Complementi di Idrologia
Appello del 22 Febbraio 2012

Problema n°1 (8 punti di cui 4 per il punto 3)

1. Si consideri la distribuzione geometrica. Motivare, appoggiandosi ad una serie di esperimenti di Bernoulli, perché la sua equazione è $P[N=n]=p_N(n)=(1-p)^{n-1}p$, dove N rappresenta il numero di esperimenti per arrivare al primo successo e p la probabilità di successo nel generico esperimento.
2. Scrivere l'espressione di $E[N]$ e $\mathrm{Var}[N]$.
3. In una stazione pluviometrica in India durante il periodo dei monsoni si osservano i seguenti giorni consecutivi di pioggia:

giorni di pioggia:	1	2	3	4	5	6	7	8
frequenza assoluta:	161	52	32	17	8	6	4	1

   Calcolare il numero di giorni di pioggia consecutivi la cui probabilità è pari a 0,05.

Risposte

1. Sia p la probabilità di successo nel generico esperimento in cui posso osservare 1 (successo) o 0 (insucesso). Il primo successo arriva all'ennesimo esperimento quando è preceduto da n-1 insuccessi. La probabilità di osservare n-1 insuccessi e un successo subito dopo è pari a $(1-p)^{n-1}p$. Pertanto la probabilità di avere il primo successo all'ennesimo esperimento è $P[N=n]=p_N(n)=(1-p)^{n-1}p$.
2. $E[N]=\dfrac{1}{p};\ Var[N]=\dfrac{1-p}{p^2}$.
3. Si indichi come successo il giorno di "non pioggia" e insuccesso il giorno di "pioggia". Per stimare la probabilità di successo al generico esperimento (ovvero di osservare non pioggia nel generico giorno) occorre prima stimare il numero medio di esperimenti che occorre fare per arrivare al primo successo ovvero il numero medio di giorni che intercorre fra un periodo di non pioggia e il successivo periodo di non pioggia.

   La stima del valore atteso di N, ovvero $\hat{E}[N]$ è data da (media pesata):

$$\hat{E}[N]=\frac{161\cdot1+52\cdot2+32\cdot3+17\cdot4+8\cdot5+6\cdot6+4\cdot7+1\cdot8}{161+52+32+17+8+6+4+1}=\frac{541}{281}=1{,}92$$

   Pertanto la stima della probabilità di successo nel generico esperimento (ovvero di osservare non pioggia nel generico giorno) è:

$$\hat{p}=\frac{1}{\hat{E}[N]}=\frac{1}{1{,}92}=0{,}52$$

   Pertanto:

$$P[N=n]=p_N(n)=(1-p)^{n-1}p=0{,}05$$

$$n=1+\frac{\ln 0{,}05-\ln p}{\ln(1-p)}=1+\frac{\ln 0{,}05-\ln 0{,}52}{\ln(1-0{,}52)}=4{,}19\cong 5$$

   Se dunque è cinque il numero di esperimenti per arrivare con probabilità 0,05 al primo successo (giorno di non pioggia) vuol dire che **quattro** sono i giorni di pioggia consecutivi aventi la probabilità assegnata.

Problema n°2 (8 punti di cui 4 per il punto 3)

1. Si enunci il teorema del limite centrale e si discuta.
2. Si derivi la distribuzione esponenziale a partire dalla distribuzione di Poisson.
3. In un corso d'acqua si osserva la presenza di una sostanza tossica. Vengono fatti diversi test chimici da cui si deduce che la concentrazione media di questa sostanza è di 2 mg/L mentre la sua varianza è pari a 4 $(mg/L)^2$. Quale è la probabilità che la concentrazione di questa sostanza sia compresa fra 1 e 3 mg/L assumendo che la distribuzione sia (a) esponenziale, (b) normale?

Risposte

1. "Sotto **condizioni molto generali**, man mano che il numero delle variabili presenti nella sommatoria diventa **grande**, la distribuzione della variabile somma **si avvicina** sempre più alla distribuzione normale".
   Commento. Il teorema vale in "condizioni molto generali" ovvero se (1) le variabili considerate nella sommatoria sono indipendenti e ugualmente distribuite, (2) se le variabili sono indipendenti e diversamente distribuite. Il termine "si avvicina" non va inteso nel senso matematico di limite e più opportuno leggerlo come "viene approssimata da". Infine il termine "grande" dipende dalla bontà dell'approssimazione richiesta e dalla natura delle distribuzioni delle variabili presenti nella somma.
2. Indichiamo con T il tempo intercorrente fra due eventi successivi. La probabilità che T ecceda un assegnato valore t equivale alla probabilità che nessuno evento si manifesti in tale intervallo. Quindi:
   $$1-F_T(t)=\frac{(\lambda t)^0 e^{(-\lambda t)}}{0!}=e^{(-\lambda t)} \qquad t\geq 0$$
   $$F_T(t)=1-\exp(-\lambda t);$$
   $$f_T(t)=\frac{dF_T(t)}{dt}=\lambda\exp(-\lambda t);$$
3. Sia X la concentrazione della sostanza tossica nell'acqua del fiume.
   (a) Distribuzione esponenziale
   $$\lambda=\frac{1}{m_X}=\frac{1}{2}=0,5;$$
   $$P[1\leq X\leq 3]=\left[1-\exp(-0,5\cdot 3)\right]-\left[1-\exp(-0,5\cdot 1)\right]=0,38$$
   (b) Distribuzione normale
   $$P[1\leq X\leq 3]=F_Z\left(\frac{3-2}{2}\right)-F_Z\left(\frac{1-2}{2}\right)=F_Z(0,5)-F_Z(-0,5)=0,6915-0,3085=0,38$$

4.

Domanda n°1 (8 punti)

1. Sia $Y = g(X)$ una funzione continua monotona decrescente. Dimostrare che:
   a. $F_Y(y)=1-F_X\left(g^{-1}(y)\right)$
   b. $f_Y(y)=\left|\frac{dg^{-1}(y)}{dy}\right| f_X\left(g^{-1}(y)\right)$
2. Si consideri la distribuzione *EV2*. Scrivere l'equazione della probabilità di non superamento, i vincoli sui parametri e il suo dominio di definizione.
3. Dimostrare che la variabile *EV2* nel piano di Gumbel ha una concavità rivolta verso l'alto.

Risposte

1. Sia g (X) una funzione continua, monotona decrescente. Allora:

$$P[Y \le y] = F_Y(y) = P\left[X \ge g^{-1}(y)\right] = 1 - F_X\left(g^{-1}(y)\right)$$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{d\left[1 - F_X\left(g^{-1}(y)\right)\right]}{dy} = -f_X\left(g^{-1}(y)\right)\frac{dg^{-1}(y)}{dy} = \left|\frac{dg^{-1}(y)}{dy}\right| f_X\left(g^{-1}(y)\right)$$

2. $F_{Z_2}(z_2) = \exp\left[-\left(1 - k\frac{(z_2 - u)}{\alpha}\right)^{\frac{1}{k}}\right]$; k<0; α>0; $u + \frac{\alpha}{k} \le z_2 \le \infty$

3. $Y_2 = 1 - k\frac{(Z_2 - u)}{\alpha} = g(Z_2)$

$$Z_2 = u + \frac{\alpha}{k} - \frac{\alpha}{k}Y_2 = g^{-1}(Y_2)$$

$$F_{Y_2}(y_2) = F_{Z_2}\left(g^{-1}(y_2)\right) = \exp\left[-y_2^{\frac{1}{k}}\right]$$

$$Y_1 = -\frac{1}{k}\ln Y_2 = g(Y_2)$$

$$Y_2 = e^{-kY_1} = g^{-1}(Y_1)$$

$F_{Y_1}(y_1) = F_{Y_2}\left(g^{-1}(y_1)\right) = e^{-e^{-y_1}}$; dunque Y_1 è effettivamente la variabile ridotta di Gumbel.

Il piano di Gumbel ha per ascissa Y_1 e per ordinata Z_2. Poiché k<0 l'equazione $Z_2 = u + \frac{\alpha}{k} - \frac{\alpha}{k}Y_2 = u + \frac{\alpha}{k} - \frac{\alpha}{k}e^{-kY_1}$ indica che la curva $Z_2=f(Y_1)$ ha una concavità rivolta verso l'alto e andamento esponenziale.

Miscellanea n°1 (8 punti)

1. La pioggia annuale su di un bacino è solitamente distribuita seconda una normale. Su di un particolare bacino, la pioggia annuale X ha media 1000 mm e deviazione standard 200 mm. Il deflusso annuale (in mm) Y è legato alla pioggia annuale sulla base della seguente relazione: Y=100+0,4X.
   a. Specificare la distribuzione di Y e valutare i suoi parametri.
   b. Quale è la probabilità che Y sia minore di 350 mm?
2. Scrivere l'equazione del modello Binomiale spiegando il significato di tutti i simboli.
3. Sia P la probabilità di avere successo in *almeno* un esperimento su n. Quale è la probabilità p di avere successo nel singolo esperimento?
4. Appoggiandosi alla distribuzione geometrica, dimostrare che $1/\left(1 - F_X(x_0)\right)$ è il numero medio di esperimenti che intercorre fra un esperimento in cui $X>x_0$ (successo) e il successivo posto che $F_X(x_0)$ $=P[X \le x_0]$.

Risposte

1. $E[X] = 1000$ mm; $Var[X] = 200^2$ mm^2; Y=100+0,4X.

$E[Y]$=100+0,4$E[X]$=500 mm
Var[Y]=0,4^{2}Var[X]=0,4^2 200^2 mm^2
$Y \sim N(500, 80^2)$
P[Y≤350] = $F_Z\left(\frac{350 - 500}{80}\right) = F_Z(-1{,}875) = 1 - F_Z(1{,}875) = 1 - 0{,}9696 = 0{,}0304$

2. $p_X(x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$. X è la variabile Binomiale e rappresenta il numero di successi. $p_X(x)$ rappresenta la probabilità di avere x successi su n esperimenti indipendenti in ciascuno dei quali la probabilità di successo è p.
3. Poiché $P = 1-(1-p)^n$ ne segue che la probabilità di avere successo nel singolo esperimento è $p = 1-(1-P)^{1/n}$.
4. La distribuzione geometrica descrive in senso probabilistico il numero di esperimenti per arrivare al primo successo ovvero, (essendo gli esperimenti indipendenti gli uni dagli altri) il numero di esperimenti che intercorre fra un successo e l'altro. Posto che p sia la probabilità di avere successo nel singolo esperimento, allora il numero medio di esperimenti che intercorre fra un successo e l'altro è $E[N] = m_N = 1/p$. Se per successo si intende che l'output del proprio esperimento sia maggiore di un assegnato valore x_0 allora p=1- $F_X(x_0)$ e quindi $E[N] = m_N = 1/p = 1/(1-F_X(x_0))$.