MATHÉMATIQUES

ÉPREUVE COMMUNE : ÉCRIT

Sylvain Arlot, Aurélien Garivier

Coefficient : 3

Durée : 4 heures

Calculatrice interdite

Commentaires généraux

Le sujet était composé de deux exercices indépendants et d'un problème, portant sur des aspects totalement différents du programme. Comme à l'habitude, le sujet était long, couvrait une très large partie du programme, et permettait à chaque candidat de valoriser ses connaissances, ses aptitudes et la qualité de ses raisonnements mathématiques sur des questions de difficultés variées. En particulier, les premières questions de chaque exercice ou problème permettaient de tester l'acquisition des connaissances de base en analyse, probabilités et algèbre.

Tous les candidats suffisamment préparés auraient dû facilement venir à bout des questions I.1, II.1 et III.4,8 ; cela assurait déjà une note supérieure à 6.5/20, comme 298 candidats sur 548 présents. Il semblait facile de résoudre en plus les questions I.4 et III.1,5,7, ce qui permettait d'atteindre au moins 11/20, comme 141 copies. En résolvant correctement les questions I.1–4, II.1–2,5 et III.1,4–5,8, on obtenait au moins 14/20, niveau que seuls 62 candidats ont atteint. Enfin, en résolvant en plus les questions I.5–7, II.3–4 et III.2–3,7, on obtenait une note supérieure à 17.5/20, comme les 13 meilleures copies cette année. Les candidats préférant prendre le temps de faire aboutir des raisonnements plus fins trouvaient dans les questions I.8, II.8–12, et III.9,11–14 matière à réflexion, et pouvaient obtenir d'excellentes notes sans en traiter un grand nombre.

Insistons sur le fait que la présence de questions relativement difficiles (I.8, II.8–12, et III.9,11–14) ne doit aucunement décourager les candidats qui n'arriveraient pas à en venir à bout : on pouvait obtenir 19.5/20 sans en résoudre aucune, à condition de savoir rédiger correctement les questions plutôt classiques du reste du sujet.

Le jury renouvelle son inquiétude de voir une proportion importante de copies totalement indigentes (192 copies sur 548 ont obtenu une note inférieure ou égale à 4/20). La remontée de la moyenne ne doit pas cacher le fait que manifestement, de trop nombreux candidats ont totalement abandonné les mathématiques avant de se présenter au concours. Signalons donc que parmi les 57 admissibles à l'ENS, seuls quatre ont obtenus une note en mathématiques strictement inférieure à 10/20, et un seul d'entre eux a été admis.

Saluons toutefois la qualité des meilleures copies, qui ont impressionné le jury par la précision de leur rédaction, la qualité des raisonnements présentés et la quantité de questions traitées correctement : on pourrait croire que leurs auteurs suivent un cursus de mathématiques.

De manière générale, il semble qu'une majorité des candidats sont troublés par le fait de devoir mener des raisonnements avec un paramètre libre. L'exercice I (qui se borne l'étude de trois fonctions d'une variable réelle) et la partie (A) du problème (où l'on demandait simplement de diagonaliser une matrice 2×2) nous semblent relever d'un niveau d'abstraction minimal exigible pour l'épreuve de mathématiques de la filière B/L.

Un nombre significatif de copies ne résolvent les questions I.1–5 et III.1–2 que pour des valeurs particulières des paramètres ($p = q = 1/2$ ou $x = 0.5$ dans l'exercice I, $p, q \in \{0, 1\}$ dans le problème), ce qui a fait perdre de nombreux points, même aux candidats que l'on devinait proches de résoudre le cas général.

Avant de rentrer dans le détail des exercices, voici quelques conseils aux candidats pour la rédaction de leur copie :

– Il est souhaitable de présenter sa copie le plus clairement possible. En particulier, le jury apprécie que les réponses à un même exercice soient présentées dans l'ordre, et qu'en tout cas les éléments de réponse à une même question soient rassemblés en un seul endroit, sauf mention explicite du contraire !
– Plusieurs candidats prennent dans leur copie des libertés avec les notations introduites dans le sujet, par exemple, appeler f la fonction nommée g à la question I.5. Cela peut être très troublant pour le correcteur, et peut conduire à pénaliser gravement le candidat.
– Pour un bon nombre de questions d'analyse, tracer un tableau de variation peut s'avérer la plus claire des justifications (ici, pour les questions I.2,4–5), à condition bien sûr de ne pas oublier de répondre à la question posée. Au sujet des questions I.4–5, il est d'ailleurs assez étrange d'avoir vu d'assez nombreuses copies mener une étude complète des fonctions proposées, mais ne pas conclure leur réponse par un graphe qui était pourtant explicitement demandé, et nécessaire à l'obtention de la majorité des points de la question.
– Recopier une question de l'énoncé ne peut jamais rapporter de point et demeure parfaitement inutile.
– Énoncer une affirmation manifestement fausse ne peut pas servir le candidat, mais seulement jeter la suspicion sur tout ce qu'il écrit.

En vue de préciser notre analyse des principales faiblesses observées dans les copies, nous indiquons cette année pour chaque question le nombre de copies ayant obtenu au moins 75% des points (sur un total de 548 candidats présents).

Commentaires détaillés sur chaque exercice

Exercice I. Le premier exercice portait sur l'étude de propriétés élémentaires de la divergence de Kullback-Leibler pour entre des variables de Bernoulli (ce qu'il était, bien évidemment, totalement inutile de réaliser !). Cet exercice d'analyse très classique ne comportait guère de difficulté insurmontable par un candidat moyen, la question I.8 seule demandant de prendre une petite initiative. Il a toutefois été diversement réussi : sur les questions I.1–7, la note médiane est de 21%, seules 39 copies obtenant plus de 75% des points.

1. [281 copies $\geqslant$ 75%] Question élémentaire, même si malheureusement la majorité des candidats ne prennent pas le temps de justifier que la quantité à l'intérieur du logarithme est bien strictement positive pour tout réel t, perdant au passage une (toute petite) fraction de point. Il était apprécié que le candidat fournisse la réponse sous la forme factorisée la plus simple possible.
2. [233 copies $\geqslant$ 75%] Question assez bien réussie dès lors que h' et h'' sont correctes, si ce n'est qu'une proportion significative de candidats mélange "$h'(x) = 0$", "h admet un maximum local en x" et "h admet un maximum global en x".
3. [197 copies $\geqslant$ 75%] Manipulations élémentaires, bien faites en général lorsque l'expression de $t^\star$ est correcte.
4. [82 copies $\geqslant$ 75%] Il ne s'agit finalement que d'une question très classique (étude de fonction et tracé de l'allure de son graphe), où la seule petite difficulté est la présence d'un paramètre libre. Malgré cela, la question a été étonnamment peu réussie,

comme si les candidats n'osaient par principe pas se lancer dans le dessin d'un graphe. Quelques erreurs graves, et malheureusement fréquentes :

– des tableaux de variations et graphes tracés hors de l'intervalle de définition $]0,1[$, alors que celui-ci était donné dans l'énoncé,
– des graphiques incohérents avec l'énoncé de la question (6) ; les candidats obtenant des valeurs strictement négatives pour $d(p,q)$ auraient pu repérer leur erreur.

Enfin, il est inutile d'écrire des dérivées partielles, f et g sont des fonctions d'une seule variable.

5. [12 copies $\geqslant 75\%$] Question encore moins bien réussie que I.4, alors qu'elle n'est pas tellement plus difficile (hormis les demi-tangentes aux bords, que quasiment personne n'a trouvées). Plusieurs copies indiquent que g ayant une limite en 0, elle y admet une asymptote horizontale.
6. [65 copies $\geqslant 75\%$] Question facile quand on avait fait I.4 ou I.5 (encore fallait-il s'en rendre compte). Bon nombre de candidats, visiblement peu sûrs d'eux, croient nécessaire d'invoquer *à la fois* I.4 *et* I.5. Quelques candidats ont aussi vu qu'on pouvait déduire le résultat de I.3, car $d(p,q) = h(t^\star) \geqslant h(0) = 0$.
7. [36 copies $\geqslant 75\%$] C'est également une conséquence directe de I.4 ou I.5, ce que tous les candidats ayant trouvé I.6 n'ont pas vu. Il est pourtant classique d'obtenir le cas d'égalité à l'aide d'une analyse attentive de la preuve de l'inégalité. Deux bons candidats ont remarqué que c'était aussi une conséquence de I.3, en utilisant l'unicité du maximum $t^\star$.
8. [aucune copie $\geqslant 75\%$] Les plus courageux n'ont pu faire mieux que commencer quelques calculs. La solution la plus simple était de fixer x et d'étudier la fonction correspondante de y, comme en I.4.

Exercice II. Dans le deuxième exercice, qui portait sur un des thèmes les plus classiques de la théorie des probabilités, on étudiait la loi du premier retour à l'équilibre d'une marche aléatoire symétrique sur $\mathbb{Z}$ via le lemme du scrutin, par une approche combinatoire. Il n'a guère été apprécié des candidats ; même les questions 1 à 7, qui n'ont pourtant rien de très difficile, ont visiblement posé problème : seuls 61 candidats y ont obtenu au moins la moitié des points ! Il semble que les candidats ne soient, pour beaucoup, pas très à l'aise avec les probabilités discrètes sous l'angle de la combinatoire. Nous insistons sur le fait que les questions combinatoires élémentaires ont toute leur place à l'épreuve de mathématiques, aussi bien à l'écrit qu'à l'oral.

Tout au long de l'exercice, beaucoup confondent n et N dans leur copie. Il faut être plus attentif à la lecture de l'énoncé, qui était pourtant aussi clair que possible, avec même une illustration graphique pour aider les candidats à s'approprier les objets étudiés.

1. [104 copies $\geqslant 75\%$] Deux raisonnements possibles : par dénombrement, ou en remarquant (avec une brève justification) que T_i est une somme de variables de Bernoulli $(1/2)$ indépendantes. Beaucoup de candidats ont tenté cette question, mais en proposant des réponses plus ou moins fantaisistes, telles que $\mathbb{P}(\omega_i = 1) = 2^{-n}$ ou 2^{-N}, ou encore que $T_n \sim \mathcal{B}(n, 2^{-N})$ ou $\mathcal{B}(n,p)$ (sans préciser la valeur de p).
2. [35 copies $\geqslant 75\%$] Cette question, qui visait essentiellement à aider les candidats pour la question 3, a été particulièrement mal traitée. En particulier, pour l'ensemble des valeurs prises par S_n, la grande majorité des candidats n'a pas vu que la parité de S_n est fixée par celle de n, même parmi ceux qui écrivent bien à la question suivante que la probabilité d'avoir $S_n = x$ est nulle lorsque $(x+n)/2$ n'est pas un entier.
On a aussi beaucoup vu la notation $\overline{T_n}$, à laquelle les correcteurs étaient censés donner un sens inspiré de celui qu'a $\overline{A}$ pour un événement A.
3. [75 copies $\geqslant 75\%$] C'est une conséquence directe des deux premières questions, bien traitée par les candidats qui avaient réussi les précédentes.

4. [84 copies $\geqslant$ 75%] Conséquence facile de la question 3. Un certain nombre de candidats ont pu se raccrocher ici en réussissant un raisonnement combinatoire direct pour calculer $p_{2k,0}$. C'est dommage qu'ils n'en aient pas profité pour réaliser comment répondre à la question 3 avec un raisonnement similaire.
5. [58 copies $\geqslant$ 75%] Ce calcul d'équivalent a été en général bien réussi par les candidats qui l'ont tenté. Les principales erreurs observée ici concernent la formule du coefficient binomial $\binom{2k}{k}$, en particulier des simplifications, telles $((2k)!)/(k!)^2 = 2/k!$, pour le moins audacieuses. Signalons également le manque d'esprit critique des candidats qui obtiennent que $q_{2k} \to +\infty$ quand $k \to +\infty$, sans s'inquiéter de voir une probabilité dépasser un. Moins grave, mais un peu dommage : de nombreux candidats laissent la réponse sous la forme $\sqrt{k\pi}/(k\pi)$, qui mérite quand même une simplification.
6. [27 copies $\geqslant$ 75%] Beaucoup ont tenté de répondre à cette question, mais rarement avec des arguments convaincants. Le plus souvent, on voit une suite de calculs formels sans aucune justification, ce qui n'était pas valorisé. Les candidats confondent souvent "incompatibles" et "complémentaires", ou encore union et intersection. Ici, il faut expliquer (avec un raisonnement, donc des mots) que (i) $S_1 > 0$ implique $S_1 = 1$ (conséquence de la question 2, par exemple), et (ii) si la séquence de terme S_k n'est pas tout le temps strictement positive pour $k = 1 \ldots n$, alors elle s'annule nécessairement (car elle varie de 1 en 1), puisqu'elle termine en $x > 0$.
7. [25 copies $\geqslant$ 75%] Première question un peu subtile de l'exercice, réussie par un nombre appréciable de candidats.
8. [1 copie $\geqslant$ 75%] Première question véritablement difficile : rares sont les candidats ayant tenté une question dans l'exercice à partir de là. Le dessin était là pour suggérer comment construire φ. Trois candidats semblent avoir compris l'idée, un seul a su la justifier correctement.

9–12. Aucun candidat n'a abordé sérieusement ces questions, mis à part trois candidats qui ont réussi à combiner les questions 6–8 pour obtenir la première partie de la question 9. Ces questions pourront servir aux futurs candidats à compléter leur préparation pour ce qui est de la combinatoire.

Problème. Ce problème d'algèbre linéaire assez classique visait à décrire l'ensemble des matrices stochastiques, et plus généralement, l'ensemble des matrices admettant un vecteur propre donné. Malgré plusieurs questions faciles (1, 5 et 7) à très faciles (4 et 8), les candidats n'ont majoritairement pas brillé : sur les questions 1, 4–5 et 7–8, la note médiane est de 25%, et 67 copies seulement ont obtenu 75% des points ou plus.

Une remarque générale concerne les manipulations de matrices : il est beaucoup plus simple de noter $M_{i,j}$ le coefficient d'indice (i, j) de la matrice M, plutôt que d'introduire des notations telles que $a_{i,j}$ ou $x_{i,j}$, qui ne font qu'ajouter un peu de confusion.

1. [82 copies $\geqslant$ 75%] Tous les candidats devraient être capables de trouver les valeurs propres d'une matrice 2×2, même avec des paramètres libres. Ce n'était clairement pas le cas cette année. Et même parmi les candidats sachant comment faire, une majorité obtiennent un résultat correct mais inutilement compliqué, par exemple
$$\frac{1 - q + p + \sqrt{(1 - q + p)^2 - 4(p - q)}}{2}$$
au lieu de 1, ce qui ne peut rapporter tous les points. Ici, il fallait reconnaître que le discriminant du trinôme obtenu est un carré, et c'était indispensable pour pouvoir discuter en fonction de p et q si l'on avait ou pas deux valeurs propres distinctes.
2. [50 copies $\geqslant$ 75%] En principe, c'est une question assez facile lorsque l'on a bien fait la précédente. Il ne fallait pas oublier de considérer le cas $p - q = 1$. Beaucoup (parmi ceux qui n'ont pas simplifié leur résultat à la question 1) semblent croire que deux

nombres s'écrivant de deux manières différentes sont nécessairement différents. Beaucoup trop de candidats écrivent des raisonnement montrant de très lourdes lacunes, comme "la matrice P n'est pas diagonalisable car elle possède 3 valeurs propres".

3. [16 copies $\geqslant 75\%$] Mis à part les meilleurs, les candidats semblent ne pas comprendre ce type de question. Précisons que "$\lambda = p - q$" (sans plus de précision sur le statut de p et q) n'est pas une réponse admissible.
4. [232 copies $\geqslant 75\%$] Question très facile, pour laquelle il suffit de savoir calculer le produit MJ. Elle a permis de discriminer entre les copies faibles et très faibles. Notons qu'on lit souvent $(MJ)_{i,j}$, les candidats ne réalisant pas que MJ est un vecteur, et plus grave $MJ = J \implies M = 1$, ou $M = J/J = 1$.
5. [148 copies $\geqslant 75\%$] Question facile si l'on pense à utiliser la question précédente. Une bonne moitié des candidats repart de zéro, avec plus ou moins de succès selon les cas. Les plus graves erreurs sont d'écrire $(MJ)^n = J^n = J$ ou $(M^n)_{i,j} = (M_{i,j})^n$. Certains candidats affirment en guise d'argument que le produit de deux matrices stochastiques est stochastique : dans le meilleur des cas, ils n'avaient pas vu que l'objet de la question était précisément de le montrer dans un cas particulier.
6. [21 copies $\geqslant 75\%$] Il faut ici faire un raisonnement plus subtil, bien que relativement classique, dont seuls les bons candidats en sont venu à bout. Deux approches au moins sont possibles : (i) prendre X_j tel que $|X_j|$ est maximale (et pas X_j, attention), ou bien (ii) voir que M^n admet λ^n comme valeur propre, tout en ayant des coefficients entre 0 et 1, ce qui implique que λ^n est bornée lorsque $n \to +\infty$. Mentionnons une grave erreur lue à de trop nombreuses reprises : "$|MX| \leqslant |\lambda||X|$ et l'on peut simplifier par $|X|$ car $X \neq 0$"...
7. [140 copies $\geqslant 75\%$] Question facile dès qu'on a compris que les éléments de E_X sont des matrices, et qu'admettre X comme vecteur propre ne signifie pas que la valeur propre associée est la même pour tous les éléments de E_X. Quelques candidats semblent confondre E_X avec la notation (qui vient peut-être de leur cours) E_λ relative à un sous-espace propre, tandis que d'autres écrivent sans gêne aucune $MX = XM$, ou bien fréquemment : $MX = 0$ et $X \neq 0 \implies M = 0$.
8. [296 copies $\geqslant 75\%$] À cet endroit se trouvait une des questions les plus faciles de l'énoncé : cela a été en tout cas la plus réussie. Quelques candidats ont réussi à donner une réponse fausse, par exemple en faisant commuter A et X, mais ils sont heureusement assez rares : une écrasante majorité des 225 candidats ayant eu zéro à cette question ne l'ont tout simplement pas tentée.
9. [5 copies $\geqslant 75\%$] Les raisonnement plus fins commencent avec cette question. Beaucoup de candidats essaient d'appliquer dès maintenant le théorème du rang (alors que c'est à la question 11 qu'il faudra le faire). Beaucoup ne voient même pas que l'espace de départ et l'espace d'arrivée n'ont pas la même dimension, quand ils n'écrivent pas que $\dim(\mathcal{M}_k(R)) = k$. Et de nombreux candidats croient prouver ici que $\ker(\varphi_X) = \{0\}$ (avec l'argument "$AX = 0$ implique $A = 0$ car X est non-nul"), ce qui rend les questions suivantes très faciles, mais aussi les réponses toutes fausses. On peut tout de même saluer les cinq candidats ayant la bonne réponse correctement justifiée.
10. [13 copies $\geqslant 75\%$] Il n'y avait pas de difficulté insurmontable ici, mais la réponse est assez longue à rédiger, et nécessite une bonne maîtrise des raisonnements logiques. Rares sont ceux qui ont bien prouvé les deux inclusions et que la somme est directe. La plupart ont prouvé que la somme est directe et que $\ker(\varphi_X) + \mathrm{vect}(I_k) \subset E_X$ en croyant avoir prouve l'égalité. On voit souvent une tentative d'utiliser un argument de dimension pour conclure, mais ce n'était pas la bonne approche ici (ce que l'on pouvait deviner au vu de la question suivante, portant sur le calcul de $\dim(E_X)$).

Les candidats procédant par analyse-synthèse obtiennent le plus souvent des arguments très confus. C'est une bonne manière de faire au brouillon, mais il est important de ne retranscrire sur la copie que la synthèse bien rédigée.
Une erreur grave (souvent lue ici) est de croire que $\mathrm{vect}(I_k)$ serait l'ensemble des matrices diagonales non-nulles.

11. [19 copies $\geqslant 75\%$] C'est donc ici qu'il fallait appliquer le théorème du rang, en faisant attention au fait que ϕ_X part d'un espace de dimension k^2, et non k. Une quarantaine de candidats semblent avoir compris l'argument, et l'ont rédigé avec plus ou moins de succès. Beaucoup d'erreurs graves dans les autres copies ayant tenté la question, la plus courante étant $\dim(\mathrm{vect}(I_k)) = k$.
12. [1 copie $\geqslant 75\%$] Question très rarement tentée, plutôt bien résolue dans une poignée de copies.
13. [7 copies $\geqslant 75\%$] Peu de candidats ont tenté cette question, alors qu'elle n'est pas si difficile à résoudre en partant de zéro. Sinon, c'est une conséquence assez directe de la question 12 (le noyau de F_J étant alors de dimension 1).
14. [aucune copie $\geqslant 75\%$] Cette question, quasiment jamais abordée, faisait le lien entre les parties (B) et (C) du problème. Un candidat obtient une réponse acceptable pour une base de E_J.