



La statistica e le "statistiche"

Maria Piera Rogantin

Dipartimento di Matematica – Università di Genova

`rogantin@dima.unige.it`

Secondo voi la Statistica serve per

- a) gestire le ricerche di Google
- b) costruire i test psicologici
- c) produrre un farmaco
- d) costruire un telefono cellulare
- e) calcolare il tasso di disoccupazione
- f) fare le leggi di una regione
- g) analizzare i censimenti
- h) fare previsioni del tempo
- i) vincere al lotto
- j) farsi delle opinioni politiche
- k) commentare una partita di calcio

Quindi ...

Chi “fa statistiche” oggi?

- gli istituti **demoscopici** (risultati: giornali, TV, ...)
- gli istituti di ricerca orientati alla **salute pubblica**
- le aziende **farmaceutiche**
- le aziende che gestiscono i **motori di ricerca online**
- gli enti che effettuano ricerche **socio-economiche**
- il **settore industriale** (gestione linee di produzione, ricerche di mercato, controllo della qualità, pianificazione degli esperimenti, ...)
- gli enti che si occupano di **analisi del territorio** in ambito **ambientale, geologico, fisico, ...**
- le **banche** e le **assicurazioni**
- gli **enti pubblici** (ISTAT, Comuni, ...) che raccolgono e elaborano i dati su **tutta la popolazione**
- ...

- Le situazioni reali sono complesse
- Le quantità di dati sono sempre più grandi

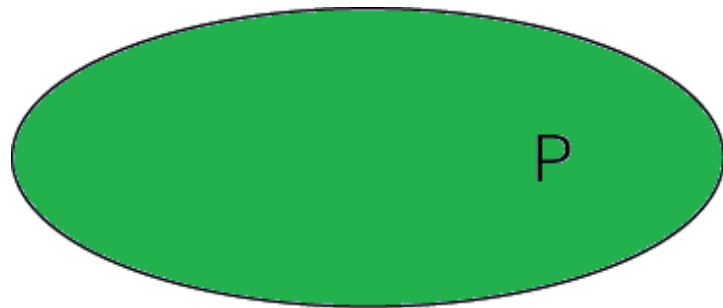
Le metodologie statistiche devono avvalersi di conoscenze matematiche approfondite e essere in grado di utilizzare strumenti informatici che permettano di trattare grandi quantità di dati.

Che cosa vuol dire fare statistica:

- **raccogliere** dati
- conservare i dati in modo opportuno
- controllare la "bontà" dei dati
- **sintetizzare** le informazioni tramite indici e rappresentazioni grafiche
- se i dati riguardano informazioni parziali del fenomeno da studiare, **generalizzare** le conoscenze e **fare previsioni** sul comportamento futuro

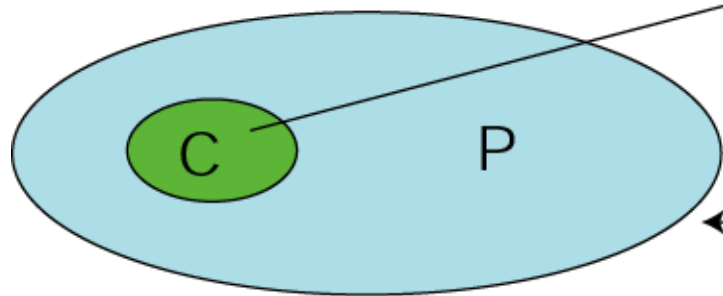
Principali metodologie statistiche "classiche"

STATISTICA DESCRITTIVA



analisi della popolazione
o del processo

STATISTICA INFERENZIALE



analisi del campione

calcolo delle probabilità

informazioni sulla
popolazione o sul processo

ESPERIMENTI SIMULATI

Un nuovo approccio: il DATA SCIENCE

Sia la statistica che il data science mirano a **estrarre conoscenza dai dati**

Statistica	data science
competenze	
Matematiche	informatiche + altro
dimensione dei dati	
relativamente piccolo	molto grande
raccolta dati	
spesso pianificata - qualità controllata	non pianificata
modo di costruire e verificare modelli	
il modello scelto si adatta ai dati? le assunzioni sono verificate? quantificazione dell'incertezza	confronto accuratezza predittiva vari modelli

ESEMPI DI USO DI METODOLOGIE STATISTICHE IN AMBITO:

- industriale
- economico
- demografico e demoscopico
- medico-farmacologico
- epidemiologico
- biotecnologico
- ambientale
- fisico-medico