

Elementi di Matematica e di Statistica

Introduzione alla Statistica

Docente: Riccardo levoli

Corso di Laurea in Biotecnologie
a.a. 2025-2026

01/10/2025

Qualche domanda

- Cos'è la statistica?
- Perché devo studiarla?
- Quanto serve sapere la matematica?

Alcuni miti da sfatare

Nell'immaginario comune gli statistici si occupano di:

- Nel calcio: possesso palla (che poi forse non serve, boh?);
- Diagrammi a torta;
- Sondaggi (ci “prendono e non ci prendono”);
- Dire cose che poi non si avverano;
- Polli (?) (vedi poesia Trilussa);
- Manipolare i dati a loro piacimento (o per soddisfare esigenze di terzi);
- Altro?

Cos'è la Statistica?

Prima di tutto è una scienza. In qualche modo, se la matematica è la scienza “dei numeri” (visti in ottica “deterministica”) la statistica è la scienza dell'**incertezza**.

Due principali prospettive della statistica:

- **Descrittiva**: dalle informazioni disponibili sotto forma di dato si propongono misure di sintesi che permettono di comprendere un fenomeno.
- **Inferenziale**: dai dati osservati si può stimare l'entità del fenomeno di interesse, estendendo le conclusioni anche a situazioni non direttamente osservate (generalizzazione).

Cos'è la Statistica?

Parte 2

Semplificando:

- Un medico si prende cura dei suoi pazienti così come un biotecnologo dei suoi strumenti, materiali di laboratorio, soluzioni innovative, ecc.
- Uno statistico si prende cura dei dati: ne comprende i limiti e cerca di ricavare informazioni sul fenomeno oggetto di studio.

Aspetti da tenere in considerazione:

- Rilevanza della qualità dei dati
- Saper cogliere i limiti dell'esperimento, della ricerca, dei dati stessi.

Cos'è la Statistica?

Un esperimento con Wooclap

`https:
//app.wooclap.com/OGDRP0?from=event-page`

Perché studiare Statistica?

Una parte delle innovazioni e delle scoperte effettuate dalle scienze (es: biologia declinata poi nelle biotecnologie) si basano su **ricerche empiriche** (prospettiva *evidence-based*).

Le fasi ideali della ricerca

- 1 Definizione teorica del **problema**
- 2 Scelta del **metodo**
- 3 Raccolta dei dati (operativizzazione + scelta degli strumenti)
- 4 **Analisi** dei dati
- 5 Presentazione e Interpretazione dei **risultati**
- 6 **Conclusioni**

Nei punti da 2 a 5 la statistica è fondamentale

Quanto serve sapere la Matematica?

Nelle 24 ore a disposizione cercheremo di comprendere alcuni concetti fondamentali legati alle tecniche statistiche di largo utilizzo anche nelle Biotecnologie:

- Il livello matematico richiesto sarà pressoché identico a quello che studieremo nelle lezioni di matematica
- Non affronteremo molti dei dettagli metodologici che si basano su una eccessiva formalizzazione
- Verrà promosso l'utilizzo di Microsoft Excel (o fogli di calcolo) per l'analisi dei dati.
- Saranno presentati altri software facoltativi (R Studio), ma tali software non saranno rilevanti ai fini del superamento dell'esame.

Obiettivo del Corso

Modulo di Statistica

L'obiettivo è quello di fornire le basi fondamentali per comprendere ed analizzare criticamente:

- Risultati provenienti da esperimenti;
- Analisi di dati (solitamente incluse nelle tesi di laurea/esame finale);

Nel vostro futuro, la statistica potrebbe assumere un ruolo sempre più rilevante.

Contenuti del Corso

Il Modulo di **Statistica** può essere suddiviso nei tre seguenti macro-argomenti:

- ➊ **Statistica descrittiva**: tecniche finalizzate a sintetizzare le informazioni presenti in una matrice di dati
- ➋ **Cenni di probabilità**: definizione e alcuni concetti chiave collegati al calcolo delle probabilità.
- ➌ **Inferenza Statistica**: partendo da un campione, si possono estendere le informazioni raccolte alla popolazione, sotto determinate assunzioni (*generalizzazione dei risultati*).

Per maggiori dettagli consultare il sito docente:

<https://docente.unife.it/riccardo.ievoli>

Sempre ricordando che

Teaching statistics



Doing Statistics



Alcuni concetti introduttivi

Esperimenti

Gli esperimenti presentano le seguenti componenti:

- **Unità statistiche:** entità su cui vengono osservate le grandezze studiate (ad esempio: pazienti, animali, provette).
- **Variabili** o caratteri: sono le grandezze o, più in generale, caratteristiche, registrate sulle unità statistiche (ad esempio: livello di pH, altezza, valori di emoglobina, genere).
- **Popolazione di interesse:** insieme di potenziali unità statistiche da cui provengono le unità inserite nello studio. Spesso è necessario fare i conti con alcuni limiti dovuti ai meccanismi con cui le unità entrano nello studio.

Alcuni concetti introduttivi

Variabili Quantitative e Qualitative

Le variabili, a seconda della loro natura, possono essere distinte in:

- **Variabili quantitative** (numeriche): caratteristiche che possono essere valutate numericamente (età, altezza, peso e altri parametri biometrici, numero di figli).
- **Variabili qualitative** (categoriche): caratteristiche che suddividono la popolazione in gruppi, presentando diverse modalità (genere, presenza malattia, titolo di studio).

Per modalità si intendono (tutti) i possibili valori che una variabile può assumere.

Alcuni concetti introduttivi

Variabili Quantitative: discrete e continue

Sono presenti due ulteriori sotto-categorie di variabili quantitative (numeriche):

- **Variabile quantitativa discreta:** assume valori appartenenti all'insieme dei numeri naturali $(0,1,2,\dots,N)$. Esempio: numero di animali in ogni allevamento.
- **Variabile quantitativa continua:** grandezza misurabile che assume valori nell'insieme dei numeri reali (altezza, temperatura, prezzo, valori di emoglobina).

Nella pratica le variabili continue vengono di fatto discretizzate (troncamento). Ad esempio, si sceglie un numero finito di decimali dopo la virgola.

Alcuni concetti introduttivi

Variabili Qualitative: nominali e ordinabili

Si possono individuare due tipologie di variabili qualitative (categoriche):

- **Variabile qualitativa nominale:** non può essere individuato un ordinamento tra le modalità (colore occhi, genere, presenza malattia, professione). Caso particolare: variabile con sole due modalità. In questo caso si parla di variabile dicotomica (binaria o *dummy*, esempio: paziente malato oppure sano).
- **Variabile qualitativa ordinabile:** si può individuare un ordinamento tra le modalità. Si precisa che la distanza tra le modalità non può essere quantificata dal punto di vista numerico. Esempi: titolo di studio, giudizio.

Alcuni concetti introduttivi

Variabili: gerarchia e notazione

Si può notare che le tipologie di variabili definite hanno la seguente gerarchia informativa crescente:

1. Variabile qualitativa nominale
2. Variabile qualitativa ordinale
3. Variabile quantitativa discreta
4. Variabile quantitativa continua

Una variabile può essere trasformata nelle tipologie che la precedono (perdendo informazione), ma non in quelle che seguono. Notazione.

Si è soliti indicare le variabili con lettere maiuscole (X , Y , Z) e le osservazioni nelle unità statistiche con lettere minuscole (x , y , z).

Alcuni concetti introduttivi

Variabili di Risposta e Variabili Esplicative

A seconda del loro ruolo all'interno dell'esperimento, è possibile inoltre distinguere tra:

- **Variabile di risposta:** è l'oggetto di studio vero e proprio.
- **Variabili esplicative:** caratteristiche delle unità che influenzano la variabile di risposta (ad esempio: età, genere, trattamento).

Esempi di **variabili di risposta**: a. prezzo (o valutazione) di un appartamento, b. comparsa (o presenza) di una determinata malattia.

- a. le variabili esplicative possono essere: m^2 , numero di camere, numero di bagni, zona (quartiere), esposizione, ...
- b. le variabili esplicative possono essere: genere, età, eccesso di fumo e/o alcol, patologie pregresse, ...

Alcuni concetti introduttivi

I concetti di Popolazione e Campione

- **Popolazione:** insieme di tutte le unità statistiche che sono oggetto di interesse
- **Campione:** sottoinsieme di unità statistiche effettivamente osservate appartenenti alla popolazione di riferimento

Il campione è necessario per studiare il fenomeno se non è possibile ottenere (censire) tutte le unità appartenenti alla popolazione.

Di solito, le caratteristiche della popolazione (parametri) vengono **stimate** attraverso l'uso di un campione.

Parlare di stima implica ammettere che i risultati ottenuti siano affetti da una sorgente di errore: l'errore di campionamento.

Più il campione è numeroso, minore sarà l'entità di questo errore.

Alcuni concetti introduttivi

La scelta del Campione

Qual è il campione migliore? Un campione estratto in maniera casuale. Ovvero, ad ogni elemento della popolazione si associa una probabilità di essere incluso nel campione.

Perché è preferibile? Previene eventuali distorsioni ed errori di campionamento. Nella pratica non è sempre possibile ottenere un campione casuale. Spesso si dispone di *campioni di comodo*.

Quali sono i problemi dei campioni di comodo? Le unità estratte sono solitamente non **indipendenti** tra loro e il criterio di inclusione adottato può portare a delle distorsioni sistematiche.

Primi cenni alle Distribuzioni Statistiche

Visualizzare la distribuzione per età del corso

LINK:

`https://app.wooclap.com/OGDRP0?from=
event-page`

Riassunto: cosa abbiamo imparato oggi?

- 1 Il ruolo della Statistica
- 2 Componenti di un esperimento
- 3 Tipologie di variabili: quantitative vs qualitative
- 4 Variabili di risposta e variabili esplicative
- 5 Relazione tra popolazione e campione

Prova introduttiva

Andate al seguente Link per svolgere un'esercitazione introduttiva di statistica.

Cliccare **QUI**

Avete a disposizione circa 30 minuti per la prova (15 domande)