

# Lezione 1

## Introduzione alla Statistica

***Stefania Capecchi***

Università degli Studi di Napoli Federico II

*stefania.capecchi@unina.it*

- 1 ***Statistica come scienza delle decisioni***
- 2 ***Dati, informazioni e conoscenza***
- 3 ***Tipologia e fasi delle analisi statistiche***
- 4 ***Un paradigma unitario per la Statistica***

# Cosa è la Statistica?

- ▶ La **Statistica** è una scienza relativamente giovane. Tale disciplina viene spesso confusa con **le statistiche**: dati, tabelle, grafici, indici, medie. In effetti, la statistica è funzionale a tutte le scienze e rappresenta oggi uno strumento essenziale per lo studio delle relazioni tra fenomeni.
- ▶ Soprattutto, la Statistica è uno strumento efficace per addivenire, minimizzando il margine di rischio, a prendere decisioni.
- ▶ Difatti, nelle situazioni nelle quali occorre assumere **decisioni in condizioni di incertezza**, la Statistica offre strumenti analitici, dai più semplici ai più avanzati, funzionali ai cosiddetti stakeholders (decisori, policy makers, parti interessate...), alla ricerca scientifica, alle applicazioni della tecnologia, al monitoraggio e alla valutazione di sistemi complessi, alla pianificazione economica, etc.
- ▶ In questi ambiti, la Statistica forma e fornisce strumenti per il *momento conoscitivo* grazie a metodologie rigorose e corroborate dalle osservazioni.

- ▶ Etimologia di **STATISTICA**. L'italiano Ghislini, intorno al 1580, introduce il termine Statistica come "*descrizione delle qualità che caratterizzano e degli elementi che compongono uno Stato*".
- ▶ In conseguenza di questa origine, si assiste spesso (soprattutto nei mezzi di informazioni) ad un uso distorto della Statistica (o delle "statistiche") a sostegno di tesi predefinite, cioè a fini ideologici.
- ▶ Invece, la Statistica è uno strumento formidabile per incrementare la conoscenza umana: in quanto tale, apporta un sostegno decisivo a una democrazia consapevole e alla cittadinanza attiva.

- Storicamente, la Statistica come disciplina è il risultato della fusione, non sempre indolore..., di due "anime":
- Statistica come *gestione delle informazioni* esistenti per finalità predefinite. Ciò deriva dalla sua vocazione pubblica a "**conoscere la realtà per decidere ed intervenire**"
  - Statistica come analisi probabilistica dei fenomeni. Ciò deriva dalla consapevolezza che la realtà è il risultato di un processo ove "**la componente casuale è ineliminabile**"

La sintesi fra queste due anime inizia a maturare all'inizio del '900 con l'**indeterminismo** nella conoscenza umana. I riferimenti, in termini di filosofia della scienza, si ritrovano a partire dalla teoria della relatività ristretta (1905) e dal principio d'indeterminazione di Heisenberg in fisica e in meccanica quantistica (1927).

# Una concezione unitaria della Statistica

- ▶ Con l'avvento di moderni computer, disponibili a costo contenuto ed estremamente efficienti, vi è la tentazione di considerare la Statistica come una procedura per elaborare masse di informazioni al di fuori di qualsiasi schema concettuale predefinito.
- ▶ La Statistica è una scienza che nasce per assumere decisioni sulla base di informazioni reali. Questo obiettivo configura la Statistica come una disciplina scientifica organizzata in più momenti:
  - Raccogliere ed esaminare le informazioni in vista delle decisioni da assumere
  - Interpretare i dati disponibili entro uno schema concettuale che non può che essere soggetto all'incertezza immanente in ogni realtà
  - Predisporre strumenti operativi per intervenire/agire in vista delle decisioni richieste

- ▶ L'analisi accurata dei dati e della struttura delle loro interrelazioni (nel tempo, nello spazio, nelle circostanze) costituiscono **valore aggiunto della conoscenza** perché inducono ed agevolano la formazione di ipotesi che la medesima Statistica consente di sottoporre a verifica scientifica.
- ▶ Tutte le informazioni umane sono parziali, limitate e provvisorie. Per questo, l'indagine statistica è un momento ineludibile della conoscenza.
- ▶ La *conoscenza statistica* procede per tappe successive esattamente come il pensiero umano che dai fatti e dall'esperienza deduce convinzioni e leggi di comportamento, di volta in volta soggetti a verifica sperimentale in un'ottica di apprendimento continuo.

# Dalle osservazioni ai problemi

- ▶ I fenomeni reali possono essere registrazioni univoche di **variabili osservabili** che si manifestano in modo esplicito: Reddito, Consumo, Produzione, Età, Resa per ettaro, Temperatura, Velocità, Voto ad un esame universitario, Numerosità di: turisti/visitatori..., contatti, etc.
- ▶ Talvolta, invece, i fenomeni di interesse possono essere il risultato di fenomeni reali non osservabili direttamente detti **variabili latenti**: Benessere, Gusto, Preferenza, Preoccupazione, Disagio, *Feeling*, Emarginazione, Pericolo, Rischio, Ansia, Stress, etc.
- ▶ In entrambi i casi, l'analisi statistica pone in essere strumenti e metodi per comparare, correlare, classificare e discriminare tali variabili in modo da pervenire a decisioni suffragate da elementi ben definiti.
- ▶ Solo in tal modo, le decisioni possono essere soggette a critica e verifica sperimentale.

- ▶ I dati e le informazioni non costituiscono verità assoluta ma solo realizzazione di una legge o di una ipotesi che quei dati devono mettere in discussione.
- *Senza modelli falsificabili dall'esperienza, i dati statistici sono numeri privi di senso.*
- ▶ Al di fuori di uno schema concettuale rigoroso, non esiste interesse scientifico né utilità pratica per qualunque analisi (non solo statistica....).



- Nell'analisi di qualsiasi problema, vale un principio generale:

*Gli obiettivi di una indagine specificano sia la natura delle informazioni da raccogliere che gli strumenti mediante i quali esaminare tali informazioni.*



- Predefinire **obiettivi** delimitati ed agevoli da investigare sul piano empirico.
- Individuare le **informazioni** necessarie per una puntuale conoscenza del fenomeno oggetto di studio. Predisporre le modalità di raccolta dei dati.
- Applicare **metodi** adeguati per il trattamento delle informazioni.
- Esaminare criticamente i **risultati** della metodologia statistica per saggiarne la coerenza, la validità, la generalizzabilità e verificarne l'utilità in rapporto agli obiettivi dell'indagine.

# Dai problemi reali ai problemi statistici

- Il passaggio logico

*problema reale* ⇒ *problema statistico*

richiede esperienza specifica nel settore da studiare e cultura statistica.

- Esso può avvenire in modo agevole solo mediante la interazione tra studiosi di discipline differenti.
- All'inizio di una indagine, è opportuno circoscrivere lo studio limitando gli obiettivi solo a quelli essenziali. Successivamente, essi possono essere arricchiti mediante nuovi particolari, man mano che l'analisi procede.

- 1. *Definizione degli obiettivi della ricerca*
- 2. *Rilevazione dei dati*
- 3. *Elaborazione metodologica*
- 4. *Presentazione ed interpretazione dei risultati*
- 5. *Utilizzazione dei risultati della ricerca*

# Esempio di analisi statistica

- 1 **Obiettivo della ricerca:** Conoscere il livello di preferenza rispetto ai servizi museali, analizzare le sue criticità in modo da poterli migliorare (CRM, customer relationship management)
- 2 **Rilevazione dei dati:** Assumere informazioni tramite questionari, interviste, etc. individuando una serie di fattori che determinano le preferenze o che ad esse possano essere collegati
- 3 **Elaborazione metodologica:** Svolgere elaborazioni ed applicare metodi statistici per derivare dai dati raccolti il massimo di informazioni utilizzabili ai fini dell'obiettivo prefissato
- 4 **Presentazione ed interpretazione dei risultati:** Utilizzare mezzi grafici e riassuntivi per presentare i risultati ed interpretare quelli di maggiore interesse operativo
- 5 **Utilizzazione dei risultati della ricerca:** Intervenire sulla struttura individuando strategie per ottimizzare la CRM

➤ *La Statistica è la scienza delle decisioni in condizioni di incertezza*

➤ Tale definizione deriva dallo sviluppo della *conoscenza umana* che si sviluppa mediante le seguenti tappe:

➤ **Osservare** i fatti nella loro struttura sperimentale

➤ **Comprendere** il loro meccanismo generatore

➤ **Agire** in vista di un obiettivo predefinito

# Il paradigma dell'analisi statistica

- **Osservare** i fatti nella loro struttura sperimentale:  
Esaminare con attenzione le osservazioni raccolte descrivendo e riassumendo le caratteristiche essenziali dei dati nonché eventuali legami tra fenomeni
- **Comprendere** il loro meccanismo generatore:  
Ipotizzare un meccanismo teorico che genera le osservazioni di interesse e verificare sul campione disponibile la coerenza tra dati ed ipotesi, assieme al rischio di possibili errori
- **Agire** in vista di un obiettivo predefinito:  
Specificare modelli e regole di comportamento per prevedere, classificare, valutare, discriminare il fenomeno di interesse in modo da intervenire e -ove possibile- migliorare la realtà ottimizzando i propri obiettivi

- 1 **Osservare** i fatti nella loro struttura sperimentale:  
ANALISI ESPLORATIVE sui dati disponibili
- 2 **Comprendere** il loro meccanismo generatore:  
*Interpretare in ambito probabilistico le osservazioni disponibili al fine di generalizzare i risultati*
  - **PROBABILITÀ**: dalla popolazione ai risultati
  - **INFERENZA**: dai risultati alla popolazione
- 3 **Agire** in vista di un obiettivo predefinito:  
MODELLI STATISTICI per intervenire e rendere operative le decisioni da assumere