Complementi di Idrologia
Appello del 13 Gennaio 2012

Problema n°1 (8 punti)

1. Si considerino k variabili gaussiane X_i (i=1,k) indipendenti tutte distribuite con media m_X e varianza σ^2_X. Sia Y la variabile somma di queste k variabili. Quale è la sua media e la sua varianza?
2. Se A e B sono due eventi indipendenti, quale è la probabilità di osservarli simultaneamente?
3. Supponiamo che X abbia media 100 e varianza 400. Calcolare la probabilità $P[X\leq 50]$, $P[X\geq 150]$ e $P[50\leq X\leq 150]$.
4. Le piogge mensili registrate da un pluviometro sono indipendente e ugualmente distribuite secondo una normale con media 20 cm e varianza 120 cm^2. Determinare la probabilità che il totale di pioggia su sei mesi consecutivi sia maggiore di 200 cm. Determinare la probabilità che in ciascun mese, per un periodo di sei mesi consecutivi, la precipitazione sia minore di 18 cm.

Risposte

1. $Y=\sum_{i=1}^{k} X_i$; $m_Y = k\cdot m_X$; $\sigma_Y^2 = k\sigma_X^2$
2. $P[A]\times P[B]$
3. $P\left[X\leq 50\right]=P\left[\frac{50-100}{\sqrt{400}}\right]=P\left[-2,5\right]=1-P\left[2,5\right]=0,5-0,4983=0,0062$

   $P\left[X\geq 50\right]=1-P\left[\frac{150-100}{\sqrt{400}}\right]=1-P\left[2,5\right]=0,0062$

   $P\left[50\leq X\leq 150\right]=1-P\left[X\leq 50\right]-P\left[X\geq 150\right]=1-2\cdot 0,0062=0,9876$
4. Sia X la variabile casuale che rappresenta il totale di pioggia gaussiano nel generico mese. Il totale di pioggia su 6 mesi è $Y=\sum_{i=1}^{6} X_i$ con media pari a $m_Y = 6\cdot m_X = 6\cdot 20 = 120$ cm e varianza $\sigma_Y^2 = 6\cdot\sigma_X^2 = 6\cdot 120 = 720$ cm^2. Pertanto la probabilità che il totale di pioggia su 6 mesi sia maggiore di 200 cm è : $P\left[Y\geq 200\right]=1-P\left[\frac{200-120}{\sqrt{720}}\right]=1-P\left[2,98\right]=0,5-04986=0,0014$.

   La probabilità che in ciascun mese, per sei mesi consecutivi, la pioggia sia inferiore a 18 cm è:

   $\left\{P\left[X\leq 18\right]\right\}^6=\left\{P\left[\frac{18-20}{\sqrt{120}}\right]\right\}^6=\left\{P\left[-0,18\right]\right\}^6=\left\{1-P\left[0,18\right]\right\}^6=0,0062$

Problema n°2 (8 punti)

1. Si consideri una macchina che produce un nastro. La probabilità di incepparsi di questa macchina è pari a 0,0002 ogni 100 metri di nastro prodotto. Quale è la probabilità che la macchina si inceppi subito dopo aver prodotto 1000 m di nastro?
2. Se la macchina impiega una settimana a produrre 1000 m di nastro, quale è la probabilità che si inceppi alla 14 settimana)
3. Mediamente ogni quante settimane si interrompe la macchina?

Risposte

1. Si utilizza la distribuzione Geometrica $p_N(n) = (1-p)^{n-1} p$ dove p rappresenta la probabilità di successo (ovvero che la macchina si inceppi) e n il numero di esperimenti che portano al primo successo. Gli esperimenti sono cadenzati ogni 100 metri di nastro prodotto. Pertanto la probabilità che la macchina si inceppi dopo aver prodotto 1000 metri di nastro equivale ad osservare il primo successo (inceppamento) all'11mo esperimento e quindi tale probabilità è: $0{,}0002 \cdot (1-0{,}0002)^{10} = 0{,}0002$.
2. Procedendo con una analogo ragionamento fatto questo volta sulle settimane si può scrivere che la probabilità è pari a $0{,}0002 \cdot (1-0{,}0002)^{13} \cong 0{,}0002$.
3. Se la probabilità di interrompersi dopo una settimana è pari a 0,0002, vuol dire che il numero medio di settimane che intercorre fra una interruzione e l'altra è 1/p=1/0,0002=5000 settimane.

Domanda n°1 (8 punti)

1. Sotto quali ipotesi la distribuzione limite della variabile X massimo di n variabili Z_i tende alla *EV1* per $n \to \infty$?
2. Scrivere la funzione di probabilità cumulata della *EV1*, l'espressione che lega i suoi parametri alla media e alla varianza e dimostrare che è possibile scrivere:
   $X(T) = m_X \left\{1 - \left[0,45 + 0,78 \ln\left(-\ln\left(1 - \frac{1}{T}\right)\right)\right] CV\right\}$; T = tempo di ritorno
3. Dimostrare che il modello *EV1* nel piano di Gumbel è una retta.
4. In che modo si possono rappresentare i valori di un campione di dati massimi annui nel piano di Gumbel?

Risposte

1. Le variabili Z_i sono:
   (a) indipendenti fra di loro
   (b) ugualmente distribuite
   (c) la coda della distribuzione di probabilità cumulata della generica Z_i può essere approssimata dalla seguente legge: $F_Z(z) = 1 - e^{-g(z)}$ con $g(z)$ monotona crescente
   (d) n tende all'infinito

2. $F_X(x) = \exp\left[-\exp\left(-\frac{x-u}{\alpha}\right)\right], \quad -\infty \le x \le +\infty; \;\; \alpha > 0$

   $$\begin{cases} m_X = u + 0,5772\alpha \\ \sigma_X^2 = \dfrac{\pi^2 \alpha^2}{6} \end{cases}$$

   Derivando da queste ultime due equazioni l'espressione di u e α e sostituendole nell'equazione della distribuzione cumulata di probabilità e applicando due volte il logaritmo naturale a sinistra e a destra dell'uguale si ottiene:
   $X(T) = m_X \left\{1 - \left[0,45 + 0,78 \ln\left(-\ln\left(1 - \frac{1}{T}\right)\right)\right] CV\right\}$, dove CV è il coefficiente di variazione ($CV=\sigma_X/m_X$) e T il tempo di ritorno ($T= 1/(1-F_X(x))$)

3. Sia $Y = \frac{X-u}{\alpha} = g(X)$. Ne segue che la distribuzione di probabilità cumulata di *Y* è:

   $F_Y(y) = F_X\left(g^{-1}(y)\right)$ dove $g^{-1}(y) = \alpha y + u$. Pertanto:

   $F_Y(y) = \exp\left[-\exp\left(-\frac{g^{-1}(y)-u}{\alpha}\right)\right] = \exp\left[-\exp(-y)\right]$. Da questo segue:

   $Y = -\ln\left(-\ln\left(1-\frac{1}{T}\right)\right)$ e $X=\alpha Y+u$. Pertanto nel piano di Gumbel la variabile *X* è legata alla variabile ridotta *Y* da una legge lineare che mostra *u* come intercetta (parametro di posizione) e α come parametro che ne descrive la pendenza (parametro di scala).
4. Per riportare i punti campionari nel piano di Gumbel si procede nel seguente modo:
   - Si ordina il campione in modo crescente;
   - Si stima la frequenza campionaria con una formula del tipo: $\hat{F}_i = \frac{i-a}{N+b}$ dove *N* è il numero dei dati che compongono il campione (*a* e *b* variano a seconda della formula: nel caso di Gringorten *a*= 0,44 e *b*=0,12) e *i* è la posizione del *i*-esimo valore nel campione ordinato;
   - Si associa la frequenza campionaria stimata alla variabile ridotta di Gumbel nel seguente modo: $y_i = -\ln\left(-\ln\left(\hat{F}_i\right)\right)$;
   - Si formano le coppie x_i,y_i dove x_i è il valore numerico del dato campionario che occupata la i-esima posizione nel campione ordinato in modo crescente.

Domanda n°2 (8 punti)

1. Che cosa rappresenta la curva di riduzione dei volumi $\rho_{D,T}$ e quale è la sua definizione formale?
2. Sotto quali assunzioni è possibile scrivere $\rho_{D,T} \approx \rho_D = m_{Q_D}/m_Q$ dove m_{Q_D} rappresenta la media dei massimi annui della portata media sulla durata *D* e m_Q la media dei massimi annui al colmo.
3. Se $\rho_D = \left(1+\frac{D}{\alpha}\right)^{\beta-1}$, come faccio a stimare i due parametri? In particolare, di quali dati ho bisogno?
4. Come faccio a costruire un'onda di progetto di assegnato tempo di ritorno *T* se assumo $\rho_D = (1+D/\alpha)^{\beta-1}$ e una forma di tipo triangolare?
5. Come la 4 ma con forma di tipo Gamma.

Risposte

1. Sia $Q_{D,T}$ la portata media massima annua sulla durata *D* di assegnato tempo di ritorno *T*. Sia Q_T la portata al picco massima annua di assegnato tempo di ritorno. La curva di riduzione dei volumi è espressa dal rapporto $\rho_{D,T} = Q_{D,T}/Q_T$.
2. E' sempre possibile scrivere:

   $$\rho_{D,T} = \frac{Q_{D,T}}{Q_T} = \frac{m_{Q_D} + K_T\sigma_{Q_D}}{m_Q + K_T\sigma_{Q_Q}}$$

   dove K_T rappresenta il fattore di frequenza. Nel caso per entrambe le variabili si adotti una distribuzione di Gumbel possiamo scrivere:

$$\rho_{D,T} = \frac{m_{Q_D}\left\{1+K_T CV_{Q_D}\right\}}{m_Q\left\{1+K_T CV_Q\right\}} = \frac{m_{Q_D}\left\{1-\left[0,45+0,78\ln\left(\ln\left(\frac{T}{T-1}\right)\right)\right]CV_{Q_D}\right\}}{m_Q\left\{1-\left[0,45+0,78\ln\left(\ln\left(\frac{T}{T-1}\right)\right)\right]CV_Q\right\}}$$

Per T tendente ad infinito possiamo scrivere:

$$\rho_{D,T} \cong \frac{m_{Q_D}\left\{\left[0,45+0,78\ln\left(\ln\left(\frac{T}{T-1}\right)\right)\right]CV_{Q_D}\right\}}{m_Q\left\{\left[0,45+0,78\ln\left(\ln\left(\frac{T}{T-1}\right)\right)\right]CV_Q\right\}} = \frac{m_{Q_D}CV_{Q_D}}{m_Q CV_Q}$$

Infine, il coefficiente di variazione non varia con la durata e quindi $CV_{Q_D} \cong CV_Q$. Quindi:

$$\rho_{D,T} = \frac{Q_{D,T}}{Q_T} \cong \frac{m_{Q_D}}{m_Q} = \rho_D$$

3. Per stimare i parametri occorre avere a disposizione un campione di dati per ogni durata *D*, ciascuno composto dai valori massimi annui delle portate medie su detta durata, e un campione di dati dei valori massimi annui delle portate al colmo. Fissato un valore *D* si calcola il valore medio del campione di dati corrispondente; si calcola inoltre il valore medio del campione dei dati relativi alle portate massime annue al colmo. Per ogni valore di *D* si stima il rapporto $\widehat{\rho}_D = r_D = \frac{\widehat{m}_{Q_D}}{\widehat{m}_Q}$. A questo punto di stimano i due coefficienti α e β con il metodo dei minimi quadrati non lineare.
4. L'onda di progetto triangolare ha tempo di base t_B, tempo al picco t_p e portata al picco Q_p. Per prima cosa si pone la portata al picco uguale alla portata al colmo di assegnato tempo di ritorno, ovvero $Q_p = Q_T$. Si impone inoltre che il volume dell'onda abbia il medesimo tempo di ritorno. Il volume dell'onda di assegnato tempo di ritorno *T* si esprime nel seguente modo:
$V_{onda,T} = Q_{t_B,T} \times t_B$ dove $Q_{t_B,T}$ è la portata media sulla durata t_B avente tempo di ritorno T che a sua volta può essere scritto come: $Q_{t_B,T} = \left(1+t_B/\alpha\right)^{\beta} Q_T$. Tenendo inoltre conto che per semplice geometria il volume dell'onda triangolare si può anche scrivere base per altezza diviso 2, si ottiene:
$V_{onda,T} = Q_{t_B,T} \times t_B = \left(1+t_B/\alpha\right)^{\beta} Q_T \times t_B = t_B \times Q_T / 2$. Da questo segue:
$t_B = \alpha\left[\left(\frac{1}{2}\right)^{1/\beta-1} - 1\right]$. Per quanto riguarda infine t_p, questo si pone uguale a $r \times t_B$ dove r è il rapporto di ripartizione della portata (solitamente minore di 1/2).
5. Nel caso si assuma una distribuzione Gamma il procedimento è il seguente. Si parte dalla distribuzione gamma con area sottesa pari a 1, ovvero:
$f_T(t) = \frac{\lambda(\lambda t)^{k-1} e^{-\lambda t}}{\Gamma(k)}$ e si cerca il valore di $t = t_p$ che ne definisce il picco q_p. Il valore di tp si ottiene facendo la derivata di $f_T(t)$ rispetto al tempo e imponendola uguale a zero. Questa operazione porta a:
$t_p = \frac{1}{\lambda}(k-1);\ q_p = \frac{\lambda(k-1)^{(k-1)} e^{-(k-1)}}{\Gamma(k)}$.
Si impone che il picco sia uguale a Q_T. Da questo segue:

$Q_T = c \times q_p$ e il valore di c risulta:

$$c = \frac{Q_T \Gamma(k)}{\lambda (k-1)^{(k-1)} e^{-(k-1)}}$$

L'onda richiesta avrà equazione:

$$Q(t) = c \times f_T(t) = \frac{Q_T \Gamma(k)}{\lambda (k-1)^{(k-1)} e^{-(k-1)}} \frac{\lambda (\lambda t)^{k-1} e^{-\lambda t}}{\Gamma(k)}$$

Al fine di stimare i parametri λ e k ho bisogno di scrivere due relazioni. La prima si basa sulla equivalenza con l'onda triangolare (che ha tempo di base t_B, tempo al picco t_p e portata al picco Q_T). In particolare impongo che il volume dell'onda triangolare sia uguale al volume dell'onda gamma:

$$\overbrace{t_B Q_T / 2 = \frac{1}{2} Q_T \alpha \left[\left(\frac{1}{2} \right)^{1/\beta - 1} - 1 \right]}^{\text{volume onda triangolare}} = \overbrace{\frac{Q_T \Gamma(k)}{\lambda (k-1)^{(k-1)} e^{-(k-1)}}}^{\text{volume onda gamma}}.$$

La seconda relazione è invece riferita al tempo al picco:

$$\overbrace{\frac{1}{\lambda}(k-1)}^{\text{tempo al picco della funzione gamma}} = \overbrace{r t_B = r\alpha \left[\left(\frac{1}{2} \right)^{\frac{1}{\beta - 1}} - 1 \right]}^{\text{tempo al picco della funzione triangolare}}$$

Da queste due relazioni posso stimare λ e k e quindi definire pienamente l'onda di forma gamma di assegnato tempo di ritorno T.