Université des Sciences et Technologies de Lille
U.F.R. de Mathématiques Pures et Appliquées

CAPES-M1 **Année 2010-2011**

Fiche de T.D. n° 2

Ex 1. *Urne à composition évolutive*

Dans une urne contenant au départ une boule verte et une rouge on effectue une suite de tirages d'une boule selon la procédure suivante. Chaque fois que l'on tire une boule verte, on la remet dans l'urne *en y rajoutant une boule rouge*. Si l'on tire une boule rouge, on arrête les tirages. On désigne par T le nombre de tirages effectués par cette procédure. On notera V_i (resp. R_i) l'événement *obtention d'une boule verte au ième tirage* (resp. *rouge*).

1) Pour $k \in \mathbb{N}^*$, donnez une expression de l'événement $\{T = k\}$ à l'aide des événements V_i $(1 \leq i \leq k-1)$ et R_k.

2) Que vaut $P(V_n \mid V_1 \cap \ldots \cap V_{n-1})$ pour $n \geq 2$?

3) Déterminez la loi de T.

4) Calculez $\mathbf{E}T$.

5) Calculez l'espérance de la variable aléatoire $\frac{1}{T}$.

6) On recommence l'expérience en changeant la procédure : à chaque tirage d'une boule verte on la remet dans l'urne *en y rajoutant une boule verte*. Comme précédemment, on interrompt les tirages à la première apparition d'une boule rouge. Soit S le nombre de tirages effectués suivant cette nouvelle procédure. Déterminez la loi de S. Que peut-on dire de l'espérance de S ? Interprétez.

Ex 2. *La formule de Poincaré*

La formule de Poincaré vue en cours a été démontrée par récurrence. Le but de cet exercice est d'en proposer une autre démonstration basée sur les propriétés de l'espérance.

1) Nous aurons besoin de la formule algébrique suivante où $x_1, \ldots, x_n$ sont des nombres réels quelconques $(n \geq 2)$.

$$\prod_{i=1}^{n}(1 - x_i) = \sum_{I \subset \{1,\ldots,n\}} (-1)^{\operatorname{card} I} \prod_{i \in I} x_i. \tag{1}$$

Dans cette formule la collection de tous les sous-ensembles I de $\{1, \ldots, n\}$ qui indexe le $\sum$ comprend en particulier $I = \emptyset$ pour lequel on pose par convention $\prod_{i\in\emptyset} x_i := 1$.

a) Vérifiez directement cette formule dans le cas $n = 3$ en développant le produit $(1 - x_1)(1 - x_2)(1 - x_3)$.

b) Démontrez la formule (1).

2) Dans toute la suite, on travaille sur un espace probabilisé $(\Omega, \mathcal{F}, P)$ fixé. On rappelle que si A est un sous-ensemble de Ω, sa fonction indicatrice $\mathbf{1}_A$ est l'application $\Omega \to \{0, 1\}$ définie par

$$\mathbf{1}_A(\omega) := \begin{cases} 1 & \text{si } \omega \in A, \\ 0 & \text{si } \omega \notin A. \end{cases}$$

Soient $B_1, \dots, B_k$, k sous-ensembles de Ω. Vérifiez l'égalité

$$\mathbf{1}_B = \prod_{j=1}^{k} \mathbf{1}_{B_j}, \quad \text{où } B := \bigcap_{j=1}^{k} B_j. \tag{2}$$

3) Montrez que pour tout entier $n > 1$ et toute suite finie $A_1, \dots A_n$ dans $\mathcal{F}$,

$$P\left(\bigcup_{i=1}^{n} A_i\right) = 1 - \mathbf{E}\left(\prod_{i=1}^{n}(1 - \mathbf{1}_{A_i})\right). \tag{3}$$

4) En utilisant tout ce qui précède, établir la formule de Poincaré, à savoir

$$\forall A_1, \dots, A_n \in \mathcal{F}, \qquad P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{\substack{I \subset \{1,\dots,n\} \\ I \neq \emptyset}} (-1)^{1+\operatorname{card} I} P\left(\bigcap_{i \in I} A_i\right). \tag{4}$$

Indiquez clairement quelles propriétés de l'espérance mathématique sont utilisées dans cette démonstration.

Ex 3. *Espérance du min et du max de deux v.a. uniformes*

Soient X et Y deux variables aléatoires *positives* définies sur le même espace probabilisé $(\Omega, \mathcal{F}, P)$, toutes deux de loi uniforme sur $[0, 1]$. Sauf dans la dernière question, on ne suppose pas X et Y indépendantes. On s'intéresse à l'espérance des variables aléatoires $\min(X, Y)$ et $\max(X, Y)$.

1) Calculez $\mathbf{E}X$ et $\mathbf{E}(X^2)$.

2) Écrivez plus simplement la somme $\min(X, Y) + \max(X, Y)$ et en déduire une relation simple entre $\mathbf{E}\min(X, Y)$ et $\mathbf{E}\max(X, Y)$. En déduire l'expression de $\mathbf{E}|X - Y|$ en fonction de $\mathbf{E}\min(X, Y)$, puis en fonction de $\mathbf{E}\max(X, Y)$.

3) Montrez (en deux lignes maximum) que l'on a toujours

$$\mathbf{E}\min(X, Y) \leq \frac{1}{2} \leq \mathbf{E}\max(X, Y).$$

Est-il possible avec un choix convenable de X et Y que ces deux inégalités soient des égalités ?

4) On suppose maintenant que X et Y sont indépendantes, ce qui implique notamment que pour tous intervalles I et J de $\mathbb{R}$ les évènements $\{X \in I\}$ et $\{Y \in J\}$ sont indépendants. Calculez $\mathbf{E}\min(X, Y)$, $\mathbf{E}\max(X, Y)$ et en déduire $\mathbf{E}|X - Y|$.
Indication : on pourra commencer par calculer $P(\min(X, Y) > t)$ ou $P(\max(X, Y) \leq t)$.

Ex 4. *Espérance des statistiques d'ordre*

1) *Question préliminaire.* Pour tout $n \in \mathbb{N}^*$ et tout $j = 0, 1, \dots, n$, on pose

$$I_{j,n} := \int_0^1 t^j (1-t)^{n-j} \, \mathrm{d}t.$$

Calculez $I_{n,n}$ et justifiez la relation : $\forall j = 0, \dots, n-1, \quad I_{j,n} = \frac{n-j}{j+1} I_{j+1,n}$. En déduire que

$$I_{j,n} = \frac{j!(n-j)!}{(n+1)!}.$$

Dans toute la suite de l'exercice, on note $X_1, \dots, X_n$ une suite de variables aléatoires *positives*, définies sur le même espace probabilisé $(\Omega, \mathcal{F}, P)$, *indépendantes et de même loi*, donc de même fonction de répartition notée F. Le but de ce problème est de calculer l'espérance des variables aléatoires $X_{n:1}, \dots X_{n:n}$ obtenues par réarrangement croissant de $X_1, \dots, X_n$ et appelées *statistiques d'ordre*. Plus précisément, pour tout $\omega \in \Omega$, on a *l'égalité des ensembles* $\{X_{n:k}(\omega);\ k = 1, \dots, n\}$ et $\{X_k(\omega);\ k = 1, \dots, n\}$ et de plus

$$\forall k = 1, \dots, n-1, \quad X_{n:k}(\omega) \le X_{n:k+1}(\omega).$$

En particulier, $X_{n:1} = \min(X_1, \dots, X_n)$ et $X_{n:n} = \max(X_1, \dots, X_n)$. Pour illustrer cette définition, on vous montre ci-dessous les valeurs des statistiques d'ordre pour $n = 5$ déterminées à partir des valeurs prises par la suite des $X_i(\omega)$ pour ω et ω' différents.

k	1	2	3	4	5
$X_k(\omega)$	$2,7$	$0,3$	6	$1,2$	$0,9$
$X_{5:k}(\omega)$	$0,3$	$0,9$	$1,2$	$2,7$	6

k	1	2	3	4	5
$X_k(\omega')$	$4,1$	$0,7$	$8,3$	$0,7$	$0,2$
$X_{5:k}(\omega')$	$0,2$	$0,7$	$0,7$	$4,1$	$8,3$

Noter que la présence d'*ex-æquo* ne pose pas de problème.

2) Soit I une partie non vide de $\{1, \dots, n\}$. En exploitant les hypothèses faites sur les X_i, exprimez à l'aide de $F(x)$, n et $\operatorname{card} I$ la probabilité

$$p_I(x) := P(\forall i \in I,\ X_i \le x \text{ et } \forall i \in \{1, \dots, n\} \setminus I,\ X_i > x).$$

3) On note F_k la fonction de répartition de la variable aléatoire positive $X_{n:k}$. Montrez que

$$\forall x \ge 0, \quad F_k(x) = P\big(X_{n:k} \le x\big) = \sum_{j=k}^{n} C_n^j F(x)^j \big(1 - F(x)\big)^{n-j}. \tag{5}$$

Indication : commencez par remarquer que « $X_{n:k} \le x$ » est équivalent à « au moins k parmi les X_i sont inférieurs ou égaux à x ».

4) Expliquez comment déduire de (5) la formule

$$\forall x \geq 0, \quad 1 - F_k(x) = \sum_{j=0}^{k-1} C_n^j F(x)^j \big(1 - F(x)\big)^{n-j}. \tag{6}$$

En déduire une formule générale exprimant $\mathbf{E}X_{n:k}$ à l'aide d'intégrales où intervient la fonction F.

5) On suppose dans cette question que les X_i suivent la loi uniforme sur $[0,1]$.

a) Explicitez alors $F(x)$ et achevez le calcul de $\mathbf{E}X_{n:k}$.

b) On pose de plus $X_{n:0} := 0$ et $X_{n:n+1} := 1$ et $Y_{n,k} := X_{n:k+1} - X_{n:k}$. Les variables aléatoires $Y_{n,k}$, $k = 0, \dots, n$ sont appelées les *espacements*. Que valent les $\mathbf{E}Y_{n,k}$?

6) On suppose dans cette question que les X_i suivent la loi exponentielle de paramètre 1. On a donc pour tout $x \geq 0$, $P(X_i > x) = \mathrm{e}^{-x}$. Calculez les $\mathbf{E}X_{n:k}$.

7) Montrez que si la variable aléatoire positive X_1 est *intégrable*, alors chacune des v.a. positives $X_{n:k}$ l'est aussi.

8) On suppose dans cette question que F est C^1 sur $[0, +\infty[$, avec $F(0) = 0$ et on note f sa dérivée. Montrez qu'alors $X_{n:k}$ admet pour densité la fonction f_k nulle sur $]-\infty, 0[$ et vérifiant

$$\forall x > 0, \quad f_k(x) = \frac{n!}{(k-1)!(n-k)!} F(x)^{k-1} \big(1 - F(x)\big)^{n-k} f(x).$$

Indication : commencez par vérifier les formules

$$jC_n^j = nC_{n-1}^{j-1}, \qquad (n-j)C_n^j = nC_{n-1}^j.$$

Ex 5. *Une inégalité de Markov à deux variables*

Dans tout l'exercice, X et Y sont deux variables aléatoires positives définies sur l'espace probabilisé $(\Omega, \mathcal{F}, P)$.

1) Montrez que

$$\forall s > 0,\ \forall t > 0, \qquad P(X > s \text{ et } Y > t) \leq \frac{1}{st} \mathbf{E}(XY). \tag{7}$$

Indication : on pourra utiliser après l'avoir soigneusement justifiée, l'inégalité

$$st \mathbf{1}_{\{X > s \text{ et } Y > t\}} \leq XY. \tag{8}$$

2) On suppose de plus que $c := \mathbf{E}(XY) < +\infty$. Pour quelles valeurs de (s,t), l'inégalité (7) a-t-elle un intérêt ? Représentez graphiquement l'ensemble de ces valeurs de (s,t).

3) On suppose maintenant que X et Y sont indépendantes. Expliquez comment on peut dans ce cas retrouver (7) en utilisant l'inégalité de Markov classique.

Ex 6. *Loi de Gumbel*

On considère la fonction $F : \mathbb{R} \to \mathbb{R}$, $x \longmapsto \exp\big(-\mathrm{e}^{-ax}\big)$, où $a > 0$ est une constante.

1) Vérifiez que F est la fonction de répartition d'une loi de probabilité sur $\mathbb{R}$. Cette loi s'appelle *loi de Gumbel de paramètre a*.

2) Expliquez pourquoi cette loi admet une densité f et calculez f.

3) Soit X une variable aléatoire réelle suivant la loi de Gumbel. Pour toute constante réelle b, $Y := \mathrm{e}^{bX}$ est une v.a. positive. Exprimez l'espérance de Y comme une intégrale généralisée dont vous donnerez l'expression à l'aide de l'intégrale généralisée plus simple $\int_0^{+\infty} u^{s-1}\mathrm{e}^{-u}\,\mathrm{d}u$. On rappelle que cette dernière définit $\Gamma(s)$ pour $s > 0$. Pour quelles valeurs de b, la v.a. Y est-elle intégrable ?

4) Soient $X_1, \dots, X_n, \dots$ des variables aléatoires indépendantes et de même loi de Gumbel de paramètre a. On pose

$$M_n := \max_{1 \le k \le n} X_k.$$

Montrez que M_n a même loi que $X_1 + c_n$, où les c_n sont des constantes que l'on calculera. *Indication :* commencez par calculer la fonction de répartition de M_n et constatez qu'elle s'exprime très simplement à l'aide de F.

Ex 7. *Médiane(s) et minimisation d'écart*

Soit X une variable aléatoire réelle de fonction de répartition F. On appelle médiane de X (ou de la loi de X), tout réel m vérifiant

$$P(X \le m) \ge \frac{1}{2} \quad \text{et} \quad P(X \ge m) \ge \frac{1}{2}.$$

On admettra qu'il y a toujours au moins une médiane. Le but de cet exercice est de prouver la propriété de minimisation suivante. Si X est intégrable et m est une médiane de X, alors pour tout réel c,

$$\mathbf{E}|X - c| \ge \mathbf{E}|X - m|. \tag{9}$$

1) Vérifiez que si m est une médiane de X, $-m$ est une médiane de $-X$.

2) On choisit une médiane m de X et on fixe $c > m$. On pose

$$Y := |X - c| - |X - m|$$

et on se propose de montrer l'inégalité $\mathbf{E}Y \ge 0$.

2.a) Justifiez l'intégrabilité de Y.

2.b) Expliquez la décomposition $Y = Y_1 + Y_2 + Y_3$, où

$$Y_1 := Y\mathbf{1}_{]-\infty,m]}(X), \quad Y_2 := Y\mathbf{1}_{]m,c]}(X), \quad Y_3 := Y\mathbf{1}_{]c,+\infty]}(X).$$

2.c) Vérifiez les égalités

$$Y_1 = (c-m)\mathbf{1}_{]-\infty,m]}(X), \quad Y_2 = (m+c-2X)\mathbf{1}_{]m,c]}(X), \quad Y_3 = (m-c)\mathbf{1}_{]c,+\infty]}(X).$$

2.d) Justifiez l'inégalité $Y_2 \geq (m-c)\mathbf{1}_{]m,c]}(X)$ et en déduire que

$$\mathbf{E}Y \geq (c-m)\big(P(X \leq m) - P(X > m)\big).$$

2.e) Conclure sur la positivité de $\mathbf{E}Y$.

3) Sans reprendre toute l'étude ci-dessus pour $c < m$, compléter la démonstration de la propriété de minimisation (9).

4) Montrez que si m et m' sont deux médianes de X, $\mathbf{E}|X-m| = \mathbf{E}|X-m'|$.