



Facoltà di Medicina e Chirurgia a.a. 2008-2009
Corso di Laurea Specialistica in Medicina e Chirurgia
Corso Integrato di Matematica, Fisica e Statistica Medica
Insegnamento di Statistica Medica (5 CFU)
docenti: prof. Lucio Torelli, dott. Massimo Borelli

Finalità dell'insegnamento

Il ciclo di lezioni si propone di introdurre gli Studenti al linguaggio della Statistica Medica, presentando quali siano le tecniche di base della statistica descrittiva e della statistica inferenziale che sono utilizzate dai ricercatori per validare e pubblicare i risultati dei loro studi.

Obiettivi dell'insegnamento

A conclusione del corso, gli Studenti saranno in grado di individuare all'interno di un articolo scientifico gli strumenti utilizzati della statistica descrittiva ed inferenziale, univariata e bivariata. Sapranno inoltre produrre analisi statistiche univariate e bivariate per mezzo del pacchetto statistico **R** (www.r-project.org), valutando criticamente i risultati ottenuti.

Dettagliatamente, gli Studenti dovranno esser in grado di:

- spiegare i concetti di: unità statistica, popolazione, caratteri statistici, tipi di variabili
- utilizzare le metodiche di descrizione dei dati: distribuzioni di frequenza, tabelle e grafici
- descrivere le misure di tendenza centrale e di posizione; individuare gli indici di variabilità delle distribuzioni; dare le definizioni di probabilità; indicare le operazioni sulle probabilità; definire il concetto di indipendenza
- conoscere l'utilizzazione del teorema di Bayes
- utilizzare dati epidemiologici per la stima della probabilità a priori di malattia
- calcolare sensibilità, specificità e valori predittivi di segni, sintomi e test diagnostici per stimare la probabilità a priori di malattia
- definire le variabili aleatorie discrete e continue; conoscere le principali distribuzioni di probabilità per l'interpretazione dei fenomeni biologici: binomiale, poisson, normale, t , Chi quadro, F ; definire il concetto di valore atteso e di varianza della variabile aleatoria
- definire la popolazione e il campione, rappresentare la distribuzione di campionamento; definire l'errore standard; illustrare le metodiche puntuali e intervallari (classiche e bootstrap) per la stima dei parametri nella popolazione
- definire la logica del test di ipotesi: ipotesi nulla e ipotesi alternativa; definire il livello di significatività e spiegare il significato di potenza di un test
- analizzare il confronto tra più gruppi indipendenti e dipendenti per dati quantitativi e qualitativi: discutere le modalità d'interpretazione dei risultati;
- analizzare la relazione tra variabili quantitative, qualitative e miste; spiegare l'utilizzabilità del modello di regressione lineare semplice
- utilizzare un word-processor e un foglio elettronico
- utilizzare il pacchetto statistico **R**
- utilizzare i motori di ricerca Google

Sito web, modalità d'esame e materiali didattici dell'Insegnamento

<http://www.dmi.units.it/~borelli>

Libri di testo consigliati

- Martin Bland, An Introduction to Medical Statistics, OUP
- Sergio Invernizzi, Maurizio Rinaldi, Andrea Sgarro, Moduli di matematica e statistica, Zanichelli
- Stefano Iacus, Guido Masarotto, Laboratorio di statistica con R, McGraw-Hill

Programma dell'insegnamento

Prerequisiti di Matematica

Le percentuali. Esponenti e logaritmi; il grafico della funzione esponenziale e il grafico della funzione logaritmo; alcune applicazioni: emivita dei radiofarmaci, la legge di percezione sensoriale di Weber-Fechner, il pH. La derivata: la pendenza di una retta, la retta tangente ad una curva è un ottimo modello in senso locale, la "formula dei tre punti". L'integrale: il metodo dei trapezi; il metodo MonteCarlo.

Statistica Descrittiva

I dati sperimentali: popolazione, campione, unità statistiche; scale di misura; dati simulati casualmente; distribuzioni uniformi e non uniformi dei dati; frequenze assolute e relative. Gli indici di centralità: la media (aritmetica), la mediana, la moda. Gli indici di dispersione: il range, i quartili, la differenza interquartile, i quantili, i decili, i percentili, la varianza della popolazione e del campione, la deviazione standard della popolazione e del campione, il coefficiente di variazione. I grafici: diagramma a barre (a bastoncini), torta, istogramma, il boxplot. Gli indici della statistica bivariata: la covarianza, il coefficiente di correlazione lineare di Karl Pearson, la correlazione secondo Kendall e Spearman; la retta di regressione lineare; la regressione nonlineare, la regressione LOESS.

Teoria della Probabilità

La probabilità intesa come stima statistica. I termini esperimento, evento, spazio campionario, evento impossibile, evento certo. La probabilità dell'evento complementare e dell'unione di due eventi: eventi compatibili ed incompatibili. La probabilità dell'intersezione di due eventi: eventi dipendenti ed indipendenti. La probabilità condizionata. Il rischio relativo. La "regola" di Bayes, probabilità a priori ed a posteriori. La probabilità ed i test diagnostici: prevalenza, incidenza, sensibilità, specificità, valori predittivi, la curva R.O.C. Il rischio relativo e lo odds ratio. Il concetto intuitivo di variabile aleatoria. La densità di probabilità e funzione di ripartizione in una variabile aleatoria discreta. La speranza matematica e la varianza di una variabile aleatoria. La variabile aleatoria binomiale di Bernoulli, la variabile aleatoria di Poisson. Le variabili aleatorie continue: funzione di densità, funzione di ripartizione, quantile. La variabile aleatoria normale, t di Student, Chi quadrato di Pearson, F di Snedecor.

Statistica Inferenziale

Le stime puntuali e le stime intervallari dei parametri statistici. La tecnica bootstrap. Il concetto di test statistico: l'ipotesi nulla, il livello di significatività, il consuntivo, il quantile, il p-value, la potenza di un test. Il test t di Student per un campione vs. la popolazione normale. Il test di Student - Welch per due campioni. Il test t di Student per dati accoppiati. Eteroschedasticità, test F e test di Fligner-Killen. Il QQ-plot. Il test di normalità di Shapiro - Wilk. Il test non parametrico dei ranghi di Wilcoxon e Mann-Whitney. Il test Chi quadrato di indipendenza di Pearson. Il test delle proporzioni. I test parametrici e non parametrici di correlazione di due variabili continue. La diagnostica del modello lineare: normalità dei residui con varianza costante e media nulla.

Metodologia Statistica.

Studi sperimentali e studi osservativi: in vitro, in vivo, in silico. Il problema dell'assegnazione nei clinical trials, la random allocation (randomization). I bias, gli esperimenti in "doppio cieco" (double blind, double masked). Studi cross-section, di coorte, caso-controllo. Studi prospettici e retrospettivi.