Corrigé de l'interrogation du 5 mars 2018

Exercice 1

1. (a) Soit $n \geq 1$ un entier naturel. pour $1 \leq k \leq n$, on définit la proposition $\mathcal{P}_k$ par :

$$\mathcal{P}_k : \text{“On a l'égalité } I_{k,n} = \frac{(k-1)!(n-k)!}{n!}\text{”}$$

La proposition $\mathcal{P}_k$ est vraie pour $k = 1$ puisque :

$$I_{1,n} = \int_0^1 (1-x)^{n-1} dx = \left[-\frac{(1-x)^n}{n} \right]_0^1 = \frac{1}{n} = \frac{0!(n-1)!}{n!}.$$

Supposons maintenant que $\mathcal{P}_k$ est vraie pour un certain $1 \leq k \leq n-1$. En intégrant par partie, on obtient :

$$\begin{aligned} I_{k+1,n} = \int_0^1 x^k (1-x)^{n-k-1} dx &= -\left[x^k \frac{x^{n-k}}{n-k} \right]_0^1 + \int_0^1 k x^{k-1} \frac{x^{n-k}}{n-k} dx \\ &= 0 + \frac{k}{n-k} I_{k,n} \\ &= \frac{k}{n-k} \times \frac{(k-1)!(n-k)!}{n!} \\ &= \frac{k!(n-k-1)!}{n!}. \end{aligned}$$

La proposition $\mathcal{P}_{k+1}$ est donc vraie.
Par récurrence sur k, on a donc montré la proposition $\mathcal{P}_k$, pour tout $1 \leq k \leq n$.

(b) Par définition de $I_{k,n}$, on a clairement $\int_{\mathbb{R}} p_{k,n}(x)dx = 1$. Pour montrer que $p_{k,n}$ est une densité de probabilité, il reste donc seulement à remarquer que la fonction $p_{k,n}$ est positive. C'est bien le cas, puisqu'elle est nulle hors de $[0,1]$, et que sur $[0,1]$ elle s'exprime comme produit de puissances des quantités positives x et $(1-x)$.

2. Comme $p_{k,n}$ est nulle hors de $[0,1]$, une variable aléatoire X de densité $p_{k,n}$ est donc presque sûrement à valeurs dans $[0,1]$. X est donc bornée, et, par suite, intégrable.

On a ensuite :

$$\mathbb{E}X = \int_{\mathbb{R}} x p_{k,n}(x) dx = \frac{1}{I_{k,n}} \int_0^1 x^k (1-x)^{n-k} dx = \frac{I_{k+1,n+1}}{I_{k,n}}.$$

D'après la question (1a), on a donc

$$\mathbb{E}X = \frac{k!(n-k)!}{(n+1)!} \times \frac{n!}{(k-1)!(n-k)!} = \frac{k}{n+1}.$$

3. (a) Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction bornée. On a

$$\begin{aligned} \mathbb{E}f(Y_K) = \mathbb{E}\left(f(Y_K) \times \sum_{k=1}^n \mathbf{1}_{\{K=k\}} \right) = \sum_{k=1}^n \mathbb{E}\left(f(Y_K) \mathbf{1}_{\{K=k\}} \right) &= \sum_{k=1}^n \mathbb{E}\left(f(Y_k) \mathbf{1}_{\{K=k\}} \right) \\ &= \sum_{k=1}^n \mathbb{E}\left(f(Y_k) \right) \mathbb{P}(K=k). \end{aligned}$$

Dans la dernière égalité, on a utilisé l'indépendance entre K et Y_k. Or, on a par hypothèses $\mathbb{P}(K=k) = 1/n$ et $\mathbb{E}(f(X_k)) = \int_{\mathbb{R}} f(x) q_k(x) dx$. Autrement dit, on a

$$\mathbb{E}(f(Y_K)) = \int_{\mathbb{R}} f(x) \times \left(\frac{1}{n} \sum_{k=1}^n q_k(x) \right) dx.$$

Cette égalité étant vraie pour toute fonction bornée f, la densité de Y_K est bien donnée par $x \mapsto \frac{1}{n} \sum_{k=1}^n q_k(x)$.

(b) Si $q_k = p_{k,n}$ la densité f de Y_K est nulle hors de $[0,1]$ et vérifie pour tout $0 \leq x \leq 1$, en utilisant l'égalité $(I_{k,n})^{-1} = n\binom{n-1}{k-1}$:

$$f(x) = \frac{1}{n}\sum_{k=1}^{n} p_{k,n}(x) = \frac{1}{n}\sum_{k=1}^{n} n\binom{n-1}{k-1}x^{k-1}(1-x)^{n-k} = \sum_{q=0}^{n-1}\binom{n-1}{q}x^q(1-x)^{n-1-q},$$

où l'on a posé $q = k-1$ dans la dernière égalité. On reconnaît le développement de Newton de $(x+(1-x))^{n-1} = 1$. La densité de Y_K est donc $\mathbf{1}_{[0,1]}$: Y_K suit la loi uniforme sur $[0,1]$.

Exercice 2

1. L'évènement $\{Y_k \leq x, Y_{k+1} > x\}$ est l'évènement sur lequel k des variables $(X_1, \ldots, X_n)$ sont inférieures ou égales à x, et les $n-k$ variables restantes sont strictement supérieures à x. On a donc
$$\{Y_k \leq x, Y_{k+1} > x\} = \bigcup_{E \subset \{1,\ldots,n\}, \operatorname{card} E = k} \{\forall i \in E, X_i \leq x, \forall i \notin E, X_i > x\}.$$
On a donc, l'union étant disjointe,
$$\mathbb{P}(Y_k \leq x, Y_{k+1} > x) = \sum_{E \subset \{1,\ldots,n\}, \operatorname{card} E = k} \mathbb{P}(\forall i \in E, X_i \leq x, \forall i \notin E, X_i > x).$$
Par symétrie de la loi du n-uplet $(X_1, \ldots, X_n)$, quel que soit $E \subset \{1, \ldots, n\}$ de cardinal k, on a
$$\mathbb{P}(\forall i \in E, X_i \leq x, \forall i \notin E, X_i > x) = \mathbb{P}(X_1 \leq x, \ldots, X_k \leq x, X_{k+1} > x, \ldots, X_n > x).$$
Comme il y a $\binom{n}{k}$ tels sous-ensembles E, on a bien l'égalité demandée.

2. Comme les Y_k vérifient $Y_1 \leq Y_2 \leq \ldots \leq Y_n$, on a l'égalité
$$\Omega = \bigcup_{q=0}^{n}\{Y_q \leq x, Y_{q+1} > x\},$$
où l'union est disjointe (q correspond au plus grand k tel que $Y_k \leq x$). En prenant l'intersection avec l'évènement $\{Y_k \leq x\}$, on obtient $\{Y_k \leq x\} = \bigcup_{q=k}^{n}\{Y_q \leq x, Y_{q+1} > x\}$, où l'union est encore disjointe. Le résultat s'obtient en passant à l'espérance.

3. On a clairement $\mathbb{P}(Y_k \leq x) = 0$ si $x < 0$ et $\mathbb{P}(Y_k \leq x) = 1$ si $x > 1$, puisque les X_k sont à valeurs dans $[0,1]$. Soit donc $0 \leq x \leq 1$. D'après les questions précédentes, on a
$$\begin{aligned}\mathbb{P}(Y_k \leq x) = \sum_{q=k}^{n}\mathbb{P}(Y_q \leq x, Y_{q+1} > x) &= \sum_{q=k}^{n}\binom{n}{q}\mathbb{P}(X_1 \leq x, \ldots, X_q \leq x, X_{q+1} > x, \ldots, X_n > x) \\ &= \sum_{q=k}^{n}\binom{n}{q}x^q(1-x)^{n-q},\end{aligned}$$
où la dernière égalité utiliser l'indépendance des $(X_k)_{1\leq k\leq n}$ et le fait que chaque X_k suive la loi uniforme sur $[0,1]$. Cette expression est de classe $\mathcal{C}^1$, vaut 0 en 0 et 1 en 1. La fonction de répartition de Y_k est donc continue et $\mathcal{C}^1$ par morceaux. Y_k admet donc une densité f_k qui s'obtient en dérivant la fonction de répartition :
$$\begin{aligned}f_k(x) &= \sum_{q=k}^{n}\binom{n}{q}\left(qx^{q-1}(1-x)^{n-q} - (n-q)x^q(1-x)^{n-q-1}\right) \\ &= \sum_{q=k-1}^{n-1}\binom{n}{q+1}(q+1)x^q(1-x)^{n-q-1} - \sum_{q=k}^{n-1}\binom{n}{q}(n-q)x^q(1-x)^{n-q-1} \\ &= \binom{n}{k}kx^{k-1}(1-x)^{n-k} + \sum_{q=k}^{n-1}0.\end{aligned}$$
À la dernière égalité, on utilise la relation $\binom{n}{q+1}(q+1) = \binom{n}{q}(n-q)$, qui s'obtient par exemple à partir de l'expression des coefficients binomiaux à partir de factorielles.