

Calcolo delle probabilità e statistica I

Prof. Camillo Fucci
Facoltà di Informatica di Ferrara
Appunti di Reitano Francesco

Questi appunti non sono stati corretti per tanto non sono responsabile di eventuali errori e/o mancanze. Qualora vogliate informarmi per correggerli potete inviare una e-mail a reitanofrancesco@hotmail.com

Per avere la versione in PDF mandatemi una e-mail



STATISTICA

Il termine “**Statistica**” deriva dal latino medievale “*Status*” ovvero stato che vuole significare l’ordinamento politico, è stata chiamata statistica l’area scientifica che si occupava di fornire descrizioni relative ad uno stato o descrizioni comparative dei vari stati.

Tali descrizioni ebbero inizio già con Aristotele nel 350 a.c. e proseguirono nelle epoche storiche successive. Pur tuttavia la parola statistica apparve per la prima volta nella Enciclopedia Britannica nel 1797, essendo stata coniata ufficialmente sembra dal professore tedesco *Goffried Achenwall* (1719-1772) verso la metà del 1700, che la definì come la descrizione delle cose più notevoli di uno stato.

Il significato che essa ha mantenuto fino agli inizi del 1800 è stato dunque quello di descrizione o censimento con successiva analisi dei dati prodotti da tali operazioni.

In seguito il significato della statistica si è esteso ben oltre l’interpretazione originale, al giorno d’oggi si parla di due tipi di statistica:

1. **una statistica descrittiva o metodologica**
2. **ed una statistica matematica o inferenziale**

Per **statistica descrittiva** si intende **il complesso delle tecniche di cui si serve lo sperimentatore per raccogliere, elaborare e rappresentare gruppi di dati osservati che sono in genere in numero elevato e in forma disordinata.**

Lo scopo finale è quello di ricavare dalle osservazioni opportuni elementi di sintesi, questi possono essere: dei numeri (come ad *es. la media, la varianza, la mediana, la moda ecc.*) o informazioni rappresentate sotto forma di *tabelle, grafici, carte (geologiche, numeriche, ecc.), diagrammi, ecc.* In questa situazione si cerca però **soltanto di descrivere il fenomeno osservato** senza cercare di interpretare o generalizzare più di tanto riguardo alla totalità dei potenziali dati osservabili di cui quelli in esame sono soltanto una piccola parte.

Al giorno d’oggi naturalmente sono utili anche queste informazioni però quasi sempre vogliamo saperne di più e spendere meno, ciò non significa dovere censire tutto per avere le informazioni che ci servono, ma estrarre le stesse dalla conoscenza dei dati osservati che sono a nostra disposizione.

In termini statistici vogliamo trarre conclusioni su tutta la popolazione partendo dal campione osservato, ciò implica che molti dati saranno a noi ignoti, e che *dovremo cercare un metodo per averne una qualche conoscenza che non sia quella diretta sperimentale.*

Le informazioni di qui sopra rappresentano quindi un passo preliminare, però molto importante per la successiva ricerca di un modello interpretativo che è un aspetto proprio della statistica matematica o inferenziale.

Si cerca dunque di studiare gruppi di dati per decidere quale tipo di modello è quello che descrive al meglio l’intera popolazione.

I modelli considerati possono essere diversi fra loro in modo sostanziale o soltanto per alcuni parametri caratteristici della popolazione; è opportuno sottolineare fin da ora che **il calcolo delle probabilità diventa essenziale nella trattazione della statistica inferenziale.**

Per quanto riguarda la **probabilità** diamo innanzitutto alcuni brevi cenni storici:

i primi documenti che si conservano sono dovuti a *Girolamo Cardano* (1501-1576) (accanito giocatore) nel suo libro “*De Ludo Aleae*” scritto forse nel 1526 e pubblicato postumo nel 1663 e a *Galileo* in un suo scritto del 1620.

Nel libro di Cardano è inserito il calcolo delle probabilità dei punteggi che si ottengono come somma nel lancio di 3 dadi, il risultato però contiene qualche inesattezza e gli sviluppi non sono molto chiari, cosicché i giudizi sul valore dell’opera sono controversi, però è senza dubbio la sua importanza storica; anche nello scritto di Galileo viene affrontato lo stesso problema del lancio di 3 dadi, però con maggiore chiarezza, ma la **nascita del calcolo delle probabilità** viene abitualmente



attribuita all'interesse di *Blais Pascal* (1623-1662) e fissata nella corrispondenza fra lui e l'altro grande matematico francese *Pierre Fermat* (1601-1665) verso la metà del 1600.

L'interesse di Pascal fu stimolato da *Cavalier de Méré*, spirito vivace, matematico discreto e accanito giocatore d'azzardo che si rivolse spesso a lui per molti problemi riguardanti il gioco.

Il calcolo delle probabilità sorto quindi, come appena accennato, verso la metà del 1600 si può ritenere oggi come una disciplina matematica che trova applicazione nello studio degli esperimenti non deterministici o casuali.

Un esperimento non deterministico è un esperimento che può dar luogo ad un risultato fra quelli possibili non determinabile a priori in modo univoco.

Un esempio è rappresentato dal lancio di un dado il cui risultato è affetto da un certo grado di incertezza, potendo essere rappresentato da una qualunque delle 6 facce del dado stesso.

Questi brevi cenni hanno lo scopo di introdurre come si presenta originariamente il concetto di probabilità e di introdurre alle definizioni o interpretazioni che ora presenteremo.

Definizioni di probabilità

Definizione classica

La prima definizione di probabilità, detta perciò classica, si ritrova già in Pascal e definisce la probabilità di un evento come il rapporto fra il numero di casi favorevoli all'evento e i numeri di casi possibili, purché questi ultimi siano tutti ugualmente possibili.

Molti però vedono in questa definizione una *tautologia*, nel senso che bisogna sapere già prima che significato dare alla probabilità perché non si vede bene la differenza fra ugualmente possibili e ugualmente probabili.

Comunque i casi risulteranno ***ugualmente possibili*** quando si ha una situazione di **perfetta simmetria fisica** come avviene ad es. per le palline nell'urna del gioco del lotto, per le facce di un dado, per le carte di un mazzo ben mescolato, ecc.

Più che una definizione questa può essere considerata una **regola per misurare la probabilità di un evento in condizioni opportune**, quando cioè vi sia un numero finito di alternative possibili che possono essere considerate ad esempio per simmetria tutte ugualmente possibili.

Tutto questo però, sapendo già che cos'è la probabilità.

In conclusione la validità di questa definizione resta discutibile.

D'altra parte è facile constatare che essa risponde al requisito di essere operativa.

Il campo nel quale più direttamente si può applicare la definizione classica è quello dei giochi di dadi, carte, ecc. In essi infatti si possono individuare con precisione le diverse alternative e si può ragionevolmente supporre che esse siano tutte ugualmente possibili.

Vediamo un esempio:

Supponiamo di voler determinare la probabilità che lanciando una moneta non truccata esca testa.

In questo esempio c'è un solo caso favorevole, cioè testa, e due casi possibili cioè testa e croce.

Si arriva pertanto alla conclusione che la probabilità è uguale a $\frac{1}{2}$.

