Année 2005 – 2006 Session de Janvier 2006
ETAPE : INF 5 UE : INF110
Épreuve de : Méthodes Statistiques pour l'Informatique
Date : 10 Janvier 2006 Durée : 1h30

Documents autorisés
Épreuve de Mr Saheb

Département Licence

Indiquez votre code d'**anonymat** : | No : |

Barème
Questions : 5 points
Exercices 1, 2, 3, 4 : chacun 5 points

Questions
(Justifier vos réponses)

1. Si A et B sont indépendants, peut-on confirmer que A et $\overline{B}$ le sont aussi ?
2. Soit $Pr(B/A) < Pr(B)$, où les événements A et B sont supposés être de probabilité non nulle. Que peut-on dire de $Pr(A/B)$ et $Pr(A)$?
3. Peut-on déduire de $Pr(B/A) = Pr(B/\overline{A})$, où A et $\overline{A}$ sont supposés de probabilité non nulle, que A et B sont indépendants ?
4. Soient X et Y deux v.a. normales indépendantes de types $\mathcal{N}(\mu_1, \sigma_1)$ et $\mathcal{N}(\mu_2, \sigma_2)$ respectivement. Pour quelles valeurs des facteurs réels α et β la v.a. $\alpha X + \beta Y$ devient une v.a. normale centrée-réduite ?

Solution

1. Oui. En effet, si A et B sont indépendants, on a $Pr(A \cap \overline{B}) = Pr(A) - Pr(A \cap B) = Pr(A) - Pr(A)Pr(B) = Pr(A)\,[1 - Pr(B)] = Pr(A)Pr(\overline{B})$.
2. On a $Pr(A/B) < Pr(A)$. En effet $Pr(A/B) = \frac{Pr(A \cap B)}{Pr(B)} = \frac{Pr(A)Pr(B/A)}{Pr(B)} < Pr(A)$.
3. Oui. On a en effet : $\frac{Pr(A \cap B)}{Pr(A)} = \frac{Pr(\overline{A} \cap B)}{Pr(\overline{A})}$. Si l'on pose $Pr(\overline{A}) = 1 - Pr(A)$ et $Pr(\overline{A} \cap B) = Pr(B) - Pr(A \cap B)$ dans cette équation, on obtient, après les simplifications nécessaires $Pr(A \cap B) = Pr(A)Pr(B)$.
4. Pour que $\alpha X + \beta Y$ devienne centrée-réduite, il faut et il suffit que $\alpha\mu_1 + \beta\mu_2 = 0$ et $\alpha^2{\sigma_1}^2 + \beta^2{\sigma_2}^2 = 1$. Un calcul simple aboutit à la solution :

$$\alpha^2 = \frac{{\mu_2}^2}{{\mu_2}^2{\sigma_1}^2 + {\mu_1}^2{\sigma_2}^2}$$

$$\beta^2 = \frac{{\mu_1}^2}{{\mu_2}^2{\sigma_1}^2 + {\mu_1}^2{\sigma_2}^2},$$

où il faut coupler une racine positive et une racine négative.

Choisissez 3 exercices parmi les 4 exercices proposés

Exercice 1. La distribution conjointe des v.a. X et Y est donnée dans le tableau suivant :

X \ Y	-1	0	1	2
-1	0	1/36	1/6	1/12
0	1/18	0	1/18	0
1	0	1/36	1/6	1/12
2	1/12	0	1/12	1/6

1. Trouver $Pr(X \geqslant 1 \cap Y \leqslant 0)$.
2. Quelle est la probabilité conditionnelle pour que $Y \leqslant 0$ sachant que $X = 2$?
3. X et Y sont-elles indépendantes ?
4. Donner la distribution de $Z = XY$.

Solution

1. $Pr(X \geqslant 1 \cap Y \leqslant 0) = \frac{1}{36} + \frac{1}{12} = \frac{1}{9}$.
2. On a $Pr(Y \geqslant 0/X = 2) = \frac{\frac{1}{12}}{\frac{1}{12}+\frac{1}{12}+\frac{1}{6}} = \frac{1}{4}$.
3. Non. Il suffit de remarquer que $Pr(X = -1, Y = -1) \neq Pr(X = -1)Pr(Y = -1)$.
4. Z prend valeur dans l'ensemble $\Omega = \{-2, -1, 0, 1, 2, 4\}$. On trouve facilement par inspection :

$$Pr(Z = -2) = Pr(X = -1)Pr(Y = 2) + Pr(X = 2)Pr(Y = -1) = 1/6$$

et

$$Pr(Z = -1) = Pr(Z = 1) = Pr(Z = 2) = Pr(Z = 4) = 1/6.$$

Elle a donc une distribution uniforme sur Ω.

Exercice 2. On tire indépendamment deux v.a. X et Y uniformément dans l'intervalle réel $[a, b]$, où $a < b$. Posons $Z = \min\{X, Y\}$.

1. Calculer $Pr(Z > z)$ pour $z \in [a, b]$.
2. En déduire la fonction de répartition et la densité de la v.a. Z.
3. Calculer l'espérance et la variance de Z.

Solution

1. On a $Pr(Z > z) = Pr(X > z \cap Y > z) = Pr(X > z)Pr(Y > z) = \frac{(b-z)^2}{(b-a)^2}$.
2. On en déduit $F_Z(z) = 1 - \frac{(b-z)^2}{(b-a)^2}$. Donc la densité de Z vaut :

$$f(z) = \begin{cases} \frac{2(b-z)}{(b-a)^2} & \text{si } \; z \in [a, b] \\ 0 & \text{sinon} \end{cases}$$

3. Un calcul standard donne l'espérance et la variance de Z :
   - $\mathbb{E}Z = \int_a^b z f(z)dz = \frac{b+2a}{3}$
   - $\mathbb{V}Z = \int_a^b z^2 f(z)dz - (\mathbb{E}Z)^2 = \frac{3a^2+b^2+2ab}{6} - (\frac{b+2a}{3})^2 = \frac{(b-a)^2}{18}$.
   (La simplification n'était pas exigée.)

Exercice 3. Soit $A = \{a, b, c\}$ un alphabet à 3 lettres. Soit r un entier positif fixé. On effectue une suite de tirages (avec remises) uniformes de lettres dans A, jusqu'à ce qu'on obtienne r occurrences de "c". L'événement associé au tirage est alors le mot tiré et l'espace des événements sera donc :

$$\Omega \;=\; \{w \in A^* \mid |w|_c = r \text{ et } w \text{ se termine par une occurrence de "}c\text{"}\}.$$

On rappelle que $|w|_c$ est le nombre d'occurrences de "c" dans w et que $|w|$ est la longueur de w.

1. Pour $r = 2$, énumérer les mots de longueur 3 de Ω. Quelle est la probabilité de ce tirage ?
2. Montrer que l'on peut munir Ω d'une mesure de probabilité : $\forall w \in \Omega,\; Pr(w) =$ probabilité du tirage de w, avec $\sum_{w \in \Omega} Pr(w) = 1$.

3. On définit la v.a. $X : \Omega \to \mathbb{N}$, par $X(w) = |w|$. Calculer $Pr(X = n)$, pour $n = r, r+1, r+2, ...$
4. Trouver la fonction génératrice de X (la méthode conseillée est l'utilisation du résultat bien connu sur la f.g. de la somme des v.a. indépendantes).

Solution

1. Les mots de longueur 3 de Ω sont acc, bcc, cac, cbc.
2. Il suffit de mettre $Pr(w) = \frac{1}{3^{|w|}}$ (c'est la probabilité de tirage de w). La somme de ces probabilités vaut 1. En effet, avec la probabilité 1, le tirage finira à un moment par l'obtention de r occurrences de "c".
3. Pour que $X = n$, il faut que les $n-1$ premiers tirages contiennent exactement $r-1$ occurrences de "c" et $n-r$ occurrences de "a" ou "b" et que le dernier tirage soit une occurrence de "c" . Il y a $2^{n-r}\binom{n-1}{r-1}$ mots de longueur n qui satisfont ces conditions. D'où $Pr(X = n) = \frac{1}{3^n}\binom{n-1}{r-1}2^{n-r}$ pour $n = r, r+1, r+2, ...$
4. Avec la méthode conseillée, soient X_1, X_2,...,X_r les nombres de tirages de lettres nécessaires pour obtenir la 1ère occurrence, la 2ème occurrence,..., la $r$ème occurrence de "c" respectivement. La v.a. X est la somme des v.a. X_1, X_2,...,X_r. Or ces v.a. sont indépendantes et identiquement distribuées. Nous avons vu que la f.g. de cette distribution (nombre de tirages nécessaire pour l'obtention d'une première occurrence de "c" est $\frac{z}{3-2z}$. D'où la f.g. de X vaut :

$$G(z) = \frac{z^r}{(3-2z)^r}.$$

Exercice 4. Soit X une v.a. réelle uniformément répartie dans l'intervalle $[\theta_1, \theta_2]$, où θ_1 et θ_2, sont deux paramètres donnés (on suppose $\theta_1 < \theta_2$).

1. Donner la densité de la v.a. X et calculer son espérance et sa variance.
2. Écrire la densité conjointe $f_{\theta_1,\theta_2}(x_1, ..., x_n)$ de $(X_1, ..., X_n)$, où $X_1, ..., X_n$ sont n réalisations indépendantes de X.
3. Trouver les valeurs de θ_1 et θ_2 qui maximisent $f_{\theta_1,\theta_2}(x_1, ..., x_n)$, pour un vecteur $(x_1, ..., x_n)$ donné ; la solution est simple et ne nécessite pas de dérivation.

Solution

1. La densité de X vaut :

$$f(x) = \begin{cases} \frac{1}{\theta_2-\theta_1} & \text{si } x \in [\theta_1, \theta_2] \\ 0 & \text{sinon} \end{cases}$$

   L'espérance la variance de X se calculent facilement :

$$\mathbb{E}X = \int_{\theta_1}^{\theta_2} x f(x)dx = \frac{\theta_1 + \theta_2}{2}.$$

   et

$$\mathbb{V}X = \int_{\theta_1}^{\theta_2} x^2 f(x)dx - (\frac{\theta_1 + \theta_2}{2})^2 = \frac{(\theta_2 - \theta_1)^2}{12}$$

   (cf. cours7)
2. La densité conjointe vaut clairement :

$$f_{\theta_1,\theta_2}(x_1, ..., x_n)\frac{1}{(\theta_2 - \theta_1)^n}\mathbf{1}_{[\theta_1,\theta_2]^n}(x_1, ..., x_n)$$

3. Pour que $f_{\theta_1,\theta_2}(x_1, ..., x_n)$ soit maximum, il faut que le pavé $[\theta_1, \theta_2]^n$ contienne le point $(x_1, ..., x_n)$ et que, sous cette condition, le dénominateur $(\theta_2 - \theta_1)^n$ soit le plus petit possible. Ceci se réalise si $\theta_1 = \min\{x_1, ..., x_n\}$ et $\theta_2 = \max\{x_1, ..., x_n\}$.