



PRÉPASIMPLE

prepasimple.fr

ESCP — 2024

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Première partie

On conserve les notations du sujet. Pour $0 < \alpha < 1$, les expressions de rang $\lfloor \alpha n \rfloor$ sont utilisées à partir du moment où ce rang vaut au moins 1. En effet, $\lfloor \alpha n \rfloor \geq 1$ dès que $n \geq 1/\alpha$, et $\lfloor \alpha n \rfloor < n$. Cette précision rectifie un défaut de définition aux premiers indices du sujet original, sans modifier les limites étudiées.

1. (a) Le maximum est inférieur ou égal à x exactement lorsque chacun des X_i l'est. L'indépendance mutuelle donne donc :

$$F_{(n,n)}(x) = P\left(\bigcap_{i=1}^n [X_i \leq x]\right) = \prod_{i=1}^n P(X_i \leq x) = F(x)^n.$$

Puisque X possède une densité, F est continue sur $\mathbb{R}$ et dérivable, de dérivée f , aux points où f est continue. Ainsi F^n est continue et sa dérivée, hors des éventuels points de discontinuité de f , vaut :

$$(F^n)'(x) = nf(x)F(x)^{n-1}.$$

Cette fonction est positive. De plus, en intégrant la dérivée de F^n entre $-R$ et R , puis en faisant tendre R vers $+\infty$, on obtient une intégrale égale à 1, car $F(-R) \rightarrow 0$ et $F(R) \rightarrow 1$. Les éventuelles valeurs aux points de discontinuité ne changent pas l'intégrale. Elle définit donc une densité de $X_{(n,n)}$:

$$f_{(n,n)}(x) = nf(x)F(x)^{n-1}.$$

Pour $n = 1$, cette formule redonne f .

1. (b) Le minimum dépasse x exactement lorsque tous les X_i dépassent x . Par indépendance :

$$P(X_{(n,1)} > x) = \prod_{i=1}^n P(X_i > x) = (1 - F(x))^n.$$

Sa fonction de répartition est donc :

$$F_{(n,1)}(x) = 1 - (1 - F(x))^n.$$

En dérivant aux points de continuité de f , on obtient :

$$f_{(n,1)}(x) = nf(x)(1 - F(x))^{n-1}.$$

Comme en 1. (a), la fonction de répartition est continue, cette dérivée est positive et son intégrale sur $\mathbb{R}$ vaut 1. Il s'agit donc d'une densité du minimum. Le cas $n = 1$ redonne également f .

2. (a) Par définition, $Y_{i,x}$ ne prend que les valeurs 0 et 1, et :

$$P(Y_{i,x} = 1) = P(X_i \leq x) = F(x), \quad P(Y_{i,x} = 0) = 1 - F(x).$$

Ainsi $Y_{i,x}$ suit $\mathcal{B}(F(x))$. Si $F(x) = 0$ ou $F(x) = 1$, cette loi est une loi certaine, respectivement en 0 ou en 1.

2. (b) Pour chaque issue ω , la somme

$$\sum_{i=1}^n Y_{i,x}(\omega)$$

compte le nombre de valeurs parmi $X_1(\omega), \dots, X_n(\omega)$ qui sont inférieures ou égales à x .

La k -ième valeur de la liste ordonnée est inférieure ou égale à x si et seulement si cette liste contient au moins k valeurs inférieures ou égales à x . Cela reste vrai en présence d'ex æquo.

Ainsi, issue par issue :

$$[X_{(n,k)} \leq x] = \left[\sum_{i=1}^n Y_{i,x} \geq k \right].$$

2. (c) Méthode 1 — Loi binomiale

Pour x fixé, les variables $Y_{i,x}$ sont des fonctions respectives des variables indépendantes X_i . Le lemme des coalitions garantit leur indépendance mutuelle. Elles suivent toutes

$\mathcal{B}(F(x))$, donc :

$$N_{n,x} = \sum_{i=1}^n Y_{i,x} \sim \mathcal{B}(n, F(x)).$$

L'égalité d'événements précédente donne :

$$F_{(n,k)}(x) = P(N_{n,x} \geq k) = \sum_{\ell=k}^n \binom{n}{\ell} F(x)^\ell (1 - F(x))^{n-\ell}.$$

Méthode 2 — Dénombrement des choix d'indices

Pour obtenir exactement ℓ valeurs inférieures ou égales à x , on choisit les ℓ indices correspondants : il y a $\binom{n}{\ell}$ choix. Pour un choix fixé, l'indépendance donne la probabilité $F(x)^\ell (1 - F(x))^{n-\ell}$. Les événements associés aux différents choix sont disjoints, d'où :

$$P(N_{n,x} = \ell) = \binom{n}{\ell} F(x)^\ell (1 - F(x))^{n-\ell}.$$

En additionnant pour $\ell = k, \dots, n$, on retrouve la formule annoncée. La première méthode reconnaît une loi usuelle ; la seconde explique directement l'origine du coefficient binomial.

Aux valeurs $F(x) = 0$ et $F(x) = 1$, on peut lire la formule comme une évaluation polynomiale, avec les puissances d'exposant nul égales à 1. Elle vaut respectivement 0 et 1.

3. (a) Pour $1 \leq \ell \leq n - 1$, la dérivation du produit donne :

$$b'_{n,\ell}(z) = \binom{n}{\ell} \ell z^{\ell-1} (1 - z)^{n-\ell} - \binom{n}{\ell} (n - \ell) z^\ell (1 - z)^{n-\ell-1}.$$

Les identités, obtenues avec les factorielles, sont :

$$\ell \binom{n}{\ell} = n \binom{n-1}{\ell-1}, \quad (n-\ell) \binom{n}{\ell} = n \binom{n-1}{\ell}.$$

En remplaçant les coefficients :

$$b'_{n,\ell}(z) = n(b_{n-1,\ell-1}(z) - b_{n-1,\ell}(z)).$$

Il reste à traiter $\ell = n$, afin de ne pas écrire une puissance d'exposant négatif multipliée par zéro. Dans ce cas $b_{n,n}(z) = z^n$, donc :

$$b'_{n,n}(z) = nz^{n-1} = n(b_{n-1,n-1}(z) - b_{n-1,n}(z)),$$

puisque $b_{n-1,n}$ est nulle par convention. La formule vaut donc pour tous les indices demandés, y compris $n = 1$.

3. (b) Il s'agit d'une somme finie, que l'on peut dériver terme à terme. Grâce à 3. (a) :

$$a'_{n,k}(z) = n \sum_{\ell=k}^n (b_{n-1,\ell-1}(z) - b_{n-1,\ell}(z)).$$

Tous les termes intermédiaires se compensent. Il reste :

$$a'_{n,k}(z) = n(b_{n-1,k-1}(z) - b_{n-1,n}(z)) = nb_{n-1,k-1}(z).$$

Ainsi les indices cherchés sont $m = n - 1$ et $\ell = k - 1$, et :

$$a'_{n,k}(z) = n \binom{n-1}{k-1} z^{k-1} (1-z)^{n-k}.$$

3. (c) D'après 2. (c), $F_{(n,k)} = a_{n,k} \circ F$. Cette fonction est continue sur $\mathbb{R}$. Aux points où f est continue, la dérivation d'une composée donne :

$$F'_{(n,k)}(x) = a'_{n,k}(F(x))f(x) = n \binom{n-1}{k-1} f(x) F(x)^{k-1} (1-F(x))^{n-k}.$$

Cette expression est positive, puisque $f \geq 0$ et $0 \leq F \leq 1$. La fonction de répartition tend vers 0 en $-\infty$ et vers 1 en $+\infty$: en effet $a_{n,k}(0) = 0$ et $a_{n,k}(1) = 1$ pour $1 \leq k \leq n$.

L'intégration de sa dérivée, avec passage aux limites aux extrémités, donne donc une intégrale égale à 1. La fonction $f_{(n,k)}$ annoncée est bien une densité de $X_{(n,k)}$.

4. (a) Posons $u = F(x)$ et

$$M_n = \frac{1}{n} \sum_{i=1}^n Y_{i,x}.$$

Comme $\alpha - u > \varepsilon$, le réel $d = \alpha - u - \varepsilon$ est strictement positif. Pour n assez grand pour que $1/n \leq d$:

$$\frac{\lfloor \alpha n \rfloor}{n} \geq \alpha - \frac{1}{n} \geq u + \varepsilon.$$

Sur l'événement

$$[\sum Y_{i,x} \geq \lfloor \alpha n \rfloor],$$

on a donc $M_n - u \geq \varepsilon$, et a fortiori $|M_n - u| \geq \varepsilon$. L'inclusion demandée est établie.

Méthode 1 — Loi faible des grands nombres

Les $Y_{i,x}$ sont indépendantes, de même loi de Bernoulli, d'espérance u et de variance $u(1-u)$. La loi faible des grands nombres donne $M_n \rightarrow u$ en probabilité. En utilisant 2. (b) et l'inclusion :

$$0 \leq F_{(n, \lfloor \alpha n \rfloor)}(x) \leq P(|M_n - u| \geq \varepsilon) \rightarrow 0.$$

Ainsi $F_{(n, \lfloor \alpha n \rfloor)}(x) \rightarrow 0$.

Méthode 2 — Majoration quantitative par Bienaymé-Tchebychev

Par linéarité et indépendance, $E(M_n) = u$ et $V(M_n) = u(1-u)/n$. L'inégalité de

Bienaymé-Tchebychev donne :

$$F_{(n, \lfloor \alpha n \rfloor)}(x) \leq \frac{u(1-u)}{n\varepsilon^2} \rightarrow 0.$$

Cette seconde méthode prouve la même limite et fournit une borne explicite de l'erreur.

4. (b) Avec les mêmes notations, sur $[|M_n - u| < \varepsilon]$:

$$M_n > u - \varepsilon > \alpha \geq \frac{\lfloor \alpha n \rfloor}{n}.$$

Donc

$$\sum_{i=1}^n Y_{i,x} \geq \lfloor \alpha n \rfloor.$$

L'inclusion d'événements est valable pour tout $n \geq 1$, même lorsque le seuil entier vaut 0.

Méthode 1 — Loi faible des grands nombres

La convergence $M_n \rightarrow u$ en probabilité implique $P(|M_n - u| < \varepsilon) \rightarrow 1$. Pour les indices où le rang est défini :

$$1 \geq F_{(n, \lfloor \alpha n \rfloor)}(x) \geq P(|M_n - u| < \varepsilon) \rightarrow 1.$$

$$\text{Ainsi } F_{(n, \lfloor \alpha n \rfloor)}(x) \rightarrow 1.$$

Méthode 2 — Majoration quantitative du complémentaire

L'inclusion obtenue donne :

$$0 \leq 1 - F_{(n, \lfloor \alpha n \rfloor)}(x) \leq P(|M_n - u| \geq \varepsilon) \leq \frac{u(1-u)}{n\varepsilon^2}.$$

La borne de droite tend vers 0, d'où la même conclusion. Comme en 4. (a), cette méthode précise l'écart à la limite.

5. (a) Sur I , F est de classe $\mathcal{C}^1$ et $F' = f > 0$. Elle est donc continue et strictement croissante.

On peut aussi vérifier directement la stricte croissance : si $a < s < t < b$, alors f est continue et strictement positive sur $[s, t]$, et :

$$F(t) - F(s) = \int_s^t f(v) dv > 0.$$

Comme f est nulle hors de I et son intégrale totale vaut 1 :

$$\lim_{x \rightarrow a^+} F(x) = 0, \quad \lim_{x \rightarrow b^-} F(x) = 1,$$

avec les interprétations usuelles aux extrémités infinies. Le théorème de la bijection donne donc une bijection de I sur $]0, 1[$. Sa réciproque F^{-1} est continue sur $]0, 1[$.

5. (b) Posons $c = F^{-1}(\alpha) \in I$. Une variable certaine égale à c a pour fonction de répartition :

$$G(x) = \begin{cases} 0 & \text{si } x < c, \\ 1 & \text{si } x \geq c. \end{cases}$$

Il faut établir la convergence des fonctions de répartition en tout point de continuité de G , c'est-à-dire pour tout $x \neq c$.

Si $x < c$ et $x \in I$, la stricte croissance donne $0 < F(x) < \alpha$, donc 4. (a) donne $F_{(n, [\alpha n])}(x) \rightarrow 0$. Si $x < c$ et $x \notin I$, alors nécessairement $x \leq a$ et $F(x) = 0$, d'où $F_{(n, [\alpha n])}(x) = 0$.

Si $x > c$ et $x \in I$, alors $\alpha < F(x) < 1$ et 4. (b) donne $F_{(n, [\alpha n])}(x) \rightarrow 1$. Si $x > c$ et $x \notin I$, alors $x \geq b$ et $F(x) = 1$, d'où $F_{(n, [\alpha n])}(x) = 1$.

La convergence est donc vérifiée pour tous les $x \neq c$. Ainsi :

$$X_{(n, [\alpha n])} \xrightarrow{\mathcal{L}} c.$$

Aucune condition n'est demandée en $x = c$, puisque G y est discontinue.

5. (c) Nous prouvons directement le passage de la convergence en loi vers une constante à la convergence en probabilité, sans supposer un théorème supplémentaire.

Posons $Q_n = X_{(n, [\alpha n])}$. Pour $\varepsilon > 0$:

$$[|Q_n - c| \geq \varepsilon] \subset [Q_n \leq c - \varepsilon] \cup [Q_n > c + \varepsilon/2].$$

En effet, si $Q_n \geq c + \varepsilon$, il dépasse strictement $c + \varepsilon/2$. Par sous-additivité :

$$P(|Q_n - c| \geq \varepsilon) \leq F_{Q_n}(c - \varepsilon) + 1 - F_{Q_n}(c + \varepsilon/2).$$

Les deux points considérés sont des points de continuité de la fonction de répartition de la constante c . Par 5. (b), le premier terme tend vers 0 et le second aussi. Ainsi :

$$\forall \varepsilon > 0, \quad P(|Q_n - c| \geq \varepsilon) \rightarrow 0,$$

ce qui est la convergence en probabilité vers $F^{-1}(\alpha)$.

6. Voici un tri par insertion en place :

Python

```
import numpy as np

def tri(T):
    s = np.shape(T)[0]
    for k in range(1, s):
        x = T[k]
        j = k - 1
        while j >= 0 and T[j] > x:
            T[j + 1] = T[j]
            j = j - 1
        T[j + 1] = x
```

Au début de l'étape k , la portion d'indices 0 à $k - 1$ est triée. On mémorise la valeur

$x = T[k]$ pour éviter de la perdre lors des décalages. On décale vers la droite tous les éléments de cette portion qui sont strictement supérieurs à x .

La variable entière j diminue à chaque passage et reste au moins égale à -1 ; la boucle termine. À sa sortie, soit $j = -1$, soit $T[j] \leq x$. L'affectation $T[j + 1] = x$ insère alors x à la bonne place. La portion d'indices 0 à k est triée et les indices au-delà de k restent inchangés. La propriété est initialement vraie pour la portion réduite à l'indice 0 , donc une récurrence prouve le tri final.

Dans la condition `and`, Python évalue d'abord $j \geq 0$; si cette condition est fausse, il n'évalue pas $T[j] > x$. Cela évite d'utiliser l'indice négatif -1 comme un élément à comparer.

Le tableau fourni est modifié directement et la fonction ne renvoie pas de nouveau tableau. La borne supérieure de `range(1, s)` est exclue : les étapes sont bien $1, \dots, s - 1$. Pour un tableau vide ou de longueur 1 , aucune étape n'est nécessaire.

Code exécuté : l'exemple fourni aboutit à $[15, 24, 28, 31, 35, 38]$. Des tableaux déjà triés, décroissants, vides et contenant des valeurs répétées ont également été vérifiés.

7. (a) Puisque les indices Python commencent à 0 , la k -ième valeur triée est à l'indice $k - 1$:

Python

```
import numpy as np

def tri(T):
    s = np.shape(T)[0]
    for k in range(1, s):
        x = T[k]
        j = k - 1
        while j >= 0 and T[j] > x:
            T[j + 1] = T[j]
            j = j - 1
        T[j + 1] = x
```

```
def X(n, k):
    res = ech(n)
    tri(res)
    return res[k - 1]
```

La fonction ech est celle supposée définie dans l'énoncé ; une définition complète est donnée en 7. (d). L'hypothèse $1 \leq k \leq n$ garantit que l'indice appartient à $\{0, \dots, n - 1\}$. Le tri conserve les valeurs de l'échantillon et les ordonne, donc la fonction renvoie bien une réalisation de $X_{(n,k)}$.

7. (b) Les abscisses sont les 20 nombres α_j régulièrement espacés de 0,01 à 0,99. À chaque abscisse correspond une réalisation de :

$$X_{(2000, \lfloor 2000\alpha_j \rfloor)}.$$

Chaque appel à X appelle ech : le programme tire un nouvel échantillon pour chaque point. Il ne trace donc pas les statistiques d'ordre d'un unique échantillon commun.

Tous les rangs sont valides : le plus petit vaut 20 et le plus grand 1980. Sous les hypothèses de 5, lorsque n est grand, ces valeurs sont proches en probabilité de $F^{-1}(\alpha_j)$. Le graphique donne ainsi une approximation expérimentale de la fonction quantile $\alpha \mapsto F^{-1}(\alpha)$.

7. (c) Question adaptée au programme actuel : la demande d'identification certaine est remplacée par la recherche d'une loi usuelle compatible et par une distinction entre observation et preuve.

Les points se placent près de la droite $y = 2 + 4\alpha$. Pour une loi uniforme sur $[2, 6]$, la fonction de répartition est :

$$F(x) = \begin{cases} 0 & \text{si } x \leq 2, \\ \frac{x-2}{4} & \text{si } 2 < x < 6, \\ 1 & \text{si } x \geq 6, \end{cases}$$

et, pour $0 < \alpha < 1$, $F^{-1}(\alpha) = 2 + 4\alpha$. La loi $\mathcal{U}([2, 6])$ est donc la loi usuelle compatible attendue.

La densité peut être choisie égale à $1/4$ sur $]2, 6[$ et nulle ailleurs, ce qui satisfait les hypothèses de 5. Les variations des points autour de la droite sont compatibles avec l'aléa des statistiques d'ordre.

Un nombre fini de réalisations ne détermine toutefois pas une loi de manière unique. Des lois proches de cette uniforme peuvent également produire un tel nuage. Le graphique suggère cette loi ; il ne prouve pas que ech l'utilise effectivement.

7. (d) Si U suit la loi uniforme sur $[0, 1]$, alors $2 + 4U$ suit la loi uniforme sur $[2, 6]$. La commande `rd.random(n)` fournit n valeurs uniformes indépendantes. Voici le script complet :

Python

```
import numpy as np

import numpy.random as rd

import matplotlib.pyplot as plt


def tri(T):
    s = np.shape(T)[0]
    for k in range(1, s):
        x = T[k]
        j = k - 1
        while j >= 0 and T[j] > x:
            T[j + 1] = T[j]
            j = j - 1
        T[j + 1] = x


def ech(n):
    return 2 + 4 * rd.random(n)


def X(n, k):
    res = ech(n)
```

```

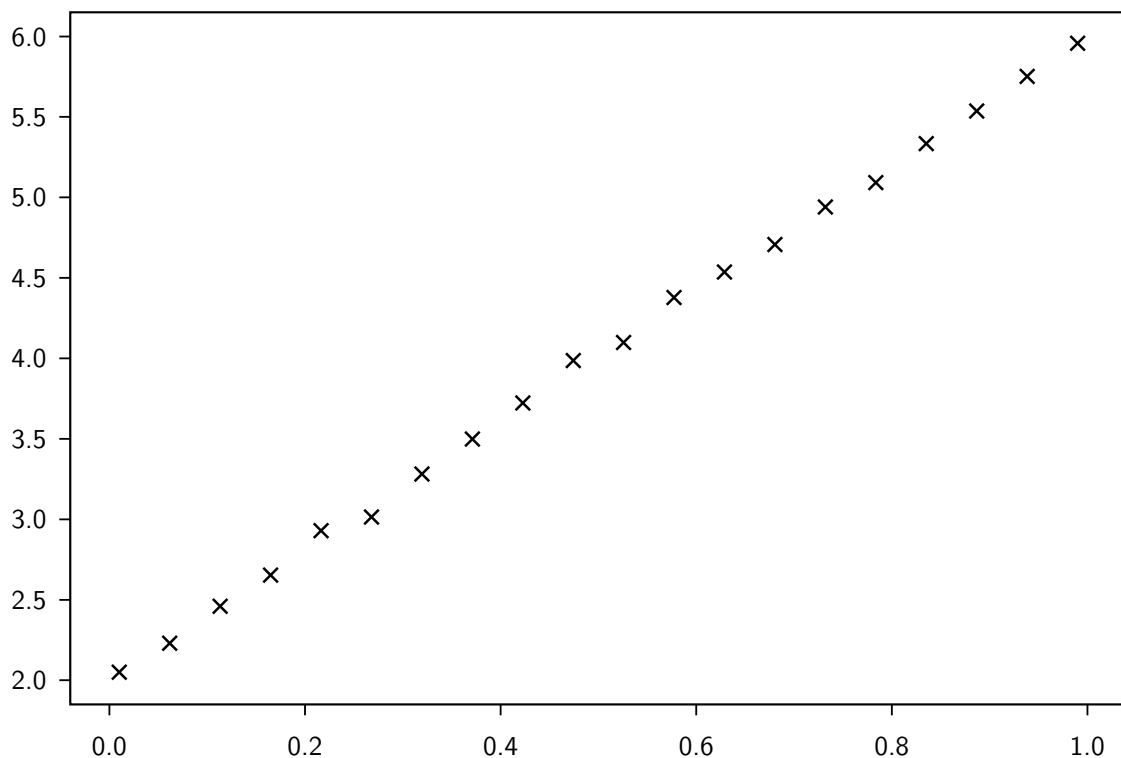
    tri(res)

    return res[k - 1]

n = 2000
na = 20
A = np.linspace(0.01, 0.99, na)
B = np.zeros(na)
for k in range(na):
    a = A[k]
    B[k] = X(n, int(np.floor(a*n)))
plt.plot(A, B, "xk")
plt.show()

```

$B[k]$ est une valeur simulée de la statistique d'ordre au rang entier demandé. Le style "xk" représente chaque point par une croix noire. Le script a été exécuté avec les paramètres de l'énoncé; les valeurs simulées sont proches de $2 + 4A[k]$. Les points changent d'une exécution à l'autre : on reproduit un comportement, pas les réalisations exactes du document.



La simulation illustre le résultat théorique de 5. (c), sans constituer une démonstration.

Deuxième partie

Dans cette partie, $\lambda > 0$ et les X_i sont indépendantes de loi exponentielle de taux λ . On choisit :

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0, \\ 0 & \text{si } x \leq 0, \end{cases} \quad F(x) = \begin{cases} 1 - e^{-\lambda x} & \text{si } x > 0, \\ 0 & \text{si } x \leq 0. \end{cases}$$

8. (a) On raisonne par récurrence, comme demandé.

Initialisation : $S_1 = T_1$ suit $\mathcal{E}(n\lambda)$. Sa densité vaut $n\lambda e^{-n\lambda x}$ pour $x > 0$, et 0 pour $x \leq 0$. C'est bien $g_{n,1}$ puisque $\binom{n-1}{0} = 1$ et $(1 - e^{-\lambda x})^0 = 1$.

Hérédité : supposons le résultat vrai à un rang $k < n$. La variable S_k , fonction de $(T_1, \dots, T_k)$, est indépendante de T_{k+1} par le lemme des coalitions. Cette dernière suit $\mathcal{E}((n-k)\lambda)$.

La somme $S_{k+1} = S_k + T_{k+1}$ admet une densité obtenue par convolution. Les deux variables étant positives, cette densité est nulle pour $x \leq 0$. Pour $x > 0$:

$$g_{n,k+1}(x) = \int_0^x g_{n,k}(t)(n-k)\lambda e^{-(n-k)\lambda(x-t)} dt.$$

La fonction intégrée est continue sur le segment considéré, en choisissant ses valeurs aux bornes par prolongement ; l'intégrale existe. En remplaçant $g_{n,k}$ et en regroupant les exponentielles :

$$g_{n,k+1}(x) = n \binom{n-1}{k-1} (n-k)\lambda^2 e^{-(n-k)\lambda x} \int_0^x e^{-\lambda t} (1 - e^{-\lambda t})^{k-1} dt.$$

La dérivée de $(1 - e^{-\lambda t})^k$ vaut $k\lambda e^{-\lambda t} (1 - e^{-\lambda t})^{k-1}$. On obtient directement la primitive et donc :

$$\int_0^x e^{-\lambda t} (1 - e^{-\lambda t})^{k-1} dt = \frac{(1 - e^{-\lambda x})^k}{k\lambda}.$$

Ainsi :

$$g_{n,k+1}(x) = n \binom{n-1}{k-1} \frac{n-k}{k} \lambda e^{-(n-k)\lambda x} (1 - e^{-\lambda x})^k.$$

Or $\binom{n-1}{k-1} \frac{n-k}{k} = \binom{n-1}{k}$. On retrouve exactement $g_{n,k+1}$, ce qui achève l'hérédité.

Par récurrence, la formule vaut pour tous les $k \in \{1, \dots, n\}$. Lorsque $n = 1$, seule l'initialisation est nécessaire.

8. (b) Pour $x > 0$, la formule de 3. (c) donne :

$$f_{(n,k)}(x) = n \binom{n-1}{k-1} \lambda e^{-\lambda x} (1 - e^{-\lambda x})^{k-1} (e^{-\lambda x})^{n-k},$$

soit :

$$f_{(n,k)}(x) = n \binom{n-1}{k-1} \lambda e^{-\lambda(n-k+1)x} (1 - e^{-\lambda x})^{k-1} = g_{n,k}(x).$$

Pour $x \leq 0$, les deux densités sont nulles. Elles sont donc égales, ce qui prouve :

$X_{(n,k)}$ et S_k ont la même loi.

Cette égalité de lois ne signifie pas que les variables sont égales sur l'espace probabilisé.

9. Chaque T_i possède des moments d'ordre 1 et 2, puisque c'est une exponentielle de taux strictement positif $(n - i + 1)\lambda$. On a :

$$E(T_i) = \frac{1}{(n - i + 1)\lambda}, \quad V(T_i) = \frac{1}{(n - i + 1)^2 \lambda^2}.$$

La somme finie S_k possède une espérance et un second moment : par exemple,

$$S_k^2 \leq k \sum_{i=1}^k T_i^2,$$

d'après Cauchy-Schwarz. L'existence de la variance est donc justifiée.

Par linéarité de l'espérance :

$$E(S_k) = \sum_{i=1}^k E(T_i) = \frac{1}{\lambda} \sum_{i=1}^k \frac{1}{n-i+1}.$$

Par indépendance, les covariances entre deux T_i distinctes sont nulles, donc :

$$V(S_k) = \sum_{i=1}^k V(T_i) = \frac{1}{\lambda^2} \sum_{i=1}^k \frac{1}{(n-i+1)^2}.$$

L'égalité de lois établie en 8. (b) transmet à $X_{(n,k)}$ l'existence et la valeur de ces moments. Les formules demandées sont donc établies.

10. (a) Pour $j \geq 1$ et $x \in [j, j+1]$, les dénominateurs sont strictement positifs et :

$$\frac{1}{j} - \frac{1}{x} = \frac{x-j}{jx} \geq 0.$$

De plus $0 \leq x-j \leq 1$ et $x \leq j+1 \leq 2j$. Ainsi :

$$x(x-j) \leq x \leq 2j.$$

En divisant par $jx^2 > 0$, on obtient :

$$\frac{x-j}{jx} \leq \frac{2}{x^2}.$$

Cela établit les deux inégalités, y compris aux bornes de l'intervalle.

10. (b) Intégrons l'inégalité de 10. (a) sur le segment $[j, j+1]$:

$$0 \leq \frac{1}{j} - \ln \left(\frac{j+1}{j} \right) \leq 2 \left(\frac{1}{j} - \frac{1}{j+1} \right).$$

Posons $a = n - k + 1 \geq 1$. La formule de 9, après changement d'indice $j = n - i + 1$, donne

$$\lambda E(X_{(n,k)}) = \sum_{j=a}^n 1/j.$$

En sommant les inégalités précédentes de $j = a$ à n , les logarithmes et la borne supérieure se télescopent :

$$0 \leq \sum_{j=a}^n \frac{1}{j} - \ln \left(\frac{n+1}{a} \right) \leq 2 \left(\frac{1}{a} - \frac{1}{n+1} \right).$$

En divisant par $\lambda > 0$ et en remplaçant a , on obtient exactement l'encadrement annoncé. Il reste valable pour $k = 1$ et $k = n$.

10. (c) Méthode 1 — Encadrement de la question précédente

Posons $k_n = \lfloor \alpha n \rfloor$ et $a_n = n - k_n + 1$. Comme $0 \leq \alpha - k_n/n < 1/n$:

$$\frac{k_n}{n} \rightarrow \alpha, \quad \frac{a_n}{n} \rightarrow 1 - \alpha > 0.$$

Par conséquent :

$$\frac{n+1}{a_n} \rightarrow \frac{1}{1-\alpha}, \quad \frac{1}{a_n} - \frac{1}{n+1} \rightarrow 0.$$

Par continuité de $\ln$ sur $]0, +\infty[$ et par l'encadrement de 10. (b), les deux bornes de l'espérance tendent vers :

$$\frac{1}{\lambda} \ln \left(\frac{1}{1-\alpha} \right) = -\frac{1}{\lambda} \ln(1-\alpha).$$

Le théorème d'encadrement donne cette limite pour l'espérance.

Méthode 2 — Sommes de Riemann

Le changement d'indice $r = i - 1$ donne :

$$\lambda E(X_{(n, k_n)}) = \frac{1}{n} \sum_{r=0}^{k_n-1} \frac{1}{1 - r/n}.$$

La fonction $h(t) = 1/(1-t)$ est de classe $\mathcal{C}^1$ sur le segment $[0, \alpha]$, et $0 < h'(t) \leq 1/(1-\alpha)^2$. Sur chaque intervalle $[r/n, (r+1)/n]$, l'inégalité des accroissements finis donne $|h(t) - h(r/n)| \leq 1/(n(1-\alpha)^2)$. L'écart entre la somme et $\int_0^{k_n/n} h(t) dt$ est donc au plus $k_n/(n^2(1-\alpha)^2) \leq \alpha/(n(1-\alpha)^2)$, qui tend vers 0.

Enfin $0 \leq \alpha - k_n/n < 1/n$ et $h \leq 1/(1-\alpha)$ sur ce segment, donc :

$$0 \leq \int_{k_n/n}^{\alpha} h(t) dt \leq \frac{1}{n(1-\alpha)} \rightarrow 0.$$

Ainsi la somme converge vers $\int_0^{\alpha} \frac{dt}{1-t} = -\ln(1-\alpha)$. Après division par λ , on retrouve la limite demandée. La première méthode exploite une borne explicite ; la seconde reconnaît l'aire sous une fonction continue.

11. Méthode 1 — Majoration de chaque terme

Posons encore $k_n = \lfloor \alpha n \rfloor$ et $a_n = n - k_n + 1$. Tous les dénominateurs de la somme de 9 sont au moins a_n , donc :

$$0 \leq V(X_{(n, k_n)}) \leq \frac{k_n}{\lambda^2 a_n^2}.$$

Or :

$$\frac{k_n}{a_n^2} = \frac{1}{n} \frac{k_n/n}{(a_n/n)^2} \rightarrow 0,$$

puisque $a_n/n \rightarrow 1 - \alpha > 0$. Par encadrement, la variance tend vers 0.

Méthode 2 — Comparaison avec une intégrale

Pour n assez grand, $a_n \geq 2$. La fonction $t \mapsto 1/t^2$ est décroissante sur $]0, +\infty[$. Pour tout entier $j \geq a_n$ et tout $t \in [j-1, j]$, $1/j^2 \leq 1/t^2$. En intégrant et en sommant :

$$\sum_{j=a_n}^n \frac{1}{j^2} \leq \int_{a_n-1}^n \frac{dt}{t^2} = \frac{1}{a_n-1} - \frac{1}{n}.$$

Cette borne tend vers 0 puisque $a_n \rightarrow +\infty$. Après division par λ^2 , on obtient encore la limite nulle. Cette comparaison est utile pour contrôler les restes de sommes de puissances.

- 12.** (a) Posons $Q_n = X_{(n, [\alpha n])}$ et $c = -\frac{1}{\lambda} \ln(1 - \alpha)$. Par 9, Q_n possède un second moment, donc la variable positive $(Q_n - c)^2$ possède une espérance finie.

L'événement $[|Q_n - c| \geq \varepsilon]$ est exactement $[(Q_n - c)^2 \geq \varepsilon^2]$. L'inégalité de Markov appliquée à $(Q_n - c)^2$ donne :

$$P(|Q_n - c| \geq \varepsilon) \leq \frac{E((Q_n - c)^2)}{\varepsilon^2}.$$

En remplaçant c , on retrouve l'inégalité demandée.

- 12.** (b) Le développement du carré, avec $E(Q_n - E(Q_n)) = 0$, donne :

$$E((Q_n - c)^2) = V(Q_n) + (E(Q_n) - c)^2.$$

Le premier terme tend vers 0 d'après 11 et le second d'après 10. (c). Pour tout $\varepsilon > 0$, 12.

(a) implique donc :

$$0 \leq P(|Q_n - c| \geq \varepsilon) \leq \frac{V(Q_n) + (E(Q_n) - c)^2}{\varepsilon^2} \rightarrow 0.$$

Ainsi Q_n converge en probabilité vers $-\frac{1}{\lambda} \ln(1 - \alpha)$. Il s'agit d'un calcul particulier de moments et d'une application de Markov, sans recours à une théorie générale d'évaluation des estimateurs.

- 13.** La question 5. (c) permet de retrouver le résultat. Pour une exponentielle, on prend $I =]0, +\infty[$ et la densité continue et strictement positive $f(x) = \lambda e^{-\lambda x}$ sur I , nulle ailleurs.

L'équation $F(x) = \alpha$ s'écrit $1 - e^{-\lambda x} = \alpha$, donc :

$$F^{-1}(\alpha) = -\frac{1}{\lambda} \ln(1 - \alpha).$$

Les hypothèses de 5 sont satisfaites ; 5. (c) donne exactement la convergence obtenue en 12. (b).

14. (a) Fixons $t > 0$ et posons $u = e^{-\lambda t} \in]0, 1[$. Introduisons $I_i = \mathbf{1}_{[X_i > t]}$ et

$$N_{m,t} = \sum_{i=1}^m I_i.$$

Par le lemme des coalitions, les I_i sont indépendantes, de même loi de Bernoulli de paramètre u .

On a $V_{m,t} = N_{m,t}/m + 1/m$. Cette variable est bornée, donc possède tous ses moments. Les formules de la binomiale donnent :

$$E(V_{m,t}) = u + \frac{1}{m} \rightarrow u, \quad V(V_{m,t}) = \frac{u(1-u)}{m} \rightarrow 0.$$

Le critère de convergence en probabilité par l'espérance et la variance donne donc :

$$V_{m,t} \xrightarrow{P} e^{-\lambda t}.$$

14. (b) La correction $+1$ garantit $V_{m,t} \geq 1/m > 0$: le logarithme est toujours défini.

Posons $u = e^{-\lambda t}$ et $\delta = u(1 - e^{-\varepsilon}) > 0$.

Si $|V_{m,t} - u| \leq \delta$, alors :

$$ue^{-\varepsilon} = u - \delta \leq V_{m,t} \leq u + \delta = u(2 - e^{-\varepsilon}).$$

Or $2 - e^{-\varepsilon} \leq e^{\varepsilon}$. Pour le vérifier, écrivons :

$$e^{\varepsilon} + e^{-\varepsilon} - 2 = (e^{\varepsilon/2} - e^{-\varepsilon/2})^2 \geq 0.$$

Donc $ue^{-\varepsilon} \leq V_{m,t} \leq ue^{\varepsilon}$. Par croissance du logarithme :

$$\ln u - \varepsilon \leq \ln V_{m,t} \leq \ln u + \varepsilon.$$

Comme $\ln u = -\lambda t$, on obtient $|\ln V_{m,t} - \lambda t| \leq \varepsilon$. Nous avons ainsi l'inclusion d'événements qui entraîne l'inégalité de probabilités demandée.

Pour conclure avec la convention $\geq$ dans la définition de la convergence en probabilité,

appliquons cette inclusion au seuil $\eta = \varepsilon/2$. Le complémentaire donne :

$$P(|-\ln V_{m,t} - \lambda t| \geq \varepsilon) \leq P(|-\ln V_{m,t} - \lambda t| > \eta) \leq P(|V_{m,t} - u| > u(1 - e^{-\eta})) \rightarrow 0.$$

La dernière limite résulte de 14. (a), avec un seuil strictement positif fixé. Ainsi :

$$-\ln V_{m,t} \xrightarrow{P} \lambda t.$$

Cette preuve par encadrement est nécessaire pour traiter le logarithme au voisinage de la constante positive u sans employer une fonction supposée continue sur tout $\mathbb{R}$.

14. (c) Les deux nombres de dépassements augmentés de 1 sont strictement positifs. En utilisant la définition de $V_{m,t}$, les facteurs m se simplifient dans le quotient :

$$W_m = \frac{\ln V_{m,t_1} - \ln V_{m,t_2}}{t_2 - t_1}.$$

Posons $A_m = -\ln V_{m,t_1}$ et $B_m = -\ln V_{m,t_2}$. Par 14. (b), $A_m \rightarrow \lambda t_1$ et $B_m \rightarrow \lambda t_2$ en probabilité. Or :

$$W_m - \lambda = \frac{(B_m - \lambda t_2) - (A_m - \lambda t_1)}{t_2 - t_1}.$$

Pour $\varepsilon > 0$, l'inégalité triangulaire entraîne :

$$[|W_m - \lambda| \geq \varepsilon] \subset \left[|A_m - \lambda t_1| \geq \frac{\varepsilon(t_2 - t_1)}{2} \right] \cup \left[|B_m - \lambda t_2| \geq \frac{\varepsilon(t_2 - t_1)}{2} \right].$$

Par sous-additivité, sa probabilité est majorée par la somme de deux probabilités tendant vers 0. L'indépendance entre A_m et B_m n'est pas nécessaire, et n'est pas affirmée.

La variable W_m est une fonction de l'échantillon et des seuils connus t_1, t_2 , sans utiliser la valeur inconnue de λ . C'est donc un estimateur de λ , et la convergence $W_m \rightarrow \lambda$ en probabilité prouve qu'il est convergent.

Troisième partie

15. Le résultat admis dans l'énoncé affirme l'existence de variables $T_{n,1}, \dots, T_{n,n}$ mutuellement indépendantes, de lois respectives $\mathcal{E}((n-i+1)\lambda)$, telles que simultanément, pour tous les $k \in \{1, \dots, n\}$:

$$X_{(n,k)} = \sum_{i=1}^k T_{n,i}.$$

La question 8. (b) établissait seulement, pour chaque k , une égalité de lois entre $X_{(n,k)}$ et une somme d'exponentielles indépendantes. Elle ne donnait pas d'égalité entre les variables elles-mêmes et ne déterminait pas la loi conjointe des statistiques d'ordre de l'échantillon.

Le théorème fournit une représentation simultanée sur le même espace probabilisé. En posant $X_{(n,0)} = 0$, on obtient nécessairement :

$$T_{n,i} = X_{(n,i)} - X_{(n,i-1)}.$$

Il affirme donc notamment l'indépendance des espacements successifs et précise leurs lois. C'est une information sur le vecteur entier des statistiques d'ordre, plus forte que l'information sur chaque loi marginale séparément.

Le résultat est admis pour cette question, conformément à l'énoncé. Les questions suivantes en démontrent le cas $n = 2$.

16. Méthode 1 — Décalage d'une loi géométrique

La variable $G = Z + 1$ prend ses valeurs dans $\mathbb{N}^*$ et, pour $j \geq 1$:

$$P(G = j) = P(Z = j - 1) = pq^{j-1} = p(1-p)^{j-1}.$$

Ainsi G suit la loi géométrique de paramètre $p = 1 - q \in]0, 1[$. Cette loi possède une espérance et une variance :

$$E(G) = \frac{1}{p}, \quad V(G) = \frac{1-p}{p^2} = \frac{q}{p^2}.$$

Par transformation affine :

$$E(Z) = E(G) - 1 = \frac{q}{p}, \quad V(Z) = V(G) = \frac{q}{p^2}.$$

Pour $\ell \in \mathbb{N}$, la série géométrique converge puisque $0 < q < 1$, et :

$$P(Z \geq \ell) = \sum_{j=\ell}^{+\infty} pq^j = \frac{pq^\ell}{1-q} = q^\ell.$$

Cela donne notamment $P(Z \geq 0) = 1$.

Méthode 2 — Séries géométriques dérivées

Les formules usuelles des deux séries géométriques dérivées sont valables pour $|q| < 1$:

$$\sum_{j=1}^{+\infty} jq^{j-1} = \frac{1}{(1-q)^2}, \quad \sum_{j=2}^{+\infty} j(j-1)q^{j-2} = \frac{2}{(1-q)^3}.$$

Elles montrent la convergence des séries positives nécessaires aux moments. Ainsi :

$$E(Z) = p \sum_{j=0}^{+\infty} jq^j = \frac{pq}{(1-q)^2} = \frac{q}{p},$$

$$E(Z(Z-1)) = p \sum_{j=0}^{+\infty} j(j-1)q^j = \frac{2pq^2}{(1-q)^3} = \frac{2q^2}{p^2}.$$

Puisque $Z^2 = Z(Z-1) + Z$, le second moment existe et :

$$E(Z^2) = \frac{2q^2}{p^2} + \frac{q}{p} = \frac{q(1+q)}{p^2}.$$

La formule de Huygens donne :

$$V(Z) = \frac{q(1+q)}{p^2} - \frac{q^2}{p^2} = \frac{q}{p^2}.$$

Enfin, la même somme géométrique que dans la première méthode donne

$P(Z \geq \ell) = q^\ell$. La première méthode reconnaît une transformation d'une loi usuelle ; la seconde illustre le calcul des moments à partir de la loi.

17. (a) Pour $\ell \in \mathbb{N}$, le minimum est au moins ℓ exactement lorsque toutes les variables le sont. L'indépendance mutuelle et 16 donnent :

$$P(U_n \geq \ell) = \prod_{i=1}^n P(Z_i \geq \ell) = (q^\ell)^n = q^{n\ell}.$$

Pour $\ell = 0$, cette formule donne bien 1.

17. (b) La variable U_n est à valeurs dans $\mathbb{N}$. Pour tout $\ell \in \mathbb{N}$:

$$P(U_n = \ell) = P(U_n \geq \ell) - P(U_n \geq \ell + 1),$$

soit :

$$P(U_n = \ell) = q^{n\ell} - q^{n(\ell+1)} = (1 - q^n)(q^n)^\ell.$$

Puisque $0 < q^n < 1$, il s'agit exactement de la loi $\mathcal{L}(r)$ avec $r = q^n$.

18. Décomposons l'événement selon la valeur j de Z_1 . Les événements correspondants sont disjoints :

$$[Z_1 \geq k_1] \cap [Z_2 - Z_1 \geq k_2] = \bigcup_{j=k_1}^{+\infty} ([Z_1 = j] \cap [Z_2 \geq j + k_2]).$$

Par additivité dénombrable puis indépendance de Z_1 et Z_2 :

$$P([Z_1 \geq k_1] \cap [Z_2 - Z_1 \geq k_2]) = \sum_{j=k_1}^{+\infty} P(Z_1 = j)P(Z_2 \geq j + k_2).$$

D'après 16, cela vaut :

$$\sum_{j=k_1}^{+\infty} pq^j q^{j+k_2} = pq^{k_2} \sum_{j=k_1}^{+\infty} (q^2)^j.$$

La série géométrique converge, car $0 < q^2 < 1$, et sa somme vaut $(q^2)^{k_1}/(1 - q^2)$. Puisque $1 - q^2 = (1 - q)(1 + q) = p(1 + q)$:

$$P([Z_1 \geq k_1] \cap [Z_2 - Z_1 \geq k_2]) = \frac{q^{2k_1} q^{k_2}}{1 + q}.$$

Le calcul reste valable lorsque k_1 ou k_2 vaut 0.

19. (a) Une exponentielle est presque sûrement positive. Les variables $Z_i = \lfloor mX_i \rfloor$ sont donc à valeurs dans $\mathbb{N}$. Pour tout $j \in \mathbb{N}$:

$$[Z_i = j] = \left[\frac{j}{m} \leq X_i < \frac{j+1}{m} \right].$$

La variable X_i possède une densité, donc les valeurs aux bornes ont une probabilité nulle. En utilisant sa fonction de répartition :

$$P(Z_i = j) = e^{-\lambda j/m} - e^{-\lambda(j+1)/m} = (1 - e^{-\lambda/m})(e^{-\lambda/m})^j.$$

Ainsi Z_1 et Z_2 suivent $\mathcal{L}(q_m)$ avec $q_m = e^{-\lambda/m} \in]0, 1[$. Elles restent indépendantes, car elles sont des fonctions respectives de X_1 et X_2 , indépendantes. Cette dernière observation permet d'appliquer 18 dans la suite.

19. (b) Posons $u = mX_1$, $v = mX_2$, $a = mx$ et $b = my$. Alors $Z_1 = \lfloor u \rfloor$, $Z_2 = \lfloor v \rfloor$, $k_1 = \lfloor a \rfloor$ et $k_2 = \lfloor b \rfloor$.

Si $Z_1 \geq k_1 + 1$, alors $u \geq Z_1 \geq k_1 + 1 > a$, donc $X_1 > x$. Réciproquement, si $X_1 > x$, alors $u > a \geq k_1$, ce qui implique $Z_1 \geq k_1$. Les deux inclusions de la première ligne sont établies.

Pour la seconde ligne, utilisons les inégalités définissant la partie entière :

$$v \geq Z_2, \quad u < Z_1 + 1.$$

Si $Z_2 - Z_1 \geq k_2 + 2$, alors :

$$v - u > Z_2 - Z_1 - 1 \geq k_2 + 1 > b,$$

donc $X_2 - X_1 > y$.

Inversement, $v < Z_2 + 1$ et $u \geq Z_1$, donc $v - u < Z_2 - Z_1 + 1$. Si $v - u > b \geq k_2$, on a :

$$Z_2 - Z_1 + 1 > k_2.$$

Comme $Z_2 - Z_1$ est un entier, cette dernière inégalité entraîne $Z_2 - Z_1 \geq k_2$. Ainsi :

$$[Z_2 - Z_1 \geq k_2 + 2] \subset [X_2 - X_1 > y] \subset [Z_2 - Z_1 \geq k_2].$$

Le décalage de 2 dans la borne inférieure est bien celui du sujet : il garantit l'inégalité stricte malgré les deux parties entières.

19. (c) En intersectant les inclusions de 19. (b), on obtient :

$$[Z_1 \geq k_1 + 1] \cap [Z_2 - Z_1 \geq k_2 + 2] \subset [X_1 > x] \cap [X_2 - X_1 > y],$$

$$[X_1 > x] \cap [X_2 - X_1 > y] \subset [Z_1 \geq k_1] \cap [Z_2 - Z_1 \geq k_2].$$

La croissance des probabilités et 18 donnent :

$$\frac{q_m^{2(k_1+1)} q_m^{k_2+2}}{1 + q_m} \leq P([X_1 > x] \cap [X_2 - X_1 > y]) \leq \frac{q_m^{2k_1} q_m^{k_2}}{1 + q_m}.$$

En remplaçant $q_m = e^{-\lambda/m}$, $k_1 = \lfloor mx \rfloor$ et $k_2 = \lfloor my \rfloor$, on retrouve exactement les deux expressions du sujet. Tous les dénominateurs sont strictement positifs.

19. (d) La définition de la partie entière donne :

$$0 \leq x - \frac{\lfloor mx \rfloor}{m} < \frac{1}{m}, \quad 0 \leq y - \frac{\lfloor my \rfloor}{m} < \frac{1}{m}.$$

Par conséquent, lorsque $m \rightarrow +\infty$:

$$\frac{k_1}{m} \rightarrow x, \quad \frac{k_1 + 1}{m} \rightarrow x, \quad \frac{k_2}{m} \rightarrow y, \quad \frac{k_2 + 2}{m} \rightarrow y, \quad q_m \rightarrow 1.$$

Par continuité de l'exponentielle, les deux bornes de 19. (c) tendent vers $\frac{1}{2}e^{-2\lambda x}e^{-\lambda y}$. La probabilité encadrée ne dépend pas de m ; le théorème d'encadrement prouve donc l'égalité exacte :

$$P([X_1 > x] \cap [X_2 - X_1 > y]) = \frac{1}{2}e^{-2\lambda x}e^{-\lambda y}.$$

Le raisonnement s'applique aussi à $x = 0$ ou $y = 0$.

20. (a) Méthode 1 — Résultat précédent et symétrie

La formule de 19. (d) avec $x = y = 0$ donne :

$$P([X_1 > 0] \cap [X_2 - X_1 > 0]) = \frac{1}{2}.$$

Comme $X_1 > 0$ presque sûrement, on a $P(X_2 > X_1) = 1/2$. Les variables étant indépendantes de même loi, le couple (X_1, X_2) a la même loi que (X_2, X_1) : les probabilités rectangulaires sont les mêmes produits de probabilités marginales. Ainsi $P(X_1 > X_2) = 1/2$.

Les événements $[X_1 < X_2]$, $[X_1 = X_2]$ et $[X_1 > X_2]$ forment une partition, d'où :

$$P(X_1 = X_2) = 1 - \frac{1}{2} - \frac{1}{2} = 0.$$

De plus, chacun des X_i est strictement positif presque sûrement, donc leur minimum $X_{(1)}$ l'est aussi. Enfin :

$$X_{(2)} - X_{(1)} = |X_2 - X_1|,$$

qui est strictement positif hors de l'événement négligeable $[X_1 = X_2]$. Les deux positivités demandées sont établies.

Méthode 2 — Densité de la différence

La variable $-X_1$ est indépendante de X_2 et possède une densité par transformation affine. La convolution donne une densité pour $D = X_2 - X_1$. En écrivant $t = X_1$, elle est :

$$h(d) = \int_{\max(0, -d)}^{+\infty} \lambda e^{-\lambda t} \lambda e^{-\lambda(t+d)} dt.$$

Cette intégrale converge car $2\lambda > 0$. En la calculant :

$$h(d) = \frac{\lambda}{2} e^{-\lambda d} e^{-2\lambda \max(0, -d)} = \frac{\lambda}{2} e^{-\lambda |d|}.$$

Une variable à densité n'a pas de masse en un point, donc $P(D = 0) = 0$, ce qui prouve $P(X_1 = X_2) = 0$. Comme les exponentielles sont positives presque sûrement, leur minimum est positif presque sûrement ; et la différence ordonnée $|D|$ est positive presque sûrement. On retrouve toutes les conclusions.

La première méthode exploite la construction du problème ; la seconde montre l'intérêt de la convolution pour exclure une égalité entre variables indépendantes à densité.

20. (b) Pour $x, y > 0$, notons $E = [X_{(1)} > x] \cap [X_{(2)} - X_{(1)} > y]$.

Sur $[X_1 < X_2]$, le minimum est X_1 et l'écart est $X_2 - X_1$: la portion de E correspond donc à :

$$E_{12} = [X_1 > x] \cap [X_2 - X_1 > y].$$

Sur $[X_2 < X_1]$, elle correspond à :

$$E_{21} = [X_2 > x] \cap [X_1 - X_2 > y].$$

L'événement E ne peut contenir d'égalité $X_1 = X_2$, puisque l'écart y est strictement positif. Ainsi $E = E_{12} \cup E_{21}$, et cette union est disjointe.

Par symétrie du couple indépendant de même loi, $P(E_{12}) = P(E_{21})$. Par additivité :

$$P(E) = 2P(E_{12}) = 2P([X_1 > x] \cap [X_2 - X_1 > y]).$$

La même décomposition reste valable pour $x = 0$ ou $y = 0$. Nous pourrions donc employer l'égalité aux bornes dans la question suivante.

20. (c) Posons $T_1 = X_{(1)}$ et $T_2 = X_{(2)} - X_{(1)}$. D'après 19. (d) et 20. (b), pour $x, y \geq 0$:

$$P(T_1 > x, T_2 > y) = e^{-2\lambda x} e^{-\lambda y}.$$

Par 20. (a), T_1 et T_2 sont strictement positives presque sûrement. On peut donc poser $y = 0$ puis $x = 0$ dans cette égalité pour obtenir :

$$P(T_1 > x) = e^{-2\lambda x} \quad (x \geq 0), \quad P(T_2 > y) = e^{-\lambda y} \quad (y \geq 0).$$

Pour des seuils négatifs, ces probabilités valent 1. Les fonctions de répartition sont donc exactement celles de $\mathcal{E}(2\lambda)$ et $\mathcal{E}(\lambda)$.

Pour démontrer l'indépendance par les fonctions de répartition, appliquons la formule du crible aux événements $[T_1 > x]$ et $[T_2 > y]$. Pour $x, y \geq 0$:

$$P(T_1 \leq x, T_2 \leq y) = 1 - e^{-2\lambda x} - e^{-\lambda y} + e^{-2\lambda x} e^{-\lambda y},$$

soit :

$$P(T_1 \leq x, T_2 \leq y) = (1 - e^{-2\lambda x})(1 - e^{-\lambda y}) = F_{T_1}(x)F_{T_2}(y).$$

Si $x < 0$ ou $y < 0$, le membre de gauche et le produit de droite sont tous deux nuls, par positivité presque sûre. La factorisation est donc vraie pour tous les réels x, y . La caractérisation de l'indépendance par la fonction de répartition conjointe prouve que T_1 et T_2 sont indépendantes.

Par définition, on a simultanément :

$$X_{(2,1)} = T_1, \quad X_{(2,2)} = T_1 + T_2.$$

Avec $T_1 \sim \mathcal{E}(2\lambda)$ et $T_2 \sim \mathcal{E}(\lambda)$ indépendantes, cela démontre précisément le théorème annoncé en 15 pour $n = 2$.

Relevé des adaptations

- Questions 4 à 5 et 10 à 13 : précision commune selon laquelle les statistiques de rang $\lfloor \alpha n \rfloor$ sont considérées pour n suffisamment grand pour que ce rang soit au moins 1. Le rang 0, possible aux premiers indices dans le PDF, n'est pas défini. Cette précision ne change aucune limite.
- Question 7. (b) : la figure est retranscrite en SVG dans l'énoncé à partir de l'original, avec les valeurs mesurées. Les 20 abscisses sont celles définies exactement par le code ; les ordonnées reproduisent les positions mesurées des croix avec la précision du document.
- Question 7. (c) : remplacement de l'identification certaine de la loi à partir d'un nuage simulé par la proposition d'une loi usuelle compatible, accompagnée d'une explication de la limite de cette observation. Un nuage fini ne détermine pas une loi unique.
- Question 12 : explicitation de la fixation de $\alpha \in]0, 1[$, déjà utilisée dans les questions précédentes et dans les expressions de cette question.