Devoir maison de probabilités — 2M231
Proposition de corrigé

Exercice 1

(a) C'est vrai. En utilisant la formule de transfert, pour tout entier n, on a :

$$\begin{aligned} P(f(X)=k) &= P\left(X \in f^{-1} < k >\right) \\ &= P\left(Y \in f^{-1} < k >\right) \qquad \text{car } X \text{ et } Y \text{ ont même loi} \\ &= P\left(f(Y)=k\right) \end{aligned}$$

(b) C'est faux. On peut par exemple considérer l'expérience aléatoire telle que :

$$\begin{array}{rcl} P\left(A \cap B \cap C\right) &=& 0,125 \\ P\left(A \cap B \cap \overline{C}\right) &=& 0,1 \\ P\left(A \cap \overline{B} \cap C\right) &=& 0,1 \\ P\left(A \cap \overline{B} \cap \overline{C}\right) &=& 0,175 \\ P\left(\overline{A} \cap B \cap C\right) &=& 0,1 \\ P\left(\overline{A} \cap B \cap \overline{C}\right) &=& 0,175 \\ P\left(\overline{A} \cap \overline{B} \cap C\right) &=& 0,175 \\ P\left(\overline{A} \cap \overline{B} \cap \overline{C}\right) &=& 0,05 \end{array}$$

Cette expérience est construite de telle sorte que $P\left(A \cap B \cap C\right) = 0,125 = 0,5^3 = P(A) \times P(B) \times P(C)$ mais $P(A \cap B) = 0,225 \neq P(A) \times P(B)$.
Ainsi, les événements A, B et C ne sont pas indépendants !

(c) On rappelle que la propriété de convergence en loi porte sur les espérances de fonctions bornées.
Or la fonction $x \mapsto x$ n'est pas bornée.
Il s'agit donc de trouver un contre exemple.
On va considérer la suite de variable $(X_n)_{n\in\mathbb{N}}$ définie par

$$\left\{ \begin{array}{rcl} P(X_n = 0) &=& 1 - \frac{1}{n} \\ P(X_n = n) &=& \frac{1}{n} \end{array} \right.$$

Pour tout n, $E(X_n) = 1$.
D'autre part, pour toute fonction continue et bornée f et pour tout n :

$$E\left(f(X_n)\right) = \left(1 - \frac{1}{n}\right) f(0) + \frac{1}{n} f(n)$$

Comme f est bornée, $\lim\limits_{n\to+\infty} E\left(f(X_n)\right) = f(0)$
On en déduit que X_n converge en loi vers 0.
Pourtant $E(0) \neq \lim\limits_{n\to+\infty} E(X_n) = 1$.
On aurait aussi pu utiliser la convergence simple de la suite des fonctions caractéristiques.

(d) C'est faux.
En effet, soit la suite de variables U_n qui suivent des lois uniformes sur $[0;\ n]$. Leurs fonctions de répartitions $\dfrac{\mathbb{1}_{[0;\ n]}}{n}$ convergent simplement vers 0 qui n'est pas une fonction de répartition.

Exercice 2

(a) X suit une loi binomiale de paramètres n et p. En effet, X est bien une somme de n Bernoulli indépendantes de paramètres p, chacune des Bernoulli représentant le fait que le pétard éclate ou non.

(b) $Y - X$ représente le nombre de pétards qui éclatent le 14 juillet exactement.
Pour X donné, $Y - X$ suit une loi de Bernoulli de paramètres $n - X$ et q.
On en déduit :

$$\begin{aligned} P(Y = k|X = l) &= P(Y - X = k - l|X = l) \\ &= \begin{cases} 0 \text{ si } k > n \text{ ou } k < l \\ \binom{n-l}{k-l} q^{k-l}(1-q)^{n-k} \text{ sinon} \end{cases} \end{aligned}$$

(c) Pour trouver la loi de Y, on utilise ce qui précède et la formule des causes (ou probabilités totales) :

$$\begin{aligned} P(Y = k) &= \sum_{0 \le l \le k} P(Y = k|X = l) \times P(X = l) \\ &= \sum_{0 \le l \le k} \binom{n-l}{k-l} q^{k-l}(1-q)^{n-k} \times \binom{n}{l} p^l (1-p)^{n-l} \end{aligned}$$

Or, un peu de calcul montre que :

$$\binom{n-l}{k-l} \times \binom{n}{l} = \binom{n}{k} \times \binom{k}{l}$$

On en déduit :

$$\begin{aligned} P(Y = k) &= \binom{n}{k} (1-p)^{n-k}(1-q)^{n-k} \sum_{0 \le l \le k} \binom{k}{l} q^{k-l}(1-p)^{k-l} p^l \\ &= \binom{n}{k} (1-p)^{n-k}(1-q)^{n-k} \left(q(1-p) + p\right)^k \\ &= \binom{n}{k} (1-p)^{n-k}(1-q)^{n-k} (q + p - pq)^k \\ &= \binom{n}{k} \left[(1-p)(1-q)\right]^{n-k} \left[1 - (1-p)(1-q)\right]^k \end{aligned}$$

On reconnaît que Y suit une loi binomiale de paramètres n et $1 - (1-p)(1-q)$.
On considérant Z le nombre de pétards qui n'ont pas du tout éclaté, il est clair que Z suit une loi binomiale de paramètres n et $(1-p)(1-q)$.
Et comme $Y = n - Z$, on peut retrouver le résultat précédent.

Exercice 3

(a) X compte le nombre d'éléments de type 1 dans un ensemble de n éléments.
Il y a $\binom{N}{n}$ manières de réaliser un tirage de n éléments parmi N.
On fixe $n \leq N$ et on détermine $P(X = m)$.
Si $m > M$, $P(X = m) = 0$.
De même, si $n - m > N - M$, c'est à dire si $m < n + M - N$, $P(X = m) = 0$.
Dans les cas contraire, si $m \in [\![n + M - N;\ M]\!]$, il y a $\binom{M}{m}$ manières de choisir m éléments de type 1 et $\binom{N-M}{n-m}$ manières de choisir $n - m$ éléments de type 2.
Finalement, la loi de X est :

$$P(X = m) = \begin{cases} \frac{\binom{M}{m}\times\binom{N-M}{n-m}}{\binom{N}{n}} & \text{si } m \in [\![n + M - N;\ M]\!] \\ 0 & \text{sinon} \end{cases}$$

(b) Pour simplifier, dans cette question, on va supposer que $n \leq N - M$ et $n \leq M$, c'est à dire que l'on tire un nombre d'éléments inférieur au stock de pièces de type 1 et au stock de pièces de type 2.
Pour établir la loi de X_k, il y a plusieurs approches. La première consiste à calculer la loi de X_k en réalisant une disjonction de cas sur le nombre de pièces de type 1 tirées lors des $k - 1$ premiers tirages. Cela conduit à une formule complexe, avec des coefficients binomiaux, difficilement exploitables.
Une autre option consiste à procéder par tâtonnement sur les premières valeurs de k. En adoptant cette approche, on constate que la loi des X_k semble indépendante des valeurs de k et que l'on a toujours $P(X_k = 1) = \dfrac{M}{N}$ et $P(X_k = 0) = \dfrac{N - M}{N}$.
On va prouver cela par récurrence.
Il est clair que $P(X_1 = 1) = \dfrac{M}{N}$ et $P(X_1 = 0) = \dfrac{N - M}{N}$, ce qui initialise la démonstration.
On suppose la propriété vraie jusqu'à un certain rang k.
Au rang $k + 1$:

$$P(X_{k+1} = 1) = P(X_{k+1} = 1|X_k = 1) \times P(X_k = 1) + P(X_{k+1} = 1|X_k = 0) \times P(X_k = 0)$$

En appliquant l'hypothèse de récurrence :

$$P(X_{k+1} = 1) = \frac{M - 1}{N - 1} \times \frac{M}{N} + \frac{M}{N - 1} \times \frac{N - M}{N} = \frac{M^2 - M + MN - M^2}{N(N - 1)} = \frac{M}{N}$$

Ainsi, X_k est une Bernoulli de paramètre $\dfrac{M}{N}$. En particulier :

$$E(X_k) = \frac{M}{N} \qquad \text{et} \qquad V(X_k) = \frac{M}{N}\left(1 - \frac{M}{N}\right)$$

(c) X_iX_j ne peut pendre que deux valeurs : 0 ou 1. Son espérance est donc :

$$E(X_iX_j) = P(X_i = 1 \text{ et } X_j = 1) = P(X_j = 1|X_i = 1)P(X_i = 1)$$

On peut supposer $i < j$. $P(X_j = 1|X_i = 1)$ correspond à la probabilité de tirer un élément de type 1 sachant que l'on a tiré un élément de type 1 précédemment.
Ainsi,

$$P(X_j = 1|X_i = 1) = \frac{M - 1}{N - 1}$$

Et donc :

$$E(X_iX_j) = \frac{M-1}{N-1} \times \frac{M}{N}$$

Finalement,

$$\operatorname{cov}(X_i;\ X_j) = \frac{M-1}{N-1} \times \frac{M}{N} - \left(\frac{M}{N}\right)^2$$

(d) On sait que $X = \sum_{1 \leq k \leq n} X_k$ et donc, par linéarité de l'espérance :

$$E(X) = \sum_{1 \leq k \leq n} E(X_k) = n \times \frac{M}{N}$$

On peut également calculer :

$$\begin{aligned} V(X) &= \sum_{1 \leq k \leq n} V(X_k) + 2 \sum_{1 \leq i < j \leq n} \operatorname{cov}(X_i;\ X_j) \\ &= n \times \frac{M}{N} \times \left(1 - \frac{M}{N}\right) + 2 \times \frac{n(n-1)}{2} \times \left(\frac{M-1}{N-1} \times \frac{M}{N} - \left(\frac{M}{N}\right)^2\right) \\ &= \frac{nM(N-M)(N-n)}{N^2(N-1)} \end{aligned}$$

Exercice 4

(a) Soient $(x;\ h) \in \mathbb{R} \times \mathbb{R}^*$.

$$\begin{aligned} \frac{1}{h}\left(F(x+h) - F(x)\right) - \rho(x) &= \frac{1}{h}\left(\int_x^{x+h} \rho(t)\mathrm{d}t - h\rho(x)\right) \\ &= \frac{1}{h}\int_x^{x+h} \left(\rho(t) - \rho(x)\right) \mathrm{d}t \end{aligned}$$

On suppose maintenant que ρ est continue en x.
Dans ce cas, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que pour tout $|h| \leq \eta$ et pour tout $t \in [x;\ x+h]$ ou $[x+h;\ x]$, $|\rho(t) - \rho(x)| \leq \varepsilon$.
En particulier, pour tout $h \neq 0$ tel que $|h| < \eta$, on en déduit :

$$\left|\frac{1}{h}\left(F(x+h) - F(x)\right) - \rho(x)\right| \leq \frac{1}{h}\int_x^{x+h} |\rho(t) - \rho(x)|\, \mathrm{d}t \leq \varepsilon$$

Cela prouve que F est dérivable en x et de dérivée ρ.

(b) (i) La fonction $x \mapsto \dfrac{1}{x}$ est mesurable et est définie sur $\mathbb{R}^*$.
Or X ne prend que des valeurs supérieures à $\frac{1}{2}$.
Donc Y est bien une variable aléatoire.

(ii) Ici, il faut calculer $F_Y(t) = P\left(\frac{1}{X} \leq t\right) = P\left(X \geq \frac{1}{t}\right) = 1 - P\left(X \leq \frac{1}{t}\right) + P\left(X = \frac{1}{t}\right)$.
Or $P\left(X = \frac{1}{t}\right) = 0$ puisque X suit une loi à densité. On obtient donc :

$$F_Y(t) = 1 - F_X\left(\frac{1}{t}\right)$$

(c) On obtient ce résultat par dérivation composée (ou par changement de variable dans une intégrale). Pour tout t, $F_Y'(t) = \frac{1}{t^2}\rho_X\left(\frac{1}{t}\right)$.

Et comme la fonction $t \mapsto \rho_X\left(\frac{1}{t}\right)$ est continue par morceaux, on peut conclure.

(d) La fonction de densité de X est $2\mathbb{1}_{[1/2;\ 1]}$ donc la fonction de densité de Y est $t \mapsto \frac{2}{t^2}\mathbb{1}_{[1/2;\ 1]}\left(\frac{1}{t}\right)$, c'est à dire $t \mapsto \frac{2}{t^2}\mathbb{1}_{[1;\ 2]}(t)$.

L'espérance de Y est donc

$$E(Y) = 2\int_1^2 \frac{\mathrm{d}t}{t} = 2[\ln(t)]_1^2 = 2\ln(2)$$

L'espérance de Y^2 est également

$$E(Y^2) = 2\int_1^2 \mathrm{d}t = 2$$

Finalement, on en déduit la variance de Y :

$$V(Y) = 2 - 4\ln(2)^2$$

(e) (i) Nous venons de prouver que $\frac{1}{X_1}$ était L^1 et que son espérance était $2\ln(2)$.

Ainsi, la loi forte des grands nombres entraîne que M_n converge presque sûrement vers $2\ln(2)$

(ii) Posons $S_n = \displaystyle\sum_{1\leq k\leq n} \frac{1}{X_k}$.

Puisque $\frac{1}{X_1}$ est L^2, nous sommes dans les conditions d'application du théorème central limite. En posant $\mu = E\left(\frac{1}{X_1}\right) = 2\ln(2)$ et $\sigma = \sqrt{V\left(\frac{1}{X_1}\right)} = \sqrt{2 - 4\ln(2)^2}$, on sait que

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} \xrightarrow[\mathcal{L}]{} \mathcal{N}(0;\ 1)$$

Or

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} = \frac{\sqrt{n}}{\sigma}(M_n - \mu)$$

Ainsi, on en déduit, que pour $a = \mu = 2\ln(2)$,

$$\lim_{n\to+\infty} P\left(\sqrt{n}(M_n - a) \in [-1;\ 1]\right) = P\left(N \in \left[\frac{-1}{\sigma};\ \frac{1}{\sigma}\right]\right)$$

Où N est une variable aléatoire qui suit une loi normale centrée réduite.