

STATISTICA
corso integrato di *Matematica e Statistica*

Codice Disciplina : BT0010 Codice Corso integrato : BT008	Settore: MED/01	N° CFU: 4
Docente: Magnani Corrado	Ore : 40	Corso di Laurea : Biotechnologie Anno : I Semestre : I

Obiettivo del modulo

Acquisire abilità in merito alla descrizione di dati statistici.
Comprendere i fondamenti della valutazione di probabilità di un evento.
Comprendere i fondamenti dell'inferenza statistica.
Saper applicare ed interpretare in modo appropriato alcuni metodi e test statistici.
In sintesi si intende fornire le basi necessarie per:

- la lettura di articoli scientifici di interesse biotecnologico;
- l'analisi e la presentazione di semplici serie di dati;
- l'interpretazione dei risultati di esami di laboratorio e test clinici

Conoscenze ed abilità attese

Le competenze matematiche e logiche previste per un diplomato di scuola superiore.

Programma del corso

Concetti generali:
Presentazione dei dati e loro classificazione.

Sintesi e presentazione dei dati categorici:
Frequenze assolute, relative, cumulative.
Tabelle di contingenza.
Rappresentazioni grafiche di distribuzioni di frequenza.

Sintesi e presentazione dei dati numerici:
Indici di posizione (media, mediana, moda) e di dispersione (deviazione standard, coefficiente di variazione, percentili).
Rappresentazioni grafiche di dati numerici.
Box-plot
Media geometrica, trasformazione logaritmica.
Calcolo della media per dati raggruppati.
Diagrammi a punti

Probabilità:
Definizione di probabilità.
Calcolo della probabilità di un evento e della probabilità congiunta di più eventi.
Probabilità condizionate
Applicazione: sensibilità, specificità e valore predittivo di un test diagnostico.
Probabilità congiunta di eventi indipendenti.
Applicazione: Valutazione del risultato di più tests diagnostici.

Distribuzioni teoriche di probabilità:
Distribuzione binomiale. Calcolo della probabilità di un evento con distribuzione di probabilità binomiale.
Distribuzione normale

Distribuzione normale standard. Uso delle tavole della distribuzione normale standard.

Campionamento e distribuzione campionaria della media:

Definizione di popolazione e campione.

Perché estrarre un campione.

Relazione tra popolazione e campione e proprietà delle statistiche campionarie.

Teorema del limite centrale.

Applicazioni del teorema del limite centrale.

Dimensione del campione

Metodi campionamento

La distribuzione t di Student:

Uso delle funzioni statistiche di EXCEL nel caso della distribuzione normale e della distribuzione t.

Intervalli di confidenza:

Calcolo dell'intervallo di confidenza per la media, nei casi con varianza nella popolazione nota ed ignota.

Generalizzazione del calcolo dell'intervallo di confidenza ad altre statistiche.

Test dell'ipotesi:

Introduzione all'inferenza statistica

Ipotesi 'di lavoro' ed ipotesi nulla

Errori di I° e II° tipo

Confronto tra due medie:

Uso del Test t-Student.

Campione con osservazioni appaiate.

Campioni indipendenti con uguale varianza.

Intervallo di confidenza in base alla distribuzione t.

Uso delle tavole della distribuzione t e delle funzioni di Excel.

Analisi dei dati in tabelle di contingenza :

Tabelle di contingenza a due righe e due colonne (2x2).

Indicatori di associazione: Odds Ratio e Rischio Relativo.

Intervallo di confidenza dell'Odds ratio

Test Chi-quadrato Uso delle tavole della distribuzione Chi-quadrato. Uso delle funzioni di Excel.

Correzione per la continuità

Test esatto di Fisher

Test di McNemar

Estensione alle tabelle n x m.

Analisi della varianza:

Analisi della varianza ad un criterio di classificazione.

Correzione di Bonferroni.

Regressione lineare semplice:

La retta di regressione.

La stima con il metodo dei minimi quadrati.

Uso dei residui per verificare la validità delle assunzioni.

Intervallo di confidenza dei valori predetti in base alla retta di regressione.

Come si legge il risultato di un'analisi multivariata

Introduzione al disegno dello studio.

Potenza statistica.

Argomenti di studio individuale sul testo consigliato

Tests 'distribution free' o non parametrici:

Introduzione generale, Vantaggi e svantaggi, Confronto con i corrispondenti tests parametrici.

Test U di Mann-Whitney

Test di Wilcoxon per dati appaiati
Correlazione per ranghi (test di Spearman).

Esercitazioni

Un primo ciclo di esercitazioni (complessive 9 ore a studente) prevede lo svolgimento di esercizi in aula, sotto la guida di un docente.

Un successivo ciclo (complessive 7 ore a studente) prevede esercitazioni pratiche al computer, relative all'applicazione a dati reali dei metodi appresi.

Attività a scelta dello studente

Gli studenti hanno la possibilità di partecipare all'attività dell'Unità di Statistica Medica e di Epidemiologia dei Tumori. Le attività possibili variano in relazione agli studi in corso ed alla competenza degli allievi. L'elenco delle attività è disponibile su sito internet.

Supporti alla didattica in uso alla docenza

Presentazioni in formato MS-WORD ed MS-Power Point, siti internet didattici, programma elementare per l'elaborazione di dati ed analisi statistiche (disponibile per ciascuno studente anche per il PC personale). Per ciascuna lezione viene messo a disposizione un breve set di esercizi risolti, che vengono discussi in aula.

Strumenti didattici

Il materiale didattico usato a lezione viene messo a disposizione sul sito Internet del docente: <http://www3.med.unipmn.it/magnani/>
Per ciascuna lezione sarà messo a disposizione un breve set di esercizi risolti.
Gli allievi sono tenuti a procurarsi il testo di statistica consigliato.

Materiali di consumo previsti

Gli studenti hanno a disposizione sul sito del docente copia delle immagini proiettate a lezione. Le stampe sono effettuate dagli studenti a loro discrezione.

Eventuale bibliografia

Testo adottato:

M. Pagano & K. Gauvreau. Biostatistica (II edizione italiana). ed. Idelson Gnocchi, Napoli 2003.

Testi utili per un approfondimento:

- Daniel G., Biostatistica. Ed. Edises.
- Armitage & Berry, Metodi statistici nella ricerca medica, 3a edizione (ed inglese Blackwell, tradotto e stampato in Italia da McGraw-Hill).
- Siegel & Castellan, Statistica non parametrica. McGraw-Hill.
- Douglas Altman, David Machin, Trevor Bryant, Martin Gardner Statistics with Confidence (2nd ed.) BMJ edition. Tradotto in italiano.
- www.publichealth.ac.nz/

Si consiglia inoltre la lettura degli articoli sui metodi statistici pubblicati periodicamente dal British Medical Journal e dal British Journal of Cancer.

Verifica dell'apprendimento

Il modulo prevede una prova scritta.

Il testo dell'esame comprende:

domande aperte esercizi con calcoli questionario, di regola con richiesta di spiegazione della risposta fornita

Sono oggetto di esame tutti gli argomenti del programma, anche se non sono stati svolti a lezione.