ANGELO LIZZIO
Socio corrispondente

I NUMERI NATURALI. MATEMATICA DEL CERTO E MATEMATICA DEL PROBABILE

1. I Numeri Naturali

Fin dai tempi più antichi l'uomo si è reso conto della necessità di contare, per esempio, gli animali del suo gregge, gli alberi del suo podere..., in generale i suoi averi; per controllare con esattezza i suoi beni, incideva delle tacche su un bastone oppure raccoglieva dei sassolini: basava il suo controllo sul concetto "tanti- quanti", cioè "tanti" erano i capi del suo bestiame "quante" le tacche sul bastone o i sassolini raccolti. Col passare del tempo scoprì anche che il contare era indispensabile ed essenziale, infatti la capacità di contare aumentava le possibilità di sopravvivenza: di fronte all'attacco del nemico era fondamentale capire se si era in inferiorità numerica per decidere se combattere o fuggire.

Oggi, spesso non ce ne rendiamo conto, ma ovunque e comunque ci si guardi attorno siamo circondati dai numeri; sono presenti quando siamo in treno o in aereo, quando guardiamo la TV o ascoltiamo un CD, quando siamo al telefono, quando andiamo in banca, o quando facciamo la spesa … costituiscono una parte essenziale della nostra vita e sono indispensabili nel nostro quotidiano. Oltre ad aiutarci a comunicare, ordinare e interpretare la realtà foggiano il nostro modo di concepire il mondo.

Dovettero passare centinaia e centinaia di anni prima che, passando dal concreto all'astratto, l'uomo parlasse di numero, concetto di fondamentale importanza per lo sviluppo dell'umanità. I numeri a cui ci si riferisce, chiaramente, sono i *numeri naturali*, cioè quelli che servono per contare.

Il numero naturale si presenta sotto due aspetti:

- *cardinale,* se risponde alla domanda: quanti? (esempio, quanti sono gli elementi di un dato insieme?).

- *ordinale*, se risponde alla domanda: quale? (esempio, qual è il posto di un fissato elemento in un dato insieme ordinato?).

In entrambi i casi risulta difficile darne *una definizione nominale*, cioè mediante concetti noti. Euclide, ad esempio, definisce il numero naturale così: "Unità è tutto ciò per cui ogni cosa è detta uno", "numero è una moltitudine di unità".

Solo recentemente è stata data una definizione rispondente a requisiti logici.

Una formulazione della teoria cardinale fu data da Giorgio Cantor (1845-1918). Egli definisce il numero per astrazione (ossia per mezzo di una relazione in cui interviene il concetto da definire) nel seguente modo: "Si dice che due insiemi (classi o aggregati) hanno lo stesso numero quando i loro elementi si possono porre in corrispondenza biunivoca". Chiaramente, in questo modo non si definisce che cos'è il "numero", ma la frase "avere lo stesso numero". In tale teoria è interessante la distinzione tra insiemi finiti e insiemi infiniti: un insieme è detto finito quando i suoi elementi non sono in corrispondenza biunivoca con quelli di una sua parte, in caso contrario è detto infinito.

Una teoria induttiva (o ordinale) fu introdotta da Giuseppe Peano (1858-1932) nei suoi celebri "*Arithmetices Principia*": egli, premesso che i numeri naturali costituiscono l'insieme N=(0,1,…), fonda tutta l'Aritmetica (studio delle proprietà elementari delle operazioni) sui tre concetti primitivi – *zero, numero naturale, successivo* – e sui seguenti cinque postulati (o proposizioni primitive, noti come *gli assiomi di Peano sui numeri naturali*):

1) *Zero è un numero.*
2) *Il successivo di un numero è un numero.*
3) *Se i successivi di due numeri sono eguali, i due numeri sono uguali.*
4) *Se un insieme contiene lo zero e se contenendo un numero, contiene anche il successivo allora l'insieme contiene tutti i numeri.*
5) *Il successivo di un numero non è mai lo zero.*

Il postulato n. 4 è noto come *Principio d'induzione*.

Peano e Alessandro Padoa (1868-1937) dimostrarono anche che i cinque sopradetti assiomi sono indipendenti. Successivamente Padoa osservò che il sistema dei concetti primitivi non è irriducibile rispetto ai cinque assiomi, potendosi lo zero definire mediante gli altri due e precisamente: *dicesi zero il numero che non è successivo di alcun numero*. Assumendo allora l'esistenza dello *zero* come proposizione primitiva in luogo della *1)* del Peano sopprime la *5)* divenuta superflua. Nel sistema assiomatico del Padoa si hanno dunque due concetti primitivi e quattro assiomi.

Sono poi definite le operazioni di *somma* e di *moltiplicazione* (a risultato unico) con le relative proprietà, le operazioni inverse (la sottrazione e la divisione) ed un *ordinamento* per cui

$$0 < 1 < 2 < 3 < \ldots\ldots\ldots\ldots$$

Tale ordinamento mi fa sovvenire alla mente la poesia "*NUMMERI"* del 1944, in dialetto romanesco, del noto poeta Carlo Alberto Salustri (1871-1950) meglio conosciuto con lo pseudonimo TRILUSSA (ottenuto anagrammando il suo cognome).

- Conterò poco, è vero:
- diceva l'Uno ar Zero -
ma tu che vali? Gnente: propio gnente.
Sia ne l'azzione come ner pensiero
rimani un coso voto e inconcrudente.
Io, invece, se me metto a capofila
de cinque zeri tale e quale a te,
lo sai quanto divento? Centomila.
È questione de nummeri. A un dipresso
è quello che succede ar dittatore
che cresce de potenza e de valore
più so' li zeri che je vanno appresso.

Per Pitagora (569-475 a.C.) *"I numeri governano l'universo"* mentre per lo scrittore statunitense di fantascienza moderna Robert Anson Heinlein (1907-1988) "*se qualcosa non può essere espresso in numeri non è scienza: è opinione",* come asserisce nel suo romanzo "*Lazarus Long l'Immortale"* (1973) .

In effetti tutti coloro che si occupano di scienze formalizzano le loro varie teorie scientifiche proprio attraverso il linguaggio dei numeri. Ad esempio, la *crittografia,* che studia i metodi che consentono la trasmissione sicura dell'informazione, poggia le sue basi fondamentali sui numeri *primi* (ossia quei numeri maggiori di 1 divisibili soltanto per se stessi e per 1).

Si può fare però scienza, in particolare matematica, senza ricorrere esplicitamente alla teoria dei numeri.

Volendo approfondire questo aspetto, ne mostriamo alcune applicazioni un po' insolite. E'da osservare che tali applicazioni sono più frequenti di quanto si possa immaginare e permettono di soddisfare, almeno in parte, alcuni dei bisogni reali dell'uomo, come ad esempio quello di aver la certezza di "*ragionar bene*" e quello di controllare ciò che non è certo, ma soltanto *possibile o probabile*.

Questo fatto ha condotto alla nascita e allo sviluppo di nuove branche della matematica, note col nome *di calcolo delle proposizioni, calcolo delle probabilità, statistica,* e costituiscono gli strumenti di lavoro veramente preziosi e indispensabili per chiunque si occupi di fisica, biologia, medicina, psicologia, sociologia, scienze politiche, economia, ingegneria.

Spesso, senza neanche accorgercene, facciamo uso di tali tecniche e di tali metodi.

2. Il Calcolo delle Proposizioni

Quando esprimiamo il nostro pensiero lo facciamo mediante un discorso scritto o orale. La parte di un discorso scritto che va da un punto a un punto si suole chiamare *periodo o frase*. Un periodo è composto da una o più *proposizioni* in corrispondenza dei verbi contenuti in esso. Ogni discorso, semplice o complesso che sia, si compone quindi di un insieme di proposizioni, così come una molecola è composta di atomi. Fra le proposizioni solo alcune si adoperano nei ragionamenti e quindi solo alcune interessano la logica, intesa come scienza del ragionamento. Le proposizioni che interessano la logica si chiamano *proposizioni logiche*: sono quelle proposizioni grammaticali (soggetto, verbo, complemento) del tipo "al soggetto x spetta il complemento y" soddisfacenti le seguenti due condizioni: a) esprimono un giudizio non soggettivo, affermano o negano qualcosa; b) ha senso porsi la domanda se è vera o se è falsa e poter rispondere univocamente senza alcuna riserva.

Una proposizione logica, dunque, può essere considerata come una grandezza che può assumere uno dei due valori "vero (V)" e "falso (F)".

Pur non essendo prive di senso non sono proposizioni logiche quelle che esprimono invocazioni o preghiere (Dacci oggi in nostro pane quotidiano, Signore ascoltaci), sentimenti (Ti voglio bene), comandi (Vai a destra! , Esci fuori!), domande (Che ora è?), previsioni (Domani nevicherà), la poesia (l'albero a cui tendevi la pargoletta mano), le definizioni. Tali proposizioni infatti non esprimono alcun giudizio, non affermano o non negano niente, e per esse non ha senso dire: è vera, o dire: è falsa.

I teoremi e i postulati (sono teoremi che non si dimostrano, si suppongono veri a priori) invece esprimendo delle proprietà sono proposizioni logiche.

Le proposizioni logiche si possono combinare fra loro mediante delle parole che si chiamano *connettivi logici o connettivi proposizionali* dando luogo a proposizioni composte il cui valore di verità (vero o falso) è determinato dal valore di verità delle preposizioni componenti. Sono connettivi logici: *e, o, non, se…allora, se e solo se*.

a) Il connettivo logico *"e"* si chiama *congiunzione* e si indica con $\cap$. Ad ogni coppia di proposizioni *A, B* associa la proposizione $A \cap B$ vera se e solo se *A* e *B* sono entrambe vere, falsa negli altri casi. Dal punto di vista semantico, il comportamento di tale connettivo si può rappresentare mediante la seguente tavola di verità.

A	B	$A \cap B$
V	V	V
V	F	F
F	V	F
F	F	F

A volte, nel linguaggio comune, le congiunzioni *"ma"* e "*mentre*" sono equivalenti alla congiunzione "*e*" come ad esempio "Carlo è a letto, ma non dorme" (Carlo è a letto e non dorme), "Luigi guarda la televisione, mentre mangia" (Luigi guarda la televisione e mangia).

b) La particella avversativa *"o"*, nella lingua italiana, viene usata con vari significati: oltre che come congiunzione può essere usata in certe esclamazioni come "o povero me!". Come connettivo può avere due significati: quello del latino *vel* se significa che delle due eventualità se ne verifica "*o l'una o l'altra o tutte e due*"; quello del latino *aut* se significa *"o l'una o l'altra ma non tutte e due"*. In matematica ci riferiamo alla prima interpretazione e la particella *o* si chiama *disgiunzione (o inclusione)* e si indica con $\cup$. Ad ogni coppia di proposizioni *A,B* associa la proposizione $A \cup B$ vera se almeno una delle due proposizioni *A* e *B* è vera, falsa se sono entrambe false. La tavola di verità di $\cup$ è:

A	B	$A \cup B$
V	V	V
V	F	V
F	V	V
F	F	F

Nella seconda interpretazione la particella o si chiama *esclusione* (XOR) e si indica con $\oplus$. La proposizione "$A \oplus B$" è falsa se entrambe le due proposizioni hanno lo stesso valore di verità, vera se una soltanto delle due proposizione ha tale valore di verità: la sua tavola di verità è quindi la seguente.

A	B	$A \oplus B$
V	V	F
V	F	V
F	V	V
F	F	F

c) Il connettivo logico "*non*" si chiama *negazione e* si indica col simbolo ¬. Ad ogni proposizione *A* associa la proposizione ¬*A* che è vera se *A* è falsa, falsa se *A* è vera. La tabella di verità di ¬ è:

A	¬*A*
V	F
F	V

E' da osservare che nell'eseguire la negazione di proposizioni del tipo "ogni…", "qualche…" per non commettere errori, basta far precedere la proposizione che si vuol negare da "non è vero che" e quindi dire "non è vero che ogni...", "non è vero che qualche..."; inoltre applicando due volte di seguito la negazione "*non*" si ottiene una affermazione equivalente a quella di partenza.

d) Il connettivo "*se…allora*" o "*implica*" si chiama *implicazione materiale* e si indica col simbolo →. Ad ogni coppia di proposizioni *A*, *B* associa la proposizione *A*→ *B, se A allora B,* che si conviene abbia lo stesso significato di "*nonA o B"*, falsa nel caso in cui *A* sia vera e *B* falsa, vera negli altri casi. La tabella di verità di → è:

A	*B*	*A*→ *B*
V	V	V
V	F	F
F	V	V
F	F	V

e) Il connettivo *"se e solo se"* o *"coimplica*" si chiama doppia *implicazione* e si indica col simbolo ↔. Ad ogni coppia di proposizioni *A, B* associa la proposizione *A* ↔ *B, "A se e solo se B*", che si conviene abbia lo stesso significato di *se A allora B e se B allora A*, vera se *A* e *B* hanno lo stesso valore di verità, falsa negli altri casi. La tabella di verità di ↔ è:

A	*B*	*A* ↔*B*
V	V	V
V	F	F
F	V	F
F	F	V

Si dicono *formule proposizionali* le proposizioni che si ottengono combinando due o più proposizioni con i connettivi logici. Si può calcolare il valore di verità di una formula proposizionale, conoscendo il valore di verità delle proposizioni che la compongono. In generale, possiamo associare ad ogni formula proposizionale una tavola di verità che dica qual è il valore di verità della formula data, in corrispondenza di tutte le possibili assegnazioni dei valori di verità, V e F, alle proposizioni che la compongono.

Due formule proposizionali *P* e *Q* sono *equiveridiche o logicamente equivalenti* se entrambe assumono gli stessi valori di verità in corrispondenza dei valori di verità delle proposizioni che le compongono.

Una formula proposizionale *P* che risulta vera qualunque sia il valore di verità delle proposizioni che la compongono si dice *tautologia*; se invece qualunque sia il valore di verità delle proposizioni che la compongono è falsa la *P* si dice *contraddizione*.

La tautologia *"A o nonA"* è il cosiddetto "*principio del terzo escluso*" (*Tertium non datur*) ed esprime il fatto che per qualunque proposizione non esiste altra alternativa dell'essere vera o dell'essere falsa. La contraddizione *"A e nonA"* è il cosiddetto "*principio di contraddizione*", cioè è contradditorio affermare che una proposizione sia simultaneamente vera e falsa. Questi due principi, che risalgono ai filosofi greci, sono stati chiamati "*principi fondamentali*" della logica classica.

In epoca moderna (sec XVIII) si è aggiunto ai due suddetti principi il cosiddetto "*principio d'identità*". Di solito nella logica delle proposizioni è sottointeso e si potrebbe enunciare affermando che una data proposizione ha un valore di verità, suo proprio, che è *identico a se stesso*, e ciò significa che tale valore non può cambiare se la proposizione viene pensata oggi o domani, in un luogo o in un altro, da una persona o da un'altra…

Facciamo un passo in altra direzione. Esaminiamo le seguenti proposizioni e chiediamoci per ciascuna di esse se è vera o falsa.

1) Il 25 dicembre è Natale.
2) La Sicilia è un'isola.
3) Il numero 8 è divisibile per 3.
4) Catania è una città del Lazio.
5) L'altezza media degli Italiani è m. 1,71.
6) Due ragazzi su cinque praticano il nuoto.
7) Se lancio in alto una moneta verrà testa.
8) Sabato vincerò all'enalotto.
9) Esci fuori.
10) Sempre mi fu caro questo ermo colle.

La 1) e la 2) sono vere. La 3) e la 4) sono false. La 5) e la 6) ci lasciano un po' perplessi; saranno vere, ma come si fa a verificarlo? La 7) e la 8) ci lasciano ancora più perplessi, ma nessuno metterà la mano sul fuoco per dire se sono vere o false. Per la 9) e la 10) l'imbarazzo è ancora più grande; in questo caso anzi non ha alcun senso chiedersi se esse sono vere o false.

Le proposizioni 1- 4 sono proposizioni logiche e la matematica che si occupa di questo tipo di affermazioni è una *matematica del certo*; quella che si occupa delle affermazioni come la 5) e la 6) presuppone ricerca e studio di dati e di campioni si chiama *statistica*, mentre la matematica che si occupa di affermazioni del tipo 7) e 8) è detta *matematica del probabile* (probabilità). Le proposizioni 9) e 10) non sono oggetto di studio della matematica.

La statistica e la probabilità sono tra loro intrecciate, i loro concetti spesso sono in correlazione e in molte esperienze sono presenti insieme; sono due aspetti diversi di un unico modo di analizzare la realtà, non legata ad un fatto ma ad una serie di fatti. Un modo convergente di operare per cui viene analizzato e classificato il presente con dati quantitativi in modo da prevedere un futuro possibile.

3. La Statistica

La *Statistica* ha una lunga storia; il suo nome deriva da "*stato*" perché inizialmente venne usata dai governanti come strumento per gestire i problemi fiscali, militari, demografici ed amministrativi.

Per TRILUSSA la statistica

……………………………… È 'na cosa
che serve pe fà un conto in generale
de la gente che nasce, che sta male,
che more, che va in carcere e che spósa.
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.

Negli ultimi secoli è stata organizzata con maggior rigore, attraverso l'applicazione sistematica del calcolo delle probabilità.

Essa ha un duplice aspetto: quello descrittivo con cui si possono dare interpretazioni di cose e fatti accaduti (*statistica descrittiva*) e quello inferenziale in cui giocano un ruolo determinante le metodologie probabilistiche con le quali si può tentare di "prevedere ciò che potrebbe accadere" *(statistica inferenziale*).

La statistica descrittiva permette di studiare, mediante particolari metodologie (raccolta, elaborazione e rappresentazioni di dati) e calcoli numerici di cui si serve uno sperimentatore, fenomeni collettivi – fisici, economici, scientifici, sociali,… – a partire da misurazioni fatte sui singoli elementi (*unità statistiche*) relative alle *caratteristiche* di una fissata "*popolazione*". Il termine "*popolazione*" designa un insieme (in generale piuttosto numeroso) di "*individui*": ma può trattarsi indifferentemente di una popolazione umana, di una colonia di batteri, delle molecole di un gas, degli esiti di un sperimento ripetuto tante volte,… Ovviamente, gli "*individui*" sono di volta in volta le singole persone, o i singoli batteri, o le singole molecole, o i singoli esiti dell'esperimento,… Spesso i dati relativi ad una determinata caratteristica della popolazione sono di tipo quantitativo, ossia *numerici*. Nel caso di una popolazione umana può trattarsi, per esempio, delle misure delle altezze, o dei pesi o delle età degli individui della popolazione. Ma non è esclusa l'eventualità che si abbia a che fare con dati di tipo qualitativo, ossia *non numerici,* quali ad esempio il tipo del gruppo sanguigno, oppure il colore degli occhi,…

Essendo, a volte, parecchio difficoltoso esaminare le varie unità statistiche relative agli individui della popolazione in esame, si è costretti a lavorare su un suo sottoinsieme detto *campione*, il quale deve essere "*rappresentativo*" e "*sufficientemente numeroso*", se si vuole che i risultati siano accettabili e significativi per l'intera popolazione.

In una indagine statistica, allo scopo di presentare in forma chiara e sintetica le principali informazioni presenti nei dati occorre riassumere mediante opportune *misure o indici numerici* le rilevazioni effettuate. Le misure che più frequentemente vengono impiegate sono di due tipi:

1) Misure di posizione o di tendenza centrale o *valor medio*.

2) Misure di dispersione o di variazione.

Gli indici di posizione o valor medi sono quei valori numerici che, in relazione all'indagine considerata, meglio si prestano a rappresentare tutti i dati da analizzare. Gli indici di dispersione permettono di conoscere se i dati da analizzare sono addensati o dispersi attorno al valor medio.

I valori medi di un insieme di dati più comunemente usati sono la *media aritmetica*, la *moda*, la *mediana*.

Dati *n* numeri, misure della grandezza presa in esame, $x_1, x_2, ...+ x_n$ la loro *media aritmetica* è il numero *Ma* dato dalla formula

$$Ma = \frac{(x_1 + x_2 + \cdots + x_n)}{n} \quad ,$$

usando il simbolo di sommatoria, la stessa formula, più sinteticamente, si scrive

$$Ma = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Se le misure non sono tutte distinte, cioè x_1 si presenta f_1 volte, x_2 si presenta f_2 volte, ..., x_i si presenta f_i volte, cioè i valori della grandezza sono distribuiti in un numero finito di classi, la classe alla quale appartiene il maggior numero di misure si chiama *moda* (non è escluso che esistano due o più classi ugualmente numerose).

Se gli *n* numeri, si *ordinano per valore crescente* si chiama *mediana o valore mediano* l'elemento *Me* che in questa successione occupa il posto centrale: più precisamente se *n* è dispari, l'indice che individua il posto centrale è $i = \frac{n+1}{2}$, mentre se *n* è pari non esiste un elemento di posto centrale e in tal caso si considerano i due elementi più prossimi al posto centrale, individuati dagli indici $i_1 = \frac{n}{2}$ ed $i_2 = \frac{n}{2}+1$ e se ne fa la semisomma. Sinteticamente la mediana di un insieme di *n* dati è espressa dalla formula:

$$Me = \begin{cases} x_{\frac{n+1}{2}} & \textit{se } n \textit{ è dispari} \\ \frac{1}{2}\left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}\right) & \textit{se } n \textit{ è pari} \end{cases}$$

Da queste definizioni si deducono facilmente le seguenti considerazioni:

- la media aritmetica tiene conto di tutti i dati, però non può essere rappresentativa dell'insieme di dati considerato se ci sono pochi valori molto alti o molto piccoli rispetto agli altri;

- la moda è invece il valore più *tipico* di tutto l'insieme di dati perché è quello che si presenta più spesso, ma non tiene conto degli altri dati; tuttavia qualche volta solo la moda è il valor medio più adatto;

- la mediana ha il vantaggio di non essere influenzata dai valori troppo piccoli o troppo grandi e ha il pregio di essere calcolata facilmente; in generale la mediana è utile quando l'insieme dei dati è molto numeroso e gli elementi intermedi sono molto vicini l'uno all'altro.

Le suddette considerazioni permettono di comprendere quanto asserito dall'umorista americano Evan Esar (1899-1995) nel suo aforisma sulla statistica a cui si deve la sua notorietà. "*La statistica è l'unica scienza che permette a esperti diversi, usando gli stessi numeri, di trarne conclusioni diverse*".

Il seguente esempio chiarisce bene il significato di tale aforisma.

La tabella delle paghe dei 50 dipendenti di una fabbrica è la seguente:

n. dipendenti	*1*	*1*	*2*	*3*	*18*	*22*	*3*
paga annua in euro	*124.000*	*48.000*	*26.000*	*14.000*	*9.000*	*7.000*	*6.000*

Il proprietario della fabbrica, che ha la retribuzione annua di 124.000 euro, dichiara che "la paga media annua della fabbrica è di 12.000 euro"; il sindacato che rappresenta i lavoratori della fabbrica afferma che "il salario medio annuo è di 7.000 euro"; l'agenzia delle entrate sostiene che " il salario medio annuo è di 8.000 euro".

Le diverse conclusioni dipendono dal tipo di valor medio che si considera per lo studio del fenomeno esaminato: il proprietario si riferisce alla media aritmetica (come se il denaro fosse distribuito in modo che ognuno riceva la stessa somma, nel nostro caso ogni dipendente 12.000 euro), il sindacato si riferisce alla moda (ossia la retribuzione più comune, nel caso in esame è di 7.000 euro) e l'agenzia delle entrate si riferisce alla mediana (la quale indica che metà degli stipendiati percepiscono più di 8.000 euro e l'altra metà meno di 8.000 euro).

4. Il Calcolo delle Probabilità

Storicamente il calcolo delle probabilità si è sviluppato fra il quindicesimo e il sedicesimo secolo, prevalentemente sulla base di studi e considerazioni teoriche riguardanti situazioni e problemi connessi ai giochi d'azzardo. Il primo libro sul gioco d'azzardo, *Liber de ludo aleae,* è stato scritto agli inizi del 1500 da Gerolamo Cardano (1501-1576?, matematico, fisico, medico, e astrologo italiano). Studi archeologici hanno mostrato addirittura che i primi giochi d'azzardo con dadi e oggetti ricavati da ossa di animali risalgono al 3600 a.C.. Lo sviluppo storico del calcolo delle probabilità è dovuto a grandi scienziati quali Galileo Galilei (1564-1642), Jacob Bernoulli (1654-1705), Blaise Pascal (1623-1662), Pierre de Fermat (1601-1665), Pierre Simon Laplace (1749-1827). Nel secolo scorso la teoria delle probabilità si è sviluppata in molte direzioni grazie al lavoro di famosi matematici fra i quali Andrey Nicolaevich Kolmogorov (1903-1987), autore dell'impostazione assiomatica della probabilità, e Bruno de Finetti (1906-1985), e poi Leonard Jimmie Savage (1917-1971), che hanno elaborato una concezione soggettiva della probabilità, che è la più generale, secondo cui essa è il *grado di fiducia* che una persona ha nel verificarsi dell'evento.

L'analisi di situazioni e problemi reali comporta l'esame di fatti e aspetti per i quali non si è in grado di predire in modo univoco l'esito di un esperimento o il verificarsi o meno di una circostanza; la mancanza di informazioni porta a dover ragionare in termini *probabilistici*, anziché *deterministic*i. I fatti incerti sono formalizzati, in modo non ambiguo, mediante proposizioni logiche, indicate col termine di *eventi,* che possono assumere il valore *Vero* o il valore *Falso*. Un evento è *certo* o *impossibile* a seconda che la proposizione logica che lo definisce è certamente Vera o Falsa.

Il calcolo delle probabilità vuol fornire una misura quantitativa dell'incertezza delle situazioni aleatorie: in altre parole vuole associare, mediante opportune regole ad ogni evento *E (casuale)* un numero reale P(*E)* che per convenzione è un numero reale compreso tra 0 e 1. Si attribuisce il valore $P(E) = 1$ ad un evento certo mentre ad un evento impossibile si attribuisce il valore $P(E) = 0$. Spesso il numero $P(E)$ viene espresso in forma percentuale: $0\% \leq P(E) \leq 100\%$.

Ad un livello più elementare, il calcolo delle probabilità cerca di misurare la plausibilità del verificarsi di eventi che possono avere un numero *finito* di esiti diversi e che per ogni prova si verifichi uno e uno solo di tali esiti, detti *eventi elementari*; l'insieme di tutti gli eventi elementari è detto spazio degli eventi. Anche un sottoinsieme dello spazio degli eventi è un evento; eventi elementari si dicono *equiprobabili* quando non sussistono motivi tali da indurre a privilegiare qualcuno di essi rispetto ai rimanenti.

Nel suo trattato "*Théorie Analytique des Probabilités*" del 1812 Laplace afferma che *se lo spazio degli eventi è costituito da n eventi elementari equiprobabili (casi possibili) e se un evento E è costituto da k eventi elementari (casi favorevoli o successi) la probabilità dell'evento E è data dal rapporto fra il numero k dei casi favorevoli e il numero n dei casi possibili*. Secondo questa definizione, detta *classica*, il calcolo elementare della probabilità si riduce al *calcolo combinatorio,* ramo importante (e spesso molto difficile) della matematica, il cui compito è quello di "contare senza fare conti". Il termine "combinatorio" in uso sin dal diciassettesimo secolo è stato introdotto dal filosofo e scienziato tedesco G. W. Leibniz in un suo classico trattato.

E' da notare che la suddetta definizione di probabilità ha senso solo nel caso in cui lo spazio degli eventi sia finito (di conseguenza non è applicabile nel caso continuo) e, limitazione ancora più forte, gli eventi siano equiprobabili (si noti che si parla di eventi equiprobabili senza ancora aver introdotto la nozione di probabilità). Queste limitazioni poste per la validità della definizione classica, la rendono poco adatta per la maggior parte delle applicazioni. Di conseguenza, per superare tali limiti fu proposto da Richard von Mises (1883-1953) di approcciarsi al concetto di probabilità da un diverso punto di vista: quello *frequentista*.

Sia E un evento aleatorio relativo ad un esperimento *ripetibile nelle medesime condizioni* quante volte si vuole. Se N è il numero complessivo delle volte che si effettua un esperimento E ed $F_{ass}(N)$ è il numero delle prove nelle quali esso si è verificato (frequenza assoluta) il rapporto $F_{rel}(N) = F_{ass}(N)/N$ dicesi frequenza relativa dei successi. *Secondo l'impostazione frequentista, dicesi probabilità dell'evento E il limite delle frequenze relative, al tendere all'infinito del numero delle prove cioè* $P(E) = \lim F_{rel}(N)$ *per N che tende a più infinito*.

L'uso di denotare con lo stesso simbolo $P(E)$ la probabilità sia dal punto di vista classico, sia da quello frequentista è giustificato dal teorema che Jakob Bernoulli ha dimostrato nella sua opera "*Ars conjectandi"* (pubblicata postuma nel 1713) noto come *legge dei grandi numeri*, secondo il quale se ad un evento E sono applicabili entrambe le definizioni di probabilità, esse danno luogo al medesimo valore di $P(E)$.

Il termine *limite* ha un significato meramente teorico, in pratica ci si accontenterà di un numero finito "*abbastanza grande*" N di prove e si identificherà convenzionalmente $P(E)$ con $F_{rel}(N)$, senza passare al limite per *N* tendente all'infinito. E' da notare inoltre che per poter applicare la suddetta definizione frequentista, sia pure ammettendo un certo margine di errore, è necessario ripetere un gran numero di volte l'esperimento al fine di procurarsi i dati necessari per il calcolo della frequenza relativa; si aggiunga il fatto che tutto ciò può essere scomodo e dispendioso oltre alla complicazione che non tutti gli esperimenti sono ripetibili nelle stesse condizioni un gran numero di volte.

Per ovviare a queste difficoltà De Finetti e Savage hanno proposto la seguente definizione di probabilità, detta *soggettiva*, applicabile quindi a esperimenti casuali i cui eventi elementari non siano ritenuti ugualmente possibili e che non siano necessariamente ripetibili più volte sotto le stesse condizioni: *la probabilità di un evento E è il prezzo p che si è disposti a pagare per ricevere l'importo 1 se E risulta vero (scommessa vinta) e l'importo 0 se E risulta falso (scommessa persa)*. Al fine di rendere concretamente applicabile la definizione, si richiede che la scommessa sia *equa*, nel senso che chi fissa il prezzo *p* deve essere disposto a fare sia da *scommettitore*, sia da *banco*. Se l'evento *E* non si verifica, l'importo *p* rappresenta il guadagno del banco; se invece l'evento *E* si verifica non bisogna confondere l'importo della vincita col guadagno dello scommettitore $q = 1\text{-}p$ che è la differenza tra la vincita e il prezzo pagato *p*.

Per un determinato evento *E*, il prezzo *p* di una scommessa equa coincide con la probabilità dell'evento *E*, *P(E),* sia che la si calcoli secondo la definizione classica che secondo la definizione frequentista. In mancanza di criteri certi e assoluti, il prezzo equo della scommessa va pattuito fra le parti interessate su basi *soggettive*; in ciò sta una limitazione della definizione soggettiva.

Poste le seguenti definizioni:

Def.1. Dato un evento E, dicesi suo evento contrario o complementare, E^c*, quell'evento che è vero se E è falso e viceversa.*

Def.2. Dati due eventi E e F, dicesi loro somma logica o unione, $E \cup F$*, quell'evento che è vero quando almeno uno dei due eventi E e F è vero.*

Def.3. Dati due eventi E e F, dicesi loro prodotto logico o intersezione, $E \cap F$*, quell'evento che è vero quando entrambi gli eventi*

E e F sono veri.

Def.4. Due eventi E e F, sono incompatibili o disgiunti se il verificarsi di uno esclude il verificarsi l'altro, cioè se la loro intersezione è l'evento impossibile.

qualunque sia la definizione di probabilità assunta (classica, frequentista, soggettiva) sussistono le seguenti proprietà:

Prop.1: $0 \le P(E) \le 1$; si ha $P(E)=0$ se E è impossibile, $P(E)=1$ se E è certo.

Prop.2: $P(E^c) = 1 - P(E)$.

Prop.3: $P(E \cup F) = P(E) + P(F) - P(E \cap F)$; chiaramente se E e F sono incompatibili allora vale la regola della somma: $P(E \cup F) = P(E) + P(F)$.

Dalla veridicità dei precedenti assunti, qualunque sia il punto di vista adottato (classica, frequentista, soggettiva), deriva che sono ugualmente valide tutte le conseguenze che se ne possono trarre indipendentemente dell'itinerario percorso.

Le suddette proprietà si sogliono assumere come una *definizione matematica o assiomatica di probabilità*. Tale definizione venne proposta da Kolmogorov in *Concetti fondamentali del calcolo delle probabilità* nel 1933 a seguito al dibattito fra quanti consideravano la probabilità come limiti di frequenze relative e quanti cercavano un fondamento logico della stessa.

E' da precisare che questa definizione assiomatica non è una *definizione operativa* e non fornisce indicazioni su *come* calcolare la probabilità.

Bibliografia

AA.VV., *Statistica Probabilità Logica,* Tringale Editore, 1983

Boyer C.B., *Storia della Matematica*, Isedi, 1976

Cipolla M., *Matematica Elementare*, Palumbo, 1962

Spiegel M. R., *Probabilità e Statistica*, McGraw-Hill, 1944

Villa M., *Repertorio di Matematiche*, Cedam, 1978