ACCADEMIA PICENO APRUTINA DEI VELATI

SCUOLA ESTIVA PER I DOCENTI DELLA SCUOLA SECONDARIA DEL SECONDO CICLO DI ISTRUZIONE

LA MATEMATICA DELL'INCERTO

INSEGNAMENTO DELLA PROBABILITÀ E DELLA STATISTICA

CASTELLAMMARE DI STABIA
ISTITUTO SUORE COMPASSIONISTE
15 - 16 - 17 - 18 LUGLIO
2018

APAV

Accademia Piceno – Aprutina dei Velati in Teramo
e
Sezioni Mathesis di Napoli e di Castellammare di Stabia

Scuola Estiva di Formazione per i docenti della Scuola Secondaria del Secondo Ciclo di Istruzione

Castellammare di Stabia, Istituto Suore Compassioniste
15, 16, 17, 18 luglio 2018

La Matematica dell'incerto
*Insegnamento della Probabilità e della Statist*ica

Presidente dell'APAV
Giuseppe Manuppella

Direttore del Corso
Salvatore Rao

Responsabile dell'organizzazione
Elisa Savarese

Comitato Scientifico
Loredana Biacino, Aniello Buonocore, Ferdinando Casolaro, Luciano Corso, Giangiacomo Gerla, Antonio Maturo, Raffaele Prosperi, Salvatore Rao, Salvatore Sessa, Massimo Squillante, Carlo Toffalori, Roberto Tortora.

Comitato organizzatore
Giovanna Della Vecchia, Celia Di Foggia, Antonio Fontana, Ciro Iovino, Mario Mandrone, Rosa Porciello, Francesco Quagliano, Alessandra Rotunno, Elisa Savarese, Alberto Trotta.

Curatori del testo
Ferdinando Casolaro, Giuseppe Manuppella, Antonio Maturo

Impaginazione: Fabio Manuppella
Sito web: www.fabiomanuppella.it
E-mail: fabiomanuppella@gmail.com

Copertina: Fabrizio Di Nicola
E-mail: fabriziopluc@gmail.com

Programma

Domenica 15 luglio 2018
16:00-16:30 *Apertura dei lavori e saluti iniziali*.
Elisa SAVARESE, Presidente sezione Mathesis di C/mare di Stabia;
Giuseppe MANUPPELLA, Presidente Accademia Piceno-Aprutina (APAV);
Giangiacomo GERLA, Università di Salerno;
Salvatore SESSA, Università di Napoli;
Massimo SQUILLANTE, Prorettore Università del Sannio;
Roberto TORTORA, Università di Napoli, Presidente CIIM;
Salvatore RAO, Presidente onorario sezione Mathesis Napoli, Direttore della Scuola.

16:30-18:30 *Primo e secondo seminario*. Presiede Elisa SAVARESE.
16:30-17:20 Salvatore RAO, "*Choice and Chance: calcolo combinatorio e probabilità discreta*".
17:20-17:30 *Discussione*.
17:30-18:20 Roberto TORTORA, *"Situazioni problematiche aperte in probabilità"*.
18:20-18:30 *Discussione*.

18:30-20:00 Organizzazione e primo incontro dei Gruppi di lavoro.
Presiede Alberto TROTTA, Fondatore sezione Mathesis Anzio-Nettuno.
Coordinatori:
Loredana BIACINO, Università di Napoli,
Giovanna DELLA VECCHIA, Mathesis Napoli,
Mario MANDRONE, Università del Sannio,
Bonaventura PAOLILLO, Mathesis Salerno,
Raffaele PROSPERI, Università di Napoli.

20:30 *Cena*

Lunedì 16 luglio 2018

9:00-10:05 *Terzo seminario*.
Presiede Giangiacomo GERLA, Università di Salerno.
9:00-9:50 Carlo TOFFALORI, Università di Camerino, Presidente Mathesis di Camerino,*"Algoritmi probabilistici e numeri primi"*.
9:50-10:05 *Discussione*.
10:05-11:15 *Laboratorio 1*.

11:15-11:30 *Pausa caffè*

11:30-13:00 *Laboratorio 1 e presentazioni di Esperienze didattiche*.

13:30 *Pranzo*

15:00-16:05 *Quarto seminario*. Presiede Aniello BUONOCORE, Università di Napoli.
15:00-15:50 Ferdinando CASOLARO, Università di Napoli, Presidente sezione Mathesis di Napoli, *"Matematica dell'incerto: quale percorso nella scuola secondaria di secondo grado?"*.
15:50-16:05 *Discussione*.
16:05-17:15 *Laboratorio 2*.

17:15-17:30 *Pausa caffè*

17:30-20:00 *Laboratorio 2 e presentazioni di Esperienze didattiche*.

20:30 *Cena*

Martedì 17 luglio 2018

9:00-10:05 *Quinto seminario*.
Presiede Antonio Maturo, Università di Pescara, Presidente sezione Mathesis di Pescara.
9:00-9:50 Luciano CORSO, Presidente sezione Mathesis di Verona, *"Correlazione e regressione, per problemi"*.
9:50-10:05 *Discussione*.
10:05-11:15 *Laboratorio 3*.

11:15-11:30 *Pausa caffè*

11:30-13:00 *Laboratorio 3 e presentazioni di Esperienze didattiche*.

13:30 *Pranzo*

15:00-16:05 *Sesto seminario*.
Presiede Luciano CORSO, Presidente sezione Mathesis di Verona.
15:00-15:50 Antonio MATURO, Università di Pescara, Presidente sezione Mathesis di Pescara, *"Sulla probabilità soggettiva"*.
15:50-16:05 *Discussione*.
16:05-17:15 *Laboratorio 4*.

17:15-17:30 *Pausa caffè*

17:30-20:00 *Laboratorio 4 e presentazioni di Esperienze didattiche*.

20:30 *Cena*

Mercoledì 18 luglio 2018
9:00-10:05 *Settimo seminario*.
Presiede Salvatore SESSA, Università di Napoli.
9:00-9:50 Aniello BUONOCORE, Università di Napoli, *"Indizi, Pregiudizi e il Ragionamento Bayesiano"*.
9:50-10:05 *Discussione*.
10:05-11:15 *Relazioni dei gruppi di lavoro*.

11:15-11:30 *Pausa caffè*

11:30-12:30 *Tavola rotonda plenaria:*
"L'insegnamento di probabilità e statistica: questioni didattiche"

12:30-13:00 *Conclusioni*
Elisa SAVARESE e Salvatore RAO

13:30 *Pranzo*

Indice

Prefazione

Insegnamento della Probabilità e della Statistica nella Scuola. Perché? Come?
La prima risposta alla prima domanda è quella ovvia (e burocratica): perché dobbiamo. Ormai da molti anni (prima Piano Nazionale Informatica e sperimentazione Brocca; poi Indicazioni Nazionali per i Licei e Linee Guida per gl'Istituti tecnici) gli argomenti di probabilità e di statistica sono compresi fra quelli di matematica (ora abbastanza dettagliatamente elencati sotto il titolo *Dati e previsioni*). Quindi, non ci sono dubbi, questi argomenti come gli altri vanno adeguatamente trattati.
Anche sul come la via da seguire è obbligata. Prevalenza di metodi induttivi su quelli deduttivi. Partenza da *significative situazioni problematiche* relative a questioni reali [analisi di dati in ambiti *deterministico* (fisico, chimico, biologico, …) e *non deterministico* (biologico, sociale, demografico, economico, ...)]; individuazione del *problema*, sua "matematizzazione" e costruzione o selezione di un *modello risolutivo*, uso di questo modello per produrre *previsioni* o risultati, rilevazione dei *dati empirici* corrispondenti a queste previsioni, *confronto* fra dati empirici e previsioni, *valutazione* di efficienza e adeguatezza del modello, eventuale modifica o sostituzione del modello e ripetizione del ciclo. Dunque conviene evitare di trattare la probabilità a partire *soltanto* dal calcolo combinatorio e da questioni più o meno intricate di dadi e lotterie. Conviene invece iniziare con tematiche di statistica, trattare poi quelle di probabilità, che a loro volta dovranno precedere l'inferenza. Insomma non *Insegnamento di Probabilità e Statistica*, ma *Insegnamento di Statistica, Probabilità e Statistica*. Naturalmente tutte queste questioni saranno al centro soprattutto delle attività di laboratorio della Scuola Estiva.
La prima domanda ("Perché?") richiede però anche altre risposte, indipendenti da (indicazioni sui) programmi scolastici.
Tre premi Nobel assegnati dall'Accademia di Svezia per l'economia:

- nel 2017, a Richard THALER (1945), per le sue ricerche sull'*economia comportamentale* (sistematica applicazione della psicologia allo studio delle decisioni economiche);
- nel 2013, a Robert SHILLER (1946), per i suoi lavori di *finanza comportamentale* (analisi empirica dei prezzi di azioni e di bond e valutazione del rischio);
- nel 2002, a Daniel KAHNEMAN (1934) fondatore con Amos TVERSKY (1937-1996) (non premiato perché precocemente scomparso) dell'*economia comportamentale*.

La teoria dell'utilità classica (sistemata da John von NEUMANN e Oskar MORGENSTERN nel 1944 con *Teoria dei giochi e comportamento economico*) viene sostituita in [1][1] con la *Teoria del prospetto: una analisi delle decisioni in condizioni di rischio*. Il modello descrittivo delle decisioni prese dagli individui mette ora in conto l'*effetto certezza* (guadagni certi sono preferiti a guadagni maggiori ma solo probabili; perdite maggiori probabili sono preferite a perdite minori certe; avversione al rischio nel

[1] [1] Kahneman D. & Tversky A., (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47 (2), pp. 263-292.

caso di guadagni e preferenza del rischio nel caso di perdite); l'*effetto isolamento* (tra due alternative si tralasciano gli aspetti comuni e si vedono soltanto le differenze, formulando così preferenze incoerenti: di fronte a due scommesse di fatto identiche, ma presentate diversamente, le persone si comportano come se le scommesse fossero diverse); l'*effetto contesto* (le preferenze variano secondo il contesto in cui le alternative vengono inquadrate); l'*effetto di dotazione* (il dispiacere di perdere una somma è maggiore del piacere di vincere la stessa somma; si attribuisce un valore assurdo ai beni posseduti, che non si ricomprerebbero mai ai prezzi attuali di mercato, e non ci si vuole liberare di tali beni).

Più in generale, nella loro vita quotidiana, le persone fanno scelte ed esprimono giudizi intuitivi secondo euristiche spesso non consapevoli e molto distanti da forme di razionalità scientifica (si vedano per esempio [3][2] e [4][3]).

D'altra parte si può assumere che una buona conoscenza delle basi del calcolo delle probabilità e della statistica potrebbe certo aiutarci a prendere decisioni più consapevoli e più vantaggiose. O almeno a renderci conto della distanza di molte delle nostre intuizioni da fatti scientificamente accertati. Con vantaggio nostro e di tutti. Ecco perché è importante sviluppare queste conoscenze nelle scuole (si pensi, per esempio, alla piaga sociale della dipendenza dal gioco).

Vi è poi almeno un terzo motivo per studiare le basi del calcolo delle probabilità e della statistica. Ed è una motivazione strettamente "culturale" e in connessione con la matematica. In proposito si veda per esempio l'intervento di David Bryant MUMFORD (1937), il famoso geometra algebrico medaglia Fields 1974, al Convegno Internazionale «Mathematics towards the Third Millennium», tenutosi a Roma dal 27 al 29 maggio 1999 (si veda [2][4]). Traduco l'Abstract:

<<La logica di Aristotele ha dettato le regole al pensiero degli intellettuali occidentali. Tutte le precise teorie, tutti i modelli scientifici, anche i modelli dello stesso processo del pensiero, si sono conformati almeno in linea di principio al corretto vestito della logica. Ma dai loro oscuri inizi, con l'invenzione di strategie per giochi d'azzardo e col conteggio di cadaveri nella Londra medievale, la teoria della probabilità e l'inferenza statistica ora emergono come la migliore fondazione per i modelli scientifici, specialmente per quelli del processo del pensiero e come ingredienti essenziali di teorie matematiche, anche come fondazioni della stessa matematica. Si propone che questo notevole mutamento di prospettiva influenzerà effettivamente tutta la matematica del prossimo secolo.>>

A tutto questo si cercherà di dare spazio nella Scuola estiva.

Salvatore Rao

[2] [3] Ovadia D. (2018). La mente probabilistica. *Mind: Mente&Cervello.* 163 (7), pp. 52-55.

[3] [4] Tversky A. & Kahneman D., (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, New Series, 185 (4157), pp. 1124-1131.

[4] [2] Mumford D., (2000). The Dawning of the Age of Stochasticity. *Rend. Mat. Acc. Lincei.* 9 (fasc. spec.), pp. 107-125.

Primi al novanta per cento…

Carlo Toffalori[1]

[1]Università di Camerino
Sezione di Matematica, Scuola di Scienze e Tecnologie
E-mail: carlo.toffalori@unicam.it

Sunto *Le proposizioni dell'aritmetica prevedono in genere due soli valori di verità, sì o no. Per esempio, un numero naturale maggiore di 1 o è primo o è composto. Ha senso essere primi, o composti, al novanta per cento? Servono a qualcosa verità parziali e insicure? La risposta è positiva, per i motivi di rapidità ed efficienza che cercheremo di illustrare in queste note.*

L'aritmetica

- l'insieme **N** dei numeri naturali 0, 1, 2, 3, 4, 5, 7, 8, 9, 10, 11, …
- le operazioni di addizione, moltiplicazione, *sottrazione* e *divisione*

Dante, *Convivio*
"L'arismetrica… del suo lume tutte si illuminano le scienze… l'occhio de lo 'ntelletto nol può mirare…"

Due citazioni (libere) più moderne

- L. Kronecker: *"i numeri interi sono i soli creati da Dio"*
- H. Weyl (o A. Weil?): *"i teoremi di incompletezza di Gödel dimostrano l'esistenza di Dio e anche quella del demonio"*

Le difficoltà della divisione: un numero naturale $N > 1$ è

- *primo* se gli unici divisori di N sono 1 e N,
- *composto* altrimenti

I numeri primi: 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, ...

Teorema fondamentale dell'aritmetica. Ogni numero naturale $N > 1$ si decompone in uno ed un solo modo (a meno dell'ordine dei fattori) nel prodotto di numeri primi.

$$15 = 3 \cdot 5, \quad 18 = 2 \cdot 3^2, \quad 24 = 2^3 \cdot 3, \quad 32 = 2^5, \ldots$$

La moltitudine dei numeri primi…

Teorema di Euclide. I numeri primi sono infiniti.
Tuttavia la successione dei primi è irregolare e imprevedibile
I primi gemelli: con differenza 2, dunque p e $p + 2$

- esempi: 3 e 5, 5 e 7, 11 e 13, 17 e 19, …
- controesempi: 7 e 11, 13 e 17, 19 e 23, 23 e 29, …

Non è chiaro quante siano le coppie di primi gemelli (infinite?).

Eppure, ***per ogni intero positivo s si possono trovare s numeri consecutivi nessuno dei quali è primo***. Basta considerare

- $(s + 1)! + 2 > 2$, divisibile per 2, dunque composto,
- $(s + 1)! + 3 > 3$, divisibile per 3, dunque composto,

… … …

- $(s + 1)! + s + 1 > s + 1$, divisibile per $s + 1$, dunque composto.

Ricordare: il teorema dei numeri primi, e molto altro…

La congettura di Goldbach, 1742

- C. Goldbach, lettera a Eulero: *ogni numero naturale $N \geq 6$ è la somma di 3 primi*
- L. Eulero, risposta a Goldbach: equivalente provare che C. Goldbach, ***ogni numero pari $M \geq 4$ è la somma di 2 primi***

La congettura di Goldbach:

- basta un controesempio per smentirla,
- non bastano miliardi di miliardi di miliardi esempi per confermarla

4 = 2 + 2, 6 = 3 + 3, 8 = 3 + 5, 10 = 3 + 7, 12 = 5 + 7, 14 = 3 + 11, …

$2p = p + p$ per ogni primo p

Due problemi classicissimi: primalità e fattorizzazione

- *Un input comune*: un naturale $N \geq 2$
- *L'output atteso*

Primalità: N è primo o composto?
Fattorizzazione: la decomposizione di N in fattori primi (il primo passo: per N composto dispari, un divisore proprio d di N diverso da 1 e N).

Lecito assumere N dispari, i numeri pari si trattano facilmente.

Un algoritmo elementare per entrambi i casi (noto già agli antichi Greci)

- Si prende $1 < d < N$ e si divide N per d.
- Se la divisione è esatta per qualche d, si deduce che N è composto (e anzi si trova un suo divisore d diverso da 1 e N)
- Se la divisione non è esatta per nessun d, si deduce che N è primo.

Il commento di Gauss, *Disquisitiones Aritmeticae*, 1801 (articolo 329)

“Il problema di separare i primi dai composti e di decomporre i secondi nei loro fattori primi è conosciuto essere uno dei più importanti e utili in Matematica. La dignità stessa della scienza sembra richiedere di esplorare ogni possibile mezzo per la soluzione di un problema così elegante e famoso”

Perché questa esigenza due millenni dopo i Greci?
Ancora Gauss: *“Le tecniche conosciute in precedenza richiederebbero una fatica intollerabile anche per i più instancabili* ***calcolatori.****”*

Il costo dell’algoritmo “elementare”

- $N - 2$ divisioni, una per ogni possibile $d = 2, \dots, N - 1$,
- da riferire non a N, ma alla ***lunghezza*** di N

Ma la lunghezza di N in base 10 (o 2) è all’incirca il suo logaritmo in quella base!

- Un miliardo 1.000.000.000 si compone di 10 cifre
- 100 è piccolo, ma 2^{100} supera abbondantemente l’età del mondo in secondi

Nel caso specifico:

- ogni singola divisione ha costo quadratico rispetto alla lunghezza di N,
- le divisioni richieste nei casi peggiori sono quasi a 10 elevati alla lunghezza di N.

Possibili scorciatoie

- Esplorare $1 < d < \sqrt{N}$: se $N = d \cdot d'$ con $d, d' > 1$, uno tra d e d’ è $\leq \sqrt{N}$.
- Se d non funziona, neppure $2d, 3d, \dots$ vanno bene.

Ma i guai restano! Sono prevedibili fino a $\sqrt{N} - 1$ divisioni, ossia una quantità ancora esponenziale rispetto alla lunghezza di N.

Scindere primalità e fattorizzazione
Teorema di Wilson: per ogni intero $N > 1$, N è primo se e solo se $(N-1)! \equiv -1 \ (mod\ N)$.
Commenti

- Si controlla la *primalità* di *N* senza bisogno di considerarne la *fattorizzazione*.
- La verifica richiede comunque il calcolo di $(N-1)!$, dunque $N-2$ moltiplicazioni (“troppe” rispetto alla lunghezza di *N*).

A proposito di congruenze

Il piccolo teorema di Fermat: se *N* è primo, allora ogni intero a primo con *N* soddisfa $a^{N-1} \equiv 1 \ (mod\ N)$.

Una dimostrazione rapida:

- si considera il gruppo moltiplicativo del campo $\mathbb{Z}_N$ che ha ordine $N - 1$ e si compone degli interi modulo N primi con N, e dunque invertibili modulo N;
- si applica il teorema di Lagrange.

Da notare: in riferimento alla lunghezza di N, è "rapido" controllare

- a, N primi tra loro (l'algoritmo di Euclide richiede tempi al più quadratici),
- $a^{N-1} \equiv 1 \ (mod\ N)$ (basta una divisione, dunque un tempo quadratico).

Pure l'elevamento a potenza $a \rightarrow a^N - 1 \ (mod\ N)$ ammette procedure veloci di calcolo.

Curiosità numero 1. Sia $N > 1$ un intero dispari. Fissiamo a primo con N, $1 < a < N$, ad esempio $a = 2$ (visto che N è dispari). Se vale $a^{N-1} \equiv 1 \ (mod\ N)$, allora N è primo?

Da ricordare: la congruenza $a^{N-1} \equiv 1 \ (mod\ N)$ è "rapida" da verificare.

La risposta: NO, ci sono numeri N. composti che soddisfano $a^{N-1} \equiv 1 \ (mod\ N)$ (pseudoprimi in base a)

Un esempio (Sarrus): 341 = 11 · 31 è pseudoprimo in base 2 (anche se non in base 3).

Curiosità numero 2. Se vale $a^{N-1} \equiv 1 \ (mod\ N)$ per **ogni** intero a primo con N, $1 < a < N$, allora N è primo?

Da notare: le congruenze da controllare possono arrivare, per N primo, a $N - 1$ (troppe rispetto alla lunghezza di N).

La risposta: NO, ci sono numero composti dispari N che soddisfano $a^{N-1} \equiv 1 \ (mod\ N)$ per ogni intero a primo con N, $1 < a < N$: gli *pseudoprimi di Carmichael* (1912)

Esempi: 561, 1105, 1729, 2465, 2821, 6601, 8911, 10585, 15841, 29341, 41041, 46657, 52633, 62745, 63973, … (**infiniti!** Alford-Granville-Pomerance, 1994)

Altri algoritmi di primalità basati sul piccolo teorema di Fermat?

Carl Pomerance: *"la congruenza di Fermat è così semplice che è un peccato doverci rinunciare"*

Un'altra premessa utile*: le radici quadrate modulo un primo N*

Per N primo dispari, ogni quadrato modulo N primo con N ha esattamente 2 radici quadrate modulo N:

$$se\ x^2 \equiv a^2 (mod\ N),\ allora\ x \equiv \pm\, a\ (mod\ N),\ e\ inoltre\ \ a \not\equiv -a\ (mod\ N).$$

Infatti un primo N che divide $x^2 - a^2 = (x - a) \cdot (x + a)$ divide $x - a$ o $x + a$. Inoltre N non può dividere $a - (-a) = 2a$ perché dispari e primo con a.

In particolare le radici quadrate di 1 modulo N sono ± 1.

Non sempre vero per N composto! Per esempio $x^2 \equiv 1 \ (mod\ 8)$

- per $x \equiv \pm 1 \ (mod\ 8)$
- ma anche per $x \equiv \pm 3 \ (mod\ 8)$.

Un approccio sorprendente e discusso: gli ***algoritmi probabilistici*** (le classi BPP, RP, $coRP$, ZPP della teoria della complessità computazionale)

L'idea: ***sacrificare la precisione per aumentare la velocità***

- Si ricorre all'aiuto di testimoni casuali
- Si ammettono risposte sbagliate o incerte purché i tempi di lavoro siano accelerati
- Si cerca di rendere minima la probabilità di errore o di insicurezza (al variare dei testimoni)

Émile Borel: *un evento che ha probabilità < 10^{-50} non accadrà mai, e se anche accade non sarà mai rilevato*.

Algoritmi probabilistici

- *tipo Montecarlo*: risposte probabilmente vere in tempi certamente rapidi
- *tipo Las Vegas*: risposte certamente vere in tempi probabilmente rapidi

In altre parole

- nel primo caso è ammessa la possibilità che un testimone sbagli,
- nel secondo caso è ammessa la possibilità che un testimone dica *"non lo so"*,

ed entrambe sono da rendere minime!

*Un esempio di algoritmo Montecarlo per i primi: l'***algoritmo di Miller-Rabin**, 1978, fallibile ma efficiente!

I prerequisiti (quelli già ricordati): se N è primo (dispari),

- le radici quadrate di 1 modulo N sono ± 1,
- (*il piccolo teorema di Fermat*) per ogni intero a primo con N, vale $a^{N-1} \equiv 1 \ (mod\ N)$.

Il procedimento
È dato l'intero ***dispari*** $N > 2$. Dunque $N - 1$ è pari. Si interpella un testimone intero a con $1 < a < N$.

Passo 1. Si calcola il massimo comune divisore d di a e N.

- Se $d \neq 1$, l'algoritmo risponde (correttamente) ***"N COMPOSTO"*** (ha trovato un divisore non banale di N, appunto d)
- Se $d = 1$, si procede con i passi successivi.

Passo 2. Si calcola a^{N-1} modulo N.

- Se $a^{N-1} \not\equiv 1 (mod\ N)$, l'algoritmo risponde (correttamente) ***"N COMPOSTO"*** (è smentito il piccolo teorema di Fermat).
- Se invece $a^{N-1} \equiv 1 (mod\ N)$, si va al passo successivo.

Passo 3. Si calcola $a^{\frac{N-1}{2}}$ modulo N (la radice quadrata di a^{N-1} modulo N).

- Se $a^{\frac{N-1}{2}} \not\equiv \pm 1$ (mod N), l'algoritmo risponde (correttamente) ***"N COMPOSTO"***.

- Se $a^{\frac{N-1}{2}} \equiv -1\ (mod\ N)$, l'algoritmo azzarda ***"N PRIMO"***.

- Idem se $a^{\frac{N-1}{2}} \equiv 1\ (mod\ N)$ ma $\frac{N-1}{2}$ non è pari, cioè non si può dimezzare ulteriormente.

- Altrimenti, se $a^{\frac{N-1}{2}} \equiv -1\ (mod\ N)$ e $\frac{N-1}{2}$ è pari, cioè esiste $\frac{N-1}{4}$ tra gli interi , si va ai passi successivi.

Passo 4. Come il passo 3, ma riferito ad $a^{\frac{N-1}{4}}$ modulo *N* (la radice quadrata di $a^{\frac{N-1}{2}}$ modulo N).

Si prosegue finché possibile.

Commenti

1. ***Attendibilità***: si noti che

- la risposta ***"N COMPOSTO"*** è sicura.
- quella ***"N PRIMO"*** può essere sbagliata.

Probabilità di errore dopo la consultazione di un testimone a: al variare di a, al più $\frac{1}{4}$
In altre parole: almeno 3 testimoni su 4 dicono la verità!
D'altra parte, dopo la consultazione di 100 testimoni a che rispondono concordemente ***"N PRIMO"***, la probabilità di errore diventa

$$\frac{1}{4^{100}} < \frac{1}{10^{50}}.$$

2. ***Tempi di lavoro***: sono da calcolare (**finché possibile**)

- dapprima un massimo comune divisore,
- poi per ogni passo un elevamento a potenza e una divisione.

Complessivamente

- I tempi limitati da un polinomio di grado al più 5 rispetto alla lunghezza di *N* (incluse le prevedibili iterazioni),
- la ripetizione per al più 100 testimoni non li aggrava irreparabilmente (per *N* grande).

Esempi

1) Sia $N = 13$, dunque $N - 1 = 12$. Scegliamo $a = 2$. Allora

- 2 è primo con 13,
- $2^{12} = 4096 \equiv 1\ (mod\ 13)$,
- $2^{6} = 64 \equiv -1\ (mod\ 13)$.

Dunque l'algoritmo dichiara ***"N PRIMO"*** e in effetti *N* lo è.

2) Sia $N = 7$, così $N - 1 = 6$. Riferiamoci sempre ad $a = 2$. Allora

- 2 è primo con 7,
- $2^{6} = 64 \equiv 1\ (mod\ 7)$,
- $2^{3} = 8 \equiv 1\ (mod\ 7)$

dopo di che non è più possibile dividere l'esponente 3 per 2. Dunque l'algoritmo dichiara ***"N PRIMO"*** e in effetti *N* lo è.

3) Sia adesso $N = 561$ (uno pseudoprimo di Carmichael – in altre parole, per ogni intero *b* primo con N, $b^{560} \equiv 1\ (mod\ 561)$). Si sceglie $a = 359$ e con un po' di pazienza si controlla

- 359 è primo con 561,
- $359^{560} \equiv 1\ (mod\ 561)$,
- $359^{280} \equiv 1\ (mod\ 561)$,

- $359^{140} \equiv 1\ (mod\ 561)$,
- $359^{70} \equiv 1\ (mod\ 561)$,
- $359^{35} \equiv 1\ (mod\ 561)$.

Siccome 35 non si può dividere ulteriormente per due, l'algoritmo, se ricorre ad $a = 359$, dichiara ***"N PRIMO"*** mentre N è composto.

Ma se il testimone 359 non è attendibile, $a = 2$ lo è. Infatti

- 2 è primo con 561,
- $2^{560} \equiv 1\ (mod\ 561)$,
- $2^{280} \equiv 1\ (mod\ 561)$,
- $2^{140} \not\equiv \pm 1\ (mod\ 561)$,

Quindi il ricorso ad $a = 2$ dichiara correttamente ***"N COMPOSTO".***

Altri algoritmi Montecarlo per i primi: ***l'algoritmo di Solovay-Strassen***, 1977
Si basa sul simbolo di Legendre-Jacobi e su un teorema di Eulero: ma l'algoritmo di Miller-Rabin è preferibile sia per affidabilità che per efficienza

Un algoritmo rapido e **infallibile***: l'***algoritmo AKS** *(Agrawal, Kayal, Saxena), 2002*
- Agrawal: informatico teorico, con interessi alla complessità computazionale
- Kayal, Saxena: matematici, dottorandi di Agrawal.

Agosto 2002: l'algoritmo AKS
- distingue se un dato intero $N > 1$ è primo o composto
- nelle migliori implementazioni lavora in tempo poco più che polinomiale di grado 6 rispetto alla lunghezza di *N*
- soddisfa finalmente le richieste di Gauss (due secoli dopo la loro formulazione).

La base lontana: ancora il ***piccolo teorema di Fermat***.

*Conseguenze di AKS: l'***algoritmo di Bernstein-Berrizbeitia**, 2004
- probabilistico di tipo Montecarlo
- lavora in tempo poco più che polinomiale di grado 4 rispetto alla lunghezza di *N*
- assolutamente affidabile quando risponde ***"N PRIMO"***
- talora fallibile (ma con bassa probabilità di errore) quando dichiara ***"N COMPOSTO"***

Da usare in combinazione con l'algoritmo di Miller-Rabin, per ottenere un procedimento di tipo Las Vegas. Per ogni input N, si eseguono in parallelo su *N* i due algoritmi, ricorrendo a opportuni testimoni.

- Se l'algoritmo di Bernstein-Berrizbeitia dichiara ***"N PRIMO"***, si conclude che N è primo.
- Se l'algoritmo di Miller-Rabin dichiara ***"N COMPOSTO"***, si conclude N composto.

Il caso disgraziato (rarissimo)

- L'algoritmo di Bernstein-Berrizbeitia dichiara ***"N COMPOSTO"***,
- quello di Miller-Rabin **"*N PRIMO*"**.

La conclusione obbligata: ***"NON LO SO"***, necessario interpellare altri testimoni.

E il problema della fattorizzazione?

Versione "ridotta"
Input: $N > 1$ composto e dispari
Output: un divisore d di N, $1 < d < N$.

In realtà la questione chiave!

Il guaio: i tempi di lavoro degli algoritmi oggi conosciuti per la versione "ridotta" sono esponenziali (o quasi) rispetto alla lunghezza dell'input N.

Gran spreco di strumenti teorici

- curve ellittiche,
- frazioni continue, ...

nessuno capace di lavorare in tempo al più polinomiale nella lunghezza di N.

Ma attenzione alla computazione quantistica (P. Shor, 1994)!

Alcune famose decomposizioni in fattori primi

- Eulero, 1742: F(5) = 641 · 6700417
- Clausen, 1856: F(6) = 274177 · 67280421310721
- Landry, 1869: M(59) = 179951 · 3203431780337
- Cole, 1903: M(67) = 193707721 · 761838257287 (convegno dell'American Mathematical Society a S. Francisco, *"tre anni di domeniche"*)
- Poulet, 1923: M(73) = 439 · 2298041 · 9361973132609 (ma il fattore 439 era già stato scoperto da Eulero).

Solo speculazioni teoriche? Applicazioni alla sicurezza della rete

Il crittosistema a chiave pubblica più popolare: RSA (Rivest, Shamir, Adleman), 1978

Si fonda proprio su
- la facilità di generare e riconoscere i primi,
- la difficoltà di fattorizzare i composti.

Bibliografia.

Leonesi S., Toffalori C. (2006), Numeri e crittografia, Springer Italia, Milano.

Toffalori C. (2015), Algoritmi, il Mulino, Bologna.

La Scuola Estiva e il Premio "Aldo Morelli"

Elisa Savarese[1]

[1]Presidente Mathesis di Castellammare di Stabia
E-mail: elisa.elsava2@gmail.com

Sunto

Si descrive il legame tra due esigenze primarie per l'insegnamento della Matematica nella società moderna: l'attrazione della disciplina come gioco e la proposta relativa all'approfondimento della Matematica dell'incerto per i rischi emersi dalla velocizzazione degli eventi negli ultimi decenni.

Parole chiave: Gioco, Scuola, Incerto.

1. Introduzione

Con molto piacere ho condiviso che la Scuola Estiva per i docenti della Scuola secondaria di secondo grado si tenesse a Castellammare di Stabia presso le Suore Compassioniste, luogo caro a tanti docenti, soprattutto a quelli che partecipano ai Giochi Matematici per la Scuola "Premio Aldo Morelli", arrivato quest'anno alla XII edizione. Un filo invisibile che lega "Logica, intuizione e fantasia" dei Giochi alla Matematica dell'Incerto "Insegnamento della Probabilità e della Statistica".
La scuola è dedicata ad un contenuto specifico del curricolo matematico e ad uno dei cosiddetti temi trasversali. Il contenuto di questa edizione è la Probabilità ed il correlato tema della Statistica, mentre il nucleo trasversale è la Modellizzazione. Si tratta, come è ovvio, di argomenti dei quali è a tutti nota l'importanza (Prosperi R., 2011).

2. Perché la Probabilità? Come decidere?

Due sono le questioni che riguardano la probabilità su cui ci sembra importante porre l'accento. Da un lato vi è la diffusa carenza di un adeguato "pensiero probabilistico" in vasti strati della nostra società, e un'altrettanta grave e correlata difficoltà di insegnamento e comprensione dei suoi elementi nella nostra Scuola.
Che cosa non va nell'attuale insegnamento? Che cosa si può fare? Come sviluppare un'intuizione probabilistica che sia utile nella vita di tutti i giorni?

Ci sono specifiche indicazioni sull'insegnamento della probabilità, diverse da quelle che riguardano l'intera matematica? Sono queste le domande che ci poniamo, e ad esse dedicheremo attenzione nella Scuola Estiva (Casolaro, Prosperi R. 2011).
L'altra questione, ma che è strettamente correlata alla precedente, è l'impatto che hanno il calcolo delle probabilità e la statistica sulla nostra vita quotidiana, sia come individui sia come società. Le nostre decisioni, piccole e grandi, dipendono tutte dalla valutazione di quali ne saranno le conseguenze. Poiché siamo convinti che una maggiore conoscenza delle basi della Probabilità potrebbe aiutare tutti i cittadini a prendere decisioni più consapevoli e più vantaggiose, deduciamo da ciò l'importanza di quanto da insegnanti di Matematica possiamo fare per lo sviluppo di tali conoscenze.
Da un maggiore e più incisivo impegno in tale direzione, ne trarrebbe beneficio anche l'opera di contrasto alla piaga sociale della dipendenza dal gioco di cui tanti giovani sono vittime.
Si riconnette bene a tutto ciò l'argomento Modellizzazione. Qui è però necessario chiarire che cosa intendiamo noi con questo termine, visto che esso è utilizzato in svariati contesti e con accezioni diverse. Seguendo le orme della più recente ricerca in Didattica della Matematica, per noi Modellizzazione, o potremmo dire con precisione forse maggiore, Matematizzazione, è ciò che si fa ogni volta che si traduce in forma matematica una situazione reale, e ciò include il caso in cui si mettono insieme 2 e 3 caramelle, così come quello in cui si studiano fenomeni oltremodo complessi, come ad esempio quelli meteorologici. Porre l'accento sulla Modellizzazione in un contesto didattico significa allora rendersi conto che la Matematica non può ridursi solo alle sue regole interne, ma deve fare sempre i conti con la realtà. Anzi forse è in questo "sporcarsi le mani" a capire se e come un costrutto astratto si adatta a descrivere e magari a risolvere una questione concreta che si mettono veramente alla prova le qualità della Matematica e dei matematici.
Nelle Indicazioni Nazionali per la Scuola secondaria di secondo grado si sottolinea l'importanza di sviluppare nell'allievo competenze relative al processo di matematizzazione, così come l'importanza di costruire nello studio un rapporto proficuo tra le diverse discipline. Ma analogo invito lo si ritrova per gli altri ordini scolastici. Ad esempio: "Nella scuola secondaria di primo grado si svilupperà un'attività più propriamente di matematizzazione, formalizzazione, generalizzazione. L'alunno analizza le situazioni per tradurle in termini matematici, riconosce schemi ricorrenti, stabilisce analogie con modelli noti, sceglie le azioni da compiere" (…).

Il complesso rapporto tra Matematica e realtà è peraltro da lunghissimo tempo al cuore della discussione nell'ambito della ricerca e della formazione in didattica della Matematica. Più di un secolo fa Enriques scriveva: "*Se le matematiche vengono così spesso riguardate come inutile peso dagli allievi, dipende in parte almeno dal carattere troppo formale che tende a prendere quell'insegnamento, da un falso concetto del rigore tutto intento a soddisfare certe minute esigenze di parole, da una critica analitica eccessiva e fuori di posto… Ma queste tendenze si riattaccano ad una causa*

più generale; cioè al fatto che le matematiche siano state studiate come un organismo a sé, riguardandone piuttosto la sistemazione astratta conseguita dopo uno sviluppo secolare, che non l'intima ragione storica. Si dimenticano per tal modo i problemi concreti che conferiscono interesse alle teorie, e sotto la formula o lo sviluppo del ragionamento non si vedono più i fatti ormai da lungo tempo acquisiti, ma soltanto la concatenazione in cui noi artificialmente li abbiamo stretti"
(F. Enriques, Sulla preparazione degli insegnanti di scienze, relazione tenuta al V Congresso degli insegnanti di scuole medie, 1906).

I processi di modellizzazione devono essere dunque seriamente presi in considerazione come questione didattica. Si tratta di capire come introdurli nella pratica; quali competenze richiedono, come si promuovono tali competenze. Inoltre il collegamento di matematica e realtà solleva il problema del rapporto della matematica con le altre discipline: bisogna vedere la Matematica come strumento per le altre discipline, o piuttosto sono queste da concepire come campo di esperienza per costruire conoscenze matematiche?

3. La Scuola estiva e i Giochi

La Scuola Estiva affronterà tali questioni, sia in termini generali, cioè con riferimento a tutta la Matematica, sia più specificamente ricollegandosi all'argomento Probabilità. Questo infatti, forse più di ogni altro argomento di Matematica, si presta come territorio in cui confrontarsi continuamente con la realtà. Non sono infatti gli aspetti della teoria astratta della probabilità ad avere interesse nel curricolo scolastico, a differenza di ciò che avviene per altri argomenti di algebra, ad esempio, o di geometria.

Come di consueto tutte le questioni saranno affrontate con l'alternanza di Seminari e laboratori, ed anche organizzando sia attività separate per i vari ordini di scuola sia attività condivise, e sempre sollecitando la discussione e la partecipazione attiva di tutti.
Certamente una maggiore conoscenza dell'applicazione della matematica ai problemi reali renderà più proficuo il contributo che i docenti potranno dare ai Giochi Matematici per la Scuola "Premio Aldo Morelli" (Savarese E., 2014).
"Logica, intuizione e fantasia" vuole comunicare con immediatezza che i "Giochi Matematici" sono delle gare matematiche ma che, per affrontarle, non è necessaria la conoscenza di nessun teorema particolarmente impegnativo o di formule troppo complicate. Occorre invece una voglia matta di giocare, un pizzico di fantasia e quell'intuizione che fa capire che un problema apparentemente difficile è in realtà più semplice di quello che si poteva prevedere.
Un gioco matematico è un problema con un enunciato divertente e intrigante, che suscita curiosità e la voglia di fermarsi un po' a pensare. Meglio ancora se la stessa

soluzione sorprenderà poi per la sua semplicità ed eleganza (Casolaro F., Prosperi R., 2014).

Concretamente, i "Giochi matematici Premio Aldo Morelli" sono una gara articolata in tre fasi:

- Prima fase: febbraio (gara matematica di tipo promozionale, nella quale si auspica la massima affluenza di studenti), da svolgersi presso l'Istituto di appartenenza degli studenti, con la supervisione del referente di istituto.

- Seconda fase: marzo, gara matematica di selezione da svolgersi nelle sedi individuate dai referenti delle varie province.

- Terza fase: maggio, gara finale per l'individuazione dei vincitori e gara a squadre degli alunni della Scuola superiore di II grado; Seminari di aggiornamento docenti con rilascio di Attestato di partecipazione; premiazione e convegno presso la sede Mathesis di Castellammare di Stabia (NA).

Il premio si rivolge a quattro diverse tipologie di studenti:

a) alunni della IV e V elementare;

b) alunni della II e III media (s.s. di primo grado);

c) allievi del biennio delle superiori (I biennio s.s. di secondo grado);

d) allievi del triennio delle superiori (II biennio e III anno s.s. di secondo grado)

e) gara a squadre (4 allievi della scuola secondaria superiore)

Alla prima fase possono partecipare tutti gli studenti delle scuole iscritte al "Premio Aldo Morelli".
Invece, alla seconda fase, potranno partecipare, per la scuola primaria e secondaria di I grado 10 alunni per scuola; per la scuola secondaria di II grado solo 5 alunni per scuola. In ogni caso, nelle situazioni di ex aequo ciascuna scuola troverà criteri interni per individuare gli alunni che parteciperanno alla seconda fase dei "Giochi matematici per la scuola".
Alla terza e ultima fase sono ammessi:

- il primo classificato di ogni scuola;

- i primi classificati in graduatoria nazionale.

Il numero totale degli alunni ammessi alla terza fase, da tenersi a Castellammare di Stabia, sarà fissato dal Comitato Scientifico e comunicato, entro il 9 febbraio alle scuole partecipanti. Ogni scuola può iscrivere una sola squadra di quattro alunni.
Per tutte le fasi la vigilanza sarà affidata a docenti non della classe mentre l'organizzazione e la supervisione sarà affidata a componenti del comitato o loro delegati (Savarese E. 2018).

Sono previsti premi per tutte le categorie.

I "Giochi matematici per la scuola" dedicati alla memoria del Prof. Aldo Morelli, per molti anni Presidente sapiente della sezione Mathesis di Napoli, convinto sostenitore del ruolo formativo, nei processi di apprendimento degli studenti, della geometria e degli aspetti storici della matematica (Casolaro F., Olivello T., 2002)
Con lo spirito rivolto a questi principi e a questi valori partecipiamo alla Scuola Estiva con entusiasmo e ottimismo.

Bibliografia

Enriques F. (1906), “*Sulla preparazione degli insegnanti di scienze*”, relazione tenuta al V Congresso degli insegnanti di scuole medie, 1906.

Casolaro F., Olivello T, (2002) “*Raccolta di scritti per l'insegnamento*” del prof. Aldo Morelli: Convegno organizzato dal Dipartimento di Matematica R. Caccioppoli dell'Università di Napoli “*Geometria e Didattica: convegno in onore di Aldo Morelli*”.

Prosperi R. (2011), *La Matematica dell'incerto: l'insegnamento del* “*Calcolo delle Probabilità*” *e della* “*Statistica*” *nel I biennio della Scuola Secondaria di II grado*. Ed. 2C Contact 2011.

Casolaro F., Prosperi R. (2011), “Atti Terni 2011: *La Matematica per la Scuola di 2° secondo grado: un contributo per il docente di Matematica*”. Ed. 2C Contact.

Savarese E., (2014), “*Giochi di Matematica 2013, Premio Morelli*” Convegno 2013 “*Premio Aldo Morelli*”, Castellammare di Stabia, pubblicato in Journal of Science & Philosophy DiSciPhil - Fascicolo 1 (2014).

Casolaro F., Prosperi R. (2014), Atti del Convegno 2013 “*Premio Aldo Morelli*”, Castellammare di Stabia, pubblicato in Journal of Science & Philosophy DiSciPhil - Fascicolo 1 (2014).

Savarese E., (2018), “*Il Premio Aldo Morelli*” Corso di Formazione APAV-Mathesis “*L'attualità degli insegnamenti dei Grandi Maestri della Mathesis nella seconda metà del XX secolo*”. Rimini, aprile 2018.

La logica della probabilità: un viaggio attraverso l'incertezza[1]

Massimo Squillante[2], Maria Incoronata Fredella[2], Maria Grazia Olivieri[2], Gaetano Vitale[3]

[2]Dipartimento DEMM, Università del Sannio, email squillan@unisannio.it
[3]Università degli studi di Salerno, email gvitale@unisa.it

Sunto

Nella vita reale ci si trova di fronte a molte situazioni caratterizzate da incertezza, imprecisione, vaghezza. Sono state introdotte diverse modellizzazioni per il trattamento di tali concetti e problemi. Ci proponiamo di esporre sinteticamente alcuni lineamenti fondamentali di Probabilità, Statistica e Fuzzy Logic.
La probabilità ha origini recenti rispetto alle altre branche della matematica che hanno profonde radici nel passato, come la geometria o l'algebra.
Possiamo dire che un passaggio iniziale importante si è avuto con Antoine Gombaud, Chevalier de Méré (1607-1684), che pose a Blaise Pascal (1623-1662) una questione riguardante il gioco dei dadi. La corrispondenza tra Pierre de Fermat e Blaise Pascal, che ha avuto inizio nel 1654, su questioni simili, ha portato all'introduzione di concetti di base, come probabilità e aspettativa. Successivamente Christian Huygens, in "De ludo Ratiociniis in aleae", ha proposto un primo studio sistematico della nuova branca della matematica. Tuttavia, la necessità di una costruzione assiomatica della teoria della probabilità sorse per l'esigenza di analizzare situazioni più generali e complesse rispetto al gioco d'azzardo. Una forte formalizzazione è stata fornita dalla monografia "Fondamenti della teoria della probabilità" (1933) di Andrey Nikolaevich Kolmogorov.
La statistica rappresenta l'applicazione più popolare della teoria della probabilità, fornendo strumenti di ricerca in diversi settori, tra cui le scienze fisiche e naturali, la tecnologia, la psicologia, l'economia e la medicina. In un certo senso essa rappresenta il ponte che collega i dati sperimentali con la teoria matematica.
La Logica Fuzzy, da non confondere con la probabilità, si occupa del trattamento formale delle proposizioni di cui non si può affermare senza ambiguità che siano vere o false; l'idea filosofica è che "tutto è una questione di gradualità" (Zadeh).

Parole chiave: Incertezza, Probabilità, Statistica, Logica fuzzy

---

[1] *Il presente articolo, pubblicato sul n. 2-2016 della Rivista Science&Philosophy, è stato gentilmente concesso dall'Editore della Rivista per la divulgazione didattica dei contenuti.*

1. Un percorso sterrato…

Il martedì non si dà principio all'arte. E neppure il venerdì; se diciassette, poi, è possibile si manifesti l'eptacaidecafobia (ovvero, la 'paura' del numero '17'): è buona prassi, a questo punto, chiudersi in casa e attendere il termine dell'infelice combinazione numerico-temporale, nella speranza di scongiurare infausti avvenimenti.
Quando, da dove e perché nascono i nostri convincimenti irrazionali?
Ci viene in aiuto, nella risposta, il concetto di *thauma* nell'accezione aristotelica, che vede alternarsi un significato primario di *sgomento*, di fronte al divenire di tutte le cose con quello di *stupore* e *meraviglia*, davanti ad un mondo da scoprire. Il divenire a cui si fa riferimento è quello di cui l'uomo non può conoscere l'evoluzione a priori, la cui imprevedibilità genera il significato di *thauma* in *fobos* (paura). La paura, dunque, suggerisce Aristotele, costituisce il principale stimolo degli individui al pensiero filosofico e all'azione per la ricerca della *aletheia* (rivelazione, verità) rispetto all'incerto.
Scriverà più avanti Blaise Pascal nel Pensiero n 72, 1669-70 (edizione Brunschvicg):
La nostra condizione ci rende incapaci sia di conoscere con piena certezza come di ignorare in maniera assoluta. Noi voghiamo in un vasto mare, sospinti da un estremo all'altro, sempre incerti e fluttuanti. Ogni termine al quale pensiamo di ormeggiarci e di fissarci vacilla e ci lascia; e, se lo seguiamo, ci si sottrae, scorre via e fugge in un'eterna fuga.
Nulla si ferma per noi. È questo lo stato che ci è naturale e che, tuttavia, è più contrario alle nostre inclinazioni. Noi bruciamo dal desiderio di trovare un assetto stabile e un'ultima base sicura per edificarci una torre che s'innalzi all'infinito; ma ogni nostro fondamento scricchiola, e la terra si apre sino agli abissi. Non cerchiamo, dunque, né sicurezza, né stabilità.
La nostra ragione è sempre delusa dalla mutevolezza delle apparenze; nulla può fissare il finito tra i due infiniti che lo racchiudono e lo fuggono.
La proiezione delle ansie in un mondo irrazionale, sin dalle civiltà pre-elleniche dell'area egea e orientale, condurrà alla ricerca di verità, metafisiche e fisiche, che facciano da guida lungo un cammino teso verso risposte all'apparenza rassicuranti rispetto all'ignoto.
<<Cosa c'entra quello che sappiamo o non sappiamo con le leggi che governano il mondo? La domanda è legittima, e la risposta è sottile>>, sentenzia Carlo Rovelli in 'Sette brevi lezioni di fisica'.
L'ignoto, l'incertezza degli eventi, la variabilità degli stessi e delle relative osservazioni, hanno accompagnato l'uomo in un percorso che attraversa la storia e l'evoluzione del suo pensiero rivolto al bisogno ancestrale di governare un potere misterioso, un potere in grado di favorire un evento positivo o di scongiurarne uno negativo. La natura finita dell'uomo non lo rende capace di comprendere le cause dei fenomeni, non almeno nell'immediato e non nella loro interezza. Ed ecco così sovrapporsi strati culturali di credenze, riti, magie, predizioni, forme e convinzioni religiose (in un linguaggio

variabile nel tempo e nello spazio). Accertata l'attendibilità delle più accreditate ipotesi etimologiche, emerge come perfino il nome della città di Benevento (in origine *Maloenton*) che ospita l'Ateneo sannita non si sottragga all'influenza della pratica di riti propiziatori, riflesso della volontà, da parte dell'uomo, di imbrigliare l'ignoto, con i suoi rischi legati all'incertezza, ed ingraziare la sorte. Questa sconosciuta.

Ebbene, se questa è la nostra condizione, possiamo in qualche modo venire a patti con essa e governarla. Se non si può effettuare una scelta senza sottomettersi all'incertezza, non vi è altra possibilità che affrontarla. Le divinazioni, le predizioni lastricano la strada che conduce alle *previsioni* o, forse più correttamente, ai *tentativi* di previsione.

Ci interroghiamo sul futuro e cerchiamo nei fenomeni naturali (la cui misura rappresenta il 'dato') i segni di ciò che ci aspetta. Il passaggio di testimone dalla pratica divinatoria allo studio di previsione rigoroso e razionale, avviene intorno al secolo XVII. In un clima di fermento scientifico e filosofico Bacone (1561-1626), con *Novum Organum,* nel 1620 da "il La" ad un metodo sperimentale basato sull'osservazione sistematica dei fenomeni finalizzata all'estrapolazione dei concetti generali che spieghino la natura. Con Galileo (1564-1642) si fa strada la sperimentazione basata, però, su ipotesi di lavoro, e tradotta in leggi matematiche 'universalmente' valide. Una nuova e più completa visione dell'universo fa capolino nel '*Discorso sul metodo*' di René Descartes (1596-1690). È il momento in cui la superstizione cede il posto alla ragione, unico 'potere' dato all'uomo per distinguere il vero dal falso e giudicare il bene. La verità scientifica si costruisce sulle evidenze. Fioriscono le scienze teoriche, separate per la prima volta dalle arti meccaniche, con conseguente istituzione di nuove cattedre. In Italia nascono l'Accademia Nazionale dei Lincei, a Roma (1603), e l'Accademia del Cimento a Firenze (1657); a Londra, nel 1660, la Royal Society, a Parigi l'Académie des Sciences (1666) dove non troveranno posto le arti, la metafisica, la teologia e la morale.

Tra i più rilevanti prodotti della rivoluzione delle scienze, il *calcolo delle probabilità*, nato per mano di Blaise Pascal (1623-1662) e Pierre de Fermat (1601-1665), più tardi rivisto da Christian Huygens (1629-1695), è di certo quello che tra magia, filosofia, religione e logica, come una barca sospinta da venti di tempesta, riprendendo l'antica metafora, ancora annaspa tra la casualità, la causalità e l'accadimento degli eventi, tra la loro prevedibilità e l'esperienza soggettiva. Questa, vincolata alla nostra natura finita ancora combatte contro la paura dell'ignoto e siamo solo "*Nel mezzo del cammin*".

2. La probabilità e *cet instinct obscur*

Che cosa significa decidere?

Al di là delle complesse definizioni forniteci dalla letteratura, decidere significa scegliere l'alternativa ritenuta più favorevole tra le opzioni possibili.

Il punto è proprio questo: a volte due alternative sembrano ugualmente favorevoli, o ugualmente sfavorevoli, oppure una delle opzioni ha alcune conseguenze positive e altre negative, o, ancora, una scelta comporta un grande cambiamento.
In alcune situazioni decidere non è semplice: significa mettersi in gioco, fare una scelta rinunciando a tutte le altre possibilità, imparare a capire cosa per noi è davvero importante in quello specifico momento della nostra vita. Nel decidere un investimento economico o il luogo dove trascorrere le vacanze, nello stipulare un mutuo o nel fornire un parere professionale, ci affidiamo soprattutto all'intuizione o, più semplicemente, al *buon senso*. Tale buon senso è il frutto della nostra esperienza passata che ci guida, attraverso innumerevoli scorciatoie, nel risolvere problemi in tempi brevi e, molto spesso, con informazioni insufficienti.
Il decisore, purtroppo, non è in grado di considerare simultaneamente l’insieme delle informazioni disponibili, le alternative, le conseguenze; egli deve fare una selezione e scegliere quelle che ritiene importanti, in quel momento, al fine del raggiungimento del suo obiettivo.
La qualità della decisione dipende, quindi, dalla qualità con la quale tali informazioni vengono elaborate e, più precisamente dalla quantità, dalla completezza e dell'affidabilità delle informazioni raccolte e dalla capacità di elaborazione razionale di tali informazioni. Purtroppo la qualità e la quantità delle informazioni, al crescere della complessità dei problemi, sono spesso insoddisfacenti; nonostante i sistemi di automazione e l'informatica, che rendono presente "ora e subito" qualsiasi tipo di informazione, sembra che i dati non bastino mai per decidere. La capacità di scegliere, tra tutte le informazioni disponibili, le informazioni giuste per decidere in tempi brevi è sempre di più una capacità critica del buon decisore.
L’incerto permea la nostra vita.
Siamo incerti sui risultati del nostro operare e, a maggior ragione, su quegli eventi che non dipendono dal nostro operare. Siamo incerti sulle ipotesi di cui tener conto quando effettuiamo le nostre scelte, in quanto esse dipendono da una analisi degli accadimenti passati e da come riteniamo che essi influenzino il futuro. Siamo incerti sui valori di grandezze fisiche, sia prima di aver effettuato le misure che dopo. E anche quando gli eventi del futuro ci appaiono certi è spesso solo perché non ci preoccupiamo dei dettagli. Sicuramente il Sole tramonterà domani, ma è meno sicuro il minuto esatto in cui ciò avverrà. Tuttavia nessuno di noi è normalmente interessato a tale precisione. Esattezze del genere non esistono nemmeno nella fisica, considerata normalmente la regina delle “scienze esatte”. E proprio il fisico per antonomasia dell’immaginario collettivo, Albert Einstein, scriveva sull’argomento, sostenendo che:
“*As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.*” (“Quando le leggi della matematica si riferiscono alla realtà, non

sono certe, e quando sono certe, non si riferiscono alla realtà.").[2]
Prendere decisioni è un impulso vitale in generale dell'individuo che si lascia, molto spesso, guidare dalla probabilità anziché basarsi sul primordiale flight or fight.
Il calcolo delle probabilità, come teoria matematica, nasce per tentare di affrontare l'incertezza nel modo più razionale possibile. Le sue origini non sono particolarmente nobili, dal momento che si collocano tra i tavoli dei giochi d'azzardo, di dadi o di carte, che di questa nuova teoria scientifica costituirono il primo laboratorio. Ma già in questa prima fase il legame della probabilità col problema della scelta di fronte all'incertezza è subito evidente: le questioni da cui presero le mosse Galileo, Pascal, Fermat per fondare il calcolo delle probabilità, riguardavano proprio quale fosse il comportamento più adeguato che i giocatori avrebbero dovuto tenere nei diversi giochi d'azzardo considerati.
Si può dire che la probabilità costituisce uno dei settori più vivaci della ricerca matematica contemporanea.
Essa, come misura di eventi aleatori, trova applicazioni vastissime, è diventata una componente essenziale della meccanica statistica e della genetica di popolazioni, e nelle scienze del comportamento ha ispirato modelli di scelta razionale in condizioni di incertezza (von Neumann e Morgenstern, 1944). Esistono diverse impostazioni della probabilità.
La prima definizione di probabilità, dovuta a Laplace, secondo la concezione classica, considera la probabilità *P(E)* di un evento *E* uguale al rapporto $\frac{m}{n}$, tra il numero *m* dei casi favorevoli al verificarsi di *E* e il numero *n* dei casi possibili, quando tutti i casi siano giudicati ugualmente possibili:

$$P(E) = \frac{m}{n}$$

Il numeratore *m* è minore, o al più uguale, al denominatore *n*, ne segue che la probabilità è un numero compreso tra 0 e 1:
- se $m = 0$, ossia non vi sono casi favorevoli al verificarsi dell'evento *E*, la probabilità $P(E) = 0$, è quindi nulla;
- se $m = n$, ossia tutti i casi sono favorevoli al verificarsi di *E*, la probabilità $P(E) = 1$.
La definizione classica consente di calcolare effettivamente la probabilità in molte situazioni. Inoltre, è una definizione operativa e fornisce quindi un metodo per il calcolo. La sua caratteristica essenziale è la condizione che tutti i casi in cui il fenomeno può manifestarsi siano giudicati ugualmente possibili, cioè che abbiano la stessa probabilità di verificarsi, quindi nella definizione si usa lo stesso concetto che si vuole definire.
Per superare la circolarità della definizione è stato introdotto il principio della ragione insufficiente, detto successivamente principio di indifferenza, che asserisce che "in

[2] Albert Einstein (27 gennaio 1921) Geometrie und Erfahrung, versione estesa di una lezione tenuta presso l'Accademia Prussiana delle Scienze di Berlino.

mancanza di ragioni che permettano di assegnare probabilità diverse a ciascuno degli eventi elementari, questi devono essere considerati ugualmente possibili".
Ad esempio, se non vi sono ragioni per affermare che un dado sia irregolare, si accetta che ogni faccia si possa presentare con la stessa probabilità. Presenta quindi i seguenti aspetti negativi, non irrilevanti: si applica soltanto a fenomeni con risultati equiprobabili; presuppone un numero finito di risultati possibili; la definizione è circolare perché utilizza la nozione di probabilità (eventi equiprobabili) per definire la probabilità stessa.
Nell'Ottocento, per poter calcolare la probabilità di eventi aleatori, che la definizione classica non consentiva di valutare, si sviluppa l'impostazione *frequentista*, o *statistica*. Essa si applica quando sull'evento si possono eseguire, nelle stesse condizioni, tante prove quante si vogliono oppure quando di un fenomeno sono disponibili tavole di rilevazioni statistiche (come le tavole di mortalità e di sopravvivenza di una popolazione).
Secondo questa impostazione, per conoscere la probabilità di un evento si deve ricorrere all'esperimento.
È importante notare che, per un cultore di questa impostazione, non ha senso calcolare la probabilità di un evento mediante una singola prova, ma, eseguendo un grande numero di prove, si riscontra una regolarità che permette di assegnare una valutazione di probabilità che l'evento si verifichi.
L'impostazione frequentista è basata sulla definizione di frequenza relativa di un evento.
Si definisce frequenza relativa di un evento in n prove, effettuate tutte nelle stesse condizioni, il rapporto $\frac{k}{n}$, tra il numero *k* delle prove nelle quali l'evento si è verificato e il numero *n* delle prove effettuate:

$$f = \frac{k}{n} \qquad con\ 0 \leq f \leq 1.$$

Anche la frequenza è compresa tra 0 e 1, ma occorre notare che:
- se $f = 0$, non si può dire che l'evento è impossibile, ma che non si è mai verificato in quelle prove;
- se $f = 1$, non si può dire che l'evento è certo, ma che, in quelle prove, si è sempre verificato.

Si deve osservare che il valore di f non dipende solo dal numero n delle prove fatte, ma per uno stesso n può variare al variare delle prove.
Ad esempio, lanciando 100 volte una moneta si presenta testa per 55 volte. Effettuando altri 100 lanci la faccia testa generalmente si presenta un numero diverso di volte, ad esempio 48. Pertanto la frequenza relativa per il primo gruppo di lanci è 55/100 , mentre per il secondo gruppo è 48/100.
È interessante constatare che, pur variando secondo il gruppo di prove, la frequenza $\frac{k}{n}$ al crescere di n tende a stabilizzarsi intorno a un valore. Questo fatto era già stato rilevato dai demografi del XVII e XVIII secolo nello studio della frequenza di decessi e delle nascite maschili in una popolazione.

Importante è sottolineare la seguente comparazione: per fenomeni in cui si può calcolare la probabilità ricorrendo alla concezione classica, si nota che, all'aumentare del numero delle prove, la frequenza tende generalmente ad avvicinarsi alla probabilità calcolata a priori.
Sono, infatti, di importanza storica gli esperimenti di Buffon e Pearson sul lancio di una moneta, i quali portarono i due studiosi a enunciare, per eventi in cui la probabilità può essere calcolata secondo la concezione classica, la cosiddetta legge empirica del caso: "In una serie di prove ripetute un gran numero di volte, eseguite tutte nelle stesse condizioni, la frequenza tende ad assumere valori prossimi alla probabilità dell'evento e, generalmente, l'approssimazione è tanto maggiore quanto più grande è il numero delle prove eseguite."
La legge empirica del caso permette di formulare la seguente definizione frequentista di probabilità per eventi ripetibili: "La probabilità di un evento è uguale alla frequenza relativa in un numero di prove ritenuto "sufficientemente elevato".
Il campo di applicazione dell'impostazione frequentista è molto vasto perché la definizione può essere applicata a fenomeni di cui si posseggano dati statistici rilevati in passato in condizioni analoghe.
In generale, tutto il campo delle assicurazioni è basato su questa concezione di probabilità. Altre notevoli applicazioni riguardano la medicina, l'economia, la meccanica quantistica e, in generale, tutte le scienze per le quali si possono utilizzare metodi statistici. In particolare, il naturalista Gregor Mendel (1822–1884), con i suoi esperimenti sullo studio della riproduzione dei piselli odorosi (*Lathyrus odoratus* L.), elaborò le prime leggi sulla trasmissione di caratteri ereditari mediante il calcolo della probabilità. Non trascurabile, per importanza ed attualità, l'applicazione delle probabilità alla gestione del rischio.
È importante osservare che la probabilità nella concezione classica è un numero determinato *a priori*, invece nell'impostazione frequentista è un numero calcolato *a posteriori* sulla base di una raccolta dati e di un'elaborazione statistica degli stessi.
Entrambi i metodi di valutazione della probabilità sono detti oggettivi, per distinguerli dalla probabilità secondo l'impostazione soggettivista.
L'impostazione *assiomatica* della probabilità venne proposta da Andrey Nikolaevich Kolmogorov nel 1933 in Grundbegriffe der Wahrscheinlichkeitsrechnung. Va notato che la definizione assiomatica non è una definizione operativa e non fornisce indicazioni su come calcolare la probabilità.
Si assume che ogni evento nello spazio campionario Ω sia associato a un numero reale P(E), chiamato probabilità di E. Questo numero soddisfa le tre seguenti condizioni:

- la probabilità è un numero non negativo: $P(E) \geq 0$;
- la probabilità dell'evento certo è unitaria: $P(\Omega)= 1$;
- dati due eventi A e B mutuamente esclusivi, allora $P(AUB) = P(A)+ P(B)$.

Si osserva che, come conseguenza degli assiomi precedenti, necessariamente, $P(E) \leq 1$.
I tre assiomi introdotti da Kolmogorov sono coerenti con la definizione empirica fornita da Von Mises e con la definizione classica enunciata da Laplace.

Le impostazioni precedenti della probabilità a cui abbiamo accennato mostrano dei limiti di applicabilità come può essere mostrato considerando i seguenti quesiti.
Qual è la probabilità per un diplomato di trovare un'occupazione entro un anno?
Qual è la probabilità che la squadra di calcio X vinca il campionato?
Qual è la probabilità che entro l'anno 2018 sia scoperto un vaccino per la cura di una grave malattia?
Per eventi come quelli indicati non è possibile valutare la probabilità né secondo la concezione classica, perché non si possono determinare i casi possibili e i casi favorevoli, né secondo l'impostazione frequentista, perché *gli eventi non sono ripetibili*.
In questi casi si può dare una stima della probabilità legata allo stato di informazione secondo la seguente definizione di probabilità nell'impostazione *soggettivista*.
La probabilità *P(E)* di un evento *E* è la misura del *grado di fiducia* che un individuo coerente attribuisce, in base alle proprie informazioni e alle proprie opinioni, al verificarsi dell'evento *E*.
Queste valutazioni di probabilità sono soggettive, ossia possono variare da individuo a individuo, ma è importante che sia rispettata la coerenza.
Per fissare il concetto di coerenza, Bruno de Finetti (1906–1985), uno dei principali attori di questa rivoluzione del pensiero, già in '*Probabilismo*' (1931), afferma in maniera provocatoria che la probabilità non esiste. Almeno non nell'accezione classica.
Egli si ricollega alle scommesse e scrive: "La probabilità di un evento E, secondo l'opinione di un certo individuo, è il prezzo p che egli ritiene equo attribuire all'importo unitario esigibile al verificarsi di E."
La probabilità è quella somma p che l'individuo è disposto a pagare per ricevere €1 nel caso che si verifichi E, ma, per coerenza, egli è disposto ad accettare la scommessa inversa, ossia ricevere p e pagare €1 al verificarsi di E.
In questa impostazione la probabilità è un numero reale compreso tra 0 e 1. Infatti se l'evento è giudicato impossibile il prezzo $p = 0$, se l'evento è giudicato certo il prezzo è $p = 1$ e negli altri casi risulta $0 < p < 1$.
Il campo di applicazione dell'impostazione soggettivista è, praticamente, illimitato; quando prendiamo una qualsiasi decisione, anche inconsciamente, attribuiamo alle varie alternative una certa probabilità. Naturalmente persone diverse, anche in possesso della stessa informazione, possono prendere decisioni diverse secondo la differente opinione di ciascuno.

3. La statistica: dal dato al modello

Un ponte tra la 'realtà che percepiamo' e la 'realtà che potrebbe essere'; il caso del rischio sismico.
Tra il *dato*, che rappresenta il fatto o la realtà che siamo capaci di percepire e misurare, e il *modello*, ossia la realtà come potrebbe essere nel suo insieme, fa da collegamento la statistica. Essa, infatti, lega due aspetti fondamentali che caratterizzano la conoscenza

del sapere umano: l'osservazione (ovvero l'esperienza) e l'astrazione matematica (esercizio della ragione).
Nel mondo della statistica il dato e il modello si confrontano in un contesto reale, caratterizzato dalla 'variabilità individuale' degli elementi costituenti un *insieme* osservato (tale variabilità rappresenta le differenze tra gli elementi del *tutto* rispetto a uno o più caratteri).
Dal momento che l'oggetto della materia di studio è rappresentato dalle proprietà dell'insieme, il suo focus si sposta necessariamente verso la ricerca di 'invarianti', o 'costanti caratteristiche' (valori medi, misure di disuguaglianza, indicatori della forma distributiva, parametri di relazione tra variabili), non trascurando il fatto che ogni costante statistica rappresenta una sola proprietà dell'*insieme*: quella analizzata.
A dare la misura e il senso delle caratteristiche di un insieme è una molteplicità di indicatori, ad esempio, un valore medio si deve sempre accompagnare al grado di allontanamento dei valori mediati da esso (la variabilità). Inoltre, lo scarto tra le situazioni individuali di una popolazione e la media per essa calcolata esprime la distanza tra distribuzione reale e distribuzione modellata, così, una media aritmetica delle differenze assolute tra i singoli valori e il loro valor medio offre una misura della disuguaglianza.
Nei diversi campi del sapere, e non, è in crescendo il confronto con i metodi e il linguaggio statistico: la capacità di raffrontare la descrizione (o misura) di un fenomeno con una stima della sua attendibilità è di fondamentale importanza per il metodo con il quale, ad esempio, la scienza trae conclusioni mirate al più alto grado di affidabilità permesso, partendo dalle regolarità sperimentali.
Marco Li Calzi, ne "La matematica dell'incertezza" (2016) scrive: *"Proviamo a immaginare il mondo come una stanza chiusa a chiave e malamente illuminata, della quale cerchiamo di indovinare l'arredamento sbirciando dal buco della serratura. In fase esplorativa, ciò che intravediamo suggerisce le ipotesi di lavoro."*
Fa capolino, ancora una volta, il concetto di incertezza che, come visto, rappresenta l'elemento distintivo tanto dell'individuo quanto della specie. Un concetto, che attraversa più Paesi, e parla lingue diverse, tutte, però, convergenti verso un termine: 'dado'.
Nella gestione dei rischi, in generale, dei rischi naturali, in particolare, a traghettarci nell'insidioso mondo dell'incertezza è il concetto di 'pericolosità'.
Prendiamo, ad esempio, il significato di 'pericolosità' nella nozione di rischio sismico.
Gli approcci di studio possono essere mirati a una *predizione deterministica* e/o a una *previsione probabilistica* dei terremoti, ossia: ad una dichiarazione deterministica per la quale un terremoto futuro accadrà (o non accadrà) in una particolare regione geografica, in una certa finestra temporale ed entro un intervallo di magnitudo, oppure ad una probabilità che tale evento accadrà.
Il primo caso rappresenta, di fatto, l'optimum teorico, ma la natura trova sempre il modo per sfuggire a leggi lineari, così, le complesse geometrie delle sorgenti sismogenetiche, la natura caotica dei processi di rottura, le variazioni in scala delle forze

che agiscono sulle faglie e l'interazione di queste con ulteriori forze dovute a cause differenti, rappresentano soltanto alcuni dei fenomeni che conferiscono al metodo deterministico un carattere di difficile attuazione. Allo stato attuale, il sapere scientifico non conduce ad una predizione affidabile dei terremoti.
Qualsiasi informazione sul verificarsi di un sisma reca in sé grandi incertezze, pertanto tale informazione può essere valutata solo in termini di probabilità. Ci si affida, così, alla previsione probabilistica, la quale mira a quantificare l'informazione circa l'eventuale accadimento di terremoti futuri.
Bisognerà, dunque, partire dal dato (o esperienza) e il dato, in questo caso, altro non è che individuazione (con relativa delimitazione) delle aree a comportamento omogeneo dal punto di vista della sismicità (zone sismogenetiche); ciò sarà possibile soltanto se saranno note le distribuzioni spaziali e temporali dei terremoti (grazie all'esame dei cataloghi storici), le geometrie e le caratteristiche delle strutture geologiche (faglie) superficiali e profonde, e i movimenti recenti che queste hanno accomodato.
Ma quando entra in gioco la statistica? Essa interviene, appunto, nel concetto di 'sismicità', che rappresenta la storia sismica di un territorio.
Nel 1954 Gutenberg e Richter proposero una relazione statistica che lega il numero di eventi sismici, in un intervallo di tempo prefissato, all'intensità (o magnitudo) registrata, tale relazione è nota come legge di occorrenza o legge di Gutenberg-Richter ed è espressa come segue:$\log N(I)=a-bI$
dove:
N è il numero di eventi; I l'intensità macrosismica; a e b rappresentano costanti dipendenti dal territorio considerato; b, che è una caratteristica della sorgente sismogenetica, indica la pendenza della retta che rappresenta graficamente la legge.
La legge chiarisce come a frequenza maggiore si verifichino eventi di bassa magnitudo, al contrario di quelli aventi un periodo di ritorno breve.
Definita la sismicità attraverso la frequenza e l'energia con cui i terremoti si verificano in una data area e attribuendo un valore di probabilità del verificarsi di un evento tellurico che superi una soglia di intensità di nostro interesse, si ottiene, seguendo il criterio di Cornell, e calcolando, quindi, gli effetti provocati in relazione alla distanza dalla sorgente, la *pericolosità sismica.*
La pericolosità è uno dei principali fattori del *rischio sismico*, definito come la misura dei danni attesi, in un intervallo di tempo, per un'area caratterizzata da una nota sismicità e da un determinato grado di resistenza delle costruzioni e di antropizzazione. O, più semplicemente, come l'esito il prodotto della combinazione tra la pericolosità, la vulnerabilità e l'esposizione.
Carl Nilsson Linnaeus (1707-1778) affermava: *<<Nomina si nescis, perit et cognitio rerum>>* **(***Se non conosci i nomi, muore anche la conoscenza delle cose*), qualche anno più tardi in 'Palombella rossa' (1984) un furioso Nanni Moretti, inveendo contro un'incredula e sprovveduta giornalista, poco avvezza alla forma, urlava: *<<le parole sono importanti!>>.*

A tal proposito, si menziona di seguito un interessante intervento sulla prevenzione del rischio sismico a cura di Emanuela Guidoboni, la quale con un originale approccio al problema, si sofferma sull'analisi dell'etimo, giocando con le lingue francese, inglese e italiana, e giungendo ad una curiosa coincidenza.
Guidoboni, inizia la sua analisi osservando che il termine A*léa sismique,* in francese, conserva il latino *ālĕa [aleă], aleae,* sostantivo femminile I declinazione, *dado*. In sanscrito (chiave di volta nella comprensione delle variazioni fonetiche delle parole nelle lingue indoeuropee, in relazione alle fasi più antiche della storia d'Europa e del vicino Oriente) il termine *aksah* associa al significato di *dado* quello di *sorte,* e nel medio francese compare, infatti, *hazard* da *zahr* (f. voce arcaica), che traduce, appunto, *dado*.
Il passaggio all'inglese, dal quale paradossalmente importiamo il termine nell'uso comune, avviene in maniera naturale: *Seismic hazard* (dal medio francese).
In italiano, invece, ci si riferisce a due termini: il primo è *açardum* (di origine latino medievale), riferito al XIII secolo che traduce azzardo, qui ritorna *a-zahr,* tradotto nel 1665 in azzardoso; il secondo termine di riferimento è *periculum* che si traduce in tentativo, prova oppure in rischio, pericolo.
Ed ecco che l'architettura del discorso e dell'itinerario qui proposti collassa su un dado, leitmotiv tra le righe di un diario di viaggio ancora aperto, il cui punto di arrivo della prima tappa coincide con il punto di partenza, ossia: l'angoscia di non sapere, il dubbio di una sorte assoggettata ad un dio ludopatico e la speranza che il nostro 'istinto oscuro' ci conduca ad una verosimile definizione del rapporto tra casualità e causalità.

4. Logica fuzzy e vaghezza

La logica fuzzy, spesso confusa con la probabilità, vuole esprimere e formalizzare tutte le frasi che non sono del tutto vere o del tutto false; Zadeh, fondatore della logica fuzzy, spiega in sintesi che " everything is a matter of degree", ponendo attenzione alle sfumature.
L'idea che porterà poi alla nascita della logica fuzzy è stata introdotta da Jan Łukasewicz in "On determinism" che nel 1946 affermò:"I am entitled not to recognize the principle of bivalence, and to accept the view that besides truth and falsehood exist other truth-values, including at least more, the third truth-value"; proponendo di non restare ancorati alla logica classica aristotelica, ma di considerare anche quei gradi di verità che non sono assoluti. Partendo da questa idea, ha avuto origine una pletora di formalizzazioni e le idee di logiche non-standard con speculazioni pratiche e filosofiche, tutte racchiuse sotto il nome di logiche fuzzy.
Si noti che la logica fuzzy fu principalmente motivata dalla necessità di descrivere, in forma formale e computabile, le conoscenze formulate da esperti in termini vaghi, utilizzando il linguaggio naturale. Questa motivazione può essere chiarita con un esempio fornito da Kreinovich: quando si parla ad un medico, è molto probabile che non dirà "se una cisti è superiore a 7 mm di diametro e il suo colore è a 500 nm,

assumere 250 mg di una medicina" ma piuttosto "se una cisti è abbastanza grande e il suo colore è rossastro, allora somministrare al paziente una piccola dose di una qualche medicina". Oggi, con la sua poliedricità di funzioni, la logica fuzzy ha ormai assunto un ruolo fondamentale e cruciale nello sviluppo dell'intelligenza artificiale ed in altre applicazioni, come ad esempio i sistemi di controllo in ingegneria, l'elaborazione di immagini, riconoscimento fisiognomico, la diagnosi medica, ottimizzazione, analisi delle politiche pubbliche, analisi di mercato e molte altre.

Bibliografia

Baldi P., (1998), Calcolo delle probabilità e statistica, McGraw Hill Education

E. R. Caianiello, P. E. Eklund, M. Squillante, A. G. S. Ventre, (1989) Formalism and Implementations of C-calculus, Computational Intelligence, A. Martelli, G. Valle (Editors), Elsevier Science Publishers B.V. (North-Holland), 15- 26

Canova P., Rizzuto D., (2016), "Fate il nostro gioco". Add. Ed. Torino

De Finetti B., (1931), Probabilismo, Libreria Editrice Franco Perrella S.A., Napoli-Città Castello

De Finetti B., (1967), Il "saper vedere" in matematica, Loescher

De Finetti B., (1970), Teoria delle probabilità, Einaudi

Li Calzi M., (2016), "La matematica dell'incertezza", il Mulino Editore

Łukasewicz J, (1946), On Determinizmie, trad. ingl. "On Determinism"

Maldonato M., (2015), Quando decidiamo. Siamo attori consapevoli o macchine biologiche?, Giunti Editore

Manara C.F. (1997), Matematica e incertezza, Milano, KOS

Monari P., (2012), Giochi d'azzardo e probabilità, Editori Riuniti

Nocenzi M., (2002), Vivere l'incertezza: sociologia, politica e cultura del rischio ambientale nelle insicurezze da inquinamento elettromagnetico, Franco Angeli Editore

Rovelli C., (2014), Sette brevi lezioni di fisica, Adelphi Editore

Scozzafava R.,(2001), Incertezza e probabilità, Zanichelli

Squillante M., Ventre, A.G.S. (1992) Generating Fuzzy Measures, Journal of Mathematical Analysis and Applications, 165, 2, 550-555

Vergineo G.,(1985), Storia di Benevento e dintorni, Benevento, Gennaro Ricolo Editore

Weisse Neil A., (2008), Calcolo delle probabilità, Pearson.

Un approccio didattico inclusivo allo studio della Probabilità e della Statistica nella scuola secondaria di II grado

Raffaele Prosperi[1]

[1]Università "Federico II" di Napoli, Dipartimento di Architettura,
E-mail: raffaele.prosperi@unina.it

Sunto
Sulla base delle varie indicazioni del MIUR e degli esperti di didattica, si propone un percorso per l'acquisizione di competenze relative alla Matematica dell'Incerto per incrementare il grado di coinvolgimento degli alunni e permettere anche gli alunni con difficoltà di apprendimento di acquisire abilità e competenze sull'argomento. In particolare, la proposta mira a far acquisire agli alunni coinvolti un toolkit di strumenti da utilizzare con cognizione di causa nella risoluzione di problemi reali, situati nella sfera di competenze e gradimento dell'allievo. L'approccio consigliato è per episodi di apprendimento situato. Si presenta, inoltre, una proposta didattica formulata prevedendo il coinvolgimento degli alunni in casi e situazioni reali e di progressivo approfondimento e curando l'ottimizzazione del tempo disponibile.

Parole chiave: didattica, EAS, inclusione, probabilità, statistica.

1. Introduzione.

L'analisi dell'attività didattica relativa all'insegnamento del Calcolo delle probabilità e della Statistica nella scuola superiore parte dalle Indicazioni Ministeriali e dalle Linee Guida al fine di identificare gli obiettivi posti dal Ministero dell'Istruzione in relazione alla Matematica dell'incerto. Le Indicazioni Ministeriali richiedono, nella parte generale, che, a causa dell' "*ampio spettro di contenuti affrontati … l'insegnante sia consapevole della necessità di un buon impiego del tempo disponibile*", con chiaro invito a non perseguire sulla linea (preferita dalla maggior parte dei docenti) della concentrazione del tempo su specifici argomenti, quasi sempre solo teorici o con applicazioni formali, generiche e avulse dai contesti della vita reale; anzi, si invita ad "*evitare dispersioni in tecnicismi ripetitivi o casistiche sterili*" che non contribuiscono in modo significativo al potenziamento delle competenze per la risoluzione dei problemi.
Al fine di ottenere un buon impiego del tempo disponibile si sottolinea la possibilità di utilizzare strumenti informatici idonei per rappresentare e manipolare oggetti

matematici; ciò favorisce il loro uso anche per il trattamento dei dati nelle altre discipline scientifiche.
L'indicazione principale è: "*pochi concetti e metodi fondamentali acquisiti in profondità*".
Nel caso specifico della Matematica dell'incerto, nell'area "Dati e previsioni" le Indicazioni Ministeriali stabiliscono che "*Lo studente dovrà essere in grado di rappresentare e analizzare in diversi modi (in particolare utilizzando strumenti informatici) un insieme di dati, scegliendo le rappresentazioni più idonee.*
Lo studente dovrà esser in grado di utilizzare strumenti di calcolo (calcolatrice, foglio di calcolo) per studiare raccolte di dati e serie statistiche.
Dovrà quindi saper distinguere tra caratteri qualitativi, quantitativi discreti e quantitativi continui, lavorare con distribuzioni di frequenze e rappresentarle.
Lo studente dovrà essere in grado di ricavare semplici inferenze dai diagrammi statistici.
Sarà introdotta la nozione di probabilità, con esempi entro un contesto classico e con l'introduzione di nozioni di statistica.
Sarà introdotto in modo rigoroso e approfondito il concetto di modello matematico, distinguendone la specificità concettuale e metodica rispetto all'approccio della fisica classica".

2. - Osservazioni.

La riflessione su alcuni punti cruciali delle Indicazioni Nazionali e sulle indicazioni didattiche sugli alunni con B.E.S. e con D.S.A. porta a consigliare una progettazione didattica della disciplina di tipo inclusivo, personalizzato per tener conto delle diversità, per incremento e potenziamento progressivo di competenze, sia di asse che di cittadinanza; dovrebbero, pertanto, essere applicate le nuove metodologie (flipped lesson, problem solving, scoperta guidata, peer tutoring ecc.) in lavori di gruppo interdisciplinari, dove finalmente la Matematica si avvicinerebbe al mondo reale, abbandonando almeno in parte l'estremo formalismo.
Gli argomenti dell'area Dati e Previsioni vengono molto spesso tralasciati o affrontati in modo approssimativo. La motivazione più frequente è legata alla necessità di escludere gli argomenti ritenuti meno importanti al fine di sfruttare il ridotto tempo a disposizione per svolgere quelli considerati di fondamentale importanza. Tale punto di vista è legato anche, in prospettiva, alle esigenze dei docenti del secondo biennio e soprattutto del quinto anno, che richiedono particolari conoscenze (più raramente competenze e abilità) per sviluppare quasi esclusivamente le parti di analisi e geometria analitica. Ciò può avere una qualche giustificazione per il Liceo Scientifico. Infatti, per come è strutturata la prova scritta di Matematica all'esame di stato, analisi e geometria analitica da sole garantiscono la possibilità di ottenere il massimo punteggio, per cui in tale ottica è possibile, anzi preferibile, tralasciare tutto il resto; solo negli ultimi problemi c'è la

necessità di calcolare un valore di probabilità, ma è sempre relativo ad una banale applicazione della definizione classica ad un rapporto tra aree o volumi. I quesiti di Probabilità e Statistica sono alternativi ad altri di Analisi e Geometria, per cui possono essere tranquillamente trascurati pur potendo ottenere la massima valutazione nella seconda prova scritta. Non ci sono giustificazioni di questo tipo, invece, per gli altri tipi di istituto. Spesso è il risultato di un insufficiente aggiornamento da parte dei docenti, soprattutto sugli argomenti introdotti più di recente nelle indicazioni, come probabilità, statistica, informatica e logica.

L'invito ad evitare dispersioni ripetitive e sterili comporta innanzitutto la necessità di non sprecare il tempo da dedicare a Dati e Previsioni, oltre che naturalmente l'inopportunità di rendere gli alunni dei puri esecutori meccanici di procedimenti di calcolo già prefissati e sempre uguali; i "pochi concetti" e i "metodi fondamentali acquisiti in profondità" raccomandati dalle Indicazioni Nazionali e dalle Linee Guida fanno pensare piuttosto all'acquisizione delle conoscenze dei metodi e degli strumenti di base di ogni settore, dello sviluppo delle relative competenze nel saperli scegliere ed applicare nell'ordine corretto per risolvere un problema: concetti e metodi dovrebbero configurarsi come una sorta di "cassetta degli attrezzi con le relative istruzioni", un *toolkit* da cui saper, caso per caso, attingere gli strumenti necessari e applicarli nel modo e nell'ordine giusto. Sarebbe, inoltre, preferibile non costringere l'alunno a ricordare a memoria formule e procedimenti, in modo che il tempo dedicato all'approfondimento sia speso per acquisire competenze, piuttosto che sterili conoscenze mnemoniche e abilità meccaniche.

Infatti, affermando che lo studente deve essere in grado di "*rappresentare e analizzare in diversi modi insiemi di dati"*, le Indicazioni e le Linee Guida richiedono che egli sia in grado di scegliere di volta in volta lo strumento di rappresentazione più adatto, più efficace per la comprensione e per l'analisi del fenomeno descritto dai dati, implicando un percorso attraverso cui l'alunno prende coscienza dei vantaggi di ogni modello, acquisendo in tal modo la conoscenza utile per leggere e interpretare i possibili diagrammi e la competenza necessaria per selezionare il metodo migliore in ciascuna situazione.

Le frasi "*Metodi fondamentali*", "*Analizzare*", "*Scegliere*" fanno ritenere necessaria solo la capacità di individuare algoritmi risolutivi selezionando formule e metodi e applicandole correttamente secondo indicazioni non mandate a memoria, ma interpretate di volta in volta da appositi "manuali d'uso".

La richiesta dell' uso di "*Strumenti Informatici in modo critico*" invita all'utilizzo di Calcolatrici e Software (come ad esempio il Foglio di Calcolo Elettronico), senza percorsi prestabiliti.

Infine la specifica richiesta di "*Collegamento con altre discipline*" richiede l'individuazione di interdisciplinarietà; questa è fondamentale con le materie scientifiche in cui è richiesto il rilevamento e l'analisi di dati per prove sperimentali (come la chimica e la fisica), ma è importante anche con materie umanistiche, rappresentando un ulteriore metodo di descrizione di realtà non prettamente scientifiche.

In particolare, nelle attività relative all'Alternanza Scuola-Lavoro e ai Progetti Interdisicplinari è richiesto un utilizzo dei concetti matematici nella risoluzione di problemi *reali, aperti e situati.*

3. Competenze disciplinari.

Relativamente all'area "Dati e previsioni", si è proceduto all'analisi delle competenze di base per il primo biennio della scuola superiore, cioè di quelle che si ritengono già acquisite nel corso della scuola primaria e secondaria di primo grado.

3.1 Competenze in ingresso e relative conoscenze e abilità.

Sulla base dei problemi proposti nelle prove Invalsi della scuola secondaria di I grado dal 2008 al 2011 e di quelli OCSE-PISA degli anni 2000, 2003 e 2006, si ricava che le conoscenze richieste in uscita dalla scuola secondaria di I grado e, quindi, di ingresso alla scuola secondaria superiore riguardano:

- Capacità di comprendere le differenze di informazioni delle varie tipologie di grafici di descrizione delle frequenze: diagramma a barre, istogramma semplice e multiplo, cartesiano, a punti, a torta, albero;
- capacità di evincere informazioni presenti nei grafici e di evincerne altre non direttamente riportate;
- capacità di utilizzare le informazioni dei grafici per prender decisioni, confrontare situazioni, effettuare calcoli, individuare leggi matematiche, descrivere percorsi, calcolare medie e probabilità, estrapolare informazioni su percorsi e mappe;
- capacità di discutere la "validità" di un campione e di ipotizzare semplici inferenze.

3.2 Competenze in uscita e relative conoscenze e abilità.

Dall'analisi delle Indicazioni Nazionali e delle Linee Guida, si ricava che, rispetto a conoscenze e abilità già considerate acquisite in ingresso, in uscita del primo biennio della scuola secondaria superiore sia richiesto:

1. *acquisire le conoscenze e le competenze sui modelli di rappresentazione grafica (compreso l'uso di strumenti informatici) per la descrizione di insiemi di dati al fine di utilizzare, secondo il caso, la rappresentazione più idonea;*
2. *conoscere i caratteri delle grandezze statistiche e saper, caso per caso, individuarne il tipo (qualitativo, quantitativo);*
3. *comprendere il significato di distribuzione di frequenza di un dato evento e saperlo rappresentare con la forma più adatta (tabella, istogramma, aerogramma, ecc.);*

4. conoscere in modo approfondito le definizioni e le proprietà delle misure di tendenza centrale (media, moda, mediana) e di variabilità (scarto semplice, varianza, deviazione standard) di una distribuzione di valori e saperne discutere i risultati;
5. conoscere le potenzialità e le peculiarità di strumenti di calcolo (calcolatrice, foglio di calcolo elettronico) ed essere in grado di utilizzarli per rilevare informazioni di sintesi da raccolte di dati e serie statistiche;
6. essere in grado di individuare possibili legami tra le variabili dall'analisi visiva dei diagrammi statistici;
7. individuare eventi probabilistici e loro insiemi; saper discutere i casi di eventi indipendenti e incompatibili;
8. conoscere le definizioni di probabilità classica e frequentista ed essere in grado di applicarle per calcolare la probabilità di eventi, in particolare nel caso classico. Anche se non esplicitamente richiesto nelle Indicazioni Nazionali e nelle Linee Guida, si ritiene ineluttabile fornire anche la definizione soggettivista di probabilità in quanto legata al fenomeno delle scommesse, sempre più diffuso in ambito giovanile;
9. acquisire conoscenze di base sulla statistica inferenziale comprendendo la differenza di obiettivi e di situazione rispetto alla statistica descrittiva.

Allo scopo di ottimizzare il tempo a disposizione, che appare al docente sempre più ridotto rispetto agli anni precedenti, soprattutto in rapporto ad un percepito aumento degli argomenti da trattare, è consigliabile procedere, per la didattica della probabilità e della statistica, in parallelo con gli argomenti della stessa o di altre discipline (come fisica, chimica, scienze e informatica, ove presenti nel curricolo) al fine di utilizzare le definizioni, gli esempi/esercizi e lo studio di situazioni reali come parte dello svolgimento di esercizi/problemi specifici del campo statistico-probabilistico.

4. Proposta didattica.

- La raccolta dei dati potrebbe essere effettuata in parallelo allo studio sperimentale di leggi chimiche o fisiche, ma anche in collaborazione per ricerche relative a materie umanistiche.
- L'introduzione degli eventi può collegarsi direttamente alla teoria degli insiemi e alla logica elementare, in modo da utilizzare direttamente i concetti di insieme, sottoinsieme, unione, intersezione, negazione, prodotto cartesiano; anzi, sarebbe preferibile, durante il percorso di logica e di insiemistica, proporre esercizi già relativi ad eventi probabilistici, evitando tutta una serie di casi sterili e/o fini a se stessi.
- Sarebbe preferibile proporre solo pochi esercizi, da ampliare e completare nel corso dell'anno scolastico ampliandoli con nuovi concetti man mano che questi vengono proposti agli alunni.

Nei paragrafi successivi sono riportate proposte per un approccio agli argomenti fondamentali.

4.1 Media e deviazione standard.

Si propone un percorso didattico per lo svolgimento della parte riguardante gli indici di posizione e di dispersione che prevede il coinvolgimento attivo degli alunni in discussioni guidate sull'analisi delle differenti situazioni che si vengono a creare facendo variare alcuni dati con criteri suggeriti dal docente:

- proporre (o, ancora meglio, far rilevare agli alunni) una sequenza di valori riguardanti un evento di cui sono pienamente a conoscenza. Ad esempio, i voti nelle prove scritte di una materia, l'altezza degli alunni, ecc.;
- proporre una seconda sequenza per lo stesso evento (ad esempio il voto alla successiva prova scritta) e coinvolgere gli alunni in una discussione su come confrontare le due sequenze, fino a condurli alla definizione di media aritmetica semplice; far calcolare le medie; visualizzare i dati con grafici;
- aggiungere ad una delle sequenze precedenti un valore coincidente con la media, mostrando che la media complessiva resta invariata;
- aggiungere alla sequenza un valore inferiore alla media e verificare la variazione del valore della media complessiva. Dedurre la regola corrispondente (ed eventualmente dimostrarla algebricamente). Ripetere con l'aggiunta di un valore superiore alla media;
- alla sequenza originaria aggiungere due valori la cui media coincida con la media della sequenza originaria e mostrare che la media non varia;
- proporre un semplice caso in cui la media non è aritmetica; ad esempio, considerare il caso di interessi in due anni conoscendo il tasso annuo e calcolando gli interessi di ciascun anno una prima volta rispetto all'anno iniziale, una volta ciascuno rispetto all'anno precedente. Definire tale media "geometrica";
- ripetere con esempi che conducano ad una media armonica e ad una quadratica;
- dare la definizione generale di media di n valori;
- costruire e proporre una nuova sequenza che abbia la stessa media della prima ma con valori molto meno distribuiti intorno alla media; confrontarla con la prima sequenza, far notare che nei due casi le medie hanno diverso grado di "rappresentatività" della sequenza e coinvolgere gli alunni nella ricerca di un criterio per misurare tale rappresentatività, fino a condurli alle deviazione standard (scarto quadratico medio), eventualmente passando per lo scarto semplice, per lo scarto assoluto e per quello quadratico (varianza);
- far visualizzare le due distribuzioni e far notare la differenza di concentrazione intorno alle rispettive medie;
- far calcolare e confrontare le deviazioni standard per le due sequenze,

discutendone le differenze.

I calcoli dovrebbero essere effettuati con calcolatrici scientifiche o con software per PC nella parte iniziale di determinazione dei valori di sintesi.
Utilizzare un foglio di calcolo elettronico consente sia di ottenere nuovi risultati rapidamente al variare dei valori, sia di tracciare grafici idonei al confronto tra le diverse situazioni proposte.
Si ritiene opportuno far notare che l'uso della deviazione standard (o scarto quadratico medio) è preferibile a quello della varianza, in quanto, pur fornendo la stessa indicazione sulla dispersione dei dati rispetto al loro valor medio, la deviazione standard ha la stessa unità di misura dei valori cui si riferisce e quindi della loro media, per cui può essere algebricamente sommata o sottratta alla media stessa per ottenere intervalli di confidenza.

5. Probabilità.

Si propone far definire in modo intuitivo agli alunni la probabilità classica e quella frequentista, in quest'ultimo caso eventualmente guidati dal docente. La definizione soggettivista, invece, dovrebbe risultare di semplice comprensione per quel gran numero di alunni che scommette sui risultati degli incontri sportivi.
Per l'applicazione della definizione classica, potrebbe essere necessario introdurre i concetti base della Combinatoria; un possibile approccio che permetta di rendere competenti gli alunni nella determinazione dei valori delle disposizioni, combinazioni e permutazioni è proposto nel paragrafo successivo.

5.1 Analisi combinatoria.

Inizialmente dovrebbe essere sufficiente presentare, anche in modo intuitivo, le disposizioni semplici di n elementi su k posti, ovviamente come numero di differenti gruppi degli n valori distinti a k a k, in cui scambiando di posizione due elementi si ottiene una disposizione diversa (stessi elementi in ordine diverso forniscono una differente disposizione); i valori numerici … successivamente è molto semplice far notare che se n coincide con k si ottengono le permutazioni di k elementi e che se l'ordine degli elementi non modifica il significato, si ottengono le combinazioni. In quest'ultimo caso è anche semplice far notare che, a partire da una singola combinazione di n elementi su k posti si ottengono k! Disposizioni.
Solo in momenti successivi si passerà ai casi con possibili ripetizioni di elementi (disposizioni, permutazioni e combinazioni con ripetizioni).

Bibliografia.

Pier Cesare Rivoltella. *Didattica inclusiva con gli EAS.* Brescia: Editrice La Scuola

Pier Cesare Rivoltella. *Didattica per competenze e metodo EAS videolezioni*. Brescia: Editrice La Scuola - La Scuola Academy.

Pier Cesare Rivoltella. *Fare didattica con gli EAS (Episodi di Apprendimento Situat*o). Brescia: Editrice La Scuola.

http://www.mathesisnazionale.it/archivio-argomenti/indicazioni-licei.pdf

http://www.mathesisnazionale.it/archivio-argomenti/indicazini-tecnici.pdf

I pre-requisiti essenziali per lo studio della Probabilità e della Statistica

Ferdinando Casolaro[1] **Antonio Fontana**[2]

[1]Dipartimento Architettura
Università di Napoli Federico II
E-mail: ferdinando.casolaro@unina.it

[2]Liceo Scientifico Mercalli Napoli
E-mail: a.fontana2004@libero.it

Sunto

In questo lavoro si sintetizzano i concetti elementari di calcolo combinatorio e semplici esempi sulle prime nozioni elementari - quali medie e indici di variabilità - che sono alla base dello studio del Calcolo della Probabilità e della Statistica. Nell'ultima parte si richiamano le definizioni di probabilità.

1. Introduzione

La Matematica dell'incerto, pur essendo inserita già nei vecchi programmi ministeriali per la Scuola Secondaria di secondo grado (Scuola superiore o liceo prima delle Indicazioni nazionali) dal 1979, non è stata quasi mai oggetto di proposta nei corsi di insegnamento.

C'è quasi un rifiuto da parte dei docenti ad affrontare i temi di Statistica e Calcolo delle Probabilità; alcuni li ritengono non essenziali rispetto agli argomenti classici di Algebra e Geometria, altri ritengono difficile l'approccio.

La motivazione principale è sicuramente dovuta ad una scarsa Formazione fornita nei corsi universitari e, come detto, alla mancanza di proposta agli stessi docenti negli anni in cui hanno frequentato la Scuola Secondaria.

I corsi di Aritmetica e di Geometria vengono affrontati già dalla Scuola Primaria, per cui al primo anno della Scuola di secondo grado lo studio dell'Algebra ed i teoremi di Geometria sintetica trovano negli allievi la base su cui poggiare argomenti più complessi.

Il punto, la retta, il piano, l'angolo, il poligono, …, sono gli elementi di ***geometria*** che i bambini assimilano già dalla Scuola di infanzia e nelle famiglie; la numerazione, le quattro operazioni elementari, l'operazione di potenza… sono i concetti di ***aritmetica*** che si propongono agli allievi dai primi anni della Scuola Primaria.

Pertanto riteniamo opportuno, prima di introdurre le definizioni di probabilità (anche di Statistica) che il docente affronti in modo esauriente e con esempi significativi i concetti di ***media***, ***indici di variabilità*** ed ***analisi combinatoria***.
Le medie classiche (aritmetica, aritmetica ponderata, geometrica, armonica), le medie di posizione (mediana, moda) e gli indici di variabilità (scarto semplice, varianza, scarto quadratico medio) sono gli elementi su cui costruire la matematica dell'incerto, oltre ai principi fondamentali del calcolo combinatorio: fattoriale di un numero intero positivo, coefficiente binomiale, binomio di Newton. Disposizioni, permutazioni, combinazioni semplici e combinazioni con ripetizioni.
Nel paragrafo che segue ci limiteremo alla descrizione di esempi sulle medie e sul calcolo combinatorio, rimandando l'approfondimento degli indici di variabilità all'articolo di questo volume: Casolaro *F.*, *Un approccio didattico all'insegnamento della Statistica*.

2. Definizioni ed esempi.

La media aritmetica tra *n* numeri interi positivi $x_1, x_2, ... x_n$ è il quoziente fra la loro somma e il loro numero n.

$$M = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

Esempio 1: Problema che si risolve con la media aritmetica
Uno studente universitario ha ottenuto nei suoi primi esami i seguenti voti: 24, 28, 25, 29, 30, 18.
Qual è la media dello studente? Quale voto avrebbe dovuto ottenere nell'ultimo esame affinché la media fosse 26?
Soluzione

$$M = \frac{24 + 28 + 25 + 29 + 30 + 18}{6} = 25,\overline{6}$$

Chiamando x il voto che avrebbe dovuto ottenere nell'ultimo esame si ha che:

$$\frac{24 + 28 + 25 + 29 + 30 + x}{6} = 26$$

da cui si ricava $x = 20$.

La media geometrica tra *n* numeri interi positivi $x_1, x_2, ... x_n$ si ottiene estraendo la radice n-sima del loro prodotto.

$$M_g = \sqrt[n]{x_1 \cdot x_2 \cdot ... \cdot x_n}$$

Esempio 2: Problema che si risolve con la media geometrica

Uno scommettitore puntando una somma iniziale pari a 2 euro, ha conseguito un capitale di 432 euro in 3 giocate successive. In particolare, ha vinto:

- Nella prima giocata, 3 volte la somma iniziale ovvero 2·3=6 euro
- Nella seconda giocata, 8 volte la somma precedentemente vinta, ovvero 6·8=48 euro
- Nella terza giocata, 9 volte la somma precedentemente vinta, ovvero 48·9=432 euro

Qual è stata la vincita media riportata?

Soluzione

La somma complessivamente vinta è ottenuta mediante una legge moltiplicativa, ossia: 2(3·8·9) = 432 euro. Per cui, determinare di quante volte in media si è moltiplicato il capitale inizialmente puntato (2 euro) occorre individuare quel valore che sostituito ai fattori moltiplicativi 3,8 e 9 nella funzione di prodotto ne lascia invariato il risultato. Tale valore è la media geometrica dei valori osservati ovvero:

$$M_g = \sqrt[3]{3 \cdot 8 \cdot 9} = 6$$

Osservazione: se fosse stata calcolata la media aritmetica dei valori osservati:

$$M_a = \frac{3+8+9}{3} = 6{,}667$$

Sulla base di questo valore si potrebbe concludere che il giocatore ha vinto mediamente 6.667 volte la somma puntata. Tuttavia tale valore non è corretto: partendo da un capitale iniziale di 2 euro in tre giocate successive non si ottiene il capitale finale di 432 euro, se ad ogni giocata si vince 6.667 volte la posta. Considerata la tipologia del problema, la media idonea a rappresentare i valori osservati è sicuramente la media geometrica.

La media aritmetica ponderata tra *n* numeri interi positivi $x_1, x_2, \ldots. x_n$ positivi, a cui sono associati i numeri $p_1, p_2, \ldots, p_n$ detti pesi, il quoziente fra la somma dei prodotti dei numeri per i loro pesi e la somma dei pesi stessi.

$$P = \frac{x_1p_1 + x_2p_2 + \cdots + x_np_n}{p_1 + p_2 + \cdots + p_n}$$

La media aritmetica è un caso particolare di media ponderata in cui tutti i pesi son o uguali a 1.

Esempio 3: Problema che si risolve con la media aritmetica ponderata

Un' azienda spende per gli stipendi del personale le seguenti quote mensili:

920 € per 12 operai,1240 € per 8 tecnici, 1350 € per 4 impiegati e 1950 € per 2 dirigenti.

Qual è lo stipendio medio pagato dall'azienda?

Soluzione
Lo stipendio medio coincide con la media ponderata dei valori dati:

$$P = \frac{920 \cdot 12 + 1240 \cdot 8 + 1350 \cdot 4 + 1950 \cdot 2}{12 + 8 + 4 + 2} = 1163{,}85\ €$$

La media quadratica tra *n* numeri interi positivi $x_1, x_2, \ldots. x_n$ positivi è uguale alla radice quadrata del rapporto della somma dei quadrati dei valori numerici ed il numero dei valori.

$$M_q = \sqrt{\frac{x_1^2 + x_2^2 + \cdots + x_n^2}{n}}$$

Esempio 4: Problema che si risolve con la media quadratica
Quattro appezzamenti quadrati di terreno misurano rispettivamente (in metri): $l_1 = 25,\ l_2 = 35,\ l_3 = 50,\ l_4 = 85.$
Questi 4 appezzamenti vengono permutati con altri 4 terreni quadrati, uguali fra di loro, in modo da compensare la superficie ceduta con quella ricevuta.
Qual è il lato di ognuno dei 4 terreni uguali ricevuti in permuta?
Soluzione
La misura del lato ricercato coincide con la media quadratica dei valori dati:

$$l = \sqrt{\frac{25^2 + 35^2 + 50^2 + 85^2}{4}} = 53{,}79\ m$$

La media armonica tra *n* numeri interi positivi $x_1, x_2, \ldots. x_n$ è un numero reale M_a tale che il suo reciproco $\frac{1}{M_a}$ è uguale alla media aritmetica dei reciproci di $x_1, x_2, \ldots. x_n$:

$$\frac{1}{M_a} = \frac{\frac{1}{x_1} + \frac{1}{x_2} + \ldots.. + \frac{1}{x_n}}{n}$$

Pertanto, la media armonica di $x_1, x_2, \ldots. x_n$ è data da:

$$M_a = \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \ldots.. + \frac{1}{x_n}}$$

Esempio 5: Problema che si risolve con la media armonica
Un'auto percorre 240 Km effettuando tre soste rispettivamente:

- al 60*esimo* Km.
- al 120*esimo* Km.
- al 180*esimo* Km.

Durante il primo percorso l'auto tiene una velocità media v_{m1} = 80 Km/h; nei percorsi successivi le velocità medie sono rispettivamente:
v_{m2} = 100 Km/h; v_{m3} = 120 Km/h; v_{m4} = 100 Km/h.
Determinare la velocità media dell'auto sull'intero percorso.

Soluzione
La velocità media coincide con la media armonica dei valori richiesti:

$$v_m = \frac{4}{\frac{1}{100}+\frac{1}{100}+\frac{1}{80}+\frac{1}{120}} = 97{,}96\ Km/h$$

Esempio 6: Sul calcolo combinatorio
Date le cinque cifre: 3, 5, 7, 8, 9, stabilire quanti numeri interi (Disposizioni) a tre cifre si possono scrivere con essi. Se si gioca al lotto, quanti terni (Combinazioni) si possono ottenere combinando queste cifre?

Nel primo caso si tratta di calcolare le disposizioni di 5 elementi a tre a tre, in quanto la variazione di posto cambia il valore del numero; nel secondo caso si tratta di combinazioni di 5 elementi a tre a tre, in quanto con la variazione di posto si ha lo stesso terno.

Numeri interi a tre cifre: $D_{5,3} = \frac{5!}{(5-3)!} = \frac{5\cdot 4\cdot 3\cdot 2\cdot 1}{2\cdot 1} = 60$

Numeri di terni al lotto $C_{5,3} = \frac{5!}{3!\cdot(5-3)!} = \frac{5\cdot 4}{2\cdot 1} = 10$

3. I concetti fondamentali della probabilità

3.1) Definizione classica di probabilità
E' il rapporto tra il numero n_f dei casi favorevoli ed il numero n dei casi possibili:

$$p\,(a) = \frac{n_f}{n}$$

Tale definizione è valida solo se i casi sono equiprobabili.

3.2) Definizione frequentista di probabilità

E' una definizione che nasce dall'esperienza, cioè dall'osservazione di una ripetizione di prove e dal numero di volte in cui si verifica l'evento richiesto.
La probabilità frequentista è, dunque, il rapporto tra la *frequenza f* con cui si è verificato l'evento richiesto in *n* osservazioni precedenti ed il numero *n* stesso.

$$p\ (a) = \frac{f}{n}$$

quando il numero delle osservazioni è "*abbastanza grande*". L'espressione "*abbastanza grande*" ha un significato relativo allo specifico evento che si sta analizzando. Guido Castelnuovo definisce la probabilità *frequentista* mediante la seguente affermazione (legge empirica del caso):
"*In una serie di prove ripetute un gran numero di volte, nelle stesse condizioni, ciascuno degli eventi possibili si manifesta con una frequenza relativa (probabilità frequentista) che è presso a poco uguale alla sua probabilità, l'approssimazione cresce al crescere del numero delle prove*".
L'affermazione di Castelnuovo esprime la cosiddetta ***legge dei grandi numeri***:
Con il crescere del numero delle prove, è sempre più probabile che la frequenza relativa di un evento si avvicini alla sua probabilità.

3.3) Definizione soggettiva di probabilità

Premettiamo il quesito assegnato nel liceo PNI nell'anno scolastico 2005/2006.
Bruno de Finetti (1906-1985), tra i più illustri matematici italiani del secolo scorso, del quale ricorre quest'anno il centenario della nascita, alla domanda: "che cos'è la probabilità?" era solito rispondere: "la probabilità non esiste!". Quale significato puoi attribuire a tale risposta? E' possibile collegarla a una della diverse definizioni di probabilità che sono state storicamente proposte?

Il significato da attribuire alla frase **"*la probabilità non esiste!*"** si evince dal concetto di *probabilità soggettiva.* La *probabilità soggettiva p* di un evento *E* è la misura del grado di fiducia espresso dal numero reale *p*, tale che una scommessa di quota *p* su *E* sia *coerente*, cioè tenga conto delle condizioni reali.
La *probabilità soggettiva* è utilizzata nel caso in cui non abbia senso considerare ciò che è avvenuto per una successione di eventi analoghi o si deve assegnare una probabilità anche agli eventi in cui i casi possibili sono *infiniti*.
Dato un numero reale p ($0 < p < 1$) ed una somma di danaro Q, diciamo che si effettua *una scommessa di quota p su un evento E* se, versando la somma pQ si riceve l'importo Q solo se si verifica l'evento *E*.
Il guadagno dello scommettitore, nel caso di vincita è:

$$Q - pQ = Q(1 - p)$$

Da cui si evince che se fosse p > 1, la scommessa sarebbe sempre in perdita.

3.4) Definizione assiomatica di probabilità

E' una definizione che si basa su un'assiomatica che presenta analogie alla struttura della <u>*geometria euclidea*</u> ed alla costruzione della <u>*teoria della misura*</u>.
Precisamente, si fissano degli assiomi su cui viene costruita una serie di operazioni che permettono l'analisi della previsione di eventi.

Ad ogni evento *A* dello spazio campione Ω ($A \in \Omega$) associamo un numero reale *p*, tale che:

a) $0 \leq p(A) \leq 1$

cioè, <u>*la probabilità è una funzione che ad ogni elemento dello spazio campione*</u> Ω <u>*associa un numero reale compreso tra 0 e 1*</u>.

b) Se *S* è l'evento certo, si ha *p(S)* = 1.

c) Se $\varnothing$ è l'evento impossibile, si ha: $p(\varnothing) = 0$.

d) Se $A_1, A_2, \ldots, A_n$ sono n eventi che si escludono a vicenda - cioè $A_i \cap A_j = \varnothing, \quad \forall i \neq j$, si ha:

$$p\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} p(A_i)$$

e) In una σ-algebra (cioè per un insieme di infinità numerabile), si ha:

$$p\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} p(A_i) .$$

La concezione assiomatica della probabilità permette di concepire <u>*l'insieme di tutti i possibili esiti*</u> che si possono verificare come uno <u>*spazio*</u> (*spazio di probabilità* o *spazio dei campioni*).
Un esito (o un evento) è detto <u>*punto campione*</u>.

4. Conclusioni

Come si è potuto constatare dalla lettura degli argomenti trattati, ci siamo limitati a quelli che comunemente vengono indicati come pre-requisiti per affrontare concetti più avanzati.

Nell'ultimo paragrafo si sono date solo le definizioni di probabilità; non si è andati oltre perché il tema nella sua completezza è proposto ed affrontato in altri articoli del presente volume.

Bibliografia

Casolaro F. (2004), Dispensa del corso di ***Statistica*** tenuto alla Facoltà di Scienze MM.FF.NN. (Laurea in Scienze Ambientali) dell'Università del Sannio nel periodo 2004-2007.

Ventre A. (2004) - *Decisioni utili* - Editori Riuniti.

Casolaro F., Prosperi R. (2011), "Atti Terni 2011: *La Matematica per la Scuola di 2° secondo grado: un contributo per il docente di Matematica*". Ed. 2C Contact.

Prosperi R. (2011), *La Matematica dell'incerto: l'insegnamento del* "*Calcolo delle Probabilità*" *e della* "*Statistica*" *nel I biennio della Scuola Secondaria di II grado*. Ed. 2C Contact 2011.

Casolaro F., Paladino L. (2012), "*Didactics of Statistics in Sociology*". First International Conference on *Recent Trends in Social Sciences: Qualitative Theories and Quantitative Models* (RTSS) - Iaşi (Romania), 2012. Pagg. 228-241.

Bergamini M., Trifone A., Barozzi G. "*Matematica.blu*" Vol.1 Ed. Zanichelli 2017.

Sulla probabilità soggettiva nella scuola secondaria

Antonio Maturo[1]

[1]Università G. d'Annunzio di Chieti-Pescara, Viale Pindaro, 42, 65127, Pescara
e-mail: antomato75@gmail.com

Sunto
Si presentano alcuni concetti di logica e probabilità in vista di una introduzione efficace della probabilità soggettiva nella scuola secondaria. L'impostazione soggettiva del Calcolo delle Probabilità è, rispetto alle altre impostazioni, meno dipendente dal calcolo e mette maggiormente in luce una serie di aspetti logici del ragionamento probabilistico. Il collegamento della probabilità soggettiva alle scommesse ed alle decisioni in condizioni di incertezza può portare i ragazzi ad immedesimarsi in ambienti ludici a loro familiari, collegati ad esempio alle partite di calcio, alle decisioni che devono prendere in ogni momento della loro vita, senza dover ricorrere ad un apparato matematico complesso. La soggettività della probabilità può far pensare ad una possibile arbitrarietà, mentre in realtà le valutazioni di probabilità soggettiva sono legate allo stato di informazione del decisore. Inoltre, il rispetto delle condizioni di coerenza limita ampiamente la possibilità di valutazioni arbitrarie.
Si mette in evidenza la necessità della logica a tre valori per un inquadramento logico del concetto di evento condizionato e si precisano le condizioni di coerenza per gli eventi condizionati e degli algoritmi per la verifica di coerenza.

Parole Chiave: Eventi, probabilità soggettiva, scommesse, coerenza, eventi condizionati, probabilità condizionata

1. Introduzione e nozioni di base sugli eventi

Spesso capita di trovare persone con notevoli pregiudizi sulla "*probabilità soggettiva*"; di solito si confonde la parola "soggettiva" con la parola "arbitraria". In realtà la parola "soggettiva" indica che c'è un soggetto informato, il *Decisore,* che deve fare delle scelte sulla base delle sue informazioni.
Le prime decisioni riguardano gli eventi. Un *evento* E può essere definito come una *proposizione della logica bivalente* in cui è ammessa la possibilità che le informazioni in possesso del *Decisore* siano incomplete e quindi non permettono di attivare il *criterio di verificazione* per stabilire se la proposizione è vera o falsa.
"*Un evento è una proposizione di cui può essere non conosciuto il valore di verità. Se tale valore è conosciuto ed è 1, l'evento si dice **certo**, se è 0, si dice **impossibile**, se non è conosciuto si dice **aleatorio***" (de Finetti, 1970, p.710).

Il primo passo, allora, è riconoscere se una data frase è una proposizione della logica bivalente, ossia se è "*una disposizione di parole e/o simboli che esprime ciò che è o vero o falso*" (Russel, 1962). La descrizione di Russel, così come le altre presenti in letteratura, presuppone l'esistenza di un individuo, il *Decisore,* che stabilisce che una disposizione di parole e/o simboli è (a suo parere) un enunciato della logica bivalente.
Riteniamo che, in ambito scolastico, questa mancanza di completa oggettività non sia un inconveniente. I ragazzi sono responsabilizzati a ricoprire il ruolo del Decisore, a riflettere e cooperare per decidere se "una data disposizione di parole e/o simboli" può essere interpretata come una domanda a cui si può dare una e una sola delle due risposte: *vero* o *falso*.
Dal punto di vista linguistico il concetto di *proposizione* è più ampio, in quanto si riferisce ad una frase di senso compiuto (con soggetto, predicato verbale o nominale, complementi, etc.) per la quale si può esprimere un giudizio di verità che non necessariamente si limita a *vero* o *falso*, ma può essere anche *quasi del tutto vero, più vero che falso, a metà fra vero e falso, più falso che vero, quasi del tutto falso, etc.* Lo studio e le applicazioni degli enunciati linguistici si trova, ad es., in (Zadeh, 1965, 1975; Fadini, 1979).
Un evento E è quindi formato da una coppia (P, I) dove P è una proposizione della logica bivalente ed I è l'insieme delle informazioni in possesso del Decisore.
Per far familiarizzare i ragazzi con il concetto di evento, essi devono:

(1) Imparare ad analizzare le frasi che il docente o altri compagni propongono, cercando di capire se sono proposizioni della logica bivalente, proposizioni linguistiche, non proposizioni linguistiche;
(2) Nel caso in cui si ritiene che una proposizione E sia un evento il ragazzo deve distinguere se l'informazione I che possiede sia completa o incompleta per stabilire il valore di verità di E;
(3) Se l'informazione è completa e il Decisore, con l'informazione che possiede, può stabilire che E è vero, allora l'evento E è "*certo*";
(4) Se il Decisore, con l'informazione che possiede, può stabilire che E è falso, l'evento E è "*impossibile*";
(5) Se le informazioni in possesso del Decisore non gli permettono di stabilire se E è vero o falso, l'evento si dice "*aleatorio*".

Usualmente si pone E = 1 se E è vero, E = 0 se E è falso. Gli eventi certi e impossibili hanno valore costante, ogni evento aleatorio è rappresentato da una variabile che può assumere i valori 0 e 1.
Per ogni evento E, *l'evento contrario* E^c è definito come l'evento vero se e solo se E è falso. In formule:

$$E + E^c = 1. \tag{1.1}$$

Fra gli eventi si considera un ordinamento parziale $\subseteq$, ponendo $E \subseteq F$ se dal fatto che E è vero segue che anche F è vero. In formule:

$$E \subseteq F \Leftrightarrow E \leq F. \tag{1.2}$$

Le operazioni logiche binarie più utilizzate sono:

- L'*unione* o *disgiunzione* $\cup$. Dati due eventi E, F, si definisce *unione*, indicata con $E\cup F$, l'evento vero se almeno uno dei due è vero. In formule:

$$E\cup F = \max\{E, F\}. \tag{1.3}$$

- L'*intersezione* o *congiunzione* $\cap$. Dati due eventi E, F, si definisce *intersezione*, indicata con $E\cap F$, o semplicemente E F, l'evento vero se entrambi E e F sono veri. In formule:

$$E\cap F = \min\{E, F\}. \tag{1.4}$$

Due eventi E e F si dicono:

- *incompatibili*, se $E\cap F = \emptyset$;
- *esaustivi*, se $E\cup F = \Omega$.

Osserviamo che se E e F sono incompatibili ed esaustivi allora $F = E^c$.

Una famiglia di n eventi aleatori $\{E_1, E_2, \ldots, E_n\}$, $n \geq 2$ si dice *partizione dell'evento certo* se gli eventi sono a due a due incompatibili e la loro unione è l'evento certo. In formule:

$$E_1 + E_2 + \ldots + E_n = 1. \tag{1.5}$$

Dati n eventi qualsiasi $E_1, E_2, \ldots, E_n$, $n \geq 1$, si dicono *atomi* o *costituenti* della famiglia $\{E_1, E_2, \ldots, E_n\}$ tutti gli eventi non impossibili del tipo $F_1\cap F_2\cap\ldots\cap F_n$, con $F_i\in\{E_i, E_i^c\}$. Evidentemente risulta:

(a) per ogni evento E_i e per ogni *costituente* C_r, $C_r \subseteq E_i$ oppure $C_r \cap E_i = \emptyset$;
(b) ogni evento E_i è l'unione dei costituenti contenuti in E_i;
(c) due costituenti distinti sono fra loro incompatibili;
(d) per $n \geq 2$, se gli eventi E_i sono aleatori allora i costituenti formano una partizione dell'evento certo;
(e) aggiungendo $\emptyset$ e Ω agli eventi della famiglia i costituenti non cambiano.

Si dice *algebra* generata dalla famiglia $\{E_1, E_2, \ldots, E_n\}$ l'insieme che ha per elementi $\emptyset$, Ω e tutte le unioni di un numero qualsiasi di costituenti.

Per poter costruire i costituenti è necessario capire i legami logici fra gli eventi della famiglia.
Due eventi A e B si dicono *logicamente indipendenti* se tutte le intersezioni AB, AB^c, A^cB, A^cB^c sono eventi possibili. Se almeno una di tali intersezioni è impossibile gli eventi sono *logicamente dipendenti*. La dipendenza logica è "semplice" se una sola di tali intersezioni è impossibile, è multipla se almeno due intersezioni sono impossibili. Definizioni analoghe si hanno per tre o più eventi.

È consigliabile far esercitare i ragazzi a capire in significato dei vari tipi di dipendenza logica ed a determinare i costituenti che si ottengono a partire da 1, 2, 3, 4 eventi. Un ulteriore esercizio è la determinazione dell'algebra associata a tali eventi.

Importante è anche evidenziare situazioni in cui bisogna prendere decisioni in condizioni di incertezza, mostrando la necessità di individuare, per tale scopo, i costituenti degli eventi che vanno presi in considerazione.

2. La probabilità soggettiva e le condizioni di coerenza

Sia E un evento. Supponiamo che due giocatori, A detto *scommettitore* e B detto *banco*, decidono di organizzare una *scommessa* sul verificarsi di E con le seguenti regole:

(a) lo scommettitore A versa a B una somma S(E), detta *puntata*;
(b) il banco B paga una somma V(E), detta *vincita*, se si verifica E.

Si ammette la possibilità che la puntata S(E) sia positiva, nulla oppure sia negativa. Se S(E) < 0 vuol dire che è B che paga la puntata. Anche la vincita V(E) può essere positiva, nulla o negativa. Se V(E) < 0 allora è A che paga la vincita, se V(E) = 0 allora la scommessa non viene effettuata. Supponiamo $V(E) \neq 0$.

La scommessa si dice *coerente* se non porta un guadagno certo (non nullo) a uno dei due giocatori.

Per ogni evento E poniamo E = 1 se E si verifica ed E = 0 se E non si verifica. Allora il guadagno di A è dato dalla formula:

$$G_A = E\,V(E) - S(E). \tag{2.1}$$

Possono verificarsi tre casi:

(1) E è l'evento certo Ω. Allora nella formula (2.1) risulta E = 1 e per evitare un guadagno certo ad uno dei due giocatori si deve porre S(E) = V(E);
(2) E è l'evento impossibile $\emptyset$. Allora nella formula (2.1) risulta E = 0 e per escludere un guadagno certo di uno dei due giocatori si deve porre S(E) = 0;
(3) E è un evento aleatorio. Allora può essere E = 1 oppure E = 0. I possibili guadagni di A sono: V(E) – S(E), - S(E). Affinché non siano entrambi positivi o entrambi negativi, S(E) e V(E) non possono avere segni discordi e deve essere $|S(E)| \leq |V(E)|$.

Il rapporto p(E) = S(E) / V(E) si dice *probabilità soggettiva* di E (concordata fra A e B). La formula (2.1) si riscrive come:

$$G_A = (E - p(E))\,V(E). \tag{2.2}$$

In termini di probabilità le condizioni di coerenza si riducono alle seguenti formule:

(C1) $p(\Omega) = 1$, $p(\emptyset) = 0$ (*prima condizione di coerenza*);
(C2) $0 \leq p(E) \leq 1$, per ogni evento E (*seconda condizione di coerenza*).

Importante far notare ai ragazzi che p(E) = 0 può valere anche se E non è l'evento impossibile, ma si ha una estrema fiducia che esso non si verifichi. Inoltre si può valutare p(E) = 1 anche se E non è l'evento certo.

Sia $\{E_1, E_2, \dots, E_n\}$, $n \geq 2$, una partizione dell'evento certo. Supponiamo che A scommetta contemporaneamente su tutti gli eventi E_i con vincita fissa $V \neq 0$, puntando S_i sull'evento E_i. Allora, ponendo $p(E_i) = S_i/V$, si vede che il guadagno di A è dato dalla formula:

$$G_A = [\sum_i(E_i - p(E_i))]\ V. \quad (2.3)$$

Poiché $\{E_1, E_2, \ldots, E_n\}$ è una partizione dell'evento certo, risulta $\sum_i E_i = 1$, per cui G_A si riduce ad un numero. Affinché non ci sia un guadagno certo per uno dei due giocatori l'espressione fra parentesi quadre deve essere nulla. Si ha quindi la condizione di coerenza:

(C3) se $\{E_1, E_2, \ldots, E_n\}$, $n \geq 2$, è una partizione dell'evento certo allora $\sum_i p(E_i) = 1$ (*terza condizione di coerenza*). In formule:

$$E_1 + E_2 + \ldots + E_n = 1 \quad \Rightarrow \quad p(E_1) + p(E_2) + \ldots + p(E_n) = 1. \quad (2.4)$$

Si può osservare che, per la proprietà associativa dell'unione, è sufficiente imporre la condizione (C3) solo per $n = 2$ e $n = 3$. Per induzione si dimostra che vale per n qualsiasi.

Bisogna evidenziare agli studenti che la formula (2.4) deve essere accompagnata un giudizio qualitativo per stabilire effettivamente le probabilità $p(E_i)$. Precisamente, per ogni coppia (E_i, E_j) di eventi della partizione bisogna prima valutare se si ritengono "ugualmente probabili" oppure uno di essi si ritiene più probabile dell'altro.
Se E_i si ritiene più probabile di E_j bisogna scegliere fra giudizi tipo: "*un po' più probabile*", "*abbastanza più probabile*", "*molto più probabile*", "*assolutamente più probabile*" e poi seguire dei criteri per ottenere probabilità $p(E_i)$ consistenti con i confronti a coppie effettuati. Ad esempio, si può applicare il metodo introdotto da Saaty (1980) che introduce matrici di confronto a coppie di cui calcola gli autovettori normalizzati associati all'autovalore principale.
Nelle scommesse ordinarie, ad esempio sulle partite di calcio, si considera spesso il caso $n = 2$, in cui un evento aleatorio E è confrontato con il suo contrario E^c. Nello stabilire le quote delle scommesse, viene preso in considerazione il rapporto $p(E) / p(E^c)$, uguale a $S(E) / S(E^c)$ nel caso di vincita fissa scommettendo su E e E^c.

Nella probabilità soggettiva non si ritiene opportuno estendere la condizione di coerenza (C3) al caso di una partizione numerabile dell'evento certo, in quanto con giudizi di "*ugualmente probabile*" o "*un po' più probabile*" non si può ottenere che la somma di una quantità numerabile di probabilità sia 1.

A partire dalle condizioni di coerenza si dimostrano immediatamente la seguenti:

Proprietà additiva Se $E_1, E_2, \ldots, E_n$ sono eventi a due a due incompatibili, allora:

$$p(E_1 \cup E_2 \cup \ldots \cup E_n) = p(E_1) + p(E_2) + \ldots + p(E_n). \quad (2.5)$$

Proprietà di monotonia Se E e F sono due eventi allora:

$$E \subseteq F \Rightarrow p(E) \leq p(F). \quad (2.6)$$

3. **La verifica di coerenza di una assegnazione di probabilità soggettiva**.

Sia $\{E_1, E_2, \ldots, E_n\}$ una famiglia qualsiasi di $n \geq 1$ eventi. Supponiamo che il Decisore D abbia assegnato la probabilità $p(E_i)$ all'evento E_i, $i = 1, 2, \ldots, n$. Per verificare che l'assegnazione di probabilità è coerente si può seguire il seguente algoritmo:

(1) Si calcolano i costituenti $C_1, C_2, \ldots, C_k$ di $\{E_1, E_2, \ldots, E_n\}$.

(2) Detta x_s la probabilità (incognita) del costituente C_s e posto $a_{is} = 1$ se $C_s \subseteq E_i$, $a_{is}=0$ se $C_s \subseteq E_i^c$, si verifica la compatibilità del sistema:

$$a_{i1}\, x_1 + a_{i2}\, x_2 + \ldots + a_{ik}\, x_k = p(E_i),\ i = 1, 2, \ldots, n \qquad (3.1)$$

con i vincoli:

$$x_1 + x_2 + \ldots + x_k = 1,\ x_s \geq 0,\ s = 1, 2, \ldots, k \qquad (3.2)$$

(3) Se il sistema è compatibile allora l'assegnazione di probabilità $\{p(E_i), E_i, i = 1, 2, \ldots, n\}$ è coerente, altrimenti il decisore D deve cambiare la sua assegnazione di probabilità.

(4) Per quanto riguarda la probabilità dei costituenti, è possibile assegnare una funzione obiettivo $y = b_1\, x_1 + b_2\, x_2 + \ldots + b_s\, x_s$, in corrispondenza di opinioni del decisore, da massimizzare o minimizzare. Il problema della verifica della coerenza si riduce allora ad un problema di programmazione lineare.

Se $\{p(E_i), E_i, i = 1, 2, \ldots, n\}$ è una assegnazione di probabilità coerente e si introduce un nuovo evento E, allora è possibile stabilire un intervallo $[p_1, p_2]$ di valori che può assumere p(E) in modo che $\{p(E_i), E_i, i = 1, 2, \ldots, n\} \cup \{p(E)\}$ sia coerente.

In particolare, una volta verificata la coerenza di $\{p(E_i), E_i, i = 1, 2, \ldots, n\}$, per ogni soluzione $(x_1, x_2, \ldots, x_s)$ del sistema esiste, in base alla (2.5), un prolungamento coerente della probabilità per ogni evento A dell'algebra generata da $\{E_1, E_2, \ldots, E_n\}$, dato dalla formula:

$$p(A) = a_1 x_1 + a_2 x_2 + \ldots + a_k\, x_k, \qquad (3.3).$$

con $a_r = 1$ se $C_r \subseteq A$ e $a_r = 0$ se $C_r \subseteq A^c$.

Per un approfondimento dei concetti svolti in questo paragrafo e delle loro applicazioni si possono consultare i testi (Coletti, Scozzafava, 2002; Scozzafava, 1996, 2001; Maturo, 2008).

Le applicazioni della probabilità soggettiva coerente alle decisioni in condizioni di incertezza sono in (Lindley, 1990).

4. **Gli eventi condizionati e la logica trivalente**

Una caratteristica fondamentale della probabilità soggettiva è la maniera in cui viene trattato il concetto di evento condizionato e la probabilità condizionata. Precisamente, per de Finetti (1970: 641-642), un evento condizionato E|H è una proposizione che può

assumere tre valori di verità: *vero* se E H è vero, *falso* se si verifica E^c H e *indeterminato* se si verifica H^c.
Il testo di base per la logica trivalente a cui si ispira de Finetti è quello di Reichenbach, (1944), in cui vengono considerate proposizioni con tre valori di verità: *vero* = 1, *falso* = 0, *indeterminato* = i, con l'ordinamento 0 < i < 1.
Al posto della parola *indeterminato* si possono avere molti altri termini, ciascuno con il suo significato dal punto di vista filosofico, come "vuoto", "indeciso", "inconoscibile", "indecidibile", con particolari significati (Fadini, 1979: 40-43).
La struttura algebrica degli eventi condizionati (o generalizzati) è studiata, oltre che in (Reichenbach, 1942) e in (de Finetti, 1970), anche in altri lavori (ad es. Gentilhomme, 1968, Fadini, 1970; Maturo 1993).
Una assiomatica della probabilità degli eventi condizionati dal punto di vista di de Finetti si trova in (Dubins, 1975), in cui, per esigenze di calcolo, si assume la condizione $H \neq \emptyset$.

Le operazioni logiche ternarie di unione e intersezione di eventi condizionati sono una immediata generalizzazione di quelle della logica binaria. Precisamente:

Unione o *disgiunzione* $\cup$. Dati due eventi condizionati E, F, si definisce unione, $E \cup F$, di E e F l'evento condizionato il cui valore di verità è dato dalla formula:

$$E \cup F = \max\{E, F\}. \tag{4.1}$$

Intersezione o *congiunzione* $\cap$. Dati due eventi condizionati E, F, si definisce intersezione, $E \cap F$, di E e F l'evento condizionato il cui valore di verità è dato dalla formula:

$$E \cap F = \min\{E, F\}. \tag{4.2}$$

Un approfondimento del concetto di evento condizionato coinvolge il concetto di implicazione. L'evento condizionato E|H può essere visto come la proposizione condizionale "se H allora E".
Nella logica binaria tale proposizione, detta *implicazione* e indicata con $H \rightarrow E$, è la proposizione vera se $H \leq E$ e falsa se $H > E$. Il concetto di evento condizionato, invece, fa riferimento ad una *implicazione alternativa* (Reichenbach, 1944), che possiamo indicare con $H \rightarrow_a E$, che è una proposizione della logica ternaria, *vera* se E = H = 1, *falsa* se E = 0, H = 1, *indeterminata* negli altri casi. La scelta di tale implicazione alternativa è giustificata dalla interpretazione della probabilità condizionata come scommessa.

5. La probabilità condizionata e le condizioni di coerenza

Nell'impostazione assiomatica della probabilità, la probabilità condizionata è definita dalla formula:

$$p(E|H) = p(E\,H)/p(H), \text{ per } p(H) \neq 0 \tag{5.1}$$

Nell'impostazione soggettiva, invece, non è richiesta la "condizione probabilistica" $p(H) \neq 0$, ma solo la condizione logica $H \neq \emptyset$. Dal punto di vista soggettivo la (5.1) è un teorema, conseguenza delle condizioni di coerenza. Altri teoremi valgono per $p(H) = 0$, con la condizione $H \neq \emptyset$.

La *probabilità condizionata* $p(E|H)$, $H \neq \emptyset$, che ha il significato di *grado di fiducia* che l'evento condizionato $E|H$ risulti *vero*, viene assunta esistente, anche se talvolta non conosciuta, per ogni evento condizionato.

In base a *principi di coerenza* che "tutti sono disposti ad accettare", vengono stabilite le relazioni fra le varie probabilità condizionate.

Lo strumento operativo è il concetto di *scommessa coerente*.

Un giocatore A punta una somma S sul verificarsi dell'evento condizionato $E|H$ e:

- riceve una vincita $V \neq 0$ se $E|H$ risulta vero;
- ottiene la restituzione della puntata S se $E|H$ è indeterminato;
- non ottiene nulla se $E|H$ è falso.

Il rapporto $p(E/H) = S/V$ è la *probabilità soggettiva* di $E|H$, valutata da A.

La scommessa si dice *coerente* se, ammesso che si verifichi H, non porta un guadagno certo (non nullo) a uno dei due giocatori.

Poniamo $E|H = 1, 0, p(E|H)$, a seconda che $E|H$ sia vero, falso o indeterminato. Il guadagno di A è dato dalla formula:

$$G_A = (E|H - p(E|H))\, V. \tag{5.2}$$

Ammesso che si verifichi H i guadagni possibili sono $(1 - p(E|H))\, V$ e $-p(E|H))\, V$.

Le condizioni di coerenza, in termini di probabilità si riducono alle seguenti formule:

(C1H) $p(H|H) = 1$, $p(\emptyset|H) = 0$, (*prima condizione di coerenza*);

(C2H) $0 \leq p(E|H) \leq 1$, per ogni evento E (*seconda condizione di coerenza*).

Sia $\{E_1, E_2, \ldots, E_n\}$, $n \geq 2$, una partizione dell'evento $H \neq \emptyset$, ossia una famiglia di eventi non impossibili, a due a due incompatibili e tale che la loro unione è uguale ad H. Supponiamo che A scommetta contemporaneamente su tutti gli eventi $E_i|H$ con vincita fissa $V \neq 0$, puntando S_i sull'evento $E_i|H$. Allora, ponendo $p(E_i|H) = S_i/V$, si vede che il guadagno di A è dato dalla formula:

$$G_A = [\textstyle\sum_i(E_i|H - p(E_i|H))]\, V. \tag{5.3}$$

Poiché $\{E_1, E_2, \ldots, E_n\}$ è una partizione di H, se si verifica H risulta $\sum_i E_i|H = 1$, per cui G_A si riduce ad un numero. Affinché non ci sia un guadagno certo per uno dei due giocatori l'espressione fra parentesi quadre deve essere nulla. Si ha quindi la condizione di coerenza:

(C3H) se $\{E_1, E_2, \ldots, E_n\}$, $n \geq 2$, è una partizione dell'evento H, allora $\sum_i p(E_i|H) = 1$ (*terza condizione di coerenza*). In formule:

$$E_1|H + E_2|H + \ldots + E_n|H = H|H \;\Rightarrow\; p(E_1|H) + p(E_2|H) + \ldots + p(E_n|H) = 1. \tag{5.4}$$

A partire dalle condizioni di coerenza si dimostra immediatamente la seguente:

Proprietà additiva degli eventi condizionati Se $E_1, E_2, \ldots, E_n$ sono eventi a due a due incompatibili, allora:

$$p(E_1|H \cup E_2|H \cup \ldots \cup E_n|H) = p(E_1|H) + p(E_2|H) + \ldots + p(E_n|H). \qquad (5.5)$$

Una quarta e ultima condizione di coerenza si ottiene considerando una scommessa su tre eventi condizionati $E|H$, $H|K$, $E|K$, con $E \subseteq H \subseteq K$, $H \neq \emptyset$. Se le vincite scommettendo su tali eventi condizionati sono, rispettivamente, V_1, V_2, V_3, il guadagno di A è dato dalla formula:

$$G_A = [(E|H - p(E|H))]\, V_1 + [(H|K - p(H|K))]\, V_2 + [(E|K - p(E|K))]\, V_3 \qquad (5.6)$$

La condizione di coerenza è che non si possono scegliere V_1, V_2, V_3 in modo che, ammesso che si verifichi K, risulti G_A sempre positivo. Con semplici calcoli si ottiene, come conseguenza della coerenza, la seguente *proprietà moltiplicativa* (o *quarta condizione di coerenza*):

(C4H) La scommessa su $E|H$, $H|K$, $E|K$, con $E \subseteq H \subseteq K$ è coerente se e solo se:

$$p(E|H)\, p(H|K) = p(H|K). \qquad (5.7)$$

Si può introdurre un ordinamento fra eventi condizionati ponendo:

$$E|H \leq F|K \Leftrightarrow (E\, H \leq FK,\ E^c\, H \geq F^cK). \qquad (5.8)$$

In particolare, se $E \subseteq H \subseteq K$, risulta $\min\{E|H, H|K\} \geq E|K$. Dalla (5.7) si ottiene la *proprietà di monotonia*:

$$E \subseteq H \subseteq K \Rightarrow \min\{p(E|H), p(H|K)\} \geq p(E|K). \qquad (5.9)$$

6. La verifica di coerenza di una assegnazione di probabilità condizionata

Sia $\{E_1|H_1, E_2|H_2, \ldots, E_n|H_n\}$ una famiglia qualsiasi di $n \geq 1$ eventi condizionati. Si definiscono *costituenti* o *atomi* della famiglia tutti gli eventi non impossibili del tipo $F_1 \cap F_2 \cap \ldots \cap F_n$, con $F_i \in \{E_i\, H_i,\ E_i^c\, H_i,\ H_i^c\}$. Supponiamo che il Decisore D abbia assegnato la probabilità $p(E_i|H_i)$ all'evento $E_i|H_i$, $i = 1, 2, \ldots, n$. Per verificare che l'assegnazione di probabilità condizionata $\{p(E_1|H_1), p(E_2|H_2), \ldots, p(E_n|H_n)\}$ è coerente bisogna prima definire la "coerenza parziale rispetto ad un sottoinsieme" della famiglia.

Sia $S = \{E_1|H_1, E_2|H_2, \ldots, E_m|H_m\}$ un sottoinsieme non vuoto di $\{E_1|H_1, E_2|H_2, \ldots, E_n|H_n\}$. Allora:

(1) Si calcolano i costituenti $C_1, C_2, \ldots, C_k$ di $\{E_1|H_1, E_2|H_2, \ldots, E_m|H_m\}$ e si calcola l'evento H unione degli eventi H_i.

(2) Detta x_s la probabilità (incognita) dell'evento condizionato $C_s|H$, si pone $a_{is} = 1$ se $C_s \subseteq E_i\, H_i$ e $a_{is} = 0$ se $C_s \subseteq H - E_i\, H_i$.

(3) Si pone $b_{is} = 1$ se $C_s \subseteq H_i$ e $b_{is} = 0$ se $C_s \subseteq H_i^c$.

(4) Si verifica la compatibilità del "sistema associato a $\{E_1|H_1, E_2|H_2, \ldots, E_m|H_m\}$":

$$a_{i1}\, x_1 + a_{i2}\, x_2 + \ldots + a_{ik}\, x_k = p(E_i|H)\, (b_{i1}\, x_1 + b_{i2}\, x_2 + \ldots + b_{ik}\, x_k),\ i = 1, 2, \ldots, m \quad (6.6)$$

con i vincoli:

$$x_1 + x_2 + \ldots + x_k = 1,\ x_s \geq 0,\ s = 1, 2, \ldots, k \quad (6.7)$$

Se il sistema (6.6), con i vincoli (6.7), ammette soluzioni allora si dice che l'assegnazione di probabilità condizionata è *parzialmente coerente* rispetto ad S. L'assegnazione di probabilità $\{p(E_1|H_1), p(E_2|H_2), \ldots, p(E_n|H_n)\}$ è coerente se è parzialmente coerente rispetto ad ogni sottoinsieme non vuoto di $\{E_1|H_1, E_2|H_2, \ldots, E_n|H_n\}$.

7. Conclusioni e prospettive di ricerca

L'approccio soggettivo al Calcolo delle Probabilità sembra particolarmente efficace da un punto di vista didattico. Esso approfondisce maggiormente, rispetto ad altre impostazioni, i concetti logici su eventi e soprattutto su eventi condizionati, stabilendo anche un legame interdisciplinare fra matematica e italiano. In particolare, nell'impostazione soggettiva si fa una netta distinzione fra concetti logici e concetti probabilistici, ad esempio evidenziando la differenza fra indipendenza logica e indipendenza probabilistica.

L'idea della scommessa è accettata con entusiasmo dai ragazzi e le condizioni di coerenza sembrano abbastanza semplici da comprendere, anche nella scuola del primo ciclo. Inoltre, è facile far vedere che le impostazioni classica e frequentistica possono essere classificate come casi particolari dell'approccio soggettivo.

L'apparato matematico è ridotto al massimo. Non vi è una partizione dell'evento certo privilegiata, né un'algebra di eventi a priori. A mano a mano che si presentano degli eventi si indaga sui loro legami logici e si costruisce l'insieme dei costituenti e da questi l'algebra generata. Tali insiemi si ampliano con la presa in considerazione di nuovi eventi.

Il concetto di evento condizionato come proposizione della logica trivalente è una caratteristica importante della probabilità soggettiva. Da un punto di vista logico fa riflettere sul concetto di implicazione e sull'opportunità di una implicazione alternativa.

Dal punto di vista probabilistico si può assegnare direttamente una probabilità ad ogni evento condizionato $E|H$, con $H \neq \emptyset$, senza doversi preoccupare se, assegnata una probabilità agli eventi non condizionati, sia $p(H) \neq 0$. Nell'assegnazione delle probabilità condizionate un ruolo importante gioca il confronto a coppie, ad esempio se A e B sono due eventi incompatibili, vanno confrontati gli eventi condizionati $A|(A\cup B)$, $B|(A\cup B)$ e le relative probabilità condizionate devono soddisfare la condizione di coerenza che la loro somma è uguale a 1.

Le condizioni di coerenza per eventi condizionati viste nel paragrafo 5 nascondono concetti di Analisi non Standard. Gli eventi si possono dividere in livelli. Se A e B sono allo stesso livello allora valgono entrambe le condizioni di positività $p(A|(A\cup B)) > 0$, $p(B|(A\cup B)) > 0$. Se invece A è ad un livello inferiore rispetto a B allora $p(A|(A\cup B) = 0$.

Un ausilio per la comprensione della probabilità soggettiva è dato dalla probabilità geometrica. Se l'evento certo Ω è rappresentato da un rettangolo R e si ammette una distribuzione uniforme, allora, per un qualsiasi poligono P contenuto in R, la probabilità di P è il rapporto area(P)/area(R). Se S e T sono due segmenti disgiunti contenuti in R, allora $p(S|(S \cup T))$ non si può ottenere come rapporto fra p(S) e $p(S \cup T)$ che sono entrambi nulli. Tuttavia, una valutazione soggettiva di buon senso, che risulta anche essere coerente, porta a valutare $p(S|(S \cup T))$ come rapporto fra le lunghezze di S e di $S \cup T$.
Un indebolimento dell'assiomatica della probabilità soggettiva si ottiene con la *probabilità fuzzy*, in cui si considerano misure non additive (Klir, Yuan,1995; Maturo et al., 2006, 2010). In particolare, la proprietà additiva (2.5) può essere sostituita dalla più generale condizione di monotonia (2.6). Analogamente la proprietà moltiplicativa della probabilità condizionata può essere sostituita dalla (5.9).

Bibliografia

Coletti G., Scozzafava R., (2002), *Probabilistic Logic in a Coherent Setting*, Kluwer Academic Publishers, Dordrecht.

De Finetti B., (1970), *Teoria delle Probabilità*, Einaudi, Torino.

Dubins, L.E., (1975), Finitely additive conditional probabilities, conglomerability and disintegrations, *The Annals of Probability*, 3, 89-99.

Fadini, A., (1979), *Introduzione alla teoria degli insiemi sfocati*, Liguori Editore, Napoli.

Gentilhomme, M.Y., (1968), Les ensembles flous en linguistiques, *Cahiers de linguistique theorique et appliquée*, Bucarest, (5) 47, pp. 47-65.

Klir G., Yuan B., (1995), Fuzzy Sets and Fuzzy Logic: Theory and Applications, Prentice Hall,New Jersey, 1995,

Lindley D. V., (1990), *La logica della decisione*, Il Saggiatore, Milano.

Maturo A., (1993), Struttura algebrica degli eventi generalizzati, *Periodico di Matematiche*, 4, 1993, p. 18-26.

Maturo A., (2008), La moderna visione interdisciplinare di Geometria, Logica e Probabilità in Bruno de Finetti**,** *Ratio Sociologica,* 2, 2008, pp. 39-62.

Maturo A., Squillante M., Ventre A. G. S., (2006), Consistency for nonadditive measures:analytical and algebraic methods, in B. Reusch, Editor, Computational Intelligence, Theory and Applications, in Advances in Soft Computing, Springer, Berlin, Heidelberg, New York.

Maturo A., Squillante M., and Ventre A.G.S., (2010), Decision Making, Fuzzy Measures, and Hyperstructure, Advances and Applications in Statistical Sciences, Volume 2(2), 2010, 233-253.

Reichenbach H., (1942), *I fondamenti filosofici della meccanica quantistica*, tr. it. Einaudi, Torino, 1952

Russell B., (1962), *Introduzione alla filosofia matematica*, Longanesi, Milano.

Saaty, T.L. (1980). The Analytic Hierarchy Process, New York: McGraw-Hill.

Scozzafava R., (1996), *La probabilità soggettiva e le sue applicazioni*, Zanichelli, Bologna.

Scozzafava R., (2001), *Incertezza e probabilità. Significato, valutazione, applicazioni della probabilità soggettiva*, Zanichelli, Bologna.

Zadeh L. A. (1965). Fuzzy Sets. Information and Control. 8. 338-358.

Zadeh L. A. (1975), The concept of a linguistic variable and its application to approximate reasoning, Inf Sci 1975;8 Part I:199–-249, Part II 301–-357, Part III. Inf Sci 1975;9: 43–-80.

Distribuzione di probabilità
Variabili casuali discrete e continue

Giovanna Della Vecchia[1]   Alberto Trotta[2]

[1]I.I.S. Minzoni - Giugliano in Campania
E-mail: giovanna.dellavecchia@gmail.com

[2]I.I.S.S. Caterina-Amendola - Salerno
Fondatore sezione Mathesis Anzio-Nettuno
E-mail: albertotrotta@virgilio.it

Sunto Si propone il semplice compito di introdurre le distribuzioni di probabilità discreta nel corso del primo biennio della Scuola secondaria di secondo grado.

1. Introduzione

La distribuzione di probabilità è uno degli argomenti che spesso viene introdotto direttamente nel caso continuo. Si presenta, cioè, la curva di Gauss con la classica relazione:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

che, ovviamente, richiede conoscenze avanzate per cui non è possibile trattarla nel corso del primo biennio della Scuola secondaria di secondo grado, dove è opportuno partire da semplici esempi numerici della vita quotidiana che richiedono esclusivamente conoscenze di aritmetica elementare.
Si riportano, pertanto, alcune considerazioni ed alcune definizioni ritenute utili ai fini della nostra esposizione.

Si chiama ***variabile casuale*** *una funzione che associa ad ogni evento elementare E dello Spazio Campionario S, uno ed un solo numero reale.*

I valori di una variabile casuale sono numeri reali determinati dal risultato di un esperimento, e quindi incerti; di conseguenza una tale variabile è sempre accompagnata da una funzione di probabilità.

I problemi che si devono affrontare nel mondo reale risulteranno sempre più semplificati se invece di parlare di eventi si fa riferimento a funzioni a valori reali, definite a partire dagli eventi stessi. Infatti i risultati di un esperimento possono essere del tipo più disparato: numeri, attributi, caratteristiche, ecc. ; per poter trattare matematicamente tali risultati sarà necessario associare ad essi dei numeri reali.
Per fare un esempio, invece di dare un giudizio qualitativo sullo stato di salute di una persona, si possono dare elementi qualitativi che esprimono tale concetto in modo più rigoroso come la pressione sanguigna, il tasso di certe sostanze nel sangue , il suo peso corporeo.

Definire una variabile casuale significa quindi trovare una regola in base alla quale associare un numero reale ad ogni risultato di un esperimento e quindi ad ogni elemento dello spazio degli eventi.

- *Una **variabile casuale** si dice **discreta** quando può assumere solo particolari valori in punti isolati di un intervallo.*

- *Una **variabile casuale** si dice **continua** se i suoi valori possono variare con continuità in un intervallo, che può essere limitato o illimitato.*

Dal punto di vista *didattico-metodologico* è opportuno introdurre questi concetti attraverso semplici esempi che possano permettere agli allievi un *apprendimento di carattere laboratoriale*.

2. La Distribuzione di probabilità discreta attraverso problemi.

Problema 1. - *Consideriamo lo spazio degli eventi (spazio campionario) associato al lancio di tre monete ponendoci il problema di determinare*
il numero totale delle volte in cui si presenta croce?

Osserviamo che, lanciando contemporaneamente le tre monete, i possibili risultati dell'esperimento sono 2^3 e che croce si può presentare:

1. zero volte (evento E0) (tre volte testa). *2. una volta (evento E1) (due volte testa).*
3. due volte (evento E2) (una volta testa). *4. tre volte (evento E3) (zero volte testa).*

E' opportuno pertanto far rappresentare agli allievi la seguente tabella, da cui si evince immediatamente che la variabile indipendente che indichiamo con X può assumere solo i valori 0, 1 , 2 , 3

Punti dello spazio campionario	Numero di volte che esce croce:
CCC	3
CCT	2
CTC	2
CTT	1
TCC	2
TCT	1
TTC	1
TTT	0

Attribuendo ad ogni valore della variabile casuale X la corrispondente probabilità, si ottiene la ***distribuzione di probabilità di X***.

Nel nostro caso lo spazio campionario ***S*** è quindi costituito da otto elementi:

$$\boldsymbol{S} \equiv \{CCC, \quad CCT, \quad CTC, \quad CTT, \quad TCC, \quad TCT, \quad TTC, \quad TTT\}.$$

Poiché la variabile X assume: {una sola volta il valore zero, tre volte il valore uno, tre volte il valore due, una volta il valore tre}, l'insieme

$$\{X_0 = 0, \quad X_1 = 1, \quad X_2 = 2, \quad X_3 = 3\}$$

individua il dominio ***D*** della nostra funzione ai cui valori sono associate, rispettivamente, le seguenti probabilità:

$$\boldsymbol{D} \equiv \{p_0=1/8, \ p_1=3/8, \ p_2=3/8, \ p_3=1/8\}.$$

La legge che associa ad ogni valore della variabile casuale X_k - k = 0…3 - il corrispondente valore della probabilità p_k è detta ***distribuzione di probabilità della variabile casuale discreta X.***

Non è superfluo far notare agli allievi che se chiedo la probabilità affinché si verifichi uno solo degli eventi dello spazio campionario ***S*** (precisamente, lanciando le monete una alla volta, chiedo la probabilità affinché si verifichi CCC, oppure TTC,…ecc.) ci troviamo ad operare nel caso classico di eventi equiprobabili, per cui la probabilità di ogni evento è 1/8.
In questo caso la corrispondenza tra gli elementi dello spazio campionario ***S*** e gli elementi del dominio ***D*** della variabile casuale X è biettiva per cui, operativamente, il dominio ***D*** si può identificare con lo spazio campionario ***S***.

Problema 2. - ***Determinare la distribuzione di probabilità della variabile X che esprime la somma dei punteggi ottenuti lanciando due dadi.***

Come è noto i risultati possibili di questo esperimento sono 2, 3, 4 5,6 , 7 , 8, 9, 10, 11, 12.
Associamo ora a questi valori di X le corrispondenti probabilità *p*(X)

X:	2	3	4	5	6	7	8	9	10	11	12
p(X)	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

L'insieme costituito dagli 11 possibili risultati dall'esperimento con le rispettive probabilità associate
rappresenta la ***distribuzione di probabilità*** della variabile casuale X che abbiamo considerato.

- E' opportuno far rappresentare agli allievi le coppie [X, *p*(X)] sul piano cartesiano perché uno degli aspetti più significativi dei problemi proposti è l'osservazione che, con i valori di X sulle ascisse e i valori di *p*(X) sulle ordinate, i punti ad ordinata massima si hanno nella parte centrale della rappresentazione e i punti ad ordinata minima agli estremi della rappresentazione, come gli stessi allievi avranno modo di constatare negli anni successivi con la rappresentazione della curva di Gauss.

Indicato con x il numero reale che rappresenta il valore della variabile casuale X, la terna $\{S, D, p(x)\}$ è detta ***Spazio di Probabilità***.
Pertanto, uno spazio di probabilità è una terna $\{S, D, p(x)\}$, dove S è un insieme qualunque (in genere pensato come l'insieme dei risultati possibili di un esperimento casuale), D è detta σ-algebra, ovvero un insieme (gli eventi) per i quali si può calcolare una probabilità, e $p(x)$ è una misura di probabilità su S; precisamente:

$$p(x) : S \to [0, 1].$$

3. Un cenno alla distribuzione di variabili casuali continue

Se i valori di una ***variabile casuale*** variano con continuità in un intervallo, essa si dice ***continua***.
I parametri essenziali nella rappresentazione analitica di una variabile casuale continua sono:

1. La ***media aritmetica μ*** (che nella curva normale, essendo simmetrica, coincide con la *moda* e la *mediana*) *corrisponde all'asse di simmetria della curva e definisce la*

posizione, sull'asse delle ascisse, della curva. Cambiando la media, la curva trasla lungo l'asse *x*.

2. La ***deviazione standard σ*** (o scarto quadratico medio) *corrisponde alla distanza tra la media e il punto di flesso della curva* (dove la curva attraversa la sua tangente) *e determina l'ampiezza della curva stessa*.

Essa è una misura della dispersione dei valori della variabile casuale. In molti casi viene preferita alla varianza perché ha le stesse misure della stessa variabile. (Casolaro F. 2004).

3. La ***varianza σ²*** fornisce una misura di quanto i valori assunti dalla variabile si discostino dalla *media, per cui* la ***varianza*** è un indice di variabilità.
Data una distribuzione di una variabile X che rappresenta una popolazione di n elementi, la ***varianza è la media aritmetica dei quadrati delle distanze dei valori dalla loro media***

$$\sigma_X^2 = \frac{\sum_i (x_i - \mu_X)^2}{n}$$

dove $\mu_X = \frac{\sum_i x_i}{n}$ è la media aritmetica di X.
Quindi la varianza è una misura della diffusione o dispersione della densità di X.
In statistica viene molto spesso utilizzata la radice quadrata della varianza, cioè lo scarto quadratico medio (o *deviazione standard* o *scarto tipo*) $\sigma_X = \sqrt{\sigma_X^2}$. Con riferimento a questa notazione la varianza si trova quindi anche indicata come σ^2.
In statistica si utilizzano solitamente due stimatori per la varianza su un campione di cardinalità n:

$$S_n^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n} \qquad\qquad S_{n-1}^2 \; \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$$

Dove $\bar{x} = \frac{x_1 + x_2 + \ldots x_n}{n}$ è la media campionaria. Il primo è detto *varianza campionaria*, mentre il secondo è detto *varianza campionaria corretta* (Alexander M., e altri, 1988).
Riteniamo introdurre alcune distribuzioni facilmente fruibili per studenti della fascia che abbiamo preso in considerazione: la *distribuzione binomiale* e la *distribuzione di Poisson*.

4. La distribuzione binomiale

La distribuzione binomiale (o distribuzione di Bernoulli) rappresenta la distribuzione di probabilità di prove ripetute indipendenti quando i risultati di ciascuna prova sono solo due. Riportiamo come esempio ancora una volta il lancio di una moneta, i cui risultati possibili sono come è noto testa o croce; in tal caso la distribuzione di probabilità relativa ad n lanci è una distribuzione binomiale e mostra la probabilità che si verifichino 0 , 1 , 2 , …., n testa, oppure 0 , 1 , 2 , …., n croce. (Boggio A., Borello G., 1992).
Questa distribuzione si può utilizzare per descrivere i casi in cui gli esiti possibili di una prova possono essere ridotti solo a due.

Ad esempio:
- *la tensione erogata da una linea elettrica può essere sufficiente oppure no per un certo numero di utenti ;*
- *in un controllo di qualità un componente può risultare difettoso oppure no.*

Tale distribuzione viene quindi utilizzata quando:
- *un esperimento è costituito da più prove indipendenti;*
- *il risultato di una prova può essere classificato in due soli modi: successo o insuccesso;*
- *la probabilità di successo è costante per tutte le prove*.

In generale, per prove ripetute indipendenti dove i possibili risultati sono riconducibili a due, possiamo considerare uno dei risultati come successo e l’altro come insuccesso.

Indichiamo con X la variabile aleatoria binomiale che indica il numero di successi in n prove; X può assumere i seguenti valori 0 , 1 , 2 …., n;
indicheremo ora con

- p la probabilità di successo in ciascuna prova, costante per tutte le prove ;
- $q = 1 - p$ la probabilità di insuccesso in ciascuna prova.

Al fine di rendere più chiaro quanto esposto viene riportato qui di seguito un esempio inerente al lancio di un dado e indichiamo con successo l’uscita della faccia 1 ed insuccesso l’uscita delle altre. Pertanto le probabilità di successo ed insuccesso saranno rispettivamente

$$p = \frac{1}{6}; \quad q = \frac{5}{6}$$

Se ora effettuiamo sette lanci, possiamo sapere quale è la probabilità di avere un certo numero di successi. Supponiamo di voler calcolare la probabilità di ottenere due successi, ossia la probabilità che *X* assuma il valore 2.
Osserviamo che lo spazio degli eventi è costituito dall'insieme (S,I) dove S indica successo ed I che indica insuccesso.
A questo punto come prima cosa bisogna calcolare il numero ν delle sequenze in cui si verificano due successi .
Tale numero ν è uguale al numero delle combinazioni di n elementi (in tal caso n=7) presi a due a due, ossia

$$\nu = \frac{7!}{(5-2)!2!} = 21$$

Osserviamo che la probabilità di ottenere due successi è uguale a $\left(\frac{1}{6}\right)^2$ e quella di ottenere cinque insuccessi è uguale $\left(\frac{5}{6}\right)^5$. Pertanto la probabilità di ogni sequenza è data da :

$$p_s = \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^5$$

Visto che le sequenze di questo genere sono 21 la probabilità di ottenere quanto richiesto è pari a

$$p = \frac{7!}{(5-2)!2!} \cdot \left(\frac{1}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^5$$

Generalizzando possiamo asserire che:

$$p(X = x) = \binom{n}{x} \cdot p^x \cdot q^{n-x}$$

Quando, oltre alle condizioni inerenti alla distribuzione binomiale, si verifica che:

- la probabilità *p* di un successo tende a zero;
- il numero *n* delle prove è molto elevato (n tende ad infinito);
- il prodotto $n{\cdot}p$ è una quantità finita,

si può far uso della distribuzione di Poisson o distribuzione degli eventi rari.

5. Distribuzione di Poisson

La distribuzione di Poisson ci dà un modello realistico di molti fenomeni casuali. Poiché i valori di una variabile casuale di Poisson sono gli interi non negativi, ogni fenomeno casuale per il quale si possa fare un qualche tipo di conteggio può essere modellizzato assumendo una distribuzione di Poisson. Tale conteggio potrebbe interessare il numero di incidenti stradali mortali per settimana in una data regione, il numero di particelle radioattive emesse per unità di tempo, il numero di chiamate telefoniche ricevute ogni ora da una grande società, il numero di organismi per unità di volume di un fluido, il numero di difetti per unità di un certo materiale. Tale distribuzione è espressa dalla relazione:

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda} \qquad (3.1)$$

dove $\lambda = n \cdot p$ è il valore medio della variabile casuale binomiale X, k è il numero di volte in cui si verifica l'evento richiesto (Orsi R., 1985).

Riportiamo ora una semplice applicazione
Ad un centralino di una ditta pervengono mediamente 10 telefonate al minuto.
Calcolare la probabilità che in un minuto pervengano 6 telefonate.

Soluzione.
Dalla (3.1) si ha:

$$p_6 = \frac{\lambda^6}{k!} e^{-10} \cong \frac{10^6}{720} \cdot 4.54 \cdot 10^{-5} \cong 6.3 \cdot 10^{-2} \cong 6.3\%$$

Bibliografia

Casolaro F. (2004), Appunti estratti dal Corso di Statistica tenuto alla Facoltà di Scienze MM.FF.NN. (laurea in Scienze Ambientali) dell'Università del Sannio nel periodo 2004-2007.

Boggio A., Borrello G. (1992), "Argomenti e applicazioni di inferenza statistica e di interpolazione e regressione". Petrini Editore.

Orsi R. (1985), "Probabilità e inferenza statistica". Edizione il Mulino, Bologna.

Alexander M., (e altri, 1988), "Introduzione alla Statistica", Editore McGRAW-Hill.

Un approccio didattico all'insegnamento della Statistica

Ferdinando Casolaro[1]

[1]Università degli Studi di Napoli Federico II
Dipartimento di Architettura - Napoli, Italy,
ferdinando.casolaro@unina.it

> La concezione deterministica della natura racchiude in sé una reale causa di debolezza nell'irrimediabile contraddizione che essa incontra con i dati più certi della nostra stessa coscienza. (*The value of Statistical Laws in Physics and in Social Sciences*).
>
> Ettore Majorana

Sunto - *Si propone un percorso didattico per l'insegnamento della Statistica nella Scuola secondaria di secondo grado che tenga conto dell'evoluzione tecnologica che, di fatto, riduce negli allievi la capacità di concentrazione sugli aspetti formali e della presenza nelle classi di studenti extra-comunitari che spesso non posseggono le conoscenze linguistiche-grammaticali necessarie per un approccio esaustivo alla comprensione della disciplina. Nella seconda parte si espongono concetti che permettono di individuare approssimativamente il grafico di una funzione per l'analisi statistica indipendentemente dalla conoscenza del calcolo differenziale e si presentano esempi di curve caratteristiche per analizzare tali modelli.*

Parole chiave: istogrammi e diagrammi, funzioni esponenziali, funzioni a campana, distribuzione normale.

1. - Introduzione

Uno degli aspetti più complessi nell'insegnamento è la comunicazione degli argomenti alle persone che:

- non hanno sufficienti conoscenze pregresse (pre-requisiti).
- non conoscono la lingua del relatore.

Lo studio della Matematica e della Statistica, per gli obiettivi proposti, è facilitato rispetto alle altre discipline perché si può comunicare attraverso l'utilizzo di formule e visualizzazioni grafiche.

Spesso, però, le rappresentazioni attraverso formule o disegni vengono proposte in modo da risultare di difficile comprensione perché non si tiene conto delle conoscenze di base dei destinatari.

Analizziamo, ad esempio, un quesito assegnato agli Esami di Stato 2007 nelle Scuole italiane anche all'estero:

"*Si consideri la funzione*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \qquad (1.1)$$

Se ne spieghi l'importanza nelle applicazioni della Matematica illustrando il significato di μ, σ, σ^2 e come tali parametri influenzino il grafico di f(x)".

E' un quesito di Statistica, che chiede di analizzare la cosiddetta "*curva degli errori*" (funzione gaussiana) che caratterizza l'andamento di tutti i fenomeni naturali. Il grafico in oggetto è il seguente:

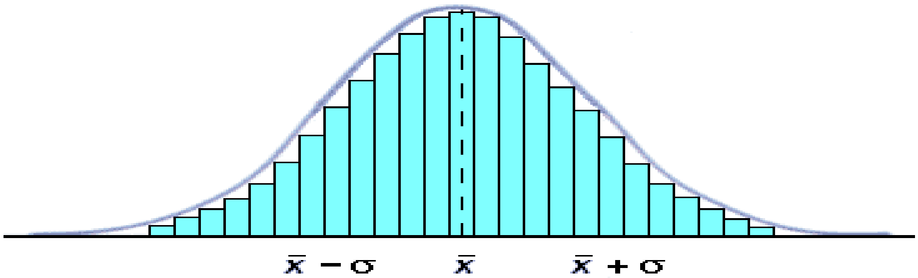


fig. 1.1

La funzione gaussiana è alla base delle analisi statistiche in ogni disciplina e la sua conoscenza permette lo studio sociologico e qualitativo di gran parte dei fenomeni attraverso la Statistica ed il Calcolo delle Probabilità, discipline essenziali per la formazione delle nuove generazioni che, a causa della velocizzazione degli eventi, vivono una realtà dominata dall'incertezza.

Questo quesito è stato molto criticato dai docenti perché ritenuto difficile ed estraneo alle indicazioni ministeriali. Per affrontarlo, l'allievo deve conoscere alcuni elementi che non si possono estrarre direttamente dalla formula, per cui riteniamo che per l'analisi dei fenomeni nelle Scienze applicate, si debba arrivare a proporre la (1.1) dopo che gli allievi abbiano maturato alcuni aspetti caratteristici dei parametri inseriti nella rappresentazione analitica.

Di seguito si richiamano alcuni concetti con i quali, partendo da semplici grafici già noti dalla Scuola secondaria di Primo grado (istogrammi e semplici rappresentazioni sul piano cartesiano), si facilita la comprensione di funzioni più complesse, come la (1.1).

2. Costruzione di semplici grafici relativi a problemi naturali

Come accennato nel paragrafo precedente, presentiamo alcune costruzioni di grafici elementari come primo approccio alla comprensione della gaussiana (Giambò A., Giambò R. (2009).

1) <u>*Con un campione di mille studenti, costruire un istogramma che presenta sull'asse delle ascisse le seguenti fasce di altezza (in centimetri):*</u>

< 150 - da 150 a 160 - da 160 a 170 - da 170 a 180 - da 180 a 190 - da 190 a 200 - > 200;

e sull'asse delle ordinate, il numero di studenti per ogni fascia.
Si vede immediatamente che la fascia centrale presenta i valori massimi, mentre gli studenti di altezza inferiore a 150 cm e gli studenti di altezza superiore a 200 centimetri presentano i valori più bassi (fig. 3.1).

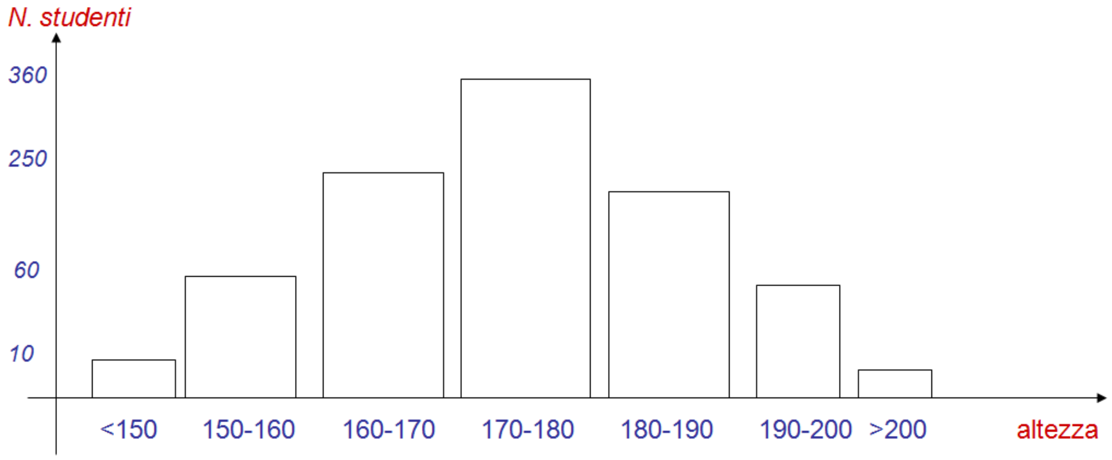


fig. 3.1

2) *Costruire l'istogramma che presenta sull'asse delle ascisse i mesi dell'anno e sulle ordinate la misura della fascia oraria compresa tra l'alba e il tramonto*.
Si hanno i seguenti dati: gennaio e dicembre: 8.00 → 17.00 - giugno-luglio: 05.00 → 21.00.
I valori massimi sono concentrati nei mesi centrali, giugno-luglio, i valori minimi agli estremi rappresentati dai mesi di gennaio e dicembre (fig. 3.2).

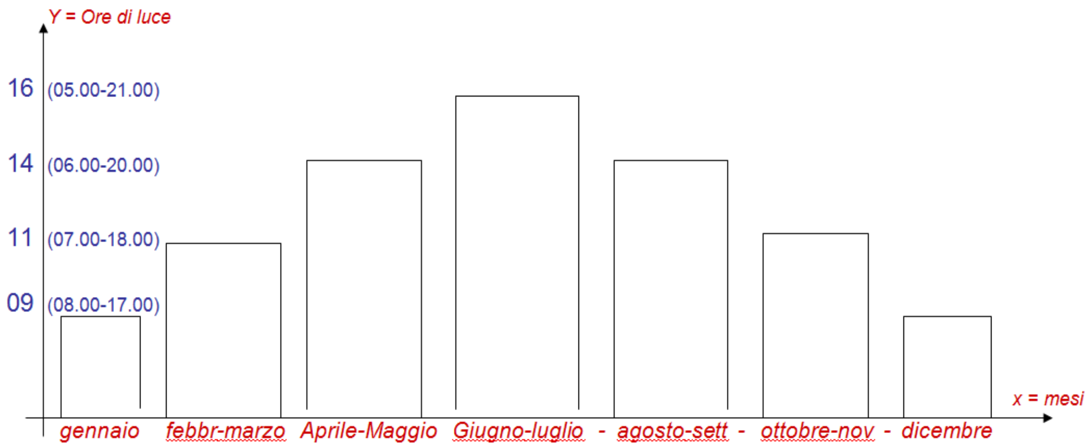


fig. 3.2

3) *Disegnare un vulcano ed indicare sulla figura i punti in cui si sviluppa la massima energia*.
E' evidente che la figura del vulcano ha l'andamento della curva di Gauss; è interessante far notare agli allievi come i valori di massima energia si hanno nella parte

centrale in cui è concentrata una massa maggiore, coerentemente al principio di equivalenza massa-energia (fig. 3.3).

fig. 3.3

Come ultimo esempio, possiamo osservare il grafico ufficiale del censimento 2001, relativo alla *distribuzione della popolazione residente in Italia per sesso e classi di età* (fig. 3.4).

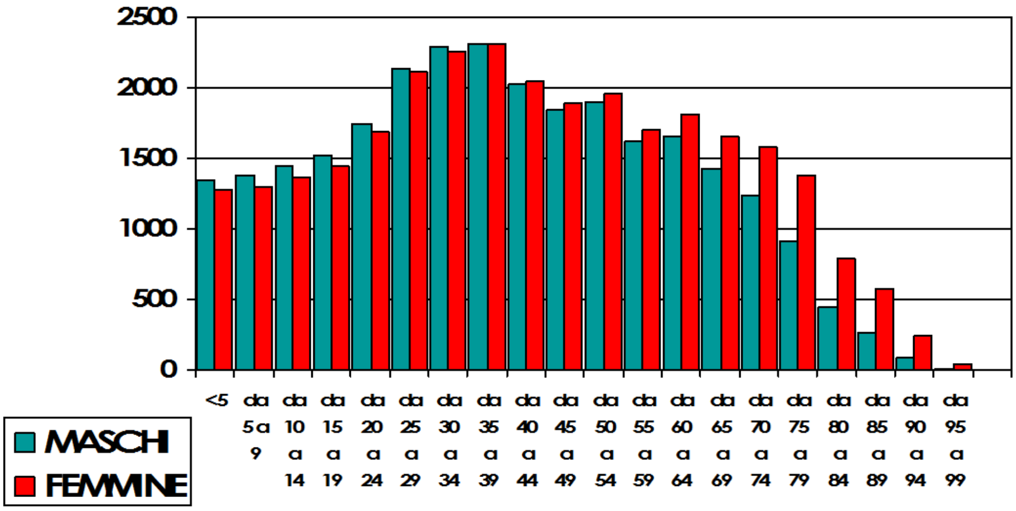


fig. 3.4

3. - Rappresentazione del grafico di una funzione. Pre-requisiti matematici.

E’ noto che per rappresentare il grafico di una funzione *f(x)* si richiede la conoscenza di:

a. equazioni, disequazioni e sistemi lineari;

b. cenni di topologia e concetto di limite;

c. concetto di derivata con relativa rappresentazione geometrica.

I concetti al punto *a.* vengono approfonditi abbastanza nelle lezioni di algebra; gli argomenti ai punti *b.* e *c.* rappresentano il nocciolo del programma di Analisi

Matematica che è oggetto di studio degli ultimi anni del corso quinquennale, per cui diventa difficile lo studio delle questioni da affrontare attraverso la Statistica nei primi anni di corso.
Ad esempio, per la crescita o la decrescita dell'andamento relativo ad un fenomeno da studiare, si opera con grandezze discrete che spesso richiedono troppi passaggi, per cui si traccia il grafico di una funzione continua per l'interpolazione con i dati a disposizione.
Come si può tracciare (almeno con un'approssimazione nei limiti di quanto basta alla comprensione del fenomeno) il grafico senza le conoscenze del calcolo differenziale?
Cerchiamo di rispondere a tale domanda pensando di proporre le seguenti tematiche nell'ordine che segue:
1. Equazione della retta: visualizzazione di equazioni e disequazioni di primo grado, con particolare riferimento ai grafici delle distribuzioni di frequenza, per meglio assimilarne l'andamento.

2. Equazione della parabola con l'asse parallelo all'asse delle ordinate: studio del segno del trinomio di secondo grado con relativa visualizzazione grafica (equazioni e disequazioni di secondo grado).

3. La funzione esponenziale e la funzione logaritmica tracciata per punti con la base 2 e generalizzazione alla base naturale *e*.

I punti 1. 2. vengono in generale trattati nei corsi del biennio per cui risultano noti agli studenti. Fissiamo, pertanto, l'attenzione sulla rappresentazione grafica delle funzioni esponenziali e logaritmiche con semplici applicazioni dell'algebra e della geometria elementare per individuarne la monotonia e la concavità.

Relativamente al punto 3. premettiamo le seguenti definizioni:

Definizione 1. La funzione $y = f(x)$, di dominio X, si dice strettamente crescente in X se

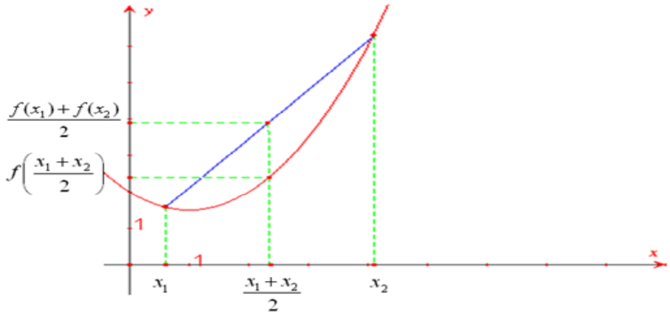
$$\forall x_1, x_2 \in X, x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$$

si dice strettamente decrescente se

$$\forall x_1, x_2 \in X, x_1 < x_2 \Rightarrow f(x_1) > f(x_2)$$

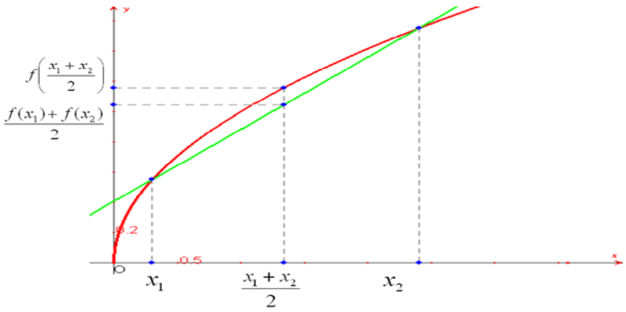
Definizione 2. *Una funzione $y = f(x)$, continua in un intervallo $[a,b] \subset R$, si dice strettamente convessa (volge la concavità verso l'alto, fig. 3.1) in tale intervallo se, $\forall x_1, x_2 \in [a, b]$, con $x_1 < x_2$, riesce:*

$$f\left(\frac{x_1+x_2}{2}\right) < \frac{f(x_1)+f(x_2)}{2}, \quad (3.1)$$



la funzione è strettamente concava (concavità verso il basso fig. 3.2) se $\forall x_1, x_2 \in [a, b]$, *con* $x_1 < x_2$, *riesce:*

$$f\left(\frac{x_1+x_2}{2}\right) > \frac{f(x_1)+f(x_2)}{2}. \qquad (3.2)$$



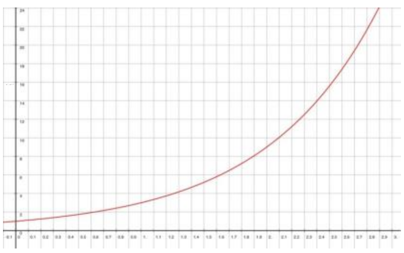
Definizione 3. *La media geometrica di due numeri reali positivi e distinti è sempre minore della media aritmetica.*

Infatti, con a >0, b > 0, risulta:

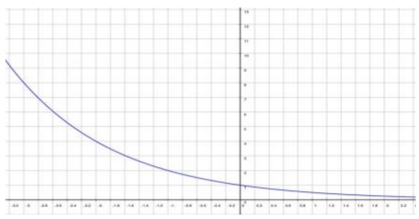
$$(a-b)^2 = a^2 - 2ab + b^2 > 0, \quad a^2 + 2ab + b^2 > 4ab, \quad \frac{(a+b)^2}{4} > a \cdot b, \frac{a+b}{2} > \sqrt{a \cdot b} \quad (3.3)$$

La funzione esponenziale *f(x) = a^x*

Teorema 1 - *La funzione* $f(x) = a^x$, $a > o$, $a \neq 1$, *è strettamente crescente se* $a > 1$, *strettamente decrescente se* $a < 1$; *è convessa (volge la concavità verso l'alto) in entrambi i casi.*



f (x) = a^x, con a > 1 - fig. 3.3



f (x) = a^x, con a < 1 - fig. 3.4

Dim. - Ci limitiamo al caso di $a > 1$. Se $x_1 < x_2$, posto: $h = x_2 - x_1$, $x_2 = x_1 + h$, si ha:

$$a^{x_2} = a^{x_1+h} = a^{x_1} \cdot a^h > a^{x_1}$$

Pertanto $f(x) = a^x$, con $a > 1$, è strettamente crescente in $[a, b]$.

Per dimostrare la concavità osserviamo che, essendo la funzione esponenziale positiva, $\forall x \in R$ risulta:

$$\forall x_1, x_2 \in [a,b], \quad a^{x_1} > 0, \quad a^{x_2} > 0 .$$

Inoltre, dalla (3.3) risulta:

$$\sqrt{a^{x_1} \cdot a^{x_2}} < \frac{a^{x_1} + a^{x_2}}{2}$$

da cui, tenendo sempre conto che a > 1, si ha:

$$\left(a^{x_1} \cdot a^{x_2}\right)^{\frac{1}{2}} = \left(a^{x_1+x_2}\right)^{\frac{1}{2}} = a^{\frac{x_1+x_2}{2}} < \frac{a^{x_1} + a^{x_2}}{2}$$

ossia:

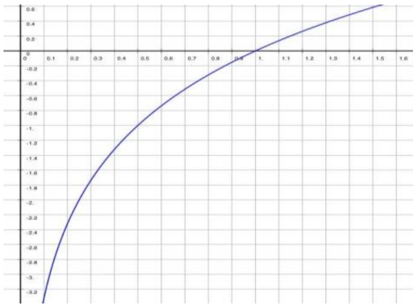
$$f\left(\frac{x_1 + x_2}{2}\right) < \frac{f(x_1) + f(x_2)}{2} .$$

per cui a^x è strettamente convessa.

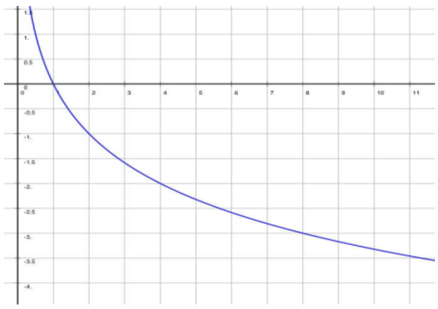
Analogamente si dimostra la proprietà per $a < 1$.

La funzione logaritmica $f(x) = log_a\, x$

Teorema 2 - *La funzione $f(x) = log_a\, x$, $a > o$, $a \neq 1$, è strettamente crescente e concava in $(0, +\infty)$ se $a > 1$, strettamente decrescente e convessa se $a < 1$*



$f(x) = log_b\, x$, con a > 1 - fig. 3.5



$f(x) = log_b\, x$, con a < 1 - fig. 3.6

Dim. - Se $a > 1$ è evidente che f(x) è strettamente crescente.

Dimostriamo la stretta concavità. Per la (3.3) e per la stretta crescenza in (0, +∞), si ha:

$$\sqrt{x_1 \cdot x_2} < \frac{x_1 + x_2}{2} \quad \Rightarrow \quad \lg_a \sqrt{x_1 \cdot x_2} < \lg_a \frac{x_1 + x_2}{2}$$

$$\frac{1}{2}\lg_a\left(x_1 \cdot x_2\right) < \lg_a \frac{x_1 + x_2}{2} \quad \text{cioè:} \quad \frac{\lg_a x_1 + \lg_a x_2}{2} < \lg_a \frac{x_1 + x_2}{2}$$

Pertanto, *f(x)* volge la concavità verso il basso in (0, +∞).

4. - Le funzioni a campana e la distribuzione normale.

Per comprendere la relazione (1.1)

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

bisogna che si chiarisca il significato dei parametri che presenta la rappresentazione analitica e l'andamento del grafico della funzione di fig. 1.1 che rientra nelle cosiddette "*funzioni a campana*" per rappresentare le "*distribuzioni normali*". Precisamente, relativamente ai parametri:

1. *la media aritmetica* μ,

2. *lo scarto quadratico medio* σ (o deviazione standard)

3. *la varianza* σ^2 [misura di quanto i valori assunti dalla variabile, si discostino dalla media]

Relativamente agli aspetti grafici, la rappresentazione della funzione $y = e^x$, trattata in modo esaustivo nel paragrafo precedente e la funzione $y = e^{-x^2}$ che rientra nelle cosiddette funzioni a campana per la *distribuzione normale*.

Pertanto, dopo aver proposto nei primi anni esempi del mondo reale attraverso semplici grafici per evidenziare come questi hanno sempre l'andamento della curva di Gauss, come indicato nel paragrafo 2 (pagg. 3-4-5) di questo articolo, è opportuno introdurre i primi elementi analitici di Statistica descrittiva e di Calcolo delle Probabilità, partendo già dal primo biennio con giochi di dadi, carte e classifiche dei campionati di calcio.

Inoltre, non riteniamo superfluo trattare in modo specifico quelle particolari curve che rappresentano la cosiddetta <*distribuzione normale*>.

Le distribuzioni normali sono una famiglia di distribuzioni che hanno le stesse caratteristiche e lo stesso andamento. Graficamente sono rappresentate da curve simmetriche rispetto ad una retta, con valori più concentrati verso il centro e meno nelle estremità laterali, che hanno un andamento di curve a campana (ma non tutte le curve a campana sono distribuzioni normali).

Una distribuzione normale può essere espressa matematicamente in funzione di due parametri: la *media* (μ) e lo *scarto tipo* (o *deviazione standard*) (fig. 4.1) come caso particolare della curva a campana

$$f(z) = e^{-z^2}$$

e del grafico della funzione

$$f(z) = k\ e^{-z^2}$$

ottenuta da:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

avendo posto:

$$k = \frac{1}{\sigma\sqrt{2\pi}} \qquad z = \frac{x-\mu}{\sqrt{2}\ \sigma}$$

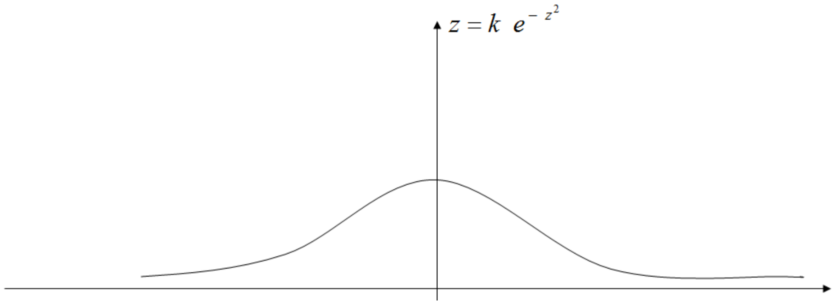


fig. 4.1

I parametri essenziali sono:

La ***media aritmetica μ*** (nella curva normale, essendo simmetrica, coincide con la *moda* e la *mediana*) che corrisponde all'asse di simmetria della curva, di cui ne caratterizza la posizione sull'asse delle ascisse. Cambiando la media la curva trasla lungo l'asse *x*.

La **deviazione standard σ** (o scarto quadratico medio) che corrisponde alla distanza tra l'asse di simmetria (la media) e il punto di flesso della curva (dove la curva attraversa la sua tangente) e determina l'ampiezza della curva stessa.

La varianza σ^2 che fornisce una misura di quanto i valori assunti dalla variabile si discostino dalla *media*. Precisamente, misura la dispersione dei dati intorno al valore atteso.

In fig. 4.2 sono rappresentati grafici di distribuzione normale in cui si può notare come le curve normali differiscano per il modo in cui i valori si distribuiscono.

In fig. 4.3 si può osservare come, al variare della media e della varianza, la curva subisca sia uno spostamento sull'asse delle ascisse, sia un appiattimento; se si fa variare solo la varianza e si tiene costante la media, la curva si appiattisce quando la varianza cresce e diventa più appuntita quando la varianza cala, mentre il centro rimane lo stesso.

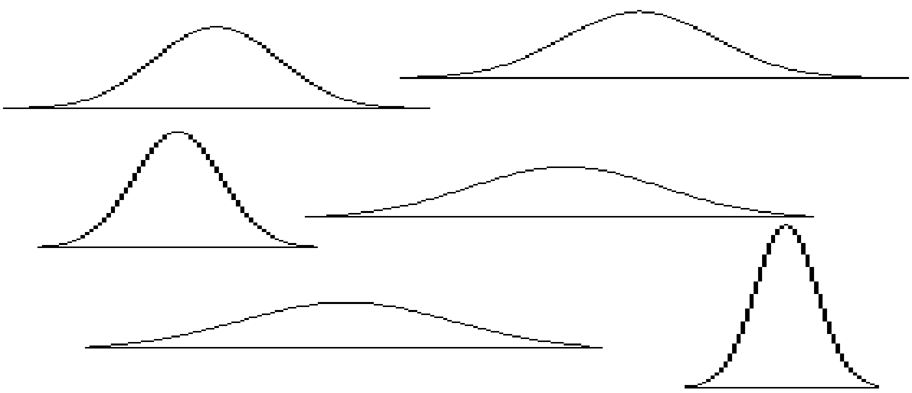


fig. 4.2

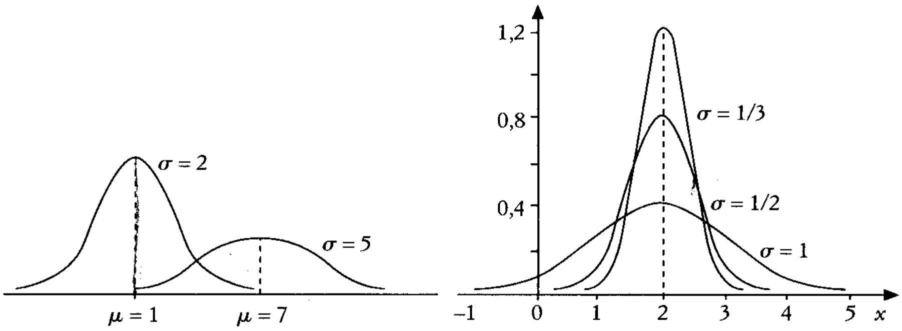


fig. 4.3

Per calcolare le probabilità che una variabile casuale X assuma valori compresi all'interno di intervalli della retta reale, si utilizza la ***distribuzione normale standardizzata***, la cui importanza sta nel fatto che le probabilità corrispondenti alle superfici racchiuse dalla curva normale sono state tabulate e vengono riportate in apposite tabelle, per cui possono essere determinate senza dover ricorrere al calcolo di integrali.

La distribuzione normale *standardizzata* presenta le stesse caratteristiche della distribuzione normale *non standardizzata*. Ciò che distingue le due distribuzioni è che la normale standardizzata ha Media = 0 e Deviazione standard = 1, per cui è rappresentata da una sola curva (fig. 4.4), mentre la distribuzione normale generale è costituita da infinite curve.

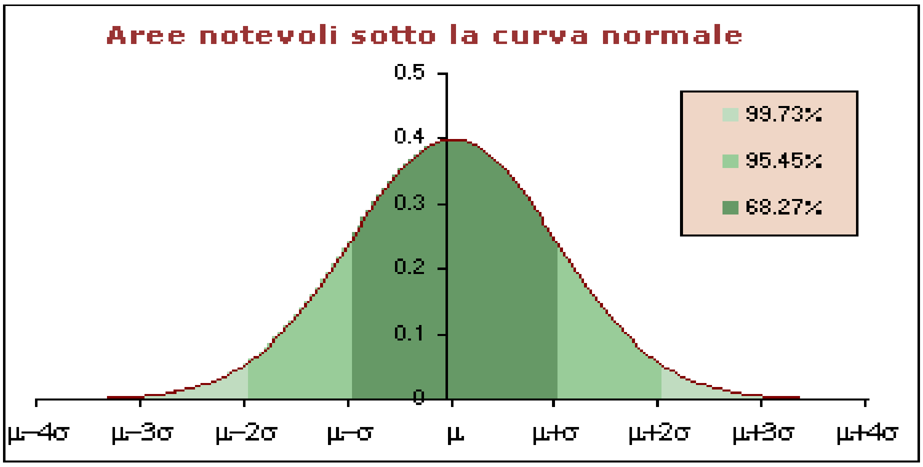


fig. 4.4

5. Conclusioni

Come si può constatare, l'esposizione di questo lavoro parte da esempi estremamente semplici che addirittura si possono proporre nella Scuola del Primo Ciclo. Si passa poi a rappresentare concetti che richiedono conoscenze più avanzate.

In realtà, è questo l'approccio che gli studenti hanno relativamente allo studio di altre branche della Matematica, come l'algebra e la geometria che nel secondo Ciclo di studi vengono proposte quando essi hanno maturato, già dai primi anni della Scuola Primaria, i concetti fondamentali di aritmetica e di rappresentazione di figure nel piano e nello spazio.

La Matematica dell'incerto, invece, viene catapultata senza alcuna progettazione in quanto la maggior parte dei docenti non ha avuto una adeguata Formazione nei corsi universitari, per cui l'obiettivo della nostra proposta è quello di razionalizzare l'intero percorso, in modo che il docente possa spalmare con gradualità i vari argomenti anno dopo anno in base alle difficoltà crescenti.

Bibliografia

Casolaro F. (1990), "*Le funzioni elementari senza l'utilizzo dell'analisi matematica*" - Convegno nazionale Mathesis 1990, Ed. Luciani - Iseo 1990.

Casolaro F. (2004), Appunti estratti dal Corso di Statistica tenuto alla Facoltà di Scienze MM.FF.NN. (laurea in Scienze Ambientali) dell'Università del Sannio nel periodo 2004-2007.

Giambò A., Giambò R. (2009), Matematica per la Scuola Superiore, Ed. Armando Scuola.

Maturo A., Squillante M., A.G. Ventre A.G.(2010). Coherence for Fuzzy Measures and Applications to Decision Making. Vol. 257, Springer-Verlag, Berlin Heidelberg 2010.

Casolaro F., Prosperi R. (2011), “Atti Terni 2011: La Matematica per la Scuola di 2° secondo grado: un contributo per il docente di Matematica”. Ed. 2C Contact.

Prosperi R. (2011), La Matematica dell'incerto: l'insegnamento del “Calcolo delle Probabilità” e della “Statistica” nel I biennio della Scuola Secondaria di II grado. Ed. 2C Contact 2011.

Casolaro F. - Paladino L. (2012): “*Analisi sociale e rigore scientifico. Scelta di equilibrio per l'ottimizzazione dei risultati nell'insegnamento della matematica*” - Logica, linguaggio e didattica della Matematica, a cura della facoltà di Scienze della Formazione dell'Università di Salerno, Edizione Franco Angeli, 2012.

Indizi, Pre-giudizi e il Ragionamento Bayesiano

Aniello Buonocore [1]

[1]Università degli Studi di Napoli Federico II
Dipartimento di Matematica e Applicazioni "Renato Caccioppoli"
Via Cintia, Complesso Monte S. Angelo, Napoli, Italy,
E-mail: aniello.buonocore@unina.it

Sunto
Attraverso tre diverse situazioni si illustrano gli ingredienti basilari e "il motore inferenziale" per prendere decisioni in condizioni di incertezza. Tale modalità di ragionamento può essere d'aiuto agli insegnanti per sollecitare gli studenti sui processi metacognitivi molto spesso richiamati nelle indicazioni nazionali.

Parole chiave: Probabilità congiunta e condizionata, verosimiglianza, probabilità a priori, probabilità a posteriori.

1. Dalle leggere alle pesanti?

Nel mese di dicembre del 1992 i professori Carlo Rossi e Romano Scozzafava inviano ad alcuni dei maggiori quotidiani nazionali una lettera dal provocatorio titolo "Dalle leggere alle pesanti?". Come si intuisce abbastanza facilmente essi intendono "intromettersi" nel dibattito, in quel frangente molto controverso della dialettica politico-sociale, avvertito di estrema importanza da parte dell'opinione pubblica sulla punibilità e sul recupero dei tossicodipendenti. Ovviamente, la loro finalità era "solo" quella di smascherare una fallacia nel ragionamento "*dalle leggere alle pesanti*":
Gli aspetti emotivi prevalgono su quelli razionali, si indaga sulle storie individuali e si interroga il consumatore di droghe per sapere, in particolare, se fumava spinelli (cioè le cosiddette droghe "leggere", come hashish e marijuana) prima di diventare tossicodipendente. La risposta è di solito affermativa, e nei dibattiti questo fatto viene regolarmente usato per sostenere la necessità di proibire anche queste sostanze, che sarebbero la "porta d'ingresso" alla tossicodipendenza.
Ora, se anche uno volesse fare l'avvocato del diavolo per dire che i fautori del proibizionismo usassero l'argomento in maniera strumentale, è stupefacente constatare come esso facesse (e fa ancora) molta presa sugli ascoltatori delle trasmissioni televisive e sui lettori dei giornali. L'argomentazione con la quale i professori Rossi e Scozzafava demoliscono il ragionamento "dalle leggere alle pesanti" è di tipo paradossale: se alle stesse persone intervistate fosse stata posta la domanda "ti piaceva mangiare le

caramelle da piccolo?" la percentuale di risposte affermative certamente sarebbe stata maggiore di quella ottenuta con la domanda sul consumo delle droghe leggere. Ma nessuno si sogna di mettere in relazione caramelle e droghe.

Alla fine della lettera gli autori forniscono una spiegazione più tecnica per la non validità del ragionamento "dalle leggere alle pesanti":

In conclusione, quindi, è fra i consumatori di spinelli che andrebbe fatta un'indagine per vedere quanti diventeranno consumatori di droghe "pesanti", e non viceversa: ma è stata mai fatta una tale indagine?

Da parte mia, faccio osservare che il disegno sperimentale (la procedura per ottenere i dati), che è del tipo "tempo corrente" - "qualche anno prima", contribuisce molto a ingenerare la convinzione della validità del ragionamento "dalle leggere alle pesanti": è agli attuali tossicodipendenti che bisogna chiedere se prima fumavano gli spinelli. È invece molto più difficile far accettare la modalità di raccolta dei dati proposta dagli autori: è agli attuali fumatori di spinelli che tra qualche anno bisognerà chiedere se sono diventati tossicodipendenti.

Ma più avanti farò vedere che ciò non è necessario.

Nel linguaggio che ho utilizzato nel titolo si potrebbe dire che sono stati confusi (volutamente o meno) gli *indizi* con le *conclusioni del ragionamento* ovvero, nel linguaggio bayesiano, le *verosimiglianze* (ovvero le probabilità tratte dai dati) con le *probabilità a posteriori*.[1]

In effetti, indicata con Ω la popolazione *obiettivo* (ben individuata e circoscritta prima dell'avvio dell'indagine) la tesi che i fautori del protezionismo vorrebbero dimostrare è che la percentuale dei tossicodipendenti (evento T) nella parte di Ω costituita dalle sole persone che prima di diventare tossicodipendenti "fumavano gli spinelli" (evento S) fosse molto più grande di quella valutata sulla totalità della popolazione Ω e rasentasse l'unità.[2] In simboli:

$$\mathbb{P}[T] \ll \mathbb{P}_S[T] \cong 1. \qquad (1)$$

Invece, il loro ragionamento è quello di "produrre" solo il risultato dell'indagine da loro individuata, ovvero la percentuale $\mathbb{P}_T[S]$ e affermare che tale quantità è "impressionantemente" grande (ovvero, abbastanza vicino a 1).

A mio avviso, c'è ancora un aspetto da sottolineare nel ragionamento fallace: la mancanza assoluta dei pre-giudizi che nel linguaggio bayesiano rappresentano le *probabilità a priori*. In altri termini, in esso, la percentuale dei tossicodipendenti nella totalità della popolazione Ω non gioca alcun ruolo. In termini più generali, nel

[1] Qui, utilizzo il concetto di probabilità nel senso frequentista o statistico: la probabilità di un evento è il rapporto tra il numero dei dati che fanno verificare l'evento e la numerosità dei dati; in tal modo, se la numerosità dei dati è sufficientemente grande, è lecito utilizzare indifferentemente la parola percentuale e la parola probabilità. Inoltre, utilizzo la lettera $\mathbb{P}$ (da leggere P contornata) per indicare la probabilità dell'evento racchiuso tra due parentesi quadrate. Infine, il pedice, sia esso A, che qualche volta è posto alla lettera $\mathbb{P}$ rappresenta la riduzione della popolazione a quella individuata dall'evento A.

[2] In contesti di questo tipo mi sembra appropriato interpretare la frase "x molto maggiore di y" se x è almeno di un ordine grandezza più grande di y.

ragionamento fallace è trascurata quella che Julia Galef asserisce essere la regola numero uno per chi voglia ragionare in modo bayesiano: tenere bene in mente la propria valutazione delle probabilità a priori.[3]

In maniera formale le cose stanno così. Indicati con A e B due qualsiasi eventi, la formula della probabilità congiunte asserisce che:

$$\mathbb{P}[B] \cdot \mathbb{P}_B[A] = \mathbb{P}[A \cap B] = \mathbb{P}[A] \cdot \mathbb{P}_A[B]. \tag{2}$$

La (2) si interpreta facilmente: la probabilità che due eventi si verifichino congiuntamente è uguale alla probabilità di uno qualsiasi di essi moltiplicata per la probabilità dell'altro valutata una volta che si è ridotta la popolazione di partenza in maniera conforme. Dalla (2) è facile ricavare la formula di Bayes:

$$\mathbb{P}_B[A] = \mathbb{P}[A] \cdot \frac{\mathbb{P}_A[B]}{\mathbb{P}[B]}. \tag{3}$$

Rientrando nel contesto del ragionamento "*dalle leggere alle pesanti"* la (3) diventa:

$$\mathbb{P}_S[T] = \mathbb{P}[T] \cdot \frac{\mathbb{P}_T[S]}{\mathbb{P}[S]}. \tag{4}$$

La (4) asserisce che, la percentuale dei tossicodipendenti tra i fumatori di spinello è proporzionale alla percentuale dei tossicodipendenti nella totalità della popolazione Ω; il fattore di proporzionalità α è il rapporto della percentuale dei fumatori di spinelli tra i tossicodipendenti e la percentuale dei fumatori di spinello nella totalità della popolazione Ω.

Allora, dal momento che i tossicodipendenti censiti annualmente dai "SerT" e "SerD" sono dell'ordine di poche centinaia di migliaia su una popolazione di qualche decina di milioni di persone, se anche ponessi $\alpha = 3$, si avrebbe che la (1) è sbagliata sia nel senso del molto minore che nel senso del circa uguale.[4]

In effetti, si può raggiungere la conclusione del ragionamento senza passare attraverso una valutazione quantitativa della percentuale dei fumatori di spinello nella totalità della popolazione Ω. Basta applicare la (3) all'evento $\bar{T}$ e rapportare membro a membro la (3) con l'equazione così ottenuta. In tal modo, si ottiene:

$$\frac{\mathbb{P}_S[T]}{\mathbb{P}_S[\bar{T}]} = \frac{\mathbb{P}[T]}{\mathbb{P}[\bar{T}]} \cdot \frac{\mathbb{P}_T[S]}{\mathbb{P}_{\bar{T}}[S]}. \tag{5}$$

Quindi, tra i fumatori di spinelli il rapporto tra la percentuale dei tossicodipendenti e quella dei non tossicodipendenti (rapporto delle probabilità a posteriori) è proporzionale al rapporto tra le percentuali degli stessi eventi valutate nella totalità della popolazione Ω (rapporto delle probabilità a priori); il fattore di proporzionalità è uguale al rapporto tra la verosimiglianza di fumare gli spinelli essendo tossicodipendente e la

[3] Julia Galef, laureata in Statistica alla Columbia University, è co-fondatrice del "Center for Applied Rationality", una organizzazione non-profit con sede a Berkeley, che organizza workshop sul miglioramento del ragionamento e del processo decisionale.

[4] I Servizi per le Tossicodipendenze (SerT), o Servizi per le Dipendenze patologiche (SerD), sono i servizi pubblici del Sistema Sanitario Nazionale italiano, dedicati alla cura, alla prevenzione ed alla riabilitazione delle persone che hanno problemi conseguenti all'abuso ed alla dipendenza di sostanze psicoattive come droghe o comportamenti compulsivi come il gioco d'azzardo patologico (Wikipedia).

verosimiglianza di fumare gli spinelli essendo non tossicodipendente. La (5) mette in risalto che in effetti la popolazione Ω è ripartita nei due eventi incompatibili T e $\overline{T}$ e che quello che è rilevante ai fini del ragionamento bayesiano (o induttivo) è la considerazione dell'evento S che nel suo primo membro si suppone verificato.

2. Salomone e le due meretrici

Mi sposto ora nel tempo e nella latitudine per considerare una situazione riportata nella Bibbia e riguardante il re Salomone in uno dei tanti episodi riguardanti la sapiente amministrazione della giustizia nel suo regno.[5]

In quel tempo vennero due donne meretrice al re e si presentarono dinanzi a lui. Una di esse disse: "Ascoltami, te ne prego, o mio signore; io e questa donna abitavamo nella medesima casa, e io partorii presso di essa nella stessa stanza. Tre giorni dopo che io ebbi partorito, anche costei ebbe un figliolo, e stavamo insieme, <u>e non vi era altri con noi nella casa all'infuori di noi due</u>. Ora morì il figliolo di questa donna durante la notte, avendolo essa soffocato mentre dormiva. Levatosi allora nel cuor della notte, di nascosto tolse il mio figlio dal fianco della tua ancella, che dormiva, e se lo collocò sul suo seno, mentre il suo figlio, che era morto, lo pose sul mio seno.

Il mattino nell'alzarmi per dare il latte al figliol mio, lo vidi morto: ma avendo guardato con maggior diligenza alla luce del giorno, m'accorsi che non era quello che io avevo generato". L'altra donna rispose: "Non è vero quanto tu dici, ma il figlio tuo è morto; il mio vive". Al contrario l'altra diceva: "Tu menti poiché il mio figlio vive e il tuo è morto"; e così litigavano alla presenza del re.

Allora il re disse: "<u>Una dice: Il mio figlio vive e il tuo figlio è morto. E l'altra risponde: "No, ma è il figlio tuo che è morto, il mio vive</u>". E il re continuò: "Portatemi una spada". Quando ebbero portato la spada davanti al re, egli soggiunse: "<u>Dividete il bambino in due parti e datene una metà all'una e una metà all'altra</u>". La donna, madre del figlio vivo, siccome si sentì commuovere le viscere per amore del proprio figliolo, disse al re: "Te ne scongiuro, o signore, dà a lei il bambino vivo e non volerlo uccidere". Al contrario l'altra diceva: "Non sia né mio, né tuo, ma sia diviso". Rispose allora il re e disse: "Date a costei il bambino vivo e non si uccida, poiché costei è la vera madre".

Tutta Israele seppe del giudizio pronunciato dal re e temette il re, vedendo che la sapienza di Dio era in lui per amministrare la giustizia (Bibbia, 1° Re 3).

Per formalizzare il ragionamento di Salomone indico con D l'evento: "la prima donna che si è rivolta al re è la madre del bambino vivo"; la prima frase sottolineata nel racconto comporta che la popolazione (in questo caso l'evento certo) è partizionata da tale evento e dal suo negato $\overline{D}$: "la prima donna che si è rivolta al re non è la madre del

[5] L'episodio è riportato anche in (Rossi, C. 1999); il lettore potrà ivi trovare sia approfondimenti relativi ai contenuti che ulteriori esempi di applicazione del ragionamento bayesiano.

bambino vivo”. La seconda frase sottolineata è la valutazione delle probabilità a priori, ovvero la probabilità dei precedenti pre-giudizi, effettuata dal re: non ha altra scelta che assegnare uguale probabilità a D e a $\overline{D}$. Con il colpo di teatro della spada, ovvero la terza frase sottolineata, Salomone si fornisce da se stesso l’indizio: indico con U l’evento: “si uccida il bambino rimasto in vita”. Dalla (3), ricordando che prima dell’indizio gli eventi D e $\overline{D}$ sono equiprobabili, ottengo:

$$\mathbb{P}_{\overline{U}}[D] = \frac{\mathbb{P}[D] \cdot \mathbb{P}_D[\overline{U}]}{\mathbb{P}[\overline{U}]} = \frac{\mathbb{P}[D] \cdot \mathbb{P}_D[\overline{U}]}{\mathbb{P}[D \cap \overline{U}] + \mathbb{P}[\overline{D} \cap \overline{U}]}$$
$$= \frac{\mathbb{P}[D] \cdot \mathbb{P}_D[\overline{U}]}{\mathbb{P}[D] \cdot \mathbb{P}_D[\overline{U}] + \mathbb{P}[\overline{D}] \cdot \mathbb{P}_{\overline{D}}[\overline{U}]} = \frac{\mathbb{P}_D[\overline{U}]}{\mathbb{P}_D[\overline{U}] + \mathbb{P}_{\overline{D}}[\overline{U}]}. \tag{6}$$

Per completare l’analisi ho bisogno solo di interpretare come il re ha valutato le verosimiglianze dell’indizio e dico che:

$$\mathbb{P}_D[\overline{U}] = 1 \quad \text{e} \quad \mathbb{P}_{\overline{D}}[\overline{U}] = 0.$$

Infatti, solo se la donna è la madre del bambino vivo allora non teme di contrastare l’autorità del re contestando la soluzione da lui prospettata.
In definitiva:

$$\mathbb{P}_{\overline{U}}[D] = \frac{\mathbb{P}_D[\overline{U}]}{\mathbb{P}_D[\overline{U}] + \mathbb{P}_{\overline{D}}[\overline{U}]} = \frac{1}{1+0} = 1. \tag{7}$$

3. Il dilemma di Monty Hall

Alla fine di questa illustrazione del ragionamento bayesiano, ritorno alla nostra quotidianità per considerare una situazione nella quale l’evento certo è partizionato in tre eventi. Nel programma televisivo americano degli anni settanta “*Let’s Make a Deal*” il conduttore Monty Hall presentava al concorrente di turno tre porte, informandolo che una di queste nascondeva un ricco premio mentre le altre due avrebbero rivelato solo premi scherzosi o di consolazione. Dopo che il concorrente aveva effettuato una prima scelta sulle tre porte Monty Hall ne apriva un’altra mostrando che essa non celava l’ambìto premio (nel seguito PREMIO). A questo punto veniva offerta al concorrente la possibilità di confermare la sua scelta o di cambiarla a favore dell’altra porta rimasta chiusa. La questione allora è la seguente: al concorrente conveniva confermare la sua scelta iniziale o cambiarla? In un problema così impostato, la risposta corretta è che al concorrente conviene sempre cambiare la sua scelta originale, ma ciò appare fortemente controintuitivo e difatti la quasi totalità delle persone, anche quelle che possono vantare una solida formazione statistica, preferisce confermare la propria scelta dato che nella fase finale del gioco attribuisce la stessa probabilità di celare il PREMIO ad entrambe le porte rimaste chiuse.

Ma le cose in realtà non stanno così e con l'aiuto dei riferimenti sopra introdotti voglio illustrarne la ragione.[6]
La situazione prima della scelta finale del concorrente è allora così schematizzabile: il concorrente sceglie la prima porta e il conduttore apre la terza porta. Se così non fosse basterebbe permutare il numero delle porte.
Considero ora gli eventi:

$$i = 1, 2, 3, \qquad \begin{cases} A_i & \text{“il PREMIO è dietro la porta } i\text{”} \\ C_i & \text{“il conduttore apre la porta } i\text{”} \end{cases}. \tag{8}$$

La prima scelta del concorrente è ininfluente e serve solo per fissare le idee: l'indizio allora è costituito dall'evento C_3.
Per quanto riguarda i pre-giudizi, osservo che essi sono costituiti dai tre eventi A_1, A_2 e A_3 che partizionano l'evento certo: per la regola numero uno di Julia Galef bisogna assegnare loro una probabilità. A questo punto ipotizzo che Monty Hall, non utilizzi una strategia deterministica per distribuire il premio tra le porte e si affidi al caso. Ad esempio, si serve di un dado non truccato e, lanciandolo, se esce un punteggio minore di 2 allora per lui si è verificato A_1, se invece esce un punteggio maggiore di 4 allora per lui si è verificato A_3, altrimenti per lui si è verificato A_2. Pertanto, a priori assegno uguale probabilità ai tre eventi.
Per le conclusioni del ragionamento bayesiano scelgo la situazione “il concorrente vince il PREMIO cambiando la scelta iniziale”; con ovvio significato dei simboli, dalla (3) segue:

$$\begin{aligned} \mathbb{P}_{C_3}[A_2] &= \frac{\mathbb{P}[A_2] \cdot \mathbb{P}_{A_2}[C_3]}{\mathbb{P}[C_3]} \\ &= \frac{\mathbb{P}[A_2] \cdot \mathbb{P}_{A_2}[C_3]}{\mathbb{P}[A_1 \cap C_3] + \mathbb{P}[A_2 \cap C_3] + \mathbb{P}[A_3 \cap C_3]} \\ &= \frac{\mathbb{P}[A_2] \cdot \mathbb{P}_{A_2}[C_3]}{\mathbb{P}[A_1] \cdot \mathbb{P}_{A_1}[C_3] + \mathbb{P}[A_2] \cdot \mathbb{P}_{A_2}[C_3] + \mathbb{P}[A_3] \cdot \mathbb{P}_{A_3}[C_3]} \\ &= \frac{\mathbb{P}_{A_2}[C_3]}{\mathbb{P}_{A_1}[C_3] + \mathbb{P}_{A_2}[C_3] + \mathbb{P}_{A_3}[C_3]}. \end{aligned} \tag{9}$$

La (9) mostra che la probabilità a posteriori dell'evento A_2 dipende solo dalle verosimiglianze; anche in questa situazione non è necessario ricorrere a dati sperimentali per la loro valutazione.
Infatti,

$$\mathbb{P}_{A_2}[C_3] = 1,$$

in quanto il concorrente ha scelto la prima porta e il premio è dietro la seconda porta: il conduttore deve necessariamente aprire la terza porta.
Inoltre,

---

[6] Un'analisi approfondita è riportata in (Gillmann L. 1992).

$$\mathbb{P}_{A_3}[C_3] = 0,$$

in quanto il concorrente ha scelto la prima porta e il premio è dietro la terza porta: il conduttore non può aprire la terza porta.

Un po' più laboriosa è la valutazione della verosimiglianza $\mathbb{P}_{A_1}[C_3]$: il concorrente ha scelto la prima porta che nasconde il PREMIO e pertanto il conduttore può, a propria scelta, aprire la seconda o la terza porta. Anche qui, per evitare di dare vantaggi a eventuali telespettatori che registrassero le sue scelte nel corso delle varie trasmissioni, ipotizzo che Monty Hall si affidi al caso (ad esempio, lanciando nuovamente il suo dado onesto) e quindi,

$$\mathbb{P}_{A_1}[C_3] = \frac{1}{2}.$$

In definitiva:

$$\mathbb{P}_{C_3}[A_2] = \frac{\mathbb{P}_{A_2}[C_3]}{\mathbb{P}_{A_1}[C_3] + \mathbb{P}_{A_2}[C_3] + \mathbb{P}_{A_3}[C_3]} = \frac{1}{1/2 + 1 + 0} = \frac{2}{3}. \qquad (10)$$

Così, con i pre-giudizi equiprobabili, l'indizio "Monty Hall apre la terza porta" sbilancia la probabilità di vincita del PREMIO a favore della seconda porta!

4. Conclusioni

Come conclusione, provo a modificare un'affermazione di Gerd Gigerenzer che con poco impegno si può reperire nella sua forma originaria.[7]

Ai nostri figli viene insegnata l'algebra, la geometria e il calcolo infinitesimale: in altre parole, la matematica della certezza. È, invece, un po' trascurata quella dell'incertezza, ovvero il pensiero statistico. Dietro questa impostazione c'è l'idea che esercitarsi in discipline astratte migliora la qualità del pensiero e la capacità di risolvere problemi. Però, purtroppo abbiamo ancora molti medici che non comprendono le statistiche sanitarie e molti avvocati che non capiscono appieno la prova del DNA… Allora, ai nostri studenti bisogna insegnare parallelamente sia quello che è importante in matematica e sia quello che potrà utile per la vita quotidiana e per la professione. Ma l'aspetto che forse è maggiormente da prendere in considerazione è la modalità dell'insegnamento: quella orientata alla soluzione di problemi richiede in maniera precipua presentazioni chiare, strumenti pratici e linee guida intelligenti. Ciò è tanto più valido per la matematica dell'incerto e credo che il ragionamento bayesiano, sollecitando oltretutto l'insegnante a prendersi carico degli standard di processo

[7] Gerd Gigerenzer è uno psicologo tedesco che ricopre le cariche di direttore emerito del "Center for Adaptive Behavior and Cognition (ABC)" presso il Max Planck Institute for Human Development e direttore del "Harding Center for Risk Literacy" entrambi in Berlino. È autore del libro "Reckoning with Risk: Learning to Live with Uncertainty", Penguin Books, 2003.

individuati dalle indicazioni nazionali, possa avere un ruolo importante a questo riguardo.
Però, per realizzare cose così nuove bisogna ricercare con convinzione l'adesione degli insegnanti.

Bibliografia

Rossi, C. 1999: "*La matematica dell'incertezza – Didattica della probabilità e della statistica*", Zanichelli, pp. 281-282.

Gillman, L. 1992: "The Car and the Goats", *American Mathematical Monthly*, **99**, 3-7.

Correlazione e Regressione, per problemi[*]

Luciano Corso[1]

[1]Direttore della rivista MatematicaMente
Presidente della sezione di Verona della Mathesis
email : lcorso@iol.it

Due variabili statistiche X e Y, su base sperimentale, hanno presentato la seguente distribuzione delle frequenze (va inteso che la coppia (x=5, y=4) è stata osservata 3 volte):

		x					
		1	2	3	5	6	8
y	1	2	1	0	0	0	0
	3	0	1	4	1	0	0
	4	0	0	1	3	1	0
	6	0	0	0	0	0	2

1) Si chiede se i due caratteri sono statisticamente indipendenti.
2) Si verifichi l'ipotesi di indipendenza statistica tra i due caratteri al livello di significa-tività del α = 5% .
3) Si calcolino medie aritmetiche e varianze totali e si valuti quale delle due variabili ha una maggiore variabilità.
4) Si determini la media aritmetica di x condizionata da y = 4 e la media aritmetica di y condizionata da x = 2.
5) Si calcolino la covarianza e il coefficiente di correlazione lineare.
6) Si determini la retta (interpolante) di regressione.
7) Si valuti con un opportuno indice la bontà dell'accostamento fatto tra fenomeno osser-vato e modello teorico.
8) Si tracci il grafico del fenomeno presentato in tabella e del modello teorico interpolato.

L'obiettivo della verifica è di valutare se uno studente possiede i concetti di indipendenza statistica di due variabili, di verifica delle ipotesi (inferenza statistica), di correlazione lineare tra due variabili, di dipendenza lineare e di regressione, di bontà di un acco-stamento tra dati sperimentali e modelli teorici, di medie totali e condizionate e di misure comparabili della variabilità.

Risoluzione

La numerosità del campione è n=16. Fissiamo come contatore di colonna j (2° pedice) e come contatore di riga i (1° pedice); le righe sono r = 4 e le colonne sono c = 6.
1) Le frequenze assolute marginali di X sono: ($n_{.1}$=2, $n_{.2}$=2, $n_{.3}$= 5, $n_{.4}$=4, $n_{.5}$=1, $n_{.6}$=2). Le frequenze assolute marginali di Y sono: ($n_{1.}$=3, $n_{2.}$=6, $n_{3.}$=5, $n_{4.}$=2). Applicando il teorema delle probabilità composte per eventi statisticamente indipendenti otteniamo le frequenze relative teoriche in ipotesi di indipendenza statistica delle variabili x e y:

$$\text{Prob}\left(x_{.j} \cap y_{i.}\right) = \text{Prob}\left(x_{.j}\right) \cdot \text{Prob}(y_{i.}),\ \forall i, j. \tag{1}$$

Per i=1 e j=1 si ha $\text{Prob}(x_{.1} \cap y_{1.}) = 6/256$ e così via fino a $\text{Prob}(x_{.6} \cap y_{4.}) = 4/256$. Comparando questi valori con quelli osservati si conclude che x e y non sono statisticamente indipendenti.
2) Per verificare se questa dipendenza è o no casuale si applica il test delle ipotesi sulla statistica

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} [\left(n_{i,j} - \hat{n}_{i,j}\right)^2 / \hat{n}_{i,j}] \tag{2}$$

dove $n_{i,j}$ e $\hat{n}_{i,j}$ sono rispettivamente le frequenze assolute osservate e teoriche in ipotesi di indipendenza; dal calcolo (r = 4, c = 6), dove r è il numero di righe e c il numero delle colonne della tabella, risulta che $\chi^2 \cong 34.9067$.

Si dimostra che χ^2 ha una distribuzione di probabilità del tipo Gamma: $\text{G}(\chi^2 |\, \lambda = 1/2, \upsilon/2)$. Presentiamo la densità di probabilità e il grafico della distribuzione della varia-bile aleatoria χ^2

$$\text{g}\left(\chi^2 \middle| \lambda = \frac{1}{2},\ \frac{\upsilon}{2}\right) = \begin{cases} \dfrac{2^{-\frac{\upsilon}{2}} \cdot e^{-\lambda \cdot \chi^2} \cdot (\chi^2)^{\left(\frac{\upsilon}{2}-1\right)}}{\Gamma\left(\frac{\upsilon}{2}\right)}, & \{\chi^2 | \chi^2 \in \mathbb{R}^+\} \\ 0 & altrove \end{cases} \tag{3}$$

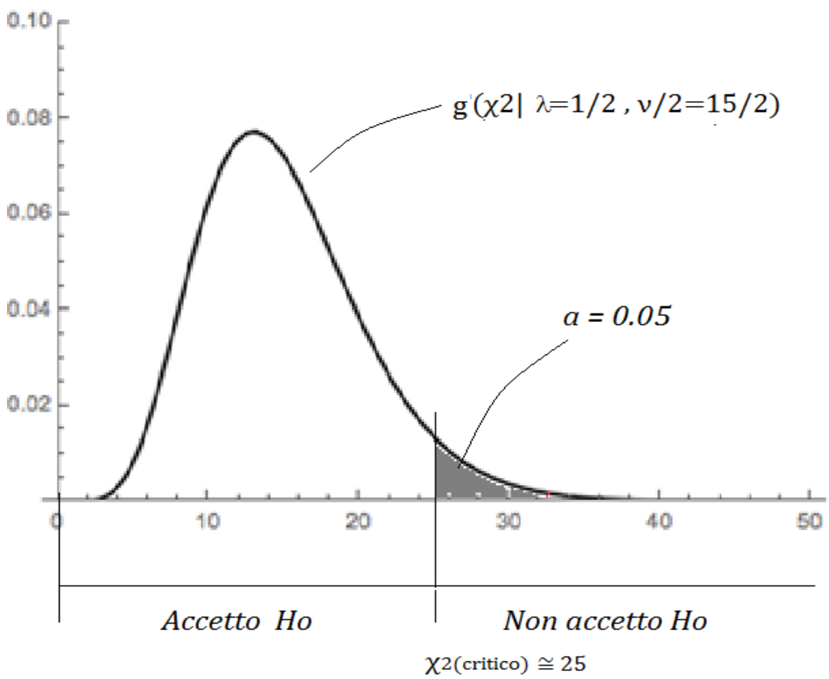


Figura 2. La funzione rappresentata rispetta i parametri del problema

Si fissano le due ipotesi alternative:

$$\begin{cases} H_0: \ n_{i,j} = \hat{n}_{i,j} & \forall\, i,j \\ H_1: \ n_{i,j} \neq \hat{n}_{i,j}\,, & \exists\, i,j \end{cases} \tag{4}$$

H_0 dichiara che non ci sono buone ragioni per dubitare che i 2 caratteri siano statisticamente indipendenti; H_1, dichiara il contrario, cioè che ci sono buone ragioni per credere che le differenze tra frequenze assolute osservate e teoriche siano dovute a fattori sistematici, non casuali quindi.

Il test d'ipotesi (induzione probabilistica) si basa sull'idea che tra i possibili valori di χ^2 ce ne sia uno, $\chi^2_{critico}$, in corrispondenza di una coda destra di $g(\chi^2|\, \lambda = 1/2,\ \upsilon/2)$ di area $\alpha = 0.05$, che divide la regione delle soluzioni possibili (nel nostro caso $\mathbf{R}^+$) in due parti: la regione di accettazione dell'ipotesi H_0 e quella di rifiuto di H_0 (e allora si accetta H_1):

Se il valore calcolato di χ^2 (nel nostro caso è 34.9067) cade nella regione di accettazione dell'ipotesi H_0 allora la si accetta e si conclude *che non ci sono buone ragioni per dubitare che i caratteri siano statisticamente indipendenti*; altrimenti si respinge H_0 e si conclude che *non ci sono buone ragioni per affermare che i due caratteri siano indipendenti statistica-mente*. Si noti che il test si fa per valutare se esiste una dipendenza sistematica tra i due caratteri, diventando irrilevante l'aspetto casuale. Dobbiamo trovare quel valore di $\chi^2_{critico}$ che rispetti la condizione:

$$\int_0^{\chi^2_{critico}} \frac{2^{-\frac{v}{2}} \cdot e^{-\lambda \cdot \chi^2} \cdot \chi^{2\left(\frac{v}{2}-1\right)}}{\Gamma\left(\frac{v}{2}\right)} \cdot d\chi^2 = 0.95. \quad (5)$$

Con il programma MATHEMATICA della Wolfram Research o con le opportune tavole si ottiene: $\chi^2_{cri}(\alpha = 0.05, \upsilon = 15) \cong 25$, dove $\upsilon = (r-1)\cdot(c-1) = (4-1)\cdot(6-1)=15$.
Si conclude che sulla base dei dati a disposizione l'ipotesi H_0 va respinta e perciò *X* e *Y* dimostrano una dipendenza sistematica, cioè non casuale.
3) Si calcolano le medie *M*(.) e le varianze *V*(.) di *x* e *y*. Si compara, quindi, la variabilità di *x* e *y* con i coefficienti di variazione υ(.) che, essendo numeri puri, permettono il con-fronto:

$$M(x) = 63/16\ ; \quad V(x) = M(x-\bar{x})^2 = 1135/256\ ; \quad \upsilon(x) = \frac{\sqrt{V(x)}}{\bar{x}} \cong 0.5348; \quad (6)$$

$$M(y) = 53/16\ ; \quad V(y) = M(y-\bar{y})^2 = 535/256\ ; \quad \upsilon(y) = \frac{\sqrt{V(y)}}{\bar{y}} \cong 0.4364.$$

In questo caso, la variabile *x* presenta una maggiore variabilità di *y*.
4) Le medie aritmetiche condizionate sono rispettivamente

$$M(x \mid y = 4) = 24/5, \qquad M(y \mid x = 2) = 2.$$

5) La covarianza viene definita nel modo seguente:

$$C(x, y) = M[(x-\bar{x}) \cdot (y-\bar{y})] \quad (7)$$

e misura la covariabilità lineare tra le due grandezze. Conviene calcolare la covarianza di *x* e *y* con il metodo indiretto:

$$C(x, y) = M(x\cdot y) - M(x)\cdot M(y) = 709/256, \quad (8)$$

poi si calcola il coefficiente di correlazione lineare (di Bravais-Pearson):

$$r(x, y) = \frac{C(x,y)}{\sqrt{V(x)}\cdot\sqrt{V(y)}} \cong 0.91. \quad (9)$$

Quest'ultimo risultato dimostra che le due variabili hanno una forte correlazione lineare positiva (è noto che $-1 \leq r(.,.) \leq +1$).
6) A questo punto siamo indotti a determinare la retta di regressione (interpolante a minimi quadrati) considerando *x* come variabile indipendente e *y* dipendente: $\hat{y} = a + b\cdot x$; si ottiene: $a = 968/1135$ e $b = 709/1135$. "*a*" e "*b*" si possono ottenere (più facilmente) anche dalle seguenti relazioni strettamente statistiche:

$$a = M(y) - b \cdot M(x) \qquad \text{e} \qquad b = C(x, y)/V(x). \tag{10}$$

La dimostrazione (facile) è lasciata al lettore.
Quindi, la retta stimata è data in forma esplicita da

$$\hat{y} = \frac{968}{1135} + \frac{709}{1135} \cdot x \tag{11}$$

7) Gli indici che scegliamo per misurare la bontà dell'accostamento fatto tra un modello interpolante teorico e il fenomeno osservato sono due, a seconda che si voglia una misura buona in generale o una misura specifica per modelli lineari. Per modelli qualsiasi si usa il coefficiente di variazione così definito:

$$I_2 = \sigma_{\hat{y}}/M(\hat{y}) = \left(\sqrt{\frac{1}{n}\sum\nolimits_{i=1}^{n}(y_i - \hat{y}_i)^2}\right) \Big/ \left(\frac{1}{n}\sum\nolimits_{i=1}^{n}\hat{y}_i\right) \tag{12}$$

Il valore che si ottiene nel nostro caso è $I_2 \cong 0.1811$. È una misura relativa buona per comparazioni. Si conviene, comunque, che un modello teorico sia bene accostato ai dati sperimentali se I_2 è molto basso (in generale, minore di 0.1). Valuterei più che sufficiente il valore ottenuto. Invece, per modelli lineari si usa come misura del grado di accosta-mento il coefficiente di determinazione. Esso è il rapporto tra la devianza di regressione e la devianza totale; cioè:

$$r^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}\ . \tag{13}$$

che, nel nostro caso è: $r^2 \cong 0.8277$. Il valore è alto ($0 \le r^2 \le 1$) e quindi la retta può essere considerata un buon modello di rappresentazione del fenomeno osservato. Osserviamo che mentre r misura la correlazione delle due variabili x e y senza esprimere alcun giudi-zio sulla dipendenza di una delle due dall'altra, r^2 misura la bontà dell'accostamento li-neare tra modello teorico e fenomeno osservato.
Se prendessimo come variabile statisticamente indipendente la y invece della x avremmo

$$\hat{x} = \alpha + \beta \cdot y\,. \tag{14}$$

Otteniamo, con ragionamento analogo al precedente, quanto segue:

$$\alpha = \bar{x} - \beta \cdot \bar{y} = -\frac{242}{535} \ ; \quad \beta = \frac{C(y,x)}{V(y)} = \frac{709}{535} \ ; \quad \hat{x} = -\frac{242}{535} + \frac{709}{535} \cdot y \tag{15}$$

8) Presentiamo, infine, il grafico 3 che descrive il comportamento dei dati sperimentali e il modello teorico interpolato. L'area di ogni punto corrisponde al valore della frequenza assoluta associata a ogni coppia (x, y) osservata.

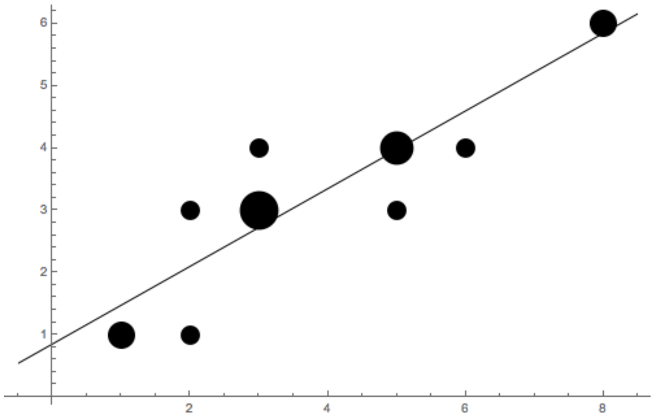


Figura 3. Sull'asse delle ascisse sta la variabile *x*, su quella delle ordinate la variabile y.

Bibliografa:

[B.1] Corso Luciano, Un problema per i lettori, MatematicaMente (ISSN: 2037-6367) n.225 e 227, giugno e agosto 2017.

[B.2] Piccolo Domenico, Statistica, Il Mulino, Bologna, 1998.

[B.3] Manzone G. A. M. & Longo S. C., Probabilità e Statistica 2, ed. Tramontana, Mila-no, 2004.

[B.4] MATEMATICA della Wolfram Research.

[*] Il presente lavoro è già stato pubblicato sulla rivista MatematicaMente (ISSN: 2037-6367) n.225 e 227, giugno e agosto 2017. È qui riportato per gentile concessione dell'autore e della Mathesis – sezione di Verona.

Un laboratorio sulla probabilità

Loredana Biacino[1]

[1]Università di Napoli Federico II
E-mail: loredana.biacino2@unina.it

Sunto – Si propone un approccio alla probabilità di tipo problem solving, pensato per coinvolgere gli alunni personalmente o in gruppo in modo attivo. Anche se quasi tutti i problemi esposti sono corredati da soluzione, l'idea è quella di fornire semplicemente una base per un laboratorio dove lanciando i dadi, una o più monete, giocando con le carte, estraendo biglie da un'urna, simulando estrazioni al lotto etc.. ci si accosti in modo quasi sperimentale ai primi elementi della teoria della probabilità.

Parole chiave: probabilità, problemi, problema della divisione della posta in giochi interrotti, il problema del Cavaliere di Méré, il problema della moneta di Buffon, eventi incompatibili, indipendenti, probabilità condizionata

1. Introduzione

Come trattare in modo matematico i fenomeni aleatori come il lancio di un dado, di una moneta, l'estrazione del lotto? Si cerca di sviluppare un tipo adeguato di matematica, scartando l'applicazione delle regole matematiche e fisiche che si adoperano per il trattamento degli eventi deterministici in quanto non rispondenti alla natura delle nuove entità in gioco. E si comincia con il descriverli. Per questo si introduce lo spazio campione: nel lancio del dado {1,2,3,4,5,6}, nel lancio della moneta {T, C}, nel gioco del lotto {1,2,3, …,90}.
Sia S lo spazio campione: i suoi elementi sono gli eventi elementari e rappresentano i diversi esiti che può avere l'esperimento aleatorio. Ad ogni prova si verifica uno ed uno solo di tali esiti. I sottoinsiemi di S sono gli eventi che constano, nel caso finito, di insiemi finiti di eventi elementari. Due eventi A e B sono incompatibili se hanno intersezione vuota. Ad es. se S={1,2,3,4,5,6} allora A={2,3} e B={4,5,6} sono eventi incompatibili, mentre C={2,4,6} non è incompatibile né con A, né con B.
Non sempre è facile determinare lo spazio campione. Ad esempio se si lancia la moneta 3 volte lo spazio campione è $S = \{T,C\}\times\{T,C\}\times\{T,C\}$, se si lancia la moneta n volte, $S = \{T,C\}^n$, se si lancia il dado due volte $S = \{1, 2, 3, 4, 5, 6\}^2$, etc….

La probabilità classica – Si basa sul concetto di equiprobabilità. Conseguentemente la probabilità di un evento E, P(E), è data dal rapporto tra casi favorevoli e casi possibili. Ad es. nel lancio di un dado non truccato si suppone che tutte le facce siano equibrobabili, cioè sostanzialmente non ci sia alcun motivo per ritenere una faccia più probabile delle altre e quindi la probabilità di ogni faccia sia 1/6. Per gli altri eventi si utilizza l'additività, ad es. si pone P({1,3,6}))≡3/6=1/2. Un'obiezione che si fa a questo punto è che per definire la probabilità si fa ricorso ad una nozione, l'equiprobabilità che presuppone eguale probabilità per gli eventi elementari, ottenendo in tal modo una definizione che si mangia la coda. Questa nozione era di fatto già conosciuta ed applicata sin dal Cinquecento, ma chi la definì in maniera rigorosa fu Pierre Simon Laplace (1749-1827) nel suo trattato *Teoria analitica delle probabilità* del 1812, il primo importante sviluppo sistematico della teoria, dove si evidenzia che la precedente non é una vera e propria definizione, ma una regola cui attenersi che si basa su un antico principio, che Laplace trasferisce in ambito probabilistico e diventa il così detto *Principio d'Indifferenza*, la regola cioè secondo la quale, se non c'é alcun motivo di preferire di due o più eventi l'uno all'altro, dobbiamo assegnare a tutti la stessa probabilità.

Probabilità frequentista – Per determinare la frequenza con cui si presenta un evento aleatorio E si ripete l'esperimento un gran numero di volte e si vede quante volte esso si presenta, si pone poi P(E) eguale al rapporto tra il numero di volte in cui si presenta E ed il numero di volte in cui si è ripetuto l'esperimento.
Questo approccio può servire a tarare un dado o una moneta. Ad esempio se lanciando un dado si vede che il {5} esce un numero triplo di volte rispetto agli altri numeri allora si può ragionare così: posto x il numero di volte in cui escono 1,2,3,4,6, sarà *3x* il numero di volte in cui esce il 5 e *8x* il numero di volte in cui si é ripetuto l'esperimento. Quindi P({1}) = P({2})= P({3}) = P({4}) = P({6}) = *x/8x*=1/8; P({5}) = *3x/8x*=3/8.

L'area del cerchio – Consideriamo un quadrato Q e un quarto di cerchio C inscritto in esso. Si lanciano sulla figura dei granelli di sabbia, a pioggia, in modo uniforme e in gran quantità. Il numero dei granelli nel cerchio (casi favorevoli) è legato al numero dei granelli nel quadrato (casi possibili) e il rapporto di questi numeri, indipendente dal raggio, è molto vicino al rapporto area C/area Q ≡π/4, cioè π/4 rappresenta la probabilità in senso frequentista che un granello di sabbia cada in un quarto di cerchio C. Ci troviamo di fronte ad uno spazio di eventi non finito, come nei seguenti esempi; il metodo precedente permette di calcolare tante cifre decimali di π quante si vuole aumentando il numero dei granelli di sabbia.

Bersaglio circolare – Qual è la probabilità di colpire a caso (ad occhi chiusi) in un bersaglio circolare, un punto più vicino al centro che alla circonferenza? In tal caso possiamo considerare come casi possibili i punti interni alla circonferenza e come casi favorevoli i punti interni alla circonferenza di raggio metà. Il rapporto tra casi favorevoli e casi possibili è $\pi r^2/4\pi r^2 = 1/4$ ed è questa la probabilità richiesta.

Il problema della moneta di Buffon – Si consideri un pavimento ricoperto di piastrelle quadrate di lato L. Qual è la probabilità dell'evento E = "una moneta di raggio R, caduta sul pavimento, intercetta il bordo di una piastrella"?

Risposta - Consideriamo un quadrato di lato L, sia Q. Q rappresenta in questo caso lo spazio campione. Perché la moneta cada al suo interno il centro della moneta deve mantenersi ad una distanza inferiore ad R dal bordo di Q. Ciò senz'altro non è possibile se R è maggiore o eguale a L/2. Sia allora R<L/2. Consideriamo un quadrato T contenuto in Q concentrico e con lati paralleli a Q di lunghezza L-2R. Se il centro della moneta cade in questo secondo quadrato, senz'altro la moneta non intercetta il bordo di Q. Quindi detto E^c il complementare dell'evento di cui vogliamo calcolare la probabilità, risulta $P(E^c)) = (L-2R)^2/L^2$ e di conseguenza $P(E) = 1-P(E^c)) = 4R(L-R)/L^2$.

La definizione assiomatica– Alla fine dell' '800 mancava ancora una definizione astratta della probabilità: la richiesta esplicita fu espressa da Hilbert che nel 1900 al Congresso di Parigi, tra i venti problemi irrisolti della matematica, pose questo al sesto posto. Solo dopo trenta anni, nel 1931 Andrei Kolmogorov propose la definizione assiomatica, usando la teoria della misura di Lebesgue. In particolare dato un insieme finito S ad ogni sottoinsieme E di S è associato un numero $P(E)\in[0,1]$ tale che

1)) $P(\emptyset) = 0$; $P(S) =1$;

2) $P(A\cup B) = P(A)+P(b)$ se A e B sono insiemi disgiunti.

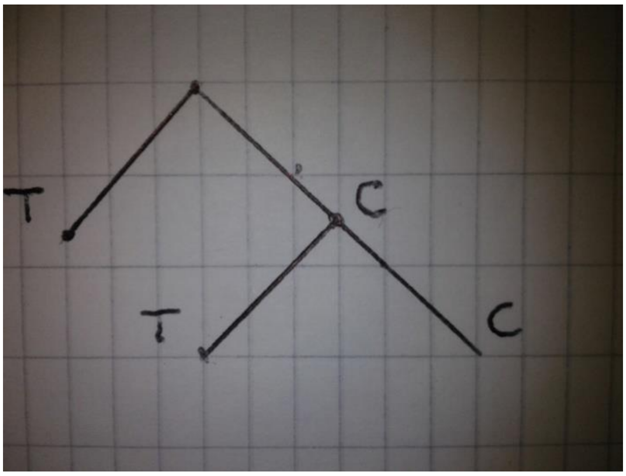
Se l'insieme non è finito, va selezionata una famiglia $\Im$ di suoi sottoinsiemi, gli *"eventi"*, cui appartengano il vuoto ed S, chiusa rispetto al complemento e le unioni numerabili; ad ogni evento E va associato un numero $P(E) \in[0,1]$ tale che valga la 1) e la numerabile additività, cioè

2') $P(\cup A_n) = \Sigma P(A_n)$ per ogni successione (A_n) di elementi a due a due disgiunti di $\Im$.

2. Problema risolto da Pascal nella lettera a Fermat del 29 luglio 1654.

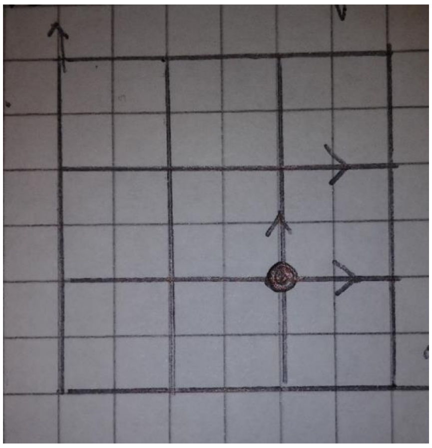
Problema - Due giocatori giocano a testa e croce con una moneta non truccata. Ogni volta che esce testa, T, il giocatore 1 si aggiudica 1 punto; se invece esce croce, C, è il giocatore 2 che si aggiudica 1 punto. Chi per primo raggiunge il punteggio di 3 punti si aggiudica la partita e vince la posta in gioco, costituita da 48 euro (ogni giocatore ha versato per giocare 24 euro). La partita ad un certo punto si interrompe prima della conclusione e si pone il problema di ripartire in modo quanto più giusto possibile la posta tra i due giocatori.

Un problema di questo tipo è già presente nel trattato del 1494 *Summa de aritmetica, geometria et proporzionalità* di Luca Pacioli. In quel caso l'autore propone di dividere la posta in gioco proporzionalmente ai risultati già riportati dai giocatori. Il problema è anche discusso da Gerolamo Cardano nel *Liber de ludo aleae* del 1526 circa, il primo trattato sul gioco dei dadi, dove si analizzano i casi favorevoli in rapporto ai casi possibili e dove è anche enunciata la regola del calcolo della probabilità congiunta di eventi indipendenti, cioè $P(A \cap B) = P(A) \cdot P(B)$: vi é espressa la massima: "l'unico vantaggio è non giocare per niente". Torniamo al procedimento usato dal Pacioli: esso fu successivamente criticato perché dà eccessiva importanza a quanto è accaduto prima dell'interruzione della partita e non tiene nel giusto peso quelle che sono possibili evoluzioni del gioco nel futuro. E' proprio in quest'ottica che va vista la soluzione del problema proposta da Pascal e da Fermat nel 1654 e che qui viene riportata con linguaggio moderno. Per risolvere il problema posto consideriamo i diversi casi che possono presentarsi:

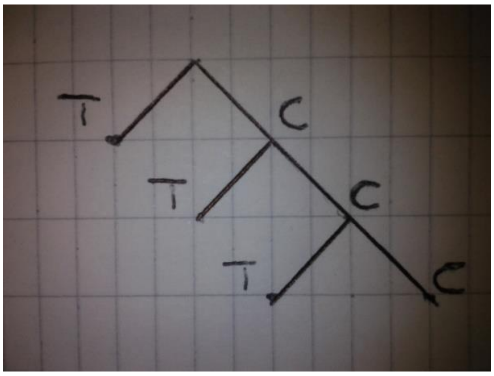


a) Quando la partita si interrompe il giocatore 1 ha totalizzato 2 punti mentre il giocatore 2 ne ha totalizzato solo 1.

Se si dividesse la posta in parti proporzionali a 2 e a 1 si dovrebbero attribuire 32 euro al giocatore 1 e 16 euro al giocatore 2. Procediamo però in altro modo osservando che per quanto riguarda gli sviluppi successivi della partita basterebbero due ulteriori giocate perché essa si concluda. Infatti se lanciando la moneta uscisse T allora 1 vincerebbe, se invece uscisse C bisognerebbe proseguire: ma allora se lanciando nuovamente la moneta uscisse T 1 vincerebbe, mentre se uscisse C vincerebbe 2 e il gioco sarebbe concluso. Quindi due sono le possibilità di vincita del giocatore 1 legate all'uscita di una T e una esclude l'altra. Ma quale è la probabilità che esca una T nei due casi? Al primo lancio è ½, al secondo lancio è ¼, perché segue l'uscita di una C. Quindi poiché i due eventi "esce T al primo lancio" e "esce T al secondo lancio" si escludono a vicenda e la loro unione fornisce l'evento "il giocatore 1 vince la partita", la probabilità di quest'ultimo evento si determina sommando le probabilità degli altri due e quindi è data da $1/2 + 1/4 = 3/4$. La probabilità di vittoria di 2 è allora $1 - 3/4 = 1/4$, come si evince anche dal fatto che una sola è la possibilità di vittoria di 2, consistente nella uscita di C sia al primo lancio che nel secondo. Quindi la posta va così divisa: al giocatore 1 vanno ¾ di 48 euro, cioè 36 euro, più quindi che nel caso di divisione proporzionale della posta, al giocatore 2 vanno i restanti 12 euro con perdita effettiva di 12 euro, meno che nel caso di divisione proporzionale della posta.



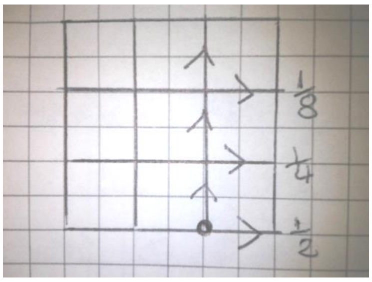
Possiamo immaginare che il gioco consista in una gara a chi esce per primo da un labirinto quadrato al cui interno ci sono 2 percorsi orizzontali e due verticali che si incrociano formando su ogni tratto verticale, e rispettivamente su ogni tratto orizzontale, 3 percorsi unitari (3 è il numero dei punti che bisogna totalizzare per vincere). Il punto di partenza è determinato dal punteggio che i giocatori hanno quando la gara precedente si interrompe, precisamente se il punteggio è (m,n) si parte dal nodo che è m-simo verso destra e n-simo verso l'alto. La regola del gioco è la seguente: scopo dei giocatori 1 e 2



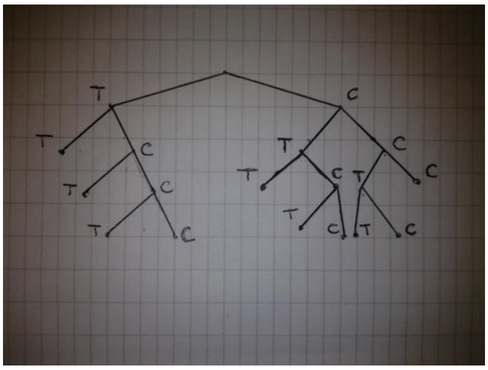
è uscire dal labirinto, il primo dal lato destro, il secondo in alto; per muoversi i giocatori ad ogni nodo debbono gettare la moneta, se esce T si muove il giocatore 1, spostandosi di un tratto unitario a destra; se esce C è il giocatore 2 che si sposta d'un tratto unitario in alto. I percorsi unitari hanno tutti probabilità ½, ma il loro concatenamento comporta il prodotto delle probabilità: una volta che siano visualizzati tutti i cammini vincenti per il giocatore 1 si determinano le loro probabilità che vanno poi sommate. La somma è la probabilità di vittoria del giocatore 1. Analogamente per il giocatore 2.

b) Quando la partita si interrompe il giocatore 1 ha totalizzato 2 punti mentre il giocatore 2 non ne ha realizzato nessuno.

In tal caso se si volesse dividere la posta proporzionalmente al risultato raggiunto il giocatore 1 si aggiudicherebbe tutta la posta, cosa che non accade invece andando a studiare le probabilità di vincita dei due giocatori. Infatti pur avendo punteggio nullo, il giocatore 2



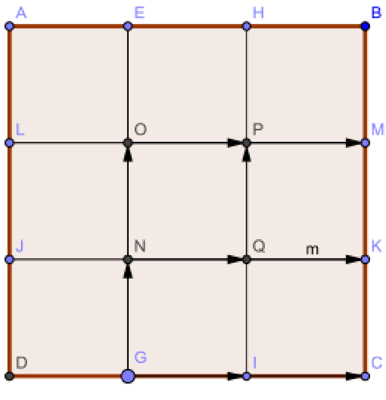
ha comunque una possibilità di vincita data dal fatto che gettando la moneta 3 volte potrebbe uscire sempre C. Questo evento ha probabilità 1/8. Quindi la probabilità di vittoria di 1 non è 1, ma 7/8 che si può calcolare anche tenendo presente che 3 sono i percorsi che assicurano a 1 la vittoria, con probabilità ½ il primo, 1/4 il secondo e 1/8 il terzo (sono i percorsi T, CT, CCT) e ½+1/4+1/8=7/8. Quindi in tal modo il giocatore 2 recupera comunque 1/8 di 48 euro, cioè 6 euro dei 24 spesi mentre a 1 vanno 42 euro con un guadagno di 42-24=18 euro.



Ovviamente le situazioni (1,1) e (2,2) sono situazioni di parità e in tal caso la posta va divisa a metà, cioè ciascuno riprende quello che ha giocato. Inoltre basta ovviamente considerare solo le situazioni in cui è in vantaggio il giocatore 1, in quanto le altre situazioni sono perfettamente simmetriche e vanno quindi trattate in modo analogo.

c) Quando la partita si interrompe il giocatore 1 ha totalizzato 1 punto e il giocatore 2 nessuno.

Dallo schema a grafo ricaviamo quali sono i possibili percorsi che si chiudono con T, in quanto corrispondono agli eventi, a due a due escludentisi il cui verificarsi permette al giocatore 1 di vincere. Ognuno di essi ha una probabilità, sommando tali probabilità si ottiene la probabilità complessiva che ha il giocatore 1 di vincere, data da: ¼+1/8+1/16+1/8+1/16+1/16=11/16. Quindi al giocatore 1 spettano 11/16.48=33euro mentre al giocatore 2 spettano5/16.48 = 15 euro.



Il metodo che è stato impiegato finora è di tipo combinatorio, abbastanza simile al metodo impiegato da Fermat nella corrispondenza con Pascal su questo tema. Pascal invece, pur accennando al metodo che Fermat gli aveva esposto in una lettera andata perduta, e sul quale egli stesso si soffermerà nella lettera del 24 agosto successivo, nella lettera del 19 luglio 1654, propone invece a Fermat un modo diverso per risolvere la questione che si può esporre alla seguente maniera. Partiamo dal caso a): Pascal osserva che il giocatore 1 potrebbe dire a 2:se lanciando la moneta uscisse C allora saremmo alla pari e dovremmo dividere la posta. Quindi dammi 24 euro; dei rimanenti 24 euro dammi però la metà che mi spetta in quanto al primo lancio potrebbe uscite T nel qual caso vincerei tutto. Quindi dammi 36 euro.

Nel caso b) il giocatore 1 potrebbe dire a 2: se lanciando la moneta uscisse C allora saremmo nel caso precedente e quindi mi spetterebbero 36 euro; ma può uscire anche T

nel qual caso vincerei tutto: dammi allora 36 euro più la metà dei restanti 12 euro, cioè 36+6 = 42 euro.
Nel caso c) il giocatore 1 potrebbe dire a 2: se lanciando la moneta uscisse T ci troveremmo su 2 a 0 e quindi mi spetterebbero 42 euro; se uscisse C saremmo alla pari e ci spetterebbero 24 euro a testa. Quindi dammi i 24 euro di cui sono certo più la metà della differenza 42-24, cioè 9 euro. In tal modo 1 si aggiudicherebbe 24+9 = 33 euro.

E' interessante il procedimento di tipo ricorsivo che impiega Pascal, contrapposto al procedimento combinatorio, ma sostanzialmente equivalente ad esso.
Abbiamo considerato un caso particolare, cioè abbiamo supposto che il punteggio da aggiudicarsi per ottenere la vittoria fosse 3. Allo stesso modo si ragiona se il punteggio da raggiungere per terminare la partita è un numero maggiore, ma ovviamente più tale numero aumenta più sono i casi da analizzare. Ad esempio si può provare a fare lo studio precedente se la vittoria si ottiene raggiungendo il punteggio 4. In tal caso si deve effettuare un'analisi analoga alla precedente supponendo che la partita possa essere stata interrotta quando i punteggi dei due giocatori erano i seguenti: 3 a 2; 3 a 1; 3 a 0; 2 a 1; 2 a 0 e infine 1 a 0.

3. Il problema posto a Pascal dal Cavaliere di Méré

Il Cavaliere de Méré (1607-1684), fu un filosofo e letterato, vissuto in Francia alla corte di Luigi XIV. Grande giocatore, il Cavaliere pose a Pascal, suo amico, il seguente quesito:
"L'uscita di un sei lanciando 4 dadi dovrebbe avere la stessa probabilità di avere almeno una coppia di sei lanciando 24 volte una coppia di dadi. Come mai invece il primo evento sembra verificarsi con maggiore frequenza del secondo?"
Sembra che per un errato calcolo delle probabilità dei due eventi il Cavaliere de Méré sia andato in rovina.
Egli calcolava infatti la probabilità nel primo caso moltiplicando per 4 la probabilità dell'uscita di un sei e otteneva così 4/6=2/3; analogamente moltiplicava per 24 la probabilità dell'uscita di un doppio 6, cioè 1/36 e otteneva 24/36 =2/3. I due valori ottenuti erano così eguali, ma il calcolo sbagliato come si rese conto giocando ripetutamente. Infatti ragionando come il Cavaliere: se si lancia 7 volte un dado la probabilità é 7/6>1?
L'errore nasce dal fatto che si possono addizionare probabilità solo all'interno di uno stesso spazio campione. Nel caso più semplice, cioè nel lancio di 4 dadi, lo spazio campione ha come eventi elementari tutte le possibili quaterne che si possono formare con i numeri 1,2,3,4,5,6, quaterne che sono in numero di 6^4.
I due eventi confrontati dal Cavaliere hanno entrambi probabilità vicine a½, lievemente superiore il primo, lievemente inferiore il secondo. Precisamente per il primo la

probabilità è data da 1-$\left(\frac{5}{6}\right)^4 \cong$ 0,518, cioè circa 0,52, in quanto coincide con la probabilità che non si verifichi l'evento contrario consistente nell'evento che in tutti e quattro i lanci esca sempre un numero diverso da 6, evento che ha probabilità $\left(\frac{5}{6}\right)^4$, in quanto lo spazio campione consiste di eventi elementari che sono quadruple ognuna con probabilità $(1/6)^4$; per il secondo caso, in modo analogo, la probabilità è data da 1-$\left(\frac{35}{36}\right)^{24} \cong$ 0,491, lievemente inferiore a 0,5.

Rivediamo il calcolo precedente in un altro modo, approfondendo l'esame degli eventi in questione, ad esempio nel primo caso (sostanzialmente equivalente al secondo). Il primo evento, diciamolo E, non consta della unione di quattro eventi di natura più semplice, quale quello del lancio una sola volta del dado, ma si compone di vari eventi di natura più complessa: precisamente esso si verifica quando il 6 esce al primo tiro, o al secondo tiro, o al terzo o al quarto, cioè quando si verifica A= "esce il 6 una sola volta nei 4 lanci"; evento che ha probabilità $4\cdot1/6\cdot(5/6)^3$ in quanto in tal caso in un lancio è uscito il 6, ma negli altri lanci il risultato è diverso da 6; E si verifica pure nel caso si verifichi B= "esce il 6 esattamente due volte nei 4 lanci"; evento che ha probabilità $6\cdot 1/6^2\cdot(5/6)^2$ in quanto in sei modi distinti il 6 può verificarsi esattamente 2 volte. E si verifica pure quando si presenta l'evento C = "il 6 si verifica esattamente 3 volte nei 4 lanci" e questo accade in 4 modi diversi e quindi con probabilità $4\cdot1/6^3\cdot\left(\frac{5}{6}\right)$, oppure quando si verifica D = "il 6 si verifica in tutti e 4 i lanci", con probabilità $1/6^4$. In definitiva E = A∪B∪C∪D e

$P(A) = 4\cdot1/6\cdot(5/6)^3 + 6\cdot1/6^2\,(5/6)^2+4\cdot1/6^3\cdot\left(\frac{5}{6}\right)+\ 1/6^4$

$= (5/6)^4 + 4\cdot\ 1/6\cdot(5/6)^3 + 6\cdot1/6^2\cdot(5/6)^2+4\cdot\ 1/6^3\cdot\left(\frac{5}{6}\right)+\ 1/6^4-(5/6)^4$

$= (1/6+5/6)^4- (5/6)^4 = 1-(5/6)^4$

che è il risultato già trovato.

Si osservi che il precedente problema è concettualmente equivalente al seguente più semplice.

Qual è la probabilità che lanciando la moneta tre volte esca due volte testa? E che esca testa una sola volta? O tre volte? O nessuna volta?

Rispondiamo alla prima domanda. Poiché la moneta è lanciata 3 volte lo spazio campione è costituito da terne del tipo (T,C,C) che si legge al primo lancio è uscita T al secondo C e al terzo C. Il numero delle possibili terne si calcola facilmente, esse sono in numero di 8 e hanno tutte la stessa probabilità, 1/8. Quelle che contengono due volte T sono (T,T, C) (T,C,T) (C,T, T), cioè 3. La probabilità richiesta è 3/8. Ragionando in modo simile si risponde alle altre domande.

Anche in questo caso si può utilizzare un diagramma ad albero, con 8 rami. I problemi del Cavaliere di Méré sarebbero allo stesso modo schematizzabili con un albero, ma il numero eccessivo di rami, nel primo caso $6^4 = 1296$, 36^{24} secondo caso, ne sconsiglia l'uso.

4. Processo di Bernoulli

Il primo libro sulla teoria della probabilità fu *l'Ars Conjectandi* di Jacques Bernoulli, pubblicato nel 1713. Vi veniva formulata *la legge dei grandi numeri*, che permette di calcolare probabilità a posteriori quando non sia possibile effettuare a priori il computo dei casi favorevoli e possibili. Questa legge fondamentale asserisce che se di un evento aleatorio A si calcola la frequenza *m* su *n* prove, tale frequenza, *m/n* al crescere di *n* ha per limite la probabilità P(A).
Ora sarà esposto il così detto processo di Bernoulli: esso contempla il caso di eventi indipendenti in successione, quali quelli che si presentano in prove o giochi ripetuti più volte.
Due eventi A e B si dicono indipendenti se P(A⋂B)=P(A).P(B).
Ad esempio gli eventi A = {2,4,6} e B = {2,3} sono indipendenti perché A⋂B = {2} e quindi P(A⋂B) = 1/6 = P(A)·P(B). Invece A={2,3} e B={2} non sono indipendenti. Infatti P({2})=1/6 mentre P((A)·P(B) = 1/3·1/6 e d'altro canto se si verifica B si verifica anche A.
Anche gli eventi {1,3} e {2} non sono indipendenti: infatti il verificarsi di uno dei due esclude l'altro.
Sussiste la seguente fondamentale:
REGOLA: La probabilità che si verifichino due eventi indipendenti in successione è data dal prodotto delle probabilità. Ad esempio la probabilità dell'evento "esce due volte T lanciando due volte i dadi" è data da 1/6 . 1/6.
Più in generale consideriamo il lancio di una moneta non truccata, cioè con P(C) = P(T) = 1/2. Qual è la probabilità che lanciando la moneta n volte esca k volte testa?
La probabilità che nei primi k lanci esca T e nei rimanenti C è $(1/2)^k(1/2)^{n-k}$, ma ovviamente, se si prescinde dall'ordine, i modi in cui si può presentare l'evento sono $\binom{n}{k}$, come si è potuto constatare in un esempio precedente con n=3 e k=2.
Se le probabilità di C e di T non sono le stesse, cioè se la moneta non è equa, posto P(T))=p e P(C)=1-p=q, si ha che la probabilità che esca k volte T è data da $\binom{n}{k}p^kq^{n-k}$.
La probabilità che la sequenza delle T si presenti in un certo prefissato ordine, ad esempio nei primi k lanci e poi esca sempre croce è p^kq^{n-k}.
In generale, sia p la probabilità di successo in un esperimento e si ripeta l'esperimento n volte. Allora:
a) la probabilità di k successi e (n-k) insuccessi in un dato ordine è p^kq^{nk};
b) la probabilità di k successi e (n-k) insuccessi è $(\genfrac{}{}{0pt}{}{n}{k})p^kq^{n-k}$.

5. Alcuni problemi del calcolo delle probabilità.

Problema 1 – In un'urna ci sono 16 palline, 10 rosse e 6 blu. Si estraggono consecutivamente, senza reintroduzione, o, ciò che è lo stesso, contemporaneamente, due palline: qual è la probabilità che siano una rossa e una blu? E che siano dello stesso colore?

Risposta - Probabilità che la prima pallina sia R: 10/16;
probabilità che la seconda pallina sia B: 6/15;
probabilità della sequenza R B: $10/16 \cdot 6/15$.
Analogamente
probabilità della sequenza B R: $6/16 \cdot 10/15$.
Quindi la probabilità richiesta è $2 \cdot 10/16 \cdot 6/15 = 1/2$.
Si osservi che la probabilità che siano dello stesso colore è di conseguenza 1/2, che si può anche ottenere sommando la probabilità che le due biglie siano entrambe rosse, data da $10/16 \cdot 9/15 = 3/8$, con la probabilità che siano entrambe blu data da $6/16 \cdot 5/15 = 1/8$.

Osservazione – Possiamo rappresentare con un grafico la precedente situazione. Consideriamo il quadrato di lato 16 disposto nel primo quadrante di un sistema di riferimento cartesiano con due lati adagiati sugli assi e un vertice coincidente con l'origine O. A partire da O su entrambi gli assi, a distanza di una unità da O e tra loro tracciamo 10 punti che rappresentano le palline rosse e poi ancora procedendo altri sei punti che rappresentano le blu. Ogni punto a coordinate intere del quadrato, che non si trovi sulla diagonale, rappresenta un'estrazione, cioè una coppia ordinata che si può pensare come una coppia di palline: ad esempio (3,4) rappresenta l'estrazione di due palline rosse, (12, 9) rappresenta l'estrazione di una pallina blu come prima estratta e di una rossa, (15,14) rappresenta l'estrazione consecutiva di due blu. Poiché non si può estrarre due volte la stessa pallina sono stati esclusi i punti della diagonale. In totale i punti da considerare sono $16 \cdot 16 - 16 = 256 - 16 = 240$. Le coppie costituite da una pallina rossa e una blu sono in numero di $10 \cdot 6 + 6 \cdot 10 = 120$, quindi la probabilità richiesta è data da: casi favorevoli/casi possibili = 120/240 = 1/2, come già determinato in precedenza. Si osservi che se si richiede la probabilità dell'uscita di due rosse, riapplicando la formula precedente troviamo per i casi favorevoli $10 \cdot 10 - 10 = 90$ e quindi probabilità uguale a 90/240 = 3/8.

Problema 2 – In un'urna ci sono 3 palline R e 5 B: calcolare la probabilità che in due estrazioni consecutive (con reintroduzione):
- escano due palline rosse;
- escano due palline blu;
- escano due palline di diverso colore.

Suggerimento - Si consiglia di ripetere lo schema precedente, senza però eliminare la diagonale nel caso della soluzione geometrica. Perché?

Problema 3 – In un'urna ci sono n biglie R e m biglie B. Si estraggono consecutivamente, (senza reintroduzione), due biglie: in quali casi la probabilità di estrarre due biglie di diverso colore è 48/100?
La probabilità di estrarre due biglie di diverso colore è $2nm/(n+m)(n+m-1)$ e imponendo che sia eguale a 48/100 si trae:
$25nm=6(n+m)^2 -6(n+m)$.
Poniamo $n+m=k$ e sostituiamo:
$25n(k-n)=6k^2-6k$ da cui : $25n^2-25nk +6k^2-6k =0$.
Risolvendo e semplificando si trova una radice quadrata il cui radicando è k^2+24k. Poiché la radice deve dare un naturale, $k^2+24k=k(k+24)$ deve essere un quadrato. Ma allora k deve essere un quadrato e deve essere tale che se a esso si somma 24 si deve ottenere ancora un quadrato. Questo avviene per $k=1$ (caso che va scartato per la natura del problema), per $k=5$ e per nessun altro valore di k, in quanto:
6^2+24, 7^2+24, …. 11^2+24 non sono quadrati e d'altro canto, $12^2+23 = 13^2$, $13^2+25=14^2$….. e pertanto i quadrati degli altri numeri a partire da quello di 13 hanno tra loro distanza sicuramente maggiore di 24 (abbiamo tenuto presente la formula:$1+3+5+ \ldots +2n-1 = n^2$, che ha una semplice dimostrazione di carattere geometrico).
Ma potrebbe anche accadere che $k(k+24)$ sia del tipo $k^2(1+s)$ con $s<24$, intero naturale. In tal caso $1+s$ dovrebbe a sua volta essere un quadrato: discutere questo caso.

Problema 4 – Quante volte si dovrebbe lanciare un dado per avere una chance del 99% che esca almeno una volta il 6?
Risposta - 26 volte. Infatti si deve determinare un k tale che lanciando k volte il dado non esca un numero diverso da 6, cioè k tale che $1-(5/6)^k = 99/100$ e quindi …. . Se il dado viene lanciato 26 volte e non è mai uscito il 6 che si dovrebbe pensare? Sapresti fare un confronto tra questo problema e il precedente? Si tratta infatti di due tipi di eventi diversi: prima come ora, si trattava di eventi in successione, ma prima il risultato della prima estrazione influenzava la seconda, ora invece ….

Problema 5 - Consideriamo il gioco che consiste, nel lancio simultaneo di due dadi, nell'indovinare la somma dei due numeri che compaiono sui dadi. Qual è il numero che ha più probabilità di essere il risultato? Qual è tale probabilità?
Risposta – I possibili esiti sono 36, i valori del gioco sono tutti i numeri da 2 a 12.
2 e 12 con probabilità 1/36; 3 e 11 con probabilità1/18; 4 e 10 con probabilità 1/12; 5 e 9 con probabilità 1/9; 6 e 8 con probabilità 5/36 e infine 7 con la maggiore probabilità, eguale a 1/6.
Si possono evidenziare tutti i precedenti risultati con uno schema simile a quello introdotto nei problemi 1 e 1'; per ogni numero n da 2 a 12 si constata che le coppie tali che la somma delle loro coordinate sia n sono allineate.

6. Probabilità condizionata

La probabilità di fare 6 lanciando due dadi è 5/6. Supponiamo di effettuare il lancio e che uno dei due dadi cada per primo mostrando il 2. Nei pochi istanti in cui aspettiamo l'uscita dell'altro dado possiamo valutare di nuovo la probabilità: solo se esce il 4 possiamo fare 6 e quindi ora la probabilità di fare 6 è 1/6>5/36. La cosa si spiega perché con l'uscita del 6 si è azzerata la possibilità che uscisse il 6 al primo lancio, la qual cosa avrebbe reso impossibile il verificarsi dell'evento. Se invece di "fare 6" consideriamo l'evento A = "la somma dei punti è 7" vediamo che le due probabilità non cambiano. Infatti P(A) = 1/6 e quindi una volta che sia uscito il 2 solo se esce 5 possiamo fare 7, con P(5) = 1/6. Se però consideriamo l'evento "la somma dei punti è 9", una volta che sia uscito il 2 ci rendiamo conto che la probabilità é zero.
Supponiamo che si voglia determinare la probabilità di un evento A, sapendo però che si é già verificato un evento B, con P(B)>0, evento che fornisce informazioni circa il verificarsi di A.
Nel 1761 il reverendo presbiterianoThomas Bayes introdusse nella memoria: *An Essay Toward Solving a Problem in the Doctrine of Chance* (Un saggio sulla soluzione di un problema relativo alla dottrina del caso) la seguente regola nota come legge di Bayes:
P(A/B) = P(A⋂B)/P(B).
Ad esempio si è lanciato un dado e si sa che il risultato é dispari. Qual è la probabilità che sia maggiore o eguale a 3?   Poniamo A ≡ {3,4,5,6}, B ≡ {1,3,5} con p(B)≡1/2.
Allora la probabilità richiesta é P(A/B) = P(A⋂B))/P(B) = P({3,5})/1/2 = 2//3.
La legge di Bayes può essere letta anche al seguente modo:
(*)   P(B/A) = P(A/B)P(B)/P(A),
In quanto   P(A/B)P(B)/P(A) = P(A⋂B)/P(B)·P(B)/P(A) = P(B/A).
Nella formula (*), A va interpretato come effetto e B come causa, la formula fornisce la probabilità che essendosi verificato l'effetto A, B ne sia la causa; é perciò una formula utile nelle diagnosi mediche.
Siano ora dati n eventi A , esaustivi e mutuamente esclusivi, con probabilità non nulle: si dice che è assegnato un *sistema di alternative.*
Il teorema delle alternative – Sia dato un sistema di alternative $A_1, \ldots, A_n$. Allora, per ogni evento A, risulta:

P(A)= $\sum_{i=1}^{n} P\left({}^{A}/_{A_i}\right) P(A_i)$.

Dim. Infatti

$\sum_{i=1}^{n} P\left({}^{A}/_{A_i}\right) P(A_i)$ = $\sum_{i=1}^{n} P(A \cap A_i)$ = P(A⋂(⋃A_i)) = P(A⋂S) = P(A), c.v.d..

Problema – Tre scatole A, B, C contengono ferri da stiro, di cui alcuni sono difettosi. A contiene 1000 ferri da stiro, di cui il 7% presentano difetti, B contiene 700 ferri da stiro, di cui il 10% sono difettosi, e C ne contiene 1500 di cui il 20% sono difettosi. Si estrae a caso un ferro da stiro da una delle tre scatole scelta a caso. Qual è la probabilità che esso non sia difettoso?
Risposta – E' un classico caso in cui si utilizza il teorema sulle alternative: definiamo

D = ”il ferro scelto è difettoso” ; A = ”Il ferro si trova nella scatola A” ; B = ”Il ferro si trova nella scatola B” , C = ”Il ferro si trova nella scatola C”.
Si ha: P(A) = P(B) = P(C) = 1/3, e inoltre P(D/A) = 7/100, P(D/B) = 10/100, P(D/C) = 20/100. Allora, senza far riferimento alle quantità dei ferri da stiro,

$$P(D) = P(D/A)P(A)+P(D/B)P(B)+P(D/C)P(C) = 1/3[7/100 +10/100 + 20/100]$$
$$= 37/300 = 0{,}1233.$$

7. Il problema di Monty Hall

Questo paradosso è legato ad un gioco televisivo statunitense. Prende il nome dal conduttore del programma, Maurice Halprin, soprannominato Monty Hall. Il gioco consiste in questo: si fanno vedere al giocatore tre porte chiuse: dietro due di esse c'é una capra, dietro la terza c'è un'automobile. Ovviamente il giocatore non conosce cosa c'è dietro ognuna di tali porte. Può vincere l'automobile se indovina dove è nascosta. Egli deve scegliere una porta. Dopo la sua scelta la porta rimane chiusa ed egli non sa cosa ha scelto. Ora il conduttore, che conosce la disposizione dei premi apre una delle due porte rimanenti, precisamente una porta dietro la quale c'è una capra e chiede al giocatore se egli intende cambiare la sua porta con quella che è rimasta chiusa. Ciò che è contro l'intuizione è il fatto che al giocatore conviene cambiare. Perché?
Valutiamo la situazione: egli inizialmente ha 1/3 di probabilità di vincere l'auto. Vediamo cosa succede dopo che il conduttore ha aperto la sua porta.
a) Il giocatore ha scelto all'inizio la porta con l'auto, cambiando perde;
b) All'inizio il giocatore ha scelto la porta con la capra, cambiando vince sicuramente, perché il conduttore ha scoperto l'altra capra.
Quindi mentre senza possibilità di cambiare porta il giocatore ha 1/3 di probabilità di vincita e 2/3 di perdita, se cambia porta la probabilità di vincita é 2/3; il paradosso consiste nel fatto che con l'espediente del cambiamento della porta la probabilità di vincita raddoppia, cioè è esattamente eguale alla probabilità di perdere nel caso non fosse possibile il cambio.
Alcuni quando viene loro proposto il gioco asseriscono che sono d'accordo sul fatto che le probabilità di successo grazie all'informazione data dal conduttore aumentano, ma ritengono che passino da 1/3 a 1/2, in quanto sostanzialmente alla fine si deve scegliere tra due opzioni fra loro equivalenti e pertanto é indifferente cambiare oppure no. Per questo motivo la soluzione che comporta il cambio della porta appare antintuitiva. Ciò che entra in gioco è il fatto che non c'è semplicemente una riduzione delle chance da 3 a 2 possibilità, ma che una precedente situazione, prima dell'intervento del conduttore c'erano 3 possibilità, é stata modificata.
E allora valutiamo la probabilità di vincita con il cambio della porta anche usando la legge di Bayes: a tale scopo, supposto che il giocatore abbia scelto la porta 1, indichiamo con A_2 l'evento "dietro la porta 2 c'è l'auto", e con C_3 l'evento "dietro la porta 3,aperta dal conduttore, c'è la capra; allora, per la legge di Bayes:

$P(A_2/C_3) = P(C_3/A_2)P(A_a)/P(C_3) = \frac{1.\frac{1}{3}}{\frac{1}{2}} = 2/3$,

avendo tenuto presente che se l'auto sta in 2 sicuramente il conduttore apre la porta 3 e che la probabilità a priori che il conduttore apra la 2 o la 3 è ½.

Nel gioco televisivo però il conduttore apriva una porta con una capra ma non dava la possibilità al giocatore di cambiare la porta che aveva scelto.

II problema delle tre porte era stato ideato dal matematico francese Joseph Louis François Bertrand nel suo libro *Calcul des Probabilités* (1889) ed era noto come problema delle tre scatole di Bertrand; fu riproposto nella rubrica Mathematical Games di Martin Gardner nel 1959 col nome di "Problema dei tre prigionieri". Esso si enuncia al seguente modo.

Tre carcerati A, B, C, in isolamento l'uno dagli altri, sono stati condannati a morte. Prima dell'esecuzione però la Corte decide di graziarne due a caso. Il prigioniero A chiede al carceriere di dirgli il nome di quello degli altri due che sarà graziato. Pensa infatti che questo non costituirà un problema per il carceriere visto che senz'altro uno degli altri due sarà graziato, ed egli non può comunicare la notizia al fortunato, per lo stato di isolamento in cui si trovano. Il carceriere però si rifiuta perché dice che non vuole alterare in senso negativo le probabilità di salvezza di A. Per quale motivo il carceriere si comporta così?

Per capire le ragioni del carceriere dobbiamo fare qualche calcolo. Indichiamo con A (risp. B,C) anche gli eventi "A (risp. B, C) è stato graziato". Allora le probabilità a priori sono date da: P(A) = P(B) = P(C) = 2/3. Consideriamo gli eventi $A\cap B$, $A \cap C$, $B \cap C$. Poiché si verifica uno e uno solo indifferentemente di essi, ognuno dei tre eventi precedenti ha probabilità a priori 1/3.

Supponiamo che il carceriere dica ad A che è stato graziato B, la probabilità a posteriori di A è allora data da $P(A/B) = \frac{P(A\cap B)}{P(B)} = \frac{\frac{1}{3}}{\frac{2}{3}} = 1/2$.

Il carceriere si è quindi rivelato un buon matematico: infatti la condizione di A peggiorerebbe con un'informazione maggiore e la sua probabilità di vita scenderebbe da 2/3 a ½. Quello che è paradossale nel precedente ragionamento è il fatto che inizialmente A già sa che uno tra B e C sarà graziato, ma non sa quale. Conoscere quale dei due sarà graziato cambia la sua probabilità di salvezza.

Bibliografia

K. Baklawski, M. Cerasoli, G. C. Rota, Introduzione alla probabilità, U.M.I., Bologna, 1984.

Progettomatematica.dm.unibo.it/ProbElem/12esercizi.html

Il problema di Monty Hall su Wikipedia.

Situazioni problematiche aperte in probabilità

Roberto Tortora[1]

[1]Università degli Studi di Napoli "Federico II"
E-mail: roberto.tortora@unina.it

Sunto
In questo intervento si presenta un esempio di problema riguardante una semplice situazione di valutazione di probabilità. Si suggerisce un articolato percorso didattico per la risoluzione del problema che consente, in un ambiente collaborativo, di confrontare ed integrare i vari approcci alla probabilità. Peraltro l'attività didattica potrebbe continuare coinvolgendo molti altri argomenti di matematica. Complessivamente la proposta e l'approccio adottato si inseriscono in una visione della matematica che privilegia gli aspetti della modellizzazione.

Parole chiave: Modellizzazione. Problem solving, Approcci alla probabilità. Inquiry based mathematical learning.

1. Introduzione

Il punto di vista generale entro il quale si inserisce il presente intervento è quello di privilegiare nella matematica, in tutta la matematica, l'aspetto della modellizzazione. Da notare che con questa parola si intende spesso un tipo particolare di costruzione matematica, spesso complessa, che serve a descrivere un fenomeno della realtà altrettanto complesso. In tali casi si fa uso quasi sempre di strumenti di analisi matematica e in particolare delle equazioni differenziali.
Qui si dà invece a questa parola un significato più generale e adatto a situazioni più svariate e di difficoltà da minima ad elevata. Parleremo di modellizzazione tutte le volte che della matematica non ci interessa sapere soltanto quali sono le sue regole interne, ma invece ci interessa altrettanto se non di più il suo rapporto con la realtà. E quindi ogni volta che usiamo i numeri naturali per contare degli oggetti stiamo facendo attività di modellizzazione, e la stiamo facendo quando misuriamo le dimensioni di una stanza e così via.
Non è certo un'idea nuova. Sentiamo anzi cosa dice in proposito F. Enriquez, in un intervento di oltre un secolo fa.

> *Se le matematiche vengono così spesso riguardate come inutile peso dagli allievi, dipende in parte almeno dal carattere troppo formale che tende a prendere quell'insegnamento, da un falso concetto del rigore tutto intento a soddisfare certe minute esigenze di parole, da una critica analitica eccessiva e fuori di posto... Ma queste tendenze si riattaccano ad una causa più generale; cioè al fatto che le matematiche siano state studiate come un organismo a sé, riguardandone piuttosto la sistemazione astratta conseguita dopo uno sviluppo secolare, che non l'intima ragione storica. Si dimenticano per tal modo i problemi concreti che conferiscono interesse alle teorie, e sotto la formula o lo sviluppo del ragionamento non si vedono più i fatti ormai da lungo tempo acquisiti, ma soltanto la concatenazione in cui noi artificialmente li abbiamo stretti*. (**F. Enriques**, *Sulla preparazione degli insegnanti di scienze*, relazione tenuta al V Congresso degli insegnanti di scuole medie, 1906).

Un ruolo speciale da questo punto di vista lo rivestono i problemi, ed a preferenza non quelli stereotipati che si trovano tipicamente fra gli esercizi dei libri di matematica della scuola, ma quelli che scaturiscono appunto dalla realtà.
In questo quadro la probabilità ha un ruolo particolare. Forse perché fra tutte le branche della matematica è nata parecchio più tardi, forse perché viene percepita come parte "meno affidabile" di altre (la distinzione di senso comune che viene fatta fra matematica del certo e dell'incerto, se pure descrive un dato di fatto, genera anche spesso una sorta di gerarchia di valori fra matematica di serie a e matematica di serie b), o forse semplicemente per l'oggetto al quale tipicamente si applica, sta di fatto che nel modo comune di essere presentata essa non ha il rango di disciplina astratta che basta a se stessa, come accade per l'aritmetica, per l'algebra, per la geometria. Infatti se queste parti più antiche della matematica si usa presentarle come autonome scienze che vivono di vita propria e che solo per chi ha voglia di occuparsene si ritrovano nella vita reale, presentare invece la probabilità come una parte astratta della matematica ovviamente anche si può, ma non si usa, almeno non nelle scuole. Questa presentazione astratta sarebbe quella che viene chiamata probabilità "assiomatica" e che infatti di solito viene presentata come uno dei possibili modi di trattare la probabilità. Ricordiamo che gli altri tre modi tradizionali sono la trattazione classica o *a priori*, quella frequentista o *a posteriori* e quella soggettiva.
A mio modo di vedere le cose non stanno così. L'unico modo di vedere la probabilità in termini matematici (e fra l'altro come capitolo di pari livello di affidabilità di tutti gli altri) è quello della sua trattazione assiomatica. Gli altri tre modi sono, come dire, adattamenti alla realtà di questo modello matematico, sono cioè il *trait d'union* fra la matematica e alcuni specifici fenomeni della realtà (quelli in cui si fanno previsioni sugli avvenimenti), sono in definitiva nient'altro che "metodi di modellizzazione".
Non sto dicendo che bisogna trattare la probabilità a scuola con metodo astratto, sto dicendo esattamente il contrario, che quello che si fa con la probabilità, e che è pratica scolastica, si dovrebbe fare anche con le altre parti, evitando di fare trattazioni astratte svincolate dalle loro applicazioni.

Ma torniamo alla probabilità. Se non si capisce come intendere "i vari modi di trattare la probabilità" si rischiano gravi **misconcezioni riguardo alla probabilità**.
La più comune, come ho già detto, è **abusare** dello slogan: "matematica del certo e dell'incerto": che cosa c'è di incerto in una parte della teoria della misura? E che cosa c'è di certo nel pretendere che nel mondo ci siano punti rette e piani? Un'altra caduta è nella ricorrente tentazione di diffidare dello strumento di modellizzazione matematica e ricorrere invece a vecchi modelli di tipo magico. Ancora un'altra difficoltà sta nel fare molta matematica difficile "interna", che riesce utile nelle scienze e in svariate applicazioni (economia, medicina, demografia, ecc.), senza aver chiarito a sufficienza gli aspetti elementari e il rapporto tra realtà e formalizzazione.

2. Il problema e la prima fase della risoluzione

Veniamo dunque al contenuto di questo intervento. Cosa vedremo? Un esempio, fra i tanti possibili, di una situazione problematica riguardante una valutazione di probabilità (preferisco questo termine – valutazione – al più riduttivo termine *calcolo*), ed una proposta di conduzione didattica di una tale situazione, diversa dall'usuale. Questo esempio, particolarmente ricco sia di contenuti matematici che di aspetti metodologici, vuole mostrare come sia possibile, nel nostro caso e a proposito di probabilità, giustificare varie regole e procedure matematiche e se del caso metterne in luce eventuali aspetti critici, proprio collegandosi ripetutamente e ostinatamente con la realtà.
Il problema trattato, che potremmo intitolare RAPPRESENTANTI DI CLASSE, è il seguente:

Una classe è composta da 25 alunni: 15 femmine e 10 maschi. Si devono estrarre a sorte 2 rappresentanti di classe. Quale è la probabilità che siano un maschio e una femmina?

Esso è tratto inizialmente da attività didattiche condotte per anni nei corsi della SICSI per insegnanti di matematica ed è poi stato utilizzato in molti altri casi simili.
Nell'affrontare il problema seguiremo una sequenza metodologica non usuale, di cui spero si capirà la ragione nel corso della trattazione. Cominceremo cioè a trattare la questione con strumenti che potremmo chiamare di probabilità soggettiva, passeremo poi alla "sperimentazione" e finiremo con una versione classica della probabilità. (Alla fine si potrebbero aggiungere un gran numero di sorprendenti sviluppi matematici). La verità è che la questione posta si presta a tutte e tre le impostazioni della probabilità (classica, frequentista e soggettiva). L'ordine è, come detto, inconsueto, ma conforme ad una visione costruttiva della conoscenza. Peraltro la stessa distinzione (classica,

frequentista, soggettiva) non sempre ha ragione d’essere o meglio non sempre può essere presentata in modo del tutto chiaro.
Per il nostro problema, ciò significa che ciascuno effettui in primo luogo una stima del valore richiesto. Confrontare queste stime, discuterne, infine chiarire l'importanza di abituarsi a effettuare stime accurate è la prima parte del lavoro, ed è una parte che deve essere trattata con tutto il tempo necessario. Dunque, per cominciare, occorre esprimere una stima della probabilità di estrarre una coppia femmina - maschio (ed anche eventualmente la probabilità degli altri due eventi FF, MM).
La risposta sia veloce, magari risultato di riflessioni ma non di calcoli o di applicazione di formule. Poi se è il caso si potrà anche discutere dei criteri con cui saranno state effettuate le stime.
Alla fine ci chiediamo: Come si può verificare l’attendibilità delle stime effettuate? E magari, come spiegare l’eventuale diversità delle varie stime?

3. La seconda fase

Si decide così di passare alla seconda fase, nella quale si procede ad effettuare delle estrazioni. Occorre ripetere le prove un numero congruo di volte, per valutare la probabilità richiesta con grado più o meno grande di affidabilità. In questa fase si incontrano nozioni delicate come quella del "tendere" della frequenza alla probabilità, dei numeri "grandi", dell'equivalenza dei vari successivi eventi.
Riporto di seguito il resoconto di questa parte dell’attività, come riportato da alcune studentesse.
“Abbiamo preso dei pezzettini di carta, su 15 vi abbiamo scritto F e su 10 M, li abbiamo arrotolati e messi in un contenitore. Infine abbiamo effettuato 100 estrazioni (con reimbussolamento) e registrato i risultati ottenuti circa il verificarsi degli eventi:

(MM) = estrazione di 2 maschi
(FF) = estrazione di 2 femmine
(MF) = estrazione di un maschio e di una femmina (indipendentemente dall’ordine)

Anche se, come ha fatto notare un nostro collega, potevamo simulare tali estrazioni con Excel, abbiamo preferito constatare, testare manualmente che cosa succede se si effettuano un numero sempre più elevato di estrazioni. In definitiva abbiamo ottenuto, su cento estrazioni, i seguenti risultati (riportati in percentuale):

MF	FF	MM
48%	37%	15%

Confrontando i risultati da noi ottenuti con quelli riportati dai nostri colleghi e mettendoli insieme abbiamo ottenuto che la probabilità di ottenere come rappresentanti di classe un maschio e una femmina si assesta intorno al 49%.
A questo punto ci siamo chieste: *"Tale probabilità su che cosa si basa, è corretto calcolarla in tal modo, è significativa?"* Rispolverando (per chi le aveva) le conoscenze in materia e documentandosi ex-novo (per chi sprovvista), possiamo affermare che i risultati dell'esperimento si basano sull'aspettativa, molto radicata nel senso comune, espressa nella **legge empirica del caso**:
"In un gran numero di prove fatte nelle stesse condizioni la frequenza relativa dei successi si avvicina alla probabilità, e l'approssimazione in genere migliora con l'aumentare del numero delle prove".
Secondo la concezione frequentista di probabilità, è possibile basarsi esclusivamente sulle frequenze per definire che cosa è la probabilità, senza alcuna valutazione preventiva (probabilità classica o a priori). Infatti secondo tale concezione la probabilità è: "*la frequenza relativa dei successi, in un gran numero di prove fatte nelle stesse condizioni".* Questa affermazione è approssimativa e si presta ad alcune osservazioni tra cui: quante prove sono necessarie per arrivare ad un valore stabilizzato che si possa chiamare probabilità?
Per ovviare a tale inconveniente, si utilizza nella definizione "il limite della frequenza", arrivando alla versione seguente (più corretta) della <u>*definizione frequentista*</u>: "*La probabilità P(E) di un evento E è il limite della frequenza (relativa) $f_n(E)$ dei successi (cioè delle prove in cui l'evento si verifica), quando il numero n delle prove tende all'infinito".*

$$P(E) = \lim_{n \to \infty} f_n(E) = \lim_{n \to \infty} \frac{n_E}{n}$$

Si osserva qui che le frequenze sono definite dal rapporto tra il numero n_E delle prove favorevoli all'evento e il numero totale n di prove effettuate.

Si è giunti così al termine della seconda fase, in cui si confronta il risultato ottenuto usando la definizione frequentista con quello fornito dalle stime precedenti. E' l'occasione per parlare di alcuni comuni problemi e alcune misconcezioni. Non è per nulla da sottovalutare l'importanza di coinvolgere gli allievi nell'esperimento sopra descritto, anzi l'approccio operativo contenuto nella definizione frequentista è da preferirsi in prima istanza, perché, "sporcandosi le mani", i ragazzi sono "costretti" a familiarizzare con strumenti empirici per risolvere problemi la cui soluzione esatta, algebrica od analitica, non sia praticabile e, aspetto non secondario, a porsi domande, e perciò a cercare risposte che contribuiscono alla costruzione della conoscenza.

4. La terza fase

Siamo ora alla terza fase: In essa si procede infine alla costruzione del modello corrispondente alla probabilità a priori. E si discute di che cosa significa casi possibili, casi favorevoli, equiprobabilità. E di quali requisiti matematici sono necessari per questi calcoli (con cenno al calcolo combinatorio).
Alla fine emerge il risultato del problema, già ragionevolmente intuito in verità alla fine della seconda fase.
Passare alle formule significa introdurre il più astratto tra gli approcci al problema, quello più "matematico", e questo non può che essere l'ultimo passo, giustificato dai tentativi precedenti e da approcci più ingenui. Si perviene così alla **Definizione classica (o a priori) della probabilità**:
"*La probabilità di un evento è il rapporto tra il numero dei casi favorevoli all'evento, e il numero dei casi possibili (purché questi ultimi siano tutti ugualmente probabili)*".
Qui si pone un'interessante questione: **Come si contano i casi possibili? E quelli favorevoli?**
Il primo problema è dunque quello di contare. Ma si fa presto a dire contare, come se fosse una cosa semplice! Bisogna imparare a contare le cose più diverse fin dalla scuola elementare. Ma altro è contare le caramelle, o le matite, altro è contare cose che non si possono toccare o perché sono lontane o perché non si vedono o perché sono troppe, o perché sono vaghe (cercare insieme tanti esempi e studiare diverse strategie per contare è un ottimo esempio di modellizzazione elementare!)

In particolare come si contano i "modi possibili"? E i "casi favorevoli"? E' qui che interviene il calcolo combinatorio che però viene presentato nei libri come un repertorio di strani nomi e di formule astratte e astruse di cui non si riesce mai a capire quale si deve usare.

Veniamo al nostro problema…
Quanti sono i casi possibili?
Attenzione, questa non è matematica, qui si pretende di catalogare le cose del mondo, e allora non ci si può (e non ci si deve!) aspettare che, come nella matematica, ci siano risposte sicure e uniche!
Si possono fare moltissimi esempi di queste difficoltà a partire dalle situazioni più comuni utilizzate in probabilità, come i lanci di dadi o di monete. O la nascita di maschi e femmine. O di fatto ogni altra cosa del mondo reale. Si ponga anche attenzione al fatto che nei giochi e nelle lotterie si tende, per ovvie ragioni, a rendere minima l'opinabilità delle classificazioni. Ma il mondo è quello che è, non quello che piacerebbe a noi: di nuovo, la matematica ne fornisce solo un modello!

Quanti sono dunque i possibili esiti dell'estrazione nel nostro problema ("i casi possibili")?
(a) Tre?
(b) Due?
(c) Quattro?
(d) Tanti quante sono le possibili coppie di persone?
E a qualcuno potrebbero venire in mente anche altre risposte. Sono in certa misura esatte tutte queste risposte! Guai a sostenere che una sola risposta è quella esatta.

Esaminiamo separatamente (a), (b) (c) e (d).
Chi sceglie (a) (tre casi) inquadra bene l'essenza del problema. Capisce che non c'è ragione di distinguere fra l'una e l'altra femmina né fra l'uno e l'altro maschio, ma che le possibili situazioni degne di essere distinte sono solo il caso FF, il caso FM ed il caso MM. Ritiene dunque, come è giusto, che l'ordine non sia importante, e quindi non distingue fra FM ed MF. Allora i casi possibili sono tre. Fra questi uno solo è "favorevole". Applicando la formula a priori si ottiene p = 1/3.
Ma il risultato è sbagliato! Come mai? Come si fa a capirlo, a capire cioè che i tre casi considerati non sono equivalenti?
La mia tesi è che bisogna maturare una capacità o un fiuto che deriva solo dal far precedere le esperienze e le stime all'uso delle formule.

Chi sceglie (b) (due casi) va al sodo. I casi sono due, infatti: o i due estratti sono dello stesso sesso oppure sono di sesso diverso. Come dargli torto? La formula qui dà subito come risposta p = 1/2. Il bello è che questa volta la risposta è esatta, ma il ragionamento è sbagliato. E' sbagliato perché sarebbe come dire che la probabilità di qualunque evento è sempre 1/2, perché è sempre possibile dire che i casi sono due: favorevole o sfavorevole. Attenzione, perché questa maniera di ragionare è tutto sommato abbastanza frequente e si adatta a molti casi (nascita di un maschio o di una femmina, numeri pari o dispari, piove o non piove), e può essere utile capire che non ci sarebbe calcolo delle probabilità se tutto l'esistente si suddivide sempre in modo dicotomico.

Chi sceglie (c) (quattro casi) ha forse cominciato a intuire una cosa importante, vale a dire che la suddivisione in casi deve tenere conto della loro equivalenza. I quattro casi per lui sono FF, FM, MF, MM. O magari ha intuito che è possibile suddividere le due estrazioni in due tappe e rendere così meglio conto di tutto ciò che potrebbe capitare: allora l'ordine è importante e occorre distinguere FM da MF. La formula qui dà come risposta p = 2/4 = 1/2. Di nuovo la risposta è esatta, ma il ragionamento è purtroppo ancora una volta sbagliato. Ma capire qui perché il ragionamento è sbagliato è più difficile, e di nuovo arriva in soccorso l'esperienza fatta nella fase precedente, quando si è visto, come è peraltro intuitivo, che il caso FF è più frequente del caso MM.

Non resta che la risposta (d), che è "quella giusta". Ma non è affatto facile arrivarci e faremmo torto all'intelligenza degli studenti, a imporgliela. Perché se facciamo così, qualcuno fra gli studenti ci sarà pure che capisce bene e ci segue, ma sarà solo quello particolarmente dotato; la gran maggioranza invece non ci riuscirà.
Per avere casi sicuramente equivalenti dunque conviene trattare tutti gli studenti come diversi uno dall'altro. Qui però diventa difficile l'operazione del conteggio. Difficile anche perché gli studenti sono tanti (25), e sarebbe arduo elencare tutte le coppie distinguendoli uno per uno con dei nomi o con altre etichette (ad esempio numeri). Se fossero molti di meno l'espediente dell'elenco puro e semplice sarebbe il migliore approccio.

Si possono invece individuare due strategie. La prima consiste nel separare la prima estrazione dalla seconda, la seconda nel trovare un modo per contare le possibili coppie.
Esaminiamo la prima strategia. Alla prima estrazione ci sono 25 studenti, di cui 15 femmine e 10 maschi. La probabilità, con la formula a priori, che si estragga una femmina alla prima estrazione è dunque $p(F_1) = 15/25$, quella che si estragga un maschio è $p(M_1) = 10/25$. Qui dovrebbe ormai essere chiaro che la formula può adoperarsi con sicurezza, perché si può supporre che tutti gli individui abbiano la medesima chance di essere estratti.

Veniamo alla seconda estrazione. Qui dobbiamo distinguere il caso in cui nella prima è stata estratta una femmina da quello in cui è stato estratto un maschio. Nel primo caso su 24 studenti ci sono ora 14 femmine e 10 maschi. La probabilità di avere una femmina è $p(F_2/F_1) = 14/24$, quella di avere un maschio è $p(M_2/F_1) = 10/24$.
Analogamente si trova per il secondo caso $p(F_2/M_1) = 15/24$ e $p(M_2/M_1) = 9/24$, perché su 24 studenti ora le femmine sono 15 e i maschi 9.
Si tratta ora di moltiplicare due a due i valori trovati, secondo il criterio della probabilità composta, che si applica quando si prende in considerazione il susseguirsi di due eventi.
Capire perché si debba moltiplicare non è banale, ma ci si può arrivare come al solito facendo prima esperienza concreta o anche usando rappresentazioni grafiche. Ma non mi ci voglio ora soffermare.
Avremo dunque quattro possibili casi: FF, FM, MF, MM, dove ora conta l'ordine delle due uscite. Le rispettive probabilità sono date da:

$$p(FF) = \frac{15}{25} \times \frac{14}{24} = \frac{7}{20} = 35\%$$
$$p(FM) = \frac{15}{25} \times \frac{10}{24} = \frac{1}{4} = 25\%$$
$$p(MF) = \frac{10}{25} \times \frac{15}{24} = \frac{1}{4} = 25\%$$
$$p(MM) = \frac{10}{25} \times \frac{9}{24} = \frac{3}{20} = 15\%$$

Esaminiamo la seconda strategia (più semplice della prima e secondo me preferibile, perché non richiede la nozione di probabilità composta). Si tratta di contare tutte le coppie di studenti e fra queste quelle costituite da (due femmine, da due maschi e da) una femmina e un maschio. Come si fa? Questo è un esercizio di conta, di quelli su cui bisognerebbe molto impratichirsi. Tutto sommato non è difficile. Si tratta solo di capire che, scelto un primo studente fra 25 (dunque in 25 modi diversi), la scelta del secondo si potrà fare in 24 modi. Vengono in tutto $25 \times 24 = 600$ coppie. Poiché però ogni coppia è presente in questo elenco due volte, bisogna poi dividere per 2, e si ottiene così il numero voluto: 300.
In modo analogo si contano le coppie costituite da due femmine, che sono $15 \times 14/2 = 105$, quelle costituite da due maschi che sono $10 \times 9/2 = 45$ e quelle costituite da un maschio e una femmina che sono tutte le rimanenti: $300 - 105 - 45 = 150$. E si conclude di nuovo che:

$$p(FF) = \frac{105}{300} = \frac{7}{20} = 35\%$$
$$p(FM) + p(MF) = \frac{150}{300} = \frac{1}{2} = 50\%$$
$$p(MM) = \frac{45}{300} = \frac{3}{20} = 15\%$$

5. Conclusioni e cenno a possibili sviluppi

Gli allievi coinvolti in processi di risoluzione che possono svilupparsi in svariati modi possono rendersi conto personalmente dei vantaggi e degli svantaggi nell'utilizzo dell'una o dell'altra "definizione" di probabilità, nelle varie situazioni problematiche, e chiedersi quale delle definizioni sia da preferire. Tipiche domande che sorgono sono: per quali fenomeni è accessibile la nozione di probabilità classica? Quelli in cui i vari eventi sono valutati come equiprobabili? E chi valuta se sono equiprobabili? Come fare a sapere se una moneta è tale che ogni faccia abbia la stessa probabilità di uscire? E' una previsione dettata dalla situazione concreta… che poi può essere verificata facendo

un numero elevato di lanci… ma così facendo ci stiamo spostando di nuovo sulle frequenze!
Su queste questioni gli studiosi continuano a riflettere e a discutere ancora oggi. Le varie strategie di soluzioni proposte per questo problema potrebbero stimolare un inquadramento storico della probabilità, con le diverse "definizioni" presentate a partire dai celebri esempi dei rispettivi caposcuola: Laplace (la probabilità che il sole sorga ancora), von Mises (la probabilità che tra *n* persone ce ne siano due nate nello stesso giorno) e De Finetti (la probabilità dei risultati 1X2 al Totocalcio).
In un liceo l'excursus storico potrebbe concludersi mostrando come la definizione assiomatica proposta da Kolmogorov nel 1933 è compatibile sia con la definizione classica, che con quella frequentista, che con quella soggettivista.
Offrire ai ragazzi la possibilità di conoscere i fondamenti su cui si basa la probabilità e di accostarsi ai contenuti sostanziali in modo intuitivo è indispensabile per entrare nel merito della disciplina, che stenta, malgrado il suo utilizzo in molti rami della scienza e della tecnica, ad entrare nel curricolo della scuola. Essa fa parte di quella cerchia di cenerentole, tra cui possiamo annoverare anche la logica, relegate in uno sparuto capitolo a se stante, dove tutto è lasciato alla sensibilità e buona volontà dell'insegnante, anche egli alle prese, in molti casi, con forti lacune in tale ambito.

Potrebbe bastare. Se non fosse che il problema è così "bello" che, una volta risolto, ne pone subito di nuovi.
Negli sviluppi possibili si incontra "tutta la matematica", dal calcolo algebrico (non banale), allo studio di funzioni, alle successioni, all'aritmogeometria, eccetera.
E alla fine c'è un "premio", cioè la scoperta di cose nuove, belle, inaspettate.
Ma di questo si parlerà, se ci sarà tempo, nell'esposizione orale.

Introduzione all'elaborazione statistica dei dati sperimentali

Mario Innocenzo Mandrone[1]

[1]Dipartimento di Scienze e Tecnologie - Università del Sannio (Bn)
E-mail: almavit@libero.it

Sunto
Il processo attraverso il quale si sviluppa la conoscenza scientifica si basa su di una continua interdipendenza tra la fase dell'osservazione sperimentale e la fase relativa alla costruzione di uno schema teorico- interpretativo. Il processo di acquisizione dell'informazione sperimentale si basa su tre stadi successivi: 1) La raccolta dei dati e la misura delle grandezze fisiche che descrivono il fenomeno che si vuole studiare; 2) L'organizzazione e l'analisi dei dati raccolti mediante l'uso di tecniche appropriate; 3) La determinazione del grado di validità statistica, e quindi di attendibilità del risultato che si è ottenuto. Lo scopo di questo lavoro è di presentare alcuni metodi di elaborazione dei dati sperimentali. Tali metodi hanno una loro precisa sistemazione nella Statistica.
Parole chiave: Grandezze fisiche, misure dirette ed indirette, caratteristiche degli strumenti di misura, errori sistematici ed errori casuali, indici di posizione, indici di dispersione, propagazione degli errori statistici

1. Introduzione

Il processo attraverso il quale si sviluppa la conoscenza scientifica si basa su di una continua interdipendenza tra la fase dell'osservazione sperimentale- tradotta nella misura delle grandezze fisiche che consentono la descrizione del fenomeno in studio- e la fase relativa alla costruzione di uno schema teorico- interpretativo. Si comprende quindi come i dati sperimentali giochino un ruolo centrale nella formulazione di ogni teoria fisica o nella descrizione quantitativa di un qualunque fenomeno fisico. Il processo di acquisizione dell'informazione sperimentale si basa su tre stadi successivi:

a) La raccolta dei dati e la misura delle grandezze fisiche che descrivono il fenomeno fisico che si vuole studiare.

b) L'organizzazione e l'analisi dei dati raccolti mediante l'uso di tecniche appropriate. Tale compito è affidato ai metodi di elaborazione statistica dei dati sperimentali.

c) La determinazione del grado di validità statistica, e quindi di attendibilità del risultato che si è ottenuto.

Lo scopo di questo lavoro è di presentare alcuni metodi di elaborazione dei dati sperimentali. Tali metodi hanno una loro precisa sistemazione nella Statistica.

Tutto ciò che è suscettibile di una determinazione quantitativa, cioè che può essere misurato, è una grandezza fisica. Ne sono esempi la massa, la lunghezza, il tempo, la velocità l'energia, ecc. Per definire le grandezze fisiche è necessario, perciò, stabilire come esse si misurano e con quali dispositivi, mettendole anche in relazione con altre grandezze già definite; in ciò consiste la "definizione operativa" di una grandezza fisica. Misurare una grandezza vuol dire metterla a confronto con il campione unitario, detto unità di misura, al fine di stabilire quanti campioni è necessario sommare per ottenere la grandezza oggetto della misura. Il confronto diretto di una grandezza con l'unità campione, oppure con una sua copia, rappresenta una **misura diretta.** Non sempre è possibile misurare direttamente una grandezza fisica. Ad esempio, non è possibile misurare con una bilancia la massa di un elettrone o della Terra. In casi come questi, è in generale ogniqualvolta la misura diretta è difficile da realizzare, si ricorre a una **misura indiretta.** Misurare indirettamente una grandezza fisica significa ricavarne il valore utilizzando opportune relazioni analitiche tra le grandezze oggetto della misura e altre grandezze di cui è possibile una misura diretta. Spesso il valore della grandezza incognita è determinato dalla posizione di un indice mobile su una scala di un apparecchio o strumento di misura precedentemente tarato. Esempi di strumenti tarati sono i termometri, i manometri, i tachimetri, i cronometri, i voltmetri, gli amperometri ecc. apparecchi come questi sono detti analogici per distinguerli da quelli digitali o numerici in cui il valore della grandezza è indicato direttamente da un numero su un display digitale. Il risultato di una misura è, però, sempre un numero che approssima il cosiddetto "valore vero" della grandezza misurata; cioè la misura è affetta da errori che si manifestano quando la loro ampiezza supera l'errore di sensibilità dello strumento.

Tutte le operazioni di misura, anche se eseguite con le tecniche più accurate e con gli accorgimenti più appropriati, non ci permettono mai di conoscere il valore esatto della grandezza misurata, ma solo di individuarlo con una certa approssimazione. Le cause responsabili di tali errori sono sempre molteplici, ma esse intervengono sulle misure essenzialmente in due modi distinti, originando errori sistematici e/o errori casuali. **Gli errori sistematici** sono dovuti ad un difetto dell'apparecchio o al modus operandi adottato. Sono caratterizzati, in generale, da un valore fisso in segno e grandezza e, di conseguenza sono difficilmente rivelabili con una semplice ripetizione delle misure. L' unica maniera per prevenire questo genere di errori consiste in una taratura periodica degli strumenti di misura e in uno studio critico dei metodi sperimentali utilizzati. Quando possibile, può anche essere istruttivo ripetere la misura con un altro apparecchio e/o un altro metodo. **Gli errori casuali** (detti anche aleatori) sono caratterizzati da una distribuzione casuale dei risultati sperimentali attorno al valore "vero" della grandezza misurata. Ripetendo le misure, questi errori non conservano sistematicamente lo stesso segno o la stessa grandezza. Si può ridurre questo genere di

errori aumentando il numero di misure. Gli errori casuali sono trattabili con metodi statistici, basati sulla teoria della probabilità.
In conclusione, l'errore associato ad una misura è funzione dello strumento che effettua la misura, dell'operatore e del metodo di misura.
La teoria statistica è in grado di esaminare ed interpretare il comportamento degli errori casuali. Si suppone, invece, che eventuali errori sistematici, principalmente dovuti agli strumenti di misura e che non possono essere interpretati statisticamente, siano eliminati grazie all'abilità dello sperimentatore stesso. Ovviamente, la misura deve essere quanto più possibile accurata e precisa. Si dice "accurata" una misura che contiene piccoli errori sistematici"; precisa una misura che contiene piccoli errori casuali.

2. Caratteristiche degli strumenti di misura

Gli strumenti di misura, dai più semplici ai più sofisticati, consentono il confronto tra la grandezza in esame e la corrispondente unità di misura, fornendo una risposta quantitativa. Le caratteristiche principali degli strumenti di misura sono:
1. ***intervallo di funzionamento***
E' individuato dalla portata e dalla soglia che lo strumento è in grado di fornire della grandezza da misurare. La portata indica il valore massimo della grandezza incognita misurabile con lo strumento e corrisponde al limite superiore assoluto del campo di misura. La soglia è il minimo valore per cui si ha una risposta dello strumento. Ad esempio, l'intervallo di funzionamento di un termometro clinico è 35° C- 43° C.
2**. *prontezza***
E' legata al tempo necessario perché lo strumento risponda ad una data variazione della sollecitazione: quanto minore è questo tempo caratteristico, tanto maggiore è la prontezza.
3) ***portata***
E' il valore massimo della grandezza che lo strumento è in grado di misurare in modo corretto e senza danneggiamenti dello strumento stesso.
4) ***sensibilità***
Si definisce sensibilità S di uno strumento di misura il rapporto tra lo spostamento Δn dell'indice (misurato in divisioni) sulla scala e il valore della grandezza ΔG che ha causato tale spostamento:

$$S = \frac{\Delta n}{\Delta G} \left(\frac{divisioni}{unità\ di\ misura}\right)$$

In pratica S rappresenta il numero di divisioni corrispondenti al valore unitario della grandezza misurata. Il suo inverso è detto "costante dello strumento K" 0 "errore di sensibilità " E_S" :

$$K = E_S = \frac{1}{S} = \frac{\Delta G}{\Delta n} \left(\frac{unità\ di\ misura}{divisioni}\right)$$

La costante K o E_S (errore di sensibilità) viene determinata anche come rapporto tra la "portata" (massimo valore misurabile dallo strumento) e "fondo scala" (numero totale N delle divisioni sulla scala).
5) ***precisione***
La precisione" è il rapporto tra la sensibilità dello strumento di misura e la sua portata. Essa dipende da due parametri: a) la fedeltà; b) la giustezza. Uno strumento si dice preciso se è contemporaneamente giusto e fedele, cioè se i valori delle misure sono tutti compresi in un piccolo intorno dell'ipotetico "valore vero". E' altresì evidente che uno strumento di misura può essere fedele, ma non giusto, oppure giusto ma non fedele.

3. Errori nelle misure dirette – Misure ed incertezze

"Quando puoi misurare ciò di cui stai parlando, ed esprimerlo in numeri, puoi affermare di sapere qualcosa; se però non puoi misurarlo, se non puoi esprimerlo con numeri, la tua conoscenza sarà povera ed insoddisfacente". (William Thomson - Lord Kelvin-1824-1907).
Le scienze sperimentali si propongono di fornire valutazioni quantitative dei fenomeni che studiano. Tuttavia nessun metodo di misura fornisce dei valori certi, ovvero non esistono misure perfette; ogni valore sperimentale è affetto da un'indeterminazione (errore) che può dipendere:
1. Dal modus operandi dello sperimentatore;
2. Dalla natura della grandezza da misurare;
3. Dalla qualità delle tecniche e degli strumenti utilizzati.

Si usa distinguere gli errori di misura in: errori sistematici ed errori casuali (detti anche aleatori), caratterizzati da una distribuzione casuale dei risultati sperimentali attorno al "valore vero" della grandezza misurata. Si può ridurre (ma mai eliminare) questo genere di errori aumentando il numero di misure.
Gli errori casuali sono trattabili con metodi statistici, basati sulla teoria della probabilità. Tali errori si dividono a loro volta in **errori massimi ed errori statistici.** Per capire la differenza tra le due tipologie, distinguiamo fra tre situazioni sperimentali:
1. Può capitare in alcune particolari situazioni di riuscire ad eseguire una sola operazione di misura perché si è deciso di non eseguire più misure per qualche ragione (per questioni di tempo o convenienza di vario genere) oppure è impossibile ripetere le misure perché il fenomeno è unico in determinate condizioni, come capita per esempio in molti fenomeni naturali (eclissi, fenomeni metereologici, etc.) e biologici (prelievi invasivi, test che distruggono il campione, ecc.) .

In entrambi i casi, se si è in assenza di errori sistematici, l'indeterminazione sulla conoscenza della grandezza sarà dovuta solo alle caratteristiche dell'apparato. Il modo corretto di fornire il risultato della misura sarà quello di dare la migliore stima della quantità in questione x_{best} insieme ad un intervallo $2\Delta x$ all'interno del quale essa si trova:

$$x_{best} \pm \Delta x$$

Δx rappresenta dunque un errore massimo ed indicherà l'attendibilità della misura stessa. Esso è l'errore di sensibilità dello strumento e, per questo motivo, viene anche detto errore strumentale. L'errore massimo ha le stesse dimensioni fisiche e quindi le stesse unità di misura della grandezza a cui si riferisce. Il doppio di tale intervallo, 2 Δx, rappresenta l'intervallo di sensibilità dello strumento di misura.

2) Se si ripete più volte la misura e si ottiene sempre lo stesso risultato vuol dire che l'errore di sensibilità è molto maggiore della larghezza della distribuzione delle misure. Cioè possiamo dire che la totalità dei valori assunti dalla grandezza è compreso nell'intervallo di sensibilità. Pertanto, anche in questo caso, l'errore di sensibilità assume il significato di errore massimo. In entrambi i casi previsti precedentemente x_{best} risulta essere il valore centrale dell'intervallo di sensibilità.

3) Quando invece, ripetendo più volte la misura della grandezza con una tecnica di elevata sensibilità, si ottengono valori diversi, l'errore di sensibilità non è più sufficiente a descrivere cosa sta accadendo. Se il numero delle misure è abbastanza grande, si assumerà l'esistenza di un **errore casuale**, detto anche **statistico**, che è responsabile della variabilità delle misure.

Nel 1809, nella memoria "Theoria motus corporum coelestium", Gauss curò la prima esposizione della Teoria degli errori di misura. Secondo Gauss la migliore stima del misurando si può ottenere elaborando il complesso dei dati di misura. Aumentare il numero si rivelerà molto importante per migliorare la nostra conoscenza della grandezza in esame.

4. Indici di posizione

Il valore di una grandezza fisica, dopo aver eseguito su di essa N misure, può essere stimato attraverso quei parametri che sono detti **indici di posizione** e che indicano il valore attorno al quale i dati del campione sono posizionati, ovvero che ragionevolmente diano indicazioni su dove si trovino i valori della variabile casuale. Un indice di posizione particolarmente utilizzato quando si hanno pochi dati è la semisomma dei valori massimo e minimo della grandezza in esame:

$$x_{best} = \frac{x_{min} + x_{max}}{2}$$

Un indice molto comune è la media aritmetica:

$$\overline{x} = \frac{x_1 + x_2 + \cdots \ldots \ldots x_N}{N} = \frac{\sum_{1=1}^{N} x_i}{N}$$

Notiamo che la media è calcolabile solo per quelle grandezze che sono manipolabili in senso algebrico, cioè per le quali ha senso definire un ' operazione di somma e di divisione. Quindi la media può essere utilizzata solo per grandezze continue o discrete.

Attribuire ad una grandezza misurata più volte il valor medio dei risultati della serie di misure è intuitivamente logico, ma vedremo anche che trova una forte giustificazione dal punto di vista statistico, in quanto il valor medio risulterà essere la migliore stima della grandezza:

$$x_{best} = \overline{x}$$

Può tuttavia capitare che non tutti i risultati misurati pesino allo stesso modo. Il tal caso si ricorrerà alla **media pesata o media ponderata** dei risultati:

$$\overline{x} = \frac{p_1x_1 + p_2x_2 + \cdots \ldots\ldots + p_Nx_N}{p_1 + p_2 + \cdots \ldots\ldots + p_N} = \frac{\sum_1^N p_ix_i}{\sum_1^N p_i}$$

dove x_i è il risultato della i-esima misura e p_i è detto il peso delle i-esima misura. Definendo la frequenza di ciascun risultato come:

$$f_i = \frac{p_i}{N}$$

ed essendo:

$$\sum_1^N p_i = N$$

risulterà

$$\overline{x} = \frac{\sum_1^N p_ix_i}{\sum_1^N p_i} = \frac{\sum_1^N p_ix_i}{N} = \sum_1^N f_i\, x_i$$

Risulta, inoltre:

$$\sum_1^N f_i = \sum_1^N \frac{p_i}{N} = \frac{N}{N} = 1$$

Qualunque insieme numerico la cui somma vale uno è detta **"normalizzata".**

Vale la pena di ricordare che, oltre alla media aritmetica e alla media pesata, esistono anche altri tipi di medie.

a) **La media geometrica**, che si ottiene come radice n-sima del prodotto di n termini:

$$x_G = \sqrt[n]{x_1 + x_2 + \cdots \ldots\ldots + x_n}$$

La media geometrica viene utilizzata quando le osservazioni raccolte sono in progressione geometrica, come spesso succede, ad esempio, nei fenomeni economici.

b) **La media armonica** si ottiene, invece, calcolando il reciproco della media dei reciproci dei dati:

$$x_A = \frac{1}{\sum_1^N \frac{1}{n} * \left(\frac{1}{x_i}\right)}$$

Tale media è utile quando il fenomeno da indagare è misurato dai reciproci dei dati statistici rilevati.

c) **La media quadratica**, data dalla radice quadrata della media aritmetica dei quadrati dei valori ottenuti in una operazione di misura:

$$x_Q = \sqrt{\frac{1}{n}\sum_{1}^{N} x_i^2}$$

La media quadratica è l'indice che in statistica ha maggiori possibilità di utilizzo. Una espressione analoga, infatti, verrà ad esempio utilizzata nel calcolo dello scarto quadratico medio.
Da un punto di vista esclusivamente descrittivo, la media non rappresenta l'unica scelta possibile. Un altro indice di posizione molto comune è la mediana. La mediana è quel valore che, una volta ordinati i dati del campione, lascia sia alla sua sinistra che alla sua destra la metà dei dati rilevati.
Formalmente la mediana si calcola con la seguente procedura. Detto N il numero dei dati, allora:

1. Si ordinano i dati in senso crescente;
2. Si calcola $\frac{N+1}{2}$;
3. Se il numero $\frac{N+1}{2}$ è intero (N dispari), la mediana è il valore della variabile casuale che corrisponde al numero trovato;
4. Se il numero $\frac{N+1}{2}$ non è intero (N pari), si calcola la media fra il valore precedente e quello successivo al valore così determinato.

L'ultimo indice di posizione che introduciamo è la moda, che è definito come il valore della variabile casuale che si ripete con maggior frequenza. Questo indice si può utilizzare anche per una variabile nominale, in quanto si può comunque individuare la caratteristica che ricorre con frequenza maggiore.
Gli indici di posizione media, mediana e moda sono a volte chiamati misure di tendenza centrale, nel senso che forniscono un valore numerico che si trova, grosso modo, al centro dell'insieme di osservazioni. La media è normalmente considerata la migliore stima di tendenza centrale. Ogni volta che i dati sono distribuiti in modo simmetrico, media, moda e mediana coincidono. Vedremo che la distribuzione dei dati è simmetrica nella maggior parte dei casi in cui i risultati provengono da esperimenti di laboratorio. Occorre tuttavia notare che fornire un solo valore, usando il più opportuno indice di posizione, per caratterizzare una serie di dati, può essere fuorviante, per quanto comodo perché sintetico. Quindi bisogna andare oltre e fornire, accanto all'indice di posizione, un altro indice che dia contemporaneamente anche l'idea di quanto la distribuzione sia sparpagliata attorno all'indice di posizione stesso. A tal proposito verranno introdotti gli indici di dispersione: semidispersione, scarto quadratico medio, range interquartile.

5. Calcolo dell'errore nelle misure dirette

Se uno strumento di misura ha scarsa sensibilità, può accadere che, nel ripetere la misura si trovi sempre lo stesso valore. In tal caso si assume come misura della grandezza il valore letto sullo strumento e come errore la semiampiezza dell'intervallo minimo che lo strumento può apprezzare, cioè l'errore di sensibilità, che assume il significato di errore massimo. Ma come dobbiamo regolarci nel caso in cui le misure ripetute diano risultati diversi? Siamo in grado di indicare un valore della grandezza e un valore dell'errore che siano attendibili? La risposta, come vedremo, è affermativa.
Ammesso di aver eliminato gli errori sistematici, supponiamo che in una serie di N misure, effettuate con lo stesso strumento e con lo stesso metodo, si siano trovati n valori:

$$x_1, x_2, \dots\dots\dots\dots . x_N$$

di una certa grandezza fisica, in generale diversi tra loro.
Si può supporre che questa serie di misure dia un'informazione più precisa rispetto alla singola misura, però è evidente che non è possibile esprimere il risultato di una misura come un insieme di tanti numeri diversi. Ci si pone quindi il problema di come esprimere il risultato di tutte le N misure. Poiché è ragionevole aspettarsi che gli errori casuali avvengano con uguale probabilità per eccesso e per difetto, si può dimostrare che la misura più attendibile della grandezza è la media aritmetica delle singole misure:

$$\overline{x} = \frac{x_1 + x_2 + \cdots \dots . . x_N}{N} = \frac{\sum_{1=1}^{N} x_i}{N}$$

Il valor medio così calcolato risulta essere la migliore stima dell'ipotetico valore vero " della grandezza fisica in esame. Assunta la media aritmetica come valore più probabile della grandezza fisica, rimane il problema di calcolare l'errore di misura detto anche "errore statistico", che può essere così classificato: a) Semidispersione; b) Errore semplice medio c) Deviazione standard

a) La semidispersione

Si può assumere come errore da cui è affetta una misura la semidispersione definita da:

$$d = \frac{1}{2}(x_{max} - x_{min})$$

dei valori trovati in una serie di misure di valor medio

$$\overline{x} = \frac{x_1 + x_2 + \cdots \dots . . x_N}{N} = \frac{\sum_{1=1}^{N} x_i}{N}$$

Se indichiamo con x la misura della grandezza, il risultato della misura andrà espresso da:

$$x = \overline{x} \pm d$$

intendendo con ciò che, presumibilmente il valore x di una misura è compreso nell'intervallo

$$[\overline{x} - d , \overline{x} + d]$$

b) Errore semplice medio

Consideriamo le differenze:

$$x_1 - \bar{x};\ x_2 - \bar{x}; \dots\dots\dots\dots\dots\dots\dots x_N - \bar{x}$$

tra le singole misure e il valor medio. Tali differenze si chiamano " scarti semplici " o scarti della serie di misure dal valor medio. L'errore semplice medio o errore medio è definito come la media aritmetica del valori assoluti degli scarti dalla media:

$$\delta = \frac{|x_1 - \bar{x}|+|x_2 - \bar{x}| + \cdots \dots\dots + |x_N - \bar{x}|}{N} = \frac{1}{N} * \sum_{1}^{N} |x_i - \bar{x}|$$

Il risultato della misura andrà espresso da:

$$x = \bar{x} \pm \delta$$

L'errore qui calcolato è l'errore medio da cui risulta presumibilmente affetta la misura x e, come tale, è sempre minore della semidispersione che rappresenta l'errore massimo. Da notare, inoltre, che la somma degli scarti dalla media è sempre uguale a zero. Questa è una proprietà generale. Difatti:

$$\bar{x} = \frac{x_1 + x_2 + \cdots \dots .. x_N}{N} = \frac{\sum_{1=1}^{N} x_i}{N}$$

Da cui segue:

$$x_1 + x_2 + \cdots \dots .. x_N = N\bar{x}$$

Consideriamo, ora, la somma degli scarti:

$$(x_1 - \bar{x}) + (x_2 - \bar{x}) + \cdots \dots\dots\dots\dots + (x_N - \bar{x}) = (x_1 + x_2 + \cdots \dots\ + x_N) - N\bar{x}$$
$$= N\bar{x} - N\bar{x} = 0$$

La media degli scarti si rivela, quindi, identicamente uguale a zero. Una grandezza più utile ai fini della stima dell'errore è la media dei quadrati degli scarti (chiamata varianza).

c) Deviazione standard

Siano ancora $x_1, x_2, \dots \dots ., x_N$ le misure di una grandezza fisica ottenute in N determinazioni effettuate dallo stesso sperimentatore e nelle stesse condizioni. Definiamo "varianza" la media aritmetica dei quadrati degli scarti dalle singole misure da valor medio. Detta σ^2 la varianza, si ha perciò:

$$\sigma^2 = \frac{1}{N} * \sum_{i=1}^{N} (x_i - \bar{x})^2$$

Non è superfluo osservare che la varianza non è una grandezza omogenea a quella oggetto di misura. La radice quadrata della varianza è invece omogenea alla grandezza misurata e si chiama "errore quadratico medio" o " deviazione standard" della serie di misure. In simboli si ha:

$$1)\ \sigma = \sqrt{\frac{1}{N} * \sum_{i=1}^{N} (x_i - \bar{x})^2}$$

La misura della grandezza fisica, assumendo come errore la deviazione standard, è espressa dalla scrittura:

$$x = (\bar{x} \pm \sigma)$$

Più precisamente si dimostra che esiste una probabilità del 68,3% che il valore di una generica misura cada nell'intervallo $(\bar{x} - \sigma,\ \bar{x} + \sigma)$. Una formula matematicamente equivalente alla 1) è:

$$\sigma = \frac{1}{N}\sum_{1}^{N}(x_i{}^2) - \bar{x}^2$$

molto più agevole da usare della relazione 1), durante l'elaborazione dei dati, quando ancora nono si conosce il numero totale di misure.

Notiamo ancora che nella definizione dello scarto quadratico medio è più corretto sostituire $N - 1\ ad\ N$, ottenendo così:

$$\sigma = \sqrt{\frac{1}{N-1} * \sum_{i=1}^{N}(x_i - \bar{x})^2}$$

La giustificazione teorica di tale sostituzione è legata al fatto che il valor medio della grandezza è stato calcolato a partire dai valori delle misure e quindi, nella sommatoria a numeratore, i termini indipendenti sono in realtà $N - 1$ e non N. Questo, in gergo tecnico, significa la perdita di un grado di libertà , dove il numero di gradi di libertà viene definito come la differenza tra il numero di misure indipendenti ed il numero di parametri calcolati da queste misure.

La deviazione standard così calcolata è, per definizione, un'indicazione di quanto mediamente ciascuna misura è distante dal valore medio e si può assumere come indicatore dell'errore della singola misura. Però ci possiamo attendere che l'errore sulla media sia minore. In effetti il motivo per cui si fanno misure ripetute è proprio per cercare di diminuire l'incertezza dovuta agli errori casuali. E' quindi da aspettarsi che, mentre l'errore sulla singola misura sia approssimativamente sempre lo stesso, l'errore sul valor medio diminuisca all'aumentare del numero delle misure effettuate.

Si introduce, pertanto, una grandezza detta **"deviazione standard della media "**definita da:

$$\sigma_{\bar{x}} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N * (N-1)}}$$

Tale relazione può essere riscritta:

$$\sigma_{\bar{x}} = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \bar{x})^2}{N * (N-1)}} = \frac{1}{\sqrt{N}} * \sqrt{\frac{1}{N-1} * \sum_{i=1}^{N}(x_i - \bar{x})^2} = \frac{\sigma}{\sqrt{N}}$$

Dal che si evince che la deviazione standard della media diminuisce all'aumentare del numero di misure: questa diminuzione non è però lineare, visto che il termine N compare sotto radice. Così ad esempio, se si vuol ridurre il valore della deviazione standard della media di un fattore 10, si dovrà aumentare di un fattore 100 il numero delle misure effettuate.

6. La propagazione degli errori

Fino ad ora abbiamo visto come determinare la precisione di una grandezza misurata direttamente. Spesso però capita che il valore della grandezza che si vuole determinare non è misurabile, ma deve essere ricavato a partire da misure di altre grandezze ad essa correlate.
Ad esempio, supponiamo di dover preparare una serie di soluzioni di diversa concentrazione. Per preparare la soluzione si deve pesare una certa massa di soluto (usando una bilancia analitica di precisione), e scogliere tale soluto in acqua in apposito contenitori tarati, che forniscono quindi la misura del volume. Abbiamo pertanto:

$$\begin{cases} m \pm \delta m \ \ (\mathit{errore\ determinato\ sulla\ base\ della\ precisione\ della\ bilancia} \\ V \pm \ \delta V \ (\mathit{errore\ determinato\ sulla\ base\ della\ precisione\ del\ contenitore}) \end{cases}$$

A questo punto è possibile calcolare la concentrazione di questa soluzione, data da:

$$c = \frac{m}{V}$$

Se si conosce l'errore su m e l'errore su , ci chiediamo quanto vale l'errore sulla concentrazione ?
Ad una prima analisi si potrebbe pensare che l'errore sulla grandezza derivata è la somma degli errori sulle diverse grandezze che compaiono nella formula. Ma no avrebbe senso sommare termini tra loro dimensionalmente diversi (la massa si esprime in Kg, il volume in litri o metri cubi).
La relazione che lega le tre variabili $c, m\ , V$ è una relazione funzionale del tipo:

$$c = f(\, m, V)$$

Generalizzando, consideriamo una relazione funzionale del tipo:

$$y = f(\, x_1, x_2, \dots \dots \dots \ x_N \,)$$

In questo contesto, la serie di variabili $x_1, x_2, \dots \dots \dots \ x_N$ non rappresentano diverse misure della stessa grandezza, bensì le N differenti quantità misurate. Considerando l'esempio precedente, y corrisponde alla concentrazione c, e le quantità misurate sono due x_1 che corrisponde alla massa m e x_2 al volume V.
Si può dimostrare che l'errore sulla grandezza derivata è dato da:

$$\delta_f = \sqrt{\left(\frac{\partial y}{\partial x_1}\delta_{x_1}\right)^2 + \left(\frac{\partial y}{\partial x_2}\delta_{x_2}\right)^2 + \cdots \dots \dots \dots \left(\frac{\partial y}{\partial x_n}\delta_{x_n}\right)^2} = \sqrt{\sum_{i=1}^{N}\left(\frac{\partial y}{\partial x_i}\delta_{x_i}\right)^2}$$

Avendo indicato con il simbolo:

$$\frac{\partial y}{\partial x_i}$$

La derivata parziale della funzione f rispetto alla i-esima variabile x_i e δ_{x_i} l'errore associato alla variabile x_i. Questa espressione è valida nel caso in cui le variabili sono tra loro indipendenti; se non lo fossero, nell'equazione precedente comparirebbero ulteriori termini, detti di covarianza.

7. Conclusioni

Per il calcolo effettivo dell'errore dobbiamo distinguere due casi:
1) Gli errori statistici sono minori del più piccolo intervallo della grandezza che lo strumento può apprezzare (cioè, ripetendo la misura si ottiene sempre lo stesso valore);
2) La sensibilità dello strumento è tale che, ripetendo la misura, si trovano valori diversi.

Nel primo caso (e anche quando la misura della grandezza viene eseguita una sola volta) si assume come errore di misura la semiampiezza o l'ampiezza (a seconda che le suddivisioni siano o no sufficientemente distanziate) dell'intervallo corrispondente a due consecutive suddivisioni della scala dello strumento.

Nel secondo caso, invece, è necessario calcolare uno dei tre errori statistici. Può assumersi come errore statistico la semidispersione se le misure effettuate sono poche (si tratta comunque di un a soluzione di ripiego), mentre negli altri casi è necessario calcolare l'errore medio oppure la deviazione standard. Quest' ultima è più significativa dell'errore medio, essendo più sensibile ad apprezzare anche lievi variazioni nelle misure.

Si dimostra, inoltre, se il numero di misure è sufficientemente elevato, che sussista la seguente relazione tra l'errore medio δ e la deviazione standard σ :

$$\delta = 0{,}79\ \sigma$$

Inoltre una qualsiasi misura ha una probabilità del 68,3 % di essere compresa nell'intervallo $(\bar{x} - \sigma;\ \bar{x} + \sigma)$ e del 99,7 % di appartenere all'intervallo di estremi $(\bar{x} - 3\sigma;\ \bar{x} + 3\sigma)$. per questo motivo si assume come errore massimo il triplo della deviazione standard e si scrive:

$$x = (\overline{x}\ \pm 3\ \sigma).$$

Bibliografia

Peter Holl (1990), Elementi di statistica- il Mulino

Romano Scozzafava (2010), Incertezza e probabilità- Zanichelli

G. Filatrella-P. Romano (2010), Elaborazione statistica dei dati sperimentali- Edises

L. D' Ambra-S. Spedaliere (2013), Statistica descrittiva- RCE Multimedia

Joseph Toscano (2012), Training autogeno in probabilità- Zanichelli

Wayne W. Daniel (2000), Biostatistics: A foundation for analysis in the health sciences- John Wiles e Sons Inc.

Editore: Accademia Piceno - Aprutina dei Velati in Teramo (APAV)
Via del Concilio n. 24, Pescara, Italy
Codice Fiscale: 92036140678 Partita IVA:02184450688

Siti web: www.apav.it ; www.eiris.it
Email: apavsegreteria@gmail.com, apavsegreteria@pec.it

www.apav.it
www.eiris.it

ISBN 978-88-94350-12-8
9 788894 350128