CORRIGÉ du SUJET
"Distance entre deux distributions de probabilités sur $\mathbb{N}$"

I. Nombre de points fixes d'une permutation.

1. Comme $\mathrm{Card}(\mathcal{S}_n) = n!$, on a $0 \leq d_n \leq n!$, donc $\left|\dfrac{d_n}{n!}\right| \leq 1$. On en déduit que $R \geq R'$, où $R' = 1$ est le rayon de convergence de la série entière $\sum x^n$.

2. Pour construire une permutation de $[\![1,n]\!]$ ayant k points fixes, on choisit l'ensemble A de ses points fixes, de cardinal k, il y a $\dbinom{n}{k}$ choix possibles. Ensuite, on choisit un dérangement du complémentaire $B = [\![1,n]\!] \setminus A$, et il y a d_{n-k} possibilités. Le nombre de permutations de $[\![1,n]\!]$ ayant exactement k points fixes est donc $\dbinom{n}{k}\, d_{n-k}$.

L'univers $\mathcal{S}_n$, de cardinal $n!$, étant muni de la probabilité uniforme P_n, on a donc

$$\forall k \in [\![0,n]\!] \qquad P_n(X_n = k) = \frac{\big|\{X_n = k\}\big|}{|\mathcal{S}_n|} = \frac{\dbinom{n}{k}\, d_{n-k}}{n!} = \frac{d_{n-k}}{k!\,(n-k)!}\ .$$

3. Comme s et $x \mapsto e^x$ sont développables en série entière sur $]-1,1[$, par produit de Cauchy, on obtient sur cet intervalle

$$\begin{aligned} s(x)\,e^x &= \left(\sum_{p=0}^{+\infty} \frac{d_p}{p!}\,x^p\right)\left(\sum_{q=0}^{+\infty} \frac{x^q}{q!}\right) \\ &= \sum_{n=0}^{+\infty}\left(\sum_{k=0}^{n} \frac{d_{n-k}}{k!\,(n-k)!}\right)x^n = \sum_{n=0}^{+\infty}\left(\sum_{k=0}^{n} P_n(X_n = k)\right)x^n \\ &= \sum_{n=0}^{+\infty} x^n = \frac{1}{1-x}\ . \end{aligned}$$

Donc, pour tout $x \in\,]-1,1[$, on a $s(x) = \dfrac{e^{-x}}{1-x}$. Comme $\lim\limits_{x\to 1^-} s(x) = +\infty$, on ne peut pas prolonger s en une fonction continue sur $]-R,R[$ avec $R > 1$. On en déduit que $R = 1$.

4. Pour $x \in\,]-1,1[$, on a donc $(1-x)\displaystyle\sum_{n=0}^{+\infty} \frac{d_n}{n!}\,x^n = \sum_{n=0}^{+\infty} \frac{(-1)^n}{n!}\,x^n$ soit, après réorganisation du premier membre,

$$d_0 + \sum_{n=1}^{+\infty}\left(\frac{d_n}{n!} - \frac{d_{n-1}}{(n-1)!}\right)x^n = 1 + \sum_{n=1}^{+\infty} \frac{(-1)^n}{n!}\,x^n\ .$$

Par identification (unicité du développement en série entière), on retrouve $d_0 = 1$ et on obtient

$$\forall n \in \mathbb{N}^* \qquad \frac{d_n}{n!} - \frac{d_{n-1}}{(n-1)!} = \frac{(-1)^n}{n!}\ .$$

Changeons de notations: fixons $n \in \mathbb{N}^*$, alors pour tout $k \in [\![1,n]\!]$, $\dfrac{d_k}{k!} - \dfrac{d_{k-1}}{(k-1)!} = \dfrac{(-1)^k}{k!}$.
En sommant ces inégalités et en observant le télescopage dans le membre de gauche, il vient

$$\frac{d_n}{n!} - d_0 = \sum_{k=1}^{n} \frac{(-1)^k}{k!}\ .$$

Comme $d_0 = 1$, on conclut que

$$\forall n \in \mathbb{N} \qquad \frac{d_n}{n!} = \sum_{k=0}^{n} \frac{(-1)^k}{k!} .$$

5. Pour $k \in [\![0, n]\!]$, on a

$$P_n(X_n = k) = \frac{1}{k!} \, \frac{d_{n-k}}{(n-k)!} = \frac{1}{k!} \sum_{i=0}^{n-k} \frac{(-1)^i}{i!} .$$

6. Comme $U_i(\mathcal{S}_n) = \{0, 1\}$, U_i est une variable de Bernoulli dont le paramètre est

$$P_n(U_i = 1) = P_n\big(\sigma(i) = i\big) = \frac{(n-1)!}{n!} = \frac{1}{n} .$$

En effet, il y a autant de permutations de $[\![1, n]\!]$ fixant l'élément i que de permutations de $[\![1, n]\!] \setminus \{i\}$, soit $(n-1)!$. Donc $U_i \sim \mathcal{B}\Big(\dfrac{1}{n}\Big)$.

Si $i \neq j$, la variable U_iU_j prend aussi ses valeurs dans $\{0, 1\}$, c'est donc encore une variable de Bernoulli, et son paramètre est

$$P_n(U_iU_j = 1) = P_n(U_i = 1, U_j = 1) = P_n\big(\sigma(i) = i, \sigma(j) = j\big) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)} .$$

En effet, ici aussi, si $i \neq j$, il y a autant de permutations de $[\![1, n]\!]$ fixant les éléments i et j que de permutations de $[\![1, n]\!] \setminus \{i, j\}$, soit $(n-2)!$.

7. Il est immédiat que $X_n = \displaystyle\sum_{i=1}^{n} U_i$.

De la question précédente, on déduit que $\mathrm{E}(U_i) = \dfrac{1}{n}$ et $\mathrm{V}(U_i) = \dfrac{n-1}{n^2}$.

Par linéarité de l'espérance, $\mathrm{E}(X_n) = \displaystyle\sum_{i=1}^{n} \mathrm{E}(U_i) = n \, \frac{1}{n} = 1$.

Ensuite, $\mathrm{V}(X_n) = \mathrm{V}\Big(\displaystyle\sum_{i=1}^{n} U_i\Big) = \sum_{i=1}^{n} \mathrm{V}(U_i) + \sum_{i \neq j} \mathrm{Cov}(U_i, U_j)$.

Et, pour $i \neq j$, on a $\mathrm{Cov}(U_i, U_j) = \mathrm{E}(U_iU_j) - \mathrm{E}(U_i)\,\mathrm{E}(U_j)$. Or, $\mathrm{E}(U_iU_j) = \dfrac{1}{n(n-1)}$ d'après la question précédente, ce qui donne $\mathrm{Cov}(U_i, U_j) = \dfrac{1}{n(n-1)} - \dfrac{1}{n^2} = \dfrac{1}{n^2(n-1)}$.

Comme il y a $n(n-1)$ couples (i, j) dans $[\![1, n]\!]^2$ avec $i \neq j$, on termine le calcul:

$$\mathrm{V}(X_n) = n \cdot \frac{n-1}{n^2} + n(n-1) \cdot \frac{1}{n^2(n-1)} = 1 .$$

8. On a $P_n(X_n = k) = \dfrac{1}{k!} \displaystyle\sum_{i=0}^{n-k} \frac{(-1)^i}{i!} \underset{n \to +\infty}{\longrightarrow} y_k = \frac{1}{k!} \sum_{i=0}^{+\infty} \frac{(-1)^i}{i!} = \frac{e^{-1}}{k!}$.

On reconnaît la distribution de probabilités d'une loi de Poisson de paramètre 1, donc $Y \sim \mathcal{P}(1)$.

9. D'après le cours (ou par un calcul immédiat), on a $G_Y(s) = \sum_{k=0}^{+\infty} P(Y = k)\, s^k = e^{s-1}$, la série entière ayant un rayon de convergence infini. D'autre part, comme X_n prend ses valeurs dans $[\![0, n]\!]$, sa fonction génératrice est polynomiale, on a pour tout s réel,

$$\begin{aligned}
G_{X_n}(s) &= \sum_{k=0}^{n} P_n(X_n = k)\, s^k = \sum_{k=0}^{n} \left(\frac{1}{k!} \sum_{i=0}^{n-k} \frac{(-1)^i}{i!} \right) s^k \\
&= \sum_{k=0}^{n} \left(\frac{1}{k!} \sum_{j=k}^{n} \frac{(-1)^{j-k}}{(j-k)!} \right) s^k \qquad \textit{translation d'indice} \\
&= \sum_{j=0}^{n} \left(\frac{1}{j!} \sum_{k=0}^{j} \binom{j}{k} (-1)^{j-k} s^k \right) \qquad \textit{interversion des sommes} \\
&= \sum_{j=0}^{n} \frac{(s-1)^j}{j!}\,.
\end{aligned}$$

On observe bien que, pour tout s réel fixé, $\lim_{n \to +\infty} G_{X_n}(s) = \sum_{j=0}^{+\infty} \frac{(s-1)^j}{j!} = e^{s-1} = G_Y(s)$.

II. Convergence en variation totale.

10. Pour tout k, on a $\left|x(k) - y(k)\right| \leq \left|x(k)\right| + \left|y(k)\right| = x(k) + y(k)$. Comme les séries $\sum_{k \geq 0} x(k)$ et $\sum_{k \geq 0} y(k)$ convergent et ont pour somme 1, par comparaison de séries à termes positifs, on déduit la convergence de la série $\sum_{k \geq 0} \left|x(k) - y(k)\right|$ et la majoration de sa somme par 2.
Il est clair, au passage, qu'une distribution de probabilités prend ses valeurs dans $[0, 1]$. Donc $d_{VT}(x, y)$ existe et appartient à $[0, 1]$.

Si $x = y$, alors $d_{VT}(x, y) = 0$. Réciproquement, comme une somme de termes positifs est nulle si et seulement si tous ses termes sont nuls, si $d_{VT}(x, y) = 0$, alors $\left|x(k) - y(k)\right| = 0$ pour tout k, donc $x = y$ (*axiome de séparation*).

La *symétrie* $d_{VT}(y, x) = d_{VT}(x, y)$ est immédiate.

Enfin pour l'*inégalité triangulaire*:

$$\begin{aligned}
2\, d_{VT}(x, z) &= \sum_{k=0}^{+\infty} \left|x(k) - z(k)\right| = \sum_{k=0}^{+\infty} \Big| \big(x(k) - y(k)\big) + \big(y(k) - z(k)\big) \Big| \\
&\leq \sum_{k=0}^{+\infty} \Big(\left|x(k) - y(k)\right| + \left|y(k) - z(k)\right| \Big) = 2\, \big(d_{VT}(x, y) + d_{VT}(y, z) \big)\,.
\end{aligned}$$

11. On a ici

$$2\, d_{VT}(p_X, p_Y) = \left|p_X(0) - p_Y(0)\right| + \left|p_X(1) - p_Y(1)\right| = \left|(1-\lambda) - (1-\mu)\right| + |\lambda - \mu| = 2\, |\lambda - \mu|\,,$$

donc $d_{VT}(x, y) = |\lambda - \mu|$.

12. Puisque $p_X(k)$ est nul pour $k \geq 2$, on calcule

$$\begin{aligned} 2\, d_{VT}(p_X, \pi_\lambda) &= \big|p_X(0) - \pi_\lambda(0)\big| + \big|p_X(1) - \pi_\lambda(1)\big| + \sum_{k=2}^{+\infty} \pi_\lambda(k) \\ &= \big|1 - \lambda - e^{-\lambda}\big| + \big|\lambda - \lambda\, e^{-\lambda}\big| + e^{-\lambda} \sum_{k=2}^{+\infty} \frac{\lambda^k}{k!} \\ &= e^{-\lambda} - (1-\lambda) + \lambda\,(1 - e^{-\lambda}) + e^{-\lambda}\,(e^\lambda - 1 - \lambda)\ . \end{aligned}$$

On a utilisé l'inégalité $e^{-\lambda} \geq 1 - \lambda$ qui résulte de la convexité de la fonction exponentielle. Après simplifications, il reste $d_{VT}(p_X, \pi_\lambda) = \lambda\,(1 - e^{-\lambda})$.

Avec le même argument de convexité, $1 - e^{-\lambda} \leq \lambda$, donc $d_{VT}(p_X, \pi_\lambda) \leq \lambda^2$.

13. Il suffit d'appliquer la définition, sachant que $p_{X_n}(k)$ est nul pour $k > n$:

$$\begin{aligned} 2\, d_{VT}(p_{X_n}, \pi_1) &= \sum_{k=0}^{n} \big|p_{X_n}(k) - \pi_1(k)\big| + \sum_{k=n+1}^{+\infty} \pi_1(k) \\ &= \sum_{k=0}^{n} \left| \frac{1}{k!} \sum_{i=0}^{n-k} \frac{(-1)^i}{i!} - \frac{e^{-1}}{k!} \right| + \sum_{k=n+1}^{+\infty} \frac{e^{-1}}{k!} \\ &= \sum_{k=0}^{n} \frac{1}{k!} \left| \sum_{i=n-k+1}^{+\infty} \frac{(-1)^i}{i!} \right| + e^{-1}\, r_n\ , \end{aligned}$$

où l'on pose $r_n = \displaystyle\sum_{k=n+1}^{+\infty} \frac{1}{k!}$. On a utilisé $e^{-1} = \displaystyle\sum_{i=0}^{+\infty} \frac{(-1)^i}{i!}$.

14. Par un décalage d'indice, on écrit $r_n = \displaystyle\sum_{k=0}^{+\infty} \frac{1}{(n+1+k)!}$. Or, pour tout $k \in \mathbb{N}$,

$$(n+1+k)! = (n+1)!\,(n+2)(n+3)\cdots(n+k+1) \geq (n+1)!\,(n+2)^k\ .$$

Donc, en reconnaissant une série géométrique,

$$r_n \leq \sum_{k=0}^{+\infty} \frac{1}{(n+1)!\,(n+2)^k} = \frac{1}{(n+1)!} \sum_{k=0}^{+\infty} \frac{1}{(n+2)^k} = \frac{n+2}{n+1}\, \frac{1}{(n+1)!}\ .$$

Comme, par ailleurs, $r_n \geq \dfrac{1}{(n+1)!}$, l'encadrement

$$\frac{1}{(n+1)!} \leq r_n \leq \frac{n+2}{n+1}\, \frac{1}{(n+1)!}\ ,$$

les termes extrêmes étant équivalents entre eux, donne $r_n \underset{n\to+\infty}{\sim} \dfrac{1}{(n+1)!}$.

15. La suite $\left(\dfrac{1}{i!}\right)_{i\in\mathbb{N}}$ étant décroissante de limite nulle, le théorème spécial des séries alternées permet de majorer en valeur absolue le reste d'ordre $n-k$ de la série $\sum \dfrac{(-1)^i}{i!}$ par le premier terme négligé, à savoir $\left|\displaystyle\sum_{i=n-k+1}^{+\infty} \frac{(-1)^i}{i!}\right| \leq \dfrac{1}{(n-k+1)!}$. On déduit

$$\begin{aligned}
\sum_{k=0}^{n} \frac{1}{k!}\left|\sum_{i=n-k+1}^{+\infty} \frac{(-1)^i}{i!}\right| &\leq \sum_{k=0}^{n} \frac{1}{k!\,(n-k+1)!} = \frac{1}{(n+1)!}\sum_{k=0}^{n}\binom{n+1}{k} \\
&\leq \frac{1}{(n+1)!}\sum_{k=0}^{n+1}\binom{n+1}{k} = \frac{2^{n+1}}{(n+1)!}
\end{aligned}$$

en concluant par la formule du binôme. Finalement, de **13.**, on déduit la majoration

$$0 \leq d_{VT}(p_{X_n}, \pi_1) \leq \frac{2^n}{(n+1)!} + \frac{r_n}{2e}\ .$$

De **14.**, on déduit que r_n est négligeable devant $\dfrac{2^n}{(n+1)!}$ lorsque $n\to+\infty$. On conclut donc que

$$d_{VT}(p_{X_n}, \pi_1) \underset{n\to+\infty}{=} O\left(\frac{2^n}{(n+1)!}\right)\ .$$

III. Autres estimations de distances en variation totale.

16. Posons $z = x * y$. On a clairement $z(k) \geq 0$ pour tout $k \in \mathbb{N}$. Puis, par produit de Cauchy,

$$\sum_{k=0}^{+\infty} z(k) = \sum_{k=0}^{+\infty}\left(\sum_{i+j=k} x(i)\,y(j)\right) = \left(\sum_{i=0}^{+\infty} x(i)\right)\left(\sum_{j=0}^{+\infty} y(j)\right) = 1\times 1 = 1\ ,$$

donc $x * y \in \mathcal{D}_{\mathbb{N}}$.

17. Pour tout k entier naturel, on a (union disjointe):

$$\{X+Y=k\} = \bigsqcup_{i+j=k} \big(\{X=i\}\ \cap\ \{Y=j\}\big)\ ,$$

donc par additivité finie, et les variables X et Y étant indépendantes,

$$\begin{aligned}
p_{X+Y}(k) &= P(X+Y=k) = \sum_{i+j=k} P(X=i, Y=j) \\
&= \sum_{i+j=k} P(X=i)\,P(Y=j) = \sum_{i+j=k} p_X(i)\,p_Y(j) \\
&= (p_X * p_Y)(k)\ .
\end{aligned}$$

18. Soit k un entier naturel. On a alors

$$\begin{aligned}
\big|(x*y)(k)-(u*v)(k)\big| &= \Big|\sum_{i+j=k}\big(x(i)y(j)-u(i)v(j)\big)\Big| \\
&= \Big|\sum_{i+j=k}\Big(y(j)\,\big(x(i)-u(i)\big)+u(i)\,\big(y(j)-v(j)\big)\Big)\Big| \\
&\le \sum_{i+j=k}\Big|y(j)\,\big(x(i)-u(i)\big)+u(i)\,\big(y(j)-v(j)\big)\Big| \\
&\le \sum_{i+j=k} y(j)\,\big|x(i)-u(i)\big|+\sum_{i+j=k}u(i)\,\big|y(j)-v(j)\big|
\end{aligned}$$

en utilisant diverses inégalités triangulaires et la positivité des nombres $u(i)$ et $y(j)$.

19. Calculons encore un peu, en utilisant la majoration obtenue en question **18.**:

$$\begin{aligned}
2\,d_{VT}(x*y,u*v) &= \sum_{k=0}^{+\infty}\big|(x*y)(k)-(u*v)(k)\big| \\
&\le \sum_{k=0}^{+\infty}\Big(\sum_{i+j=k}y(j)\,\big|x(i)-u(i)\big|\Big)+\sum_{k=0}^{+\infty}\Big(\sum_{i+j=k}u(i)\,\big|y(j)-v(j)\big|\Big)\ .
\end{aligned}$$

Ce calcul peut être fait dans $[0,+\infty]$ puisque tous les termes sont positifs, ce qui évite de se poser des problèmes de convergence a priori. On reconnaît alors de nouveau des produits de Cauchy. Donc

$$\begin{aligned}
2\,d_{VT}(x*y,u*v) &\le \Big(\sum_{i=0}^{+\infty}\big|x(i)-u(i)\big|\Big)\Big(\sum_{j=0}^{+\infty}y(j)\Big)+\Big(\sum_{i=0}^{+\infty}u(i)\Big)\Big(\sum_{j=0}^{+\infty}\big|y(j)-v(j)\big|\Big) \\
&= 2\,d_{VT}(x,u)\quad\times\quad 1\quad+\quad 1\quad\times\quad 2\,d_{VT}(y,v)\ ,
\end{aligned}$$

ce qui est bien la majoration souhaitée.

20. Soit par un calcul direct, soit en considérant la somme de deux variables de Poisson indépendantes sur un espace probabilisé, on a facilement la relation

$$\forall(\alpha,\beta)\in(\mathbb{R}_+^*)^2\qquad \pi_\alpha*\pi_\beta=\pi_{\alpha+\beta}\ .$$

Sur un espace probabilisé $(\Omega,\mathcal{A},P)$, considérons n variables de Bernoulli indépendantes, de même paramètre λ, notées $X_1,\,\cdots,\,X_n$. Pour tout $k\in[\![1,n]\!]$, soit $S_k=\displaystyle\sum_{i=1}^{k}X_i$. Il est alors connu que $S_k\sim\mathcal{B}(k,\lambda)$. Donc S_n a la même loi que U, et $p_{S_n}=p_U$.

On peut montrer par récurrence finie sur $k\in[\![1,n]\!]$ que $d_{VT}(p_{S_k},\pi_{k\lambda})\le k\lambda^2$.

- initialisation pour $k=1$: c'est la question **12.** puisque $S_1=X_1$ est une variable de Bernoulli ;

- hérédité: soit $k\in[\![1,n-1]\!]$, supposons la propriété vraie au rang k. Comme $S_{k+1}=S_k+X_{k+1}$ avec $S_k\perp\!\!\!\perp X_{k+1}$ d'après le lemme des coalitions, on a

$$p_{S_{k+1}} = p_{S_k} * p_{X_{k+1}} = p_{S_k} * p_{X_1} \ .$$

Par ailleurs, $\pi_{(k+1)\lambda} = \pi_{k\lambda} * \pi_\lambda$. De la question **19.** et de l'hypothèse de récurrence, on déduit alors

$$\begin{array}{rcl} d_{VT}(p_{S_{k+1}}, \pi_{(k+1)\lambda}) & = & d_{VT}(p_{S_k} * p_{X_1}, \pi_{k\lambda} * \pi_\lambda) \\ & \leq & d_{VT}(p_{S_k}, \pi_{k\lambda}) + d_{VT}(p_{X_1}, \pi_\lambda) \\ & \leq & k\lambda^2 + \lambda^2 \ = \ (k+1)\,\lambda^2 \ , \end{array}$$

ce qui achève la récurrence, et donne la majoration souhaitée pour $k = n$.

21. On peut faire un calcul direct (et classique, je ne détaille pas) pour montrer que

$$\lim_{n\to+\infty} P(B_n = k) = \lim_{n\to+\infty} \binom{n}{k} \left(\frac{\alpha}{n}\right)^k \left(1 - \frac{\alpha}{n}\right)^{n-k} = e^{-\alpha}\,\frac{\alpha^k}{k!} \ .$$

Mais on peut aussi exploiter ce qui précède puisque, en posant $\lambda = \dfrac{\alpha}{n}$, on a la majoration

$$d_{VT}(p_{B_n}, \pi_\alpha) \leq n \left(\frac{\alpha}{n}\right)^2 = \frac{\alpha^2}{n} \ .$$

Pour tout $k \in \mathbb{N}$ fixé, on a clairement

$$\left|p_{B_n}(k) - \pi_\alpha(k)\right| \leq \sum_{j=0}^{+\infty} \left|p_{B_n}(j) - \pi_\alpha(j)\right| = 2\,d_{VT}(p_{B_n}, \pi_\alpha) \leq 2\,\frac{\alpha^2}{n} \quad \underset{n\to+\infty}{\longrightarrow} \quad 0 \ ,$$

donc par majoration de la valeur absolue, $\displaystyle\lim_{n\to+\infty} P(B_n = k) = \pi_\alpha(k) = e^{-\alpha}\,\frac{\alpha^k}{k!}$.

22. Pour $n \in \mathbb{N}^*$ et $\lambda \in \,]0,1[$, notons $\theta_{n,\lambda}$ la **distribution binomiale de paramètres** n et λ, c'est-à-dire l'application $\theta_{n,\lambda} : \mathbb{N} \to [0,1]$ définie par

$$\forall k \in \mathbb{N} \qquad \theta_{n,\lambda}(k) = \begin{cases} \dbinom{n}{k}\,\lambda^k\,(1-\lambda)^{n-k} & \text{si} \quad 0 \leq k \leq n \\ 0 & \text{si} \qquad k > n \end{cases} .$$

Ainsi, $\theta_{1,\lambda}$ est la "distribution de Bernoulli" de paramètre λ.

Par un calcul analogue à celui de la question **20.**, on montre par récurrence que, si λ et μ sont dans $]0,1[$, alors pour tout $n \in \mathbb{N}^*$, on a $d_{VT}(\theta_{n,\lambda}\ ,\ \theta_{n,\mu}) \leq n\,|\lambda - \mu|$, l'initialisation étant donnée par la question **11.**

En particulier, pour $n > \max\big\{\lfloor\alpha\rfloor, \lfloor\beta\rfloor\big\}$, on a $d_{VT}\big(\theta_{n,\frac{\alpha}{n}}\ ,\ \theta_{n,\frac{\beta}{n}}\big) \leq |\beta - \alpha|$.

La question **20.** montre aussi que $d_{VT}(\theta_{n,\lambda}\ ,\ \pi_{n\lambda}) \leq n\lambda^2$ si $\lambda \in \,]0,1[$.

Par l'inégalité triangulaire et la symétrie (question **10.**), pour $n > \max\big\{\lfloor\alpha\rfloor, \lfloor\beta\rfloor\big\}$, on a

$$\begin{array}{rcl} d_{VT}(\pi_\alpha, \pi_\beta) & \leq & d_{VT}\big(\theta_{n,\frac{\alpha}{n}}\ ,\ \pi_\alpha\big) + d_{VT}\big(\theta_{n,\frac{\alpha}{n}}\ ,\ \theta_{n,\frac{\beta}{n}}\big) + d_{VT}\big(\theta_{n,\frac{\beta}{n}}\ ,\ \pi_\beta\big) \\ & \leq & n\left(\dfrac{\alpha}{n}\right)^2 + |\beta - \alpha| + n\left(\dfrac{\beta}{n}\right)^2 \ = \ \dfrac{\alpha^2 + \beta^2}{n} + |\beta - \alpha| \ . \end{array}$$

En faisant tendre n vers $+\infty$, on obtient $d_{VT}(\pi_\alpha, \pi_\beta) \leq |\beta - \alpha|$

Variante. En notant δ_0 la "distribution de Dirac" donnée par $\delta_0(k) = \delta_{0,k}$ (symbole de Kronecker) pour tout $k \in \mathbb{N}$, on vérifie que $u * \delta_0 = u$ pour tout $u \in \mathcal{D}_{\mathbb{N}}$, puis on écrit, en supposant $\beta > \alpha$,

$$\begin{aligned}
d_{VT}(\pi_\alpha, \pi_\beta) &= d_{VT}(\pi_\alpha * \delta_0, \pi_\alpha * \pi_{\beta-\alpha}) \\
&\leq d_{VT}(\pi_\alpha, \pi_\alpha) + d_{VT}(\delta_0, \pi_{\beta-\alpha}) \ = \ d_{VT}(\delta_0, \pi_{\beta-\alpha}) \\
&= 1 - e^{\alpha-\beta} \quad (\textit{calcul facile}) \ \leq \ \beta - \alpha \ = \ |\beta - \alpha| \ .
\end{aligned}$$

~~Les questions sont de difficultés variées. Certaines sont très proches du cours, d'autres demandent une bonne maîtrise des théorèmes, d'autres enfin sont vraiment difficiles. Elles ont permis aux candidats de montrer leurs diverses qualités. L'échelonnement des notes est très satisfaisant.~~
~~Une analyse détaillée des questions est présentée dans l'annexe E.~~

~~1.7.2 Conclusion~~

~~La maîtrise des techniques et des résultats du cours est indispensable pour réussir les concours. C'était particulièrement le cas pour ce sujet, qui demandait en particulier une bonne maîtrise de l'algèbre bilinéaire. Beaucoup de candidats ont su montrer leurs qualités sur des questions assez techniques.~~

~~Rappelons pour terminer que la qualité de la rédaction et la présentation sont prises en compte dans l'évaluation des copies. Les correcteurs apprécient notamment que les résultats soient soulignés, que les copies ne soient pas un jeu de piste et que les ratures soient propres ! Enfin, certaines copies écrites avec une encre gommable sont un peu difficiles à lire ; ce type de stylo est donc à éviter.~~

1.8 Mathématiques 2 - filière PSI

1.8.1 Présentation générale et intérêt scientifique du sujet

Le sujet avait trait à plusieurs modes d'approximation des lois de Poisson par des lois à support fini. Dans un premier temps (partie **I**), on étudiait la probabilité qu'une permutation d'un ensemble à n éléments soit un dérangement, par la méthode des séries entières génératrices, puis la loi du nombre X_n de points fixes d'une permutation d'un ensemble à n éléments. On démontrait, lorsque n tend vers $+\infty$, la convergence en loi de X_n vers la loi de Poisson de paramètre 1.
La deuxième partie du sujet étudiait une mesure effective de l'écart entre deux lois sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$: la mesure en question est la **distance de variation totale**, dont on montrait à la question 10 qu'elle vérifiait les axiomes d'une distance sur l'ensemble des familles positives sommables de somme 1 (que l'on peut identifier à l'ensemble des lois de probabilité sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$).
L'objectif essentiel, dans le reste du problème, était de quantifier la convergence observée en partie **I** au sens de cette mesure (questions 14 et 15), puis de faire de même pour l'approximation de la loi de Poisson de paramètre λ par la loi binomiale $\mathcal{B}(\backslash, \lambda/\backslash)$ (question 20) et l'approximation d'une loi de Poisson par une autre (question 22).
La stratégie, dans cette dernière partie, était de contrôler la distance de variation totale entre deux produits de convolution (opération sur les lois correspondant à l'addition de deux variables aléatoires entières indépendantes) en fonction des distances facteur à facteur. Les derniers résultats était obtenus par écriture de la loi binomiale $\mathcal{B}(\backslash, \lambda/\backslash)$ comme produit de convolution de n lois de Bernoulli de paramètre $\frac{\lambda}{n}$, et de la loi de Poisson $\mathcal{P}(\lambda)$ comme produit de convolution de n lois de Poisson de paramètre $\frac{\lambda}{n}$.

1.8.2 Structure du sujet et questions souvent abordées

Le sujet pouvait être intégralement traité dans le temps imparti. Plusieurs candidats y sont essentiellement parvenus, ne lâchant des points que par quelques erreurs d'étourderie et imprécisions dans l'argumentation.
Les trois grandes parties présentaient des dépendances significatives entre elles. Le résultat de la question 5 intervenait dans les questions 12 à 15, tandis que l'objet introduit dans la partie **II**, à savoir la distance en variation totale, était d'usage constant dans la partie **III**.
La plupart des candidats ont tenté de traiter les toutes premières questions, puis la question 6, la question 8, les questions 10 à 12, puis 16 et 17. Le reste a moins souvent été abordé.
Beaucoup de candidats ont essayé de traiter un très grand nombre de questions, mais en survolant absolument tout. En général, ces copies ont reçu une note très faible. Le picorage est très fortement déconseillé : on attend au contraire un réel investissement des candidats dans le sujet.
Le jury relève unanimement un important relâchement dans la présentation des copies par rapport aux éditions précédentes. On déplore de nombreuses copies à la limite de la lisibilité, des copies dont la rédaction et les justifications sont quasi-absentes, d'autres truffées de fautes d'orthographe.
Dans certaines copies, il est rare de voir correctement quantifiées les propositions mathématiques énoncées : les objets du discours ne sont pas fixés.
Une analyse détaillée des questions est présentée dans l'annexe F.

1.8.3 Conclusion

L'observation générale est le manque de soin dans la vérification des hypothèses des théorèmes manipulés. L'exemple le plus flagrant de cette tendance est les produits de Cauchy (de séries entières ou de séries numériques), dont les hypothèses ne sont qu'épisodiquement rappelées, et encore plus épisodiquement vérifiées avec rigueur.
La rédaction des raisonnements de dénombrement met en général les candidats en grande difficulté. On est loin en la matière de l'excès de formalisme, bien au contraire : le plus souvent ces questions révèlent un franc manque de rigueur dans l'usage du vocabulaire.
Les questions de rayon de convergence sont rarement satisfaisantes : presque tous les candidats oublient de considérer la valeur absolue du terme général, beaucoup tentent de manipuler des inégalités entre les sommes de séries entières (sans que rien n'ait vraiment été bien justifié) pour conclure sur les rayons de convergence. On attendait des candidats une distinction claire entre la série entière (qui est une série de fonctions, ce qui équivaut à la donnée d'une suite de fonctions) et sa somme.
Enfin, on note beaucoup de passages en force de la part des candidats : sauts abrupts à la conclusion, étapes de calcul sans justification claire etc.
En définitive, les candidats disposant d'une maîtrise suffisante du programme et ayant bien intégré les attendus en termes de rédaction et de justification se sont très facilement détachés des autres.