



Quando si parla di statistica, mi è capitato numerose volte di dover affrontare i versi di Trilussa.

*“[...] Ma pè me la statistica curiosa
è dove c’entra la percentuale,
pè via che, lì, la media è sempre eguale
puro co’ la persona bisognosa.*

*Me spiego: da li conti che se fanno
seconno le statistiche d’adesso
risurta che te tocca un pollo all’anno:*

*e, se nun entra nelle spese tue,
t’entra ne la statistica lo stesso
perch’è c’è un antro che ne magna due.”*

È La poesia chiamata “La Statistica” che rende difficile il compito dello statista in quanto mette in dubbio l’affidabilità della materia stessa. Trilussa va direttamente al punto: è giusto dire che in media ogni cittadino mangia un pollo anche se in realtà, il cittadino ricco ne mangia due e quello povero nessuno? La risposta che avremmo dovuto dare a Trilussa sarebbe stata: “Perché non hai tenuto conto della **varianza?**”

La varianza fornisce una misura che descrive quanto i dati in nostro possesso si discostano dalla media. In altri termini, più sarà alto questo valore e più la variabilità dei valori sarà alta, dando quindi maggiore instabilità ai dati finali. Poniamo di avere una classe di 2 alunni. Affermare che la media dei voti della classe è 24, quando in realtà il primo alunno ha ottenuto 30 e il secondo 18, non ha la stessa attendibilità rispetto a una classe in cui entrambi gli studenti hanno ottenuto come punteggio 24. Inoltre la varianza a sua volta dipende dalla numerosità del campione. Possedere un campione formato da 2 unità non è altrettanto affidabile rispetto a un campione molto più ampio. Questo è solo uno dei tanti strumenti che si utilizzano nelle analisi statistiche per certificare la precisione delle affermazioni che ne derivano.

È interessante notare che nonostante la statistica sia una scienza, molto deve al suo interprete. Le statistiche infatti sono nient’altro che numeri e senza una persona che li **legga** e li **trasformi** in parola, l’analisi non è efficace. L’interprete o statista quindi ha una grande responsabilità. Con accanto la sua cassetta degli attrezzi, in cui negli anni ha collezionato scariati strumenti matematici e statistici, egli dovrà trovare risposte a domande concrete. Domande che possono

essere semplici o difficili, specifiche o ampie, e dovrà sapersi destreggiare nella foresta dei dati, caratterizzati da imprevisti, dati anomali e missing value.

“Questo prodotto posizionato su questo scaffale, vende di più?”, “Come posso ridurre i costi di produzione?”, “Come si può ridurre la spesa pubblica?”, e così via.

La scarsità dei dati non è più un problema, anzi l’abbondanza di dati (Big Data) sembra essere la nuova preoccupazione. Un’eccedenza di dati che, se non utilizzata, risulterebbe un vero spreco. Perché non trasformare questi dati in informazioni? E perché non usare queste informazioni per migliorare e accelerare le decisioni, evidenziare le tendenze, e fare previsioni?

Non bisogna dimenticare che il mero dato nient’altro è che una molecola d’acqua all’interno di una cascata. La dinamicità dell’acqua, l’altezza del dislivello, il grado di erosione del terreno, sono caratteristiche che solo un individuo con una visione olistica può catturare. Se l’interprete non si soffermerà solo sul dato e farà un passo indietro, non vedrà più un insieme disordinato di gocce ma potrà mirare un unico flusso d’acqua che segue un particolare andamento. Molte volte tuttavia avere una visione d’insieme non è semplice, in quanto la realtà è più complessa della teoria e in molti casi sarà compito dell’interprete **colmare le lacune** usando il buon senso per dare un significato fluido ai dati. Ed è in questa fase che si deve porre maggiore attenzione.

Trilussa (volutamente) ha preso un campione con varianza massima, ovvero con dati completamente opposti e con numerosità minima (solo 2 cittadini). Infatti egli voleva sollevare un tema già all’epoca importante ma che ai giorni nostri è ancora più rilevante. L’utilizzo della statistica ai fini di propaganda. Nella promozione sia politica sia economica, la statistica è uno strumento molto potente che può diventare un’arma decisiva per affermare la propria tesi sia in buona fede sia in malafede. Infatti nella statistica di propaganda alcuni analisti non danno **voce ai dati** ma cercano di plasmarli a loro piacimento per raggiungere un particolare scopo o affermazione. Questo fenomeno si manifesta soprattutto per dati con scarse informazioni dettagliate e bisogna saper diffidare da essi.

La statistica è un insieme di strumenti complessi, studiati e rappresenta un mondo affascinante che non prescinde dalla conoscenza della materia che si intende esplorare. Se utilizzato con onestà e correttezza sarà d’aiuto in ogni nostra piccola realtà, che essa sia domestica, lavorativa e perché no, anche nazionale.

Erika Agostinelli