

Esercitazione del 28/10/2011

Calcolo delle probabilità

Distribuzione di una funzione di una variabile aleatoria discreta.

Sia X una variabile aleatoria discreta, sia f una funzione da $\mathbb{R}$ in $\mathbb{R}$, se $Y := f(X)$ allora Y è una variabile aleatoria discreta e la sua densità è data da:

$$P(Y = y) = \sum_{x|f(x)=y} P(X = x) \quad \forall y \quad (1)$$

Esercizio 1

Sia X una variabile aleatoria uniformemente distribuita sull'insieme $\{-2, -1, 0, +1, +2\}$ (cioè $P(X = -2) = P(X = -1) = P(X = 0) = P(X = 1) = P(X = 2) = \frac{1}{5}$) e sia f la funzione definita da $f(x) = x^2$ per ogni x in $\mathbb{R}$. Calcolare la distribuzione di $Y := f(X)$.

Soluzione: $P(Y = 0) = \frac{1}{5}, P(Y = 1) = P(Y = 4) = \frac{2}{5}$

Talvolta la funzione f è definita solo su un sottinsieme di $\mathbb{R}$, se tale supporto include il supporto della variabile aleatoria X allora ha comunque senso definire $Y := f(X)$ (a patto di escludere al più un insieme di misura nulla) ed è possibile usare l'uguaglianza (1) per calcolare la distribuzione di X .

Esercizio 2

Sia X una variabile aleatoria geometrica di parametro $\frac{1}{2}$ e sia f la funzione definita da $f(n) = (-1)^n$ per ogni n in $\mathbb{N}$. Calcolare la distribuzione di $Y := f(X)$.

Soluzione: $P(Y = -1) = \frac{2}{3}, P(Y = 1) = \frac{1}{3}$

Vettori aleatori discreti.

Sia (X, Y) un vettore aleatorio discreto, con densità f allora X e Y sono v.a. discrete e la loro distribuzione è data da:

$$P(Y = y) = \sum_{x|f(x,y)>0} f(x, y) \quad \text{e} \quad P(X = x) = \sum_{y|f(x,y)>0} f(x, y) \quad (2)$$

Esercizio 3

Sia (X, Y) un vettore aleatorio con densità f data da:

$$f(m, n) = \left(\frac{2}{2 + 3^m} \right)^n \quad \forall n, m \text{ interi e positivi}$$

(cioè $P(X = m, Y = n) = \left(\frac{2}{2 + 3^m} \right)^n$).

(a) Calcolare la distribuzione di X .

(b) A quale distribuzione appartiene X ? (Indicare gli eventuali parametri.)

Soluzione: (a) $P(X = m) = \frac{2}{3^m}$ per ogni m intero e positivo. (b) La v.a. X è una v.a. geometrica di parametro $\frac{2}{3}$.

Variabili aleatorie assolutamente continue.

Una variabile aleatoria X è assolutamente continua se esiste una funzione f_X da $\mathbb{R}$ in $\mathbb{R}$ non negativa tale che per ogni insieme A misurabile¹ si abbia:

$$P(X \in A) = \int_A f_X(x) dx \quad (3)$$

la funzione f_X è detta la densità della v.a. X . Dalla (3) risulta $P(a \leq X \leq b) = P(a < X < b) = \int_a^b f_X(x) dx$ e anche $P(X = a) = 0$, per ogni a e b con $a < b$. Inoltre si ha:

$$F_X(a) = \int_{-\infty}^a f_X(x) dx \quad \forall a \quad \text{e} \quad f_X(a) = \frac{d}{da} F_X(a) \quad \text{per quasi ogni}^2 a$$

Valore atteso di variabili aleatorie assolutamente continue.

Sia X una variabile aleatoria assolutamente continua e sia f_X la sua densità.

$$\mathbb{E}[X] = \int_{\mathbb{R}} x f_X(x) dx \quad \text{se} \quad \int_{\mathbb{R}} |x| f_X(x) dx < \infty$$

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}} g(x) x f_X(x) dx \quad \text{se} \quad \int_{\mathbb{R}} |g(x)| f_X(x) dx < \infty$$

Cosa succede se la condizione $\int_{\mathbb{R}} |x| f_X(x) dx < \infty$ oppure la condizione $\int_{\mathbb{R}} |g(x)| f_X(x) dx < \infty$ non è verificata? Vediamo il seguente esercizio.

Esercizio 4 Sia X una v.a. uniforme sull'intervallo $(0, 2)$, sia g la funzione definita da

$$g(x) = \begin{cases} -\frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Calcolare $\mathbb{E}[g(X)]$.

Soluzione: $\mathbb{E}[g(X)] = -\infty$

Variabili aleatorie normali.

Proposizione 1. Se X è una v.a. normale, con distribuzione $X \sim N(\mu, \sigma^2)$ allora anche $Y := aX + b$ è normale e vale $Y \sim N(a\mu + b, a^2\sigma^2)$.

Proposizione 2. Se $X_1, X_2, \dots, X_n$ sono v.a. normali indipendenti con distribuzione $X_i \sim N(\mu_i, \sigma_i^2)$ allora anche $Y := a_1X_1 + a_2X_2 + \dots, a_nX_n$ sarà

¹Sono insiemi misurabili tutti gli intervalli, tutti gli insiemi finiti o numerabili e tutti gli insiemi che sono differenze, unioni o intersezioni finite o numerabili di insiemi misurabili.

²L'uguaglianza è vera a meno di un insieme di misura di Lebesgue nulla, in particolare se la funzione F è C^1 a tratti allora l'uguaglianza sarà vera ovunque all'interno di tutti i tratti della F_X in cui è C^1

normale e vale $Y \sim N(a_1\mu_1 + a_2\mu_2 + \dots, a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \dots + a_n^2\sigma_n^2)$.

Una variabile aleatoria Z normale con distribuzione $Z \sim N(0, 1)$ è detta *normale standard*, la sua funzione di ripartizione è indicata generalmente con la lettera Φ e si ha:

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$$

Vale $\Phi(-x) = 1 - \Phi(x)$ e se $X \sim N(\mu, \sigma^2)$ allora si ha

$$P(X \leq x) = \Phi\left(\frac{x - \mu}{\sigma}\right)$$

Approssimazione Normale. Le ragioni del *perché* molte distribuzioni siano approssimativamente normali verranno illustrate quando verrà studiato il teorema del limite centrale. In questa lezione vedremo il *come* si applica il procedimento di approssimazione Normale (**Gaussiana**). L'ambito di applicazione maggiore è dato dall'approssimazione di v.a. $X \sim \text{Bin}(n, p)$ binomiali con il parametro n grande, nel caso della binomiale l'approssimazione sarà tanto migliore quanto maggiore è la varianza di X , $\text{VAR}(X) = np(1 - p)$ nella maggior parte delle applicazioni una varianza maggiore di 10 sarà considerata sufficiente per poter procedere con l'approssimazione normale.

Procedimento di Approssimazione Normale. Sia X la variabile che vogliamo approssimare (**non necessariamente binomiale**). Siano $\mu = \mathbb{E}[X]$ e $\sigma^2 = \text{VAR}[X]$. Allora la distribuzione di X verrà approssimata da quella di una v.a. Y normale con la stessa media e la stessa varianza di X : $Y \sim N(\mu, \sigma^2)$.

Correzione di continuità. La correzione di continuità permette di avere stime più precise e si applica quando la variabile da approssimare X è una v.a. binomiale (vedere esempio 4f del libro di testo pag 219). La correzione di continuità è cruciale per avere una buona approssimazione quando la varianza non è molto grande. Illustriamo il procedimento con un esempio. Sia $X \sim \text{Bin}(n, p)$, $\mu = np$ e $\sigma^2 = np(1 - p)$. Le seguenti probabilità sono uguali

$$P(X \leq 60) = P(X \leq 60.5) = P(X < 61) \quad (4)$$

mentre per $Y \sim N(\mu, \sigma^2)$ si ha

$$P(Y \leq 60) \neq P(Y \leq 60.5) \neq P(Y < 61)$$

La approssimazione migliore per la probabilità (4) è data da:

$$P(X \leq 60) = P(X \leq 60.5) = P(X < 61) \simeq P(Y \leq 60.5) = \Phi\left(\frac{60.5 - \mu}{\sigma}\right)$$

Variabili aleatorie esponenziali.

Minimo di v.a. esponenziali indipendenti.

Ricordiamo innanzitutto che due variabili aleatorie X_1 e X_2 sono indipendenti se e solo se comunque scelti A e B sottoinsiemi misurabili di $\mathbb{R}$ si ha $P(X \in A, Y \in B) = P(X \in A) \cdot P(Y \in B)$.

Siano X e Y due variabili esponenziali $X \sim \text{Esp}(\lambda_1)$ e $Y \sim \text{Esp}(\lambda_2)$.

Proposizione Se $X \sim \text{Esp}(\lambda_1)$, $Y \sim \text{Esp}(\lambda_2)$, X e Y sono indipendenti e $Z := \min\{X, Y\}$ allora anche Z è una variabile aleatoria esponenziale e vale $Z \sim \text{Esp}(\lambda_1 + \lambda_2)$.

Esercizio. Dimostrare la proposizione precedente.

La distribuzione Beta.

Una variabile aleatoria X si dice avere una distribuzione Beta se è assolutamente continua e ha densità f_X data da:

$$f_X(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} & 0 < x < 1 \\ 0 & \text{altrimenti} \end{cases}$$

dove i parametri della distribuzione a e b sono entrambi maggiori di zero. La costante di normalizzazione $B(a, b)$ è data dal seguente integrale.

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx.$$

La media è data da:

$$\mathbb{E}[X] = \frac{a}{a+b}.$$

Esercizi Vari

Esercizio

Calcolo delle Probabilità 18/11/2010

Sia X una variabile aleatoria normale con distribuzione e $X \sim N(0, 9)$. Sia $Y = \max\{X, 0\}$.

(a) Dimostrare che Y ha funzione di ripartizione:

$$F_Y(t) = \begin{cases} 0 & t < 0 \\ \Phi\left(\frac{t}{3}\right) & t \geq 0 \end{cases}$$

(b) Calcolare la derivata: $\frac{d}{dt} F_Y(t)$ per $t > 0$.

(c) Calcolare $\mathbb{E}[Y]$.

Esercizio

Calcolo delle Probabilità 03/11/2009

Un contadino si affida alla previsioni metereologiche secondo le quali vi è una probabilità del 90% che la prossima settimana piova. Lui sa che se concimerà il suo campo, allora ci saranno un 50% di piante che seccheranno in caso che non piova mentre tale probabilità scende al 10% in caso di pioggia.

Se invece decide di non concimare il suo campo ci saranno un 30% di piante che seccheranno nel caso che non piova e un 20% in caso di pioggia.

(a) Se decide di concimare il suo terreno, qual è la percentuale media di piantine che sopravviveranno?

(b) Cosa gli conviene fare se vuole massimizzare il numero medio di piantine che non seccheranno?