



UNIVERSITÀ  
DEGLI STUDI DI MILANO-BICOCCA

## SYLLABUS DEL CORSO

### Analisi dei Dati

2021-1-F0901D043

---

#### Obiettivi

Il candidato sarà in grado di: comprendere aspetti basilari del disegno dello studio, implementare autonomamente analisi statistiche di base, leggere con spirito critico la letteratura scientifica che presenti analisi statistiche descrittive e inferenziali.

Il candidato sarà in grado di: acquisire le conoscenze ed i concetti di base relativi alle metodologie e alle tecniche computazionali per la raccolta, la gestione e l'analisi di dati in biologia molecolare, come i dati di sequenze generati dalle tecnologie Next Generation Sequencing (NGS), e la padronanza dei principali strumenti computazionali necessari per estrarre informazioni di interesse per la ricerca biomedica dalle principali banche dati di sequenze.

#### Contenuti sintetici

I due moduli di cui si compone il corso si propongono di contribuire alla formazione di un biotecnologo medico che sia in grado di:

- 1) comprendere i principi del disegno sperimentale in medicina e biologia
- 2) conoscere le principali tecniche di analisi statistica dei dati
- 3) utilizzare un software per l'elaborazione dei dati
- 4) compiere l'interpretazione critica dei risultati presentati nella letteratura scientifica.
- 5) essere introdotto alla bioinformatica: motivazioni, problemi e metodologie.

- 6) conoscere le tecnologie NGS
- 7) conoscere le principali basi di dati; accesso, interrogazione, inserimento dati
- 8) conoscere le principali tecniche di analisi dei dati: ricostruzione e annotazione di genomi; confronto di sequenze: algoritmi di allineamento globale, locale e multiplo; ricostruzione di filogenie; analisi del trascrittoma.

## **Programma esteso**

Il modulo di Biostatistica si articola in tre parti: la prima relativa alla statistica descrittiva, la seconda alla statistica inferenziale, la terza relativa all'interpretazione di articoli scientifici. Le parti prima e seconda presentano le seguenti caratteristiche:

- 1) includono aspetti metodologici di disegno dello studio e programmazione dell'esperimento
- 2) vengono erogate nella forma di riflessione su particolari esempi applicativi
- 3) prevedono l'uso del pacchetto applicativo per l'analisi dei dati STATA

---

Parte seconda - Generalità sulla statistica inferenziale, Verifica di ipotesi nulle relative alla media di variabili continue, Test T in disegno semplice ed appaiato, Verifica di ipotesi nulle relative alla associazione per variabili categoriali, Test chi quadrato, Verifica di ipotesi nulle relative alla proporzione di variabili dicotomiche: Test McNemar, Cenni all'analisi della varianza, Studio della potenza del test e calcolo della dimensione del campione.

Parte terza - Lettura, interpretazione, riflessione metodologica basata su articoli scientifici.

---

1. La gestione dei dati nelle scienze della vita

2. L'informatica essenziale

---

---

---

---

---

---

---

---

---

---

—

—

—

—

—

—

—

—

—

—

—

—

—

—

—

—

—

—



[https://www.pearson.it/opera/pearson/0-6471-fondamenti\\_di\\_statistica\\_per\\_le\\_discipline\\_biomediche](https://www.pearson.it/opera/pearson/0-6471-fondamenti_di_statistica_per_le_discipline_biomediche)

M. Helmer Citterich, F. Ferrè, G. Pavesi, C. Romualdi, G. Pesole, Fondamenti di bioinformatica (Zanichelli editore)

Dispense fornite dai docenti

### **Periodo di erogazione dell'insegnamento**

Primo semestre

### **Modalità di verifica del profitto e valutazione**

Prova scritta (Biostatistica) e Prova orale (Bioinformatica)

### **Orario di ricevimento**

Da definire con lo studente via email

---