

## Correction détaillée

BCPST – 0426 – Mathématiques

Méthodes de calcul et raisonnement

### Partie I : Approximation de la loi de Poisson

Dans cette partie,  $\lambda > 0$  est fixé. Pour chaque bactérie, l'événement "devenir résistante et survivre" a pour probabilité  $\lambda/N$ , et les bactéries sont supposées indépendantes.

**Question 1.** La variable  $X_N$  compte le nombre de succès parmi  $N$  expériences indépendantes, chacune de probabilité de succès  $\lambda/N$ . Ainsi, pour  $N$  assez grand pour que  $\lambda/N \leq 1$ ,

$$X_N \sim \mathcal{B}\left(N, \frac{\lambda}{N}\right).$$

Donc, pour  $k \in \{0, \dots, N\}$ ,

$$\mathbb{P}(X_N = k) = \binom{N}{k} \left(\frac{\lambda}{N}\right)^k \left(1 - \frac{\lambda}{N}\right)^{N-k}.$$

**Question 2.** On écrit

$$\left(1 - \frac{\lambda}{N}\right)^{N-k} = \left(1 - \frac{\lambda}{N}\right)^N \left(1 - \frac{\lambda}{N}\right)^{-k}.$$

Or

$$\left(1 - \frac{\lambda}{N}\right)^N \xrightarrow{N \rightarrow +\infty} e^{-\lambda}, \quad \left(1 - \frac{\lambda}{N}\right)^{-k} \xrightarrow{N \rightarrow +\infty} 1.$$

Donc

$$\boxed{\left(1 - \frac{\lambda}{N}\right)^{N-k} \xrightarrow{N \rightarrow +\infty} e^{-\lambda}.}$$

**Question 3.** Pour  $k \in \mathbb{N}$  fixé, et pour  $N \geq k$ ,

$$\mathbb{P}(X_N = k) = \binom{N}{k} \left(\frac{\lambda}{N}\right)^k \left(1 - \frac{\lambda}{N}\right)^{N-k}.$$

Or

$$\binom{N}{k} \left(\frac{\lambda}{N}\right)^k = \frac{N(N-1)\cdots(N-k+1)}{k!} \frac{\lambda^k}{N^k} = \frac{\lambda^k}{k!} \prod_{j=0}^{k-1} \left(1 - \frac{j}{N}\right).$$

Comme  $k$  est fixé,

$$\prod_{j=0}^{k-1} \left(1 - \frac{j}{N}\right) \xrightarrow{N \rightarrow +\infty} 1.$$

En utilisant la question précédente,

$$\boxed{\mathbb{P}(X_N = k) \xrightarrow{N \rightarrow +\infty} e^{-\lambda} \frac{\lambda^k}{k!}.}$$

On retrouve donc la loi de Poisson de paramètre  $\lambda$ .

**Question 4.** Si  $X$  suit une loi de Poisson de paramètre  $\lambda$ , alors

$$\mathbb{E}(X) = \lambda \quad \text{et} \quad \mathbb{V}(X) = \lambda.$$

Comme  $\lambda > 0$ ,

$$\boxed{\frac{\mathbb{V}(X)}{\mathbb{E}(X)} = 1.}$$

## Partie II : Inégalités de concentration

---

On considère une variable aléatoire réelle positive  $X$ , admettant une espérance.

**Question 5.** Pour  $a > 0$ , on a l'inégalité d'indicatrices

$$a \mathbf{1}_{\{X > a\}} \leq X.$$

En prenant l'espérance des deux membres,

$$a \mathbb{P}(X > a) \leq \mathbb{E}(X).$$

Donc

$$\boxed{\mathbb{P}(X > a) \leq \frac{\mathbb{E}(X)}{a}.}$$

C'est l'inégalité de Markov.

**Question 6.** Soit  $t > 0$ . On applique l'inégalité de Markov à la variable positive

$$Z_t = e^{tX}.$$

Comme la fonction exponentielle est croissante,

$$\{X > a\} = \{e^{tX} > e^{ta}\}.$$

Ainsi,

$$\mathbb{P}(X > a) = \mathbb{P}(Z_t > e^{ta}) \leq \frac{\mathbb{E}(Z_t)}{e^{ta}}.$$

Or  $\psi(t) = \ln(\mathbb{E}(Z_t))$ , donc  $\mathbb{E}(Z_t) = e^{\psi(t)}$ . Finalement,

$$\boxed{\mathbb{P}(X > a) \leq e^{\psi(t) - ta}.}$$

**Question 7.a.** On suppose que  $X \sim \mathcal{P}(\lambda)$ . Pour tout  $t > 0$ ,

$$\mathbb{E}(e^{tX}) = \sum_{k=0}^{+\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{(\lambda e^t)^k}{k!} = e^{-\lambda} e^{\lambda e^t}.$$

Donc

$$\mathbb{E}(e^{tX}) = e^{\lambda(e^t - 1)}.$$

Cette quantité est finie pour tout  $t > 0$ , donc  $\psi$  est bien définie sur  $\mathbb{R}_+^*$ , et

$$\boxed{\psi(t) = \lambda(e^t - 1).}$$

**Question 7.b.** D'après la question 6, pour tout  $t > 0$ ,

$$\mathbb{P}(X > a) \leq \exp(\lambda(e^t - 1) - ta).$$

On cherche à optimiser le majorant lorsque  $a \geq \lambda$ . Posons

$$f(t) = \lambda(e^t - 1) - ta.$$

Alors

$$f'(t) = \lambda e^t - a.$$

Si  $a \geq \lambda$ , le choix

$$t_0 = \ln\left(\frac{a}{\lambda}\right)$$

est positif ou nul. Lorsque  $a > \lambda$ ,  $t_0 > 0$  et c'est le minimum de  $f$ . Lorsque  $a = \lambda$ , on obtient le résultat par passage à la limite  $t \rightarrow 0^+$ .

On a alors

$$\begin{aligned} f(t_0) &= \lambda\left(\frac{a}{\lambda} - 1\right) - a \ln\left(\frac{a}{\lambda}\right) \\ &= a - \lambda - a \ln a + a \ln \lambda. \end{aligned}$$

Ainsi,

$$\boxed{\mathbb{P}(X > a) \leq \exp(a \ln \lambda - a \ln a + a - \lambda) \quad (a \geq \lambda).}$$

**Question 7.c.** L'inégalité de Markov donne ici

$$\mathbb{P}(X > a) \leq \frac{\lambda}{a}.$$

La borne de la question 7.b vaut

$$\exp(a \ln \lambda - a \ln a + a - \lambda) = \exp\left(-a \ln\left(\frac{a}{\lambda}\right) + a - \lambda\right).$$

Lorsque  $a \rightarrow +\infty$ , cette borne tend vers 0 beaucoup plus rapidement que  $\lambda/a$ . Plus précisément,

$$\frac{\exp(a \ln \lambda - a \ln a + a - \lambda)}{\lambda/a} \xrightarrow{a \rightarrow +\infty} 0.$$

Donc, pour  $a$  grand, la meilleure borne est celle de la question 7.b.

## Partie III : Étude d'un endomorphisme

---

On considère une fonction polynomiale  $P$ . Pour  $v > 0$ , on définit

$$L_P(v) = \frac{1}{v} \int_0^{+\infty} P(s) e^{-s/v} ds.$$

**Question 8.** Soit  $v > 0$ . Comme  $P$  est un polynôme,

$$|P(s)| e^{-s/(2v)} \xrightarrow{s \rightarrow +\infty} 0.$$

Donc il existe  $M > 0$  tel que, pour tout  $s > M$ ,

$$|P(s)| e^{-s/(2v)} \leq 1.$$

En multipliant par  $e^{-s/(2v)}$ , on obtient

$$|P(s)|e^{-s/v} \leq e^{-s/(2v)}.$$

Ainsi, pour  $s > M$ ,

$$\boxed{|P(s)e^{-s/v}| \leq e^{-s/(2v)}}.$$

La fonction  $s \mapsto P(s)e^{-s/v}$  est continue sur  $[0, M]$ , donc intégrable sur ce segment. Sur  $[M, +\infty[$ , elle est dominée en valeur absolue par  $s \mapsto e^{-s/(2v)}$ , qui est intégrable. Par conséquent, l'intégrale est convergente et  $L_P(v)$  est bien défini.

**Question 9.** Soit  $k \in \mathbb{N}$ . On effectue le changement de variable  $s = vu$ , donc  $ds = v du$ . Alors

$$\begin{aligned} \frac{1}{v} \int_0^{+\infty} s^k e^{-s/v} ds &= \frac{1}{v} \int_0^{+\infty} (vu)^k e^{-u} v du \\ &= v^k \int_0^{+\infty} u^k e^{-u} du. \end{aligned}$$

Or, par intégrations par parties successives,

$$\int_0^{+\infty} u^k e^{-u} du = k!.$$

Donc

$$\boxed{\frac{1}{v} \int_0^{+\infty} s^k e^{-s/v} ds = k!v^k}.$$

Si

$$P(X) = \sum_{k=0}^d a_k X^k,$$

alors, par linéarité de l'intégrale,

$$\boxed{L_P(v) = \sum_{k=0}^d k! a_k v^k}.$$

On fixe désormais  $d \in \mathbb{N}^*$  et l'on définit l'application  $L$  par

$$L\left(\sum_{k=0}^d a_k X^k\right) = \sum_{k=0}^d k! a_k X^k.$$

**Question 10.** L'application  $L$  envoie bien  $\mathbb{R}_d[X]$  dans  $\mathbb{R}_d[X]$ . De plus, si

$$P = \sum_{k=0}^d a_k X^k, \quad Q = \sum_{k=0}^d b_k X^k,$$

et si  $\alpha, \beta \in \mathbb{R}$ , alors

$$\begin{aligned} L(\alpha P + \beta Q) &= L\left(\sum_{k=0}^d (\alpha a_k + \beta b_k) X^k\right) \\ &= \sum_{k=0}^d k! (\alpha a_k + \beta b_k) X^k \\ &= \alpha L(P) + \beta L(Q). \end{aligned}$$

Donc  $L$  est linéaire. Comme son espace de départ et d'arrivée est le même,

$$\boxed{L \text{ est un endomorphisme de } \mathbb{R}_d[X]}.$$

**Question 11.** La base canonique de  $\mathbb{R}_d[X]$  est

$$(1, X, X^2, \dots, X^d).$$

Elle possède  $d + 1$  éléments. Donc

$$\dim(\mathbb{R}_d[X]) = d + 1.$$

**Question 12.** Par définition,

$$\ker(L) = \{P \in \mathbb{R}_d[X] \mid L(P) = 0\}.$$

Soit

$$P = \sum_{k=0}^d a_k X^k.$$

Alors

$$L(P) = 0 \iff \sum_{k=0}^d k! a_k X^k = 0.$$

Par unicité des coefficients d'un polynôme,

$$\forall k \in \{0, \dots, d\}, \quad k! a_k = 0.$$

Or  $k! \neq 0$ , donc  $a_k = 0$  pour tout  $k$ . Ainsi  $P = 0$ , et

$$\ker(L) = \{0\}.$$

**Question 13.** L'endomorphisme  $L$  est injectif puisque son noyau est réduit à  $\{0\}$ . Or  $L$  est un endomorphisme d'un espace vectoriel de dimension finie. Ainsi, par le théorème du rang, injectivité et surjectivité sont équivalentes. Donc

$$L \text{ est surjectif.}$$

On peut aussi voir directement que, si

$$Q(X) = \sum_{k=0}^d b_k X^k,$$

alors

$$P(X) = \sum_{k=0}^d \frac{b_k}{k!} X^k$$

vérifie  $L(P) = Q$ .

**Question 14.** Lorsque  $d = 3$ , dans la base canonique  $(1, X, X^2, X^3)$ ,

$$L(1) = 1, \quad L(X) = X, \quad L(X^2) = 2X^2, \quad L(X^3) = 6X^3.$$

Donc

$$\text{Mat}_{(1, X, X^2, X^3)}(L) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix}.$$

**Question 15.** Pour tout  $k \in \{0, \dots, d\}$ ,

$$L(X^k) = k!X^k.$$

Ainsi chaque monôme  $X^k$  est un vecteur propre de  $L$  associé à la valeur propre  $k!$ .

Comme  $0! = 1! = 1$ , la valeur propre 1 possède pour espace propre

$$E_1 = \text{Vect}(1, X).$$

Pour  $m \in \{2, \dots, d\}$ , les valeurs  $m!$  sont deux à deux distinctes, et

$$E_{m!} = \text{Vect}(X^m).$$

Les valeurs propres de  $L$  sont donc

$$\text{Sp}(L) = \begin{cases} \{1\}, & \text{si } d = 1, \\ \{1, 2!, 3!, \dots, d!\}, & \text{si } d \geq 2. \end{cases}$$

**Question 16.** La base canonique

$$(1, X, X^2, \dots, X^d)$$

est formée de vecteurs propres de  $L$ . Donc  $L$  est diagonalisable pour tout  $d \in \mathbb{N}^*$ . Ainsi,

$$\text{pour tout } d \in \mathbb{N}^*, \quad L \text{ est diagonalisable.}$$

**Question 17.a.** On considère

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad C(A) = \{M \in \mathcal{M}_3(\mathbb{R}) \mid AM = MA\}.$$

L'application

$$\Phi : \mathcal{M}_3(\mathbb{R}) \rightarrow \mathcal{M}_3(\mathbb{R}), \quad \Phi(M) = AM - MA$$

est linéaire. Alors

$$C(A) = \ker(\Phi).$$

Un noyau d'application linéaire est un sous-espace vectoriel. Donc

$$C(A) \text{ est un sous-espace vectoriel de } \mathcal{M}_3(\mathbb{R}).$$

**Question 17.b.** Soit

$$M = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & k \end{pmatrix}.$$

Alors

$$AM = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ g & h & k \end{pmatrix}, \quad MA = \begin{pmatrix} 0 & 0 & c \\ 0 & 0 & f \\ 0 & 0 & k \end{pmatrix}.$$

Ainsi

$$AM = MA \iff c = f = g = h = 0.$$

Donc

$$M \in C(A) \iff M = \begin{pmatrix} a & b & 0 \\ d & e & 0 \\ 0 & 0 & k \end{pmatrix}.$$

**Question 17.c.** D'après la question précédente, une base de  $C(A)$  est

$$\boxed{(E_{11}, E_{12}, E_{21}, E_{22}, E_{33}),}$$

où  $E_{ij}$  désigne la matrice dont le seul coefficient non nul est un 1 en position  $(i, j)$ . Donc

$$\boxed{\dim C(A) = 5.}$$

**Question 17.d.** La matrice

$$B = E_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

appartient à  $C(A)$  d'après la question 17.b.

En revanche,

$$\text{Vect}(I_3, A) = \{\alpha I_3 + \beta A \mid \alpha, \beta \in \mathbb{R}\} = \left\{ \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha + \beta \end{pmatrix} \right\}.$$

Toutes les matrices de  $\text{Vect}(I_3, A)$  sont diagonales, alors que  $B$  ne l'est pas. Donc

$$\boxed{B \in C(A) \quad \text{et} \quad B \notin \text{Vect}(I_3, A).}$$

## Partie IV : Loi de Yule-Simon

Dans les parties IV et V, on note  $\nu$  le paramètre strictement supérieur à 1 de la loi de Yule-Simon. Pour tout  $k \in \mathbb{N}$ , on pose

$$p_k = \nu \int_0^{+\infty} (1 - e^{-s})^k e^{-(\nu+1)s} ds.$$

**Question 18.** Pour tout  $s \geq 0$ ,

$$0 \leq 1 - e^{-s} \leq 1.$$

Donc

$$0 \leq (1 - e^{-s})^k e^{-(\nu+1)s} \leq e^{-(\nu+1)s}.$$

Or  $\nu + 1 > 0$ , donc

$$\int_0^{+\infty} e^{-(\nu+1)s} ds$$

est convergente. Par comparaison, l'intégrale définissant  $p_k$  est convergente.

**Question 19.** On effectue le changement de variable

$$t = e^{-s}.$$

Alors  $dt = -e^{-s} ds = -t ds$ , donc  $ds = -\frac{dt}{t}$ . Lorsque  $s = 0$ ,  $t = 1$ , et lorsque  $s \rightarrow +\infty$ ,  $t \rightarrow 0$ . Ainsi,

$$\begin{aligned} p_k &= \nu \int_0^{+\infty} (1 - e^{-s})^k e^{-(\nu+1)s} ds \\ &= \nu \int_1^0 (1 - t)^k t^{\nu+1} \left(-\frac{dt}{t}\right) \\ &= \nu \int_0^1 t^\nu (1 - t)^k dt. \end{aligned}$$

Donc

$$\boxed{\forall k \in \mathbb{N}, \quad p_k = \nu \int_0^1 t^\nu (1 - t)^k dt.}$$

**Question 20.a.** Pour  $t \in ]0, 1]$ ,

$$\sum_{k=0}^N (1-t)^k = \frac{1 - (1-t)^{N+1}}{1 - (1-t)} = \frac{1 - (1-t)^{N+1}}{t}.$$

Pour  $t = 0$ , la somme vaut  $N + 1$ . Dans les intégrales qui suivent, la formule sur  $]0, 1]$  suffit puisque le point  $t = 0$  est isolé.

**Question 20.b.** Comme  $\nu > 1$ , on a  $\nu - 1 > 0$ . Pour tout  $t \in [0, 1]$ ,

$$0 \leq t^{\nu-1} \leq 1.$$

Donc

$$0 \leq t^{\nu-1}(1-t)^{N+1} \leq (1-t)^{N+1}.$$

En intégrant sur  $[0, 1]$ ,

$$\boxed{\int_0^1 t^{\nu-1}(1-t)^{N+1} dt \leq \int_0^1 (1-t)^{N+1} dt.}$$

**Question 20.c.** D'abord, chaque  $p_k$  est positif puisque l'intégrande est positif.

Pour  $N \in \mathbb{N}$ , posons

$$S_N = \sum_{k=0}^N p_k.$$

D'après la question 19,

$$\begin{aligned} S_N &= \nu \int_0^1 t^\nu \sum_{k=0}^N (1-t)^k dt \\ &= \nu \int_0^1 t^\nu \frac{1 - (1-t)^{N+1}}{t} dt \\ &= \nu \int_0^1 t^{\nu-1} dt - \nu \int_0^1 t^{\nu-1}(1-t)^{N+1} dt. \end{aligned}$$

Or

$$\nu \int_0^1 t^{\nu-1} dt = 1.$$

En notant

$$R_N = \int_0^1 t^{\nu-1}(1-t)^{N+1} dt,$$

la question 20.b donne

$$0 \leq R_N \leq \int_0^1 (1-t)^{N+1} dt = \frac{1}{N+2} \xrightarrow{N \rightarrow +\infty} 0.$$

Donc

$$S_N \xrightarrow{N \rightarrow +\infty} 1.$$

Ainsi,

$$\sum_{k=0}^{+\infty} p_k = 1.$$

La famille  $(p_k)_{k \in \mathbb{N}}$  définit donc une loi de probabilité sur  $\mathbb{N}$ . Il existe alors une variable aléatoire réelle discrète  $X$  telle que

$$\boxed{\forall k \in \mathbb{N}, \quad \mathbb{P}(X = k) = p_k.}$$

On dit que  $X$  suit une loi de Yule-Simon de paramètre  $\nu$ .

## Partie V : Calculs de moments

Dans cette partie, on suppose  $\nu > 2$ . On considère une variable aléatoire  $X$  suivant la loi de Yule-Simon de paramètre  $\nu$ .

**Question 21.a.** Pour  $n, p \in \mathbb{N}$ , on pose

$$I_{n,p} = \int_0^1 t^n (1-t)^p dt.$$

On calcule  $I_{n,p+1}$  par intégration par parties :

$$I_{n,p+1} = \int_0^1 t^n (1-t)^{p+1} dt.$$

On pose

$$u(t) = (1-t)^{p+1}, \quad dv = t^n dt.$$

Alors

$$u'(t) = -(p+1)(1-t)^p, \quad v(t) = \frac{t^{n+1}}{n+1}.$$

Les termes de bord sont nuls, donc

$$\begin{aligned} I_{n,p+1} &= \frac{p+1}{n+1} \int_0^1 t^{n+1} (1-t)^p dt \\ &= \frac{p+1}{n+1} I_{n+1,p}. \end{aligned}$$

Ainsi,

$$I_{n,p+1} = \frac{p+1}{n+1} I_{n+1,p}.$$

**Question 21.b.** On prouve la formule par récurrence sur  $p$ .

Pour  $p = 0$ ,

$$I_{n,0} = \int_0^1 t^n dt = \frac{1}{n+1} = \frac{n!0!}{(n+1)!}.$$

La formule est donc vraie au rang  $p = 0$ .

Supposons-la vraie pour un certain  $p \in \mathbb{N}$ . Alors, d'après la question 21.a,

$$\begin{aligned} I_{n,p+1} &= \frac{p+1}{n+1} I_{n+1,p} \\ &= \frac{p+1}{n+1} \cdot \frac{(n+1)p!}{(n+p+2)!} \\ &= \frac{n!(p+1)!}{(n+p+2)!}. \end{aligned}$$

La formule est donc vraie au rang  $p+1$ . Ainsi, pour tout  $(n, p) \in \mathbb{N}^2$ ,

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

**Question 21.c.** On utilise la question précédente avec  $n = 1$  et  $p = N$  :

$$\int_0^1 t(1-t)^N dt = I_{1,N} = \frac{1!N!}{(N+2)!} = \frac{1}{(N+1)(N+2)}.$$

Donc

$$(N+1) \int_0^1 t(1-t)^N dt = \frac{1}{N+2} \xrightarrow{N \rightarrow +\infty} 0.$$

Ainsi,

$$\boxed{\lim_{N \rightarrow +\infty} (N+1) \int_0^1 t(1-t)^N dt = 0.}$$

**Question 22.a.** Pour  $t \in ]0, 1]$ , on part de la somme géométrique

$$\sum_{k=0}^N r^k = \frac{1-r^{N+1}}{1-r}$$

avec  $r = 1 - t$ . En dérivant par rapport à  $r$ ,

$$\sum_{k=1}^N k r^{k-1} = \frac{1-r^{N+1} - (N+1)r^N(1-r)}{(1-r)^2}.$$

En remplaçant  $r$  par  $1 - t$ , on obtient

$$\boxed{\sum_{k=1}^N k(1-t)^{k-1} = \frac{1}{t^2} - \frac{(1-t)^{N+1}}{t^2} - \frac{(N+1)(1-t)^N}{t}.$$

**Question 22.b.** Comme  $\nu > 2$ , l'intégrale est convergente et

$$\begin{aligned} \int_0^1 \nu t^\nu (1-t) \frac{1}{t^2} dt &= \nu \int_0^1 t^{\nu-2} (1-t) dt \\ &= \nu \left( \int_0^1 t^{\nu-2} dt - \int_0^1 t^{\nu-1} dt \right) \\ &= \nu \left( \frac{1}{\nu-1} - \frac{1}{\nu} \right) \\ &= \frac{1}{\nu-1}. \end{aligned}$$

Donc

$$\boxed{\int_0^1 \nu t^\nu (1-t) \frac{1}{t^2} dt = \frac{1}{\nu-1}.$$

**Question 22.c.** Pour  $N \geq 1$ , posons

$$S_N = \sum_{k=1}^N k p_k.$$

D'après la question 19,

$$\begin{aligned} S_N &= \nu \int_0^1 t^\nu \sum_{k=1}^N k(1-t)^k dt \\ &= \nu \int_0^1 t^\nu (1-t) \sum_{k=1}^N k(1-t)^{k-1} dt. \end{aligned}$$

En utilisant la question 22.a,

$$\begin{aligned} S_N &= \nu \int_0^1 t^\nu (1-t) \left[ \frac{1}{t^2} - \frac{(1-t)^{N+1}}{t^2} - \frac{(N+1)(1-t)^N}{t} \right] dt \\ &= \frac{1}{\nu-1} - \nu \int_0^1 t^{\nu-2} (1-t)^{N+2} dt - \nu(N+1) \int_0^1 t^{\nu-1} (1-t)^{N+1} dt. \end{aligned}$$

Les deux derniers termes sont positifs. Notons-les  $R_N$  et  $T_N$ .

D'une part,

$$0 \leq R_N \leq \nu \int_0^1 (1-t)^{N+2} dt = \frac{\nu}{N+3} \xrightarrow{N \rightarrow +\infty} 0.$$

D'autre part, puisque  $\nu > 2$ , on a pour  $t \in [0, 1]$

$$t^{\nu-1} (1-t)^{N+1} \leq t(1-t)^N.$$

Donc, d'après la question 21.c,

$$0 \leq T_N \leq \nu(N+1) \int_0^1 t(1-t)^N dt \xrightarrow{N \rightarrow +\infty} 0.$$

Ainsi,

$$S_N \xrightarrow{N \rightarrow +\infty} \frac{1}{\nu-1}.$$

La série  $\sum_{k \geq 1} k p_k$  converge, donc  $X$  possède une espérance, et

$$\boxed{\mathbb{E}(X) = \frac{1}{\nu-1}}.$$

**Question 23.a.** On admet que  $(X-1)X$  possède une espérance et que

$$\mathbb{E}((X-1)X) = \nu \int_0^1 t^\nu (1-t)^2 \frac{2}{t^3} dt.$$

Alors

$$\begin{aligned} \mathbb{E}((X-1)X) &= 2\nu \int_0^1 t^{\nu-3} (1-t)^2 dt \\ &= 2\nu \int_0^1 t^{\nu-3} (1-2t+t^2) dt \\ &= 2\nu \left( \frac{1}{\nu-2} - \frac{2}{\nu-1} + \frac{1}{\nu} \right). \end{aligned}$$

En simplifiant,

$$\boxed{\mathbb{E}((X-1)X) = \frac{4}{(\nu-1)(\nu-2)}}.$$

**Question 23.b.** Comme

$$X^2 = X(X-1) + X,$$

on a

$$\begin{aligned} \mathbb{E}(X^2) &= \mathbb{E}(X(X-1)) + \mathbb{E}(X) \\ &= \frac{4}{(\nu-1)(\nu-2)} + \frac{1}{\nu-1} \\ &= \frac{\nu+2}{(\nu-1)(\nu-2)}. \end{aligned}$$

Donc  $X$  possède une variance, et

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 \\ &= \frac{\nu + 2}{(\nu - 1)(\nu - 2)} - \frac{1}{(\nu - 1)^2} \\ &= \frac{\nu^2}{(\nu - 1)^2(\nu - 2)}.\end{aligned}$$

Ainsi,

$$\boxed{\mathbb{V}(X) = \frac{\nu^2}{(\nu - 1)^2(\nu - 2)}}.$$

**Question 23.c.** On a

$$\frac{\mathbb{V}(X)}{\mathbb{E}(X)} = \frac{\nu^2}{(\nu - 1)^2(\nu - 2)} \cdot (\nu - 1) = \frac{\nu^2}{(\nu - 1)(\nu - 2)}.$$

Or

$$\nu^2 - (\nu - 1)(\nu - 2) = 3\nu - 2 > 0$$

car  $\nu > 2$ . Donc

$$\boxed{\frac{\mathbb{V}(X)}{\mathbb{E}(X)} > 1}.$$

**Question 24.** Pour choisir entre le modèle de Poisson de la partie I et le modèle de Yule-Simon de la partie V à partir d'un échantillon, on peut comparer la moyenne empirique et la variance empirique.

Si l'échantillon est  $(x_1, \dots, x_n)$ , on calcule par exemple

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2.$$

Pour une loi de Poisson,

$$\frac{\mathbb{V}(X)}{\mathbb{E}(X)} = 1.$$

Pour une loi de Yule-Simon avec  $\nu > 2$ ,

$$\frac{\mathbb{V}(X)}{\mathbb{E}(X)} > 1.$$

Ainsi, si le rapport empirique

$$\frac{s^2}{\bar{x}}$$

est proche de 1, on privilégiera le modèle de la partie I, c'est-à-dire le modèle de Poisson. S'il est nettement supérieur à 1, on privilégiera plutôt le modèle de Yule-Simon, qui rend compte d'une surdispersion.