

# Calcolo delle Probabilità e Statistica

Stefania Bartoletti

28 Febbraio 2021

# Informazioni

► **Docente:** `stefania.bartoletti@unife.it`

# Informazioni

- ▶ **Docente:** stefania.bartoletti@unife.it
- ▶ **Orario lezioni:**

# Informazioni

- ▶ **Docente:** stefania.bartoletti@unife.it
- ▶ **Orario lezioni:**
  - ▶ Lunedì 11:30 – 13:00

# Informazioni

- ▶ **Docente:** stefania.bartoletti@unife.it
- ▶ **Orario lezioni:**
  - ▶ Lunedì 11:30 – 13:00
  - ▶ Mercoledì 9:00 – 11:00

# Informazioni

- ▶ **Docente:** stefania.bartoletti@unife.it
- ▶ **Orario lezioni:**
  - ▶ Lunedì 11:30 – 13:00
  - ▶ Mercoledì 9:00 – 11:00
- ▶ **Tutorato:** eddie.conti@edu.unife.it

# Informazioni

- ▶ **Docente:** stefania.bartoletti@unife.it
- ▶ **Orario lezioni:**
  - ▶ Lunedì 11:30 – 13:00
  - ▶ Mercoledì 9:00 – 11:00
- ▶ **Tutorato:** eddie.conti@edu.unife.it
- ▶ **Ricevimento:** Hangouts/Skype su appuntamento

# Probabilità vs. Statistica



# Probabilità vs. Statistica

- ▶ Il calcolo delle probabilità e la statistica sono discipline distinte ma profondamente connesse: i concetti e le definizioni della statistica si basano su concetti e definizioni di probabilità.

# Probabilità vs. Statistica

- ▶ Il calcolo delle probabilità e la statistica sono discipline distinte ma profondamente connesse: i concetti e le definizioni della statistica si basano su concetti e definizioni di probabilità.
- ▶ Il calcolo delle probabilità è **logicamente autosufficiente**. Si basa su poche leggi formali da seguire logicamente.

# Probabilità vs. Statistica

- ▶ Il calcolo delle probabilità e la statistica sono discipline distinte ma profondamente connesse: i concetti e le definizioni della statistica si basano su concetti e definizioni di probabilità.
- ▶ Il calcolo delle probabilità è **logicamente autosufficiente**. Si basa su poche leggi formali da seguire logicamente. Il calcolo tuttavia può risultare abbastanza complesso.

# Probabilità vs. Statistica

- ▶ Il calcolo delle probabilità e la statistica sono discipline distinte ma profondamente connesse: i concetti e le definizioni della statistica si basano su concetti e definizioni di probabilità.
- ▶ Il calcolo delle probabilità è **logicamente autosufficiente**. Si basa su poche leggi formali da seguire logicamente. Il calcolo tuttavia può risultare abbastanza complesso.
- ▶ La statistica si occupa della raccolta dei dati, della loro descrizione e della loro analisi. I concetti e le definizioni della statistica ci guidano nel trarre conclusioni dai dati. Viene pertanto definita **l'arte di apprendere dai dati**.

# Esempio di calcolo delle Probabilità

- ▶ Si lancia una moneta 100 volte. Sappiamo che ad ogni lancio, la probabilità che risulti “testa” o “croce” è uguale.

# Esempio di calcolo delle Probabilità

- ▶ Si lancia una moneta 100 volte. Sappiamo che ad ogni lancio, la probabilità che risulti “testa” o “croce” è uguale.
- ▶ Qual è la probabilità che risulti 60 volte “testa”?

# Esempio di calcolo delle Probabilità

- ▶ Si lancia una moneta 100 volte. Sappiamo che ad ogni lancio, la probabilità che risulti “testa” o “croce” è uguale.
- ▶ Qual è la probabilità che risulti 60 volte “testa”?
- ▶ **Esiste una sola risposta esatta** (circa 0.028444) e impareremo a calcolarla

# Esempio di Analisi Statistica

- ▶ Si lancia una moneta di provenienza incognita



# Esempio di Analisi Statistica

- ▶ Si lancia una moneta di provenienza incognita
- ▶ Per capire se la moneta è truccata si lancia 100 volte e si conta il numero di “teste”

# Esempio di Analisi Statistica

- ▶ Si lancia una moneta di provenienza incognita
- ▶ Per capire se la moneta è truccata si lancia 100 volte e si conta il numero di “teste”
- ▶ Immaginiamo di contare 60 “teste”.

# Esempio di Analisi Statistica

- ▶ Si lancia una moneta di provenienza incognita
- ▶ Per capire se la moneta è truccata si lancia 100 volte e si conta il numero di “teste”
- ▶ Immaginiamo di contare 60 “teste”.
- ▶ Uno statistico deve trarre una conclusione (**inferire**) da questo dato.

# Esempio di Analisi Statistica

- ▶ Si lancia una moneta di provenienza incognita
- ▶ Per capire se la moneta è truccata si lancia 100 volte e si conta il numero di “teste”
- ▶ Immaginiamo di contare 60 “teste”.
- ▶ Uno statistico deve trarre una conclusione (**inferire**) da questo dato.
- ▶ **Esistono diversi modi per procedere**: diversi calcoli di probabilità che possiamo applicare per giustificare diverse conclusioni. In generale, diversi statistici possono inferire conclusioni diverse.

# Probabilità vs. Statistica

- ▶ Nel primo esempio, le leggi che regolano l'esperimento sono note (**probabilità testa = 0.5**). Lo scopo è calcolare la probabilità che un determinato risultato (**almeno 60 teste**) possa accadere.

# Probabilità vs. Statistica

- ▶ Nel primo esempio, le leggi che regolano l'esperimento sono note (**probabilità testa = 0.5**). Lo scopo è calcolare la probabilità che un determinato risultato (**almeno 60 teste**) possa accadere.
- ▶ Nel secondo esempio, il risultato è noto (**60 teste**). Lo scopo è inferire le leggi che regolano l'esperimento (**probabilità testa**)

# Cenni Storici

- ▶ Il termine statistica deriva dalla parola **stato** e nella sua concezione originale indicava una raccolta di fatti di interesse per lo stato

# Cenni Storici

- ▶ Il termine statistica deriva dalla parola **stato** e nella sua concezione originale indicava una raccolta di fatti di interesse per lo stato
- ▶ In Egitto si rilevava l'ammontare della popolazione già ai tempi della prima dinastia e durante la seconda si rilevavano vari beni a fini fiscali.



# Cenni Storici

- ▶ Tale concezione della statistica è tuttora usata quotidianamente per descrivere i dati raccolti in un'indagine, esperimento o studio.

# Cenni Storici

- Tale concezione della statistica è tuttora usata quotidianamente per descrivere i dati raccolti in un'indagine, esperimento o studio.



HOME > DATI ANALISI E PRODOTTI > STATISTICHE A-Z

## STATISTICHE A-Z



### DATI ANALISI E PRODOTTI

#### BANCHE DATI

#### TAVOLE DI DATI

Bollettino mensile di statistica on line

#### MICRODATI

Riconoscimento

Normativa

#### COMUNICATI STAMPA

#### PUBBLICAZIONI

#### VISUALIZZAZIONI

Grafici interattivi

#### CONTENUTI INTERATTIVI

Contanomi

Questionari



[Abbigliamento](#) - 24 documenti

[Abitazioni](#) - 173 documenti

[Aborto](#) - 16 documenti

[Acqua](#) - 56 documenti

[Acque reflue](#) - 38 documenti

[Acquisti mercato elettronico](#) - 13 documenti

[Anagrafe](#) - 96 documenti

[Animali](#) - 6 documenti

[Annali](#) - 2 documenti

[Annuario Istat-Ice](#) - 7 documenti

[Annuario statistico italiano](#) - 8 documenti

[Anziani](#) - 43 documenti

# Statistica descrittiva

- ▶ La parte della statistica che si occupa di illustrare e sintetizzare i dati è detta **statistica descrittiva**

# Statistica descrittiva

- ▶ La parte della statistica che si occupa di illustrare e sintetizzare i dati è detta **statistica descrittiva**
- ▶ Si può partire da un campione di dati definito oppure può essere necessario **ideare il processo di raccolta dei dati**

# Statistica descrittiva

- ▶ La parte della statistica che si occupa di illustrare e sintetizzare i dati è detta **statistica descrittiva**
- ▶ Si può partire da un campione di dati definito oppure può essere necessario **ideare il processo di raccolta dei dati**
- ▶ La raccolta dei dati è spesso un problema difficoltoso; condurla in maniera appropriata è importante affinché l'indagine abbia senso

# Statistica Inferenziale

- ▶ Durante il XIX secolo, la teoria della probabilità era stata sviluppata da matematici come *Jacob Bernoulli*, *Karl Friedrich Gauss* e *Pierre-Simon Laplace* ma non era usata per studiare risultati statistici

# Statistica Inferenziale

- ▶ Durante il XIX secolo, la teoria della probabilità era stata sviluppata da matematici come *Jacob Bernoulli*, *Karl Friedrich Gauss* e *Pierre-Simon Laplace* ma non era usata per studiare risultati statistici
- ▶ Molti statistici di quel tempo sostenevano l'**autoevidenza dei dati**

# Statistica Inferenziale

- ▶ Durante il XIX secolo, la teoria della probabilità era stata sviluppata da matematici come *Jacob Bernoulli*, *Karl Friedrich Gauss* e *Pierre-Simon Laplace* ma non era usata per studiare risultati statistici
- ▶ Molti statistici di quel tempo sostenevano l'**autoevidenza dei dati**
- ▶ Alla fine del secolo, la statistica iniziò ad occuparsi di inferire conclusioni a partire da dati numerici



“La statistica è una disciplina scientifica che si occupa della raccolta, analisi ed **interpretazione** dei dati ottenuti da osservazioni sperimentali. Questa materia ha una struttura coerente che **si basa sulla teoria della probabilità** e include molte tecniche differenti che si affiancano alla ricerca e allo sviluppo **in tutti i campi della Scienza e della Tecnologia.**”  
(*E. Pearson, 1936*)

# Applicazioni

- ▶ I campi di utilizzo del calcolo delle probabilità e della statistica sono innumerevoli: ingegneria, medicina, scienze sociali, scienze della vita, economia e **informatica**

# Applicazioni

- ▶ I campi di utilizzo del calcolo delle probabilità e della statistica sono innumerevoli: ingegneria, medicina, scienze sociali, scienze della vita, economia e **informatica**
- ▶ Test dei farmaci, misure di legami genetici, modelli climatici, previsioni economiche, epidemiologia

# Applicazioni

- ▶ I campi di utilizzo del calcolo delle probabilità e della statistica sono innumerevoli: ingegneria, medicina, scienze sociali, scienze della vita, economia e **informatica**
- ▶ Test dei farmaci, misure di legami genetici, modelli climatici, previsioni economiche, epidemiologia
- ▶ Data science: data analytics, machine learning, artificial intelligence

# Diagramma di Venn per Data Science

For four years in a row, data scientist has been named the number one job in the U.S. by Glassdoor. What's more, the U.S. Bureau of Labor Statistics reports that the demand for data science skills will drive a 27.9 percent rise in employment in the field through 2026. Not only is there a huge demand, but there is also a noticeable shortage of qualified data scientists.<sup>1</sup>

---

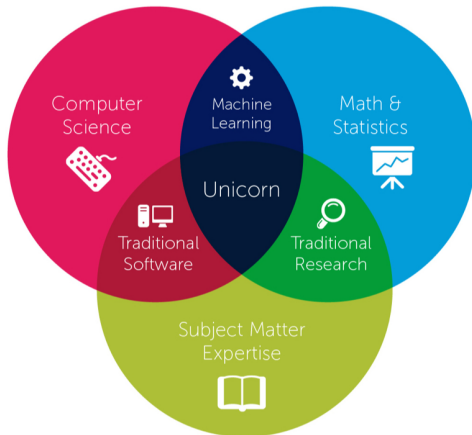
<sup>1</sup><https://www.northeastern.edu/>

# Diagramma di Venn per Data Science

**Diagramma di Venn (Eulero-Venn):** Diagramma che mostra tutte le possibili relazioni logiche tra una collezione finita di insiemi differenti

# Diagramma di Venn per Data Science

**Diagramma di Venn (Eulero-Venn):** Diagramma che mostra tutte le possibili relazioni logiche tra una collezione finita di insiemi differenti



# Toy Examples

- ▶ Data questa incredibile lista di applicazioni, come mai parleremo per settimane di **monete, dadi, urne e palline?**



# Toy Examples

- ▶ Data questa incredibile lista di applicazioni, come mai parleremo per settimane di **monete, dadi, urne e palline**?
- ▶ Perché ci permettono di rappresentare in maniera semplice e intuitiva moltissimi problemi complessi del mondo reale.

# Toy Examples

- ▶ Data questa incredibile lista di applicazioni, come mai parleremo per settimane di **monete, dadi, urne e palline**?
- ▶ Perché ci permettono di rappresentare in maniera semplice e intuitiva moltissimi problemi complessi del mondo reale.
- ▶ Una semplice moneta ci permette di rappresentare innumerevoli esperimenti con due possibili risultati: es. successo o fallimento di una terapia (o di un esame).

# Toy Examples

- ▶ Data questa incredibile lista di applicazioni, come mai parleremo per settimane di **monete, dadi, urne e palline**?
- ▶ Perché ci permettono di rappresentare in maniera semplice e intuitiva moltissimi problemi complessi del mondo reale.
- ▶ Una semplice moneta ci permette di rappresentare innumerevoli esperimenti con due possibili risultati: es. successo o fallimento di una terapia (o di un esame).
- ▶ A volte un problema è così complesso che è difficile trovare un modello per rappresentarlo. Una soluzione alternativa sono le simulazioni al computer.

# Toy Examples

- ▶ Data questa incredibile lista di applicazioni, come mai parleremo per settimane di **monete, dadi, urne e palline**?
- ▶ Perché ci permettono di rappresentare in maniera semplice e intuitiva moltissimi problemi complessi del mondo reale.
- ▶ Una semplice moneta ci permette di rappresentare innumerevoli esperimenti con due possibili risultati: es. successo o fallimento di una terapia (o di un esame).
- ▶ A volte un problema è così complesso che è difficile trovare un modello per rappresentarlo. Una soluzione alternativa sono le simulazioni al computer.
- ▶ In particolare, si simula la realizzazione di diversi esperimenti e si stimano le probabilità a partire dai risultati della simulazione.

# Toy Examples

- ▶ Data questa incredibile lista di applicazioni, come mai parleremo per settimane di **monete, dadi, urne e palline**?
- ▶ Perché ci permettono di rappresentare in maniera semplice e intuitiva moltissimi problemi complessi del mondo reale.
- ▶ Una semplice moneta ci permette di rappresentare innumerevoli esperimenti con due possibili risultati: es. successo o fallimento di una terapia (o di un esame).
- ▶ A volte un problema è così complesso che è difficile trovare un modello per rappresentarlo. Una soluzione alternativa sono le simulazioni al computer.
- ▶ In particolare, si simula la realizzazione di diversi esperimenti e si stimano le probabilità a partire dai risultati della simulazione.
- ▶ In questo corso conosceremo il linguaggio R.

# Programma del Corso

- ▶ Statistica descrittiva

# Programma del Corso

- ▶ Statistica descrittiva
- ▶ Calcolo delle probabilità

# Programma del Corso

- ▶ Statistica descrittiva
- ▶ Calcolo delle probabilità
- ▶ Statistica inferenziale



# Programma del Corso

- ▶ Statistica descrittiva
- ▶ Calcolo delle probabilità
- ▶ Statistica inferenziale
- ▶ Regressione

# Materiale

- ▶ Testi di statistica (consigliati):
- ▶ in generale ogni testo di statistica può andare bene
  - ▶ Sheldon M. Ross, “Probabilità e statistica per l'ingegneria e le scienze”
  - ▶ D.M. Levine, T.C. Krehbiel, M.L. Berenson, “Statistica”, Pearson.
  - ▶ Casella-Berger, “Statistical Inference”, Springer.
  - ▶ R. V. Hogg, “Probability and statistical inference”, Prentice Hall.
  - ▶ Lecture notes from MIT 18.05.

# Materiale

- ▶ Testi di statistica (consigliati):
- ▶ **in generale ogni testo di statistica può andare bene**
  - ▶ Sheldon M. Ross, “Probabilità e statistica per l'ingegneria e le scienze”
  - ▶ D.M. Levine, T.C. Krehbiel, M.L. Berenson, “Statistica”, Pearson.
  - ▶ Casella-Berger, “Statistical Inference”, Springer.
  - ▶ R. V. Hogg, “Probability and statistical inference”, Prentice Hall.
  - ▶ Lecture notes from MIT 18.05.
- ▶ Laboratorio:
  - ▶ <https://www.r-project.org/>
  - ▶ An introduction to R <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>
  - ▶ RStudio <https://rstudio.com/products/rstudio/>