2.3 Mathématiques 2 - filière MP

2.3.1 Généralités et présentation du sujet

Ce sujet traite de marche aléatoire dans $\mathbb{Z}^d$, c'est-à-dire d'une suite de variables aléatoires $(S_n)_n$ où $S_0 = 0_d$ et pour tout entier $n \geq 1$, $S_n = \sum_{k=1}^{n} X_k$ avec $(X_k)_{k\in\mathbb{N}^*}$ une suite de variables aléatoires mutuellement indépendantes et de même loi à valeurs dans $\mathbb{Z}^d$. Au temps k, on fait un pas dans la direction X_k. Ainsi, plusieurs questions se posent naturellement : Revient-on en 0 ? Au bout de combien de temps en moyenne, revient-on en 0 pour la première fois ? Après n pas, combien avons-nous visité en moyenne de positions différentes ? Ce sujet propose donc d'étudier ces questions. Pour cela, outre les probabilités, les séries entières sont utilisées.

2.3.2 Commentaires généraux

Le jury tient à souligner l'impolitesse dont ont fait preuve un certain nombre de candidats qui ont rendu une copie illisible, ou trop raturée, ou aux questions traitées dans le désordre, ou à l'encre trop pâle, ou encore comportant trop d'abréviations. Le correcteur attend de la part du candidat l'écriture en toutes lettres des notions utilisées, étant toléré le fait d'expliciter une abréviation pour l'employer par la suite dans le devoir.
Un certain nombre de questions donnaient la formule à démontrer, ce qui peut guider les candidats dans leur réflexion. Mais, il faut alors avoir en tête que les points ne seront pas accordés pour avoir trouvé la formule mais pour la rigueur de la démarche. Malheureusement, certains candidats ne jouent pas la « carte de l'honnêteté » et partant dans une mauvaise direction, ils truquent au fur et à mesure leurs calculs pour aboutir au résultat demandé. De plus, les candidats affirment, sans preuve, un résultat alors que les difficultés (et donc les points) se trouvent justement dans la preuve de ce résultat, parfois même ce résultat est faux. Rappelons que « *des affirmations extraordinaires nécessitent des preuves extraordinaires* » (Carl Sagan).
Le jury déplore le manque de rigueur avec lequel sont traitées l'analyse et les probabilités. Beaucoup d'erreurs ont été commises dans les questions de convergence de séries. La notion d'indépendance a donné lieu à bien des errements. Outre sa confusion encore fréquente avec celle d'incompatibilité, elle est souvent vue comme un mot clé à insérer à intervalles réguliers dans les raisonnements. Or des variables aléatoires ont souvent été affirmées indépendantes alors qu'en réalité elles ne le sont pas, simplement parce que cela permettait aux calculs d'aboutir. Une bonne compréhension de cette notion est indispensable pour maîtriser le fil des raisonnements de chaque question de probabilités. Le jury conseille donc aux élèves d'être attentifs aux exigences de rigueur de leurs professeurs.

2.3.3 Analyse détaillée des questions

- La **première question** n'a pas eu un assez grand succès, il suffisait pourtant de développer l'expression proposée avec le binôme de Newton (sans se tromper dans les indices), de savoir effectuer un produit de deux polynômes puis d'identifier les coefficients devant X^n. Des candidats, ont substitué 1 à l'indéterminée ce qui ne pouvait pas aboutir. Cela n'a pas empêché certains de pouvoir en déduire la formule demandée au prix d'escroqueries successives.
- À la **question 2**, la part des élèves ne connaissant pas la formule de Stirling est de 28%. Or cette formule figure au programme.
- La **question 3** n'a pas eu le succès espéré. Une méthode consistait à établir des encadrements en procédant par étapes :
  — Donner un encadrement propre de $f(t)$ pour $t \in [k, k+1]$ avec $k \in \mathbb{N} \setminus \{0, 1\}$, par décroissance de $f : x \mapsto x^{-\alpha}$ sur $\mathbb{R}_+^*$ (et non sur $\mathbb{R}$ ou $\mathbb{R}^*$, ce que le jury a pu lire).
  — Par croissance de l'intégrale, obtenir un encadrement de $f(k)$.
  — Sommer ces inégalités pour $k \in [\![2 ; n]\!]$, pour obtenir un encadrement de l'intégrale avec des sommes partielles (ou l'inverse).

— Calculer l'intégrale et montrer que la majoration et la minoration de la somme partielle (ou du reste) sont bien équivalents, la position de α par rapport à 1 jouait, bien entendu, un rôle.

Certains écrivent directement l'encadrement final, comme s'il allait de soi. Malheureusement, ces encadrements étaient parfois faux dus aux problèmes d'indices et de bords : $\int_0^n f(t)\,\mathrm{d}t$ n'a du sens que si on l'a vérifié, car f n'est pas définie en 0. Parfois, la minoration proposée était plus grande que la majoration. Les deux membres étaient affirmés être équivalents sans justification (alors que cela dépend de α). Il serait bien de préciser où vivent les variables utilisées par les candidats : cette inégalité annoncée est-elle vraie pour $k \in \mathbb{N}$, $k \geq 1$ ou $k \geq 2$? Ce n'est pas au jury de le deviner.

Citons d'autres méthodes utilisées qui ont été acceptées si elles étaient traitées avec rigueur :

— La sommation des relations de comparaison appliquée à $n^{-\alpha} \sim \int_n^{n+1} t^{-\alpha}\,\mathrm{d}t$.

— La comparaison de la série de Riemann à une série télescopique.

— Une part infime des candidats a utilisé le théorème de comparaison série-intégrale qui fournit la convergence de la série de terme général $f(n) - \int_n^{n+1} f(t)\,\mathrm{d}t$ pour obtenir rapidement le résultat de la première partie de la question.

En revanche, certains ont fait appel à un théorème plus précis que le théorème de comparaison série-intégrale, mais **hors programme**, fournissant un équivalent de la somme partielle.

- Beaucoup de candidats ont effectué une intégration par parties pour la **question 4**, mais tous n'ont pas justifié sa validité en disant que les fonctions utilisées étaient bien de classe $\mathcal{C}^1$ sur le segment d'intégration. Curieusement, bien rares sont les candidats qui ont remarqué que les deux membres sont nuls pour $x = 2$ et ont la même dérivée, ce qui permettait rapidement de conclure. Pour la relation de négligeabilité, utiliser le théorème d'intégration des relations de comparaison suffisait, encore fallait-il en vérifier toutes les hypothèses. Comme les intégrales de Bertrand ne sont pas au programme, le jury attendait une preuve que $\int_2^{+\infty} \frac{1}{\ln(t)}\,\mathrm{d}t$ diverge ce qui pouvait se faire aisément par comparaison avec une intégrale de Riemann. Seul un candidat sur cinq a montré proprement cette relation de négligeabilité.
- Pour la **question 5**, le développement en série entière de $x \mapsto (1+x)^\alpha$ n'est pas toujours connu. Certains confondent les développements limités et les développements en série entière. D'autres donnent le développement demandé à l'aide d'un coefficient mystérieux : $\binom{\alpha}{k}$. Lorsque $\alpha \notin \mathbb{N}$, cette notation, certes pratique, n'est pas au programme de MP. De plus, la deuxième partie de la question a parfois montré que parmi ceux qui l'utilisaient, tous ne l'avaient pas vraiment comprise. En tout cas, écrire α ! (factorielle α) pour $\alpha \notin \mathbb{N}$ n'a aucun sens, ce que le correcteur sanctionne immanquablement. Pour obtenir le développement de $x \mapsto \frac{1}{\sqrt{1-x}}$, certains commettaient plusieurs erreurs d'indices ou de signe. Au prix de quelques arnaques sur plusieurs lignes, ils ont réussi à aboutir mais ils ont surtout réussi à perdre la confiance du correcteur ...
- La **question 6** a été révélatrice de toutes les incompréhensions sur les séries et les séries entières. Certains ont utilisé le critère de D'Alembert pour obtenir un rayon de convergence. Cette stratégie était vouée à l'échec. En effet, sans aucune expression de $P(S_n = 0)$ ou de $P(R = n)$, comment savoir si ces quantités s'annulent et comment étudier la limite éventuelle du quotient ? Cela n'a pas empêché des candidats de prétendre y arriver. De plus, certains, pour essayer de montrer que $\sum P(S_n = 0)x^n$ converge, majorent

$$\left|\sum_{n=0}^{+\infty} P(S_n = 0)x^n\right|$$

Cela n'a aucun sens, car précisément on ne sait pas si cette série converge.

Pour montrer que F est définie en 1, certains utilisent $F(1)$ comme un objet (volant) identifié. Certains se sont fourvoyés en affirmant que $\sum_{n=0}^{+\infty} P(S_n = 0) = 1$. Beaucoup pensent qu'une série entière converge uniformément ou normalement sur son disque ouvert de convergence. Le développement en série entière de la fonction $x \mapsto \frac{1}{1-x}$ ou de la fonction exponentielle fournissent pourtant des contre-exemples. Certains affirment que «*puisque G est définie en 1, elle y est continue*» : doit-on en conclure que la partie entière, définie en 1, est continue en 1 ?

- La **question 7** était la première question réellement difficile du sujet (une copie sur dix a donc donné une réponse satisfaisante à la première partie de cette question). Parmi les réponses fausses, certains ont transformé l'évènement $(R = n)$ en l'évènement $(S_n = 0_d)$ alors que ce sont des évènements différents, de même $(S_n - S_k = 0)$ est remplacé par $(S_{n-k} = 0)$ (confusion entre égalité d'évènements et évènements de même probabilité). D'autres ont écrit une formule de probabilité conditionnelle (avec le risque de faire une division par zéro) :

  $$P_{(R=k)}(S_n - S_k = 0_d) \qquad \text{ou} \qquad P(S_n - S_k = 0_d | R = k)$$

  Étant gênés par $(R = k)$, ils l'ont ensuite fait disparaître mystérieusement sans explication. Certains candidats invoquent un mystérieux évènement conditionnel $(A|B)$ (l'évènement «A sachant B») ce qui n'a aucun sens. Pour traiter la question correctement, il fallait justifier que les évènements $(R = k)$ et $(S_n - S_k = 0_d)$ étaient indépendants en invoquant les variables aléatoires en jeu et le résultat communément désigné sous le nom de lemme des coalitions.

  Pour la seconde partie, un certain nombre de candidats ont affirmé que les évènements $(R = k)$ pour $k \in [\![1\,; n]\!]$ forment un système complet d'évènements. Or c'est faux, car $R(\Omega) = \mathbb{N}^* \cup \{+\infty\}$. Encore une fois, affirmer quelque chose parce que cela permet d'arriver au résultat demandé ne permet pas d'avoir de points si l'affirmation en question est fausse.
- Beaucoup ont reconnu un produit de Cauchy à la **question 8**, mais tous ne l'ont pas bien justifié. De plus, il y avait une difficulté (surmontable) avec les indices, certains se sont donc trompés mais ont miraculeusement trouvé la formule demandée au prix de modifications des expressions au gré des lignes, le jury n'a évidemment pas été dupe. Pour discuter de la limite, les candidats utilisent la relation algébrique qui relie F et G pour en déduire la même relation algébrique entre les limites de F et de G en 1 sans savoir si ces limites existent (sans compter le risque de tomber sur une forme indéterminée). C'est comme si ces candidats écrivaient :

  $$\text{«}\forall x \in \mathbb{R} \quad \cos(x)^2 + \sin(x)^2 = 1 \qquad \text{donc} \qquad \lim_{x\to+\infty} \cos(x)^2 + \lim_{x\to+\infty} \sin(x)^2 = 1\text{»}$$

  Il suffisait pourtant d'exprimer $F(x)$ en fonction de $G(x)$ puis de passer à la limite grâce à la continuité de G en 1 (peu souvent mentionnée), ainsi que les opérations sur les limites usuelles (quasiment jamais mentionnées). Pour certains candidats, il n'y a étrangement que deux possibilités : $G(1) = 0$ ou $G(1) = 1$, ce qui a laissé le jury dubitatif.
- La **question 9** n'était pas facile. Beaucoup ont cru la court-circuiter grâce à un théorème de double limite avec de la convergence uniforme/normale sur un voisinage de 1 (relire à ce sujet le commentaire de la question 6).
- Une fois encore, certains n'ont pas vu la difficulté de la **question 10** et considèrent comme évident que $\sum P(S_n = 0)$ diverge si et seulement si $\lim\limits_{x\to1^-} F(x) = +\infty$, comme si ces deux objets avaient la même nature alors que l'un est la limite d'une suite et l'autre est celle d'une fonction.
- La première partie de la **question 11** était une question difficile, seul un candidat sur vingt a donné une réponse satisfaisante. Beaucoup ont voulu montré que les évènements $(Y_i = 1)$ et $(R > i)$ sont égaux, alors que ce résultat est faux. En effet, on peut visiter à la i-ième étape une nouvelle position tout en étant déjà retourné à 0, par exemple si $d = 1$ avec $S_0 = 0$, $S_1 = 1$, $S_2 = 0$, $S_3 = -1$, alors l'évènement $(Y_3 = 1)$ est réalisé contrairement à $(R > 3)$. La deuxième partie n'a pas eu beaucoup plus de succès, beaucoup écrivent l'égalité demandée avec d'obscures explications, alors que pour la prouver, il aurait été plus sage de justifier que

  $$N_n = 1 + \sum_{k=1}^{n} Y_k$$

  puis d'utiliser la linéarité de l'espérance, puis que $E(Y_i) = P(Y_i = 1)$. Finalement, là où il faut trois étapes, les candidats la présentent en une seule.

- Lors de la **question 12**, les candidats ont quasiment tous vu que pour conclure, il suffisait d'appliquer le théorème de Cesàro. Ils affirment donc, **sans preuve** pour la plupart, que

$$\lim_{n\to+\infty} P(R>n) = P(R=+\infty)$$

  Alors qu'il y a évidemment un résultat de cours qui permet de déterminer la limite de $(P(A_n))_n$ pour $(A_n)_n$ une suite décroissante d'évènements.
- La démarche suivie par les candidats à la **question 13** était généralement la bonne. Cependant, les explications n'étaient pas toujours convaincantes. L'indépendance mutuelle des variables aléatoires $(X_n)_n$ n'a pas été souvent invoquée alors qu'elle est indispensable pour le calcul de $P(S_{2n}=0)$.
- La **question 14** a été peu souvent abordée. Pourtant, en utilisant la question précédente, on pouvait exprimer F comme une somme. On reconnaissait le développement en série entière de $\sqrt{1-4pqx^2}$. Attention à ne pas oublier de vérifier que $|4pqx^2|<1$! En utilisant le résultat de la question 8, on obtenait l'expression de $G(x)$ pour $x\in]-1,1[$. Les candidats oublient trop souvent de mentionner la continuité de G et de $x\mapsto\sqrt{1-4pqx^2}$ en 1, pour en déduire la valeur de $P(R=+\infty)$.
- La **question 15** a aussi été peu abordée car il fallait avoir résolu la question 14 pour la traiter.
- La **question 16** était d'une simplicité désarmante, il suffisait pour obtenir la première inégalité de minorer a_k par a_n dans l'égalité $\sum_{k=0}^{n} a_k b_{n-k}=1$ en précisant bien que la multiplication par b_{n-k} conserve l'inégalité car $b_{n-k}\geq 0$, ce dernier point étant trop peu souvent cité.
- La **question 17** était aussi abordable. Le jury espérait un encadrement propre de a_nB_n à l'aide de la question 16, puis un passage à la limite des deux côtés de l'encadrement et une conclusion par le théorème d'encadrement. Hélas, les candidats ont préféré user de leur imagination pour étaler leur mauvaise compréhension des petits o, des équivalents et des limites. Petit florilège de pseudo-résultats utilisés pour conclure :
  — $u_n\leq v_n$, donc $\lim\limits_{n\to+\infty} u_n \leq \lim\limits_{n\to+\infty} v_n$.
  — $u_n\sim v_n$ et $v_n\leq 1$, donc $u_n\leq 1$.
  — $v_n\leq u_n\leq w_n$ et $w_n\sim v_n$, donc à partir d'un certain rang, $v_n\leq u_n\leq v_n$.
  — $u_n\to 0$ et $v_n\sim w_n$, donc $u_n+v_n\sim 0+w_n$.
  — $u_n\sim v_n$, donc $\lim\limits_{n\to+\infty} u_n = v_n$.
  On a écrit ici les résultats avec $(u_n)_n$, $(v_n)_n$ et $(w_n)_n$ au lieu des (a_n), (B_n) et (B_{m-n}) utilisés dans les copies. Peut-être, les candidats verront-ils mieux leurs erreurs ainsi. Exercice conseillé : justifier que toutes ces affirmations sont fausses.

Les quatre dernières questions ont donné des résultats satisfaisants sur environ une copie sur cent.

- La **question 18** était difficile. Néanmoins, l'équivalent de b_n permettait facilement, via un théorème de sommation des relations de comparaison, d'affirmer que $B_n\sim C\ln(n)$, en vérifiant, **bien sûr**, les hypothèses dudit théorème. Le cœur de la question consistait à trouver une suite $(m_n)_n$ telle que m_n-n soit assez proche de n pour avoir $B_{m_n-n}\sim B_n$, et en même temps que m_n soit proche de m_n-n pour avoir $B_{m_n}-B_{m_n-n}\to 0$. En particulier, $m_n=2n$ ne fonctionnait pas car $B_{m_n}-B_{m_n-n}$ est de l'ordre de $\sum_{k=n+1}^{2n} k^{-1}$ qui ne tend pas vers 0, contrairement à ce que croient de nombreux candidats, mais vers une limite finie non nulle bien connue. Très peu de candidats ont donc proposé une suite $(m_n)_n$ convenable.
- À l'instar des questions de probabilités 7 et 11, la **question 19** était difficile, mais a été résolue avec brio par quelques candidats, soit par raisonnement probabiliste direct, soit par récurrence sur n.
- La **question 20** a été traitée par beaucoup de candidats, qui ont cru, à tort, pouvoir utiliser la question 13. En notant A_n (respectivement B_n) l'abscisse (respectivement l'ordonnée) de S_n, ces candidats ont affirmé, sans preuve ni explication, que «*A_n et B_n sont indépendantes*». Pourquoi une telle affirmation ? Parce que cela arrangeait bien les candidats, pardi ! Malheureusement pour ces candidats, A_n et B_n n'étaient pas indépendantes (laissé en exercice au lecteur).

2.3.4 Conclusions

Cette épreuve était progressive, les six premières questions étaient abordables et permettaient aux candidats sérieux de gagner des points. Malheureusement, par manque de connaissances ou par volonté d'aller vite, un certain nombre de candidats n'a pas traité ces questions assez sérieusement. Nous conseillons donc aux candidats d'être attentifs à ce premier groupe de questions plutôt que d'aller tenter une ou deux questions faisables mais éparpillées dans le sujet ou de ne prendre son temps que pour les questions difficiles.
Ce sujet permettait de faire une distinction entre les excellents candidats et les élèves sérieux, mais aussi entre les candidats sérieux et ceux dont le travail pendant deux ans a pu manquer d'intensité. À l'opposé de copies très faibles, certaines excellentes, mais très rares, ont abordé tout le sujet de façon correcte.

2.4 Mathématiques 1 - filière PC

2.4.1 Présentation du sujet

Soient $\mathbb{K}$ un corps, $n \in N^*$. Le sous-espace $\mathcal{N}_n(\mathbb{K})$ de l'algèbre $\mathcal{M}_n(\mathbb{K})$ constitué des constitué des matrices triangulaires supérieures strictes est d'une part constitué de matrices nilpotentes, d'autre part de dimension $\binom{n}{2}$. Les deux résultats suivants ont été établis par Gerstenhaber en 1958 si $\mathbb{K}$ est de cardinal supérieur ou égal à n, puis étendus par Serezhkin au cas général en 1982.

Théorème A. Tout sous-espace de $\mathcal{M}_n(\mathbb{K})$ constitué de matrices nilpotentes est de dimension au plus $\binom{n}{2}$.

Théorème B. Tout sous-espace de $\mathcal{M}_n(\mathbb{K})$ constitué de matrices nilpotentes et de dimension $\binom{n}{2}$, est conjugué à $\mathcal{N}_n(\mathbb{K})$ dans l'algèbre $\mathcal{M}_n(\mathbb{K})$.

En revanche, dès que $n \geq 3$, il existe des sous-espaces de $\mathcal{M}_n(\mathbb{K})$ constitués de matrices nilpotentes qui ne sont conjugués à aucun sous-espace de $\mathcal{N}_n(\mathbb{K})$. Autrement dit, il existe des sous- espaces de $\mathcal{M}_n(\mathbb{K})$ constitués de matrices nilpotentes, maximaux au sens de l'inclusion pour cette propriété, mais de dimension strictement inférieure à $\binom{n}{2}$.
La démonstration du théorème A pour $\mathbb{K} = \mathbb{R}$ a fait l'objet d'un problème au CCMP (PSI, Maths II, 2016). Le but du présent texte est d'établir le théorème B, toujours pour $\mathbb{K} = \mathbb{R}$. En fait, la démonstration du théorème A est implicite dans le présent sujet. L'argumentation suit celle donnée dans les parties 2 et 3 de l'article *The structured Gerstenhaber problem (II), Linear Algebra and its Applications*, vol. 569, 2019, pp. 113-145.
La limitation à $\mathbb{K} = \mathbb{R}$ permet de contourner les arguments de dualité via la structure euclidienne canonique de $\mathbb{R}^n$; avec des modifications évidentes, la preuve s'adapte au cas traité par Gerstenhaber [1].

2.4.2 Commentaires généraux

Le sujet, d'un grand intérêt mathématique, met en jeu une partie très significative des programmes d'algèbre de PCSI et PC : algèbre linéaire élémentaire, espaces euclidiens, réduction. Il est très long, d'un niveau conceptuel élevé, notamment par une rédaction privilégiant systématiquement le point de vue géométrique sur celui des matrices et la nécessité qu'il impose d'absorber un grand nombre de notations. Il comporte cependant un certain nombre de questions simples. Il a permis de tester la connaissance et la compréhension du cours des candidats, leur niveau d'abstraction, ainsi que, pour les meilleurs d'entre-eux, leur capacité à rentrer dans une démonstration subtile.
Les meilleurs candidats ont bien compris le problème. Une partie significative a produit une copie substantielle. L'étalonnage des notes est assez satisfaisant. Les correcteurs déplorent cependant un

1. La restriction sur le corps vient de l'utilisation d'arguments polynomiaux.