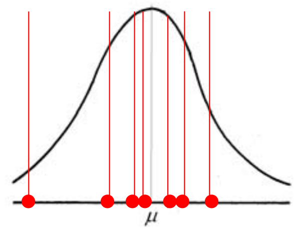
VERIFICA DI IPOTESI

VERIFICA DI IPOTESI STATISTICA

- La prima cosa da capire è *perché* serva la statistica per verificare una ipotesi
- Una ipotesi scientifica è un enunciato, una affermazione, su un fenomeno osservabile direttamente o indirettamente, almeno in linea di principio, nel mondo fisico
- Una ipotesi scientifica quindi deve potere essere oggetto di verifica empirica: occorre verificare se i dati osservativi risultano in accordo con quanto previsto dall'ipotesi (teoria, modello, ...)
- Popper K. nel 1934 chiarisce definitivamente la questione scienza/non-scienza, dettando quello che diventa il principio di scientificità della moderna epistemologia: il principio di falsificabilità
  - solo le ipotesi falsificabili sono di interesse della scienza, tutte le altre teorie appartengono alla metafisica (es. astrologia, mitologia, religioni, ...)
  - alla metafisica, cioè alle non-scienze, non si applica la categoria del vero/falso: se un enunciato **non può** essere falso, **non può nemmeno essere vero**
- La verifica dell'ipotesi presenta una **asimmetria logica**:
  - una ipotesi non può mai essere definitivamente verificata: perché le sue conseguenze sono logicamente infinite ed esiste sempre la possibilità che una nuova osservazione risulti in disaccordo
  - invece basta una sola osservazione contraria all'ipotesi perché essa sia, dal punto di vista logico, definitivamente falsificata

VERIFICA DI IPOTESI

- Dunque una ipotesi può essere temporaneamente confermata (*vera*) o definitivamente falsificata (*falsa*): dunque cosa c'entra la statistica ?
- Il principio di Popper è logicamente ineccepibile: ma la sua verifica di ipotesi è deterministica, assume di essere in presenza di dati certi, ovvero in assenza di errori di misura/rilevazione/osservazione
- Purtroppo però anche i dati empirici non sono mai certi: sono affetti dall'errore di misura, e sono in una certa misura determinati dal **caso**: quel complesso di piccoli fattori indipendenti e singolarmente irrilevanti, che però in alcuni casi (con bassa probabilità) possono anche produrre un dato *osservato* lontano dal *vero*



$$\hat{x} = \mu + \varepsilon \quad con \quad \varepsilon \sim N(0,\sigma)$$

- Si pone quindi il problema della **validità statistica** dei risultati empirici: è il grado di probabilità che i risultati osservati non siano casuali, cioè dovuti a coincidenze fortuite (fluttuazioni, dello strumento, errore di campionamento, ...) anziché riflettere il vero andamento del fenomeno

VERIFICA DI IPOTESI

- Esempio. Le oche sono bianche ?
  - Occorrerebbe esaminare tutte le oche del mondo...
  - Supponiamo di osservare un campione di 1000 oche: 999 B e 1 NON B
- Per Popper l'ipotesi è falsificata, e questo sarebbe logicamente ineccepibile, se i dati fossero assolutamente certi: ma anche i risultati dell'osservazione possono risultare falsati, ovvero non corrispondenti alla realtà
  - l'osservatore o lo strumento può avere sbagliato nel determinare il colore (es. può essere stato tratto in inganno dalle condizioni di luce)
  - può essere stato commesso un errore nel registrare il dato sulla scheda cartacea o nel trascriverlo sul supporto magnetico (file)
  - o magari nel campione è entrata per sbaglio un'anatra
  - le possibilità di errore nel processo di rilevazione dei dati empirici sono molto numerose ...
- Allora ci chiediamo: è davvero sufficiente un caso contrario su 1000 per abbandonare una ipotesi che potrebbe in realtà essere vera ?

VERIFICA DI IPOTESI

- La **verifica statistica** delle ipotesi è un metodo per valutare se ritenere accettabile una determinata ipotesi sulla base dell'evidenza empirica disponibile
- Per verificare una ipotesi occorre un metodo *statistico* perché la conoscenza empirica non è assolutamente certa (è sempre di natura campionaria):
  - sia perché non è possibile osservare tutta la "popolazione"
  - sia perché lo strumento stesso produce misure affette da errore casuale
- Lo scostamento del risultato empirico da quello atteso può quindi:
  - essere dovuto al fatto che l'ipotesi è sbagliata
  - oppure può essere dovuto al caso, benché l'ipotesi sia effettivamente vera
- Tuttavia, se il risultato dell'osservazione è "lontano" da quello previsto secondo l'ipotesi, allora è improbabile che l'ipotesi sia vera, cioè che tale scostamento sia imputabile al caso
- Il problema diventa quindi come *quantificare* il grado di accordo/disaccordo del risultato osservato con quanto previsto dall'ipotesi, per formulare una regola di accettazione/rifiuto dell'ipotesi

VERIFICA DI IPOTESI

- **Fasi della verifica di ipotesi statistica**
  - formulazione del **sistema di ipotesi**: è costituito dall'ipotesi da verificare, detta ipotesi nulla (Ho), e dall'ipotesi alternativa (H1), generalmente la negazione logica della prima
  - scegliere la **statistica test**: una quantità calcolata sui dati osservati, che sintetizza l'informazione portata dal campione ai fini dell'inferenza
  - esplicitare le **assunzioni**: ipotesi ausiliarie che *non* vengono sottoposte a verifica, ma si rendono necessarie per lo sviluppo formale del metodo: di solito riguardano la distribuzione della variabile dipendente (ovvero del processo di misura che genera i dati osservati)
  - determinare la **distribuzione campionaria** della statistica test: immaginando di ripetere il test infinite volte (principio del campionamento ripetuto), la statistica test assumerà valori diversi, descrivendo una propria distribuzione
  - prefissare il **livello di significatività** del test: il test può anche condurre a rifiutare una ipotesi vera, ma questo deve avvenire "raramente"; il livello di significatività stabilisce quante volte al max il test potrà condurre ad una decisione sbagliata
  - ⇨ determinare la **regione di rifiuto** per l'ipotesi Ho: è il punto di arrivo, ci permette di decidere se accettare o rifiutare l'ipotesi

VERIFICA DI IPOTESI

- **Formulazione dell'ipotesi**
- L'ipotesi deve essere formulata come un enunciato di natura *quantitativa* su una o più caratteristiche di un fenomeno o di una popolazione: in ultima analisi, è sempre riconducibile ad una affermazione su un *parametro* di una distribuzione
  - ipotesi sulla media, la frequenza di un carattere, la variabilità, ...
  - ipotesi sul confronto tra (le medie di) due o più popolazioni
  - ipotesi sulla dipendenza tra due variabili (correlazione, associazione, ...)
- Dunque, quale che sia l'ipotesi da verificare, anche espressa inizialmente in termini qualitativi (es. "I cigni sono bianchi") occorre formalizzarla in termini quantitativi, trasformandola in una affermazione sul valore di un parametro di una distribuzione
- L'ipotesi da verificare (o falsificare) viene detta **ipotesi nulla** e indicata con **Ho**, mentre l'ipotesi alternativa viene indicata con H1
  - Ho viene detta ipotesi *nulla* perché si preferisce formulare come Ho l'ipotesi che descrive una situazione di riferimento, o che rappresenta un valore base, rispetto alla quale evidenziare una differenza o un effetto
  - Ho deve essere una ipotesi *puntuale* (cioè una singola affermazione, ovvero un valore ben determinato), mentre H1 può essere una ipotesi *complessa* (un insieme di valori): si tratta di un requisito per lo sviluppo formale del metodo

VERIFICA DI IPOTESI

- **Il Sistema di Ipotesi**
- Ho e H1 costituiscono il **sistema di ipotesi**
- Il sistema di ipotesi si dice di tipo **bilaterale** (o bidirezionale, o a due code) quando H1 è una ipotesi complessa (cioè descrive più valori) e comprende sia i valori minori che quelli maggiori rispetto al valore puntuale previsto da Ho:

$$\begin{cases} H_0 : \ \mu = \mu_o \\ H_1 : \ \mu \neq \mu_o \end{cases}$$

dove μ_o è un valore qualsiasi (es. 100)

- Il sistema di ipotesi è invece **unilaterale** (o unidirezionale, o ad una coda) nei seguenti casi:

$$\begin{cases} H_0 : \ \mu = \mu_o \\ H_1 : \ \mu > \mu_o \end{cases} \quad \text{oppure} \quad \begin{cases} H_0 : \ \mu = \mu_o \\ H_1 : \ \mu < \mu_o \end{cases}$$

VERIFICA DI IPOTESI

- Esempio
  Gli studenti dell'università di Verona sono più intelligenti della media ?
  - Supponiamo di decidere di misurare l'intelligenza con il test del QI (ammesso che il QI misuri l'intelligenza...)
  - Per costruzione, nella popolazione di riferimento il QI ha media 100, quindi possiamo formulare il sistema di ipotesi:

$$\begin{cases} H_0 : \mu = 100 \\ H_1 : \mu > 100 \end{cases}$$

  - l'ipotesi nulla afferma che il QI medio degli studenti di Verona è uguale a quello della popolazione (100), cioè che *non sono più* intelligenti (non c'è differenza)
  - l'ipotesi alternativa H1, in questo caso, prende in considerazione solo il caso che la classe abbia una media superiore: stiamo cioè escludendo *a priori* la possibilità che possa invece essere inferiore...
  - La scelta di una alternativa unilaterale anziché bilaterale incide sul risultato e deve essere operata con molta cautela: infatti se le nostre aspettative risultassero sbagliate, potremmo sbagliare conclusione

VERIFICA DI IPOTESI

- Esempio
  Ritorniamo all'ipotesi "I cigni sono bianchi": come possiamo formulare in termini statistici, quantitativi, questa ipotesi espressa in forma qualitativa ?
  - dobbiamo formalizzarla trasformandola in una affermazione sul parametro di una distribuzione
  - in questo problema, possiamo considerare il colore come una variabile qualitativa a due livelli (dicotomica), che può assumere solo le due modalità: bianco e non-bianco
  - la distribuzione di questa variabile è caratterizzata dalla percentuale di unità bianche (che sappiamo essere la media per una variabile dicotomica 0/1)
  - l'ipotesi da verificare afferma che *tutti* i cigni sono bianchi, contro l'alternativa che ci siano anche cigni non-bianchi
  - dunque, se indichiamo con pi-greco la percentuale di cigni bianchi "nel mondo", il sistema di ipotesi può essere scritto:

$$\begin{cases} H_0 : \pi = 1 \\ H_1 : \pi < 1 \end{cases}$$

  - L'alternativa è qui unilaterale: in questo caso è una scelta obbligata, dato che una percentuale non può essere maggiore di 1

VERIFICA DI IPOTESI

- Esempio
  Vogliamo verificare se c'è una differenza di altezza (o qualunque altra variabile quantitativa) tra due popolazioni. Come possiamo formulare l'ipotesi ?
- Le due popolazioni, A e B, avranno ovviamente due diverse distribuzioni per la variabile altezza:
  - diciamo di accontentarci di confrontare le medie delle due distribuzioni, supponendo che presentino uguali variabilità (o molto simili) e andamento cioè forma Normale (ipotesi questa che può reggere per l'altezza, ma non per altre variabili, ad es. il reddito)
  - ecco allora che abbiamo già individuato come formalizzare l'ipotesi in termini statistici: può essere espressa come confronto tra le medie delle due distribuzioni
- Ricordiamo che Ho deve essere una ipotesi puntuale, cioè una affermazione secca sulla situazione prevista, non un range di possibilità: per questo dobbiamo formulare come ipotesi nulla Ho che le due medie siano uguali, e come H_1 che siano diverse:

$$\begin{cases} H_0 : \mu_A = \mu_B \\ H_1 : \mu_A \neq \mu_B \end{cases}$$

VERIFICA DI IPOTESI

- **La Statistica Test (il Test)**
- La **statistica test**, o semplicemente **il Test**, è una quantità che viene calcolata a partire dai dati osservati, in grado di riassumere l'informazione campionaria rilevante ai fini dell'inferenza
- La statistica test da utilizzare varia a seconda del problema, cioè del sistema di ipotesi (e delle assunzioni ausiliarie): la scelta è in realtà più semplice di quanto si possa temere, almeno nei problemi standard che si incontrano più frequentemente
- I principali problemi di verifica di ipotesi hanno infatti una soluzione nota, già sviluppata, ovvero un Test pronto all'uso: il lavoro diventa quello di cercare di ricondurre il problema reale ad una di queste situazioni standard
- Esempio:

$$\begin{cases} H_0 : \mu = \mu_o \\ H_1 : \mu \neq \mu_o \end{cases} \rightarrow \boxed{t = \frac{\bar{x} - \mu_o}{s / \sqrt{n}}}$$

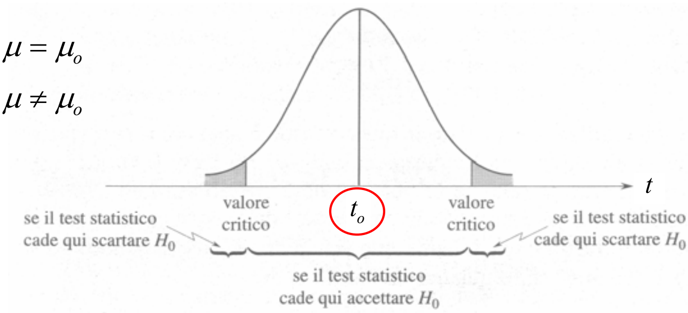
test T di Student

- La statistica test osservata (a posteriori) non è altro che un numero, calcolato sui dati campionari: esattamente come la media o la deviazione standard campionarie (anzi spesso si basa proprio su tali statistiche campionarie)

VERIFICA DI IPOTESI

- **La Regione di rifiuto**
- La **Regione di rifiuto** è l'insieme dei valori che la statistica test *non dovrebbe* assumere, se è vera l'ipotesi nulla, se non per effetto del caso e con una probabilità molto bassa

$$\begin{cases} H_0 : \mu = \mu_o \\ H_1 : \mu \neq \mu_o \end{cases}$$

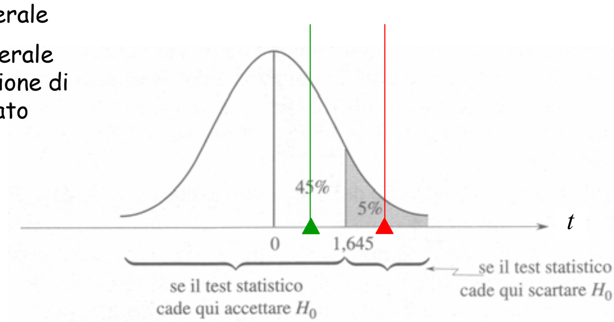


- Se il valore assunto dalla statistica test cade nella regione di rifiuto, il risultato campionario risulta **significativamente** lontano dall'atteso
- Quando il valore osservato del Test cade nella regione di rifiuto, questo conduce al rifiuto dell'ipotesi nulla, perché il risultato empirico è in disaccordo con quanto previsto dall'ipotesi

VERIFICA DI IPOTESI

- Nella maggior parte delle applicazioni pratiche, la regione di rifiuto consisterà in un intervallo, o nell'unione di due intervalli, a seconda che il test sia unilaterale o invece bilaterale
- Nel caso di un test unilaterale (es. coda a destra) la regione di rifiuto sarà tutta da un lato

$$\begin{cases} H_0 : \mu = \mu_o \\ H_1 : \mu > \mu_o \end{cases}$$



- Arrivare a determinare la regione di rifiuto significa anche avere la regola di accettazione/rifiuto dell'ipotesi Ho: tutto il metodo consiste effettivamente nella determinazione della regione di rifiuto per il Test
- La determinazione della regione di rifiuto richiede che si conosca la *distribuzione* della statistica Test

VERIFICA DI IPOTESI

- **La distribuzione campionaria della statistica Test**
- A posteriori, dato un campione, la statistica test *osservata* è un numero.
  A priori, se immaginiamo di ripetere infinite volte il campionamento, la statistica test assume valori sempre diversi, descrivendo una distribuzione tipica di quel test
- Ogni test ha cioè una propria distribuzione specifica, con una forma caratteristica (tanto che in molti casi prende il nome dal test stesso), che deve essere determinata per poter procedere alla determinazione della regione di rifiuto
- La statistica test è una trasformazione delle n variabili indipendenti X(i): quindi la sua distribuzione dipende da quella delle variabili elementari X(i) di cui si compone, e dal loro numero (n), oltre ovviamente che dalla sua espressione analitica
- Esempio: il test T di Student

$$t = \frac{\bar{x} - \mu_o}{s / \sqrt{n}} \quad \sim \quad t_{n-1}$$

- Ricavare la distribuzione della statistica test è un problema matematico che richiede nozioni avanzate di calcolo delle probabilità
- Per tutti i problemi standard, questo lavoro è già stato fatto e quindi già conosciamo le distribuzioni dei test che si usano più frequentemente; oltre alla Normale, le distribuzioni che ricorrono maggiormente sono:
  T di Student, F di Snedecor, Chi-Quadrato

VERIFICA DI IPOTESI

- **Test sulla media di una popolazione**
- Uno dei problemi più semplici è quello dell'ipotesi sulla media di una popolazione, che si può presentare con due varianti:
  - la varianza della popolazione è nota
  - la varianza della popolazione è ignota
- **Primo caso: Varianza nota**
- Quando la varianza è nota, per testare una ipotesi sulla media si usa il test z :

$$\begin{cases} H_0 : \mu = \mu_o \\ H_1 : \mu \neq \mu_o \end{cases} \rightarrow \boxed{z = \frac{\bar{x} - \mu_o}{\sigma_{nota} / \sqrt{n}}} \sim N(0,1) \,|\, H_o \text{ è vera}$$

- che si distribuisce normalmente, quando Ho è vera, infatti:

$$x_i \sim N(\mu, \sigma) \quad i.i.d. \, \forall i \quad \Rightarrow \quad \bar{x} \sim N(\mu, \frac{\sigma}{\sqrt{n}})$$

$$\text{Quando } H_0 \text{ è vera}: \ \mu = \mu_o \quad \text{e se inoltre } \sigma \text{ è nota}: \quad z = \frac{\bar{x} - \mu_o}{\sigma_{nota} / \sqrt{n}} \sim N(0,1)$$

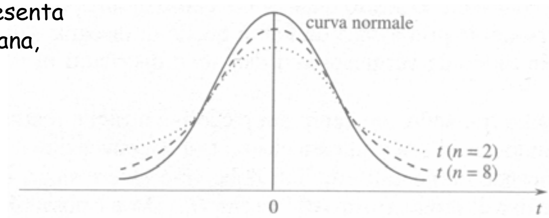
VERIFICA DI IPOTESI

$$\boxed{\hat{\sigma}^2 = s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}}$$

- **Secondo caso: Varianza ignota**
- Quando la varianza è ignota, si perviene al **test t di Student** :

$$\begin{cases} H_0 : \ \mu = \mu_o \\ H_1 : \ \mu \neq \mu_o \end{cases} \rightarrow \boxed{t = \frac{\bar{x} - \mu_o}{s / \sqrt{n}}} \sim t_{n-1} \,|\, H_o \text{ è vera}$$

- La statistica test t di Student presenta una distribuzione a forma di campana, simile alla distribuzione Normale



- **La Distribuzione t di Student**
- La forma della curva t di Student è caratterizzata da un unico parametro, detto *gradi di libertà*
- Al crescere del valore di questo parametro (cioè al crescere dei gradi di libertà), la forma della distribuzione t si avvicina sempre più a quella della Normale:
  - per n > 30 si può utilizzare direttamente la tavola della Normale
  - per n <= 30 esistono e si devono usare le tavole specifiche per la t di Student
- Perché valga questo risultato è necessario che siano vere alcune assunzioni: le X(i) devono essere i.i.d. cioè indipendenti e identicamente distribuite Normalmente

VERIFICA DI IPOTESI

- ***Assunzioni***
- Le assunzioni sono ipotesi accessorie che si rendono necessarie per lo sviluppo formale del metodo, cioè per determinare la distribuzione del test statistico
- Le assunzioni generalmente riguardano la distribuzione della variabile osservata, o meglio del processo di misura che genera i dati osservati
- Le assunzioni, in quanto tali, non sono necessariamente vere, ma non sono oggetto di verifica nel problema in questione:
  - possono essere esse stesse sottoposte a verifica preliminarmente, cioè prima di affrontare il problema di verifica che ci interessa
  - altrimenti devono comunque essere giustificate caso per caso
- I risultati della verifica di ipotesi statistica possono essere seriamente compromessi dalla violazione di una delle assunzioni su cui si basa lo specifico test, e magari condurre al rifiuto di una ipotesi in presenza di una evidenza empirica in realtà insufficiente per farlo
- Alcuni test risentono meno di altri della violazione degli assunti, e per questo sono detti *robusti* : in tutti gli altri casi dobbiamo preoccuparci che le assunzioni siano verosimili

VERIFICA DI IPOTESI

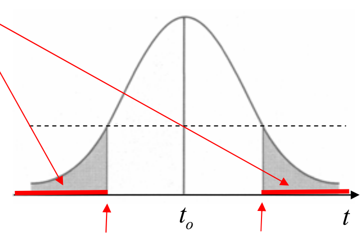
- **Il livello di Significatività del test**
- Un criterio per determinare la regione di rifiuto consiste nel prefissare la probabilità (il rischio) di rifiutare erroneamente l'ipotesi Ho nel caso che sia vera
- La probabilità di commettere tale errore, detto **errore di I tipo**, viene chiamata **livello di significatività *del test*** e indicata con α :

$$\alpha = P\{rifiutare\ H_o \mid H_o\ vera\}$$

- Se, ad esempio, si fissa α =0,05 significa che si accetta il rischio di sbagliare conclusione, nel senso di rifiutare una ipotesi nulla vera, 5 volte su 100
- Il livello di significatività prefissato determina la "dimensione" della regione di rifiuto (e di quella di accettazione):

$$\alpha = P\{\, t \in [regione\ di\ rifiuto]\,\}$$

- maggiore è il valore di α tollerato e maggiore sarà la dimensione della regione di rifiuto
- al contrario, minore è il valore prefissato di α e maggiore sarà la dimensione della regione di accettazione: quindi il test risulterà più *conservativo* nei confronti dell'ipotesi Ho



VERIFICA DI IPOTESI

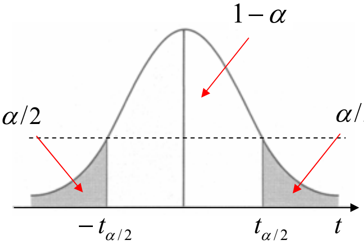
- Nel caso di test bilaterale, o a due code:

$$\begin{cases} H_0 : \ \mu = \mu_o \\ H_1 : \ \mu \neq \mu_o \end{cases}$$

si individua una regione di rifiuto del tipo

$$(-\infty\,, -t_{\alpha/2}] \cup [t_{\alpha/2}\,, +\infty) \ \ tale\ che$$

$$\alpha = P\{\, t \leq -t_{\alpha/2} \cup t \geq t_{\alpha/2}\,\}$$



$$ovvero\ una\ regione\ di\ accettazione: \ (-t_{\alpha/2}\ ,\ t_{\alpha/2}\ ) \ \ per\ la\ quale:$$

$$P\{\,-t_{\alpha/2} < t < t_{\alpha/2}\,\} = 1 - \alpha$$

dove $-t_{\alpha/2}$ e $t_{\alpha/2}$ sono i valori critici a sinistra e a destra della statistica test ottenuti in corrispondenza del livello di significatività α prefissato, equiripartito (α/2) sulle due code della distribuzione della statistica test

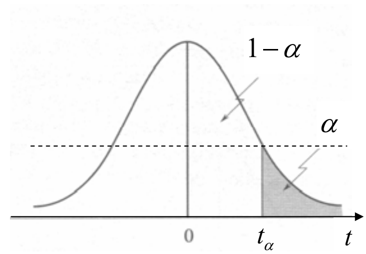
- Se il valore osservato della statistica test cade in tale regione

=> l'ipotesi nulla viene rifiutata

VERIFICA DI IPOTESI

- Nel caso di test unilaterale, o ad una coda:

$$\begin{cases} H_0 : \mu = \mu_o \\ H_1 : \mu > \mu_o \end{cases}$$



si individua una regione di rifiuto del tipo

$$[t_\alpha , +\infty) \;\; tale\ che \;\; P\{\ t \geq t_\alpha \} = \alpha$$

$$ovvero\ una\ regione\ di\ accettazione\ per\ cui\ sia \quad P\{ t < t_\alpha \} = 1 - \alpha$$

t_α è il valore critico della statistica test, ottenuto in corrispondenza del livello di significatività α prefissato

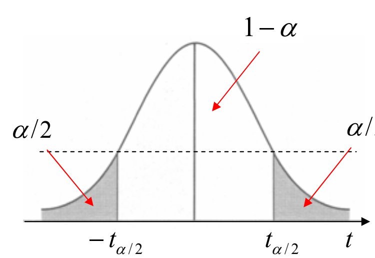
- Se il valore campionario osservato di t è maggiore del valore critico, l'ipotesi nulla viene rifiutata, altrimenti viene accettata

VERIFICA DI IPOTESI

- Test bilaterale (a due code):

$$\alpha = P\{ t \leq -t_{\alpha/2} \, , t \geq t_{\alpha/2} \}$$

$$P\{ -t_{\alpha/2} < t < t_{\alpha/2} \} = 1 - \alpha$$

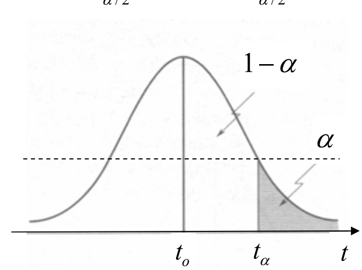


l'area α viene divisa a metà e così si determinano i due valori critici per il test: $-t_{\alpha/2}$ e $t_{\alpha/2}$

- Nel caso di test unilaterale (a una coda) l'area di dimensione α sarà tutta da una parte (es. a destra):

$$\alpha = P\{\ t \geq t_\alpha \}$$

$$P\{ t < t_\alpha \} = 1 - \alpha$$



- Comunemente, i valori utilizzati per α sono 0,05 o 0,01 : 0,01 è più conservativo nei confronti di Ho, che viene rifiutata solo di fronte ad un risultato empirico più netto
- Queste soglie sono arbitrarie e puramente orientative: i numeri indicati non hanno altra proprietà che quella di essere numeri tondi, ma sono quelli generalmente adottati

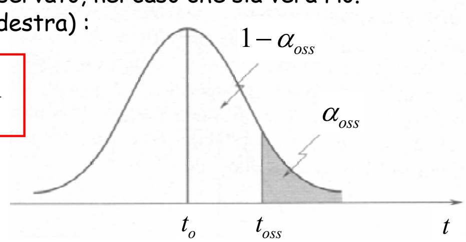
VERIFICA DI IPOTESI

- **Significatività Osservata (o p-value)**
- Il livello di significatività **osservato** (α **osservato**) è la probabilità di commettere un errore di I tipo (rifiutare Ho quando è vera) dato il risultato campionario effettivamente osservato:

$$\alpha_{oss} = P\{rifiutare\ H_o \mid risultato\ campionario\ osservato\}$$

- In pratica, l' α osservato è la probabilità che la statistica test produca un valore ancora più lontano dall'atteso di quello osservato, nel caso che sia vera Ho.
  Ad es. nel caso di un test ad una coda (a destra) :

$$\begin{cases} H_0 : \mu = \mu_o \\ H_1 : \mu > \mu_o \end{cases} \qquad \boxed{\alpha_{oss} = P\{ t \geq t_{oss} \}}$$



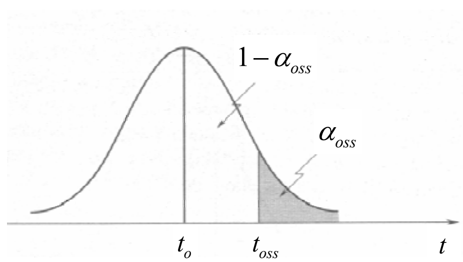
- L' α osservato è una misura della *verosimiglianza* dell'ipotesi nulla, in base al risultato osservato nel campione:
  - maggiore è l' α osservato, e più l'ipotesi nulla è verosimile e quindi risulta "confermata" dalle osservazioni, cioè compatibile con il risultato empirico
  - viceversa minore è l' α osservato, e più è improbabile che Ho sia vera:
    il risultato osservato risulta troppo (*significativamente*) diverso dall'atteso, cioè da quanto previsto da Ho

Prof. Claudio Capiluppi - Facoltà di Scienze della Formazione - A.A. 2006/07

VERIFICA DI IPOTESI

- In pratica, se $\alpha_{oss} < \alpha_{prefissato}$

  => si rifiuta Ho



- L' α osservato permette di trarre immediatamente la conclusione sull'accettazione o il rifiuto dell'ipotesi, indicandoci anche *quanto* l'ipotesi è verosimile (e quindi confermata) sulla base dei dati osservati
- Per questa ragione tutti i programmi software utilizzano in realtà questo secondo metodo, invece di calcolare i valori critici per il test, e forniscono quindi l' α osservato, spesso indicato come **p-value**
- Esempio
  Se risulta α osservato = 0,045 che conclusione traiamo sull'ipotesi ?
  Al 5% di significatività, Ho viene rifiutata: ma la sua verosimiglianza è davvero molto vicina alla soglia del 5% (e ancora compatibile con la soglia del 1%): quindi avremmo bisogno di un campione più numeroso per essere più sicuri

VERIFICA DI IPOTESI

- **Accettazione/rifiuto dell'ipotesi**
- Ricapitolando, dopo aver scelto il test appropriato al problema, esistono dunque due metodi equivalenti di procedere
- Primo metodo: classico (o dei valori critici del test)
  - prefissare il valore arbitrario di α (es. 0,05 oppure 0,01), al di sotto del quale si vuole decidere per il rifiuto di Ho
  - determinare di conseguenza, conoscendo la distribuzione della statistica test (sotto Ho), il valore critico (test ad una coda) o i due valori critici (test a due code) che individuano le regioni di accettazione e di rifiuto
  - verificare in quale regione cade il valore osservato della statistica test
- Secondo metodo: **p-value** (α osservato)
  - calcolare il livello di significatività α osservato (p-value)
  - si rifiuta l'ipotesi nulla se l' α osservato è minore di una soglia prefissata
  - con questo metodo abbiamo anche una indicazione di *quanto* l'ipotesi nulla è confermata (o invece falsificata) dai dati osservati: quanto più α osservato è piccolo, tanto più il risultato osservato risulta significativamente diverso da quello atteso

VERIFICA DI IPOTESI

- **Gradi di libertà**
- La distribuzione della statistica test dipende dalla numerosità campionaria e dai parametri che devono essere stimati per costruire la statistica test:
  - la varianza delle stime dei parametri è funzione (inversa) della numerosità campionaria
  - il numero dei parametri da stimare è esso stesso rilevante: ogni parametro stimato introduce un vincolo per il sistema dei dati
- La differenza tra numerosità campionaria e numero di parametri stimati per costruire il test (meno uno), viene chiamato numero di gradi di libertà del problema:

$$gl = n - (k - 1) = n - k + 1$$

  - questo numero caratterizza la forma delle principali distribuzioni utilizzate (χ^2, t di Student, F di Snedecor)
  - ad eccezione della Normale, che però si usa quando la statistica test (z) prevede la stima di un solo parametro, la media

VERIFICA DI IPOTESI

- **Robustezza del test. Test parametrici e non parametrici**
- In molti test lo sviluppo formale del metodo richiede di fare delle assunzioni circa la forma della distribuzione che ha generato le osservazioni, che si considera quindi nota a meno di un certo numero di parametri: l'ipotesi viene formulata sui valori di tali parametri. Questo approccio è detto **parametrico**.
- In particolare sia il test t di Student che il test F di Snedecor sono test parametrici e assumono che le popolazioni siano distribuite normalmente.
- I test parametrici sono i test più potenti a disposizione, nel senso che riescono a discriminare, cioè ad individuare differenze, anche disponendo di una piccola evidenza empirica. Tuttavia, al venir meno delle condizioni su cui si basano, i risultati di questa classe di test diventano inaffidabili.
- Una procedura statistica che, al contrario, sia relativamente insensibile al venir meno delle assunzioni, viene detta "robusta".
- I test detti **non parametrici** abbandonano le assunzioni circa la distribuzione di provenienza dei dati, e restano quindi validi per una più ampia classe di problemi. Sono meno potenti, circa il 90-95% della potenza dei parametrici quando le popolazioni sono Normali, ma più robusti cioè più attendibili quando non sono rispettate le condizioni di validità dei test parametrici.

VERIFICA DI IPOTESI

- **Fattori di compromissione della validità statistica:**
- Tutti i fattori che aumentano l'errore di misura, riducendo la precisione delle stime, infatti all'aumentare della varianza di stima:
  - l' α osservato cresce ovvero aumenta la verosimiglianza dell'ipotesi nulla, e quindi risulterà non significativa la differenza osservata del risultato empirico dal valore atteso
  - aumenta anche la probabilità di commettere un errore di 2° tipo, cioè il test perde potenza e non riesce più a discriminare tra le due ipotesi, a tutto vantaggio dell'ipotesi nulla.
- La violazione degli assunti sui quali si basa lo sviluppo formale del metodo e, in particolare, il calcolo della distribuzione della statistica test: in questo caso il risultato è la sovrastima della potenza del test e la sottostima dell' α osservato, che induce a ritenere significative differenze dall'atteso che non lo sono.
- Ad esempio, nel caso dell'analisi della varianza classica, devono essere rispettati i seguenti assunti:
  - Normale distribuzione della variabile dipendente in ciascun gruppo
  - indipendenza delle osservazioni dei diversi gruppi
  - omoschedasticità dei gruppi a confronto

VERIFICA DI IPOTESI

- Esempio. Metodo classico
- Gli studenti dell'università di Verona sono più intelligenti della media ?
- Avevamo già formalizzato il sistema di ipotesi:

$$\begin{cases} H_0 : \ \mu = 100 \\ H_1 : \ \mu > 100 \end{cases}$$

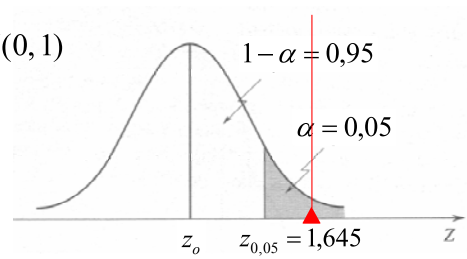
- Selezioniamo casualmente 40 studenti e ne misuriamo l'intelligenza con il test del QI, ottenendo una media campionaria pari a 107
- Sapendo che il QI nella popolazione si distribuisce come una N(100, 15), ci chiediamo: una media campionaria di 107 può essere dovuta al caso, oppure è una differenza troppo netta, e indica quindi un'intelligenza realmente superiore ?
- Essendo nota la varianza della popolazione, il test da usare è :

$$z = \frac{\bar{x} - \mu_o}{\sigma_{nota} / \sqrt{n}} \sim N(0,1)$$

- Nel nostro caso:

$$z = \frac{107 - 100}{15/\sqrt{40}} = \frac{7}{15/6{,}33} = 2{,}95$$

- Il valore critico di z (α =0,05) = 1,645



LA DISTRIBUZIONE NORMALE

- Il valore critico della statistica test, per (α =0,05), risulta

$$z_{0,05} = 1{,}645$$

- Il valore osservato della statistica test è pari a 2,95: quindi cade a destra del valore critico, nella regione di rifiuto
- Quindi l'ipotesi nulla deve essere rifiutata:
- il risultato campionario è troppo distante dall'ipotesi Ho perchè tale scostamento sia imputabile al caso
- possiamo concludere che gli studenti di Verona hanno un QI superiore alla media

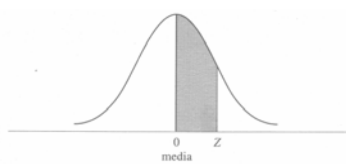


Tavola A.1 La distribuzione normale.

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4022	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990

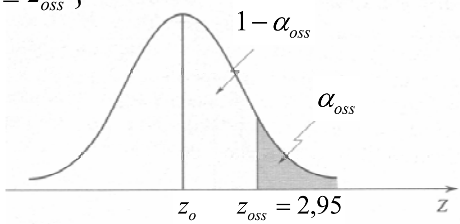
VERIFICA DI IPOTESI

- Esempio. Metodo dell' α osservato
- Nel caso del nostro problema, cioè di un test z unilaterale con coda a destra:

$$\begin{cases} H_0 : \ \mu = 100 \\ H_1 : \ \mu > 100 \end{cases}$$

la significatività osservata si calcola come

$$\alpha_{oss} = P\{ z \geq z_{oss} \}$$



- Il valore campionario della statistica test come abbiamo appena visto è:

$$z_{oss} = \frac{107-100}{15/\sqrt{40}} = \frac{7}{15/6,33} = 2,95$$

- Sulle tavole della Normale cerchiamo la probabilità che z sia maggiore di 2,95: utilizzando le solite tavole, che ci forniscono l'area complementare rispetto a quella che cerchiamo, otterremo:

$$\alpha_{oss} = P\{ z \geq 2,95\} = 0,5 - 0,4984 = 0,0016$$

- La verosimiglianza dell'ipotesi Ho in base al risultato campionario è molto bassa, minore di 0,05 e anche di 0,01, quindi rifiutiamo senz'altro Ho
- Questo significa che gli studenti sono significativamente più intelligenti della media (100) della popolazione

LA DISTRIBUZIONE NORMALE

- Avendo a disposizione la solita tavola, determiniamo l'α osservato come:

$$\alpha_{oss} = P\{ z \geq 2,95\} =$$
$$= 0,5 - 0,4984 = 0,0016$$

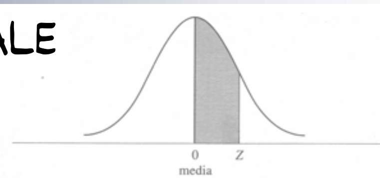


Tavola A.1 La distribuzione normale.

Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4022	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990

VERIFICA DI IPOTESI

- **Test sulle percentuali**
- Un caso particolare di media, come abbiamo visto, è quello della variabile dicotomica: vediamo come si procede per verificare una ipotesi su una percentuale

$$\begin{cases} H_0 : \pi = \pi_0 \\ H_1 : \pi \neq \pi_0 \end{cases}$$

- Ricordiamo che nel caso della variabile dicotomica, la varianza è data da p (1-p) , il test z quindi diventa:

$$z = \frac{p - \pi_o}{\sqrt{\frac{\pi_o(1-\pi_o)}{n}}}$$

- Questa formulazione del test utilizza il valore previsto da Ho anche per calcolare il valore della varianza della popolazione
- In alternativa, è possibile stimare invece la varianza della popolazione con quella campionaria, ottenendo il test t:

$$t = \frac{p - \pi_o}{\sqrt{\frac{p(1-p)}{n}}}$$

VERIFICA DI IPOTESI

- Esempio
  Ritorniamo al nostro problema dei cigni. Dunque dobbiamo decidere se i cigni sono bianchi…

$$\begin{cases} H_0 : \pi = 1 \\ H_1 : \pi < 1 \end{cases}$$ percentuale di cigni bianchi

- Supponiamo di avere effettuato un campione di 100 cigni e averne trovati 99 bianchi e 1 non bianco: tale evidenza empirica è sufficiente per considerare falsificata l'ipotesi ?
- Per la particolarità dell'ipotesi, che descrive un caso estremo, preferiamo stimare la varianza con quella campionaria e adottare il test t :

$$t = \frac{0{,}99-1}{\sqrt{\frac{0{,}99(1-0{,}99)}{100}}} = \frac{-0{,}01}{\sqrt{\frac{0{,}99 \cdot 0{,}01}{100}}} = \frac{-0{,}01}{\sqrt{\frac{0{,}0099}{100}}} \cong \frac{-0{,}01}{\sqrt{\frac{0{,}01}{100}}} = -\frac{0{,}01}{0{,}01} = -1 \quad (-1{,}005)$$

- Per n=100 possiamo utilizzare il valore critico della normale
- Il valore di t campionario cade nella regione di accettazione, quindi non consideriamo Ho falsificata
- 2/100 basterebbero ? e 1/1000 ?

$t_{0,05} = -1{,}645$