Esercitazione del 21/10/2011
Calcolo delle probabilità

Funzione di ripartizione
Sia F_X una funzione da $\mathbb{R}$ in $\mathbb{R}$. consideriamo le seguenti condizioni:

$$(*) \begin{cases} F_X \quad \text{è non decrescente} \\ \lim_{x\to\infty} F_X(x) = 1 \\ \lim_{x\to-\infty} F_X(x) = 0 \\ F_X \quad \text{è continua a destra} \end{cases}$$

La funzione F_X è una funzione di ripartizione di una variabile aleatoria se e solo se soddisfa le condizioni $(*)$. Se F_X è una funzione definita a tratti tale che all'interno di ciascun intervello di definizione è non decrescente e continua a destra allora affinché sia non decrescente e continua a destra su tutto $\mathbb{R}$ sarà sufficiente verificare che per ogni x_0 estremo degli intervalli di definizione valga la relazione:

$$\lim_{x\to x_0^-} F_X(x) \leq F_X(x_0) = \lim_{x\to x_0^+} F_X(x)\,.$$

Esercizio 1

Per quali valori di α la funzione F_X è la funzione di ripartizione di una variabile aleatoria.

$$F_X(x) = \begin{cases} 0 & x < -1 \\ \frac{1}{2} + \alpha x & -1 \leq x < 1 \\ 1 & x \geq 1 \end{cases}$$

Soluzione: $\alpha \in [0, \frac{1}{2}]$

Esercizio 2

Per quali valori di α la funzione F_X è la funzione di ripartizione di una variabile aleatoria.

$$F_X(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{\pi}(x + \alpha \sin(x)) & -0 \leq x < \pi \\ 1 & x \geq \pi \end{cases}$$

`Sugg: Poiché la funzione` F_X `è derivabile in` $(0,\pi)$ `allora si può studiare la monotonia in` $(0,\pi)$ `attraverso lo studio del segno della sua derivata prima.`
Soluzione: $\alpha \in [-1, +1]$

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 vaolri attesi.

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

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

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 $f : \mathbb{N} \to \mathbb{R}$ 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

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

Eserizio 3 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":

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

Variabili aleatorie Poissoniane.
Una variabile aleatoria discreta X è poissoniana di parametro λ se $P(X=n)=e^{-\lambda}\frac{\lambda^n}{n!}$ per ogni n intero non negativo. Ricordiamo la seguente uguaglianza:

$$e^t=\sum_{n=0}^{\infty}\frac{t^n}{n!} \qquad \forall\, t\in\mathbb{R}$$

Esercizio 4 Verificare che $\sum_{n=0}^{\infty}P(X=n)=1$ e $\mathbb{E}[X]=\lambda$.

$$\sum_{n=0}^{\infty}P(X=n)=\sum_{n=0}^{\infty}e^{-\lambda}\frac{\lambda^n}{n!}=e^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^n}{n!}==e^{-\lambda}e^{\lambda}=1$$

$$\mathbb{E}[X]=\sum_{n=0}^{\infty}nP(X=n)=\sum_{n=0}^{\infty}ne^{-\lambda}\frac{\lambda^n}{n!}=e^{-\lambda}\sum_{n=1}^{\infty}n\frac{\lambda^n}{n!}=e^{-\lambda}\sum_{n=1}^{\infty}\frac{\lambda^n}{(n-1)!}$$

ponendo $m=n-1$ si ha:

$$\mathbb{E}[X]=e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^{m+1}}{m!}=e^{-\lambda}\lambda e^{\lambda}=\lambda\,.$$

Sia ora λ una costante positiva fissata. Siano X_n variabili aleatorie binomiali di parametri $(n,\frac{\lambda}{n})$ quindi tali che $\mathbb{E}[X_n]=\lambda$ per ogni n. Ora mostreremo che per ogni $k\in\mathbb{N}$ si ha $\lim_{n\to\infty}P(X_n=k)=e^{-\lambda}\frac{\lambda^k}{k!}$. Cioè mostreremo che la densità discreta delle X_n tende a quella di una variabile aleatoria poissoniana di media λ. (Nella seconda parte del corso vedremo che si tratta di una convergenza in distribuzione.)

$$\begin{aligned}
\lim_{n\to\infty} P(X_n = k) &= \lim_{n\to\infty} \frac{n!}{(n-k)!k!}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k!} \lim_{n\to\infty} \frac{n}{n}\frac{n-1}{n}\cdots\frac{n-k+1}{n}\left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k} \\
&= \frac{\lambda^k e^{-\lambda}}{k!}
\end{aligned}$$

dove nell'ultima uguaglianza abbiamo usato il limite notevole
$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}$.

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:

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

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

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

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

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.

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

Esercizio 5 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

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

Siano N, n e m costanti positive con $N > n, m$. Consideriamo una prima urna con N pallina 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 6

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) $I = \{0, 3, 6, 9, 12, 15\}$, $\quad P(X = k) = \frac{\binom{5}{\frac{k}{3}}\binom{85}{10-\frac{k}{3}}}{\binom{90}{10}}$ per ogni $k \in I$
(b) $\mathbb{E}[X] = \frac{5}{3}$
(c) 0