ISIMA 1ère année — mardi 16 juin 2015

Documents et calculatrice autorisés — Durée : 2 heures

PROBABILITES

Exercice 1 :

Partie A : 10 points

I. Soit X une variable aléatoire continue de densité f_X, définie par $f_X(x) = k\,x$ sur le support $X(\Omega) = [\,0\,;\,2\,]$, avec $k > 0$.

1°) Déterminer la constante k.

2°) Déterminer la fonction de répartition F_X de X.

3°) Calculer $\mathrm{E}[X]$.

4°) Calculer $\mathrm{V}[X]$.

II. Soit $X' = \dfrac{X}{2}$.

1°) Déterminer la loi de X'.

2°) Calculer $\mathrm{E}[X']$.

3°) Calculer $\mathrm{V}[X']$.

III. Soient U une variable aléatoire uniforme sur $[\,0\,;\,1\,]$, indépendante de X, et $Y = \dfrac{X}{2} + U$.

1°) Déterminer la loi de Y.

2°) Calculer $\mathrm{E}[Y]$. Justifier vos calculs.

3°) Calculer $\mathrm{V}[Y]$. Justifier vos calculs.

Partie B : 10 points (on reprend les notations de la Partie A)

I. Soit $V = \begin{pmatrix} X \\ U \end{pmatrix}$.

1°) Déterminer la loi de V.

2°) Déterminer $\mathrm{E}[V]$.

3°) Déterminer la matrice de variance-covariance de V.

II. Soit $W = \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} X \\ X/2+U \end{pmatrix}$.

1°) Déterminer la loi de W et représenter graphiquement $W(\Omega)$.

Indication : On peut remarquer que $W = \begin{pmatrix} 1 & 0 \\ 1/2 & 1 \end{pmatrix}\begin{pmatrix} X \\ U \end{pmatrix} = \mathrm{M}\,V$ avec $\mathrm{M} = \begin{pmatrix} 1 & 0 \\ 1/2 & 1 \end{pmatrix}$.

2°) Déterminer $\mathrm{E}[W]$.

3°) Déterminer la matrice de variance-covariance de W.

III. 1°) Déterminer la loi de $Y_{|X=x}$.

2°) Déterminer $\mathrm{E}[Y_{|X=x}]$.

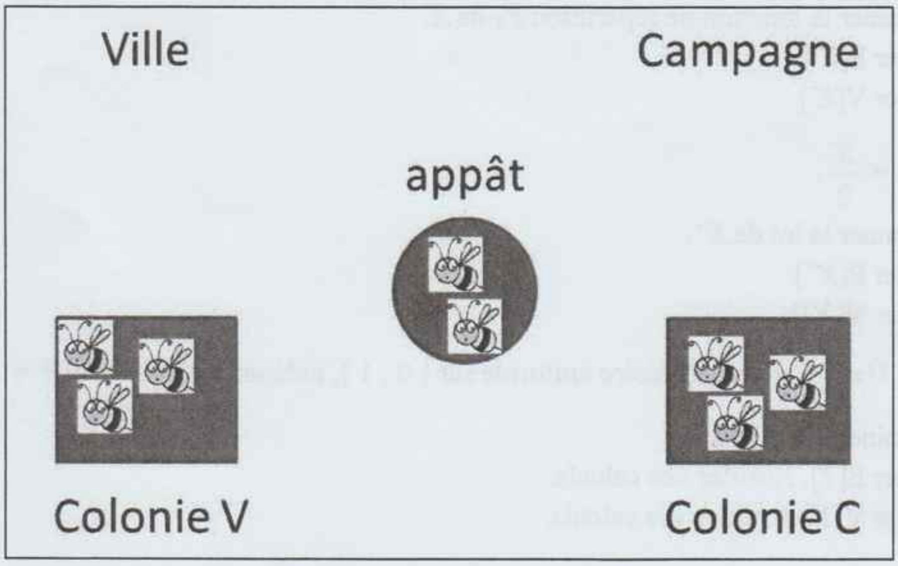
3°) En déduire $\mathrm{E}[Y \mid X]$.

IV. 1°) Déterminer la loi de $X_{\mid Y=y}$.

2°) Application : Calculer $\mathbf{P}\left(X>1 \,\middle|\, \frac{X}{2}+U=\frac{3}{4}\right)$.

Exercice 2 : 10 points

On considère 2 colonies d'abeilles, la première, V, dans une ville, et la seconde, C, à la campagne. On place un appât constitué d'une soucoupe remplie de liquide sucré à la lisière entre la ville et la campagne.



On suppose que chaque abeille de la colonie V a la même probabilité p_V d'être attirée par l'appât. On suppose également que chaque abeille de la colonie C a la même probabilité p_C d'être attirée par l'appât. (mais p_V et p_C ne sont pas forcément identiques)

Soient Nv et Nc le nombre d'abeilles de la colonie V et C respectivement, et soient b_V et b_C le nombre d'abeilles de la colonie V et C attirées par l'appât.

1°) a) Quelle distribution de probabilité suit b_V ? Nv pv bv

1°) b) Même question pour b_C.

2°) si on fait l'hypothèse que Nv et Nc sont très grands par rapport à b_V et b_C, quelles distributions peut-on utiliser comme approximations des lois de probabilité mentionnées à la première question ?

NB : On n'utilisera pas ces approximations pour les questions suivantes, et on s'en tiendra donc aux lois exactes établies à la première question pour la suite de l'exercice.

3°) En utilisant les valeurs $p_V=p_C=0.01$, $N_V=N_C=250$ et en supposant que b_V et b_C sont independants, calculer :

a) la probabilité $P(b_V = 3)$. Donner la formule, puis faire l'application numérique.

b) la probabilité $P(b_C = 0)$. Donner la formule, puis faire l'application numérique.

c) la probabilité $P(b_V = 3 \cap b_C = 0)$. Donner la formule, puis faire l'application numérique.

d) la probabilité $P(b_V = 3 \mid b_V+b_C=3)$. Donner la formule, puis faire l'application numérique.

4°) Y-a-t-il une différence entre $P(b_V = 3 \cap b_C = 0)$ et $P(b_V = 3 \mid b_V+b_C=3)$? Pourquoi ?

5°) Refaire l'application numérique de $P(b_V = 3 \mid b_V+b_C=3)$ avec $p_V=0.1$ et $p_C=0.01$.

6°) Discuter la différence de résultat entre les questions 3d et 5.

Exercice 3 :

Partie A : 6 points

On considère les séries de valeurs de températures journalières au sommet du puy de dôme (PDD), sur le campus des Cézeaux et à Brest pour les 16 derniers jours du mois de mai 2015 :

Lieu\Date	16/05	17/05	18/05	19/05	20/05	21/05	22/05	23/05	24/05	25/05	26/05	27/05	28/05	29/05	30/05	31/05
PDD	10	11	11	9	5	5	9	11	12	11	9	9	12	13	13	14
Campus	16	18	19	17	13	13	15	17	19	18	15	16	19	21	18	21
Brest	17	14	13	12	13	15	12	21	17	17	19	17	16	13	14	14

On notera ces valeurs de températures $T^{(PDD)}_i$, $T^{(Campus)}_i$ et $T^{(Brest)}_i$ pour $1 \leq i \leq 16$.

1°) Donner la formule d'un estimateur non biasé de la variance de la série de valeurs de températures sur le campus de Cézeaux et faire l'application numérique.

2°) En déduire un estimateur de la corrélation entre les séries de températures au sommet du puy de dôme et sur le campus des Cézeaux et faire l'application numérique.

3°) Même question pour la corrélation entre les températures sur le campus de Cézeaux et celles à Brest.

4°) Comparer les valeurs de corrélations trouvées aux questions 2 et 3 et discuter ces résultats.

Partie B : 4 points

On considère la série de valeurs des températures moyennes au cours du mois de mai à Clermont-Ferrand au cours des 15 dernières années :

Lieu\ Année	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
Clermont-Ferrand	15	14	16	13	15	15	16	10	12	12	17	16	12	14	16

On notera ces valeurs de températures $T^{(clermont)}_i$ pour $1 \leq i \leq 15$. On fera l'hypothèse que la distribution de ces valeurs de température est normale et stationnaire dans le temps (même si ce n'est pas le cas).

1°) Donner la formule de l'intervalle de confiance à 95% de la moyenne des températures en mai $\overline{T(clermont)}$. Faire l'application numérique.

2°) On souhaite tester si la distribution de température est stationnaire dans le temps. Proposer une façon de tester cette hypothèse et l'appliquer à ces données.