

Docente: Dott. Ing. Andrea Ghedi

Ingegnere Biomedico, specialista in Ingegneria Clinica
Ingegnere clinico c/o Azienda Ospedaliera di Desenzano del Garda

Recapiti:

Ufficio: Dietro direzione ospedaliera, palazzina Villa Andreis 2
c/o Servizio Ingegneria clinica (una volta alla sbarra seguire le
freccie)

Tel: 0309145436

Mail: andrea.ghedi@aod.it

Internet: www.andreaghedi.it/corsi.htm

Ricevimento: Al bisogno dello studente, da prenotare via mail
almeno due giorni lavorativi prima

Ing. Andrea Ghedi AA 2009/2010

*"Educare significa aiutare l'animo dell'uomo
ad entrare nella totalità della realtà.*

*Non si può però educare se non
rivolgendosi alla libertà, la quale definisce
il singolo, l'io.*

*Quando uno dice "io", la libertà è in tutto
questo dire "io"."*

Ing. Andrea Ghedi AA 2009/2010

MATERIALE DIDATTICO

- Dispense in formato elettronico da scaricare dalla pagina web del corso
- Esercizi in formato elettronico da scaricare dalla pagina web del corso
- Libri di testo reperibili presso la biblioteca del corso di laurea per eventuali chiarimenti/approfondimenti
- Appunti ispirati a: P. Armitage- G. Berry- Statistica medica. McGraw-Hill
- F. di Orio-Statistica Medica- Editore Nis

Ing. Andrea Ghedi AA 2009/2010

CHE COS' E' LA STATISTICA

**SCIENZA CHE STUDIA CON METODI
MATEMATICI FONDATI SUL CALCOLO
DELLE PROBABILITA' FENOMENI
COLLETTIVIE DI MASSA.**

(G.DEVOTO & G.C. OLI, Il dizionario della
lingua italiana, Le Monnier S.p.A., Firenze
1995).

Ing. Andrea Ghedi AA 2009/2010

CHE COS' E' LA STATISTICA

Da: "status" = condizione

COME CONCETTO GENERALE:

**"E' L'ARTE DI DESCRIVERE TUTTI GLI
OGGETTI IN RAGIONE DELLA LORO
QUALITA' "**

COME DEFINIZIONE PARTICOLARE:

**"DESCRIZIONE DELLE QUALITA' CHE
CARATTERIZZANO E DEGLI ELEMENTI
CHE COMPONGONO UNO STATO".** (Dal
greco: "statizein" = constatare, Larousse)

Ing. Andrea Ghedi AA 2009/2010

CHE COS' E' LA STATISTICA-CALCOLO DELLE PROBABILITA'

G. Cardano (1501-1576): **Liber de Ludo Aleae**
(1663)

- G. Galilei (1564-1642): **Sopra le Scoperte dei Dadi** (1613-23)
- Blaise Pascal (1623-1662): **Lettres a Pierre de Fermat** (1601-1665).
- **COME MAI OTTENERE UN "6" LANCIANDO 4 VOLTE UN SOLO DADO E' PIU' "FAVOREVOLE" CHE OTTENERE UN "DOPPIO 6" LANCIANDO 24 VOLTE DUE DADI ?** (1654) -[4: 6 = 24: 36] -Cavaliere de Méré

Ing. Andrea Ghedi AA 2009/2010

CHE COS' E' LA STATISTICA MEDICA

**SCIENZA / METODO IL CARATTERE
INDUTTIVO DELLA STATISTICA LA PONE
COME IL M E T O D O ESCLUSIVO PER
INVESTIGARE I FATTI NATURALI
QUALUNQUE SIA LA LORO SPECIE.
UN RAMO DEL SAPERE ESSENZIALMENTE
TEORICO CHE IMPIEGA MEZZI DI NATURA
LOGICA E MATEMATICA PER LA
RACCOLTA, IL "RAGGRUPPAMENTO",
L'INTERPRETAZIONE DEI DATI DI FATTO.**

Ing. Andrea Ghedi AA 2009/2010

CHE COS' E' LA STATISTICA MEDICA

Le statistiche mediche sono mere
affermazioni di carattere medico: quante
persone muoiono per una certa causa
ogni anno, quanti posti letto sono
disponibili in un ospedale in una certa
area, quanti migliorano dopo il trattamento
con un certo farmaco.

Ing. Andrea Ghedi AA 2009/2010

CHE COS' E' LA STATISTICA MEDICA

La statistica medica può essere definita
come la disciplina che concerne la
trattazione dei dati numerici derivanti dai
gruppi di singole unità statistiche o di
osservazione. Questi possono essere
persone affette da una certa malattia o
residenti in una certa area, possono
essere anche differenti momenti in cui una
determinata misurazione è stata effettuata
(prima e dopo una cura).

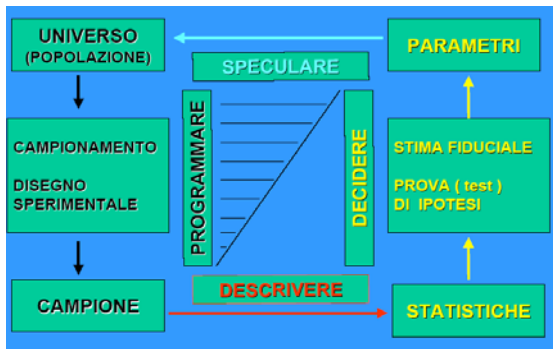
Ing. Andrea Ghedi AA 2009/2010

CHE COS' E' /NON E' LA STATISTICA MEDICA

- Se il medico in corsia scopre che in una certa circostanza un paziente con emicrania migliora dopo aver bevuto succo di prugne, da questa singola osservazione non deve scaturire che il succo di prugne è terapia utile per l'emicrania.
- Il medico ha bisogno di informazioni statistiche che dimostrino, per esempio, se in un gruppo di pazienti (campione) il miglioramento si registra con più frequenza dopo la somministrazione del succo di prugne rispetto a qualche altro trattamento alternativo (p.e. placebo)

Ing. Andrea Ghedi AA 2009/2010

FASI DEL LAVORO STATISTICO



Ing. Andrea Ghedi AA 2009/2010

FASI DEL LAVORO STATISTICO: PROGRAMMARE

- Disegnare e programmare lo studio
 - Scegliere dall' universo (o popolazione) il campione più adatto (campionamento)
- I criteri di selezione del campione possono essere vari ma tutti fondati sul principio secondo cui ogni unità della popolazione ha la stessa probabilità di entrare a far parte del campione
- Non sempre è possibile compiere osservazioni di una determinata caratteristica su tutto l'universo e ciò perché il collettivo è infinito o per ragioni di tempo o di costo (può essere necessario per l'osservazione distruggere l'unità)

Ing. Andrea Ghedi AA 2009/2010

FASI DEL LAVORO STATISTICO: STATISTICA DESCRITTIVA

- Sono le tecniche e gli strumenti fondamentali per la raccolta e la presentazione dei dati numerici
- Metodi per organizzare e sintetizzare l'informazione contenuta nei dati raccolti.
- Applicazione di " algoritmi " (= formule matematiche: + ; - ; x ; / ; $\sqrt{\quad}$).

Ing. Andrea Ghedi AA 2009/2010

FASI DEL LAVORO STATISTICO: STATISTICA INFERENZIALE

- Metodi per produrre inferenza, conclusioni di carattere generale oltre i limiti spazio-temporali in cui i dati sono stati generati. Il processo logico che consente di estendere le conclusioni ottenute dal campione alla popolazione è: **L'INFERENZA**
- Assunto di base perchè tale operazione sia realizzabile (non vi siano distorsioni - bias) è che il campione sia casuale.

Ing. Andrea Ghedi AA 2009/2010

DEFINIZIONI FONDAMENTALI

Popolazione statistica: l'insieme degli elementi che si vogliono considerare

Unità statistica: ogni elemento della popolazione statistica

Campione statistico: un sottoinsieme di unità statistiche, ottenuto dalla popolazione secondo una appropriata metodologia

Variabile statistica: un qualsiasi carattere rilevato sul campione che si è interessati a valutare per poter fare delle "considerazioni" (inferenza) sulla popolazione.

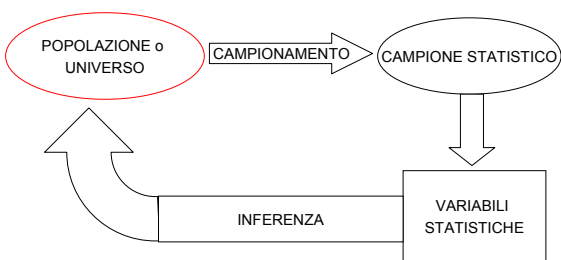
Ing. Andrea Ghedi AA 2009/2010

STATISTICA INFERENZIALE: L'INFERENZA STATISTICA

- Si definisce **INFERENZA STATISTICA** quell'insieme di tecniche che costituiscono il processo logico-operativo mediante il quale si perviene a certe conclusioni la cui validità per un collettivo più ampio (popolazione) è espresso in termini probabilistici.
- Le generalizzazioni statistiche riguardano certe caratteristiche della popolazione che definiremo **parametri** basati su caratteristiche del campione che chiameremo **statistiche**

Ing. Andrea Ghedi AA 2009/2010

TERMINOLOGIA STATISTICA: LA POPOLAZIONE



Ing. Andrea Ghedi AA 2009/2010

TERMINOLOGIA STATISTICA: LE VARIABILI

Si dicono variabili le caratteristiche rilevabili di un dato sistema oggetto di studio

- Es:
- Sesso e età di pazienti inclusi in una sperimentazione clinica controllata
 - Peso delle compresse di una specialità farmaceutica
 - Numero di fiale di una preparazione iniettabile

Ing. Andrea Ghedi AA 2009/2010

TERMONOLOGIA STATISTICA:LE VARIABILI

- I **valori numerici** (anni, grammi) o le **modalità** (maschio, femmina/vivo,morto) assunti dalle variabili costituiscono i **dati**, ricavabili da esperimenti scientifici o da indagini epidemiologiche
- I dati sperimentali si presentano sotto diverse forme essi possono essere sia di tipo **quantitativo** sia di tipo **qualitativo** ed essere espressi secondo **scale continue** o **scale discrete**

Ing. Andrea Ghedi AA 2009/2010

LE VARIABILI: VARIABILI CONTINUE/VARIABILI DISCRETE

- Si dice **continua** una variabile che può assumere un **qualsiasi valore reale** in un certo ambito (altezza, peso, età, pressione sanguigna)
- Esempio Paolo è alto 1.7577865680 m, i valori che la variabile continua può assumere sono legati alla precisione dello strumento di misura (se misuro Paolo con il metro da sarta avrò una misura di 1,77 m + e1+e2 in cui e1 è l'errore di misura dovuto alla misura e2 l'errore di misura dovuto alla precisione dello strumento)

Ing. Andrea Ghedi AA 2009/2010

LE VARIABILI: VARIABILI CONTINUE/VARIABILI DISCRETE

- Si dicono discrete quelle variabili che possono assumere solo certi valori cioè un numero finito
- Esempio: il numero di flaconi difettosi di un certo medicinale tra 100 flaconi, il numero di cavie sopravvissute ad un certo farmaco tra 20 cavie, il numero di attacchi di angina di un determinato paziente in una settimana
- Discrete sono anche le variabili che esprimono qualità o modalità: il sesso (femmina/maschio), il gruppo sanguigno (O,A,B,AB)

Ing. Andrea Ghedi AA 2009/2010

LE VARIABILI: VARIABILI ORDINABILI

- Non sempre le variabili continue sono misurabili in modo quantitativo anche se i loro valori possono essere disposti in ordine di grandezza
Esempio: Intensità del dolore toracico nel paziente infartuato

Nulla<lieve<moderato<forte

Alle modalità posso assegnare un punteggio (nulla=0 lieve=1 moderato=2 forte=3)

che non ha però significato quantitativo, il dolore 2 non è il doppio del dolore 1.

Ing. Andrea Ghedi AA 2009/2010

LA RACCOLTA DEI DATI

Come si può raccogliere il dato statistico?

- Interviste faccia a faccia (FTF Face to face)
- Interviste telefoniche
- Questionari AUTOCOMPILATI (Questionari postali e questionari elettronici)

Il modo di somministrare un questionario condiziona in parte lo strumento

Ing. Andrea Ghedi AA 2009/2010

LA RACCOLTA DEI DATI:INTERVISTE FACCIA A FACCIA (FTF)

Nell'intervista faccia a faccia una posizione assolutamente centrale è assunta dall'intervistatore. Dalla sua prestazione dipende buona parte della qualità dell'intervista.

E' quindi necessario un approccio oggettivistica (intervistatore neutrale e distaccato) e un questionario standardizzato.

Si deve limitare l'effetto dell'intervistatore, standardizzandone il comportamento, limitandone i margini di discrezionalità attraverso una fase di addestramento.

Non si deve trasmettere all'intervistato l'idea che esistano risposte giuste o sbagliate

Ing. Andrea Ghedi AA 2009/2010

LA RACCOLTA DEI DATI:INTERVISTE TELEFONICA, VANTAGGI

A causa di una serie di fattori concomitanti negli ultimi anni si sono sempre più diffuse le interviste telefoniche, esse presentano **vantaggi**:

Rapidità di rilevazione

Costi ridotti

Minori resistenze alla concessione

Possibilità di raggiungere tutta la popolazione (interviste FTF spesso sono nel centro delle città)

Inserimento immediato dei dati a computer

Ing. Andrea Ghedi AA 2009/2010

LA RACCOLTA DEI DATI:INTERVISTE TELEFONICA, SVANTAGGI

E svantaggi:

Manca il contatto personale: l'intervistato è meno coinvolto

Non si possono raccogliere dati non verbali (abitazione, abbigliamento ecc)

Settori marginali della popolazione non sono raggiungibili

La limitatezza del tempo costringe a domande rapide e quindi elementari

Va tenuto presente che non esistendo un contatto fisico l'intervistato ha maggior tendenza a rispondere a caso

Ing. Andrea Ghedi AA 2009/2010

LA RACCOLTA DEI DATI:IL QUESTIONARIO AUTOCOMPILATO

E' il caso del questionario che il soggetto compila da solo senza che vi sia un intervistatore.

Il primo vantaggio è decisamente l'economia, sia di tempo che di risorse

Svantaggio è il fatto che l'intervistato compila da solo e può commettere errori

Il questionario autocompilato deve essere **breve, conciso e semplice**.

La rilevazione può essere di gruppo (compilato in classe tra studenti) o una rilevazione individuale dove è importante porre dei vincoli sul ritiro per aumentare la percentuale di risposte

Ing. Andrea Ghedi AA 2009/2010

LA RACCOLTA DEI DATI:IL QUESTIONARIO AUTOCOMPILATO POSTALE

E' un caso particolare di autocompilazione:

VANTAGGI:

Contenimento dei costi, comodità di tempo e luogo di compilazione per l'intervistato, garanzia di anonimato, quindi assenza di distorsioni dovute all'intervistatore (bias)

SVANTAGGI:

Bassa percentuale di risposte, distorsione del campione dovuta all'autoselezione (chi risponde non è mai un campione casuale ma autoselezionato (istruzione alta,età media ecc.)

Ing. Andrea Ghedi AA 2009/2010

DISTRIBUZIONE DI FREQUENZA

Vi è ora il problema di riassumere i dati qualitativi raccolti. Il compito principale è quello di contare le osservazioni nelle varie categorie e, al più, esprimerle come proporzione o percentuali dei totali al fine di facilitarne il confronto fra distribuzioni provenienti da diverse indagini.

Questi conteggi vengono chiamati **frequenze o frequenze relative**

Ing. Andrea Ghedi AA 2009/2010

Le possibili determinazioni di una variabile quantitativa continua o discreta possono essere organizzate dividendo l'intervallo dei valori in **categorie** o intervalli di classi o **modalità** dando così origine alle tabelle di frequenza

Ing. Andrea Ghedi AA 2009/2010

RIPASSO: LE PROPORZIONI

Regola N°1: il prodotto dei medi è sempre uguale al prodotto degli estremi

$$A:B=C:D \iff B \times C = A \times D$$

Esempio:

$$A:x=C:D \quad x=?????$$

$$A \bullet D = x \bullet C$$

$$x = \frac{A \bullet D}{C}$$

Ing. Andrea Ghedi AA 2009/2010

DISTRIBUZIONE DI FREQUENZA

Esempio n°1:
 Risultato degli esami sull'espettorato dopo 3 mesi di trattamento con streptomina su un gruppo di pazienti.

	Frequenza	Percentuale
Striscio negativo, coltura negativa	141	45
Striscio negativo, senza coltura	90	28.8
Striscio o coltura positivi	82	26.2
Totale	313	100

$$141 : x = 313 : 100 \quad x = \frac{141 \bullet 100}{313} = 45,047$$

Ing. Andrea Ghedi AA 2009/2010
