Calcolo delle probabilitá II

November 24, 2017

1 Correlazione positiva e negativa

Definizione 1.1 *Dati due eventi H_1 ed H_2 diremo che l'evento H_2 è* **correlato positivamente** *con l'evento H_1 se la probabilitá che esso si verifichi aumenta una volta che si è verificato H_1 cioè :*

$$p(H_2|H_1) > p(H_2).$$

Esempio 1.2 $U = \{TT, TC, CT, CC\}$, $H_1 = \{TC, CT, CC\}$*: è uscito almeno una volta croce.* $H_2 = \{CC\}$*: uscita di due volte croce.*

$p(H_2) = 1/4$,

$$p(H_2|H_1) = \frac{p(H_1 \cap H_2)}{p(H_1)} = \frac{1/4}{3/4} = \frac{1}{3}$$

quindi $p(H_2|H_1) > p(H_2)$.

E' facile verificare che se $p(H_2|H_1) > p(H_2)$, anche H_1 è correlato positivamente con H_2 cioè $p(H_1|H_2) > p(H_1)$

Definizione 1.3 H_2 *è* **correlato negativamente con** H_1 **se** $p(H_2|H_1) < p(H_2)$**.**

Se H_2 è correlato negativamente con H_1 lo stesso accade per H_1 con H_2.

2 Formula delle probabilitá totali

Sia S l'insieme universo, cioè l'insieme degli eventi elementari. Sia E un evento qualunque ($E \in \mathcal{P}(S)$). Si consideri una partizione di S, $\{C_1, C_2, \ldots, C_n\}$. L'evento E si puó scrivere come somma logica di pezzettini dei C_i con $i = 1, \ldots, n$. Allora

$$p(E) = p(\cup_i (E \cap C_i)) = \sum_i p(E \cap C_i) = \sum_i p(C_i)p(E|C_i)$$

ove la seconda uguaglianza si ha perchè gli eventi $E \cap C_i$ sono incompatibili, cioè ad intersezione vuota; la terza uguaglianza si ha per il teorema della probabilitá composta.

Esempio 2.1 *Calcolare la probabilitá che il secondo estratto del gioco del lotto sia 12. Risulta* $p(x_2 = 12) = 1/90$ *intuitivamente perchè non conosco il primo estratto. Verifichiamolo. Consideriamo una partizione di* $S = \{1, \ldots, 90\}$.

$$H = \{x_1 = 12\}, \text{ cioè esca 12 al primo estratto. } p(H) = 1/90$$

$$\overline{H} = \{x_1 \neq 12\} \text{ allora } p(\overline{H}) = 89/90.$$

Sia $E = \{x_2 = 12\}$, *cioè esca 12 al secondo estratto, allora*

$$p(E|H) = 0 \text{ mentre } p(E|\overline{H}) = 1/89$$

per la formula delle probabilitá totali ho:

$$p(E) = p(H)p(E|H) + p(\overline{H})p(E|\overline{H}) = \frac{1}{90} \cdot 0 + \frac{89}{90} \cdot \frac{1}{89} = \frac{1}{90}$$

3 La formula di Bayes

Sia E un evento tale che $p(E) > 0$, Si consideri una partizione di S, $\{C_1, C_2, \ldots, C_n\}$. Per il teorema della probabilitá condizionata ho:

$$p(C_i|E) = \frac{p(C_i \cap E)}{p(E)} = \frac{p(C_i)p(E|C_i)}{\sum_j p(C_j)p(E|C_j)}.$$

Esempio 3.1 *Un negozio è fornito di lampadine da due ditte A e B mescolate in due scatote.*

Si sceglie a caso tra le due scatole e viene estratta una lampadina della ditta A.

Nella Scatola 1 ci sono 30 lampadine di A e 15 di B. Nella Scatola 2 ci sono 20 lampadine di A e 30 di B.

Calcolare la probabilitá che la lampadina scelta provenga dalla Scatola 1.

E: estrazione di una lampadina di A.

H_1 : *Scatola 1.*

H_2 : *Scatola 2.*

H_1 *ed* H_2 *sono una partizione dell'insieme delle lampadine. Allora:*

$$p(H_1|E) = \frac{p(H_1)p(E|H_1)}{p(H_1)p(E|H_1) + p(H_2)p(E|H_2)} = \frac{1/230/45}{1/230/45 + 1/220/50} = \frac{5}{8}$$

ove $p(H_1) = 1/2$ *è la probabilitá di scegliere la scatola 1 tra le due scatole.*

4 Le variabili casuali

Una **variabile aleatoria (o casuale)** X è una grandezza che è ben definita in un range di possibili valori, ma non so quale di questi valori assumerá . Il valore da essa assunto dipenderá dai possibili risultati aleatori di un esperimento.

Dato un generico esperimento, consideriamo l'insieme universo U costituito dagli eventi elementari. Siano poi E_i eventi aleatori qualsiasi tali che $\{E_1, \ldots, E_n\}$ sia una partizione di U. A ciascun evento E_i, il cui verificarsi determina quel particolare valore della variabile aleatoria X, si puó associare il corrispondente valore della probabilitá $p_i = p(E_i)$.

L'insieme dei valori di p_i prende il nome di **distribuzione di probabilitá .**

Ovviamente risulta

$$\sum_i p_i = 1$$

perchè gli E_i sono una partizione di U.

Se l'insieme dei valori assunti da una variabile casuale è costituito da punti isolati (in numero finito) diremo che la variabile casuale è **discreta**. In caso contrario diremo che è **continua.**

5 Variabili casuali discrete

Una variabile casuale discreta X è una quantitá variabile che puó assumere gli n valori $x_1, \ldots, x_n$ al realizzarsi di n eventi, incompatibili la cui unione è tutto U, $E_1, \ldots E_n$ di probabilitá rispettivamente $p_1, \ldots, p_n$.

Esempio 5.1 *Nel gioco del lancio di un dado supponiamo fissate le seguenti modalitá :*

E_1*: esce un numero maggiore di 4 (si vince 100 lire* $x_1 = 100$*).*

E_2*: esce il numero 1 (si vince 300 lire* $x_2 = 300$*).*

E_3 *: uscita di un numero diverso dai precedenti (si perde 100 lire* $x_3 = -100$*).*

La vincita del giocatore è una variabile casuale X *che possiamo rappresentare in questo modo:*

X	-100	100	300
p	$1/2$	$1/3$	$1/6$

6 Funzione di distribuzione e di ripartizione di una variabile casuale

L'insieme delle probabilitá degli n eventi E_i prende il nome di **funzione di distribuzione di probabilitá .**

Nel caso precedente la rappresentazione grafica è la seguente:

Data una variabile casuale X si consideri la probabilitá che la variabile assuma un valore non superiore ad x, cioè considero $p(X \leq x)$. Ad ogni $x \in R$ associo $F(x) = p(X \leq x)$ ottenendo cosí la **funzione di ripartizione di una variabile casuale discreta.**

$$\begin{array}{cccccccc} x_1 & x_2 & \dots & x_i & x & x_{i+1} & \dots & x_n \\ p_1 & p_2 & \dots & p_i & & p_{i+1} & p_n & \end{array}$$

Allora

$$p(X \leq x) = p_1 + p_2 + \ldots + p_i = \sum_{k=1}^{i} p_k = F(x)$$

ove si considera la somma dei p_i essendo gli eventi E_i a due a due disgniunti.
La funzione di ripartizione $F(x)$ gode delle seguenti proprietá :

- $F(x) = 0$ per $x < x_1$;
- $F(x) = 1$ per $x \geq x_n$;
- $F(x)$ è cresente ed è $0 \leq F(x) \leq 1$;
- Il grafico di $F(x)$ è una funzione a gradini.

Esempio 6.1 *Considerando la funzione di distribuzione dell'esempio precedente, la funzione di ripartizione $F(x)$ è :*

X	-100	100	300
$\sum p_i$	$1/2$	$5/6$	1

infatti il primo valore di probabilitá non essendo preceduto da altri rimane inalterato, il secondo si ottiene sommando p_1 e p_2 ossia $1/2 + 1/3$, il terzo vale 1, essendo la somma di tutte le p_i.

Ci proponiamo ora di ricavare la funzione di distribuzione partendo dalla conoscenza della funzione di ripartizione $F(x)$.

La funzione di ripartizione delle probabilitá $F(x)$ è chiaramente discontinua proprio nei punti x_i.

Qual è il significato probabilistico di una discontinuitá ?

La misura del salto di discontinuitá in un punto x_i non è altro che la probabilitá dell'evento E_i a cui è associato il valore x_i. Infatti tale misura si ottiene calcolando $F(x_i) - F(x_{i-1}) = p(X \leq x_i) - p(X \leq x_{i-1}) = p(X = x_i) = p_i$.

Nei punti x di continuitá di $F(x)$ tale salto è nullo e quindi è nulla la probabilitá . Pertanto, nella funzione di distribuzione della probabilitá , la variabile aleatoria X non assumerá mai un tale valore x.

7 Valor medio di una variabile casuale

Il **valor medio** di una variabile casuale X è dato dalla somma dei prodotti di ciascun valore della variabile per la rispettiva probabilitá . Quindi

$$M(X) = x_1 p_1 + \ldots + x_n p_n.$$

Ad esempio nel caso precedente è $M(X) = -1001/2 + 1001/3 + 3001/6 = 100/3$.

Il valor medio $M(X)$ rappresenta il valore che mediamente si realizza eseguendo un grande numero di prove.

Infatti supponiamo che su n prove il valore x_1 si presenti n_1 volte, il valore x_2 si presenti n_2 volte, ecc., ..., il valore x_m si presenti n_m volte. Allora $n = n_1 + \ldots + n_m$ e

$$M(X) = \frac{x_1 n_1 + \ldots + x_m n_m}{n} = x_1 \frac{n_1}{n} + \ldots x_m \frac{n_m}{n}$$

I rapporti n_i/n sono le frequenze relative dei valori x_i pertanto è

$$M = x_1 f_1 + \ldots + x_n f_n.$$

Se n è abbastanza grande, per la legge dei grandi numeri, posso confondere le frequenze con le probabilitá .

Osserviamo che ovviamente $M(a) = a$, il valor medio di una costante è uguale alla costante stessa.

Operazioni con le variabili casuali.

a) *somma di una variabile casuale X e di una costante a.*

Data la variabile X che assume i valori $x_1, \ldots, x_n$ aventi rispettivamente probabilitá $p_1, \ldots, p_n$ la variabile casuale

$$X + a$$

assume valori

$$x_1 + a, \ldots, x_n + a$$

con le stesse probabilitá dei corrispondenti valori di X. Inoltre

$$M(X + a) = (x_1 + a)p_1 + \ldots (x_n + a)p_n = M(X) + a.$$

b) *Prodotto di una variabile casuale per una costante.*

$$M(aX) = aM(X).$$

c) *Quadrato di una variabile casuale*

$$M(X^2) = x_1^2 p_1 + \ldots x_n^2 p_n.$$

d) *Variabile casuale $aX + b$.*

$$M(aX + b) = aM(X) + b.$$

e) *Somma di due variabili casuali X ed Y.*

$$M(X + Y) = M(X) + M(Y).$$

8 Varianza e scarto quadrato medio.

Sia $M(X) = m$ il valor medio della variabile aleatoria X che puó assumere i valori $x_1, \ldots, x_n$. Consideriamo la varibile $(X-m)^2$, detta **varibile casuale scarto al quadrato**, che potrá assumere i valori:

$$(x_1 - m)^2, \ldots, (x_n - m)^2$$

con probabilitá

$$p_1, \ldots, p_n.$$

Allora il valor medio della variabile aleatoria $(X-m)^2$ è detto **varianza** di X ed è uguale a

$$var(X) = M((X-m)^2) = (x_1-m)^2 p_1 + \ldots (x_n-m)^2 p_n.$$

Si dimostra che:

$$var(X) = M(X^2) - M(X)^2.$$

Si definisce **scarto quadratico medio** o **deviazione standard**

$$\sigma(X) = \sqrt{var(X)}.$$

9 Distribuzione uniforme

Si ha una variabile casuale con distribuzione uniforme quando la variabile X assume i valori

$$1, 2, \ldots, n$$

e le probabilitá sono tutte uguali tra loro ed uguali a $1/n$.

Il valor medio di X è

$$M(X)\frac{1+n}{2}.$$

Infatti

$$M(X) = 1\frac{1}{n} + 2\frac{1}{n} + \ldots + n\frac{1}{n} = \frac{1}{n}(1+2+\ldots+n) = \frac{1}{n}\frac{n(n+1)}{2} = \frac{1+n}{2}.$$

La varianza di X è :

$$var(X) = \frac{n^2-1}{12}.$$

10 Distribuzione binomiale o di Bernoulli

Bisogna premettere alcune cose.

10.1 Il problema delle prove ripetute

Un problema molto frequente in probabilitá è quello di chiedersi:

Qual è la probabilitá che un dato evento E si presenti k volte su n prove effettuate $(n \geq k)$.

Per esempio, lanciando 4 volte un dado calcolare la probabilitá che il numero 6 si presenti esattamente 3 volte considerando i due casi:

a) Che si tenga conto dell'ordine con cui le facce si presentano.

b) Che non si tenga conto dell'ordine.

Caso a) I casi possibili sono:

$$666f, 66f6, 6f66, f666$$

dove f indica un numero diverso da 6. Pertanto il numero dei casi possibili è

$$\binom{4}{3} = 4.$$

In questo caso, l'evento E, uscita del numero 6 tre volte su quattro lanci tenendo conto dell'ordine, è uno soltanto dei quattro casi possibili e risulta

$$p(666f) = p(66f6) = p(6f66) = p(f666) = p(f)p(6)p(6)p(6) = p(f)p(6)^3 = p(f)^{4-3}p(6)^3 = 5/6(1/$$

trattandosi in ciascun caso di eventi indipendenti.

Caso b) In questo caso l'evento E, cioè l'uscita di tre volte il numero 6 su quattro lanci senza tener conto dell'ordine, è dato da:

$$666f \text{ oppure } 66f6 \text{ oppure } 6f66 \text{ oppure } f666$$

e risulta

$$p(E) = p(6)^3p(f)^{4-3} + p(6)^3p(f)^{4-3} + p(6)^3p(f)^{4-3} + p(6)^3p(f)^{4-3} = 4p(6)^3p(f)^{4-3}$$

e poichè

$$\binom{4}{3} = 4$$

si puó scrivere che

$$p(E) = \binom{4}{3} p(6)^3p(f)^{4-3}.$$

In generale, indicando con p la probabilitá dell'evento elementare in un singolo lancio (ad es. uscita 6) e con q la probabilitá dell'evento contrario (ad es. uscita un numero $\neq 6$) i due casi sono:

a) Se si tiene conto dell'ordine

$$p_{n,k} = p^k q^{n-k}.$$

b) Se non si tiene conto dell'ordine

$$p_{n,k} = \left(\begin{array}{c} n \\ k \end{array} \right) p^k q^{n-k}.$$

Esempio 10.1 *Lanciando 10 volte un dado calcolare la probabilitá che la faccia 6 si presenti le solo prime quattro volte. Risulta*

$$p_{10,4} = \frac{1}{6}^4 \frac{5}{6}^6.$$

Riprendiamo a parlare della distribuzione binomiale.

Se E è l'evento che si deve ripetere, su n lanci, esso puó presentarsi

$$0, 1, 2, \ldots, k, \ldots, n-1, n$$

volte e le rispettive probabilitá sono:

$$p_{n,0}, p_{n,1}, p_{n,2} \cdots, p_{n,k}, \cdots p_{n,n}$$

i cui valori sono calcolati tramite la formula:

$$p_{n,k} = \left(\begin{array}{c} n \\ k \end{array} \right) p^k q^{n-k}.$$

Si tratta di una variabile casuale perchè :

1) gli n eventi sono incompatibili ed esauriscono tutte le possibilitá ;

2) la somma delle probabilitá è uguale ad 1.

Infatti

$$p_{n,0} + p_{n,1} + \ldots + p_{n,k} = \sum_{k=0}^{n} p_{n,k} = \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) p^k q^{n-k} = (p+q)^n.$$

Ma $q = 1 - p$ pertanto $(p+q)^n = 1^n = 1$. Pertanto la somma delle probabilitá è uguale ad 1 ossia:

$$p_{n,0} + p_{n,1} + \ldots + p_{n,k} = (p+q)^n = 1$$

Per questa particolaritá tale distribuzione è detta **binomiale**.

Si puó dimostrare che:

$$M(X) = np$$

$$var(X) = npq.$$

Inoltre i valori di k (numero di volte in cui l'evento si presenta) aventi maggiore probabilitá sono compresi nell'intervallo:

$$[np - 3\sigma, np + 3\sigma]$$

ove $\sigma = \sqrt{var(X)} = \sqrt{npq}$.

Esempio 10.2 *Si lanci quattro volte un dado e si consideri l'evento E: uscita numero pari. Scrivere la variabile casuale, calcolarne il valor medio e la varianza e rappresentare graficamente la distribuzione.*

E' $p(E) = p = 1/2$ e quindi $p(\overline{E}) = q = 1/2$.

La variabile casuale è :

$$\begin{array}{llllll} X & 0 & 1 & 2 & 3 & 4 \\ p & 0,0625 & 0,25 & 0,375 & 0,25 & 0,0625 \end{array}$$

$$p_{4,0} = \left(\begin{array}{c} 4 \\ 0 \end{array} \right) \frac{1}{2}^0 \frac{1}{2}^4 = 0,0625, \textit{ecc.}$$

Il valor medio è $M(X) = 41/2 = 2$, la varianza è $var(x) = 41/21/2 = 1$.

11 Distribuzione di Poisson

Si variabile casuale con **distribuzione di Poisson** quando i valori assunti da questa sono tutti i numeri naturali:

$$0, 1, 2, \ldots, k\, ldots$$

e la distribuzione è espressa da

$$p_k = e^{-\delta} \frac{\delta^k}{k!}$$

dove δ è una costante reale positiva ed e è il numero di Nepero.

Caratteristica della distribuzione di Poisson è che $M(X) = var(x) = \delta$.

Si puó dimostrare che, al crescere di n, il valore della formula binomiale approssima sempre di piú il valore di p_k, calcolato con la formula di Poisson, cioè :

$$lim_n \left(\begin{array}{c} n \\ k \end{array} \right) p^k q^{n-k} = e^{-\delta} \frac{\delta^k}{k!}$$

con $\delta = np$.

L'approssimazione migliore si ha con p molto piccolo e n molto grande.

12 Distribuzione normale o di Gauss

Le formule relative alla distribuzione binomiale sono di facile applicazione fin quando i valori di n e k si mantengono piccoli. Per valori di n e k grandi, il calcolo sia del coefficiente binomiale, sia delle potenze di p e q risulta praticamente impossibile.

Esempio 12.1 *Lanciando 1000 volte una moneta la faccia croce si presenti 300 volte, senza tener conto dell'ordine di uscita.*

$$p_{1000,300} = \binom{1000}{300} \frac{1}{2}^{300} \frac{1}{2}^{700}$$

di difficile calcolo.

Si puó dimostrare che la seguente formula fornisce un valore sufficientemente approssimato di $p_{n,k}$:

$$p_{n,k}(x) = \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(x-np)^2}{2npq}}$$

In questa formula x rappresenta il generico valore che puó essere assunto dalla variabile casuale; esso nel caso di distribuzione binomiale con grandi valori di n e k è il numero k delle volte che si presenta l'evento.

La differenza fondamentale fra questa formula e quella della distribuzione binomiale consiste nel fatto che, per quest'ultima la variabile casuale è discreta, puó cioè assumere un numero finito anche se molto grande di valori; per la prima invece la variabile casuale puó assumere gli infiniti valori reali compresi tra $-\infty$ e $+\infty$ per tale motivo è detta continua; la rappresentazione grafica della distribuzione è una curva continua.

Ricordando che $\sigma = \sqrt{npq}$ e $m = np$ si ha che:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

La distribuzione definita dalla precedenta formula viene detta **normale** o **gaussiana**; graficamente abbiamo una cirva a campana detta di **Gauss.**

La curva normale standard.

Mediante la trasformazione

$$z = \frac{x-m}{\sigma} \text{ e } p(z) = \sigma p(x)$$

si ottiene

$$p(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}.$$

La distribuzione normale standard è un caso particolare della distribuzione normale in cui $m = 0$ e $\sigma = 1$.

13 Variabili casuali continue

Come giá detto, se la variabile casuale puó assumere in modo continuo tutti i valori reali compresi in un dato intervallo (finito o infinito) essa viene detta **continua.**

Se le variabili casuali sono continue non è possibile definire il valore della probabilitá in un punto dell'intervallo, cioè non è possibile definire una funzione di distribuzione come per le variabili casuali discrete.

Infatti se nel generico punto x, $p(x)$ fosse > 0 la somma delle probabilitá sarebbe infinita, mentre per definizione di variabile casuale essa deve valere 1. Allora per ogni x nell'intervallo è $p(x) = 0$, il che è altrettanto impossibile per lo stesso motivo.

Si fa allora ricorso alla **funzione di densitá di probabilitá** che rappresenta la probabilitá che il valore della variabile casuale continua cada in $[x, x + \Delta x]$.

Indicando come al solito con $F(x)$ la funzione di ripartizione della variabile casuale x si ha che $F(x) = p(X \leq x)$, quindi la probabilitá che la variabile aleatoria assuma un valore compreso nell'intervallo $[x, x + \Delta x]$ è :

$$p(x < X \leq x + \Delta x) = F(x + \Delta x) - F(x).$$

Se $F(x)$ è derivabile si ha che:

$$lim_{\Delta x \to 0} \frac{p(x < X \leq x + \Delta x)}{\Delta x} = lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x} = F'(x).$$

La derivata prima $F'(x)$ della funzione di ripartizione rappresenta la funzione densitá di probabilitá $f(x)$, che è l'equivalente della funzione di distribuzione per le varibili discrete. Quindi:

$$f(x) = F'(x)$$

da cui si ricava la relazione che lega la densitá di probabilitá e la funzione di ripartizione:

$$F(x) = \int_{-\infty}^{x} f(z)dz.$$