Mathématiques 2

Présentation du sujet

Le sujet propose l'étude d'une suite particulière de fonctions qui converge uniformément sur $\mathbb{R}$ vers la fonction gaussienne $t \mapsto \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}$. On en déduit deux applications : une démonstration de la première version du théorème central limite (avec des variables aléatoires de Bernoulli) et une démonstration d'une inégalité de concentration. Les questions sont réparties en trois parties décrites ci-après.

La partie I a pour but de donner une démonstration de l'égalité classique $\int_{-\infty}^{+\infty} \mathrm{e}^{-t^2/2}\,\mathrm{d}t = \sqrt{2\pi}$ (*via* une utilisation adéquate du théorème de la convergence dominée) ainsi que celle de l'équivalent

$$\int_x^{+\infty} \mathrm{e}^{-t^2/2}\frac{\mathrm{d}t}{\sqrt{2\pi}} \underset{x\to+\infty}{\sim} \frac{\mathrm{e}^{-x^2/2}}{x\sqrt{2\pi}}.$$

On y montre également une inégalité probabiliste qualifiée d'inégalité maximale qui énonce dans les grandes lignes que borner toutes les sommes partielles $\sum_{i=1}^{p} Z_i$ d'une somme de n variables aléatoires indépendantes Z_i est essentiellement équivalent à borner la somme totale $\sum_{i=1}^{n} Z_i$.

La partie II définit une suite de fonctions $B_n : \mathbb{R} \to \mathbb{R}$ qui sont constantes par morceaux et étudie la convergence uniforme de (B_n) sur $\mathbb{R}$ vers la fonction gaussienne $t \mapsto \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-t^2/2}$.

Enfin la partie III s'attelle à prouver les applications énoncées ci-dessus : théorème central limite et inégalité de concentration (appelée dans le sujet « critère de tension »).

Analyse globale des résultats

Le sujet couvre un spectre assez large des connaissances du programme : analyse, probabilités, intégration, comportement asymptotique. Concernant la prestation des candidats, le bilan est globalement mitigé.

Les questions plutôt élémentaires (voire reposant sur une connaissance basique du cours) ont été bien traitées bien que, comparativement aux années précédentes, les questions de nature asymptotique et probabilistes sont moins bien maîtrisées. Le jury espère que cela reste exceptionnel et que cela n'est pas un signe de baisse de niveau sur ces points du programme.

Commentaires sur les réponses apportées et conseils aux futurs candidats

La partie I a naturellement été abordée dans la totalité des copies tout comme la partie II.

La partie I est globalement bien réussie (hormis la partie probabiliste I.C). On note cependant beaucoup de lacunes sur les arguments qui concernent des comportements asymptotiques (et donc qui nécessitent une maitrise des symboles O et $\sim$ qui dépassent la simple connaissance de leur définition). En pratique, montrer un équivalent $I_n \sim K_n$ revient presque systématiquement à prouver la limite $\frac{I_n}{K_n} \xrightarrow[n\to+\infty]{} 1$ et cela semble inconnu de beaucoup trop de candidats.

La partie II a été moyennement réussie. De nouveau, ce sont souvent les questions d'étude asymptotique qui ont posé problème. La question **Q22** n'a quasiment jamais été réussie (ce qui est compréhensible vu sa difficulté). En revanche, le jury est un peu déçu concernant la question **Q23** qui nécessite de connaitre le développement limité $\frac{1}{1+z} = 1 + O(z)$ au voisinage de $z = 0$.

La partie III, ayant trait aux applications, n'a été abordée que dans 33 % des copies environ. Bien que cette partie ne contienne que six questions, ce pourcentage est assez faible par comparaison aux années précédentes concernant une dernière partie. La raison est sans doute la difficulté des questions (trois questions, voire quatre, sont relativement délicates).

On ne commente ci-après que la plupart des questions majoritairement traitées (celles de la partie III ont été très peu abordées).

Q1 et **Q2**. Ces questions ont globalement été bien faites. On note que de nombreuses copies oublient que c'est la continuité (ou continuité par morceaux) de l'intégrande qui assure que la borne $+\infty$ est le seul problème d'intégration. Signalons que le calcul de l'intégrale $\int_0^{+\infty} \frac{\mathrm{d}t}{1+t^2}$ a parfois posé quelques soucis (ou bien fausse primitive ou bien mauvaise valeur de $\lim_{t\to+\infty} \arctan t$).

Q3. On peut soit minorer $1+t^2 \geqslant 2t$ soit $1+t^2 \geqslant 1+t$. Signalons que l'inégalité $1+t^2 \geqslant 2t$, trivialement équivalente à $(t-1)^2 \geqslant 0$, a parfois été démontrée géométriquement en invoquant la position de la tangente au point $(1,2)$ du graphe de la fonction convexe $t \mapsto 1+t^2$.

Cette question a souvent fait l'objet d'une confusion assez grave : on souhaite comparer, pour $n \to +\infty$, une intégrale (dépendante de n) avec $\frac{1}{n2^n}$, or certaines copies ont comparé l'intégrande avec $t \to +\infty$ puis ont intégré. Cela fait perdre toutes les informations de l'intégrale et ne peut pas aboutir.

Signalons au passage que, *stricto sensu*, la question nécessite de majorer $|\int_0^1 \frac{1}{(1+t^2)^n}\,\mathrm{d}t|$ et qu'une mention de la positivité de l'intégrale est bienvenue. Le jury n'a cependant pas pénalisé ce défaut d'inattention mineur.

Q4. Dans de nombreuses copies, l'équivalence $I_n \sim K_n$ est reformulée en $\lim_{n\to+\infty} I_n - K_n = 0$ ou $I_n - K_n = O(I_n)$, ce qui est faux.

Q5. Cette question a été très bien réussie alors qu'elle ne comporte aucune indication. En effet, la formule $K_n = K_{n+1} + \frac{1}{2n} K_n$ s'obtient grâce à une intégration par parties. Signalons cependant que cette intégration par parties doit être effectuée sur un intervalle non borné et le calcul nécessite quelques explications : par passage à des bornes finies ou bien, comme autorisé dans le programme, par légitimité des limites dans le crochet (le programme assure en effet que *l'existence des limites du produit fg aux bornes de l'intervalle assure que les intégrales de fg' et de $f'g$ sont de même nature*).

Q6. Question de synthèse moyennement réussie. Dans la plupart des copies, les questions précédentes permettent une bonne compréhension de la question : il suffit d'obtenir un équivalent de (K_n), puis la relation de récurrence précédente amène facilement à la formule

$$K_n = \frac{\prod_{s=1}^{n-1} 2s-1}{\prod_{s=1}^{n-1} 2s} \underbrace{K_1}_{=\pi/2}.$$

Mentionnons que le jury n'a pas jugé utile d'exiger la formalisation d'une récurrence aussi immédiate (avec mise en évidence d'une hypothèse de récurrence, d'une preuve d'initialisation et d'une preuve d'hérédité). En revanche, la suite a parfois posé des problèmes. L'idée est alors de faire intervenir des factorielles puis la formule de Stirling :

$$K_n = \frac{(2n-2)!}{\left((n-1)!2^{n-1}\right)^2}\frac{\pi}{2} \sim \frac{(2n-2)^{2n-2}\mathrm{e}^{-2n+2}\sqrt{2\pi(2n-2)}}{\left((n-1)^{n-1}\mathrm{e}^{-n+1}\sqrt{2\pi(n-1)}2^{n-1}\right)^2}\frac{\pi}{2} \sim \frac{\sqrt{2}}{\sqrt{2\pi}\sqrt{n}}\frac{\pi}{2}$$

et donc $K_n \sim \frac{\sqrt{\pi}}{2\sqrt{n}}$. Signalons que les calculs sont légèrement plus simples en cherchant un équivalent de K_{n+1}. Certaines copies ont d'ailleurs confondu l'expression de K_{n+1} avec celle de K_n mais trouvent le bon équivalent par pure chance (sans remarquer que l'on a bien $K_{n+1} \sim K_n$ à priori grâce à la formule de récurrence ou à posteriori grâce à l'expression finale de l'équivalent).

Q7. Question très bien réussie. L'explicitation du changement de variable linéaire (ou affine) suffit largement pour convaincre le jury.

Q8. Le théorème de la convergence dominée est à priori attendu.

- Concernant l'hypothèse de limite simple, une preuve rigoureuse de la limite $\lim_{n\to+\infty} \frac{1}{(1+u^2/n)^n} = \mathrm{e}^{-u^2}$ est attendue. Ajoutons que l'intégrale $\int_0^{\sqrt{n}} \frac{1}{(1+u^2/n)^n}\,\mathrm{d}u$ doit préalablement être reformulée $\int_0^{+\infty} f_n(u)\,\mathrm{d}u$ avec des bornes fixes avant d'espérer appliquer le théorème de la convergence dominée.
- Le jury a accordé des points pour les copies ayant tenté de prouver la majoration (en réalité fausse) $\frac{1}{(1+u^2/n)^n} \leqslant \mathrm{e}^{-u^2}$ car cela constitue un angle d'attaque totalement naturel pour valider l'hypothèse de domination. Seules les meilleures copies ont pu résoudre cette question par exemple *via* la majoration (obtenue grâce au binôme de Newton)

$$\frac{1}{(1+u^2/n)^n} = \frac{1}{\sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\frac{u^{2k}}{n^k}} = \frac{1}{1+\frac{nu^2}{n} + \text{termes positifs}} \leqslant \frac{1}{1+u^2}.$$

  Mentionnons également une erreur souvent vue dans de très bonnes copies. En vu de démontrer une formule de la forme $\lim_{n\to+\infty} \int_0^{+\infty} f_n(u)\,\mathrm{d}u = \int_0^{+\infty} \lim_{n\to+\infty} f_n(u)\,\mathrm{d}u$ avec le théorème de la convergence dominée, l'hypothèse de domination s'écrit $|f_n(u)| \leqslant g(u)$ avec $g : [0,+\infty[\to \mathbb{R}$ intégrable. Or certaines copies ont vérifié une hypothèse plus faible :

$$\forall u \in [0,+\infty[\qquad \exists n_u \in \mathbb{N} \qquad \forall n \geqslant n_u \qquad |f_n(u)| \leqslant g(u).$$

  Il est plausible qu'une mauvaise maitrise des quantificateurs soit à l'origine de cette confusion.

Q11. Question très difficile et globalement ouverte quant au choix de la fonction auxiliaire à proposer. Un exemple de choix adéquat de fonction auxiliaire est

$$\Psi : x \mapsto \int_x^{+\infty} \varphi(t)\,\mathrm{d}t - \frac{x}{x^2+1}\varphi(x).$$

Q13. Question moyennement réussie. Dans la quasi-totalité des copies ayant abordé cette question, la bonne réponse $A = \bigcup_{p=1}^{n} A_p$ est trouvée. En revanche, la preuve est souvent absente voire très sommaire. Une réponse satisfaisante consiste par exemple à considérer, pour tout $\omega \in A$, le plus petit $p \in \{1, ..., n\}$ vérifiant $|R_p(\omega)| \geqslant 3x$.

Q17. Le point clé est de montrer l'inégalité

$$\max_{1\leqslant p\leqslant n} \mathbb{P}\big(\{|R_n - R_p| > 2x\}\big) \leqslant 2 \max_{1\leqslant p\leqslant n} \mathbb{P}\big(\{|R_p| \geqslant x\}\big).$$

Dans de nombreuses copies, cette inégalité est considérée comme une évidence. Le jury attendait une argumentation. Par exemple, on peut exploiter les inclusions

$$\{|R_n - R_p| > 2x\} \subset \{|R_n| > x\} \cup \{|R_p| > x\} \subset \{|R_n| \geqslant x\} \cup \{|R_p| \geqslant x\}.$$

Q18. Question très facile et réussie dans la quasi-totalité des copies. On trouve $-x_{n,k} = x_{n,n-k}$. De rares copies ont mal interprété la question et ont tenté d'obtenir des inégalités reliant $-x_{n,k}$ et $x_{n,n-k}$ sans doute à cause du verbe *comparer* présent dans l'énoncé.

Q19. Les réponses ont été systématiquement justes une fois comprise la définition de B_n (qui est à valeurs dans un ensemble fini donc est bornée). D'excellentes copies ont proposé la réponse expéditive :

$$\Delta_n \leqslant \max_{0\leqslant k\leqslant n} \frac{\sqrt{n}}{2^{n+1}}\binom{n}{k} + \sup_{x\in\mathbb{R}}|\varphi(x)| \leqslant \max_{0\leqslant k\leqslant n} \frac{\sqrt{n}}{2^{n+1}}\binom{n}{k} + \frac{1}{\sqrt{2\pi}}.$$

De rares copies ont mal interprété la question : au lieu de montrer la finitude du nombre Δ_n, ces copies ont seulement justifié qu'une borne supérieure existait (éventuellement valant $+\infty$). Cette approche n'apporte malheureusement aucune information pertinente et n'a pas été valorisée.

Q20. La question semble à priori très facile mais est sans doute l'une des plus difficiles du sujet non par sa difficulté intrinsèque mais plutôt parce qu'elle nécessite une sérieuse attention. En effet, vu l'énoncé, l'angle naturel est évidemment de justifier que B_n est une fonction paire pour en déduire immédiatement

$$\Delta_n = \sup_{x\in\mathbb{R}}|B_n(x) - \varphi(x)| = \sup_{x\geqslant 0}|B_n(x) - \varphi(x)|.$$

Or la fonction B_n n'est pas paire ! Par exemple, on a

$$B_n(-\sqrt{n} - \frac{1}{\sqrt{n}}) = \frac{\sqrt{n}}{2}\frac{1}{2^n} \qquad \text{et} \qquad B_n(n + \frac{1}{\sqrt{n}}) = 0.$$

Le jury a tout de même valorisé les copies ayant tenté cette approche (qui aurait sans doute pu fonctionner si B_n avait été définie de façon légèrement différente). Le jury est néanmoins satisfait de constater que de nombreuses copies ont décelé cette absence de parité (ce qui a été naturellement valorisé) et ont essayé de contourner cet obstacle.

Q21. Certaines excellentes copies ont traité directement les cas n pair et impair : il s'agit de démontrer la décroissance de la suite $k \mapsto \frac{\sqrt{n}}{2}\binom{n}{k}\frac{1}{2^n}$ pour les indices k tels que $0 \leqslant x_{n,k} + \frac{1}{\sqrt{n}}$, c'est-à-dire $n \leqslant 2k+1$. Cette décroissance s'obtient directement en examinant le quotient :

$$\frac{\binom{n}{k}}{\binom{n}{k+1}} = \frac{(k+1)!(n-k-1)!}{k!(n-k)!} = \frac{k+1}{n-k} \geqslant 1.$$

Le jury attendait une rédaction soignée pour un seul des deux cas (n pair ou n impair) et une mention que le second cas est similaire.

Q22. Cette question nécessite deux types d'arguments :

- une argumentation formelle (grâce à la formule de Stirling) pour retrouver le terme principal ;
- une argumentation asymptotique qui nécessite de gérer des termes du type $O(\frac{1}{k})$ et $O(\frac{1}{n-k})$. Il faut alors comprendre que l'hypothèse $x_{n,k} \in [0, \ell+1]$ implique que les deux précédents termes sont en fait de la forme $O(\frac{1}{n})$.

L'argumentation formelle est obtenue dans la quasi-totalité des copies tandis que l'argumentation asymptotique (plus difficile) est rarement écrite de façon satisfaisante.

Q23. Des calculs algébriques faciles amènent à des facteurs de la forme $\frac{1}{1+O(1/n)}$. La justification suivante a rarement été mentionnée :

$$\frac{1}{1+O\left(\frac{1}{n}\right)} = 1 + O\left(\frac{1}{n}\right).$$

De façon précise, vu l'énoncé de la question, la plupart des copies font apparaitre directement $1+O(1/n)$ au numérateur.

Conclusion

Le jury se permet de réitérer un passage du rapport de la session 2022 : il est attendu qu'une copie normale soit lisible, claire et propre. Beaucoup de copies corrigées n'ont pas respecté ces critères et ont été pénalisées par l'application d'un malus. Le jury conseille notamment de mettre en avant les hypothèses d'un résultat du cours nécessaire pour répondre à une question. Il arrive en effet fréquemment que des réponses soient partiellement valorisées sur la simple connaissance d'un théorème même si ce dernier est *in fine* mal appliqué.

Concernant le sujet, les candidats ont assez bien assimilé les questions d'analyse (calcul algébrique et inégalités simples) mais les questions probabilistes (peut-être moins faciles que celles des sujets des années précédentes) ont causé beaucoup de difficultés de même que les questions d'analyse ayant trait à un comportement asymptotique.