Mathématiques 1

Présentation du sujet

Le problème propose la démonstration du théorème de Johnson et Lindenstrauss, qui est un énoncé de nature géométrique, par une méthode probabiliste. Plus précisément, si $\varepsilon > 0$, on cherche une ε-isométrie entre espaces de dimensions différentes ; pour cela, on démontre que, si la dimension d'arrivée n'est pas trop petite, la probabilité qu'une application linéaire *au hasard* convienne est strictement positive.

Notons que la minoration de la dimension d'arrivée proposée par le sujet est optimale ; mais la preuve de ce fait aurait trop rallongé l'énoncé.

La première partie, plutôt longue, compte pour 29 % des points de barème. Elle est constituée de blocs de questions indépendants les uns des autres et vise à fournir des outils techniques pour aborder cette démonstration. Les résultats sont tous donnés, pour ne pas bloquer les candidats.

Dans la deuxième partie, comptant pour 42 % du total, on entre dans le vif du sujet : il s'agit d'établir l'inégalité de concentration de Talagrand, étape majeure. On commence par aborder des cas particuliers simples, avant de s'attaquer au cas général, par récurrence sur la dimension de l'espace, en utilisant les outils de la première partie.

Dans la dernière partie, évaluée sur 29 % du total, on étudie un ensemble fini de matrices aléatoires et on trouve à l'aide de l'inégalité de Talagrand une borne inférieure pour la dimension d'arrivée, garantissant qu'un tel ensemble contient une ε-isométrie.

Analyse globale des résultats

Une remarque positive pour commencer : une proportion croissante de candidats est capable d'aborder avec succès des questions abstraites de probabilités. Certains manquent parfois un peu d'efficacité, d'aisance, ou de rigueur, à cause d'un manque de pratique, mais leurs raisonnements sont pour l'essentiel corrects et intelligents.

Ce satisfecit n'empêche pas de remarquer qu'une majorité de candidats continue à avoir beaucoup de difficultés avec des notions probabilistes pourtant simples, parfois introduites dès le lycée : la loi binomiale, le théorème de transfert et, surtout, la formule des probabilités totales. Les propriétés utiles de l'espérance, surtout la croissance ou la positivité, ne sont pas toujours identifiées ni citées. L'indépendance des variables ou des événements, propriété opératoire essentielle, n'est pas souvent signalée dans la suite des calculs effectués.

Le traitement des inégalités, et plus généralement les manipulations algébriques élémentaires, sont assez souvent source d'erreurs facilement évitables.

Une partie des candidats estime inutile de justifier une suite de calculs, voire de les présenter et d'encadrer les résultats. Quand le résultat est fourni par l'énoncé, il est quand même évident que les correcteurs attendent une justification précise, explicite et lisible de la démarche qui permet de l'obtenir. Et pourtant, les opérations usuelles (multiplication, division, logarithme, dérivation) sont effectuées sans aucune rigueur (signe ? non nullité ? stricte positivité ? dérivabilité ?) et conduisent à des formules fausses ou mal démontrées.

Une autre partie des candidats se contente de quelques mots-clés comme justification. Si cela permet de comprendre, voire de valider la démarche, c'est encore insuffisant ; par exemple, citer le nom d'un mathématicien célèbre (Pythagore, Fubini, Markov) ne dispense pas de vérifier les hypothèses du théorème qui lui est couramment associé.

Commentaires sur les réponses apportées et conseils aux futurs candidats

I Préliminaires

La première question a été plutôt réussie, sauf par certains candidats qui ne savent pas développer le carré d'une norme. L'interprétation géométrique est souvent erronée ou absente ; écrire le nom de la formule prouvée n'est pas suffisant. La question suivante a également été réussie, par les candidats qui utilisaient **Q1.**.

La question **Q3.** n'a pas toujours été traitée correctement ; il y a quelques confusions sur les notions de compacité ou de borne inférieure. Quelques candidats ont cru bon de citer le cours, alors que ce résultat ne figure plus au programme ; d'autres ont cru que F devait être, de par son nom, un sous-espace vectoriel de E, et ont utilisé le projecteur orthogonal sur F.

La question **Q4.** contenait une coquille, souvent signalée par les candidats, et qui ne semble pas les avoir perturbés.

La rédaction de la question **Q5.** a été souvent bâclée : peu de candidats prennent le soin d'écarter rapidement les cas triviaux $a = 0$ ou $b = 0$, ils utilisent à la place le logarithme sans se soucier de la nullité ou non des variables. La question **Q6.** a rarement été complètement traitée : si le cas $\mathbb{E}(|X|^p) = \mathbb{E}(|Y|^q) = 1$ est souvent abordé, par transfert ou croissance de l'espérance, le passage au cas général est souvent absent, et peu de candidats mentionnent que $\mathbb{E}(|X|^p) = 0$ équivaut à $X = 0$ presque surement, sans parler de le démontrer.

La question **Q7.** est facile, et pourtant elle a été très discriminante, à cause de la pauvreté des justifications fournies : l'interversion des sommes n'a pas toujours été expliquée, ou alors avec des arguments excessifs (sommabilité !), et surtout, la formule des probabilités totales, dans sa version avec probabilités conditionnelles, a été utilisée sans en vérifier les hypothèses (système complet d'événements non négligeables).

La question **Q8.** était difficile, et certains candidats, ayant une intuition correcte de la démarche à adopter, ont choisi malheureusement d'en négliger grossièrement les détails techniques (gestion des indices, justification de l'intégrabilité, télescopie bien mise en évidence) ; ils ont été sanctionnés.

Un argument d'intégrabilité ou de positivité était attendu pour la question **Q9.**.

L'étude du signe d'un trinôme dans **Q11.** pose des problèmes à beaucoup de candidats, qui font à cette occasion preuve d'inventivité pour trouver des arguments faux, mais rapides. La question **Q13.** aurait pu servir d'alerte à ceux qui n'avaient pas utilisé l'hypothèse de **Q12.** pour mener le calcul, mais il n'en a rien été.

II L'inégalité de concentration de Talagrand

La question **Q15.** a révélé le manque d'aisance des candidats avec la loi binomiale. Après avoir montré plus ou moins correctement qu'il s'agissait de calculer la loi d'une somme de variables de Bernoulli, beaucoup de candidats se lancent dans un dénombrement verbeux qui est forcément faux dès qu'on ne mentionne pas l'indépendance ; et ils oublient de simplement citer leur cours (de Terminale), qui permet de conclure en une ligne.

Le théorème de transfert, nécessaire pour résoudre **Q16.**, est parfois mal utilisé, et l'expression obtenue n'est pas reconnue et simplifiée. La question **Q17.** demandait de donner les arguments dans le bon ordre (la croissance de l'espérance avant sa majoration, par exemple) ; et cela n'a pas toujours été le cas dans les copies.

L'équivalence de la question **Q19.** n'a pas toujours été décomposée en deux implications, ce qui rendait souvent le raisonnement faux.

Les propriétés topologiques nécessaires pour traiter **Q20.** sont très mal maitrisées. On a l'impression que les candidats croient que les douze énoncés : « l'image directe/réciproque d'une partie fermée/convexe/non vide par une application continue/linéaire l'est encore » sont tous vrais (et dans le cours).

Dans la question **Q21.**, il s'agissait d'appliquer la formule des probabilités totales, en explicitant le système complet d'événements correct. L'indépendance des ε_i était essentielle pour obtenir la forme demandée par l'énoncé, elle n'a pas souvent été citée.

Dans la question **Q23.**, il fallait mentionner l'orthogonalité pour appliquer le théorème de Pythagore (et non pas l'inégalité triangulaire...) ; et il fallait utiliser la convexité de la fonction carré pour conclure. Signalons à ce sujet que $x \mapsto \|x\|^2$ est certes une application convexe, mais ceci est un résultat hors programme qu'il convient de redémontrer.

Le calcul de la question **Q25.** était particulièrement délicat ; la difficulté était bien de faire disparaitre le conditionnement par l'événement $[\varepsilon_n = -1]$. Il fallait déjà identifier cette difficulté, procéder par étapes et surtout ne pas tenter d'abuser le correcteur, qui repère très facilement les tentatives de passage en force.

De nombreux candidats ont cru résoudre la question **Q26.** en invoquant l'inégalité de Cauchy-Schwarz ou l'indépendance des variables considérées. Il fallait plutôt faire explicitement référence à **Q6.**, en précisant le lien entre les notations des deux questions (λ, p, q, X, Y).

Il est regrettable que les candidats fassent autant d'erreurs de dérivation dans la question **Q30.**, et que les inégalités utilisées soient fausses ou non démontrées.

Les rares candidats qui ont abordé la question **Q33.** de façon pertinente n'ont pas justifié l'utilisation de l'inégalité de Markov.

III Démonstration du théorème de Johnson-Lindenstrauss

La question **Q34.** a été très souvent abordée, mais la justification de la continuité de g a souvent été incomplète, voire fausse : g n'est ni linéaire, ni polynomiale. Peu de candidats ont donné l'expression correcte de $\|M \cdot u\|^2$ dans la question **Q35.**, qui permettait l'utilisation de l'inégalité de Cauchy-Schwarz.

Dans la question **Q36.**, très peu de candidats ont pensé à dire que C n'était pas vide avant d'utiliser la notion de projeté.

Dans la question **Q39.**, l'utilisation de la question **Q37.** nécessitait d'être précis sur les valeurs de r utilisées : les candidats doivent être rigoureux dans leur raisonnement même en fin de composition.

Conclusion

La présentation d'une copie de mathématiques doit permettre de comprendre la démarche effectuée, et d'en vérifier la solidité. À travers un problème difficile, mais guidé par des questions intermédiaires, on cherche à évaluer la capacité des candidats à s'approprier une problématique complexe et produire des raisonnements élaborés à partir des théorèmes du programme.

Il convient donc de bien mettre en évidence sa démarche, et de justifier chaque étape de raisonnement ou de calcul par des arguments tirés du cours, ou des questions précédentes. S'intéresser aux cas particuliers dans l'obtention d'un résultat général est souvent une preuve de recul qui est valorisée par les correcteurs.

Rappelons pour finir quelques évidences. Ratures, calligraphie minuscule ou difficilement déchiffrable, aucune mise en valeur des résultats intermédiaires ou finaux, utilisation abusive d'abréviations : ces défauts se retrouvent, à des degrés variés, dans beaucoup de copies. Si on peut comprendre la volonté d'aller le plus loin possible dans le temps imparti, il ne faudrait pas oublier qu'une copie a vocation à être lue en l'état par le correcteur, et qu'il ne s'agit pas d'un brouillon. Ces défauts donnent ainsi lieu à une minoration de la note, ce qui est contraire au but recherché par le candidat.