Devoir maison n° 10

(corrigé)

Table des matières

1 Commentaires

Le premier exercice est un classique posé un peu partout (le premier exemple qui me vient en tête est le Concours Commun Mines-Ponts, Mathématiques II, filière PSI, année 2020 – mais ce n'est pas de là que je tire mon énoncé, qui est « fait maison »). Le second exercice est extrait de l'épreuve de Mathématiques I de Centrale-Supélec, filière PSI, année 2022.

Enfin, le problème est une adaptation de l'épreuve de Mathématiques I du Concours Commun Mines-Ponts, filière MP, année 2021.

Il y aurait beaucoup de choses à commenter concernant ces deux exercices et le problème, mais je suis comme Évariste Galois un certain 29 mai : je n'ai pas le temps (et c'est bien la seule fois où j'ai un point commun avec lui).

Je n'ai corrigé que les deux exercices : le corrigé du problème est dû à Jérémy Larochette (professeur en MPI au lycée Leconte de Lisle de Saint-Denis, à La Réunion) et Rémi Souveton (professeur en MP au lycée Albert Châtelet de Douai).

Questions faciles ou classiques à retravailler

— PREMIER EXERCICE : questions 1 à 4 ;

— DEUXIÈME EXERCICE : questions 1 à 4 ;

— PROBLÈME : question 5, question 12 (la seconde partie de la question), question 14.

2 Rapport officiel de l'épreuve (deuxième exercice)

2.1 Présentation du sujet

Le sujet de Maths 1 de la filière PSI 2022 a pour objectif la démonstration d'un résultat de géométrie dans un espace euclidien. Pour cela, plusieurs résultats intermédiaires sont démontrés (...) sur les variables aléatoires discrètes.

(...) Il était attendu des candidats qu'ils maîtrisent bien leur cours d'algèbre linéaire pour traiter ce problème (...). Une bonne maîtrise des raisonnements élémentaires de probabilités était également indispensable : indépendance ou incompatibilité d'évènements, propriétés de l'espérance et de la variance. Enfin, quelques autres chapitres (espaces euclidiens, études de fonctions) rentraient également en jeu.

2.2 Analyse globale des résultats

(...) [Le] cours n'est pas toujours bien appris et certains résultats, pourtant très importants, ne sont parfois pas cités correctement (propriétés de la trace ou formule du binôme de Newton par exemple).

(...) La dernière partie a été moins abordée par les candidats, sans doute à cause de sa position dans le problème, mais aussi peut-être parce qu'elle nécessitait de combiner habilement plusieurs résultats d'analyse, d'algèbre et de probabilités.

Concernant la présentation des copies, une majorité est assez clairement présentée, avec des questions numérotées correctement, traitées dans l'ordre et des résultats encadrés. Ceux qui dérogent à ces règles de base font tout de suite mauvaise impression et prennent le risque d'être moins bien compris par les correcteurs.

2.3 Commentaires sur les réponses apportées et conseils aux futurs candidats

Le jury souhaite insister sur un certain nombre de points qui ont souvent posé problèmes aux candidats :

- Les candidats doivent faire un effort de présentation des copies, numéroter les questions, les traiter dans l'ordre (quitte à laisser des blancs pour y revenir) et encadrer leurs résultats.
- L'utilisation des abréviations doit être limitée : si certaines (CNS, SSI...) sont très couramment utilisées, d'autres (FPT pour formule des probabilités totales...) le sont nettement moins. De même, l'emploi d'abréviations telles que $\forall$, $\iff$ doit être modéré dans des explications, *et ces symboles ne doivent figurer que dans des assertions ne contenant que des symboles mathématiques.*
- Un raisonnement doit être articulé avec des mots clés (considérons, or, donc, car, en effet) : les hypothèses et les objectifs doivent être clairement identifiés.
- Lorsqu'une égalité entre deux ensembles est demandée et qu'un raisonnement par « double inclusion », est choisi, il est important de bien démontrer les deux inclusions, ou à défaut, de signaler que l'une d'entre elles est évidente si tel est le cas.
- Pour démontrer une équivalence entre deux propriétés, on peut raisonner directement par équivalence, ou raisonner par double implication. Mais montrer une seule implication ne suffit pas.
- Lorsqu'un résultat précédemment démontré est utilisé, il est important de le signaler. De même, le lemme des coalitions (...) devait être rappelé par tout candidat qui souhaitait l'utiliser.
- Dans la question 4, il était attendu que les candidats donnent précisément les hypothèses de l'inégalité de Markov. Par ailleurs, il ne fallait pas oublier de traiter le cas où $t = 0$.

Conclusion

Le sujet était plutôt long mais la progressivité du texte et la diversité des chapitres mathématiques nécessaires (...), ont permis à tous les candidats de traiter de nombreuses questions et de mettre en évidence leurs compétences. Quelques lacunes sur des notions de base ont malheureusement aussi été repérées.

De nombreux candidats ont su montrer leur maîtrise du langage mathématique en général, et plus spécifiquement des points qui étaient nécessaires pour aborder les diverses parties de ce problème : le langage des probabilités (...) Quelques candidats ont même abordé avec succès les questions plus difficiles qui parsemaient le sujet, et les correcteurs tiennent à les en féliciter.

Les correcteurs ont constaté cette année une bonne maîtrise de la rédaction (logique, double implication, clarté des calculs entrepris...). Une partie non négligeable des copies propose une rédaction très agréable à lire en mêlant rigueur, justesse et clarté. Les correcteurs encouragent par ailleurs vivement les candidats à utiliser un brouillon et à ne pas commencer systématiquement la rédaction aussitôt l'énoncé lu. De nombreuses erreurs grossières pourraient ainsi être évitées.

3 Rapport officiel de l'épreuve (problème)

3.1 Généralités et présentation du sujet

Le sujet portait sur l'analyse et les probabilités, avec des utilisations conséquentes de techniques d'analyse asymptotique au programme de première année.

On trouvait également une application du théorème de convergence dominée et des questions sur les théories de l'intégration et des probabilités, le tout assurant une bonne couverture du programme d'analyse et de probabilités de la filière MP.

Le problème consistait, pour sa partie principale, en la démonstration du théorème de Moivre-Laplace, théorème important de probabilités qui décrit le comportement limite d'une suite de variables aléatoires $(X_n)_{n\in\mathbb{N}}$ qui suivent des lois binomiales $\mathcal{B}(n,p)$.

Cela donnait un sujet manifestement trop long, surtout pour une épreuve de trois heures, mais comme il était progressif, un barème adapté, attribuant peu de points aux dernières questions, a permis d'obtenir une moyenne d'épreuve correcte et un bon étalement des notes.

Cette répartition des points n'a lésé personne, les meilleures copies contenaient une résolution correcte des dix-sept premières questions et quelques éléments épars dans les cinq dernières, ce qui permettait d'avoir une excellente note.

Par contre, les candidats qui ont bâclé les premières questions pour aller grappiller quelquefois jusqu'à la dernière n'y ont pas gagné.

Nous n'avons pas eu de cas de candidats qui négligent les premières questions et traite correctement les dernières, et qui auraient donc été défavorisés par le barème.

3.2 Analyse détaillée des questions

Q1. Cette question a été souvent abandonnée et traitée correctement par moins de 5% des candidats. Il suffisait de penser à étudier le bon quotient, mais cela n'avait en effet rien d'évident surtout dans le stress d'une épreuve en temps limité.

Q2. La première limite demandée se déduisait directement de la Q2, par contre la deuxième demandait un raisonnement plus soigneux, soit par une minoration, soit par une utilisation des o.

Q3-4. Ces questions étaient assez calculatoires et ne présentaient pas de difficultés importantes, elles ont été traitée plutôt correctement, mais souvent avec beaucoup de ratures.

Q5. Dans cette question on entrait dans le vif du sujet avec les probabilités, les performances restaient correctes, il fallait bien préciser les valeurs de l'espérance et de la variance d'une loi binomiale puis les propriétés de l'espérance et de la variance utilisées.

Q6. La question ne présentait pas de difficulté, mais la qualité de la rédaction était variable.

Q7. On attendait, pour justifier que la fonction était en escalier, que l'on donne explicitement les intervalles sur lesquels elle était constante. On retrouvait dans la fin de la question les problèmes de confusion entre passage à la limite dans les inégalités et théorème d'encadrement.

Q8. En général il n'y avait que la première partie de cette question qui était abordée, avec un passage à la limite assez rarement justifié correctement. La deuxième partie ne se trouvait que dans les très bonnes copies.

Q9. La question pouvait effrayer, mais elle se résumait à bien utiliser les questions précédentes.

Q10. La question ne demandait que de la précision dans l'utilisation des équivalents.

Q11. La question présentait bien des analogies avec la question 3, mais dire « on fait comme à la question 3 » ne suffisait pas.

Q12. La question demandait une rédaction rigoureuse dans l'utilisation du développement limité de Ψ.

Q13. Dans cette question, l'utilisation du théorème de convergence dominée semblait évidente, pourtant elle n'a pas fait l'unanimité et la condition de domination, qu'il fallait aller chercher à la question 4, a été très rarement bien traitée.

À partir de là nous n'avons plus trouvé de résolution complète, quelques très rares copies donnaient des éléments de réponses intéressants, la plupart ne contenaient que des tentatives éventuellement de grappillage.

3.3 Conclusions

En conclusion, on peut évoquer, comme l'année dernière, la présentation des copies. Du fait de la numérisation, certaines techniques de correction sont interdites, mais l'utilisation du brouillon reste autorisée, et même recommandée.

Il n'est pas question de faire une résolution complète au brouillon puis de recopier, mais il faudrait éviter de faire figurer toutes les tentatives sur la copie en raturant les échecs.

On peut imaginer ce que cela peut donner sur un calcul un peu long de développement limité, comme il y en avait dans le sujet de cette année.

4 Corrigé

Premier exercice

1. Par le théorème de transfert, on a :

$$\forall x \in [0,1], \quad B_n(f)(x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \mathrm{P}(S_n(x) = k) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} x^k (1-x)^{n-k}.$$

   On reconnaît l'application polynomiale associée à $\displaystyle\sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) X^k (1-X)^{n-k}$, d'où le résultat.

2. Soit $n \in \mathbb{N} \setminus \{0\}$. On a par linéarité de l'espérance et l'inégalité triangulaire :

$$\begin{aligned} |B_n(f)(x) - f(x)| = \left| \mathrm{E}\left(f\left(\frac{S_n(x)}{n}\right)\right) - \mathrm{E}\left(f(x)\right)\right| &= \left| \mathrm{E}\left(f\left(\frac{S_n(x)}{n}\right) - f(x)\right)\right| \\ &\leqslant \mathrm{E}\left(\left| f\left(\frac{S_n(x)}{n}\right) - f(x)\right|\right). \end{aligned}$$

   Il suffit alors d'écrire : $1 = \mathbb{1}_{A_{n,\eta}(x)} + \mathbb{1}_{\overline{A_{n,\eta}(x)}}$, et d'encore utiliser la linéarité de l'espérance pour avoir le résultat voulu :

$$|B_n(f)(x) - f(x)| \leqslant \mathrm{E}\left(\left| f\left(\frac{S_n(x)}{n}\right) - f(x)\right| \mathbb{1}_{A_{n,\eta}(x)}\right) + \mathrm{E}\left(\left| f\left(\frac{S_n(x)}{n}\right) - f(x)\right| \mathbb{1}_{\overline{A_{n,\eta}(x)}}\right).$$

3. Soit $n \in \mathbb{N} \setminus \{0\}$. Comme f est continue sur le compact $[0,1]$, par le théorème de Heine elle est uniformément continue. Soit $\eta > 0$ un module d'uniforme continuité de f associé à ε. Alors η est effectivement indépendant de x, et on a :

$$\left| f\left(\frac{S_n(x)}{n}\right) - f(x)\right| \mathbb{1}_{\overline{A_{n,\eta}(x)}} \leqslant \varepsilon.$$

En effet, soit $\mathbb{1}_{\overline{A_{n,\eta}(x)}}(\omega) = 0$ (pour ω dans l'ensemble de départ de $S_n(x)$), auquel cas l'inégalité est trivialement vraie ; soit $\mathbb{1}_{\overline{A_{n,\eta}(x)}}(\omega) = 1$, ce qui veut dire que l'évènement $\left(\left|\frac{S_n(x)}{n} - x\right| \leqslant \eta\right)$ est réalisé, et donc par définition de η on a :

$$\left|f\left(\frac{S_n(x)(\omega)}{n}\right) - f(x)\right| \mathbb{1}_{\overline{A_{n,\eta}(x)}}(\omega) \leqslant \left|f\left(\frac{S_n(x)(\omega)}{n}\right) - f(x)\right| \leqslant \varepsilon.$$

Bref, l'inégalité est vérifiée pour tout ω. Par croissance de l'espérance :

$$\mathrm{E}\left(\left|f\left(\frac{S_n(x)}{n}\right) - f(x)\right| \mathbb{1}_{\overline{A_{n,\eta}(x)}}\right) \leqslant \mathrm{E}(\varepsilon) = \varepsilon,$$

d'où le résultat.

4. Comme : $\forall n \in \mathbb{N}\setminus\{0\}$, $S_n(x) \sim \mathcal{B}(n,x)$, on a : $\forall n \in \mathbb{N}\setminus\{0\}$, $\mathrm{E}\left(\frac{S_n(x)}{n}\right) = \frac{1}{n}\mathrm{E}(S_n(x)) = \frac{nx}{n} = x$.
Par la loi faible des grands nombres (une variable de loi binomiale étant au fond une somme de variables indépendantes de même loi de Bernoulli, quitte à avoir un univers suffisamment grand), on a donc :

$$\forall n \in \mathbb{N}\setminus\{0\}, \quad \mathrm{P}\left(A_{n,\eta}(x)\right) = \mathrm{P}\left(\left|\frac{S_n(x)}{n} - x\right| > \eta\right) \leqslant \frac{x(1-x)}{n\eta} \leqslant \frac{1}{n\eta},$$

où l'on a utilisé le fait que le maximum de la fonction $x \mapsto x(1-x)$ soit en $x = \frac{1}{2}$. Cette majoration est indépendante de $x \in [0,1]$. Par propriété de la borne supérieure :

$$\forall n \in \mathbb{N}\setminus\{0\}, \quad 0 \leqslant \left\|\mathrm{P}\left(A_{n,\eta}\right)\right\|_\infty \leqslant \frac{1}{n\eta} \underset{n\to+\infty}{\longrightarrow} 0,$$

donc par le théorème des gendarmes : $\lim\limits_{n\to+\infty} \left\|\mathrm{P}\left(A_{n,\eta}\right)\right\|_\infty = 0$, et la suite de fonctions $\left(\mathrm{P}\left(A_{n,\eta}\right)\right)_{n\geqslant 1}$ converge uniformément sur $[0,1]$ vers la fonction nulle.

5. On reprend l'inégalité de la question 2, en exploitant la majoration de la question 3. On a :

$$\begin{aligned}\forall n \in \mathbb{N}\setminus\{0\},\ \forall x \in [0,1], \quad |B_n(f)(x) - f(x)| &\leqslant \mathrm{E}\left(\left|f\left(\frac{S_n(x)}{n}\right) - f(x)\right| \mathbb{1}_{A_{n,\eta}(x)}\right) + \varepsilon \\ &\leqslant \mathrm{E}\left(\left(\left|f\left(\frac{S_n(x)}{n}\right)\right| + |f(x)|\right) \mathbb{1}_{A_{n,\eta}(x)}\right) + \varepsilon \\ &\leqslant 2\|f\|_\infty \mathrm{E}\left(\mathbb{1}_{A_{n,\eta}(x)}\right) + \varepsilon \\ &= 2\|f\|_\infty \mathrm{P}\left(A_{n,\eta}(x)\right) + \varepsilon \\ &\leqslant 2\|f\|_\infty \left\|\mathrm{P}\left(A_{n,\eta}\right)\right\|_\infty + \varepsilon.\end{aligned}$$

Or, par la question précédente : $\lim\limits_{n\to+\infty} 2\|f\|_\infty \left\|\mathrm{P}\left(A_{n,\eta}\right)\right\|_\infty = 0$, donc il existe $N \in \mathbb{N}\setminus\{0\}$ tel que pour tout entier $n \geqslant N$, on ait : $2\|f\|_\infty \left\|\mathrm{P}\left(A_{n,\eta}\right)\right\|_\infty \leqslant \varepsilon$. On peut conclure :

$$\exists N \in \mathbb{N}\setminus\{0\},\ \forall n \geqslant N,\ \forall x \in [0,1], \quad |B_n(f)(x) - f(x)| \leqslant 2\varepsilon,$$

donc $\left(B_n(f)\right)_{n\geqslant 0}$ converge uniformément sur $[0,1]$ vers f et est une suite d'applications polynomiales : cela démontre le théorème d'approximation de Weierstraß dans le cas du segment $[0,1]$.
Pour le cas d'une fonction $f : [a,b] \to \mathbb{R}$ définie sur un segment $[a,b]$ quelconque (avec $a < b$ puisque le cas $a = b$ est trivial), il suffit d'approcher uniformément l'application définie sur $[0,1]$ par $x \mapsto f(a+x(b-a))$ par une suite d'applications polynomiales $\left(P_n\right)_{n\geqslant 0}$, et de constater qu'alors la suite $\left(x \mapsto P_n\left(\frac{x-a}{b-a}\right)\right)_{n\geqslant 0}$ converge uniformément vers f sur $[a,b]$. D'où le résultat.

Deuxième exercice

1. Soit $(i,j) \in I^2$ tel que $i \neq j$. Comme $\vec{u}_i$ et $\vec{u}_j$ sont unitaires, on a par l'inégalité de Cauchy-Schwarz :
$$|\langle \vec{u}_i, \vec{u}_j\rangle| \leqslant \|\vec{u}_i\| \cdot \|\vec{u}_j\| = 1.$$
Cette majoration est indépendante de (i,j). Par propriété de la borne supérieure, on en déduit :
$$C(u) = \sup_{\substack{(i,j)\in I^2 \\ i\neq j}} \{|\langle \vec{u}_i, \vec{u}_j\rangle|\} \leqslant 1.$$
D'où le résultat, l'inégalité $C(u) \geqslant 0$ étant évidente (en effet $|\langle \vec{u}_i, \vec{u}_j\rangle| \geqslant 0$ pour tous i et j).

2. On a, en minorant la factorielle de n par ses facteurs pairs :
$$\forall n \in \mathbb{N}, \quad (2n)! = \prod_{k=1}^{2n} k \geqslant \prod_{\ell=1}^{n} (2k) = 2^n n!,$$
On en déduit : $\forall t \in \mathbb{R}, \forall n \in \mathbb{N}, \dfrac{t^{2n}}{(2n)!} \leqslant \dfrac{t^{2n}}{2^n \cdot n!}$, parce que $t^{2n} \geqslant 0$ pour tout $t \in \mathbb{R}$ et tout $n \in \mathbb{N}$. On écrit ensuite $t^{2n} = (t^2)^n$ pour avoir :
$$\forall t \in \mathbb{R}, \quad \mathrm{ch}(t) = \sum_{n=0}^{+\infty} \frac{t^{2n}}{(2n)!} \leqslant \sum_{n=0}^{+\infty} \frac{(t^2)^n}{2^n \cdot n!} = \sum_{n=0}^{+\infty} \frac{(t^2/2)^n}{n!} = e^{\frac{t^2}{2}},$$
d'après les développements en série entière en 0 des fonctions usuelles exponentielle et cosinus hyperbolique. D'où le résultat.

3. Soit $t \in \mathbb{R}$. Les variables $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ sont indépendantes, donc $e^{\frac{tX_1Y_1}{n}}, \ldots, e^{\frac{tX_nY_n}{n}}$ le sont par le lemme des coalitions. Passons au calcul de $\mathrm{E}\left(e^{t\langle X,Y\rangle}\right)$. On a par définition du produit scalaire usuel :
$$\mathrm{E}\left(e^{t\langle X,Y\rangle}\right) = \mathrm{E}\left(e^{\sum_{i=1}^{n} t\frac{X_i}{\sqrt{n}}\frac{Y_i}{\sqrt{n}}}\right) = \mathrm{E}\left(\prod_{i=1}^{n} e^{\frac{tX_iY_i}{n}}\right).$$
Comme les variables aléatoires $e^{\frac{tX_1Y_1}{n}}, \ldots, e^{\frac{tX_nY_n}{n}}$ sont indépendantes, l'espérance du produit est le produit des espérances, donc :
$$\mathrm{E}\left(e^{t\langle X,Y\rangle}\right) = \prod_{i=1}^{n} \mathrm{E}\left(e^{\frac{tX_iY_i}{n}}\right) = \left(\mathrm{E}\left(e^{\frac{tX_1Y_1}{n}}\right)\right)^n,$$
la dernière égalité étant vraie parce que les X_iY_i suivent toutes la même loi de Rademacher, et donc les $e^{\frac{tX_iY_i}{n}}$ suivent tous la même loi. Éclairons la première affirmation : puisque X_i et Y_i sont à valeurs dans $\{-1,1\}$, leur produit X_iY_i aussi (puisque $(\pm 1) \times (\pm 1) = \pm 1$), donc : $X_iY_i(\Omega) = \{-1,1\}$. Calculons $\mathrm{P}(X_iY_i = 1)$ et $\mathrm{P}(X_iY_i = -1)$. Puisque $X_iY_i = 1$ si et seulement si $X_i = Y_i = 1$ ou $X_i = Y_i = -1$, on a : $(X_iY_i = 1) = (X_i = 1, Y_i = 1) \cup (X_i = -1, Y_i = -1)$. Ces évènements étant disjoints, on a :
$$\mathrm{P}(X_iY_i = 1) = \mathrm{P}(X_i = 1, Y_i = 1) + \mathrm{P}(X_i = -1, Y_i = -1).$$
Or X_i et Y_i sont supposées indépendantes, donc :
$$\mathrm{P}(X_iY_i = 1) = \mathrm{P}(X_i = 1)\mathrm{P}(Y_i = 1) + \mathrm{P}(X_i = -1)\mathrm{P}(Y_i = -1) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}.$$
Alors : $\mathrm{P}(X_iY_i = -1) = 1 - \mathrm{P}(X_iY_i \neq -1) = 1 - \mathrm{P}(X_iY_i = 1) = \frac{1}{2}$. On a montré que X_iY_i suit la loi de Rademacher pour tout $i \in [\![1, n]\!]$.

Par le théorème de transfert et par définition de la loi de Rademacher, on a : $\mathrm{E}\left(e^{\frac{tX_1Y_1}{n}}\right) = e^{\frac{t}{n}} \cdot \frac{1}{2} + e^{-\frac{t}{n}}\frac{1}{2} = \mathrm{ch}\left(\frac{t}{n}\right)$. D'où :

$$\mathrm{E}\left(e^{t\langle X,Y\rangle}\right) = \left(\mathrm{ch}\left(\frac{t}{n}\right)\right)^n.$$

On conclut avec la question précédente :

$$\mathrm{E}\left(e^{t\langle X,Y\rangle}\right) = \left(\mathrm{ch}\left(\frac{t}{n}\right)\right)^n \leqslant \left(e^{\frac{(t/n)^2}{2}}\right)^n = e^{\frac{t^2}{2n}}.$$

4. Supposons $\varepsilon > 0$. Soit $t \in \mathbb{R}_+$. Posons : $Z = \langle X, Y\rangle$. Si : $Z \geqslant \varepsilon$, alors après multiplication par $t \geqslant 0$ on a : $tZ \geqslant t\varepsilon$, puis, par composition avec l'exponentielle : $e^{tZ} \geqslant e^{t\varepsilon}$. Ainsi l'évènement $(Z \geqslant \varepsilon)$ implique $(e^{tZ} \geqslant e^{t\varepsilon})$, d'où :

$$\mathrm{P}(Z \geqslant \varepsilon) \leqslant \mathrm{P}(e^{tZ} \geqslant e^{t\varepsilon}).$$

(En vérité on a même l'égalité si $t \neq 0$). Or, d'après l'inégalité de Markov (qu'on peut appliquer parce que e^{tZ} est positive) et par la question précédente, on a :

$$\mathrm{P}(e^{tZ} \geqslant e^{t\varepsilon}) \leqslant \frac{\mathrm{E}(e^{tZ})}{e^{t\varepsilon}} \leqslant \frac{e^{\frac{t^2}{2n}}}{e^{t\varepsilon}} = e^{\frac{t^2}{2n}-\varepsilon t},$$

d'où :

$$\mathrm{P}(Z \geqslant \varepsilon) \leqslant \mathrm{P}(e^{tZ} \geqslant e^{t\varepsilon}) \leqslant e^{\frac{t^2}{2n}-\varepsilon t}.$$

Ensuite : $(|Z| \geqslant \varepsilon) = (Z \geqslant \varepsilon) \cup (Z \leqslant -\varepsilon) = (Z \geqslant \varepsilon) \cup (-Z \geqslant -\varepsilon)$, et si $\varepsilon > 0$ alors les évènements sont disjoints donc :

$$\mathrm{P}(|Z| \geqslant \varepsilon) = \mathrm{P}(Z \geqslant \varepsilon) + \mathrm{P}(-Z \geqslant \varepsilon).$$

Justifions que ce qui précède peut aussi s'appliquer à $-Z$, afin d'en déduire une majoration de $\mathrm{P}(-Z \geqslant \varepsilon)$. Pour cela, il suffit d'appliquer l'inégalité $\mathrm{E}\left(e^{tZ}\right) \leqslant e^{\frac{\sigma^2 t^2}{2}}$, vérifiée par hypothèse sur Z, avec $-t$ au lieu de t ; c'est possible puisque l'inégalité est supposée vérifiée pour tout réel t. On a alors :

$$\forall t \in \mathbb{R}, \quad \mathrm{E}\left(e^{t(-Z)}\right) = \mathrm{E}\left(e^{(-t)Z}\right) \leqslant e^{\frac{(-t)^2}{2n}} = e^{\frac{t^2}{2n}}.$$

Ainsi $-Z$ vérifie la même hypothèse que Z, donc $\mathrm{P}(Z \geqslant \varepsilon)$ et $\mathrm{P}(-Z \geqslant \varepsilon)$ se majorent semblablement. On en déduit :

$$\forall t \in \mathbb{R}_+, \quad \mathrm{P}(-Z \geqslant \varepsilon) \leqslant e^{\frac{t^2}{2n}-\varepsilon t}.$$

Ainsi, pour tout $t \in \mathbb{R}_+$:

$$\mathrm{P}(|Z| \geqslant \varepsilon) = \mathrm{P}(Z \geqslant \varepsilon) + \mathrm{P}(-Z \geqslant \varepsilon) \leqslant 2e^{\frac{t^2}{2n}-\varepsilon t}.$$

Pour en déduire la majoration de l'énoncé, cherchons t tel que : $\frac{t^2}{2n} - \varepsilon t = -\frac{\varepsilon^2 n}{2}$. Cela revient à résoudre une équation polynomiale du second degré. Son unique solution est : $t_0 = \varepsilon n$. Comme ε est supposé strictement positif, on peut bien poser $t = t_0 \in \mathbb{R}_+$, et la majoration ci-dessus donne :

$$\mathrm{P}(|Z| \geqslant \varepsilon) \leqslant 2e^{-\frac{\varepsilon^2}{2\sigma^2}},$$

d'où le résultat (si $\varepsilon = 0$ alors la majoration est triviale parce qu'une probabilité est inférieure ou égale à 1).

5. Par sous-additivité, on a :

$$\mathrm{P}\left(\bigcup_{1\leqslant i<j\leqslant N}\left(\left|\left\langle X^i, X^j\right\rangle\right| \geqslant \varepsilon\right)\right) \leqslant \sum_{1\leqslant i<j\leqslant N} \mathrm{P}\left(\left|\left\langle X^i, X^j\right\rangle\right| \geqslant \varepsilon\right).$$

Par hypothèse, pour tout $(i,j) \in [\![1,n]\!]^2$ tel que $i \neq j$, les variables aléatoires $X_1^i, \dots, X_n^i, X_1^j, \dots, X_n^j$ sont indépendantes et de même loi de Rademacher, donc d'après la question précédente :

$$\mathrm{P}\left(\left|\left\langle X^i, X^j \right\rangle\right| \geqslant \varepsilon\right) \leqslant 2e^{-\frac{n\varepsilon^2}{2}}.$$

On en déduit :

$$\mathrm{P}\left(\bigcup_{1\leqslant i<j\leqslant N}\left(\left|\left\langle X^i, X^j \right\rangle\right| \geqslant \varepsilon\right)\right) \leqslant \sum_{1\leqslant i<j\leqslant N} 2e^{-\frac{n\varepsilon^2}{2}} = 2e^{-\frac{n\varepsilon^2}{2}} \sum_{1\leqslant i<j\leqslant N} 1.$$

Pour calculer cette dernière somme, il faut dénombrer combien il y a de couples $(i,j) \in [\![1,N]\!]^2$ tels que : $i < j$. On note que cela revient à dénombrer le nombre de choix de deux éléments parmi N (on appelle alors i le plus petit des deux éléments choisis et j le plus grand) : il y a $\binom{N}{2} = \frac{N(N-1)}{2}$ possibilités. Donc :

$$\mathrm{P}\left(\bigcup_{1\leqslant i<j\leqslant N}\left(\left|\left\langle X^i, X^j \right\rangle\right| \geqslant \varepsilon\right)\right) \leqslant 2e^{-\frac{n\varepsilon^2}{2}} \times \frac{N(N-1)}{2} = N(N-1)e^{-\frac{n\varepsilon^2}{2}},$$

d'où le résultat.

6. Puisque : $n \geqslant \frac{4\ln(N)}{\varepsilon^2}$, on a : $-\frac{n\varepsilon^2}{2} \leqslant -2\ln(N)$. Par croissance de l'exponentielle, cela implique :

$$\mathrm{P}\left(\bigcup_{1\leqslant i<j\leqslant N}\left(\left|\left\langle X^i, X^j \right\rangle\right| \geqslant \varepsilon\right)\right) \leqslant N(N-1)e^{-\frac{n\varepsilon^2}{2}} \leqslant N(N-1)e^{-2\ln(N)} = \frac{N(N-1)}{N^2} = \frac{N-1}{N} < 1.$$

Par passage au complémentaire :

$$\mathrm{P}\left(\bigcap_{1\leqslant i<j\leqslant N}\left|\left(\left\langle X^i, X^j \right\rangle\right| < \varepsilon\right)\right) = 1 - \mathrm{P}\left(\bigcup_{1\leqslant i<j\leqslant N}\left(\left|\left\langle X^i, X^j \right\rangle\right| \geqslant \varepsilon\right)\right) > 0.$$

Puisque l'évènement $\bigcap_{1\leqslant i<j\leqslant N} |(\langle X^i, X^j\rangle| < \varepsilon)$ est de probabilité strictement positive, il n'est pas l'évènement impossible : il existe donc $\omega \in \bigcap_{1\leqslant i<j\leqslant N} |(\langle X^i, X^j\rangle| < \varepsilon)$. Un tel élément ω vérifie (par définition de cette intersection), pour tout $(i,j) \in [\![1,N]\!]^2$ tel que $i<j$:

$$\left|\left\langle X^i(\omega), X^j(\omega) \right\rangle\right| < \varepsilon.$$

Si l'on pose : $\forall i \in [\![1,N]\!]$, $\vec{u}_i = X^i(\omega)$, alors $\vec{u}_i$ est bien un vecteur unitaire pour tout $i \in [\![1,N]\!]$, puisque :

$$\|X^i\| = \frac{1}{\sqrt{n}}\sqrt{\sum_{j=1}^n (X_j^i)^2} = \frac{1}{\sqrt{n}}\sqrt{\sum_{j=1}^n (\pm 1)^2} = \frac{1}{\sqrt{n}}\sqrt{n} = 1.$$

On a alors, pour tout $(i,j) \in [\![1,N]\!]^2$ tel que $i<j$, d'après la majoration ci-dessus : $|\langle \vec{u}_i, \vec{u}_j\rangle| < \varepsilon$. C'est aussi vrai si $j<i$, par symétrie du produit scalaire. Par propriété de la borne supérieure, on a donc : $C(u) < \varepsilon$. D'où le résultat.

PROBLÈME

Partie A. Étude asymptotique d'une suite.

1. Pour montrer que p_n est le plus grand des $\mathrm{P}(X_n = k)$ pour $k \in [\![0, n]\!]$, il suffit de montrer que pour tout $k \in [\![0, n]\!]$, on a : $\mathrm{P}(X_n = k) \leqslant \mathrm{P}(X_n = x_n) = p_n$.

   Ayant en tête l'allure en cloche de la loi binomiale, nous montrons que $(\mathrm{P}(X_n = k))_k$ est croissante jusqu'à x_n puis décroissante ce qui permet bien de conclure.

   Or, les probabilités avec la loi binomiale étant strictement positives, si $1 \leqslant k \leqslant n$:

$$\frac{\mathrm{P}(X_n = k)}{\mathrm{P}(X_n = k-1)} = \frac{\dbinom{k}{n} p^k q^{n-k}}{\dbinom{k-1}{n} p^{k-1} q^{n-k+1}} = \frac{(k-1)!(n-k+1)!}{k!(n-k)!} \cdot \frac{p}{q} = \frac{n-k+1}{k} \cdot \frac{p}{q} = \left(\frac{n+1}{k} - 1\right) \frac{p}{q}$$

   et :

$$\left(\frac{n+1}{k} - 1\right) \frac{p}{q} \geqslant 1 \Longleftrightarrow \frac{n+1}{k} - 1 \geqslant \frac{q}{p} \Longleftrightarrow \frac{n+1}{k} \geqslant \frac{1}{p} \Longleftrightarrow k \leqslant (n+1)p.$$

   Or : $x_n = \lceil np - q \rceil \leqslant np - q + 1 = np + p = (n+1)p$ (l'inégalité est même stricte), donc, pour $1 \leqslant k \leqslant x_n$:

$$\mathrm{P}(X_n = k-1) \leqslant \mathrm{P}(X_n = k).$$

   Et : $x_n + 1 = \lceil np - q \rceil + 1 \geqslant np - q + 1 = (n+1)p$, donc pour $x_n + 1 \leqslant k \leqslant n$:

$$\mathrm{P}(X_n = k-1) \geqslant \mathrm{P}(X_n = k).$$

   Le maximum des $\mathrm{P}(X_n = k)$ est donc atteint pour $k = x_n$ et vaut p_n.

2. En utilisant l'encadrement : $\forall u \in \mathbb{R},\ u - 1 < \lfloor u \rfloor \leqslant u$ et le théorème des gendarmes, on montre aisément : $x_n = \lceil np - q \rceil \underset{n\to+\infty}{\sim} np \underset{n\to+\infty}{\longrightarrow} +\infty$ donc : $\lim\limits_{n\to+\infty} x_n = +\infty$.

   Et comme : $np - q \leqslant x_n < np - q + 1$, on a :

$$n - np + q - 1 = (n+1)q - 1 < n - x_n \leqslant n - np + q = (n+1)q$$

   donc, toujours par le même raisonnement : $n - x_n = \lfloor (n+1)q \rfloor \underset{n\to+\infty}{\sim} (n+1)q \underset{n\to+\infty}{\sim} nq$, donc : $\lim\limits_{n\to+\infty} (n - x_n) = +\infty$.

   Ainsi, en utilisant trois fois la formule de Striling :

$$\begin{aligned}
\sqrt{npq}\,p_n &= \sqrt{npq} \frac{n!}{x_n!(n-x_n)!} p^{x_n} q^{n-x_n} \\
&\underset{n\to+\infty}{\sim} \sqrt{npq} \frac{\sqrt{2\pi n}\left(\dfrac{n}{e}\right)^n}{\sqrt{2\pi x_n}\left(\dfrac{x_n}{e}\right)^{x_n}\sqrt{2\pi(n-x_n)}\left(\dfrac{n-x_n}{e}\right)^{n-x_n}} p^{x_n} q^{n-x_n} \\
&= \sqrt{\frac{n^2 pq}{x_n(n-x_n)}} \frac{n^n p^{x_n} q^{n-x_n}}{\sqrt{2\pi} x_n^{x_n} (n-x_n)^{n-x_n}} \\
&\underset{n\to+\infty}{\sim} \sqrt{\frac{n^2 pq}{np \cdot nq}} \frac{n^n p^{x_n} q^{n-x_n}}{\sqrt{2\pi} x_n^{x_n} (n-x_n)^{n-x_n}}
\end{aligned}$$

   donc : $\sqrt{npq}\,p_n \underset{n\to+\infty}{\sim} \dfrac{n^n p^{x_n} q^{n-x_n}}{\sqrt{2\pi} x_n^{x_n} (n-x_n)^{n-x_n}}$.

3. On suppose : $n > \max\left(\frac{p}{q}, \frac{q}{p}\right)$. Alors : $np - q > 0$, donc : $x_n = \lceil np - q \rceil \geqslant 1$, et : $0 < nq - p = nq + q - 1 = (n+1)q - 1$, donc : $n - x_n = \lfloor (n+1)q \rfloor \geqslant 1$. On peut donc écrire, d'une part,
$$\frac{n^n p^{x_n} q^{n-x_n}}{x_n^{x_n}(n-x_n)^{n-x_n}} = e^{n\ln(n)+x_n\ln(p)+(n-x_n)\ln(q)-x_n\ln(x_n)-(n-x_n)\ln(n-x_n)}.$$
D'autre part (le fait que $0 < x_n < n$ légitime l'utilisation de Ψ) :
$$np\,\Psi\left(\frac{x_n - np}{np}\right) = x_n \ln\left(\frac{x_n}{np}\right) = x_n\ln(x_n) - x_n\ln(n) - x_n\ln(p),$$
et :
$$nq\,\Psi\left(\frac{np - x_n}{nq}\right) = (n-x_n)\ln\left(\frac{n-x_n}{nq}\right) = (n-x_n)\ln(n-x_n) - (n-x_n)\ln(n) - (n-x_n)\ln(q),$$
ce qui permet finalement d'aboutir à l'égalité voulue :
$$\frac{n^n p^{x_n} q^{n-x_n}}{x_n^{x_n}(n-x_n)^{n-x_n}} = e^{-np\Psi\frac{x_n-np}{np} - nq\Psi\frac{np-x_n}{nq}}.$$

4. Un développement limité élémentaire permet de montrer :
$$\Psi(x) = (x+1)\ln(x+1) = (x+1)\left(x - \frac{x^2}{2} + \underset{x\to 0}{o}\left(x^2\right)\right) = x + \frac{x^2}{2} + \underset{x\to 0}{o}\left(x^2\right).$$
Pour l'utiliser dans l'expression de la question précédente, on vérifie que : $\dfrac{x_n - np}{np} \underset{n\to+\infty}{\longrightarrow} 0$, et : $\dfrac{np - x_n}{nq} \underset{n\to+\infty}{\longrightarrow} 0$. Comme n est au voisinage de $+\infty$, on peut supposer : $n > \max\left(\dfrac{p}{q}, \dfrac{q}{p}\right)$. On a par définition :
$$np - q \leqslant x_n \leqslant np - q + 1 = (n+1)p,$$
donc :
$$-\frac{q}{np} \leqslant \frac{x_n - np}{np} \leqslant \frac{p}{np} = \frac{1}{n}$$
donc, par encadrement : $\lim\limits_{n\to+\infty} \dfrac{x_n - np}{np} = 0$. Puis :
$$-\frac{p}{nq} \leqslant \frac{np - x_n}{nq} \leqslant \frac{q}{nq} = \frac{1}{n}$$
donc, par encadrement : $\lim\limits_{n\to+\infty} \dfrac{np - x_n}{nq} = 0$. Alors, d'après le développement limité de Ψ ci-dessus :
$$np\,\Psi\left(\frac{x_n - np}{np}\right) = x_n - np + \frac{(x_n - np)^2}{2np} + \underset{n\to+\infty}{o}\left(\frac{1}{n}\right)$$
(avec $(x_n - np)_{n\geqslant 0}$ bornée, intégré au petit o) et de même
$$nq\,\Psi\left(\frac{np - x_n}{nq}\right) = np - x_n + \frac{(x_n - np)^2}{2nq} + \underset{n\to+\infty}{o}\left(\frac{1}{n}\right)$$
et donc :
$$-np\,\Psi\left(\frac{x_n - np}{np}\right) - nq\,\Psi\left(\frac{np - x_n}{nq}\right) = -\frac{1}{p} + \frac{1}{q}\frac{(x_n - np)^2}{2n} + \underset{n\to+\infty}{o}\left(\frac{1}{n}\right) \underset{n\to+\infty}{\longrightarrow} 0$$
car $(x_n - np)_{n\geqslant 0}$ bornée et $\left(\dfrac{1}{n}\right)_{n\geqslant 1}$ converge vers 0.

Ainsi, en utilisant les deux questions précédentes et la continuité de l'exponentielle en 0, on a : $\lim\limits_{n\to+\infty} \sqrt{npq}\,p_n = \dfrac{1}{\sqrt{2\pi}}$.

Partie B. Convergence en loi.

5. Comme X_n suit une loi binomiale de paramètre (n, p), les valeurs prises par les Y_n sont les $\tau_{n,k}$ pour $k \in [\![0, n]\!]$, et pour un tel k on a : $\mathrm{P}(Y_n = \tau_{n,k}) = \mathrm{P}(X_n = k) = \binom{k}{n} p^k q^{n-k}$.

   De plus, np est l'espérance de X_n et npq sa variance, donc Y_n est la variable aléatoire centrée réduite associée à X_n : on a directement : $\mathrm{E}(Y_n) = 0$, et : $\mathrm{V}(Y_n) = 1$.

6. Comme : $\lim_{n\to+\infty} \frac{1}{\sqrt{npq}} = 0$, on a un rang $N_1 \in \mathbb{N} \setminus \{0\}$ à partir duquel : $\frac{1}{\sqrt{npq}} \leqslant b - a$.

   Comme $\tau_{n,0} = -\sqrt{\frac{np}{q}} \underset{n\to+\infty}{\longrightarrow} -\infty$, on a un rang $N_2 \in \mathbb{N} \setminus \{0\}$ à partir duquel : $\tau_{n,0} \leqslant a$.

   Comme $\tau_{n,n} = \sqrt{\frac{nq}{p}} \underset{n\to+\infty}{\longrightarrow} +\infty$, on a un rang $N_3 \in \mathbb{N} \setminus \{0\}$ à partir duquel : $\tau_{n,n} \geqslant b$.

   Soit $N = \max(N_1, N_2, N_3) \in \mathbb{N} \setminus \{0\}$. Alors pour tout entier $n \geqslant N$, on a : $[a, b] \subseteq [\tau_{n,0}, \tau_{n,n}]$ et : $\frac{1}{\sqrt{npq}} \leqslant b - a$.

7. Par définition et croissance de la partie entière, k_n est une fonction croissante en escalier de subdivision adaptée $\left(\frac{k - np}{\sqrt{npq}}\right)_{k\in\mathbb{Z}} = (\tau_{n,k})_{k\in\mathbb{Z}}$. C'est donc aussi le cas de $e_n : t \mapsto \frac{k_n(t) - np}{\sqrt{npq}}$.

   De plus, par définition de la partie entière, pour tout $t \in \mathbb{R}$, on a : $k_n(t) \leqslant \sqrt{npq}t + np < k_n(t) + 1$ donc :
$$e_n(t) \leqslant t < e_n(t) + \frac{1}{\sqrt{npq}}.$$

   Ainsi, pour $t \in \mathbb{R}$ fixé, pour tout $n \in \mathbb{N} \setminus \{0\}$, on a : $t - \frac{1}{\sqrt{npq}} < e_n(t) \leqslant t$, donc par le théorème d'encadrement : $\lim_{n\to+\infty} e_n(t) = t$: la suite $(e_n)_n$ converge simplement sur $\mathbb{R}$ vers $\mathrm{id}_{\mathbb{R}}$ (et même uniformément, avec l'encadrement précédent).

8. Soit $n \in \mathbb{N} \setminus \{0\}$. On remarque que : $\tau_{n,k_n(a)} = e_n(a)$, et :
$$\tau_{n,k_n(b)+1} = \frac{k_n(b) + 1 - np}{\sqrt{npq}} = e_n(b) + \frac{1}{\sqrt{npq}}.$$

   On a alors :
$$\int_{\tau_{n,k_n(a)}}^{\tau_{n,k_n(b)+1}} e^{-\frac{t^2}{2}} \mathrm{d}t = \int_{e_n(a)}^{e_n(b)+\frac{1}{\sqrt{npq}}} e^{-\frac{t^2}{2}} \mathrm{d}t = \int_0^{e_n(b)+\frac{1}{\sqrt{npq}}} e^{-\frac{t^2}{2}} \mathrm{d}t - \int_0^{e_n(a)} e^{-\frac{t^2}{2}} \mathrm{d}t.$$

   Or : $\lim_{n\to+\infty} e_n(a) = a$, et : $\lim_{n\to+\infty} \left(e_n(b) + \frac{1}{\sqrt{npq}}\right) = b$, et $t \mapsto e^{-\frac{t^2}{2}}$ est continue sur $\mathbb{R}$ donc l'application $x \mapsto \int_0^x e^{-\frac{t^2}{2}} \mathrm{d}t$ qui en est une primitive l'est aussi et :
$$\lim_{n\to+\infty} \int_{\tau_{n,k_n(a)}}^{\tau_{n,k_n(b)+1}} \Phi(t)\mathrm{d}t = \int_0^b \Phi(t)\mathrm{d}t - \int_0^a \Phi(t)\mathrm{d}t = \int_a^b \Phi(t)\mathrm{d}t.$$

   Puis, d'une part :
$$\mathrm{P}(e_n(a) \leqslant Y_n \leqslant e_n(b)) = \sum_{\substack{k\in[\![0,n]\!] \\ e_n(a)\leqslant\frac{k-np}{\sqrt{npq}}\leqslant e_n(b)}} \mathrm{P}(X_n = k) = \sum_{k=k_n(a)}^{k_n(b)} \mathrm{P}(X_n = k)$$

et d'autre part, avec le changement de variable $u = \sqrt{npq}t + np$:

$$\begin{aligned}
\int_{\tau_{n,k_n(a)}}^{\tau_{n,k_n(b)+1}} f_n(t)\mathrm{d}t &= \sqrt{npq}\int_{\tau_{n,k_n(a)}}^{\tau_{n,k_n(b)+1}} \mathrm{P}(Y_n = e_n(t))\mathrm{d}t \\
&= \sqrt{npq}\int_{\tau_{n,k_n(a)}}^{\tau_{n,k_n(b)+1}} \mathrm{P}(X_n = k_n(t))\mathrm{d}t \\
&= \int_{k_n(a)}^{k_n(b)+1} \mathrm{P}(X_n = \lfloor u \rfloor)\mathrm{d}u \qquad \text{(fonction en escalier)} \\
&= \sum_{k=k_n(a)}^{k_n(b)} 1\cdot \mathrm{P}(X_n = k)
\end{aligned}$$

et donc finalement : $\mathrm{P}(e_n(a) \leqslant Y_n \leqslant e_n(b)) = \displaystyle\int_{\tau_{n,k_n(a)}}^{\tau_{n,k_n(b)+1}} f_n(t)\mathrm{d}t$.

9. Soit $n \in \mathbb{N}\setminus\{0\}$ et $k \in [\![1, n-1]\!]$. On remarque que : $e_n(\tau_{n,k}) = \tau_{n,k_n(\tau_{n,k})} = \tau_{n,k}$ car $k \in \mathbb{Z}$. Alors :

$$\begin{aligned}
f_n(\tau_{n,k}) = \sqrt{npq}\mathrm{P}(Y_n = \tau_{n,k}) &= \sqrt{npq}\mathrm{P}(X_n = k) \\
&= \sqrt{npq}\frac{n!}{k!(n-k)!}p^k q^{n-k} \\
&= \sqrt{npq}\frac{\sqrt{2\pi n}\left(\frac{n}{e}\right)^n}{\sqrt{2\pi k}\left(\frac{k}{e}\right)^k\sqrt{2\pi(n-k)}\left(\frac{n-k}{e}\right)^{n-k}}p^k q^{n-k}\frac{1+\varepsilon_n}{(1+\varepsilon_k)(1+\varepsilon_{n-k})}
\end{aligned}$$

et après simplifications : $f_n(\tau_{n,k}) = \sqrt{npq}\sqrt{\dfrac{pqn^2}{k(n-k)}}\dfrac{p^k q^{n-k}}{\left(\frac{k}{n}\right)^k\left(\frac{n-k}{n}\right)^{n-k}}\dfrac{1+\varepsilon_n}{(1+\varepsilon_k)(1+\varepsilon_{n-k})}$.

10. Soit $t \in [a, b]$. On a :

$$np + \sqrt{npq}t - 1 \leqslant k_n(t) \leqslant np + \sqrt{npq}t,$$

donc par encadrement : $k_n(t) \underset{n\to+\infty}{\sim} np$, et :

$$nq - \sqrt{npq}t \leqslant n - k_n(t) \leqslant nq - \sqrt{npq}t + 1,$$

donc par encadrement : $n - k_n(t) \underset{n\to+\infty}{\sim} nq$.

Donc : $k_n(t)(n-k_n(t)) \underset{n\to+\infty}{\sim} n^2pq$, et : $\displaystyle\lim_{n\to+\infty}\sqrt{\frac{pqn^2}{k_n(t)(n-k_n(t))}} = 1$.

Puis, comme : $\displaystyle\lim_{n\to+\infty}\varepsilon_n = 0$ (c'est une conséquence facile de la formule de Stirling), on a : $\displaystyle\lim_{n\to+\infty} k_n(t) = +\infty$, et : $\displaystyle\lim_{n\to+\infty}(n - k_n(t)) = +\infty$ (par exemple avec les équivalents ci-dessus), donc : $\displaystyle\lim_{n\to+\infty}\frac{1+\varepsilon_n}{(1+\varepsilon_k)(1+\varepsilon_{n-k})} = 1$.

11. Soient $n \in \mathbb{N}\setminus\{0\}$ et $k \in [\![0, n]\!]$ tel que : $\max\left(\sqrt{\frac{q}{np}}, \sqrt{\frac{p}{nq}}\right)\cdot|\tau_{n,k}| < 1$, alors on peut reprendre les calculs faits à la question 3 en remplaçant x_n par k, la condition imposée à k permettant d'avoir des quantités strictement positives dans tous les logarithmes, avec : $\frac{k-np}{np} = \sqrt{\frac{q}{np}}\tau_{n,k}$, et : $\frac{np-k}{np} = -\sqrt{\frac{p}{nq}}\tau_{n,k}$:

$$\frac{p^k q^{n-k}}{\left(\frac{k}{n}\right)^k\left(\frac{n-k}{n}\right)^{n-k}} = \frac{n^n p^k q^{n-k}}{k^k(n-k)^{n-k}} = e^{-np\Psi\left(\frac{k-np}{np}\right) - nq\Psi\left(\frac{np-k}{nq}\right)},$$

donc : $\dfrac{p^k q^{n-k}}{\left(\frac{k}{n}\right)^k\left(\frac{n-k}{n}\right)^{n-k}} = e^{-np\Psi\left(\sqrt{\frac{q}{np}}\tau_{n,k}\right) - nq\Psi\left(-\sqrt{\frac{p}{nq}}\tau_{n,k}\right)}$.

12. Comme dans la question 4 avec $k_n(t)$ à la place de x_n, on a bien : $\sqrt{\frac{q}{np}}\tau_{n,k_n(t)} = \sqrt{\frac{q}{np}}e_n(t) \underset{n\to+\infty}{\longrightarrow} 0\cdot t = 0$, et : $-\sqrt{\frac{p}{nq}}\tau_{n,k_n(t)} = -\sqrt{\frac{p}{nq}}e_n(t) \underset{n\to+\infty}{\longrightarrow} 0$ et donc à partir d'un certain rang, la condition de la question précédente est bien assurée, ce qui permet d'écrire :

$$\begin{aligned}
& -np\,\Psi\left(\sqrt{\frac{q}{np}}\tau_{n,k_n(t)}\right) - nq\,\Psi\left(\sqrt{\frac{p}{nq}}\tau_{n,k_n(t)}\right) \\
&= -np\left(\sqrt{\frac{q}{np}}e_n(t) + \frac{qe_n^2(t)}{2np} + \underset{n\to+\infty}{o}\left(\frac{1}{n}\right)\right) - nq\left(-\sqrt{\frac{p}{nq}}e_n(t) + \frac{pe_n^2(t)}{2nq} + \underset{n\to+\infty}{o}\left(\frac{1}{n}\right)\right) \\
&= -\frac{1}{2}(p+q)e_n^2(t) + \underset{n\to+\infty}{o}(1) \\
&\underset{n\to+\infty}{\longrightarrow} -\frac{t^2}{2}
\end{aligned}$$

donc, par continuité de l'exponentielle et avec la question précédente :

$$\lim_{n\to+\infty} \frac{p^{k_n(t)}q^{n-k_n(t)}}{\left(\frac{k_n(t)}{n}\right)^{k_n(t)}\left(\frac{n-k_n(t)}{n}\right)^{n-k_n(t)}} = e^{-t^2/2}.$$

Ensuite, on a déjà remarqué que pour tout $k\in\mathbb{Z}$, on a : $e_n(\tau_{n,k}) = \tau_{n,k}$. Ainsi, e_n est idempotente :

$$e_n(e_n(t)) = e_n(\tau_{n,k_n(t)}) = \tau_{n,k_n(t)} = e_n(t).$$

Donc :

$$f_n(t) = \sqrt{npq}\ \mathrm{P}(Y_n = e_n(t)) = \sqrt{npq}\ \mathrm{P}(Y_n = e_n(e_n(t))) = f_n(e_n(t)) = f_n(\tau_{n,k_n(t)}).$$

Mais on ne peut utiliser la question 9 qu'à condition que : $k_n(t)\in[\![1,n-1]\!]$.
Or comme $\tau_{n,1}\to-\infty$ et $\tau_{n,n}\to+\infty$, on a un rang $N'\in\mathbb{N}\setminus\{0\}$ à partir duquel : $\tau_{n,1} < t < \tau_{n,n}$, donc : $1 < \sqrt{npq}t + np < n$, et : $k_n(t)\in[\![1,n-1]\!]$.
Alors, avec les questions 9, 10 et le calcul de limite qui précède :

$$f_n(t) = f_n(\tau_{n,k_n(t)}) \underset{n\to+\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}}e^{-t^2/2}.$$

On cherche ensuite à appliquer le théorème de convergence dominée :

— pour tout $n\in\mathbb{N}$, l'application f_n est continue par morceaux sur $[a,b]$ en tant que fonction en escalier.

— la suite de fonctions $(f_n)_{n\geqslant1}$ converge simplement sur $[a,b]$ vers $t\mapsto \dfrac{1}{\sqrt{2\pi}}e^{-t^2/2}$ qui est continue sur $[a,b]$.

— pour tout $t\in[a,b]$ et tout $n\in\mathbb{N}\setminus\{0\}$, en utilisant la question 4 :

$$|f_n(t)| = f_n(t) = \sqrt{npq}\ \mathrm{P}(Y_n = e_n(t)) = \sqrt{npq}\mathrm{P}(X_n = k_n(t)) \leqslant \sqrt{npq}p_n\,;$$

or d'après la question 4, la suite $(\sqrt{npq}p_n)_{n\geqslant1}$ converge donc elle est majorée par un réel M donc pour tout $t\in[a,b]$ et $n\in\mathbb{N}\setminus\{0\}$, $|f_n(t)\leqslant\varphi(t)=M$, où φ est une fonction positive, continue par morceaux et intégrable sur $[a,b]$.

D'après le théorème de convergence dominée : $\displaystyle\lim_{n\to+\infty}\int_a^b f_n(t)\mathrm{d}t = \int_a^b \frac{1}{\sqrt{2\pi}}e^{-t^2/2}\mathrm{d}t.$

13. D'après la question 8 :

$$\begin{aligned}
\mathrm{P}(e_n(a) \leqslant Y_n \leqslant e_n(b)) &= \int_{\tau_{n,k_n(a)}}^{\tau_{n,k_n(b)+1}} f_n(t)\mathrm{d}t \\
&= \int_{e_n(a)}^{e_n(b)+\frac{1}{\sqrt{npq}}} f_n(t)\mathrm{d}t \\
&= \int_{e_n(a)}^{a} f_n(t)\mathrm{d}t + \int_a^b f_n(t)\mathrm{d}t + \int_b^{e_n(b)+\frac{1}{\sqrt{npq}}} f_n(t)\mathrm{d}t.
\end{aligned}$$

Or on a déjà : $\lim\limits_{n\to+\infty} \int_a^b f_n(t)\mathrm{d}t = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2}\mathrm{d}t$ d'après la question précédente puis, pour tout $t \in \mathbb{R}$:

$$0 \leqslant f_n(t) = \sqrt{npq}\mathrm{P}(Y_n = e_n(t)) = \sqrt{npq}\mathrm{P}(X_n = k_n(t)) \leqslant \sqrt{npq}p_n \leqslant M$$

comme dans la question précédente, d'où, les bornes étant dans le bon ordre vu la question 7 :

$$0 \leqslant \int_{e_n(a)}^{a} f_n(t)\mathrm{d}t \leqslant M(a - e_n(a)) \underset{n\to+\infty}{\longrightarrow} 0$$

donc : $\lim\limits_{n\to+\infty} \int_{e_n(a)}^{a} f_n(t)\mathrm{d}t = 0$, et de même :

$$0 \leqslant \int_b^{e_n(b)+\frac{1}{\sqrt{npq}}} f_n(t)\mathrm{d}t \leqslant Me_n(b) + \frac{1}{\sqrt{npq}} - b \underset{n\to+\infty}{\longrightarrow} 0$$

donc : $\lim\limits_{n\to+\infty} \int_b^{e_n(b)+\frac{1}{\sqrt{npq}}} f_n(t)\mathrm{d}t = 0$.

Finalement : $\lim\limits_{n\to+\infty} \mathrm{P}(e_n(a) \leqslant Y_n \leqslant e_n(b)) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2}\mathrm{d}t$.

Ensuite, on a : $e_n(a) \leqslant a$ et $e_n(b) \leqslant b$ et comme : $\lim\limits_{n\to+\infty} e_n(b) = b > a$, on a un rang N_0 à partir duquel : $e_n(b) > a$, ce qui donne : $e_n(a) \leqslant a < e_n(b) \leqslant b$. Donc, si $n \geqslant N_0$:

$$\mathrm{P}(a \leqslant Y_n \leqslant b) = \mathrm{P}(e_n(a) \leqslant Y_n \leqslant e_n(b)) - \mathrm{P}(e_n(a) \leqslant Y_n < a) + \mathrm{P}(e_n(b) < Y \leqslant b).$$

Mais avec la question 7 :

$$\begin{aligned}
0 \leqslant \mathrm{P}(e_n(a) \leqslant Y_n < a) &\leqslant \mathrm{P}\left(e_n(a) \leqslant Y_n < e_n(a) + \frac{1}{\sqrt{npq}}\right) \\
&= \mathrm{P}(k_n(a) \leqslant X_n < k_n(a) + 1) \\
&= \mathrm{P}(X_n = k_n(a)) && \text{(car } X_n \text{ à valeurs entières)} \\
&= \mathrm{P}(Y_n = e_n(a)) \\
&= \frac{f_n(a)}{\sqrt{npq}} \\
&\underset{n\to+\infty}{\longrightarrow} 0 \cdot \Phi(a) = 0,
\end{aligned}$$

donc : $\lim\limits_{n\to+\infty} \mathrm{P}(e_n(a) \leqslant Y_n < a) = 0$ et de la même manière : $\lim\limits_{n\to+\infty} \mathrm{P}(e_n(b) < Y \leqslant b) = 0$.

Ainsi : $\lim\limits_{n\to+\infty} \mathrm{P}(a \leqslant Y_n \leqslant b) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2}\mathrm{d}t$.

Partie C. Applications.

14. Soit $T>0$. On applique l'inégalité de Bienaymé-Tchebychev à Y_n qui est centrée et réduite, et qui dit que : $\mathrm{P}(|Y_n|\geqslant T)\leqslant \dfrac{\mathrm{V}(Y_n)}{T^2}=\dfrac{1}{T^2}$, donc :
$$\mathrm{P}(-T\leqslant Y_n\leqslant T)\geqslant \mathrm{P}(-T<Y_n<T)=\mathrm{P}(|Y_n|<T)\geqslant 1-\frac{1}{T^2},$$
puis avec la question précédente en faisant tendre n vers $+\infty$ avec la question 13 :
$$\frac{1}{\sqrt{2\pi}}\int_{-T}^{T}e^{-t^2/2}\mathrm{d}t\geqslant 1-\frac{1}{T^2}.$$
Comme par ailleurs : $\mathrm{P}(-T\leqslant Y_n\leqslant T)\leqslant 1$, en faisant $n\to+\infty$: $\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-T}^{T}e^{-t^2/2}\mathrm{d}t\leqslant 1$.
Enfin, en faisant tendre T vers $+\infty$ dans l'encadrement : $1-\dfrac{1}{T^2}\leqslant\dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-T}^{T}e^{-t^2/2}\leqslant 1$ (la question de l'existence de l'intégrale ne se posant pas puisque l'intégrande est positif), on trouve :
$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty}e^{-t^2/2}\mathrm{d}t=1,$$
donc : $\displaystyle\int_{-\infty}^{+\infty}e^{-t^2/2}\mathrm{d}t=\sqrt{2\pi}$.

15. On montre que : $\displaystyle\lim_{n\to+\infty}\mathrm{P}(Y_n\leqslant b)=\int_{-\infty}^{b}e^{-t^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}}$. Soit $\varepsilon>0$. On a pour tout $a<b$, et tout $n\in\mathbb{N}\setminus\{0\}$:
$$\begin{aligned}\left|\mathrm{P}(Y_n\leqslant b)-\int_{-\infty}^{b}e^{-t^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}}\right|&=\left|\mathrm{P}(a\leqslant Y_n\leqslant b)-\int_{a}^{b}e^{-t^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}}-\mathrm{P}(Y_n<a)-\int_{-\infty}^{a}e^{-t^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}}\right|\\&\leqslant\left|\mathrm{P}(a\leqslant Y_n\leqslant b)-\int_{a}^{b}e^{-t^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}}\right|+\mathrm{P}(Y_n<a)+\int_{-\infty}^{a}e^{-t^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}}.\end{aligned}$$
Comme : $\displaystyle\int_{-\infty}^{a}e^{-t^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}}\underset{a\to-\infty}{\longrightarrow}0$, on a $A\in\mathbb{R}$ tel que si $a\leqslant A$, alors : $\displaystyle\int_{-\infty}^{a}e^{-t^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}}\leqslant\frac{\varepsilon}{3}$. Puis, avec $a<0$:
$$\mathrm{P}(Y_n<a)\leqslant\mathrm{P}(Y_n\leqslant a)\leqslant\mathrm{P}(|Y_n|\geqslant -a)=\mathrm{P}\left(Y_n^2\geqslant a^2\right)\leqslant\frac{\mathrm{E}\left(Y_n^2\right)}{a^2}$$
d'après l'inégalité de Markov. Or Y_n est centrée et réduite donc : $1=\mathrm{V}(Y_n)=\mathrm{E}(Y_n^2)-\mathrm{E}(Y_n)^2=\mathrm{E}(Y_n^2)$ donc :
$$\mathrm{P}(Y_n<a)\leqslant\frac{1}{a^2}.$$
Comme : $\dfrac{1}{a^2}\underset{a\to-\infty}{\longrightarrow}0$, on a $A'\in\mathbb{R}_*^-$ tel que si $a\leqslant A'$, alors : $\mathrm{P}(Y_n<a)\leqslant\dfrac{\varepsilon}{3}$.
On fixe alors $a\leqslant\min(A,A')$. Pour ce a, on a d'après la question 13, un rang N à partir duquel :
$$\left|\mathrm{P}(a\leqslant Y_n\leqslant b)-\int_{a}^{b}e^{-t^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}}\right|\leqslant\frac{\varepsilon}{3},$$
et alors pour tout $n\geqslant N$, $\left|\mathrm{P}(Y_n\leqslant b)-\displaystyle\int_{-\infty}^{b}e^{-t^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}}\right|\leqslant\varepsilon$.
On a montré : $\displaystyle\lim_{n\to+\infty}\mathrm{P}(Y_n\leqslant b)=\int_{-\infty}^{b}e^{-t^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}}$.
On obtient de manière analogue : $\displaystyle\lim_{n\to+\infty}\mathrm{P}(Y_n\geqslant a)=\int_{a}^{+\infty}e^{-t^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}}$.

Partie D. Généralisation.

16. L'application $\varphi : \mathbb{R} \to \mathbb{R}$ est de classe $\mathcal{C}^1$ et φ' ne s'annule pas, donc par continuité de φ', celle-ci garde un signe constant et φ est strictement monotone et continue, induisant une bijection de $\mathbb{R}$ dans l'intervalle $\varphi(\mathbb{R})$.
Si de plus : $\varphi(\mathbb{R}) = \mathbb{R}$ et, par exemple, φ est croissante, alors avec : $-\infty \leqslant \alpha \leqslant \beta \leqslant +\infty$, en notant $-\infty = \varphi(-\infty)$ et $+\infty = \varphi(+\infty)$:
$$\mathrm{P}(\alpha \leqslant Z_n \leqslant \beta) = \mathrm{P}(\varphi^{-1}(\alpha) \leqslant Y_n \leqslant \varphi^{-1}(\beta)) \underset{n\to+\infty}{\longrightarrow} \int_{\varphi^{-1}(\alpha)}^{\varphi^{-1}(\beta)} \Phi(t)\mathrm{d}t = \int_\alpha^\beta \Psi(u)\mathrm{d}u,$$
avec $\Phi : t \mapsto \dfrac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}$ et $\Psi = \dfrac{\Phi\circ\varphi^{-1}}{\varphi^{-1\prime}} = \dfrac{\Phi\circ\varphi^{-1}}{\varphi'\circ\varphi^{-1}}$, *via* le changement de variable de classe C^1 et strictement croissant $u = \varphi(t)$. Ensuite :
    — le cas où α et β sont finis distincts vient de la question 13 ;
    — celui où ils sont distincts et éventuellement infinis vient de la question 15,
    — celui où ils sont égaux finis vient aussi de la question 15 en remarquant que :
$$\begin{aligned}\mathrm{P}(Y_n = a) = \mathrm{P}((Y_n \leqslant a) \cap (Y_n \geqslant a)) &= \mathrm{P}(Y_n \leqslant a) + \mathrm{P}(Y_n \geqslant a) - \mathrm{P}(\Omega)\\ &\underset{n\to+\infty}{\longrightarrow} \int_{-\infty}^a \Phi(t)\mathrm{d}t + \int_a^{+\infty} \Phi(t)\mathrm{d}t - 1 = 0\,;\end{aligned}$$
    — dans le dernier cas où α et β sont égaux et infinis, les probabilités et l'intégrale sont nulles donc le résultat reste valable.
Dans le cas où φ est décroissante, on a : $\mathrm{P}(\alpha \leqslant Z_n \leqslant \beta) = \mathrm{P}(\varphi^{-1}(\beta) \leqslant Y_n \leqslant \varphi^{-1}(\alpha))$, ce qui permet d'aboutir au résultat avec : $\Psi = -\dfrac{\Phi\circ\varphi^{-1}}{\varphi^{-1\prime}} = \dfrac{\Phi\circ\varphi^{-1}}{|\varphi'\circ\varphi^{-1}|}$.
Enfin, si : $\varphi(\mathbb{R}) \neq \mathbb{R}$, alors le reste reste valable seulement si α, β sont dans l'intervalle $\varphi(\mathbb{R})$ ou une de ses éventuelles bornes infinies.