APPUNTI DI BIOSTATISTICA

master COPY

Corso del prof S. Presciuttini

AA 2016-2017

master COPY

master COPY

DISCLAIMER: Questi sono appunti tratti dal corso di Biostatistica del professor Presciuttini (ormai in pensione). Utili per integrare e per chiunque sia negato in statistica.

Argomenti trattati:

1. Introduzione
2. Il test di significatività
3. Il test delle ipotesi di Pearson-Neyman
4. I risultati "non significativi" e la potenza statistica
5. Tabelle di contingenza
6. Il X-quadro
7. I dati della statistica
8. Distribuzioni statistiche
9. La scelta del test statistico
10. Media e varianza
11. La distribuzione normale
12. La normale standardizzata
13. L'intervallo di confidenza
14. T di student
15. Confronti fra due medie
16. Confronti multipli e *data dredging*
17. L'analisi della varianza
18. Test non parametrici – reprise
19. Correlazione
20. Regressione lineare

master COPY
master COPY
master COPY

1 INTRODUZIONE

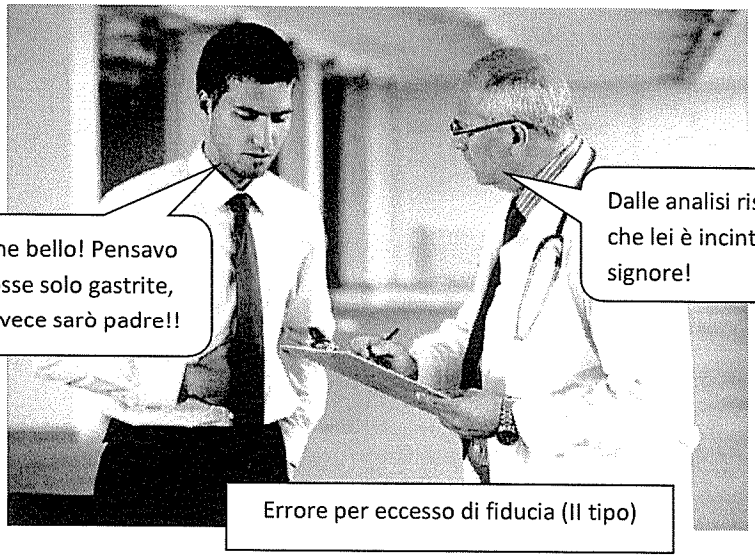
Le informazioni che noi essere umani ci scambiamo sono riconducibili ad **ASSERZIONI**, cioè espressioni linguistiche dotate di significato. Nella logica matematica, prendono il nome di **PROPOSIZIONI** e possono essere stabilite come "vere" o "false".

"Mi sento male" (asserzione) ---> "Il soggetto X è affetto da malessere" (proposizione)

La questione del vero/falso è stata, anche nel mondo animale, di vitale importanza nella sopravvivenza -> alcuni animali imitano il verso di altri per ingannarli.

Si possono commettere, in questa ottica, due tipi di errore:

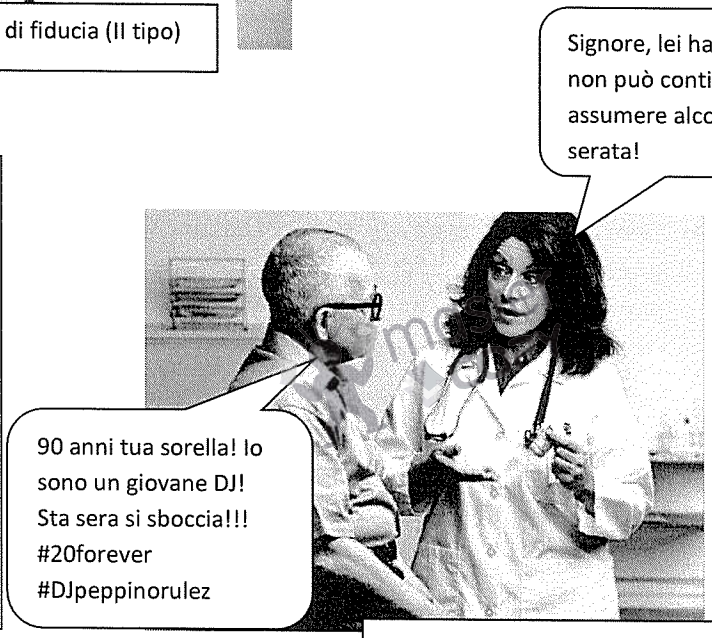
- Errore per eccesso di scetticismo/incredulità (rifiuto una cosa vera, ritenendola falsa).
- Errore per eccesso di ingenuità/fiducia (accetto una cosa falsa, ritenendola vera).



Errore per eccesso di fiducia (II tipo)

	LA PROPOSIZIONE È FALSA	LA PROPOSIZIONE È VERA
RIFIUTO LA PROPOSIZIONE	OK	Errore I tipo
ACCETTO LA PROPOSIZIONE	Errore II tipo	OK

NB. Il concetto di errore verrà rispiegato più avanti in relazione al concetto di "ipotesi nulla (H0)"



Errore per eccesso di scetticismo (I tipo)

I due tipi di errore possono portare a conseguenze diverse, a seconda del contesto:

- Se testando un farmaco si commettesse errore del I tipo (eccesso di scetticismo), potremmo ritenerlo totalmente inefficace (prendiamo per falsa la proposizione "il farmaco è efficace", quindi lo scartiamo dai trials successivi).
- Se un giudice commettesse errore del II tipo (eccesso di fiducia), potrebbe condannare un innocente (prende per vera la proposizione "x è colpevole", quando invece X è innocente).

Nella valutazione dei dati disponibili, si può ricadere in uno dei due tipi di errori per diversi motivi:

- I dati pervenuti contengono **errori di tipo metodologico**.
- I dati sono stati ottenuti correttamente, ma il caso ha voluto che fossero "strani" (ossia che andassero in senso opposto a quanto ci aspettavamo). In questo caso si parla di **errore di tipo statistico**.

L'errore di tipo metodologico si può ridurre al minimo, se non eliminare del tutto.

L'errore di tipo statistico invece è sempre presente, in quanto deriva dal processo stesso di campionamento e dalla variabilità dei dati (il caso non si controlla insomma!) -> è qui che entra in gioco la **statistica**: il suo **scopo è valutare l'errore statistico** e **quindi valutare la possibilità di compiere errori del I o del II tipo**.

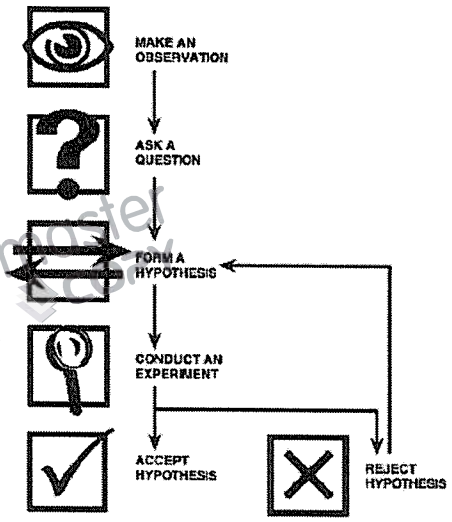
1.1 Il metodo scientifico ed il ruolo della biostatistica

Il metodo scientifico è la modalità tipica con cui la scienza procede per raggiungere una conoscenza della realtà oggettiva, affidabile, verificabile e condivisibile.

Consiste, da un lato, nella raccolta di evidenze empiriche e misurabili, dall'altro nella formulazione di ipotesi e teorie più generali, spesso sotto forma di **LEGGI UNIVERSALI**.

Nelle discipline biomediche la questione dell'estrapolazione di leggi matematiche è praticamente inesistente: solo dopo aver raccolto una serie di fatti che consideriamo accertati, ci si pone il compito di darne un'interpretazione razionale e consistente.

"L'essenza del metodo scientifico sta, nelle scienze biomediche, nell'asserire la verità o la falsità delle affermazioni empiriche con cui via via troviamo a confrontarci"



Nonostante il processo del metodo scientifico, ogni conclusione è sempre soggetta ad un **LIVELLO DI INCERTEZZA** -> pochi dati potrebbero confermare un'ipotesi, ma poi con l'aumentare delle osservazioni l'ipotesi potrebbe essere smentita!

Ad esempio, mettiamo che normalmente alle 10 arrivi un SMS pubblicitario sul cellulare. Con il passare dei giorni, uno arriverà a pensare, quando gli arriva un SMS alle 10, che probabilmente sarà pubblicità.

In questo caso dall'osservazione di un fenomeno ripetuto (SMS pubblicitario alle 10) si è arrivati all'ipotesi "*gli SMS che arrivano alle 10 probabilmente sono pubblicità*" -> dietro questa affermazione c'è sempre un po' di incertezza, visto che potrebbe capitare una volta che l'SMS ce lo abbia invece inviato un nostro amico!

La certezza dell'ipotesi aumenterà ogni qualvolta il fenomeno si ripete, ma comunque non si potrà mai avere la certezza al 100%!

La scienza biomedica procede soprattutto mediante il processo di **INDUZIONE** -> osservando casi particolari, se una tale cosa si verifica sempre con le stesse condizioni, si tende a generalizzare -> più sono i dati, più siamo sicuri, ma non potremo mai essere sicuri al 100% (dal caso particolare, alla regola generale).

CASO: Prendo i calzini dal cassetto

RISULTATO: I calzini sono bianchi.

REGOLA: Tutti i calzini nel cassetto sono bianchi.

La logica matematica, dal canto suo, procede per **DEDUZIONE** -> da una o più premesse date, si arriva ad una conclusione che è una conseguenza necessaria (dalla regola generale, al caso particolare).

REGOLA: Tutti i calzini nel cassetto sono bianchi

CASO: Prendo i calzini dal cassetto.

RISULTATO: I calzini sono bianchi.

Alla fine della fiera, il compito del metodo scientifico, nelle discipline biomediche, è di provvedere agli strumenti adatti per decidere se una data ipotesi è accettabile o meno.

Visto che non si può parlare di certezze assolute (), il ruolo della biostatistica è di ***fornire gli strumenti adatti a quantificare il nostro livello di incertezza sulle affermazioni nate dall'esperienza (in ambito medico).***

master COPY

master COPY