Ottavio Serra

FORMULA DI BAYES

Introduzione. Definizione di probabilità e proprietà.

La formula di Bayes è un'applicazione del concetto di ***Probabilità condizionata.***

Ricordo che la probabilità di un evento ***A*** è relativa a un insieme di eventi **S**, detto ***spazio degli eventi***, di cui ***A*** fa parte. Introdotta ***una m*** su **S**, $\geq \mathbf{0}$, la probabilità p di A è definita come esegue:
Un evento è un sottoinsieme di S; S stesso e l'insieme vuoto V sono particolari sottoinsiemi di S.
Definiamo ora la probabilità p di A come segue.
[1] p(A)=m(A)/m(S).

Ovviamente, p(S)=1, **S è l'evento certo, e P(V)=0 (l'insieme vuoto è l'evento impossibile.**

Probabilità totale. Siano E ed F due eventi di S. La probabilità dell'unione di E ed F non è la somma delle singole probabilità, perché in tal caso gli eventuali elementi comuni verrebbero contati due volte. Perciò

[2] $\mathbf{p}(\mathbf{E} \cup \boldsymbol{F}) = \boldsymbol{p}(\boldsymbol{E}) + \boldsymbol{p}(\boldsymbol{F}) - \boldsymbol{p}(\boldsymbol{E} \cap \boldsymbol{F})$.

Esempio 1. Sia S l'insieme dei numeri naturali da 0 a 9 (estremi inclusi), la misura m sia il conteggio: m(S)=10. Si E l'insiemi dei numeri primi di S: m(E)=4; sia F l'insieme dei numeri pari di S: m(F)= 5.

Estraendo un numero a caso da S, vogliamo calcolare la probabilità che esca un numero primo o un numero pari. La formula [2] fornisce p=4/10+5/10-1/10=8/10. Infatti, E ed F hanno il numero **2** in comune.

Esempio 2. In una regione piana S di 1000 cm^2 sono disegnati due quadrati di lato 10 cm e 15 cm, che hanno in comune un quadratino avente il lato di a cm. La probabilità che un punto scelto a caso in S cada in uno dei due quadrati è $\mathbf{p=(100+225-a^2)/1000 = 0{,}325 - a^2/1000.}$ Calcolare p nel caso che a assuma il valore massimo o minimo

Se nella [2] $\boldsymbol{E} \cap \boldsymbol{F} = \boldsymbol{V}$, i due eventi si dicono $\boldsymbol{incompatibili}$ *e* $p(\mathbf{E} \cup \boldsymbol{F}) = \boldsymbol{p}(\boldsymbol{E}) + \boldsymbol{p}(\boldsymbol{F})$.

Esercizio. Calcolare la probabilità dell'unione di tre eventi di S.

Probabilità condizionata. Siano A e B due eventi di S. La probabilità p(A)=m(A)/m(S) e analogamente p(B)=m(B)/m(S). Si chiama **probabilità di A condizionata a B,** e si indica con **p(A/B),** la probabilità di A quando si sa, o **si fa l'ipotesi**, che si è verificato B**.** Le cose vanno come se lo spazio S si fosse contratto riducendosi a B e quindi non tutto A può oramai verificarsi, ma solo la parte di A inclusa in B.

Perciò, $p(A/B) = m(A \cap B)/m(B)$. Dividendo poi sopra e sotto per m(S), si ha infine

$$[3]\, p(A/B) = p(A \cap B)/p(B).$$

Scambiando le veci di A e B, si ottiene analogamente

$$[4]\, p(\mathrm{B}/A = p(\mathrm{B} \cap A)/p(\mathrm{A}).$$ Siccome l'intersezione (la congiunzione) è commutativa, segue

$$[5]\, p(A/B).\mathrm{p(B)} = p(\mathrm{B/A}).\mathrm{p(A)}.$$

Questa formula è molto importante.

Se p(A/B)=p(A), si dice che A è indipendente da B; in tal caso la [5] ci dice che anche B è indipendente da A: A e B si dicono indipendenti e $p(A \cap B) = p(A).p(B).$ (vedi la [3] e la [4]).

Esempio 1. (a**)** Calcolare la probabilità di estrarre un ***Re*** da un mazzo di carte napoletane. (R. 1/10).

(b) Come in (a), sotto la condizione che la carta sia di spade. (R. 1/10)).

(c) Come in (a), condizionata all'ipotesi che la carta estratta sia una figura. (R. 4/12=1/3).

(d) Probabilità di estrarre il, ***Re*** di spade. (R. 1/40).

(e) Come in (d), se la carta estratta è una figura. (R. 1/12).

(f) Come in (d), se la carta estratta è di coppe. (R. 0).

Esempio 2. Un poliziotto va ad allenarsi al poligono di tiro. Si supponga che la probabilità che colpisca il bersaglio al primo tiro sia p. **Qual è la probabilità** che, su tre tiri, colpisca il bersaglio

(a) Tre volte. (b) Nessuna volta. (c) Una sola volta. (d) Almeno una volta. (e) Al più una volta.

(f) Esattamente due volte. (g) Almeno due volte. (h) Al più due volte.

Fate delle *buone* ipotesi sulla probabilità di successo nei tiri successivi, basandovi sul **buon senso**, sulla **semplicità** e sull'**ignoranza** dei dettagli.(Per me, l'ipotesi più semplice è che p sia la stessa in tutti i tiri. Un'altra ipotesi, suggerita implicitamememte, è che i casi siano due: il bersaglio è colpito, Prob.=p) o non è colpito, Prob.=q=1-p).

Esempi 3. Su un pavimento orizzontale è poggiato un tavolo rettangolare con dimensioni di 4 m e 5 m, nel quale è tracciata una linea chiusa molto complicata, che delimita una parte F del tavolo. Si lancia una pallina (a caso) nel rettangolo. Calcolare la **probabilità** che la pallina si fermi nell'interno della **figura F**. (**Suggerimento.** Se F riempisse l'intero rettangolo, se occupasse la metà, se fosse trascurabilmente piccola rispetto al rettangolo, …).

Se si ripete il lancio della pallina **molte volte,** la frequenza relativa dei successi (numero f delle volte in cui la pallina si ferma in F diviso il numero totale n dei lanci) si approssima **con probabilità crescente** alla **probabilità** che si chiede di calcolare. ***Legge dei grandi numeri. Al crescere di n la frequenza relativa dei successi f/n tende alla probabilità p.*** Queste considerazioni conducono a metodi ***statistici*** per calcolare integrali complicati e in spazio multidimensionali, che si chiama, per ovvi motivi, ***metodo Monte Carlo***. (Quello suggerito è il più semplice: *Hit or Miss*).

La formula di Bayes. *Imparare dall'esperienza*

Partizione di un insieme S: è una famiglia $\{A_1, A_2, ..., A_n\}$ di sottoinsiemi di S che gode delle seguenti proprietà: nessun A_k è vuoto, gli A_k a due a due sono disgiunti; l'unione degli A_k è S. Da tali proprietà segue, per ogni sottoinsieme B di S

$$[6] B = B \cap S = B \cap \{A_1 \cup A_2 \cup ... \cup A_n\} = (B \cap A1) \cup \cup (B \cap A2) \cup .... \cup (B \cap An)$$

6Interpretiamo, ora, S come spazio di eventi, a partizione come una partizione di S come una partizione di eventi, B come un evento speciale e supponiamo di conoscere, per ogni indice k da 1 a n, la probabilità $p(A_k)$ e la probabilità condizionata $p(B/A_k)$. Da questi dati vorremmo ricavare la probabilità condizionata **$p(A_k/B)$**. Ciò è possibile utilizzando le formule [4], [5] e [6].

La formula che otterremo è la formula di Bayes.

$$[7]p(A_k / B) = \frac{p(B / A_k).p(A_k)}{p(B)} = \frac{p(B / A_k).p(A_k)}{\sum_{i=1}^{n} p(B / A_i).p(A_i)}.$$

Esempio 1. Si hanno 200 mele e 100 banane, ma sono acerbe 90 mele e 30 banane. Si estrae a caso un frutto e risulta acerbo. Qual è la probabilità che si tratti di una mela?

Schematizziamo i dati: n=2, A_1 sono le mele, A_2 le banane; B i frutti acerbi.

$p(A_1)$ =2/3, $p(A_2)$=1/3, $p(B/A_1)$=9/20, $p(B/A_2)$=3/10. La probabilità richiesta è

$p(A_1/B)$=(9/20).(2//3)/(120/300)=**3/4.**

Verificare che la probabilità che si tratti di una banana, sapendo che è acerba è **¼.**

Esempio 2. Un’industria automobilistica ha tre stabilimenti: in una settimana il primo produce 1200 auto, il econdo1000 e il terzo 800. Risultano difettose, in media, 30 auto prodotte nel primo stabilimento, 22 nel secondo e 19 nel terzo. Un cliente acquista un’auto che risulta difettosa. Calcolare la probabilità che provenga dal primo stabilimento, dal secondo, dal terzo.

Schema: S=3000 (auto/settimana), $p(A_1)$ =1200/3000=2/5=0, 4; $p(A_2)$ =1000/30000=1/3=0,333; $p(A_3)$=800/3000=4/15=0,2666. Assumendo le frequenze di auto difettose come probabilità condizionate, abbiamo $p(B/A_1)$=30/1200=1/40=0,025; $p(B/A_2)$=22/1000=0,022; $p(B/A_3)$=19/800=0,02375.

Infine, p(B)=0,025.0,4+0,022.0,3333+0,02375.0,2555=0,02366. Si trova, perciò

$p(A_1/B)$=0,025.0,4/0,02366=**0,4225** (il **42 %**);

$p(A_2/B)$=0,022.0,3333/0,02366=**0,30986 (il 31 %);**

$p(A_3/B)$=0,02375.0,2666/0,023666=0,2675 (il **27 %**).

Notare che la somma delle tre probabilità condizionate è 1 (100%), come è giusto. Se un’auto ***Spellantis*** è difettosa, deve ***certamente*** provenire da uno degli stabilimenti ***Spellantis.***

<u>Per approfondire:</u> veda nel sito ***http://digilander.libero.it/ottavioserra0*** la cartella *Lezioni allo Scorza*, in particolare le sezioni Matematica 2010 e Probabilità 2016.