

Statistica medica MED/01 per il Corso di Laurea in Medicina e Chirurgia LM/41

Docente

Dott. Vincenzo Guardabasso

Azienda Ospedaliero-Universitaria "Policlinico-Vittorio Emanuele"

Il corso si propone di partire dalle conoscenze di base di statistica, approfondendo la sintesi riassuntiva di dati di vario tipo.

Vengono presentati i principali test statistici utilizzati nella statistica biomedica.

Al termine del corso lo studente sarà in grado di interpretare i risultati statistici riportati nella letteratura biomedica e di utilizzare semplici tecniche di analisi statistica e rappresentazione dei dati.

1

La statistica è quella dei polli?

LA STATISTICA

*Sai ched'è la statistica? È 'na cosa
che serve pe' fa' un conto in generale
de la gente che nasce, che sta male,
che more, che va in carcere e che sposa.*

*Ma pe' me la statistica curiosa
è dove c'entra la percentuale,
pe' via che, lì, la media è sempre eguale
puro co' la persona bisognosa.*

*Me spiego: da li conti che se fanno
secundo le statistiche d'adesso
risurta che te tocca un pollo all' anno;*

*e, se nun entra ne le spese tue,
t'entra ne la statistica lo stesso
perché c'è un antro che ne magna due.*

Trilussa (Carlo Alberto Salustri), 1871-1950

I polli di Trilussa

Quasi tutti abbiamo sentito citare almeno una volta i famosi "polli" di Trilussa discutendo di statistica. Pochi, però, conoscono il testo originale da cui è tratta questa famosa massima.

http://www.efira.com/spigolature/trilussa_statistica.htm



2

Scopi della statistica medica

◆ Perché la statistica:

- variabilità biologica
 - tra individui e/o nel tempo
- errori

DEFINIZIONE:

Disciplina che riguarda il trattamento di dati derivanti da un gruppo di individui

◆ Scopi

- riassumere: statistica **descrittiva**
- sperimentare: statistica **inferenziale**

◆ Utilità

- eseguire studi (clinici, di qualità, di gestione, in lab.)
- capire (dati, statistiche, diagrammi, risultati)

3

Definizioni

◆ Unità statistica

- l'oggetto elementare di analisi/misura/studio

◆ Popolazione (statistica)

- l'insieme degli oggetti da studiare

◆ Campione

- il gruppo di unità scelte appropriatamente per rappresentare la **Popolazione**

4

Strumenti

- ◆ Numeri
 - Interi e decimali
 - Arrotondamento
 - Cifre significative
- ◆ Variabili numeriche
 - Unità di misura
 - Campo - intervallo - *range*
 - Ordine di grandezza
 - Notazione scientifica
- ◆ Formule
 - Operatori e Parentesi
 - Sommatoria
 - Trasformazioni
- ◆ Strumenti di misura
 - accuratezza
 - precisione
 - influenza dell'osservatore
- ◆ Raccolta dei dati
 - metodi manuali
 - contatori e strumenti
 - elaborazione automatica
 - tabelle
- ◆ Diagrammi
 - punti e linee
 - segmenti e barre
 - aree

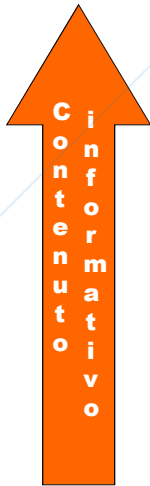
5

1: Statistica descrittiva

L'insieme dei metodi, numerici e grafici, per riassumere e presentare i dati

6

Dati: tipologia delle variabili



- ◆ **Numeriche (Quantitative)**
 - a scala intervallare o rapportabile
 - ➔ **continue** (da misura)
 - esempi: temperatura, altezza, peso
 - ➔ **discrete** (da conteggi)
 - esempi: posti letto, numero di pazienti
- ◆ **Ordinali (Categoriche)**
 - ➔ per Ranghi: graduatorie
 - esempi:
 - ordine di arrivo di una gara
 - gradimento di un servizio
 - scala del coma; indice di Apgar
- ◆ **Nominali (Categoriche) o Qualitative**
 - esempi: colore, etnia
 - ➔ **dicotomiche** se solo due modalità
 - Esempi: sesso, morte

7

Frequenze

- ◆ **Conteggio:**
 - ➔ contare il numero delle volte in cui si presenta una modalità o valore di una variabile
 - ➔ risultato: **Frequenza assoluta f**
 - è un numero intero
- ◆ **Riferimento del conteggio al totale N**
 - ➔ risultato: **Frequenza relativa f / N**
 - **Proporzione:** è un valore decimale compreso tra 0 e 1
- ◆ **Riferimento a 100**
 - ➔ risultato: **Frequenza percentuale $f / N \times 100$**
 - è indicata con il segno percentuale %
 - **attenzione:** è diversa da variazione percentuale
- ◆ **Somma delle frequenze =**
 - **Frequenza cumulativa e percentuale cumulativa**

8

Raccolta e rappresentazione di dati

◆ Frequenze

- Tabella (ad una entrata)
- Tabella a due entrate (*Tabella di contingenza*)

◆ Distribuzione di frequenza

- per dati numerosi, variabili continue, ecc.
- costruzione di classi di frequenza
 - perdita di informazione, ma migliore visione di insieme
 - caratteristiche:
 - limite inferiore, limite superiore
 - intervallo di classe, ampiezza
 - classi aperte
 - valore centrale
 - frequenza di classe

9

Misure riassuntive

◆ Necessità di sintetizzare l'informazione contenuta nei dati raccolti

- per esprimere una posizione di riferimento
 - a che serve?
 - esempio: possiamo dire che un paziente è iperglicemico se conosciamo il valore "normale" della glicemia. Ma come troviamo il valore normale?
- per esprimere la variabilità
 - a che serve?
 - esempio: di quanto può variare il valore "normale" della glicemia, senza uscire dalla "normalità"

10

Indici di posizione (tendenza centrale)

- ◆ **Necessità:** riassumere i dati in un indice rappresentativo (comprensibile, confrontabile)
 - ➔ compreso tra minimo e massimo
 - ➔ che rappresenta bene i dati più frequenti
 - Media (analitica)
 - valida solo per variabili quantitative
 - Mediana
 - indice valido anche per variabili ordinali
 - Moda
 - indice valido anche per variabili categoriche
 - nota su distribuzioni bimodali

11

Le medie (analitiche)

◆ **media aritmetica**

$$\bar{x} = \sum_{i=1}^N x_i / N$$

- semplice
- pesata (ad es.: usando i crediti formativi)
- **Limitazioni all'uso**
 - dati non quantitativi
 - diversi ordini di grandezza
 - presenza di valori estremi molto "distanti", indeterminati o infiniti
- **media geometrica**
 - dati quantitativi su diversi ordini di grandezza
- **altre medie**
 - media armonica, ...

12

Altri indici di posizione

◆ Mediana

→ valore centrale di una serie ordinata

- 1: ordinare gli N dati disponibili
- 2: prendere
 - valore centrale se N dispari
 - media dei valori centrali se N pari

→ Vantaggi

- non risente dei valori estremi

→ Limitazioni

- non tiene conto della grandezza delle osservazioni
- non si presta ad elaborazioni matematiche

◆ Moda

13

I quantili

◆ estensione del concetto di mediana

◆ utilizzati se i dati sono numerosi

◆ maniera per rappresentare la variabilità in serie di dati di tipo quantitativo o ordinale

→ quartili

- differenza interquartile

→ percentili

- 5%-95% percentile delle curve di accrescimento

14

Indici di variabilità (dispersione)

- ◆ Un criterio per esprimere la variabilità deve sempre essere associato ad un indice di posizione, per esprimere compiutamente il fenomeno studiato e i dati osservati

◆ Variabili quantitative

→ Deviazione standard

- usa le stesse unità della media
- si calcola dagli "scarti" elevati al quadrato
- il denominatore N-1 è chiamato "Gradi di libertà"

$$s = \sqrt{\frac{\sum_{i=1}^N (x_i - \bar{x})^2}{N-1}}$$

→ Coefficiente di variazione percentuale CV%

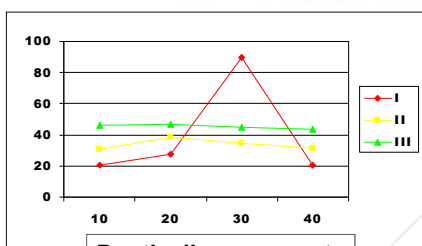
- è un indice di dispersione relativa

◆ Variabili ordinali

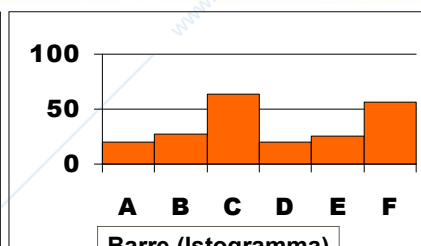
→ Differenza interquartile

15

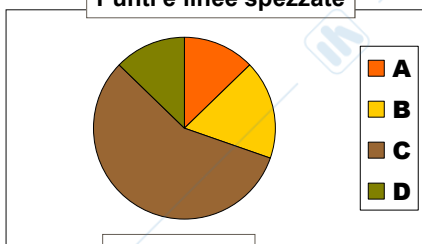
Rappresentazioni grafiche



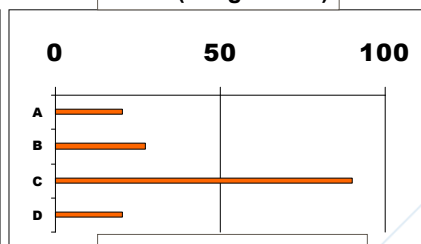
Punti e linee spezzate



Barre (Istogramma)



Torta (area)



Segmenti (orizzontali)

16

Altre rappresentazioni grafiche



FIGURA 2
Esempio di ideogramma (dati della Tabella 7.4).

Ideogramma

Elementi di Statistica medica
Pasquale Bruno Lantieri,
Domenico Riso, Giambattista Ravera
© 2007 - The McGraw-Hill Companies s.r.l.

Pictogramma

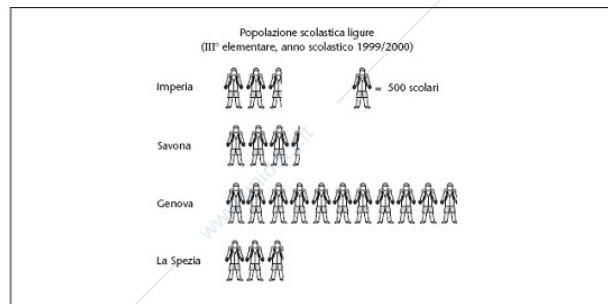
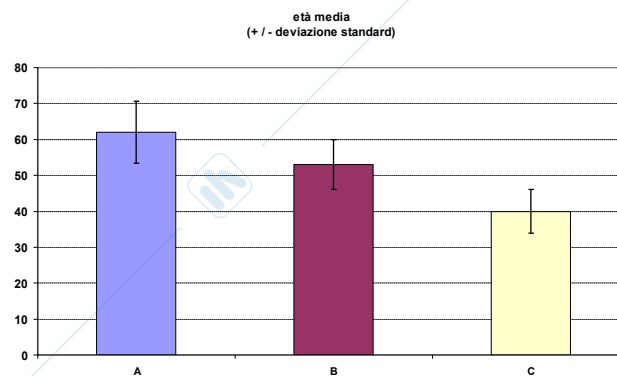


FIGURA 3
Esempio di pictogramma (dati della Tabella 7.4).

17

Rappresentazione grafica - media

- ◆ medie e deviazioni standard:
→ grafici con "baffi" che esprimono la variabilità



18

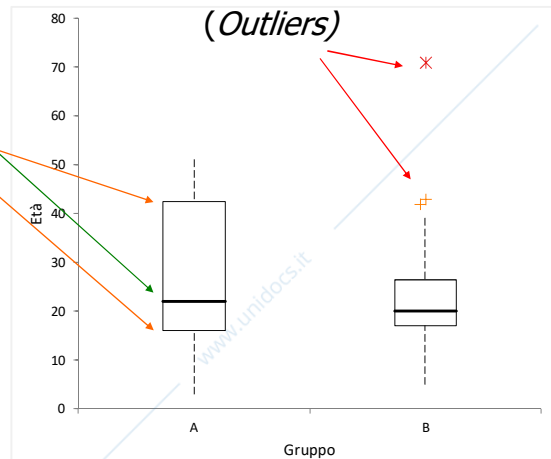
Rappresentazione grafica - mediana

◆ Box Plot:

→ Mediana

→ Quartili

→ Valori estremi
(Outliers)



19

2: Probabilità e statistica inferenziale

Metodi per capire e per scoprire, basati sulla probabilità

20

Statistica inferenziale

- ◆ Teoria e metodologie per fare affermazioni sulle caratteristiche di una popolazione a partire da osservazioni su un gruppo (campione)

➤ Esempio:

- Studio di un farmaco nuovo su un gruppo di pazienti
- Il farmaco è efficace?
- Sarà efficace nella popolazione?

- ◆ Si basa sulla teoria della probabilità:

➔ Dai risultati osservati nel campione l'**inferenza statistica** permette di **indurre** (estrapolare) le caratteristiche della popolazione, con un certo **grado di incertezza**, espresso da una certa **probabilità**

21

La Teoria della Probabilità

- ◆ Studio dei "fenomeni aleatori" (casuali)

➔ **Evento o esito:** risultato di un fenomeno aleatorio

➔ **Probabilità di un evento**

➤ Ad ogni evento è associato un numero compreso tra 0 e 1, detto probabilità dell'evento

- Evento certo
 - Probabilità = 1
- Evento impossibile
 - Probabilità = 0
- Evento possibile
 - Probabilità p tale che $0 < p < 1$

➔ **Spazio campione (delle possibilità)**

➤ insieme di tutti gli eventi (esiti, risultati) possibili

22

Il concetto di probabilità

- ◆ **Uso empirico della probabilità**
 - Concetto usato per prendere decisioni in condizioni di incertezza
- ◆ **Idea "classica" (per risultati con la stessa probabilità...)**
 - Probabilità di un risultato rispetto a tutti i risultati possibili
 - $P(E) = k/N$
 - esempio: la probabilità di avere un numero pari tirando un dado è $3/6 = 0,5$
- ◆ **Idea frequentista**
 - La frequenza relativa di un risultato corrisponde alla probabilità
 - $P(E) = n / N$, per N molto grande
 - esempio: probabilità di complicanze calcolata da numero di complicanze / N di casi di malattia (per N grande)
- ◆ **Idea soggettivista**
 - Misura del grado di fiducia che quel risultato si verifichi
 - esempio: probabilità di successo di un intervento su un paziente espressa come "aspettativa"

23

Calcolo della probabilità

- ◆ **Principio della somma**
 - La probabilità complessiva di alcuni esiti "mutuamente esclusivi" è uguale alla somma delle probabilità
 - $P(A \cup B) = P(A) + P(B)$
 - Esempio: pari al lancio di un dado
- ◆ **Principio del prodotto**
 - La probabilità complessiva di alcuni esiti di eventi "indipendenti" è uguale al prodotto delle probabilità
 - $P(A \cap B) = P(A) \times P(B)$
 - Esempio: pari al lancio di due dadi

24

Probabilità condizionata

- ◆ Definizione: la probabilità P dell'evento B , dopo che è avvenuto l'evento A

$$\rightarrow P(B | A) = P(A \cap B) / P(A)$$

- ◆ Teorema di Bayes

→ Probabilità delle cause

- La probabilità che un evento sia associato ad una certa causa, tra le N cause possibili
- Esempio: test diagnostici e probabilità di malattia

25

Distribuzioni di probabilità

- ◆ Una distribuzione di probabilità descrive il comportamento di una **variabile** applicando la teoria della probabilità
- ◆ Variabile discreta
 - La distribuzione di probabilità specifica i possibili valori insieme alle loro probabilità
 - Esempio: **Distribuzione binomiale** $p + q = 1$
 - La distribuzione binomiale ha una espressione matematica determinata dalla probabilità p di un evento, e dal numero N delle prove; media = Np , deviazione standard = $\sqrt{Np(1-p)}$
- ◆ Variabile continua
 - La distribuzione di probabilità esprime la "densità di probabilità" in modo continuo
 - Esempio: **Distribuzione normale** di Gauss
 - La distribuzione gaussiana ha una espressione matematica determinata dalla media μ e dalla deviazione standard σ

26

Distribuzione normale

◆ E' una funzione di *densità di probabilità*

- di una variabile quantitativa continua
- ha forma simmetrica, a campana
- esprime la probabilità in funzione di media μ e deviazione standard σ
- l'area sotto la curva (la probabilità totale) è 1
- Media = mediana = moda

◆ Distribuzione normale standardizzata

$$z = \frac{x - \mu}{\sigma}$$

- media $\mu = 0$, deviazione standard $\sigma = 1$

27

Campionamento ed inferenza

◆ Gli studi biomedici sono in genere effettuati su campioni, non su popolazioni

◆ Problemi

- Scelta del campione per garantire la rappresentatività
- Vizi ed errori di campionamento

◆ La statistica inferenziale permette

- di **stimare** le caratteristiche della popolazione
- di verificare **ipotesi** specifiche

28

Campionamento

1. Individuare bene la popolazione
 - criteri di inclusione
 - criteri di esclusione
2. Selezionare un sottoinsieme che rappresenti **adeguatamente** la popolazione
 - Numerosità
 - Metodo di Campionamento
 - Selezione casuale: "Randomizzazione"
 - La randomizzazione, quando è possibile, **garantisce** la validità statistica delle **inferenze** (conclusioni) tratte da un'indagine!
 - Stratificazione

29

Randomizzazione

- ◆ Dall'inglese *random* = casuale
 - un processo di selezione in cui ogni elemento di un insieme ha uguale probabilità di essere scelto.
- ◆ Procedura
 - lista numerata di elementi tra cui scegliere il campione
 - estrarre un numero casuale



30

Vizi ed errori casuali

- ◆ Vizi o distorsioni (in inglese = *bias*)
 - sono errori sistematici, prevalentemente in un senso o direzione, forniscono risultati distorti, non accurati, quindi non validi
 - non possono essere misurati
- ◆ Errori casuali
 - sono errori dovuti a fattori diversi ed imprevedibili, agiscono in sensi diversi e con intensità varia, diminuiscono la precisione ma non l'accuratezza
 - possono essere misurati con metodi statistici
- ◆ E' necessario adottare accorgimenti per evitare i vizi e ridurre gli errori casuali

31

Dimensione del campione

- ◆ La possibilità di trarre conclusioni fondate dipende dal numero di unità coinvolte nello studio, e dalla modalità di selezione
- ◆ In genere negli studi sperimentali si vuole conoscere un risultato entro margini predefiniti di incertezza ("*con una certa precisione*")
- ◆ La dimensione del campione dipende da questi margini, non dalla dimensione della popolazione!

32

Stima della media

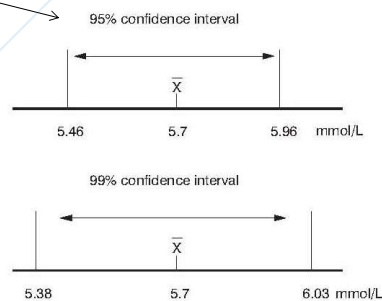
- ◆ La **media** calcolata su un **campione** si discosta dalla media vera della popolazione: $\varepsilon = m - \mu$ (dove ε è l'errore)
- ◆ All'aumentare del numero delle misure la differenza tra media del campione (stima) m e media vera (ignota) μ si riduce
- ◆ Campioni della stessa numerosità forniscono stime diverse della stessa media μ
- ◆ Teorema del limite centrale:
 - la distribuzione delle medie è gaussiana con media μ e deviazione standard $\sigma_{\mu} = \frac{\sigma}{\sqrt{N}}$

Errore Standard della Media = ES

33

Intervallo di confidenza della media

- ◆ Intervallo costruito intorno alla media campionaria, compreso tra i "limiti di confidenza", utilizzando media m e deviazione standard s del campione, al livello di probabilità desiderato (ad esempio 95%)



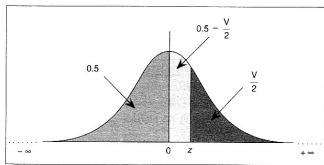
34

l'Intervallo di confidenza ...

- ◆ fornisce una stima intervallare della media della popolazione, poiché include la media vera (ignota) della popolazione; ad esempio, al livello di probabilità del 95% si avrebbe:

→ Limite inferiore: $m - 1,96 \sigma / \sqrt{N}$
 → Limite superiore: $m + 1,96 \sigma / \sqrt{N}$

dove il valore 1,96 è il cosiddetto **valore critico z** preso dalla distribuzione normale standardizzata Z e σ è la deviazione standard della popolazione (ma al suo posto possiamo usare la deviazione standard **s** del campione)



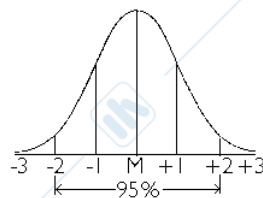
distribuzione
 tabella

TABELLA 2.1 AREA ALLA DESTRA DI z NELLE DUE CODE DELLA DISTRIBUZIONE
 Il valore di ricerca dall'elenco tra righe (colita e prima decimale di z) e
 di colonne (seconda decimale)

Distribuzione z	
z	Area alla destra di z
0.0	0.5000
0.1	0.4602
0.2	0.4207
0.3	0.3821
0.4	0.3438
0.5	0.3065
0.6	0.2704
0.7	0.2354
0.8	0.2019
0.9	0.1700
1.0	0.1398
1.1	0.1112
1.2	0.0844
1.3	0.0594
1.4	0.0363
1.5	0.0155
1.6	0.0075
1.7	0.0044
1.8	0.0027
1.9	0.0016
2.0	0.0009
2.1	0.0005
2.2	0.0003
2.3	0.0002
2.4	0.0001
2.5	0.0000
2.6	0.0000
2.7	0.0000
2.8	0.0000
2.9	0.0000
3.0	0.0000

Costruzione dell'Intervallo di confidenza

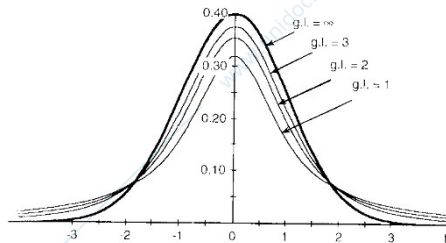
- ◆ ad esempio, al livello 95% si avrebbe:
- limite inferiore: $m - 1,96 s / \sqrt{N}$
 → limite superiore: $m + 1,96 s / \sqrt{N}$



ES

Costruzione dell'Intervallo di confidenza per "piccoli campioni"

- ◆ Se N è minore di 30, è indispensabile usare la distribuzione **t di Student** invece del "valore critico" preso dalla distribuzione gaussiana
 - Limite inferiore: $m - t_{(\alpha, \nu)} s / \sqrt{N}$
 - Limite superiore: $m + t_{(\alpha, \nu)} s / \sqrt{N}$
- ◆ La distribuzione **t** è determinata dai "gradi di libertà", quindi il **valore critico** che si usa cambia col numero di dati disponibili
 $\nu = N - 1$



37

Intervalli di confidenza per proporzioni

- ◆ **Proporzione:** rapporto tra una parte ed un totale
 - f = numero di eventi della variabile dicotomica
 - N = numero totale di eventi
 - $p = f / N$
 - $q = 1 - p = (N-f) / N$
 - $p + q = 1$
 - p e q sono proporzioni
 - p e q possono essere interpretate come probabilità;
 - **intervallo di confidenza della proporzione**
 - al livello 95% si ha:
 - limite inferiore: $p - 1,96 \sqrt{p(1-p) / N}$
 - limite superiore: $p + 1,96 \sqrt{p(1-p) / N}$

38

Verifica di ipotesi: 1

◆ Test di confronto tra due medie

- Domanda: La differenza osservata tra le medie calcolate in due campioni è dovuta al caso (espressione di errore casuale di campionamento), oppure i due campioni sono tratti da due popolazioni diverse?
- ovvero: la differenza osservata tra i gruppi è *statisticamente significativa*?

◆ Esempio: *test t di Student*:

- 1 formulare l'*ipotesi nulla* H_0 e l'*ipotesi alternativa* H_A
- 2 scegliere un *livello di significatività*
- 3 calcolare la statistica t
- 4 verificare la probabilità p del valore t ottenuto

39

Verifica di ipotesi: 2

◆ Test di confronto tra frequenze

- Domanda: La differenza tra le frequenze *assolute* osservate in una tabella di contingenza e quelle attese è dovuta al caso, oppure ad un fattore di classificazione della tabella?
- ovvero: la differenza tra frequenze osservate e frequenze attese è *statisticamente significativa*?

◆ Esempio: *test chi quadro* χ^2 :

- 1 formulare l'*ipotesi nulla* H_0 e l'*ipotesi alternativa* H_A
- 2 scegliere un *livello di significatività*
- 3 calcolare la statistica χ^2
- 4 verificare la probabilità p del valore χ^2 ottenuto

40

Ma gli Intervalli di confidenza...

- ◆ Fanno pure parte della statistica inferenziale
- ◆ Sono più informativi del valore p da un test
- ◆ Presentano i risultati nella stessa scala dei dati
- ◆ Esprimono anche la significatività statistica
- ◆ Possono essere costruiti per varie stime statistiche campionarie:
 - medie, proporzioni, mediane, tassi, ecc.

41

Errori di I e II tipo

- ◆ Nell'applicazione della tecnica del test di ipotesi possono essere commessi due tipi di errore

	Test significativo (si rifiuta H_0)	Test non significativo (non si rifiuta H_0)
H_0 vera	Errore di I tipo (prob. α) Falso Positivo	Conclusione corretta (Confidenza $1 - \alpha$)
H_0 falsa	Conclusione corretta (Potenza $1 - \beta$)	Errore di II tipo (prob. β) Falso Negativo

Analogia giudiziaria

◆ Ipotesi nulla H_0 : l'imputato è innocente

	Imputato condannato (si rifiuta H_0)	Imputato assolto (non si rifiuta H_0)
Imputato innocente (H_0 vera)	Errore di I tipo (prob. α) Falso Positivo	Conclusione corretta (Confidenza $1 - \alpha$)
Imputato colpevole (H_0 falsa)	Conclusione corretta (Potenza $1 - \beta$)	Errore di II tipo (prob. β) Falso Negativo

Metodi statistici avanzati

◆ Per variabili quantitative:

- Analisi della correlazione e regressione
 - per studiare l'associazione tra una variabile "dipendente" ed una o più variabili "indipendenti"
- Analisi della varianza (ANOVA)
 - per il confronto di più di 2 campioni
- Entrambi i metodi sono basati su assunzioni:
 - distribuzione normale (o che può essere trasformata in normale)
 - variabilità (Varianza) simile

Correlazione

- ◆ Scopo:
 - studiare associazione tra due variabili (quantitative)
- ◆ Attenzione:
 - associazione non significa causalità
- ◆ Strumento grafico
 - Diagramma di dispersione (XY)
- ◆ Correlazione lineare: proporzionalità
 - coefficiente di correlazione = r
 - coefficiente di determinazione = r^2
- ◆ Strumento statistico:
 - Regressione lineare (minimi quadrati)

45

Verifica di ipotesi: 3

- ◆ Coppie di variabili quantitative
 - Domanda: Esiste una relazione per cui al crescere del valore di una variabile il valore dell'altra variabile cresce (o decresce), oppure si tratta di fluttuazioni casuali?
 - ovvero: la correlazione tra le due variabili è *statisticamente significativa*?
- ◆ Esempio: **correlazione**:
 - 1 formulare l'**ipotesi nulla** H_0 e l'**ipotesi alternativa** H_A
 - 2 scegliere un *livello di significatività*
 - 3 calcolare la statistica r
 - 4 verificare la probabilità p del valore r ottenuto

46

Analisi della Varianza (ANOVA)

- ◆ Confronto di medie ottenute secondo una classificazione; esempi di classificazioni
 - a) per dose = ordine
 - b) per categoria = criterio graduale (es. sforzo)
 - c) per reparto = casuale, non ordinato

Dati k gruppi, ognuno di numerosità n_i

Media generale = senza tenere conto dei gruppi

Media di gruppo = per ogni gruppo della classificazione

SS totale =

SS per ogni gruppo +
SS delle medie di gruppo dalla media generale

“entro”

“tra”

Varianze = SS / numerosità

Rapporto tra le varianze:

$F = \text{varianza "tra"} / \text{varianza "entro"}$

dopo l'ANOVA possono essere eseguiti test di confronto tra singoli gruppi, ma..

esiste il problema dei confronti multipli!

tutta un'altra storia...

47

Metodi non parametrici

◆ Per:

- Variabili quantitative con
 - distribuzione non normale o ignota
 - eterogeneità della variabilità (Varianza)
- Variabili ordinali
- Variabili categoriche

- ◆ Si utilizzano metodi statistici che non presuppongono una distribuzione determinata da **parametri** (come μ e σ per la distribuzione normale di Gauss)

48

Test non parametrici

- ◆ Due gruppi
 - indipendenti: test della somma dei ranghi di Mann-Whitney
 - stessi soggetti: test di Wilcoxon
- ◆ Più di due gruppi
 - indipendenti: test di Kruskal-Wallis
 - stessi soggetti: test di Friedman
- ◆ Conteggi (variabili categoriche)
 - indipendenti: test del "chi quadro" χ^2
 - stessi soggetti: test di Mc Nemar
- ◆ Correlazione
 - coefficiente di Spearman basato sui ranghi

49

Esempio: Test di Wilcoxon

- ◆ "Test dei ranghi con segno"
- ◆ Test valido per due gruppi di dati "appaiati"
- ◆ Per situazioni prima-dopo
 - ad esempio: cambiamento dei giudizi prima e dopo un corso
- ◆ Vale per variabili quantitative o ordinali
- ◆ Non richiede la distribuzione normale dei dati
- ◆ Permette di gestire tanti valori uguali

50