Sommes de variables aléatoires indépendantes

Les exercices et questions précédés d'une étoile (⋆) peuvent être passés dans un premier temps.

Si f et g sont deux fonctions définies sur $\mathbb{R}$, on rappelle que le produit de convolution de f et g est défini par $f \star g\,(z) = \displaystyle\int_{-\infty}^{+\infty} f(x)g(z-x)\,dx$ pour tout $z \in \mathbb{R}$ (pourvu que l'intégrale converge).
On admet que si des variables aléatoires indépendantes X et Y admettent respectivement pour densité f et g, alors $X+Y$ admet pour densité $f \star g$.

Exercice 1 : *Stabilité par somme*

1. Quelle est la loi d'une somme de n variables aléatoires indépendantes et suivant la même loi de Bernoulli de paramètre p ?
2. Soient X_1 et X_2 deux variables aléatoires discrètes indépendantes. Pourquoi la loi de $S = X_1 + X_2$ ne dépend-elle que des lois de X_1 et X_2 ? Donner l'espérance et la variance de S ?
   En déduire, à l'aide de la première question, que si X_1 et X_2 suivent respectivement les lois binomiales $\mathcal{B}(n,p)$ et $\mathcal{B}(n',p)$, alors $X_1 + X_2$ suit la loi binomiale $\mathcal{B}(n+n',p)$. Généraliser au cas de N variables binomiales indépendantes ayant même paramètre p.
3. Montrer que, si X et Y sont des variables aléatoires indépendantes suivant respectivement les lois de Poisson $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$, alors $X+Y$ suit la loi de Poisson $\mathcal{P}(\lambda+\mu)$. Généraliser au cas de la somme de N variables aléatoires de Poisson indépendantes.

Exercice 2 : *(Oral Véto 2008)*
Soient X, Y des variables aléatoires indépendantes de même loi donnée par :

$$\forall n \in \mathbb{N},\, \mathbb{P}(X=n) = \mathbb{P}(Y=n) = \frac{1}{4}\frac{1+a^n}{n!},$$

où $a \in \mathbb{R}$.

1. Déterminer a.
2. Calculer l'espérance de X.
3. Déterminer la loi de $X+Y$.

Exercice 3 : *Cas de deux lois discrètes uniformes*

1. Donner la loi de la somme de deux dés (on pourra faire un tableau et considérer les diagonales correspondant aux valeurs de la somme).
2. Dans une urne contenant n boules numérotées de 1 à n on prélève successivement deux boules avec remise. On note X_1 et X_2 les numéros des boules tirées. Quelles sont les lois de X_1 et X_2 ?
3. On pose $S = X_1 + X_2$. Préciser les valeurs possibles de S et calculer $P(S=k)$ (en distinguant le cas $k \leq n+1$ du cas $k \geq n+2$).

⋆ Exercice 4 : *Cas de deux lois géométriques*
Soient X et Y deux variables aléatoires indépendantes suivant respectivement les lois géométriques de paramètres p et p' (par exemple, X représente le rang du premier succès sur une suite d'épreuves de Bernoulli indépendantes avec, pour chacune, probabilités de succès p et d'échec $q = 1-p$).
Déterminer la loi de $X+Y$ (distinguer les cas $p = p'$ et $p \neq p'$).

Exercice 5 : *Cas de deux exponentielles*

Soient X et Y sont des variables aléatoires indépendantes et suivant la loi exponentielle de paramètre λ. Calculer une densité de probabilité de $X+Y$.

Exercice 6 : *(Partiel 2010)*

Soient X et Y deux variables aléatoires indépendantes, suivant toutes deux la loi exponentielle de paramètre 1. Le but de l'exercice est de déterminer la loi de la variable aléatoire $Z = \dfrac{X}{X+Y}$.

Pour tout réel t, on note $F_Z(t) = P(Z \leq t)$.

1. (a) Que vaut $P([X > 0] \cap [Y > 0])$?
   (b) Pourquoi la variable aléatoire Z est-elle bien définie ?
2. On suppose t fixé, $0 < t < 1$, et on pose $U = (t-1)X$, $V = tY$ et $W = U + V$.
   (a) Vérifier que $F_Z(t) = P(W \geq 0)$.
   (b) Déterminer les densités f_U et f_V des variables aléatoires U et V.
3. Déterminer une densité de W.
4. En déduire la loi de Z.

Exercice 7 :

1. (a) Soit f la fonction définie par $f(x) = ae^{-|x|}$ pour tout $x \in \mathbb{R}$. Déterminer a pour que f soit une densité de probabilité.
   (b) Soit Z une variable aléatoire admettant f pour densité. On dit alors que Z suit la loi de Laplace. Déterminer la fonction de répartition de Z, son espérance et sa variance.
2. Dans cette question, X et Y sont deux variables aléatoires indépendantes suivant toutes les deux la loi exponentielle de paramètre 1. Soit $Z = X - Y$.
   (a) Déterminer une densité de la variable aléatoire $-Y$.
   (b) Déterminer la loi de Z.
   (c) Déterminer la loi de la variable aléatoire $T = |Z|$.

Exercice 8 : *(Examen 2011)*

On considère deux variables aléatoires indépendantes X et Y suivant des lois exponentielles de paramètres respectifs 1 et 2. Déterminer la loi de $X+Y$.

Exercice 9 :

Soient X et Y deux variables aléatoires indépendantes de lois respectives la loi uniforme sur $[0,1]$ et la loi uniforme sur $[0,2]$ On pose $S = X + Y$.

1. Déterminer f_X et f_Y.
2. Pourquoi a-t-on sans calcul $f_S(s) = 0$ si $s \notin [0,3]$?
3. Déterminer la loi de S.

Exercice 10 : *Attente aux guichets (Agro concours A 2008)*

Trois personnes, Messieurs A, B et C, se présentent à l'ouverture d'un bureau de poste comportant deux guichets. Messieurs A et B accèdent immédiatement à un guichet, tandis que Monsieur C attend qu'un des deux guichets se libère et il y va aussitôt.

On note respectivement X, Y et Z les durées de passage aux guichets de Messieurs A, B et C. On suppose que X, Y et Z sont des variables aléatoires indépendantes et qu'elles suivent chacune la loi uniforme sur $[0,1]$.

1. Déterminer une densité de $U = X - Y$.
2. Montrer que $V = |U|$ est une variable à densité dont on donnera une densité.
3. Quelle est la probabilité pour que Monsieur C soit la dernière personne à sortir du bureau de poste ?

Exercice 11 : *Agro concours A 2008*

On choisit 2 nombres aléatoires X et Y dans $[0, 1]$ de façon uniforme et indépendante. On définit W par $W = 1$ si $XY \leq 1/2$ et $W = 0$ sinon.

1. Déterminer une densité de $\ln X + \ln Y$.
2. En déduire la probabilité pour que $\ln X + \ln Y \leq - \ln 2$. Préciser alors la loi de W.

Exercice 12 : *Stabilité des lois normales*

Soient X et Y des variables aléatoires gaussiennes indépendantes et de lois $\mathcal{N}(m, \sigma^2)$ et $\mathcal{N}(m', \sigma'^2)$. On pose $Z = X + Y$.

1. $\star$ On suppose dans cette question que $m = m' = 0$ et $\sigma' = 1$. Calculer la densité de Z et en déduire en particulier que Z est une variable aléatoire gaussienne.
   Indication : dans l'intégrale $\displaystyle\int_{-\infty}^{+\infty} f_X(x) f_Y(z - x)\, dx$, on pourra faire le changement de variable $u = x\dfrac{\sqrt{\sigma^2 + 1}}{\sigma} - z\dfrac{\sigma}{\sqrt{\sigma^2 + 1}}$.
2. Dans le cas général, on pose $A = \dfrac{X - m}{\sigma'}$ et $B = \dfrac{Y - m'}{\sigma'}$. Quelles sont les lois de A et B ? En déduire que $A + B$ est une variable aléatoire gaussienne.
3. Montrer que $Z = X + Y$ est une fonction affine de $A + B$ et en déduire que Z est une variable aléatoire gaussienne. Préciser alors sa loi.
4. Généraliser au cas de la somme de n variables aléatoires gaussiennes indépendantes. Que peut-on dire d'une combinaison linéaire de ces variables aléatoires ?

$\star$ Exercice 13 : *Lois Gamma et lois du chi-deux*

On dit qu'une variable aléatoire suit la loi Gamma $G(b, \nu)$ si elle admet pour densité la fonction f définie par $f(x) = 0$ si $x < 0$ et $f(x) = C\, x^{\nu-1} e^{-bx}$ si $x \geq 0$ (avec $C > 0$ tel que $\int_{-\infty}^{+\infty} f(x)\, dx = 1$).

1. Soient X et Y des variables aléatoires indépendantes et de lois respectives $G(b, \nu)$ et $G(b, \nu')$; notons f et g leurs densités. Montrer que $X + Y$ suit la loi $G(b, \nu + \nu')$.
   Indication : effectuer le changement de variable $x = tz$ dans l'intégrale $\int_0^z f(x) g(z - x)\, dx$ (on ne cherchera pas à calculer la valeur de la constante $\int_0^1 t^{\nu-1}(1 - t)^{\nu'-1} dt$).
2. Soit $Z = X_1^2 + \cdots + X_d^2$ où les X_i sont des variables aléatoires indépendantes normales centrées réduites.
   – Quelle est la densité de probabilité de X_1^2 (voir exercice 9 TD3) Remarquer que cette loi coïncide avec la loi $G(\frac{1}{2}, \frac{1}{2})$.
   – Montrer que la densité de la loi χ_d^2 vaut, pour $z > 0$, $h(z) = C_d\, z^{\frac{d-2}{2}} e^{-\frac{z}{2}}$ (on ne cherchera pas à préciser la constante C_d).
   – Quelle loi retrouve-t-on pour la loi du χ_2^2 ? En déduire la valeur de C_2

Exercice 14 : *Familiarisation avec la moyenne empirique*

On reprend la situation de l'exercice 10 du Td 2. Rappelons qu'on considérait que la distribution de l'âge des femmes dans la population française était gaussienne de moyenne 39 ans et d'écart-type 23 ans.
On demande leur âge à un échantillon de 15 françaises. On modélise cette situation en disant que l'âge X_i de la ième personne dans l'échantillon est une v.a. gaussienne de moyenne 39 ans et d'écart-type 23 ans et que les X_i sont indépendantes.

1. Donner, en fonction des X_i, la v.a. qui donne la moyenne d'âge des personnes choisies dans l'échantillon.

2. On appelle cette v.a. la **moyenne empirique**, on la note $\bar{X}_{15}$. Quelle est la loi de $\bar{X}_{15}$?
3. Quelle est la probabilité pour que la moyenne d'âge sur notre échantillon soit inférieure à 35 ans ?
4. Quelle est la probabilité pour que la moyenne d'âge sur notre échantillon soit supérieure à 60 ans ?
5. Quelle est la probabilité pour que la moyenne d'âge sur notre échantillon soit entre 35 et 60 ans ?
6. Donner une fourchette, centrée autour de la moyenne, dans laquelle on a 90% de chances de trouver la moyenne d'âge de notre échantillon.
7. Comparez qualitativement le comportement de X et celui de $\bar{X}_{15}$. Que se passerait-il si on augmentait la taille de l'échantillon ?

EXERCICE 15 : *Inégalité de Bienaymé-Tchebichev*

On cherche à observer des mutations sur des cellules identiques en les soumettant à une source de radioactivité. À l'issue de chaque expérience, la cellule testée a une probabilité p d'avoir muté.

1. Si on a réalisé n expériences indépendamment les unes des autres, quelle est la loi du nombre S_n de cellules mutées parmi les n testées ?
2. Quelle est l'espérance de S_n/n ? Sa variance ?
3. On admet que, si X est une variable aléatoire d'espérance m, alors pour tout $t > 0$,

$$P(|X - m| \geq t) \leq \frac{Var(X)}{t^2}.$$

   À l'aide de cette inégalité, dite de Bienaymé-Tchebichev, donner une majoration, en fonction de n et de p, de la probabilité que l'écart entre la proportion observée S_n/n et la proportion théorique p soit supérieure à ε. Application numérique : $p = 8\%$, $n = 50$ et $\varepsilon = 0,05$.
4. Pour ces valeurs de p, n et ε, à l'aide d'une approximation de la loi de S_n, calculer la probabilité que l'écart entre la proportion observée S_n/n et la proportion théorique p soit supérieure à ε. Que conclut-on pour la majoration obtenue à la question précédente ?