

Lezione del 06/02/2013

Istituzioni di Calcolo delle Probabilità

David Barbato

Variabili aleatorie discrete

Una variabile aleatoria X è discreta se “può assumere solo un’infinità al più numerabile di valori”. Cioè se esiste un insieme I al più numerabile tale che $P(X \in I) = 1$. Data una variabile aleatoria discreta X e l’insieme I dei valori che può assumere con probabilità positiva valgono le seguenti formule per il calcolo dei valori attesi.

$$\mathbb{E}[X] = \sum_{x \in I} xP(X = x) \quad \text{Se la serie è assolutamente convergente.}$$

$$\mathbb{E}[f(X)] = \sum_{x \in I} f(x)P(X = x) \quad \text{Se la serie è assolutamente convergente.}$$

Cosa succede se la serie non è assolutamente convergente?

Se la serie non è assolutamente convergente possono succedere tre cose:

- 1) La parte positiva diverge e la parte negativa converge allora si dirà che la media è più infinito.
- 2) La parte negativa diverge e la parte positiva converge allora si dirà che la media è meno infinito.
- 3) Sia la parte negativa che quella positiva divergono allora si dirà che non ammette media.

Esempio. Sia X una variabile aleatoria geometrica di parametro $p = \frac{1}{2}$ e sia $f : \mathbb{N} \rightarrow \mathbb{R}$ la funzione definita da $f(n) = 2^n$. Quanto vale $\mathbb{E}[f(X)]$?

$$\begin{aligned} \mathbb{E}[f(X)] &= \sum_{x \in I} f(x)P(X = x) = \sum_{n=1}^{\infty} f(n)P(X = n) \\ &= \sum_{n=1}^{\infty} 2^n \left(\frac{1}{2}\right)^n = \sum_{n=1}^{\infty} 1 = \infty \end{aligned}$$

In questo caso la parte negativa è nulla perché $f(X)$ è sempre non negativa e dunque si ha $\mathbb{E}[f(X)] = +\infty$

Ricordiamo le formule per il calcolo delle serie geometriche. Sia $\alpha \in (0, 1)$ allora

$$\sum_{k=1}^{\infty} \alpha^k = \frac{\alpha}{1 - \alpha} \quad \sum_{k=0}^{\infty} \alpha^k = \frac{1}{1 - \alpha}$$

Esercizio 1 Nelle ipotesi del precedente esempio sia X v.a. geometrica di parametro $\frac{1}{2}$ e sia $g(n) := \left(\frac{3}{2}\right)^n$ calcolare $\mathbb{E}[g(X)]$.

$$\begin{aligned} \mathbb{E}[g(X)] &= \sum_{x \in I} g(x)P(X = x) = \sum_{n=1}^{\infty} g(n)P(X = n) \\ &= \sum_{n=1}^{\infty} \left(\frac{3}{2}\right)^n \left(\frac{1}{2}\right)^n = \sum_{n=1}^{\infty} \left(\frac{3}{4}\right)^n = \frac{\frac{3}{4}}{1 - \frac{3}{4}} = 3 \end{aligned}$$

Assenze di memoria per le variabili aleatorie geometriche.

Siano n e M interi positivi sia X una variabile aleatoria geometrica allora vale

$$P(X = n + m | X > m) = P(X = n).$$

Dimostrazione della proprietà “assenza di memoria per variabili aleatorie geometriche”:

$$\begin{aligned} P(X = n + m | X > m) &= \frac{P(X = n + m \cap X > m)}{P(X > m)} = \frac{P(X = n + m)}{P(X > m)} \\ &= \frac{p(1-p)^{n+m-1}}{P(X = m + 1 \cup X = m + 2 \cup \dots)} \\ &= \frac{p(1-p)^{n+m-1}}{\sum_{k=1}^{\infty} P(X = m + k)} = \frac{p(1-p)^{n+m-1}}{\sum_{k=1}^{\infty} p(1-p)^{k+m-1}} \\ &= \frac{p(1-p)^{n+m-1}}{\sum_{k=1}^{\infty} p(1-p)^k (1-p)^{m-1}} \\ &= \frac{p(1-p)^{n+m-1}}{p(1-p)^{m-1} \sum_{k=1}^{\infty} (1-p)^k} \\ &= \frac{(1-p)^n}{\frac{(1-p)}{1-(1-p)}} = p(1-p)^{n-1} = P(X = n) \end{aligned}$$

Variabili aleatorie binomiali negative.

Consideriamo una sequenza di lanci di un dado regolare a sei facce. Sia X_1 la prima volta che esce il “5”, sia X_2 la seconda volta che esce il “5” e sia in generale X_r l’ r -esima volta che esce il “5”. Quali sono le distribuzioni di X_r ? X_1 indica il primo istante di successo di un evento (esce il “5”) di probabilità $p = \frac{1}{6}$ dunque X_1 si distribuisce come una variabile aleatoria geometrica di parametro $p = \frac{1}{6}$. Cosa succede in generale per X_r ? Consideriamo la probabilità $P(X_r = n)$ con $n \geq r$. $X_r = n$ vuol dire che nei primi $n - 1$ lanci ci sono stati esattamente $r - 1$ successi e che l’ n -esimo lancio è un successo. La probabilità che nei primi $n - 1$ lanci ci sono stati esattamente $r - 1$ successi è data da $\binom{n-1}{r-1} p^{r-1} (1-p)^{n-r}$ mentre la probabilità che l’ n -esimo sia un successo è p dunque:

$$P(X_r = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r} \quad (1)$$

Una variabile aleatoria con la distribuzione data da (1) è detta variabile binomiale negativa di parametri r e p . La media e la varianza delle v.a. binomiali negative sono date da:

$$\mathbb{E}[X_r] = \frac{r}{p} \quad \text{VAR}[X_r] = \frac{r(1-p)}{p^2}$$

Osservazione 1. Se $r = 1$ allora la variabile aleatoria binomiale negativa di parametri r e p altro non è che una v.a. geometrica di parametro p .

Osservazione 2. Valgono le uguaglianze

$$\mathbb{E}[X_r] = r \cdot \mathbb{E}[X_1] \text{ e } \text{VAR}[X_r] = r \cdot \text{VAR}[X_1].$$

Variabili aleatorie ipergeometriche.

Un'urna contiene N palline di cui m sono bianche e le restanti $N - m$ sono nere. Ne estraiamo n con $n \in \{0, 1, \dots, N\}$. Sia X il numero di palline bianche estratte. Qual è la distribuzione di X ? Consideriamo d'ora in poi $0! = 1$ e $\binom{r}{k}$ uguale a zero se $k < 0$ oppure $r < k$. La distribuzione di X è data dalla seguente formula.

$$P(X = i) = \frac{\binom{n}{i} \binom{N-n}{m-i}}{\binom{N}{m}} = \frac{\binom{m}{i} \binom{N-m}{n-i}}{\binom{N}{n}} \quad (2)$$

Esercizio 2 Verificare algebricamente l'ultima uguaglianza della (2).

La media e la varianza di una variabile aleatoria ipergeometrica X di parametri N , m e n sono date da

$$\mathbb{E}[X] = \frac{nm}{N} \quad \text{VAR}[X] = \frac{nm(N-n)(N-m)}{N^2(N-1)}$$

Siano N , n e m costanti positive con $N > n, m$. Consideriamo una prima urna con N palline di cui m bianche ne estraiamo n e indichiamo con X il numero di palline bianche estratte. Consideriamo una seconda urna sempre con N palline ma con n palline bianche ne estraiamo m e sia Y il numero di palline estratte bianche. Si verifica facilmente utilizzando la (2) che X e Y hanno la stessa distribuzione, come si giustifica questo fatto?

Esercizio 3

Giovanni scommette sulla ruota di Venezia, vincerà 3 euro per ogni numero estratto superiore ad 80. Sia X il numero di euro vinti da Giovanni.

- (a) Qual è la distribuzione di X ?
- (b) Qual è la vincita media di Giovanni?
- (c) Qual è la vincita più probabile?

Soluzione:

(a) $I = \{0, 3, 6, 9, 12, 15\}$, $P(X = k) = \frac{\binom{5}{\frac{k}{3}} \binom{85}{10 - \frac{k}{3}}}{\binom{90}{10}}$ per ogni $k \in I$

(b) $\mathbb{E}[X] = \frac{5}{3}$

(c) 0