no – ma non abbiamo modo di saperlo.
- La situazione è simile a quella che si dava a proposito dell'interpretazione degli intervalli di fiducia, laddove il livello di fiducia non era uguale alla probabilità che il parametro μ fosse contenuto nello specifico intervallo considerato.

Distribuzione dell'indice statistico

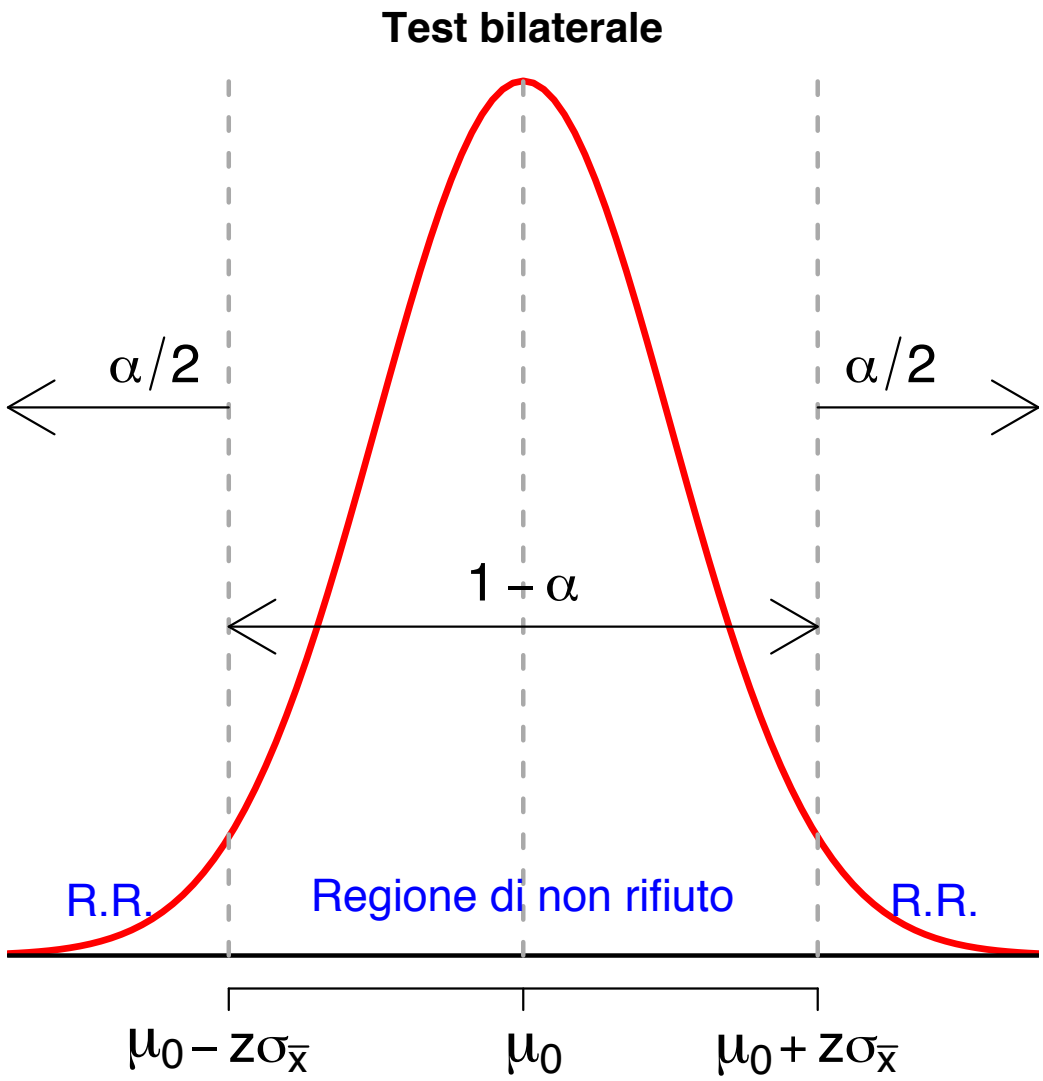
Il $p-$valore rappresenta invece la probabilità di ottenere per caso (a causa della variabilità campionaria) il valore osservato della statistica test, o un valore ancora più estremo, nella condizione che l'ipotesi nulla sia vera.

Distribuzione dell'indice statistico

L'ipotesi alternativa H_a, da verificare con un test, può essere bilaterale o unilaterale.

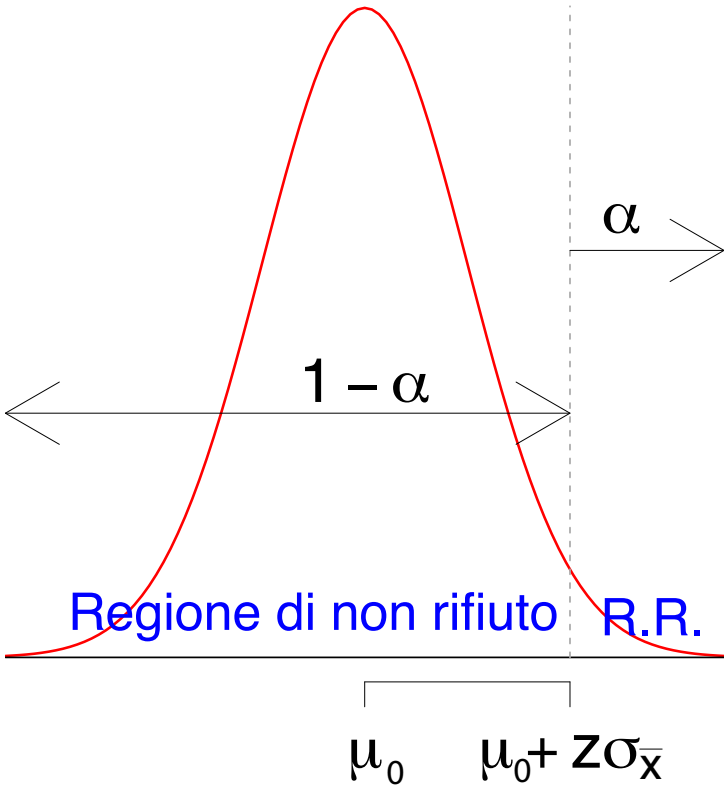
- E' bilaterale quando ci si chiede se tra la media del gruppo A (nel caso presente, metodo computer) e quella del gruppo B (metodo tradizionale) esiste una differenza significativa, senza sapere a priori quale sia maggiore (o minore).
- E' unilaterale quando è possibile escludere a priori, come privo di significato e risultato solo di errori nella conduzione dell'esperimento, il fatto che la media di B possa essere minore o maggiore della A.

Distribuzione dell'indice statistico

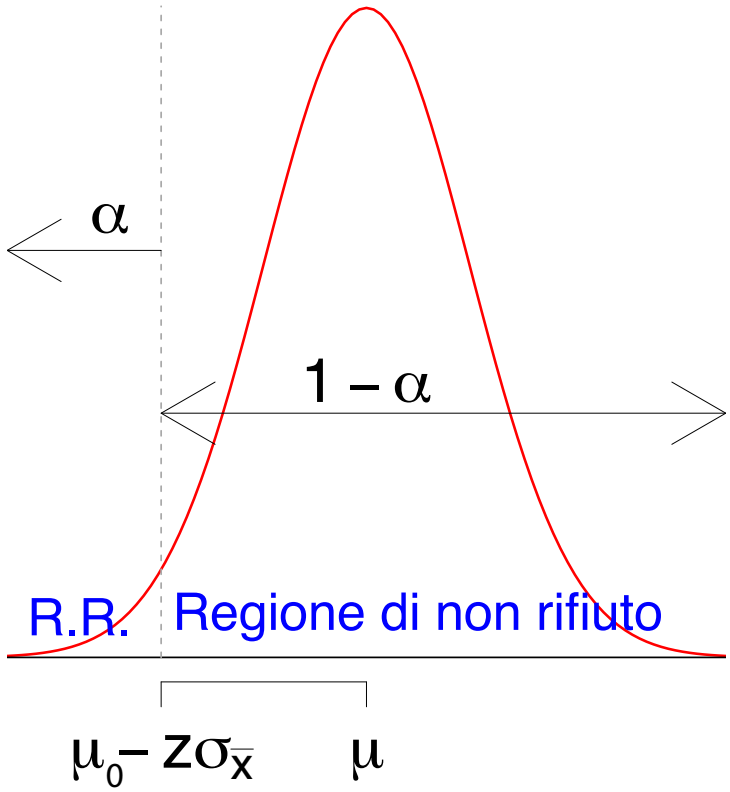


Distribuzione dell'indice statistico

Test unilaterale destro



Test unilaterale sinistro



Test unilaterale

Si ha un test unilaterale, per esempio, quando si confrontano i risultati di un farmaco con il placebo.

- L'unica domanda razionale è se gli individui ai quali è stato somministrato il farmaco abbiano risultati migliori nella sopravvivenza di coloro ai quali è stato somministrato il placebo.
- Se il gruppo al quale è stato somministrato il farmaco avesse un risultato medio minore dell'altro, sarebbe illogico e quindi sarebbe inutile proseguire l'analisi, con qualunque test statistico.

Test bilaterale

Si ha un test bilaterale, per esempio, quando si confrontano i risultati di due metodi di insegnamento, come nel caso presente, per valutare quale abbia l'effetto migliore.

- In questo caso ci sono due alternative, che lo sperimentatore ritiene ugualmente logiche e possibili.

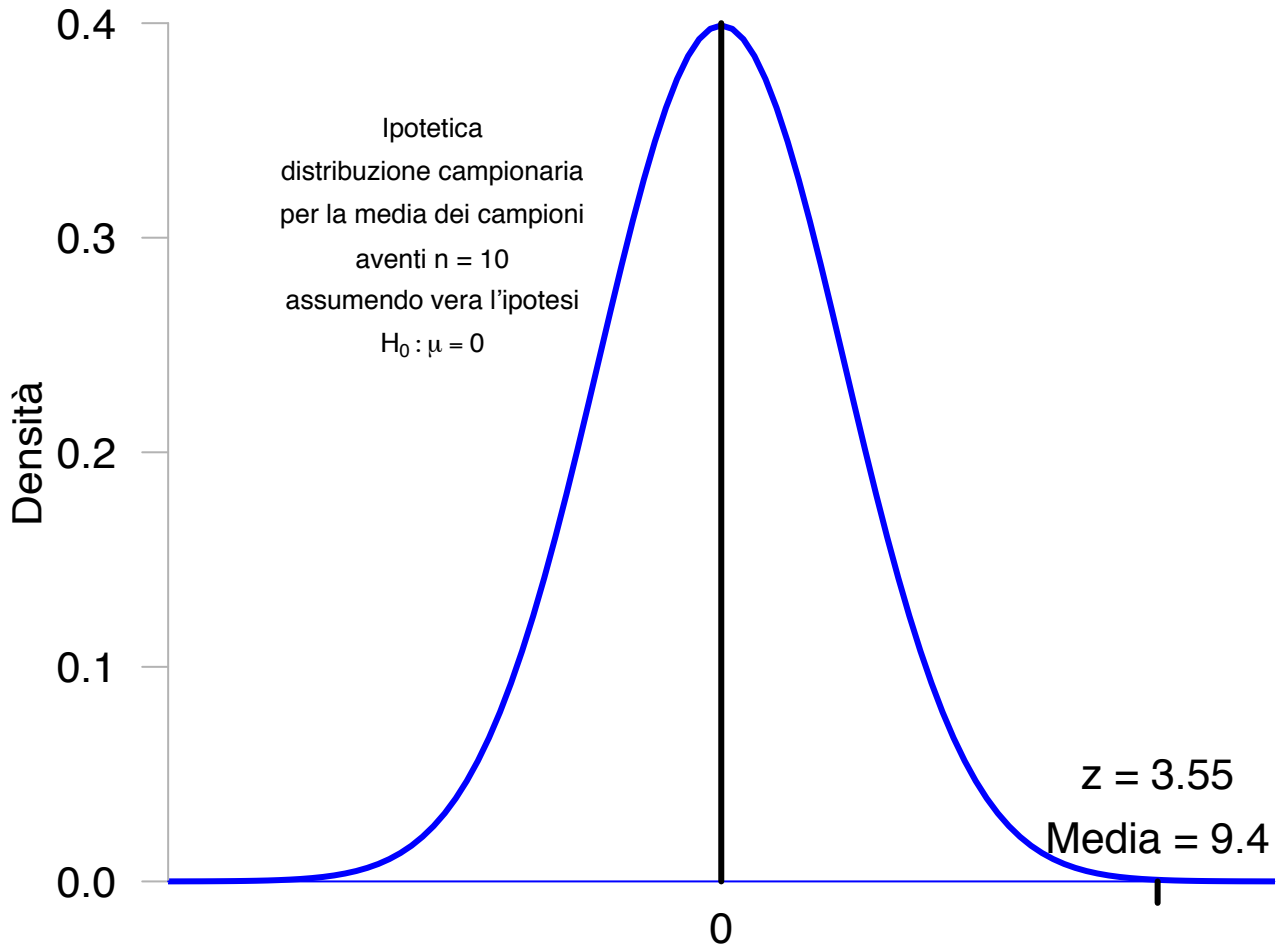
Test unilaterale

- In un test unilaterale la zona di rifiuto dell'ipotesi nulla è solamente in una coda della distribuzione;
- in un test bilaterale essa è equamente divisa nelle due code della distribuzione.
  - Per l'esempio presente, $z = 3.55$ e il $p-$valore di un test unilaterale $(H_a : \mu > 0)$ è
$$P\left[Z \geq 3.55\right] = 0.0002$$
  - il $p-$valore di un test bilaterale $(H_a : \mu \neq 0)$ è
$$P\left[(Z \leq -3.55) \cup (Z \geq 3.55)\right] 2 \times 0.0002 = 0.0004$$

Test unilaterale

- Se l'ipotesi alternativa fosse stata $H_a : \mu < 0$ – il metodo d'insegnamento tradizionale produce risultati migliori –, allora avremmo dovuto calcolare l'area sottesa alla normale nell'intervallo $[-\infty, 9.4]$.
- In questo caso, il $p-$valore sarebbe molto grande, ovvero $P = 0.9998$, il che ci avrebbe portato a non rifiutare H_0: il metodo d'insegnamento che fa uso del computer produce risultati uguali a quello tradizionale (o addirittura migliori).

Test unilaterale



Valori critici

- I valori critici per l'area in una coda alla probabilità α coincidono con quelli della probabilità $2 \times \alpha$ nella distribuzione a due code.
- Viceversa, quelli associati alla probabilità α in due code coincidono con i valori associati alla probabilità $\alpha/2$ nella distribuzione a una coda.
- Si dice che il test a due code è più conservativo, mentre quello ad una coda è più potente.

Significatività statistica

Significatività statistica

Il risultato di un test si dice statisticamente significativo quando il $p-$valore è sufficientemente piccolo da consentire di rifiutare l'ipotesi nulla.

- Quanto piccolo deve essere il $p-$valore ?

  Convenzionalmente, i livelli di soglia delle probabilità ai quali di norma si ricorre sono tre: $0.05(5\%)$; $0.01(1\%)$; $0.001(0.1\%)$.

- La probabilità prescelta per rifiutare l'ipotesi nulla si chiama livello di significatività, indicato generalmente con α.

Significatività statistica

Si parla spesso di significatività statistica e dunque è necessario capire il significato di questo concetto.

> *Un risultato statisticamente significativo è un risultato estremamente improbabile, nell'eventualità che l'ipotesi nulla H_0 sia vera.*
>
> *Ciò non significa, però, che sia necessariamente di una qualche importanza pratica.*

In un campione molto grande, per esempio, anche uno scostamento molto piccolo di $\bar{x}$ da μ_0 può facilmente risultare statisticamente significativo.

Una nota di cautela

Non è corretto ripetere più volte il test descritto in precedenza, considerando varie ipotesi statistiche.

- Supponete di eseguire 100 test, ciascuno al livello $\alpha = 0.05$.
- Benché per ciascun test la probabilità di rifiutare l'ipotesi nulla, quando H_0 è vera, sia uguale a 0.05, la probabilità di rifiutare almeno un'ipotesi nulla tra le 100 verificate è molto più grande del 5%.
- Esamineremo in seguito l'analisi della varianza che fornisce una risposta a questo problema.

Riportare il $p-$valore di un test

- I software riportano la probabilità esatta della statistica test.
- Tuttavia, quando il ricercatore riporta i risultati di un test statistico, per convenzione il $p-$valore viene approssimato per eccesso utilizzando uno dei seguenti livelli di soglia:

$$0.05 \quad 0.01 \quad 0.001$$

- Per l'esempio precedente, nel caso di un test bilaterale, si dirà:

La differenza tra i due metodi d'insegnamento è risultata significativa ($z = 3.55$, $p < 0.001$).

Test di ipotesi e intervalli di confidenza

Test di ipotesi e intervalli di confidenza

C'è una stretta relazione tra verifica delle ipotesi e intervalli di confidenza.

- La stima dell'intervallo di confidenza della media sconosciuta della popolazione fornisce due valori (indicati con L_1 e L_2) detti limiti fiduciali che comprendono l'intervallo di confidenza.
- Il calcolo dell'intervallo di confidenza serve anche per il test d'inferenza e fornisce esattamente le stesse conclusioni.

Test di ipotesi e intervalli di confidenza

• Se l'intervallo di confidenza al livello $1 - \alpha$ contiene il valore μ_0 espresso nell'ipotesi nulla $H_0 : \mu = \mu_0$ di un test bilaterale, non esistono evidenze sufficienti per respingere H_0 al livello di significatività α.

• Se l'intervallo di confidenza non contiene il valore μ_0 espresso nell'ipotesi nulla, esistono elementi sufficienti per rifiutare H_0 al livello di significatività α.

Test di ipotesi e intervalli di confidenza

- **Test d'ipotesi**: l'ipotesi nulla $H_0 : \mu = \mu_0$ non viene rifiutata ($\alpha = 0.05$, test bilaterale) se

$$-1.96 < \frac{\bar{x} - \mu_0}{\hat{\sigma}_{\bar{x}}} < 1.96$$

- **Intervallo di confidenza**: l'ipotesi nulla $H_0 : \mu = \mu_0$ non viene rifiutata se

$$\bar{x} - 1.96\hat{\sigma}_{\bar{x}} < \mu_0 < \bar{x} + 1.96\hat{\sigma}_{\bar{x}}$$

Le due asserzioni precedenti sono equivalenti dato che la prima può essere trasformata nella seconda:

$$-1.96 < \frac{\bar{x} - \mu_0}{\hat{\sigma}_{\bar{x}}} < 1.96$$

$$-1.96\hat{\sigma}_{\bar{x}} < \bar{x} - \mu_0 < 1.96\hat{\sigma}_{\bar{x}}$$

$$-\bar{x} - 1.96\hat{\sigma}_{\bar{x}} < -\mu_0 < -\bar{x} + 1.96\hat{\sigma}_{\bar{x}}$$

$$\bar{x} + 1.96\hat{\sigma}_{\bar{x}} > \mu_0 > \bar{x} - 1.96\hat{\sigma}_{\bar{x}}$$

Verifica di ipotesi e decisione

Verifica di ipotesi e decisione

- La teoria statistica consente di prendere decisioni razionali in una situazione di incertezza. Ciò non significa che tali decisioni siano immuni da errori.
- Piuttosto, la teoria statistica esplicita la probabilità di errare, associata ad ogni scelta.
- Mediante un test statistico non si perviene dunque ad una affermazione assoluta, ma ad una probabilità conosciuta di poter commettere un errore.

Verifica di ipotesi e decisione

Supponiamo che l'ipotesi nulla sia $H_0 : \mu = \mu_0$, laddove μ_0 è un qualsiasi valore (per esempio, 0).

Ci sono due possibilità: l'ipotesi nulla H_0 può essere vera oppure falsa. Ovviamente non sappiamo quale delle due possibilità sia quella corretta, altrimenti non avremmo bisogno di ricorrere ad un test statistico.

Sulla base delle informazioni fornite da un campione, lo sperimentatore può decidere di

non rifiutare H_0 (se la statistica z non eccede il valore prefissato, diciamo 1.96 per un test bilaterale al livello $\alpha = 0.05$);

rifiutare H_0 (se la statistica test eccede il valore critico).

Verifica di ipotesi e decisione

- Due tipi di errore sono possibili.
  - L'errore di I tipo consiste nel rifiutare l'ipotesi nulla H0, quando in realtà essa è vera.
  - L'errore di II tipo consiste nel non rifiutare l'ipotesi nulla H_0, quando in realtà essa è falsa.
- Un test statistico conduce ad una conclusione esatta in due casi:
  - se non rifiuta l'ipotesi nulla, quando in realtà è vera;
  - se rifiuta l'ipotesi nulla, quando in realtà è falsa.

Verifica di ipotesi e decisione

- La probabilità di commettere l'errore di I tipo è chiamata livello di significatività ed è indicata convenzionalmente con α.

  • Essa corrisponde alla probabilità che il valore campionario dell'indice statistico cada nella zona di rifiuto, quando l'ipotesi nulla è vera.

- La probabilità di commettere l'errore di II tipo, indicato convenzionalmente con β, è la probabilità di estrarre dalla popolazione un campione che non permette di rifiutare l'ipotesi nulla, quando in realtà essa è falsa.

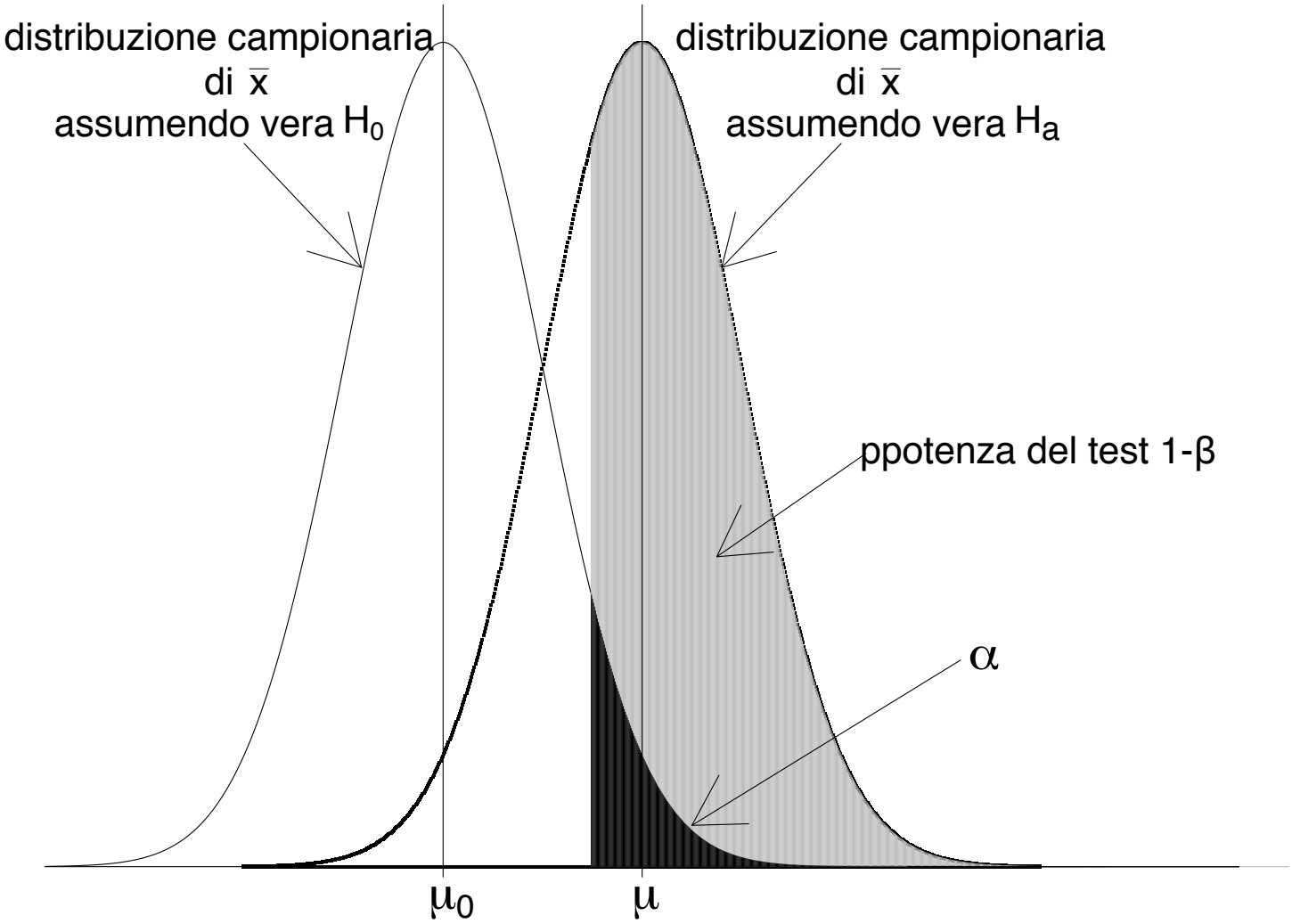
Verifica di ipotesi e decisione

- Dai concetti di errore di I e II tipo derivano direttamente anche quelli di livello di protezione e di potenza di un test.
- Sono concetti tra loro legati, secondo lo schema riportato nella tabella seguente, che confronta la realtà con la conclusione del test.

Verifica di ipotesi e decisione

conclusione	realtà: H_0 *vera*	realtà: H_0 *falsa*
H_0 *vera*	decisione corretta $P = 1 - \alpha$ Protezione	errore di II tipo $P = \beta$
H_0 *falsa*	errore di I tipo $P = \alpha$ Significatività	decisione corretta $P = 1 - \beta$ Potenza

Verifica di ipotesi e decisione

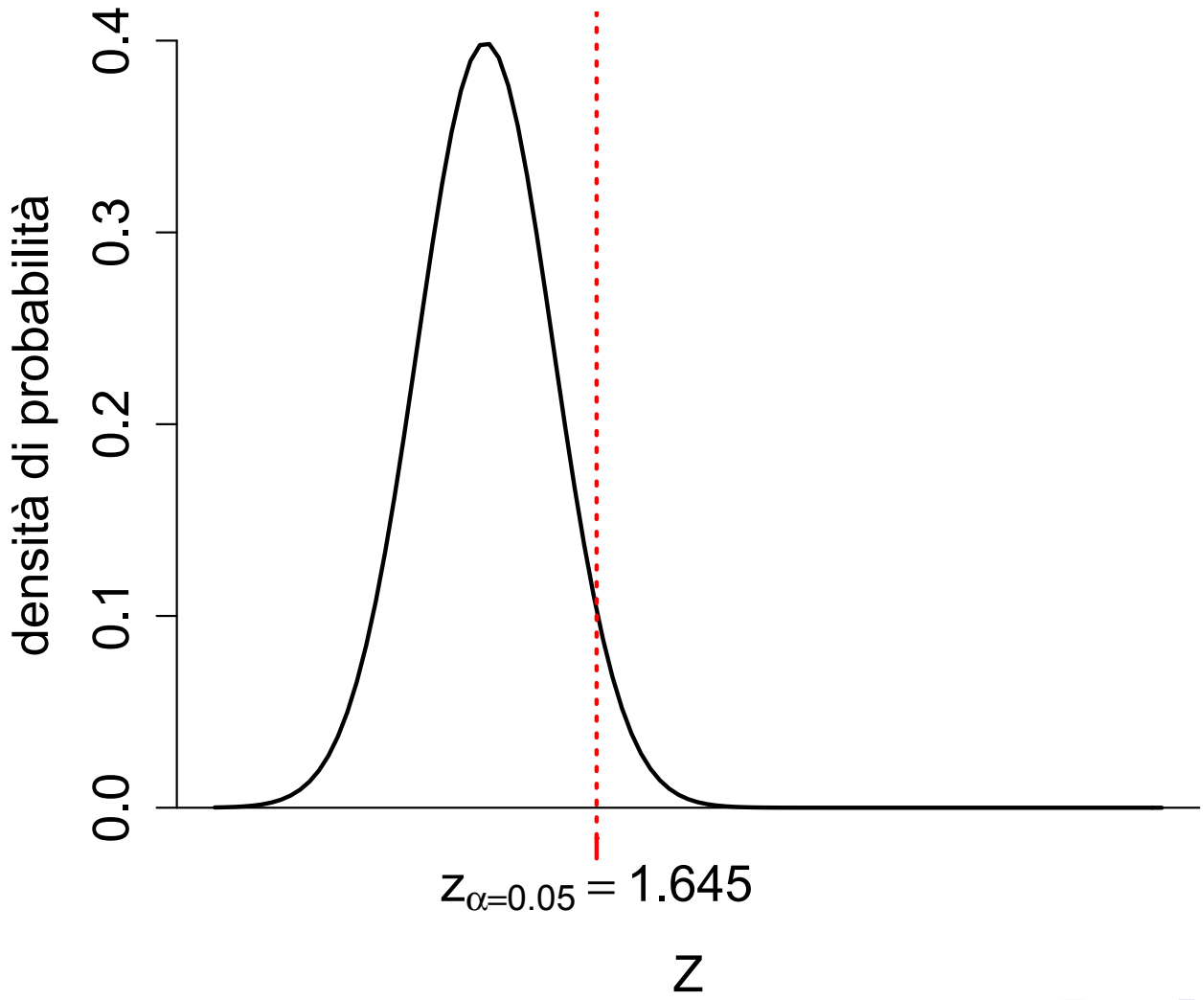


Verifica di ipotesi e decisione

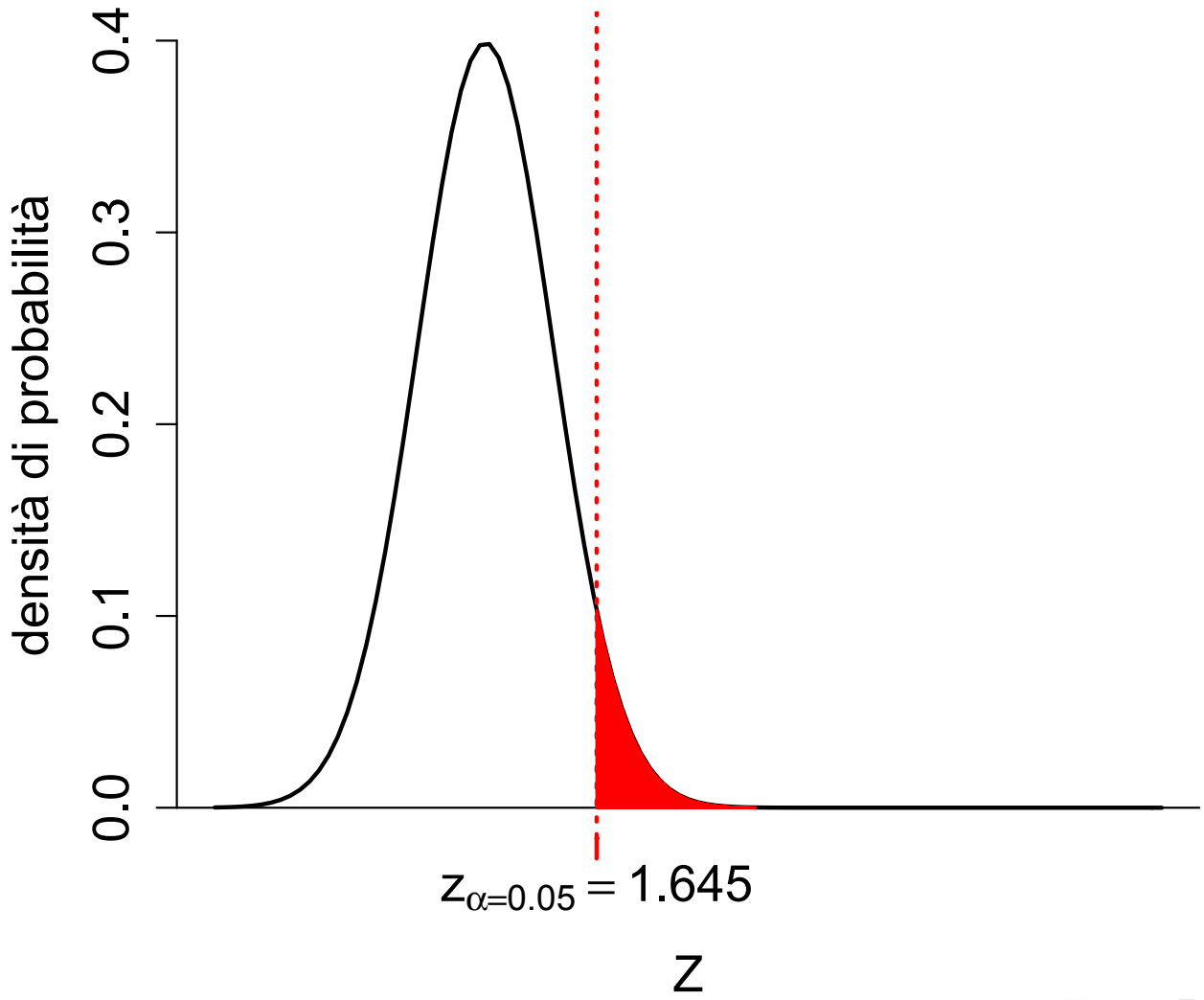
La potenza del test aumenta al crescere

1. *del livello α del test,*
2. *della differenza tra μ_a e μ_0,*
3. *della numerosità n del campione.*

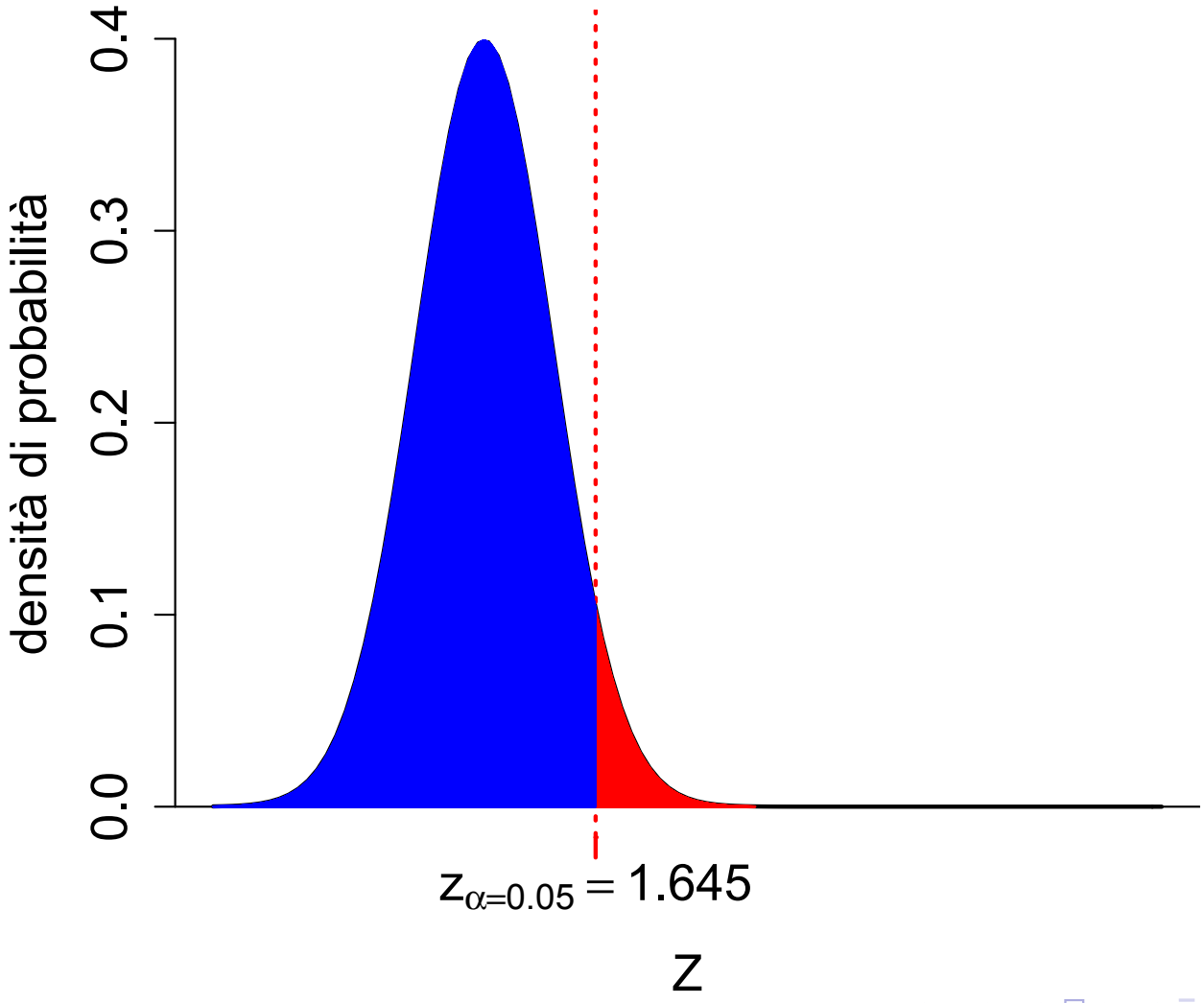
Verifica di ipotesi e decisione



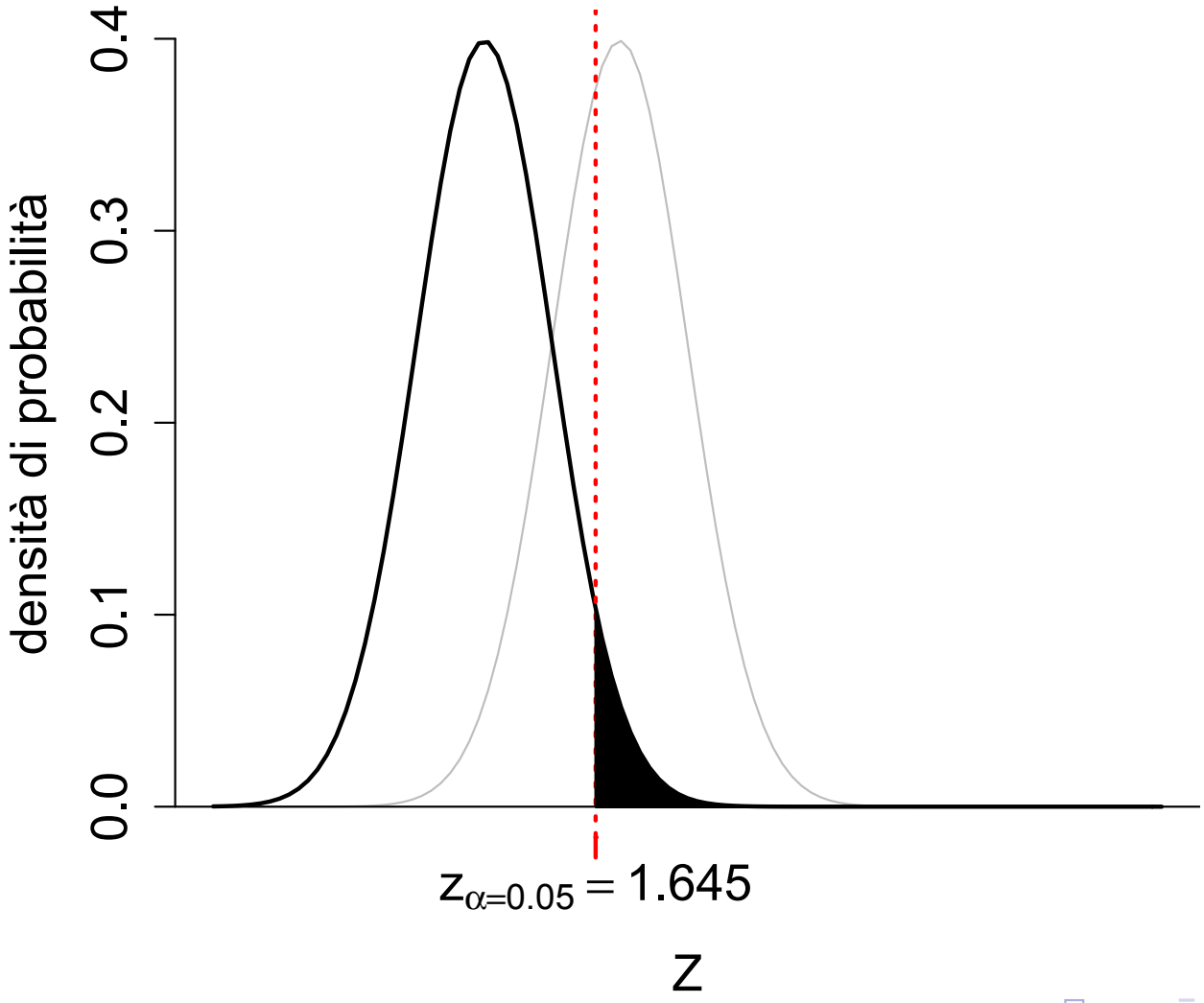
Verifica di ipotesi e decisione



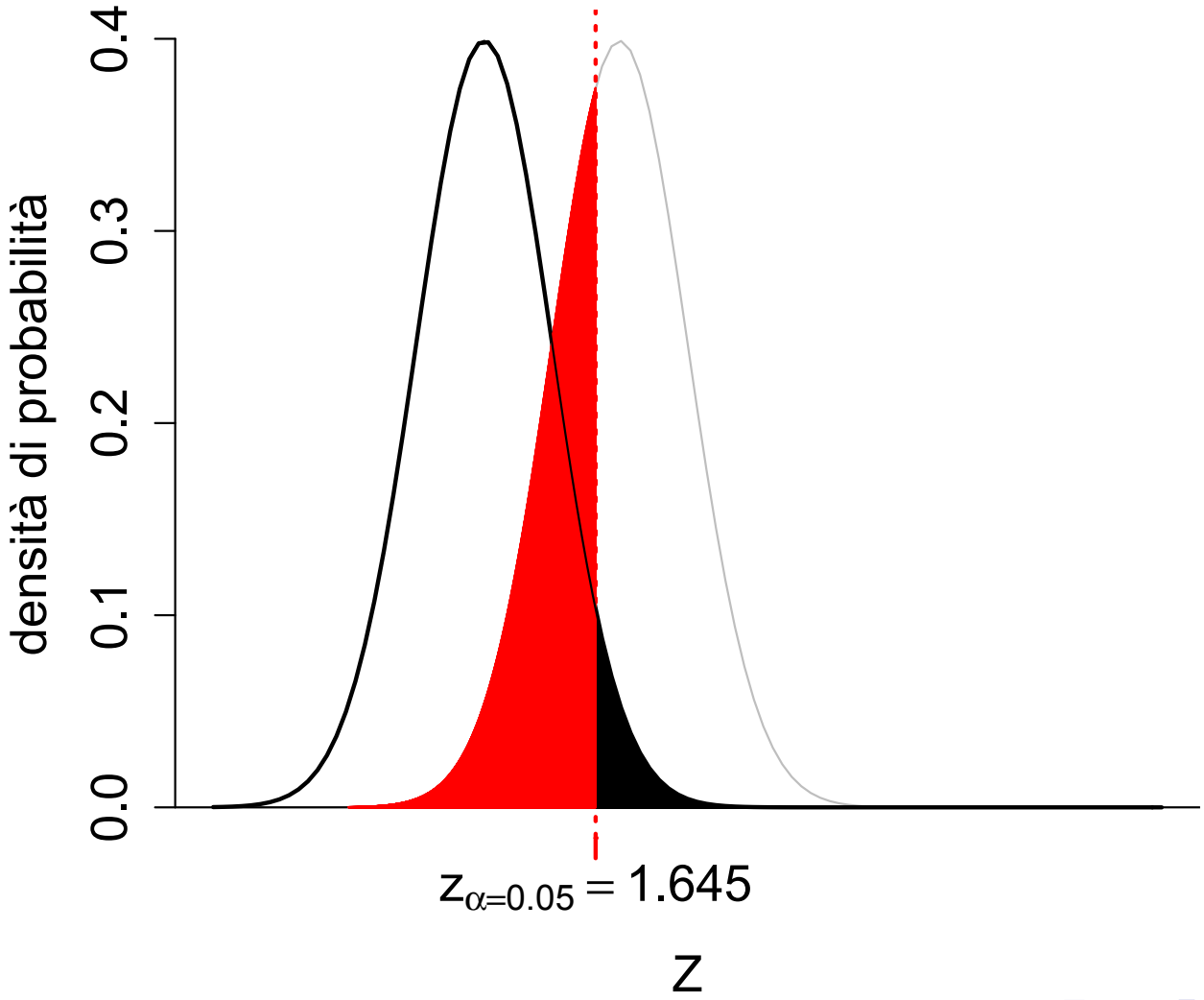
Verifica di ipotesi e decisione



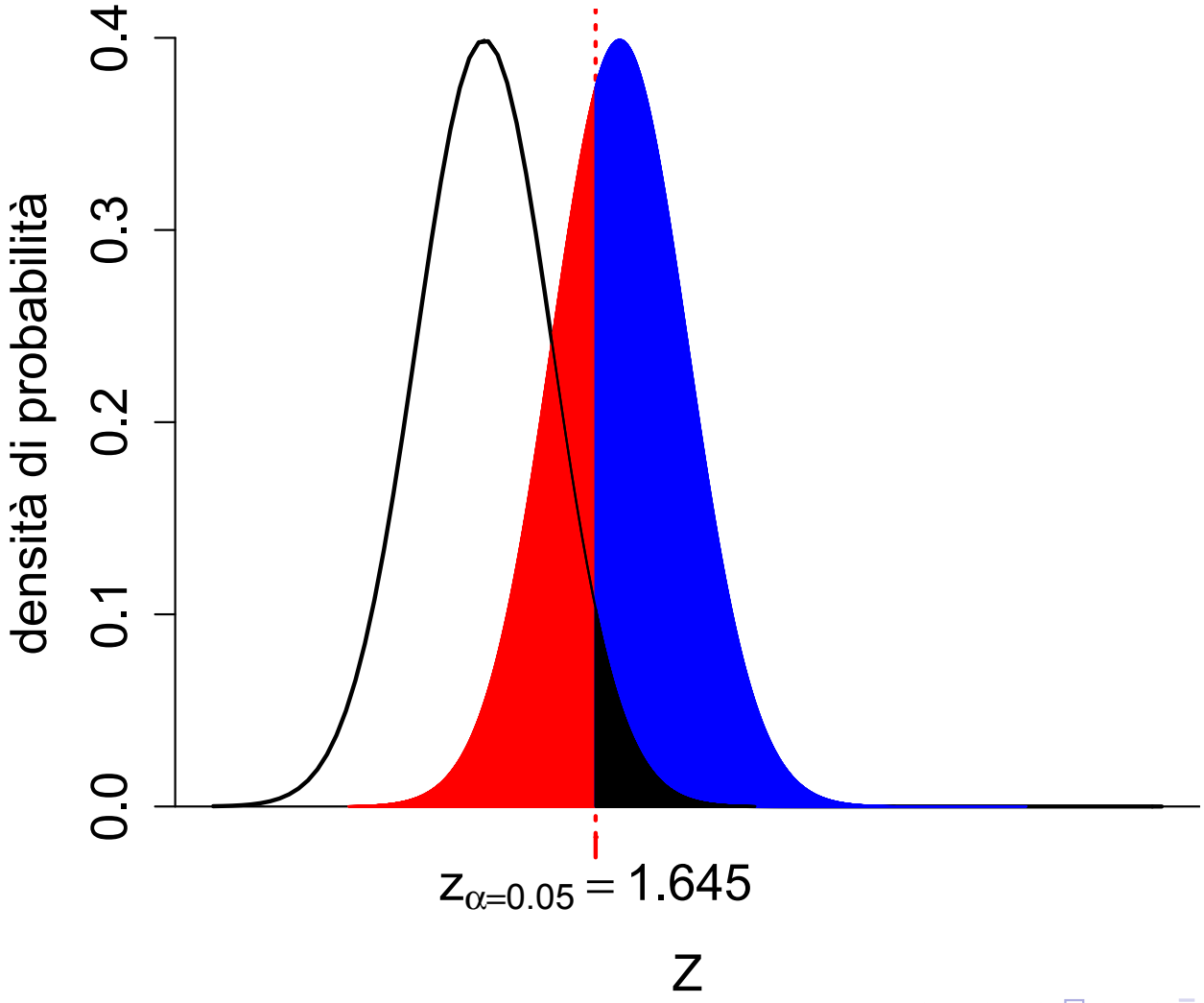
Verifica di ipotesi e decisione



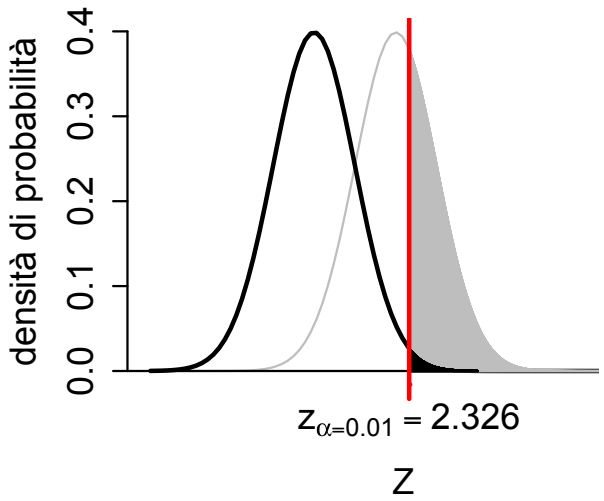
Verifica di ipotesi e decisione

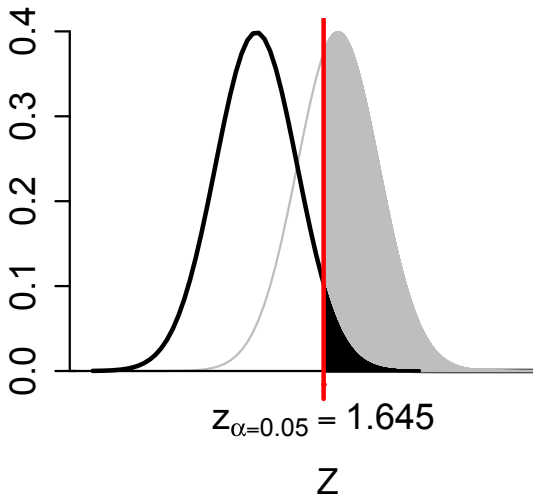


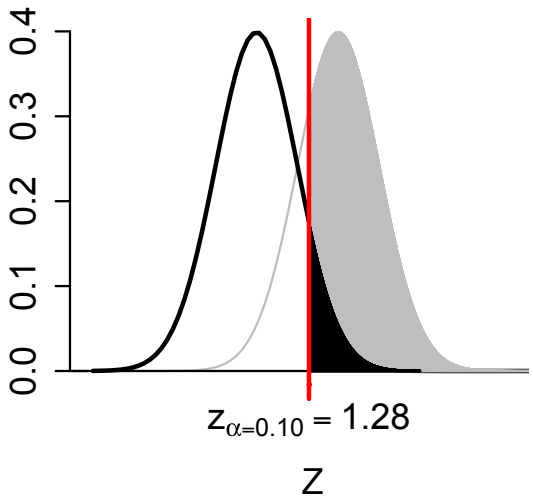
Verifica di ipotesi e decisione



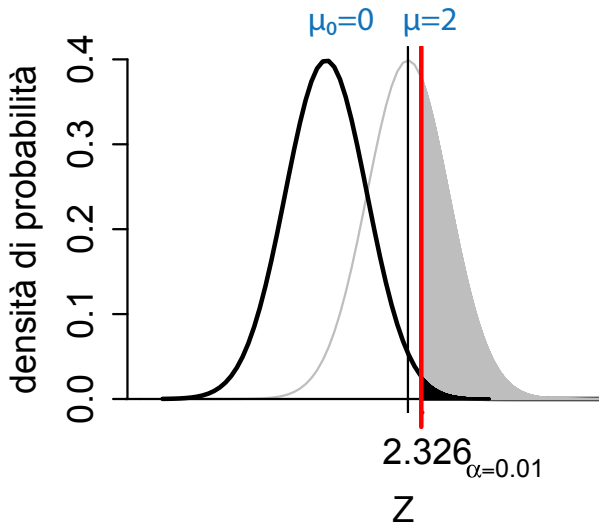
Verifica di ipotesi e decisione

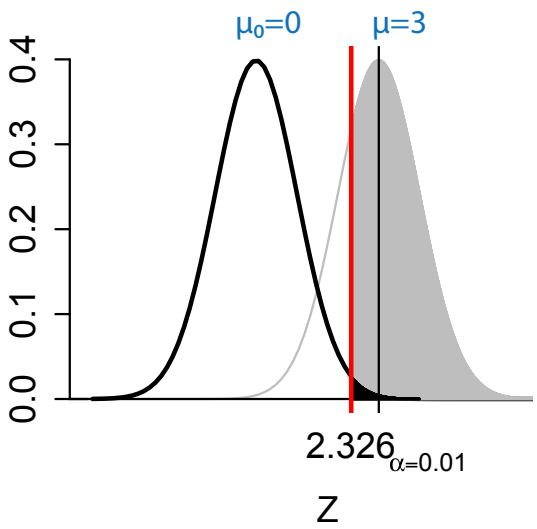


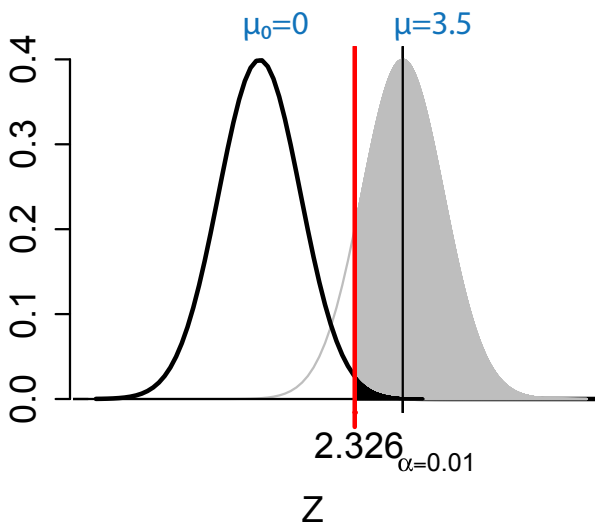




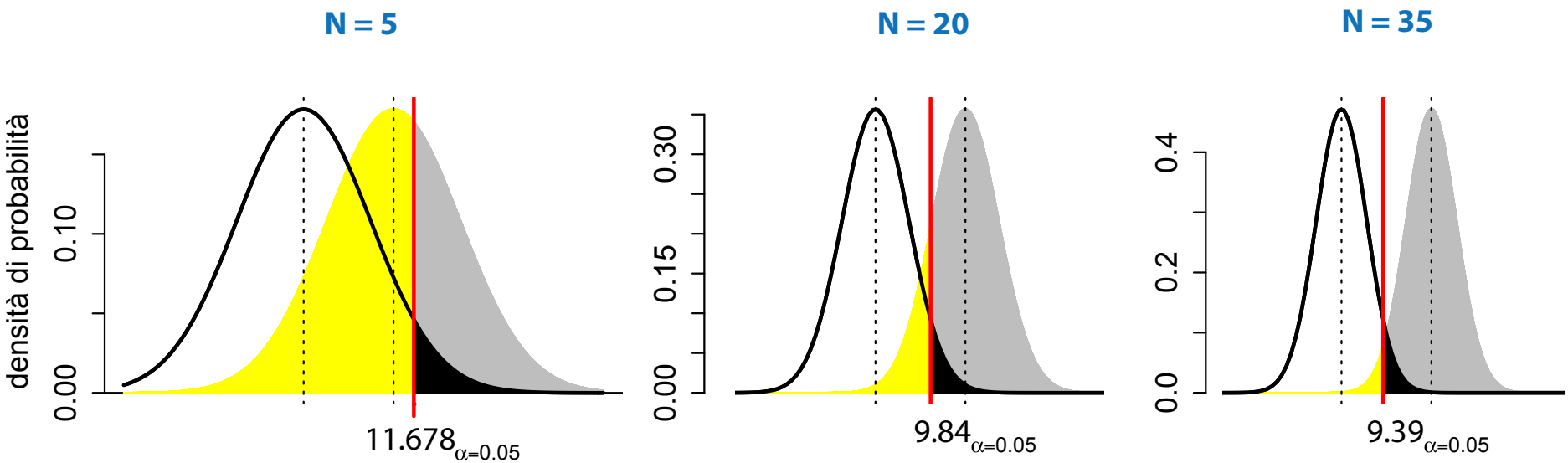
Verifica di ipotesi e decisione







Verifica di ipotesi e decisione



Conclusioni

-L'inferenza statistica produce delle predizioni sui parametri (sconosciuti) della popolazione sulla base delle informazioni fornite da un campione casuale di osservazioni.

- I test statistici sono composti da cinque elementi:
—- assunzioni sul tipo di campionamento, la forma della distribuzione della popolazione, la scala di misura delle variabili;
—- ipotesi nulla e alternativa;
—- statistica test;
—- p-valore della statistica test;
—- conclusione.

Conclusioni

- Quando una decisione viene presa, sono possibili due tipi di errore:
1. errore di I tipo (rifiutare H_0 quando è vera);
2. errore di II tipo (non rifiutare H_0 quando è falsa).

- La stima di un intervallo di confidenza è più informativa del test di un'ipotesi statistica:
—- un test statistico consente di decidere se è ragionevole assegnare un certo valore ad un parametro;
—- un intervallo di confidenza individua un intervallo di valori entro il quale si trova il parametro alla probabilità $(1 - \alpha)$.

Conclusioni

- Le dimensioni del campione sono un fattore critico per gli intervalli di confidenza e per i test statistici.

- Se n è piccolo,
  - la stima è imprecisa (l'intervallo di confidenza è grande);
  - è difficile rifiutare H_0 a meno che μ_0 non sia molto diversa da μ_a;
  - la probabilità di commettere un errore di II tipo è grande.

Conclusioni

- Gli intervalli di confidenza servono anche per il test d'inferenza e forniscono esattamente le stesse conclusioni.

—- Per un test bilaterale, se l'intervallo di confidenza alla probabilità α contiene il valore espresso nell'ipotesi nulla H_0, non esistono prove sufficienti per respingerla, ad una probabilità $P < \alpha$.
—- Se l'intervallo di confidenza non lo contiene, esistono elementi sufficienti per rifiutare l'ipotesi nulla H_0, con una probabilità $P < \alpha$.