Corrigé et barème – Sommes de variables aléatoires et loi des grands nombres — niveau ★★★

Durée : 120 min – Calculatrice autorisée – Sur 20 points · Spécialité Mathématiques Terminale

Exercice 1 – Automatismes *(3 pts)*

a) D'après la formule de König-Huygens, $E(X^2) = V(X) + E(X)^2 = 5 + 16 = 21$. Par linéarité, $E\big((X-1)(X+3)\big) = E(X^2 + 2X - 3) = E(X^2) + 2E(X) - 3 = 21 + 8 - 3 = 26$. (1 pt) (1)

b) $V(Y) = a^2 \times 3^2 = 9a^2 = 1$ avec $a > 0$ donne $a = \dfrac{1}{3}$. Puis $E(Y) = \dfrac{1}{3} \times 12 + b = 4 + b = 0$, donc $b = -4$: $Y = \dfrac{X - 12}{3}$. (1 pt : 0,5 + 0,5) (1)

c) On veut $P(|M_n - \mu| \geqslant 0{,}1) \leqslant 0{,}01$; il suffit que $\dfrac{2}{n \times 0{,}1^2} \leqslant 0{,}01$, soit $n \geqslant \dfrac{2}{0{,}01 \times 0{,}01} = 20\,000$. Alors $P(|M_n - \mu| < 0{,}1) \geqslant 0{,}99$: $n = 20\,000$. (1 pt) (1)

Astuce — La formule de König-Huygens se lit dans les deux sens : connaissant $E(X)$ et $V(X)$, on retrouve $E(X^2) = V(X) + E(X)^2$, utile pour l'espérance d'un polynôme en X.

Exercice 2 – Démontrer — Une inégalité qu'on ne peut pas améliorer *(4 pts)*

a) Si $|X - \mu| \geqslant \delta$, alors $Y = \delta^2 \leqslant (X-\mu)^2$; sinon $Y = 0 \leqslant (X-\mu)^2$. La variable $(X-\mu)^2 - Y$ ne prend donc que des valeurs positives, et son espérance est positive : par linéarité, $E\big((X-\mu)^2\big) - E(Y) \geqslant 0$, soit $E(Y) \leqslant V(X)$. Or Y vaut δ^2 avec la probabilité $P(|X-\mu| \geqslant \delta)$ et 0 sinon, donc $E(Y) = \delta^2\, P(|X-\mu| \geqslant \delta)$. Ainsi $\delta^2\, P(|X-\mu| \geqslant \delta) \leqslant V(X)$, et l'on divise par $\delta^2 > 0$. (1 pt : 0,5 pour la comparaison des variables, 0,5 pour la conclusion) (1)

b) Par linéarité, $E(Z) = a\big(E(B_1) - E(B_2)\big) = a(r - r) = 0$. B_1 et B_2 étant indépendantes, $V(B_1 - B_2) = V(B_1) + (-1)^2 V(B_2) = 2r(1-r)$, donc $V(Z) = a^2 \times 2r(1-r) = 2a^2 r(1-r)$. (1 pt : 0,5 + 0,5) (1)

c) $Z = a$ si $B_1 = 1$ et $B_2 = 0$, de probabilité $r(1-r)$ par indépendance ; $Z = -a$ si $B_1 = 0$ et $B_2 = 1$, de probabilité $(1-r)r$; $Z = 0$ sinon, de probabilité $1 - 2r(1-r)$. Donc $P(|Z| \geqslant a) = 2r(1-r)$, et le majorant vaut $\dfrac{V(Z)}{a^2} = \dfrac{2a^2 r(1-r)}{a^2} = 2r(1-r)$: il y a égalité. Si l'on ne connaît que l'espérance et la variance, l'inégalité de Bienaymé-Tchebychev ne peut donc pas être améliorée : son majorant est atteint par certaines lois. (1 pt : 0,5 pour la loi, 0,5 pour l'égalité et la conclusion) (1)

d) Pour 0,3, non : l'inégalité impose $P(|W| \geqslant 2) \leqslant \dfrac{1}{2^2} = 0{,}25$. Pour 0,25, oui : on prend $W = 2(B_1 - B_2)$ avec $2 \times 2^2 \times r(1-r) = 1$, soit $r^2 - r + \dfrac{1}{8} = 0$, d'où $r = \dfrac{2 - \sqrt{2}}{4} \approx 0{,}146$, qui est bien dans $]0\,;1[$. Alors $E(W) = 0$, $V(W) = 1$ et $P(|W| \geqslant 2) = 2r(1-r) = \dfrac{1}{4}$. (1 pt : 0,5 pour l'impossibilité, 0,5 pour la construction) (1)

Méthode — Pour montrer qu'une inégalité ne peut pas être améliorée, on exhibe un cas d'égalité : ici une loi à trois valeurs symétriques dont tout l'écart est concentré à la distance δ.

Exercice 3 – Problème — Estimer π au hasard *(5 pts)*

a) `c` compte les points tombés dans le quart de disque parmi les n points tirés : $\dfrac{c}{n}$ est la fréquence F_n de cet événement, de probabilité $\dfrac{\pi}{4}$. D'après la loi faible des grands nombres, F_n se rapproche de $\dfrac{\pi}{4}$ quand n grandit, donc la valeur renvoyée, $4F_n$, se rapproche de π. (1 pt) (1)

b) F_n est la moyenne d'un échantillon de taille n de la loi de Bernoulli de paramètre p : $E(F_n) = p$ et $V(F_n) = \dfrac{p(1-p)}{n}$. Donc $E(Q_n) = 4p = \pi$ et $V(Q_n) = 4^2 \times \dfrac{p(1-p)}{n} = \dfrac{16p(1-p)}{n}$. (1 pt) (1)

c) Comme $\pi > 3$, $p = \dfrac{\pi}{4} > \dfrac{3}{4}$, donc $p - \dfrac{1}{2} > \dfrac{1}{4} > 0$ et $\left(p - \dfrac{1}{2}\right)^2 > \dfrac{1}{16}$. Or $p(1-p) = \dfrac{1}{4} - \left(p - \dfrac{1}{2}\right)^2$, d'où $p(1-p) < \dfrac{1}{4} - \dfrac{1}{16} = \dfrac{3}{16}$ et $V(Q_n) = \dfrac{16p(1-p)}{n} < \dfrac{3}{n}$. D'après l'inégalité de Bienaymé-Tchebychev appliquée à Q_n, d'espérance π : $P(|Q_n - \pi| \geqslant \delta) \leqslant \dfrac{V(Q_n)}{\delta^2} \leqslant \dfrac{3}{n\delta^2}$. (1 pt : 0,5 pour $p(1-p) < \dfrac{3}{16}$, 0,5 pour la conclusion) (1)

d) On veut $P(|Q_n - \pi| \geqslant 0{,}01) \leqslant 0{,}05$; il suffit que $\dfrac{3}{n \times 0{,}01^2} \leqslant 0{,}05$, soit $n \geqslant \dfrac{3}{0{,}05 \times 0{,}0001} = 600\,000$ points. Avec la seule majoration $p(1-p) \leqslant \dfrac{1}{4}$, on aurait trouvé 800 000 points : savoir que $\pi > 3$ en fait économiser un quart. (1 pt) (1)

e) On veut $\dfrac{3}{10\,000\,\delta^2} \leqslant 0{,}1$, soit $\delta^2 \geqslant 0{,}003$, c'est-à-dire $\delta \geqslant \sqrt{0{,}003} \approx 0{,}0548$. Avec 10 000 points, on garantit seulement un écart inférieur à 0,055 avec une probabilité d'au moins 0,9 : à peine le premier chiffre après la virgule. La précision garantie varie comme $\dfrac{1}{\sqrt{n}}$: pour la diviser par 10, il faut 100 fois plus de points. La méthode est simple mais lente. (1 pt : 0,5 pour δ, 0,5 pour le commentaire) (1)

Piège — Multiplier une fréquence par 4 multiplie sa variance par 16, pas par 4 : l'erreur commise sur π est quatre fois celle commise sur $\dfrac{\pi}{4}$.

Exercice 4 – Problème — Surréservation *(4 pts)*

a) $N_n = X_1 + \cdots + X_n$, où $X_i = 1$ si le i-ième passager se présente et $X_i = 0$ sinon ; les X_i sont indépendantes et suivent la loi de Bernoulli de paramètre 0,9. Donc N_n suit la loi $\mathcal{B}(n\,;0{,}9)$, avec $E(N_n) = 0{,}9n$ par linéarité et $V(N_n) = n \times 0{,}9 \times 0{,}1 = 0{,}09n$ par indépendance. (1 pt) (1)

b) Pour $201 \leqslant n \leqslant 223$, on a bien $0{,}9n < 201$. Si $N_n \geqslant 201$, alors $N_n - 0{,}9n \geqslant 201 - 0{,}9n > 0$, donc $|N_n - 0{,}9n| \geqslant 201 - 0{,}9n$. D'après l'inégalité de Bienaymé-Tchebychev, $P(N_n \geqslant 201) \leqslant g(n) = \dfrac{0{,}09n}{(201 - 0{,}9n)^2}$. (1 pt : 0,5 pour l'inclusion, 0,5 pour le majorant) (1)

c) Sur $\{201, \ldots, 223\}$, g est croissante : son numérateur croît et son dénominateur, positif, décroît. $g(202) = \dfrac{18{,}18}{19{,}2^2} = \dfrac{18{,}18}{368{,}64} \approx 0{,}0493 \leqslant 0{,}05$ et $g(203) = \dfrac{18{,}27}{18{,}3^2} = \dfrac{18{,}27}{334{,}89} \approx 0{,}0546 > 0{,}05$. Pour $n \geqslant 224$, $E(N_n) \geqslant 201{,}6$ et l'inégalité ne fournit plus de majorant de ce type. La méthode garantit donc la contrainte jusqu'à 202 réservations. (1 pt : 0,5 pour la monotonie, 0,5 pour les deux valeurs et la conclusion) (1)

d) $P(N_{215} \geqslant 201) = 1 - P(N_{215} \leqslant 200) \approx 0{,}0496 \leqslant 0{,}05$ et $P(N_{216} \geqslant 201) \approx 0{,}0782 > 0{,}05$. Comme cette probabilité augmente avec n, la compagnie peut en réalité accepter 215 réservations, contre 202 avec c). L'inégalité de Bienaymé-Tchebychev n'utilise que l'espérance et la variance et vaut pour toutes les lois : elle est vraie mais très prudente. Quand la loi est connue, le calcul exact autorise 13 réservations de plus. (1 pt : 0,5 pour les calculs, 0,5 pour la comparaison) (1)

Erreur fréquente — Croire que le nombre trouvé avec Bienaymé-Tchebychev est le meilleur possible : il est seulement suffisant, et le calcul exact avec la loi binomiale autorise souvent beaucoup plus.

Exercice 5 – Pour aller plus loin — Combiner deux instruments *(4 pts)*

a) Par linéarité, $E(Z_t) = t\mu + (1-t)\mu = \mu$. Comme X et Y sont indépendantes, $V(Z_t) = t^2 V(X) + (1-t)^2 V(Y) = t^2 a + (1-t)^2 b$. (1 pt) (1)

b) $V(Z_t) = (a+b)t^2 - 2bt + b$ est un trinôme en t de coefficient $a+b > 0$: il est minimal en $t_0 = \dfrac{2b}{2(a+b)} = \dfrac{b}{a+b}$. Alors $1 - t_0 = \dfrac{a}{a+b}$ et $V(Z_{t_0}) = \dfrac{b^2 a + a^2 b}{(a+b)^2} = \dfrac{ab(a+b)}{(a+b)^2} = \dfrac{ab}{a+b}$. Comme $\dfrac{ab}{a+b} = a \times \dfrac{b}{a+b} < a$ et, de même, $\dfrac{ab}{a+b} < b$, la combinaison optimale est plus précise que chacun des deux instruments. (1 pt : 0,5 pour t_0, 0,5 pour la valeur et la comparaison) (1)

c) $t_0 = \dfrac{12}{16} = 0{,}75$ et $V(Z_{t_0}) = \dfrac{4 \times 12}{16} = 3$. $V\left(Z_{\frac{1}{2}}\right) = \dfrac{1}{4} \times 4 + \dfrac{1}{4} \times 12 = 4$: c'est la variance du premier instrument seul, donc la moyenne simple n'apporte rien par rapport au meilleur instrument ; il faut donner plus de poids au plus précis. D'après l'inégalité de Bienaymé-Tchebychev, $P(|Z_{t_0} - \mu| \geqslant 3) \leqslant \dfrac{3}{3^2} = \dfrac{1}{3}$. (1 pt) (1)

d) $\overline{X}_k$ et $\overline{Y}_k$ sont indépendantes, d'espérance μ et de variances $\dfrac{4}{k}$ et $\dfrac{12}{k}$. D'après b), leur combinaison optimale a pour variance $\dfrac{\frac{4}{k} \times \frac{12}{k}}{\frac{4}{k} + \frac{12}{k}} = \dfrac{48/k^2}{16/k} = \dfrac{3}{k}$. On veut $\dfrac{3/k}{0{,}5^2} = \dfrac{12}{k} \leqslant 0{,}05$, soit $k \geqslant 240$: il faut 240 mesures avec chaque instrument. (1 pt) (1)

Astuce — Pour combiner deux mesures sans biais, on pondère chacune par l'inverse de sa variance : $t_0 = \dfrac{b}{a+b}$ donne le plus grand poids à l'instrument de plus petite variance.

Ce document est la propriété intellectuelle de Ben, professeur particulier (benprofparticulier.fr), protégée par le Code de la propriété intellectuelle (art. L. 111-1 et L. 122-4). Il est mis à disposition pour un usage personnel de révision. Toute reproduction, diffusion, mise en ligne ou revente, totale ou partielle, sans autorisation écrite est interdite et constitue une contrefaçon (art. L. 335-2).