La storia della matematica in classe

Convegno nazionale. L'Aquila 15/10/2015

Il biliardo di Bayes
tra riferimenti storici e spunti di didattica della statistica

Vittorio Colagrande

Liceo Scientifico "G. Galilei" Lanciano

Introduzione

Nel 2013 si celebrava a livello mondiale il 250-mo anniversario della prima formulazione del *Teorema di Bayes*. L'importanza di una tale ricorrenza risiede nel fatto che, a partire dal ragionamento bayesiano, si è sviluppato quell'approccio metodologico all'analisi dei dati costituito dalla *statistica bayesiana.*

In riferimento a questo evento, durante il pi greco day del 2014, organizzato dal Liceo Scientifico di Lanciano e dalla sezione locale della Mathesis, sono state proposte agli studenti del Liceo alcune riflessioni relative all'inferenza statistica secondo la metodologia bayesiana

Scopo di questa comunicazione è quello di riformulare tali considerazioni per fornire uno stimolo per la costruzione di un possibile percorso didattico rivolto agli studenti degli ultimi anni delle scuole superiori su temi di approfondimento di statistica e probabilità, facendo riferimento *alla situazione problematica del biliardo* introdotta da Thomas Bayes nel suo lavoro scientifico principale.

Saggio sulla soluzione di un problema nella dottrina del caso

Nel 1763 viene pubblicato su

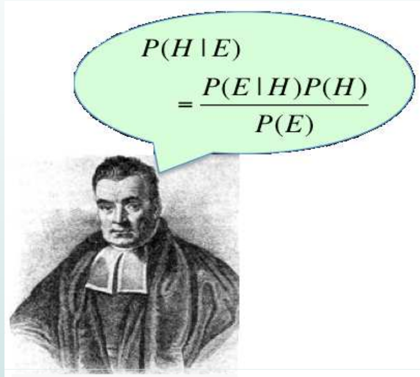
Philosophical Transactions of the
Royal Society of London

il lavoro postumo

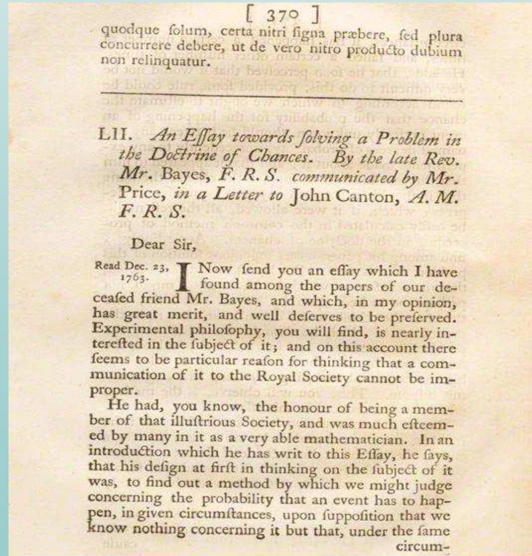
An Essay toward solving a Problem in the Doctrine of Chances

del ministro presbiteriano

Thomas Bayes



Thomas Bayes
(1702-1762)



[370]

quodque solum, certa nitri signa præbere, sed plura concurrere debere, ut de vero nitro producto dubium non relinquatur.

LII. *An Essay towards solving a Problem in the Doctrine of Chances. By the late Rev. Mr.* Bayes, *F. R. S. communicated by Mr.* Price, *in a Letter to* John Canton, *A. M. F. R. S.*

Dear Sir,

Read Dec. 23, 1763.

I Now send you an essay which I have found among the papers of our deceased friend Mr. Bayes, and which, in my opinion, has great merit, and well deserves to be preserved. Experimental philosophy, you will find, is nearly interested in the subject of it; and on this account there seems to be particular reason for thinking that a communication of it to the Royal Society cannot be improper.

He had, you know, the honour of being a member of that illustrious Society, and was much esteemed by many in it as a very able mathematician. In an introduction which he has writ to this Essay, he says, that his design at first in thinking on the subject of it was, to find out a method by which we might judge concerning the probability that an event has to happen, in given circumstances, upon supposition that we know nothing concerning it but that, under the same circum-

Fig. 1 - Prima pagina della lettera di R. Price

Essay

L'*Essay* è una delle due memorie trovate tra gli appunti di Bayes dopo la morte e pubblicate dalla Royal Society per interessamento dell'amico reverendo Richard Price .

Price nella parte iniziale della comunicazione inserisce alcune osservazioni personali per evidenziare l'importanza della soluzione di Bayes al fine di fondare razionalmente ogni ragionamento relativo a fatti passati e a previsioni del futuro.

Ma si spinge ancora oltre, ritenendo il lavoro dell'amico in grado di offrire valide argomentazioni per dimostrare l'esistenza di Dio:

"The purpose I mean is, to shew what reason we have for believing that there are in the constitution of things fixt laws according to which things happen, and that, therefore, the frame of the world must be the effect of the wisdom and power of an intelligent cause; and thus to confirm the argument taken from final causes for the existence of the Deity".

Essay

Nello specifico, l'*Essay* presenta inizialmente una definizione di probabilità:

la probabilità di un evento qualunque è il rapporto tra il valore atteso, calcolato in dipendenza del verificarsi dell'evento, e la probabilità (chance) della cosa attesa al verificarsi dell'evento.

Successivamente tratta di alcune leggi della probabilità affrontando anche la problematica dell'indipendenza degli eventi e della probabilità condizionata.

Essay

Ma scopo dichiarato dell’autore è quello di trattare il problema:

Dato il numero di volte in cui un evento incognito si è verificato e il numero di volte in cui non si è verificato: *è richiesta* la probabilità (*chance*) che la probabilità (*probability*) del suo verificarsi in una singola prova sia compresa tra due gradi di probabilità (“*degrees of probability*”) prefissati.

Singolare risulta il fatto che, nella presentazione e discussione della soluzione, Bayes fa riferimento alla situazione fisica di un tavolo da biliardo.

Essay (sezione II)

Postulato1. Si supponga che il tavolino o il piano quadrato ABCD sia costruito e livellato in modo tale che, se una delle biglie O oppure W è gettata su di esso, sia uguale la probabilità che si fermi in una qualsiasi parte del piano e che si fermi necessariamente in una di esse.

Postulato 2. Si supponga che la biglia W sia gettata per prima; nel punto in cui si ferma si tracci la parallela os al lato AD, che incontra CD e AB in s e o rispettivamente. Poi si getti la biglia O p+q = n volte e, se O si ferma tra os e AD in un singolo lancio, si dice che si è verificato l'evento M in una singola prova .

Essay

Il piano *ABCD* viene suddiviso in rettangoli con un lato su *AB* e l'altro parallelo ad *AD*, in uno di essi, diciamo *bfFL*, cade la linea *os* e il rapporto tra l'area di tale rettangolo e l'area del quadrato, secondo Bayes, può essere assunto come misura della probabilità che il segmento *os* cada tra *bL* e *fF*. La probabilità, poi, dell'evento M, una volta individuata la linea *os*, è data dal quoziente tra l'area del rettangolo *AosD* e quella del quadrato.

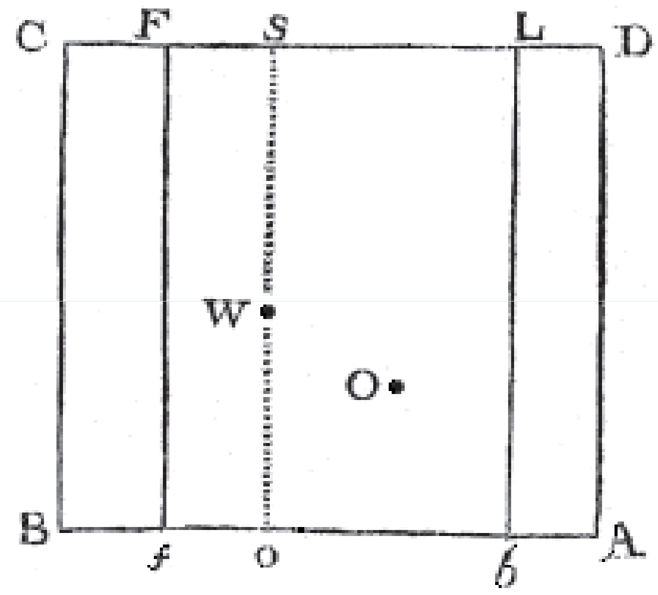


Fig. 2 - Tavolo da biliardo e biglie W e O

Essay

Lemma 1. *La probabilità che il punto **o** cada tra due punti qualsiasi sulla linea AB è pari al quoziente tra la distanza tra questi punti e la lunghezza dell'intero segmento AB.*

Lemma 2. *Gettata la biglia W e tracciata la linea **os**, la probabilità dell'evento M in una singola prova è il quoziente tra Ao e AB.*

L'evento M rappresenta il *successo* in un singolo lancio di biglia e la sua probabilità corrisponde a quella *incognita* relativa alla posizione del punto o.

Dopo alcune costruzioni grafiche e dimostrazioni di tipo geometrico e per assurdo, l'autore è in grado di provare la proposizione fondamentale del lavoro.

Essay

Proposizione 9. *Se, prima di conoscere la posizione del punto o, si ammette che l'evento M si è verificato p volte e non si è verificato q volte in p+q prove, allora posso congetturare che il punto o giace tra due punti qualsiasi f e b della linea AB e conseguentemente che la probabilità dell'evento M in una singola prova è compresa tra il rapporto di Ab su AB e di Af su AB. Tale probabilità è il rapporto di quella parte della figura AiB che è compresa tra le perpendicolari condotte sopra AB in corrispondenza dei punti f e b, rispetto all'intera figura AiB*.

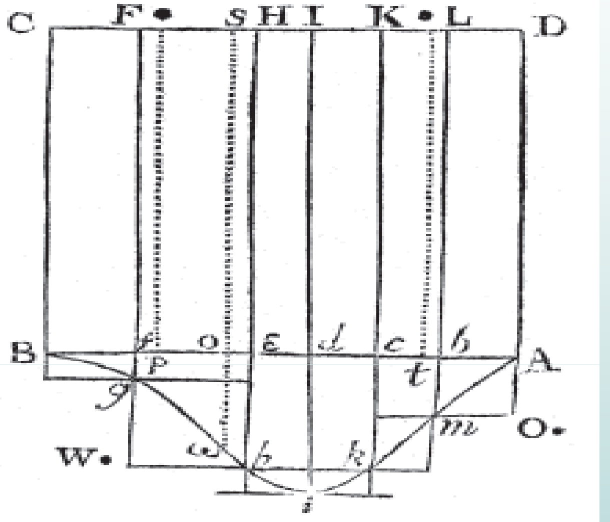


Fig. 3 - Costruzione geometrica sul quadrato ABCD

La figura *AiB* è individuata dal lato AB del quadrato e dalla curva ottenuta considerando un generico punto b su AB, il rapporto Ab/AB e calcolando

$$y = \frac{bm}{AB} = \binom{p+q}{p} x^{p} (1-x)^{q}.$$

Essay

Allo scopo di risolvere il problema di partenza, Bayes, ragionando in termini di rapporti di aree di rettangoli e operando con le serie, dimostra la proposizione conclusiva.

Proposizione 10. Se un evento incognito si è verificato p volte in p+q prove e da due punti f e t della base AH si tracciano le perpendicolari fC e tF fino alla curva di figura 4, la probabilità che la probabilità dell'evento sia compresa tra il rapporto Af su AH e il rapporto At su AH è data dal rapporto tra la parte tFCf di figura compresa tra le ordinate e l'intera figura ACFH.

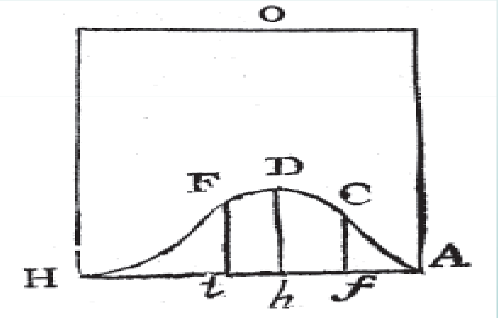


Fig. 4 - Illustrazione per la dimostrazione della prop. 10

La curva di figura è individuata dalla equazione:

$$y = x^p (1 - x)^q$$

Essendo y e x rispettivamente i rapporti tra le misure dei segmenti *hD* e *Ah* e la base *AH*

Essay

L'originalità della proposta consiste nell'esprimere una stima iniziale della (distribuzione di) probabilità dell'evento incognito (*probabilità "a priori"*) e, successivamente, aggiornare tale stima (*probabilità "a posteriori"*) sulla base di nuove informazioni legate all'evento. Nel biliardo viene ipotizzata una misura iniziale della probabilità relativa alla posizione della prima biglia e, sulla base delle posizioni assunte dalla seconda, tale stima viene revisionata producendo la probabilità a posteriori.

Il risultato, noto come *Teorema di Bayes*, deriva essenzialmente dalla *proposizione 9.* Esso permette di affrontare e risolvere il problema dell'*inferenza inversa*: calcolare la probabilità di un evento (ipotesi) sulla base di informazioni note e osservare di quanto cambia tale probabilità alla luce di nuove informazioni.

Essay

Con le applicazioni di tale risultato alle scienze fisiche e naturali effettuate successivamente da *Pierre Simone Laplace* si può dire che il calcolo delle probabilità sia entrato ufficialmente a far parte della metodologia induttiva della scienza.

Nel corso del 19° secolo e nella prima metà del 20° il metodo di Bayes-Laplace è caduto in disuso e anche osteggiato.

Il pensiero bayesiano è stato ravvivato da studiosi quali *Bruno De Finetti* e *Jimmy Jeffreys* e ha dato origine, nella seconda metà del 20° secolo, al moderno movimento della statistica bayesiana guidato in particolare da *Leonard Jimmie Savage* e *Dennis Victor Lindley*.

Tuttavia tale paradigma statistico è risultato molto difficile da mettere in atto fino agli anni ottanta del novecento a motivo soprattutto di complessità computazionali.

Esperimento del biliardo

Riformulando il problema del biliardo

Si consideri un tavolo da biliardo. Dalla sponda AB viene lanciata sul tavolo una biglia W, libera di rotolare e di fermarsi in una qualsiasi posizione, supposta incognita. Un'altra biglia, poi, viene lanciata ripetutamente sul tavolo, allo stesso modo della W; si tiene conto di quante volte essa si ferma a sinistra della prima biglia ("successo") e di quante alla destra ("insuccesso"). Sulla base di tali eventi osservati per la seconda biglia, si stima la (distribuzione di) probabilità della posizione orizzontale (parallelamente ad AB) assunta dalla prima.

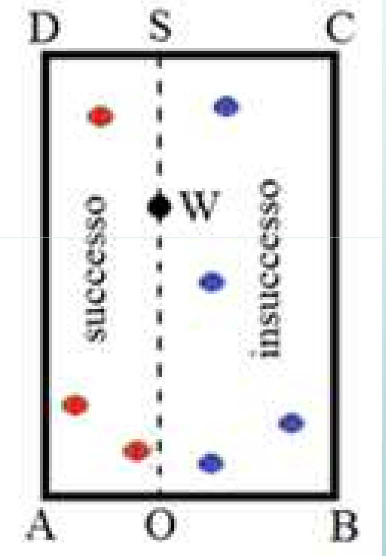


Fig. 5 - Tavolo da biliardo e biglie

Il biliardo

La probabilità della posizione orizzontale di W è la variabile aleatoria

$$\theta = \frac{\overline{AO}}{\overline{AB}}$$

con *densità a priori uniforme* in [0,1]

$$\pi(\theta) = I_{[0,1]}(\theta)\,.$$

Le osservazioni sperimentali relative alla seconda biglia, lanciata ripetutamente ***n*** volte, sono realizzazioni **x** = $(x_1, x_2, \ldots, x_n)$ del vettore aleatorio **X**= $(X_1, X_2, \ldots, X_n)$ dato da:

$$X_i = \begin{cases} 1 & \text{biglia a sinistra} \\ 0 & \text{biglia a destra} \end{cases}$$

Il biliardo

La variabile aleatoria “n° di successi su n prove”

$$Z = \sum_{i=1}^{n} X_i$$

per *θ fisso,* ha *distribuzione binomiale* e, quindi, la *probabilità di z successi sugli n lanci* della seconda biglia è data da:

$$f(z|\theta) = \binom{n}{z} \theta^z (1-\theta)^{n-z}.$$

Osservati Z=z successi, attraverso il *Teorema di Bayes* nella formulazione per *variabili casuali continue*, si ricava la *densità di probabilità a posteriori* di *θ*:

$$\pi(\theta|z) = \frac{\pi(\theta) \cdot f(z|\theta)}{P(Z=z)} = \binom{n}{z} \frac{\theta^z (1-\theta)^{n-z} I_{[0,1]}(\theta)}{P(Z=z)}$$

$$P(Z=z) = \binom{n}{z} \int_0^1 \psi^z (1-\psi)^{n-z} \, d\psi$$

Il biliardo

In primo approccio didattico per il calcolo dell'integrale

$$\Psi(\theta) = \int_0^1 \psi^z (1-\psi)^{n-z} d\psi$$

si può osservare che

$$\Psi(z) = \begin{cases} \dfrac{1}{n+1} & \text{se } \ z = 0 \\ \dfrac{z}{n-z+1}\Psi(z-1) & \text{se } z \geq 1 \end{cases}$$

e quindi $\Psi(z) = \dfrac{z!(n-z)!}{(n+1)!}$.

Pertanto la densità a posteriori è data da

$$\pi(\theta|z) = \frac{(n+1)!}{z!(n-z)!}\theta^z(1-\theta)^{n-z}$$

Il biliardo

Introducendo, poi, le funzioni gamma e beta:

$$\Gamma(t) = \int_0^{+\infty} e^{-t} y^{t-1} dy \qquad B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}$$

risulta $\Psi(z) = B(z+1, n-z+1)$ e quindi

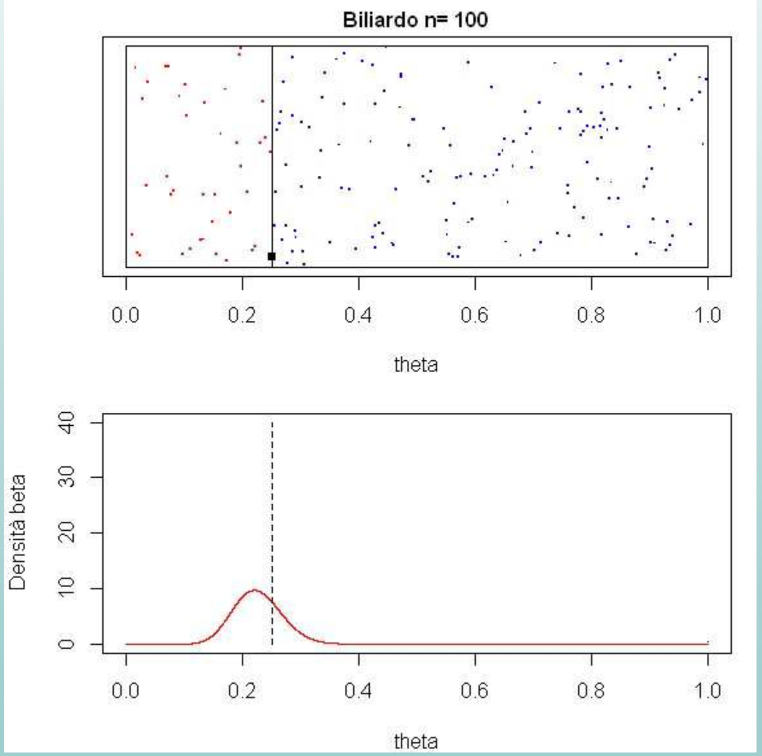
$$\pi(\theta|z) = \frac{\theta^{z}(1-\theta)^{n-z}}{\mathrm{B}(z+1, n-z+1)}$$

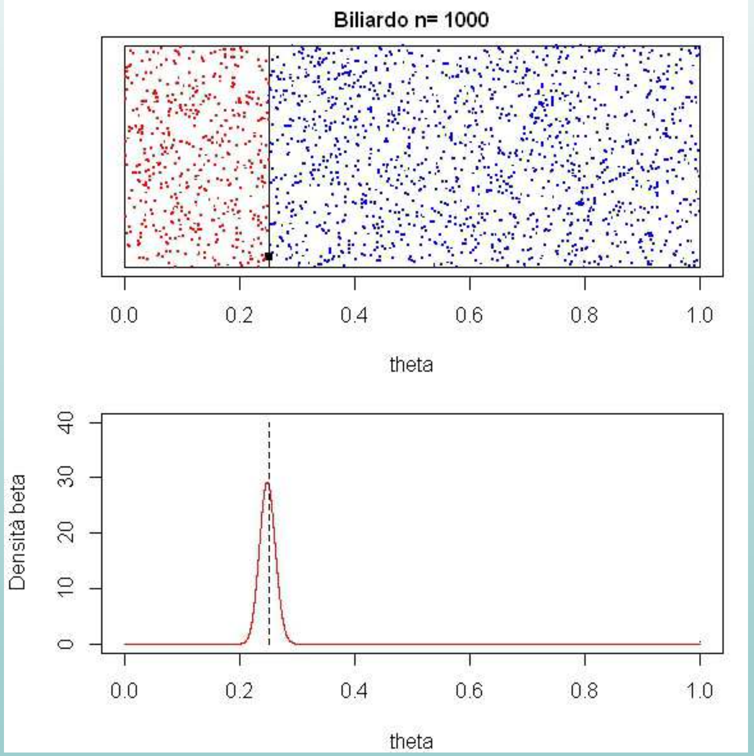
e si tratta di una *densità di probabilità beta* di parametri

$\alpha = z+1$ e $\beta = \mathrm{n} - z + 1.$

Simulazione “didattica” del biliardo

Si riportano di seguito i risultati di una simulazione effettuata attraverso un programma (*script*) nell’*ambiente statistico* ***R*** per un numero di lanci della seconda biglia pari a n=100 ed n =1000.





n = 100 z = 22

n = 1000 z = 247

Il biliardo

Per quanto attiene la richiesta del problema di Bayes, va calcolato l'integrale della *densità a posteriori* tra due prefissati valori θ_1 e θ_2 del parametro θ

$$P(\theta_1 < \theta < \theta_2 | z) = \int_{\theta_1}^{\theta_2} \pi(\theta | z)\, d\theta$$

Tralasciando il calcolo analitico, per un esempio numerico si possono utilizzare i risultati della simulazione precedente per n=1000 e la funzione *pbeta*(x,*α*,*β*) del software ***R*** calcolata in corrispondenza di $\theta_1 = 0.221$ e $\theta_2 = 0.274$

$$P(0.221 < \theta < 0.274 | 247) = P(0 < \theta < 0.274 | 247) - P(0 < \theta < 0.221 | 247) =$$

$$pbeta(0.274, 248, 754) - pbeta(0.221, 248, 754) \approx 95\%.$$

Il biliardo

In vari contesti operativi è opportuno sintetizzare la distribuzione ottenuta con degli indici, quali la moda, il valore medio o la mediana. Spesso si fa riferimento al *valor medio*

$$E(\theta|z) = \frac{(n+1)!}{z!\,(n-z)!}\int_0^1 \theta^z (1-\theta)^{n-z}\,\theta\, d\theta = \frac{z+1}{n+2}.$$

Nel caso delle simulazioni, per n = 100 e per n = 1000 il valor medio è risultato pari, rispettivamente, a 0.2255 e 0.2475, mentre il valore "vero" del parametro θ è 0.2503.
Si tratta di una stima puntuale, *ma* l'informazione risulterebbe più completa se venisse considerata una stima intervallare attraverso gli *intervalli di credibilità* $I = (\theta_1, \theta_2)$

$$P(\theta \in I|z) = P(\theta_1 < \theta < \theta_2|z) = probab.\ prefissata.$$

Il biliardo

Si osservi ora che la distribuzione uniforme assunta come distribuzione a priori è un caso particolare della beta per n=z=0.

Partendo dalla distribuzione a posteriori ottenuta in predendenza, il processo può essere reiterato assumendo tale distribuzione come informazione iniziale, effettuando ulteriori lanci di biglie e aggiornando nuovamente la distribuzione di probabilità.

La densità a posteriori che si ottiene è ancora una distribuzione beta dipendente dalla distribuzione assunta come a priori, dal numero di lanci n_1 effettuati e dal numero di successi z_1

$$\pi\left(\theta \middle| z, z_1\right) = \frac{\theta^{z+z_1}(1-\theta)^{n+n_1-(z+z_1)}}{\mathrm{B}(z+z_1+1, \mathrm{n}+\mathrm{n}_1-z-z_1+1)}.$$

Cenni di inferenza statistica

$\theta \in \Omega$ variabile aleatoria (parametro).

L'informazione iniziale su di essa può essere espressa mediante una "opportuna" *densità di probabilità a priori*

$$\pi(\theta).$$

L'informazione campionaria viene analizzata considerando realizzazioni **x** di un opportuno vettore casuale **X** e individuando la *funzione di verosimiglianza:*

$$f(\mathrm{x}|\theta).$$

Se *f(**x**,θ)* è la *densità di probabilità congiunta* di **X** e *θ*, per la probabilità condizionata risulta

$$f(\mathrm{x},\theta) = f(\mathrm{x}|\theta)\cdot \pi(\theta) = \pi(\theta|\mathrm{x})\cdot f(\mathrm{x}).$$

Cenni di inferenza statistica

Si ricava, quindi, la formula di Bayes che fornisce la *densità di probabilità a posteriori*

$$\pi\left(\theta \middle| \mathrm{x}\right) = \frac{\pi(\theta)\cdot f\left(\mathrm{x}\middle|\theta\right)}{f(\mathrm{x})}$$

essendo $f(x) = \int_{\Omega} f(x|\theta)\cdot\pi(\theta)\, d\theta$.

Si può quindi vedere come il livello di conoscenza ad un dato istante si incrementa attraverso l'acquisizione di nuovi dati (le informazioni campionarie), dati che permettono l'aggiornamento della conoscenza mediante il passaggio dalla distribuzione di probabilità a priori alla distribuzione di probabilità a posteriori, che a sua volta può costituire la distribuzione a priori in una fase successiva del processo di apprendimento dall'esperienza.

Conclusione

Aspetti inferenziali specifici, senza dubbio, possono difficilmente entrare a far parte di argomenti curriculari o anche di approfondimento delle scuole superiori, salvo in quegli indirizzi di studio che prevedono l'insegnamento specifico della statistica. In questi può essere significativo introdurre elementi di statistica bayesiana anche per un confronto con l'approccio classico, frequentista all'inferenza.

In ogni caso, la proposta presentata può avere una valenza didattica in diversi corsi di studio ai fini soprattutto di favorire ulteriormente lo sviluppo da parte dello studente della logica dell'incerto ma anche per fornire esempi applicativi di metodiche acquisite in altri contesti matematici, quale quello dell'analisi.

Bibliografia

- D.M. Cifarelli, P. Muliere, *Statistica Bayesiana: Appunti ad uso degli studenti*, G. Iuculano, 1989.
- G. D'Agostini, D. Esposito, *Così è… probabilmente, il saggio, l'ingenuo e la signorina Bayes*, www.ilmiolibro.it, 2014.
- A.I. Dale, *A History of Inverse Probability: From Thomas Bayes to Karl Pearson*, Springer-Verlag, 1991.
- *International Society for Bayesian Analysis*, URL https://bayesian.org/
- R Core Team, *R: A language and environment for statistical computing*, URL http://www.R-project.org/, 2015.
- S.M. Stigler , *The True Title of Bayes's Essay*, Statistical Science, 28,3,283,288.
- M. Valli, *La Rivoluzione di Bayes*, Il Saggiatore, 2013.