Prevedere catastrofi: come si regolano gli scienziati

La Rivista, Numeri, La terra trema...

 Alessandro Giuliani | 26 Ottobre 2016

La prevenzione si deve basare sulla previsione ‘statistica’ di ciò che ‘potrebbe succedere’ non di cosa succederà di sicuro (compito impossibile), dimensionando le nostre azioni sull’analisi attenta dei dati di natura...

La scienza si occupa di regolarità e correlazioni, cioè di classi di eventi e non di eventi singoli. Ogni evento singolo, se preso nella sua unicità, è del tutto imprevedibile, anche quando ci sembra (apparentemente) del tutto ovvio e comune.

E’ del tutto ovvio e comune che Paola Ceccarelli e Enrico Balducci (nomi di fantasia) si siano sposati l’anno scorso a Roma, ma se analizziamo le contingenze successive che hanno reso possibile questo particolare evento ci rendiamo facilmente conto che la sua probabilità di occorrenza tende a zero. Infatti dovremmo calcolare la probabilità che quel particolare etrusco si sia invaghito, molte generazioni fa, di quella particolare sannita incontrata a una fiera e via via per generazioni, con una catena lunghissima di eventi, la cui composizione esatta (basta che salti un anello e tutto va a rotoli) è unica e quindi ‘impossibile’ dal punto di vista del calcolo delle probabilità. Se invece limitiamo le nostre pretese, sfuochiamo un pochino l’obiettivo, e ci chiediamo quale è la probabilità che due ragazzi di circa trenta anni di età si sposino a Roma, possiamo fare dei calcoli ragionevoli e arrivare a una probabilità statistica abbastanza precisa. Paola ed Enrico sono scomparsi nella loro unicità e sono stati dissolti nella classe dei ‘cittadini romani dai 25 ai 35 anni’: solo così la previsione ha senso.

L’efficacia della scienza è quindi inversamente proporzionale all’informazione rilevante persa nella procedura di raggruppamento di eventi singoli in classi, l’efficacia di una diagnosi dipende da quante siano le caratteristiche comuni ai pazienti della ‘malattia X’, laddove la sua incertezza è determinata dall’ influenza delle condizioni al contorno (e.g. particolari sotto-tipi di X, influenza di variabili demografiche ..).
La termodinamica consente previsioni molto più accurate della medicina (di fatto deterministiche) in quanto si basa su proprietà collettive di grandi popolazioni di enti

(molecole) a tutti gli effetti identici tra loro.
Queste semplici considerazioni sono alla base dell'arte dell'analisi statistica dei dati da cui traiamo, per i nostri scopi, la definizione di due semplici indici: la media e la deviazione standard.
Senza entrare in particolari, qui ci basti dire che la media corrisponde alla somma dei valori presi da una misura di interesse X (e.g. peso, altezza, temperatura..) su un insieme di numerosità N di osservazioni indipendenti, diviso per il valore N, in formule:

rn
Media (X) = SUMn (Xi) / N

Dove il pedice n di SUMn significa la somma di n termini, e il pedice i di Xi indica i differenti valori individuali. La media è quindi un indice di 'valore centrale', ci parla insomma dell'ordine di grandezza della variabile X.

La deviazione standard è invece un 'indice di dispersione': ci fornisce delle informazioni sull' entità delle deviazioni dalla media all'interno della popolazione. Insomma se una popolazione di altezza media 1.75 è formata da individui tutti più o meno alti uguale (bassa deviazione standard) oppure risulta da una popolazione molto eterogena che comprende anche soggetti molto alti e molto bassi. In formule:

rn
Dev. Stand.(X) = √ SUMn (Xi – Media(X))2/N

Dalla formula vediamo come la deviazione standard non sia altro che la media delle differenze (Xi – Media(X)) dei singoli individui dal valore centrale della loro popolazione, l'apparente complicazione dell'elevamento al quadrato dei singoli addendi e successiva radice della somma è necessaria altrimenti la somma di scarti positivi e negativi si annullerebbe.
A questo punto il lettore e la lettrice di BeneComune sanno tutto ciò che c'è da sapere per seguire il resto della trattazione di un caso paradigmatico di prevenzione basata sulla previsione che risale agli anni cinquanta del secolo scorso ma che è un capolavoro insuperato che permette di comprendere nel profondo la natura della prevenzione basata sulla previsione.
All'inizio degli anni 50 del secolo scorso, l'idrologo inglese Edwin Hurst, affrontò il problema di valutare quanto dovesse essere alta una diga da costruire sul fiume Nilo. La diga doveva naturalmente essere dimensionata non a partire da indici di portata 'media' o di 'variabilità del flusso' (informazioni ottenibili da indici di 'comportamento medio' come appunto la media e la deviazione standard) ma, affinchè l'acqua non esondasse, bisognava farsi una idea dell'entità di una 'osservazione unica' cioè la 'più grande inondazione mai

vista', un evento catastrofico che, proprio in quanto catastrofico, non è attingibile da una media di popolazione.

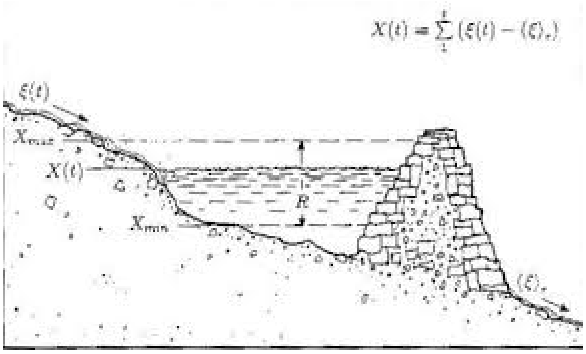
Ciò nonostante Hurst riesce a risolvere in maniera razionale *il problema del dimensionamento della diga.*

L'Egitto era detto dagli antichi 'Il dono del Nilo', il paese era infatti reso fertile dalle inondazioni del grande fiume che, circa due volte l'anno, trasformavano le sue aride sponde in campi fertilissimi grazie alla deposizione del 'limo', terra umida e grassa che permetteva abbondanti raccolti.

L'entità di ogni inondazione era proporzionale all'estensione del limo depositato, questa estensione era accuratamente misurata dagli impiegati governativi che, su tale base valutavano l'entità delle riserve alimentari (cfr. Genesi,41:17, dove l'interpretazione data da Giuseppe al sogno del Faraone delle sette vacche grasse seguite dalle sette vacche magre risulta coerente con le periodicità delle fluttuazioni delle inondazioni del Nilo poi verificate da Hurst).

L'importanza delle inondazioni del Nilo garantì ad Hurst la consultazione di una serie temporale continua di valori che andava dal 625 Avanti Cristo al 1429, cioè dall'epoca degli antichi Egizi alla fine dell'impero Bizantino.

Analizzando l'accurata documentazione storica *dei livelli di massima e di minima delle piene del fiume*, Hurst scoprì un interessante fenomeno: invece di una alternanza casuale di annate buone e annate cattive, sorprendentemente, rilevò che l'instaurarsi di una buona annata era seguita da altri anni con buon livello di acqua e, analogamente, l'inizio di scarsità di acqua persisteva per alcuni anni successivi (le vacche magre e grasse del sogno del Faraone, appunto). La figura di seguito, è un disegno originale di Hurst.



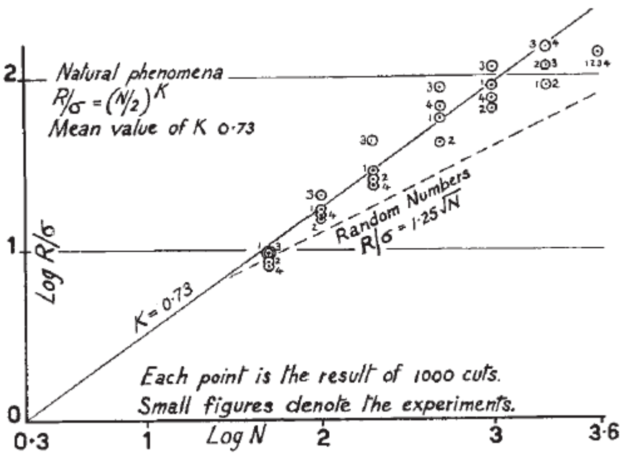
La grandezza R corrisponde all''evento eccezionale, pari al valore più grande di escursione tra minima (Xmin) e massima (Xmax) portata del fiume, mentre in alto a destra riconosciamo una nostra conoscenza (la deviazione standard o).

Hurst, invece di affrontare l'impossibile impresa di prevedere l'accadimento della piena eccezionale, prova a stimare la sua entità. Tutto sommato si trattava di costruire una diga che, se ben dimensionata, sarebbe stata in grado di contenere una piena eccezionale in qualsiasi momento fosse arrivata.

rn

Dalla statistica di base sappiamo che all'aumentare del campionamento, cioè all'aumentare del valore di N (numero di osservazioni), l'accuratezza della nostra conoscenza delle proprietà medie della distribuzione (e.g. media e deviazione standard) aumenta di conseguenza.

Questo però non è vero per i valori estremi che, a differenza dalle proprietà medie, non si calcolano sull'intera distribuzione, ma da poche (anche una sola) osservazioni eccezionali. In questo caso l'aumentare di N non provoca nessuna convergenza verso un valore unico ma, al contrario, il continuo "spostarsi dell'asticella" della misura: più osservazioni raccolgo, maggiore sarà la probabilità di osservare eventi eccezionali. In altre parole, se aumento il numero n delle osservazioni, dopo un po' (n abbastanza grande) il valore di media e deviazione standard della serie rimarranno sostanzialmente identici (è quella che popolarmente si chiama 'legge dei grandi numeri') mentre il valore del Range (R = Xmax – Xmin) tende a crescere per motivi puramente matematici (ogni osservazione 'più estrema' fa crescere R, che altrimenti resta identico, ma per definizione non può diminuire).



E' il processo alla base della figura qui riportata tratta dall'articolo originale di Hurst del 1957 che ha come ascissa la lunghezza della serie (in unità logaritmiche) e in ordinata il rapporto tra Range e Deviazione Standard (R/σ) sempre in unità logaritmiche.

Hurst si rende subito conto che all'aumentare di N il rapporto R/σ cresce con una legge molto precisa (i punti si distribuiscono su una retta) consentendo quindi di stimare per quanto tempo (proporzionale al valore di N) il valore R (determinato dall'impredicibile evento eccezionale) sarebbe rimasto comunque entro un certo valore, consentendo la prevenzione delle sue conseguenze (dimensionamento della diga).

rn

Hurst scopre una legge empirica esprimibile come: log(R/□)= Klog(N) corrispondente alla retta che passa per I circoletti della figura precedente. La cosa stupefacente della faccenda è che Hurst si toglie lo sfizio di ripetere l'analisi per le serie naturali più disparate (dalla durata dei battiti cardiaci agli anelli di accrescimento degli alberi) scoprendo che il valore (trovato empiricamente !!) del coefficiente K è sempre lo stesso, pari circa a 0.7 e molto diverso da ciò che si sarebbe atteso per puro effetto del caso (linea tratteggiata nella figura).
Questa è una delle regolarità più stupefacenti della natura, che probabilmente il curioso lettore avrà incontrato sotto il nome astruso di 'organizzazione frattale'.

Ma qui andiamo fuori dal seminato (anche se è la parte che interessa di più al vostro articolista), quello che conta è che la programmazione razionale dell'altezza della diga non si basa sul controllo assoluto del futuro (come certo pensiero dominante vorrebbe far credere) ma sulla gestione intelligente dell'incertezza che invece di lavorare contro di noi, ci fornisce una valida mano.

A questo punto il lettore ha in mano gli strumenti concettuali per:

1. **Spiegare educatamente a chi** (magari in buona fede) *pensa che qualsiasi guaio che ci succede sia opera dell'uomo* e va in cerca di un colpevole, ricordargli *che la natura esiste*, è al di fuori di noi anche se ne facciamo parte, è a noi impredicibile al dettaglio dell'evento singolo, ma ci offre delle preziose regolarità.

2. **Che la prevenzione si deve basare sulla previsione 'statistica'** di ciò che 'potrebbe succedere' non di cosa succederà di sicuro (compito impossibile), dimensionando le nostre azioni sull'analisi attenta dei dati di natura.