

# Perché studiare statistica?

Perché viviamo in un mondo complesso, in cui si incrociano tante necessità e tanti problemi. Siamo tutti tenuti a prendere continue **decisioni**: cosa fare, come farlo, quando, perché. Possiamo scegliere in modo casuale, arbitrario, certo, è un nostro diritto. Ma quando le decisioni che prendiamo non coinvolgono solo noi ma tante altre persone, risorse, circostanze, beh allora è assurdo pensare di poter fare le cose *così, come capita*. Diventa necessario avere gli strumenti per compiere una **scelta informata**. E non è semplice, perché ogni scelta si compie in **condizioni di incertezza** riguardo al contesto e all'influire di tanti altri fattori. Scegliere razionalmente, dunque, richiede capacità di previsione.

Studiare statistica significa acquisire strumenti di studio e analisi della realtà.

Si sceglie bene solo quando si conoscono i fatti; lo statistico ha proprio il compito di saper capire e interpretare la realtà per **offrire soluzioni** a problemi concreti.

## E cosa c'entra l'informatica?

Le conoscenze informatiche sono fondamentali per uno statistico, perché aumentano enormemente la sua capacità di analisi, catalogazione, esposizione dei dati. Semplificando al massimo, possiamo dire che lo statistico raccoglie i dati, poi li organizza, li analizza, infine li presenta a chi ha commissionato lo studio. In ognuna di queste fasi l'informatica ha un ruolo di supporto importantissimo:

- nella **raccolta** dei dati, perché se i dati sono in formato digitale (come quando si fanno delle **interviste telefoniche** e gli operatori possono memorizzare le risposte degli utenti su un terminale) si può passare immediatamente alla loro organizzazione
- nell'**organizzazione** dei dati raccolti, che per essere compresi e analizzati devono essere prima ordinati, *puliti* e organizzati in funzione di quanto si sta studiando
- nell'**analisi**, perché uno studio statistico si basa su calcoli matematici che il

computer svolge con maggiore rapidità ed efficienza dell'essere umano

- nella **presentazione dei risultati**, perché tutto il lavoro svolto sarà utile solo se e quando il committente capirà i risultati. Lo statistico è capace di maneggiare numeri e concetti complessi, di chi commissiona il lavoro non sempre si può dire altrettanto. Ciò significa che l'informatica permette di creare grafici, schemi e tabelle capaci di raggiungere anche un ascoltatore poco preparato.

### **Qualche esempio?**

Che tipo di problemi può incontrare uno statistico? Di seguito, alcune situazioni-tipo in cui l'intervento di uno statistico può dimostrarsi essenziale alla risoluzione di un problema.

#### **Caso 1: l'amministratore pubblico**

Il sindaco Bianchi, del Comune di Belpaese in provincia di Letizia, è alle prese con il problema dell'inquinamento del centro storico. L'area, infatti, è aperta al traffico e nelle ore di punta viene spesso superato il limite di sostanze tossiche nell'aria previsto dalle norme UE. Esistono diverse possibili soluzioni: la chiusura completa al traffico, la chiusura parziale, la chiusura in determinati orari, la creazione di un sistema di incentivi all'utilizzo dei mezzi pubblici sono alcune di queste. Il sindaco sa che prendere la decisione sbagliata significa inimicarsi la cittadinanza, ma al contempo capisce che la situazione è diventata insostenibile. Dunque, si rivolge ad uno statistico perché faccia uno studio basato su interviste e sondaggio, per sondare l'umore dei suoi concittadini.

#### **Caso 2: la piccola impresa**

Il Cavalier Rossi è titolare della Rossi-Rossi spa, un'azienda del Nord-est che produce sellini per biciclette. Gli è stato proposto di convertire tutti gli impianti all'utilizzo di un nuovo materiale, il Fantamatix, una resina di nuovissima invenzione, molto resistente ed elastica, prodotta a Taiwan. Gli dicono che costa meno dei materiali che utilizza oggi, e che entro due anni tutti gli altri produttori

passeranno al Fantamatix. Il Cavalier Rossi è tentato di fare il salto, ma ha qualche perplessità: quale sarà il costo complessivo per l'aggiornamento degli impianti? Qual'è la resa del Fantamatix? I costi di produzione aumenteranno o diminuiranno? Dovrà assumere nuovo personale? Il Cavalier Rossi è confuso e capisce di aver bisogno di una consulenza che gli permetta di fare chiarezza sulla situazione.

### **Caso 3: la media impresa**

La tipografia

### **Caso 4: la grande impresa**

La Vedoeprevedo è una delle principali compagnie assicurative in Italia. Moltissimi statistici ci lavorano, e per un'ottima ragione: la capacità di prevedere gli eventi futuri è alla base del successo dell'azienda. Chi stipula un'assicurazione sta concretamente *spostando il rischio* da sé stesso alla società. Se non succede niente di male, l'assicurazione "vince" perché riceve del denaro senza dover prestare in cambio un servizio; se invece l'evento contro cui si è assicurati avviene (furto, malattia, incidente e così via), l'assicurazione è tenuta a risarcire il danno.

La Vedoeprevedo può rimanere sul mercato in un solo modo: investendo nelle ricerche statistiche necessarie a valutare le probabilità degli eventi. Perché un'assicurazione sulla vita ha un costo maggiore se a stipularla è un settantenne piuttosto che un ventenne? Perché le donne pagano generalmente meno degli uomini le loro assicurazioni auto RCA? Perché il sistema è fondato sull'analisi dei dati statistici e sulla costruzione di modelli capaci di prevedere gli andamenti futuri.

La Curotutto è un'azienda farmaceutica internazionale da sempre specializzata in prodotti che ...

### **Caso 5: l'Università**

Anche l'Università è un'azienda: un'azienda che produce cultura, conoscenza, competenze, e che funziona grazie al contributo di tante persone diverse.

Il prodotto dell'Università è per sua natura intangibile - non è semplice "quantificare" il valore di un laureato, di un corso di studi, parlare di profitti e di introiti. Eppure, anche l'Università ha la necessità di misurare ciò che fa per fissare nuovi obiettivi e poter offrire di anno in anno un servizio migliore ai suoi studenti, e alle aziende in cui i laureati andranno a lavorare. Come *misurare* il valore di quanto l'Università produce? Lo statistico programma, progetta ed esegue una serie di analisi statistiche mirate che, basandosi su parametri di Qualità, interviste, focus group e quant'altro permettono di creare dei documenti di valutazione della didattica e dell'organizzazione dell'Ente. Un esempio? [L'indagine sui fabbisogni di alta formazione in Sicilia](#) promossa dall'Università di Palermo.

### **Il Sigad: cos'è?**

Statistica e informatica per l'analisi dei dati: questo è il Sigad dell'Università di Palermo. Il corso ha una durata di tre anni e prevede tre percorsi didattici:

- statistico sociale e sanitario
- statistico economico
- statistico generale e per le scienze sperimentali

Gli obiettivi formativi specifici sono la predisposizione di indagini statistiche; l'applicazione di metodologia statistica di base in vari campi applicativi; la classificazione, elaborazione, rappresentazione e analisi dei dati; il coordinamento dei rilevatori nelle indagini statistiche finalizzate; la formazione, organizzazione, analisi e diffusione di dati statistici.

Sbocchi occupazionali: uffici di statistica della pubblica amministrazione; aziende private; società di consulenza; centri ed istituti di ricerca pubblici e privati

### **Come posso capire se il Sigad è il corso che fa per me?**

Alcuni di noi hanno un'idea chiara sul loro futuro professionale; molti altri no.

Mi piace?

Ho degli sbocchi professionali?

Riuscirò a portare a termine gli studi?

Cerchiamo di affrontare il problema con gli occhi di uno statistico.

Per scegliere la facoltà e il corso di laurea bisogna riuscire a proiettarsi nel **futuro**: chi sarò io fra 10 anni? In che città abiterò? Come saranno la mia famiglia, la mia casa, i miei interessi, il mio tempo libero? In altre parole, *che persona voglio diventare?* Certo, è difficile riuscire a immaginare tutto questo quando si è appena usciti dalle scuole superiori. E naturalmente non tutto dipende dalla nostra volontà, non tutto si può prevedere. Ciò che si può e si deve fare è impegnarsi per capire, prendere una decisione che tenga conto delle proprie **aspettative** sul futuro. Chi arriva alla laurea, in genere, si appassiona delle sue materie. Le ha studiate a lungo, le ha scelte, è più che naturale che desideri di poter **applicare** quanto ha appreso all'Università nella sua futura attività lavorativa. Chi studia medicina vuol fare il medico, chi studia filosofia il filosofo, chi lettere antiche l'archeologo o lo studioso... ognuno desidera poter mettere a frutto le conoscenze acquisite.

*Eppure*, per moltissimi laureati le cose vanno diversamente. Il mercato del lavoro non offre abbastanza sbocchi professionali attinenti e si finisce col fare un lavoro che c'entra poco, o niente, con ciò che si è studiato.

Chi studia statistica, invece, può essere *statisticamente certo* di un fatto: potrà applicare ciò che ha imparato all'Università, perché troverà lavoro nel suo campo in tempi brevissimi e non conoscerà l'assillo della disoccupazione. Mentre le professioni si evolvono, la società si muove e il lavoro cambia, gli studenti di statistica possono stare relativamente tranquilli: la domanda di professionisti della statistica è da anni e anni superiore all'offerta, sia nel nostro paese che all'estero. In altre parole, laurearsi in Statistica significa aumentare enormemente le possibilità di poter **scegliere** nel proprio futuro lavorativo. Nel contesto in cui viviamo, questo è un lusso che pochissimi possono permettersi.

**INSERIRE DATI** I laureati in statistica trovano lavoro nel giro di XXX mesi dalla laurea, ...



Lo statistico, questo sconosciuto

Può essere di un sesso qualsiasi e di un'età qualsiasi.

Con o senza occhiali, alto o basso, magro o grasso.

Eppure!

Lo statistico è riconoscibile a gran distanza e senza bisogno di binocolo perché...

lo statistico è una persona che guarda il mondo con un occhio tutto particolare.

Lo statistico è uno scienziato.

Non studia però la cellula, le stelle, o il moto perpetuo: lo statistico lavora sui **fenomeni collettivi che coinvolgono le persone**. L'obiettivo è **capire**. Il risultato sperato, **migliorare ciò che esiste**.

**Lo statistico è uno studioso a prova di paradosso!**

Come saprai, *paradoxon* significa "oltre l'opinione comune". Lo statistico, a ben vedere, lavora proprio in questo campo: quello di chi supera pregiudizi e verità presunte per analizzare i dati e trarne delle conclusioni utili.

Lo statistico è un artigiano.

Parte da una massa informe, i **dati grezzi**, e li trasforma in materiale compiuto. Che può essere rielaborato, raffigurato, analizzato ulteriormente.

Lo statistico gioca coi numeri.

Quando parliamo di fenomeni collettivi, dietro i numeri ci sono **persone**, facce,



La probabilità  $B(c)$  che la prima cifra di un valore scelto a caso sia  $c$  è data da:

$$B(c) = \log_{10} (1 + 1/c)$$

Cerchiamo di spiegarci meglio.

Immaginiamo di voler tenere sotto controllo le spese mensili e di raccogliere in una scatola tutti gli scontrini della spesa per un anno. Successivamente trascriveremo su un file al computer le cifre spese (scontrino del 12 marzo: 12,45 euro; scontrino del 15 marzo; 31, 20 euro, e così via). Potremo quindi ordinare questo elenco di numeri secondo la prima cifra: prima tutti i numeri che iniziano per 1, poi quelli che iniziano per 2, sino al 9. Questo elenco avrà con ogni probabilità una particolarità: **i numeri che iniziano per 1 saranno molto più frequenti di quelle che iniziano per 5, o per 6, mentre i numeri che iniziano per 9 saranno i meno frequenti di tutti.**

Lo stesso fenomeno si riscontra in moltissimi altri casi: la popolazione di una regione, divisa per paese; il numero di fogli stampati in una grande azienda ogni giorno. Anche i **numeri di Fibonacci** seguono la legge di Benford.

Ma **perché** questo avviene?

Dato che le cifre iniziali di un numero qualsiasi vanno da 1 a 9, sarebbe naturale aspettarsi che esse siano equiprobabili - quindi che dato un numero a caso, esso abbia la stessa probabilità di avere come prima cifra 1, 2, 3, 4, 5, 6, 7, 8 o 9.

Per molte serie casuali, invece, la distribuzione delle prime cifre è più simile a questa:

<b>Cifra iniziale</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
<b>Frequenza</b>	30,10 %	17,61 %	12,49 %	9,69 %	7,92 %	6,70 %	5,80 %	5,11 %	4,5 %

Ma quali sono le distribuzioni che rispettano la legge di Benford?

Possiamo dire innanzitutto quali sono quelle che *non* la rispettano: quelle che seguono la **curva di Gauss**.

Moltissimi eventi naturali hanno una distribuzione di probabilità che segue l'andamento a parabola rappresentato dalla curva di Gauss - ma naturalmente questi numeri non rispettano la legge di Benford.

Ci viene in aiuto il matematico Ted Hill, che [in uno studio](#) nel 1996 ha spiegato che queste serie numeriche si determinano ogni volta che abbiamo una *distribuzione di distribuzioni*, cioè una raccolta casuale di campioni casuali di dati. Infatti, Hill ha dimostrato che se si scelgono a caso un insieme di distribuzioni di probabilità, e poi all'interno di ciascuna di esse si sceglie a caso un numero, l'insieme di questi valori che potremmo chiamare "casualmente casuali" obbedirà... alla legge di Benford !

La legge di Benford, come preannunciato nel titolo, si può applicare ai bilanci aziendali per capire rapidamente se i numeri sono stati "truccati". Se numeri che iniziano per 9, ad esempio, appaiono troppo spesso, è facile pensare che i dati siano stati manipolati. Questo risvolto pratico della legge di Benford è stato messo in evidenza per la prima volta da Mark J. Nigrini, che [nel suo sito](#) offre risorse utili a chiunque voglia stanare delle frodi fiscali. La qual cosa mantiene una certa utilità in paesi, come gli Stati Uniti, in cui il falso in bilancio sia ancora considerato un **reato**, perseguibile per legge.

IN CANTIERE:

**Proporzionale o maggioritario (con gli occhi uno statistico)**

**La curva di Gauss**

**I numeri di Fibonacci**