Composition de Mathématiques B, Filière MP
(X)

1 Présentation du sujet

Cette épreuve portait sur une marche aléatoire $(X_n)_{n\in\mathbb{N}}$ à pas positifs, indépendants, identiquement distribués. L'objectif était d'estimer combien de temps il faut, en moyenne, pour traverser un intervalle $[x, x+\ell]$. Le pas n'étant pas minoré uniformément par un réel strictement positif, il était parfaitement envisageable que la marche mette "beaucoup de temps" à traverser un intervalle donné. On démontrait, à un niveau de généralité assez grand, que l'espérance du temps de traversée de l'intervalle $[x, x+\ell]$ est asymptotiquement égale à ℓ/E où E est l'espérance d'un pas de la marche. Ce résultat porte le nom de *petit théorème de renouvellement*.

L'objet de la première partie était de mettre en place les notations du problème. Les questions 1a-3b pouvaient vraiment être considérées comme des questions de compréhension directe du cours : savoir calculer la probabilité d'un événement ; savoir calculer une espérance ; reconnaître et savoir redémontrer l'inégalité de Chebychev ; savoir utiliser l'inégalité de Markov ; savoir reconnaître une suite géométrique ; connaître le critère de convergence des suites géométriques. La première "vraie" difficulté apparaissait à la dernière question de cette partie (question 3c) où un raisonnement probabiliste était indispensable.

La deuxième partie visait à étudier la convergence et l'équation fonctionnelle satisfaite par la fonction f définie pour tout réel x par :

$$f(x) = \sum_{k=0}^{\infty} \mathbb{E}\left[g\left(x - \sum_{i=1}^{k} X_i\right)\right]$$

lorsque g est une fonction positive à support dans $[0, K]$. On démontrait que f est l'unique solution bornée et supportée dans $\mathbb{R}_+$ de l'équation

$$f = g + \sum_{i=0}^{\infty} p_i f(\cdot - x_i)$$

où les x_i sont les valeurs prises (presque sûrement) par chaque pas de la marche, avec leurs probabilités respectives $\mathbb{P}(X_k = x_i) = p_i$. On démontrait aussi que, pour tout nombre réel x,

$$f(x) = \sum_{i=0}^{\infty} q_i g(x - y_i)$$

où les y_i sont les valeurs prises (presque sûrement) par la marche $S_k = \sum\limits_{i=1}^{k} X_i$ et les q_i des coefficients que l'on calculait.

La troisième partie était complètement indépendante des parties 1 et 2. On y étudiait les parties non-vides de $]0,+\infty[$ qui sont stables par addition en utilisant des techniques similaires à celle de l'étude des sous-groupes additifs de $\mathbb{R}$. L'objectif était de démontrer que les parties non-vides de $]0,+\infty[$ qui sont stables par addition sont soit des sous-ensembles de $d\mathbb{Z}$ pour un certain $d > 0$, soit des ensembles qui sont asymptotiquement denses (l'ensemble est arbitrairement proche de tout nombre suffisamment grand). Dans ce dernier cas, on peut calculer la limite en $+\infty$ d'une fonction uniformément continue sur $]0,+\infty[$ en l'évaluant uniquement sur l'ensemble en question.

La dernière partie démontrait que, sous l'hypothèse

$$\forall d \in [0,+\infty[, \qquad \mathbb{P}(X \in d\mathbb{Z}) = \sum_{x_i \in d\mathbb{Z}} p_i < 1$$

l'espérance du nombre d'étapes nécessaires pour traverser $[x, x+\ell]$ est asymptotiquement égale à ℓ/E où E est l'espérance d'un pas de la marche. Le point clé était l'identité :

$$\lim_{x\to+\infty} \left(\sum_{k=0}^{\infty} \mathbb{E}\left[g\left(x - \sum_{i=1}^{k} X_k \right) \right] \right) = \frac{1}{E} \int_0^{\infty} g(t)dt$$

qu'on démontrait pour une classe suffisamment large de fonctions tests g supportées dans $[0,+\infty[$.

2 Statistiques générales et indications sur le barème

Les résultats de l'épreuve sont en accord avec les directives statistiques générales du concours, ce qui garantit l'influence respective des différentes épreuves.

$0 \leq \text{N} < 4$	106	7,10 %
$4 \leq \text{N} < 8$	600	40,21 %
$8 \leq \text{N} < 12$	524	35,12 %
$12 \leq \text{N} < 16$	188	12,60 %
$16 \leq \text{N} \leq 20$	74	4,96 %
Total	1492	100 %
Nombre de copies : 1492		
Note moyenne : 8,71		
Écart-type : 3,75		

L'histogramme de répartition des notes est représenté sur la figure 1. La ligne verte continue indique le nombre de places offertes dans la filière MP (environ 180) et la ligne verte discontinue indique le nombre de candidats admissibles dans la filière MP (environ 400). La boite verte identifie le peloton de tête. L'histogramme cumule les candidats français et étrangers.

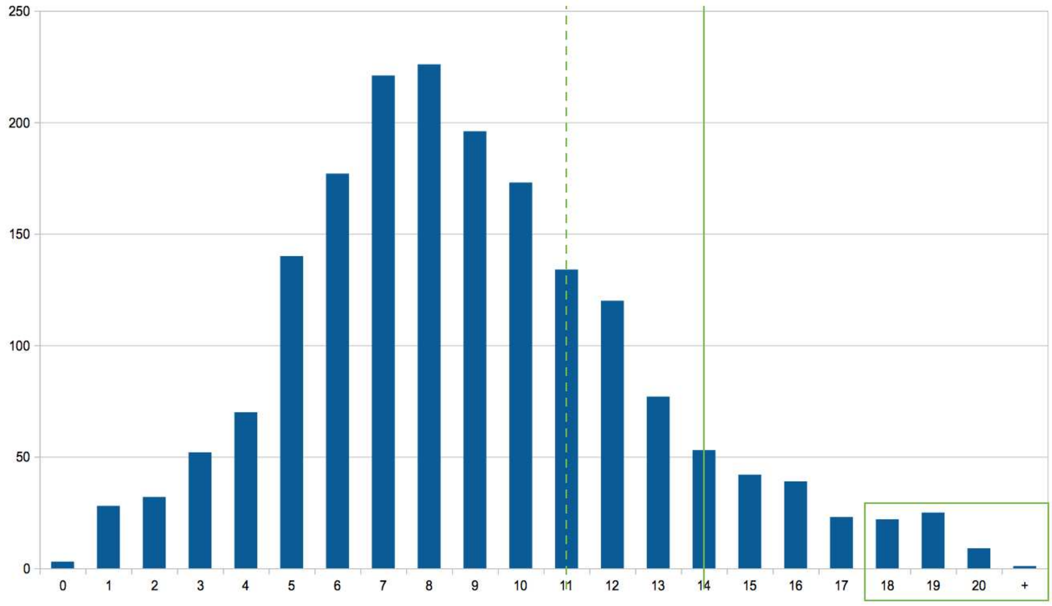


Fig.1 – En abscisse la note et en ordonnée le nombre de candidats ayant obtenu cette note.

Le barème a été conçu selon trois objectifs : séparer les candidats dans la zone d'admissibilité, éviter les effets de seuil et minimiser les points de "grappillage". En particulier, les questions "faciles" de la partie 4 n'étaient que très peu valorisées, surtout si elles étaient abordées isolément.

Le tableau suivant indique la proportion de points attribués à chaque partie. Nous avons décidé d'attribuer la note de 18/20 à un candidat ayant obtenu environ 50% des points du barème. Les 57 candidats ayant obtenu une note entre 18 et 20 ont donc fourni une quantité de travail gigantesque par rapport aux autres. Par exemple, le meilleur candidat a traité 93% du sujet.

	Proportion dans le barème	Barème cumulé
Partie 1	17%	17%
Partie 2	33%	50%
Partie 3	17%	67%
Partie 4	33%	100%

En traitant toutes les questions des parties 1 et 2, un candidat pouvait obtenir 17.7 points sur 20. En traitant correctement toutes les questions de la partie 1 puis de la partie 3, sans aborder la partie 2, un candidat pouvait obtenir 12.4 points sur 20.

3 Commentaires sur les résultats

Contrairement à l'an dernier, la proportion de copies correctes est redevenue raisonnable. Les candidats ont donc eu les moyens *d'acquérir des méthodes et une certaine intuition probabiliste* au cours de leur préparation. Ce point est encourageant quant à l'intégration des nouveaux programmes dans la formation des candidats.

Par contre, la *faiblesse sur les questions élémentaires d'analyse* reste préoccupante. La figure 2 illustre ce problème : parmi les 293 meilleurs candidats (ayant obtenu une note au moins égale à 12/20), moins de 40% d'entre eux ont obtenu un score supérieur à 60% sur les trois questions d'analyse 9a-9c. Pourtant, la plupart ont continué le sujet bien au-delà.

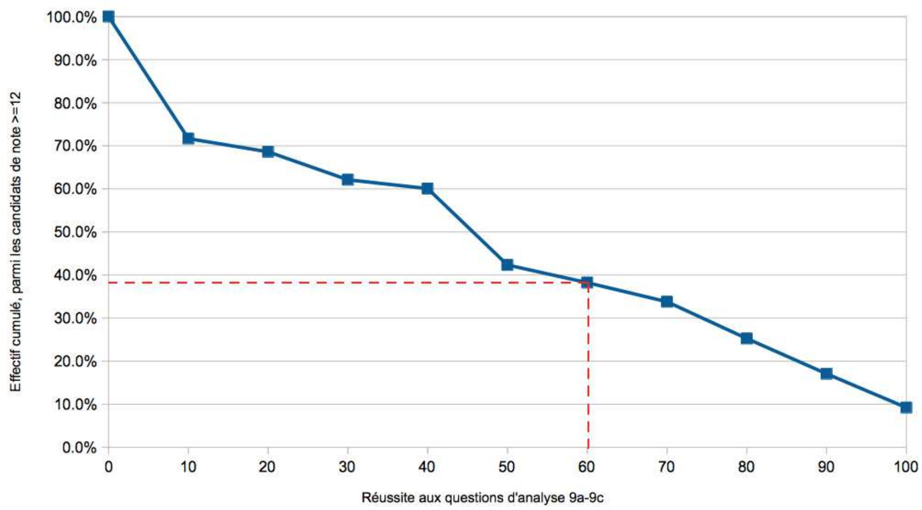


Fig. 2 – Effectif cumulé de réussite aux questions d'analyse 9a-9c, parmi les candidats ayant obtenu une note au moins égale à 12/20.

Nous avons décidé cette année de pousser l'analyse statistique un peu plus loin en tentant de *corréler l'apprentissage du cours à la note finale du candidat*. La connaissance du cours est estimée par le taux de réussite au groupe de questions 1a-3b qui sont, soit explicitement un point du programme (question 1b), soit des questions d'application directe du cours et des méthodes vues en cours.

La corrélation est présentée sur la figure 3 et appelle plusieurs commentaires.

— Une corrélation positive existe clairement entre la connaissance du cours et la réussite globale : améliorer sa connaissance du cours de 10% rapporte en moyenne 1 point sur 20 au candidat.

— La "zone impossible" est clairement visible en rouge : il est impossible d'espérer valider l'épreuve si le cours n'est pas maîtrisé. La connaissance du cours est indispensable à la réussite. Les 295 candidats (16% des présents) ayant obtenu un score inférieur à 40% en connaissance du cours devraient peut-être se poser des questions quant à la pertinence de leur candidature...

— La "zone verte" indique une estimation du nombre d'admissibles (400 candidats) et d'admis (180 candidats) dans la filière MP. Bien sûr, l'admissibilité et l'admission ne se jouent pas sur une seule épreuve de l'écrit, mais il est néanmoins très clair que le niveau requis pour l'admissibilité et l'admission représente une valeur ajoutée qui va bien au-delà de la simple connaissance du cours. Par exemple, 240 candidats (13% des présents ; boite jaune) ont obtenu un score supérieur à 95% aux questions de cours mais n'ont pourtant pas pu assurer une note globale supérieure à 11.5, ce qui leur laisse peu de chance d'être admissible.

— Chaque point de la figure 3 représente le profil d'au moins un candidat. Par contre, la multiplicité d'un même profil n'est pas représentée (plusieurs candidats par point). Il est bon de prendre conscience que 189 candidats (soit 10% des présents) ont obtenu un score supérieur à 95% aux questions de cours tout en assurant une note globale supérieure à 11.5/20 (bande vert foncé) ; ce vivier est suffisant pour pourvoir les places disponibles...

Enfin, comme les correcteurs sont exaspérés par le nombre de candidats qui prennent le Concours d'Admission pour un jeu de Loto®, nous avons souhaité analyser précisément *le biais qualité / quantité* dans le travail fourni par les candidats lors de l'épreuve. La quantité est simplement estimée par le nombre de points de contrôle abordés par le candidat, la qualité est estimée par la note finale obtenue par le candidat.

La corrélation est présentée sur la figure 4 où plusieurs zones se distinguent nettement.

— Les candidats qui privilégient la qualité sur la quantité (zone verte) sont en fait assurés de réussir au concours. Un candidat situé dans cette zone a le bon état d'esprit face au travail et ne peut que progresser. Il a l'humilité de reconnaître qu'il ne sait pas traiter certaines questions et se concentre sur un exposé clair et précis des questions qu'il sait être à sa portée. L'Ecole Polytechnique recherche ce profil d'étudiant, humble, honnête et sérieux.

— Les candidats de la zone jaune ont fait le choix de traiter un maximum de questions, au dépend de la qualité de chaque réponse individuelle. Cette attitude est en fait un frein très sérieux à la capacité de progression du candidat. Ne sachant pas faire lui même la différence entre une bonne réponse et un argument incomplet, un candidat de ce profil aura du mal à bénéficier pleinement de l'excellence des enseignements de l'Ecole. Ces candidats doivent impérativement recentrer leurs priorités et adopter une stratégie privilégiant la qualité des réponses sur la quantité.

— Les 450 candidats de la zone rouge (24% des présents) : ces candidats n'ont tout simplement pas leur place au concours et gaspillent nos ressources. A un candidat qui n'obtient presque aucun point sur les questions de cours mais va "traiter" toutes les questions de la fin du sujet en trois lignes chacune nous aimerions demander : « Pensez-vous vraiment qu'il suffit de si peu d'effort pour résoudre un problème complexe ? Pensez-vous vraiment que vous pouvez intégrer une école d'excellence sans savoir distinguer le vrai du faux et sans avoir fourni le minimum de travail personnel nécessaire ? » Franchement, ce n'est pas de l'ambition, c'est de l'arrogance.

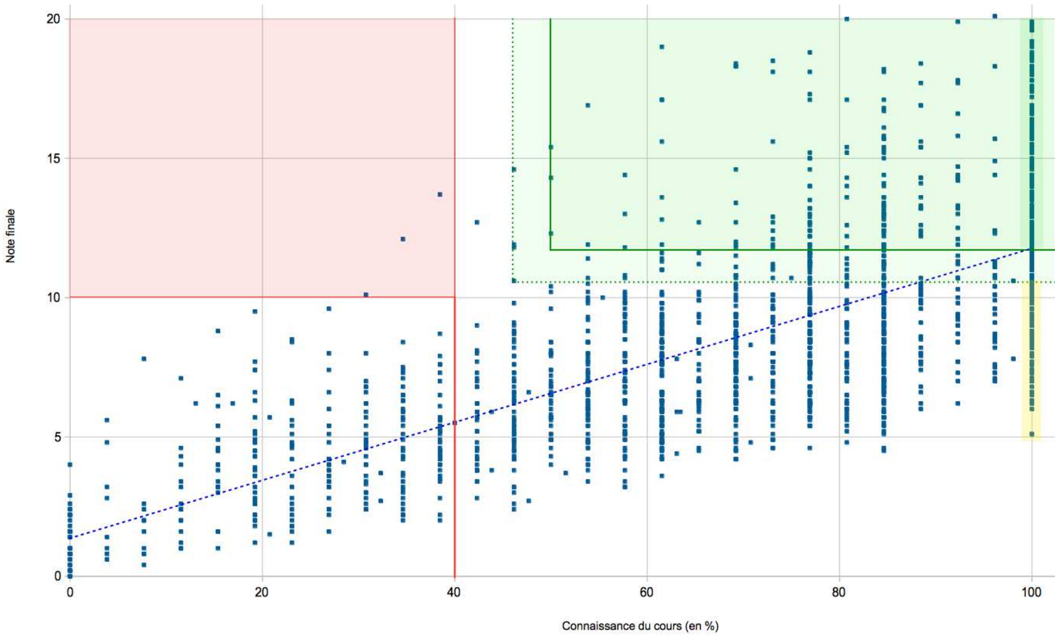


Fig. 3 – Corrélation entre la connaissance du cours et la note finale du candidat.

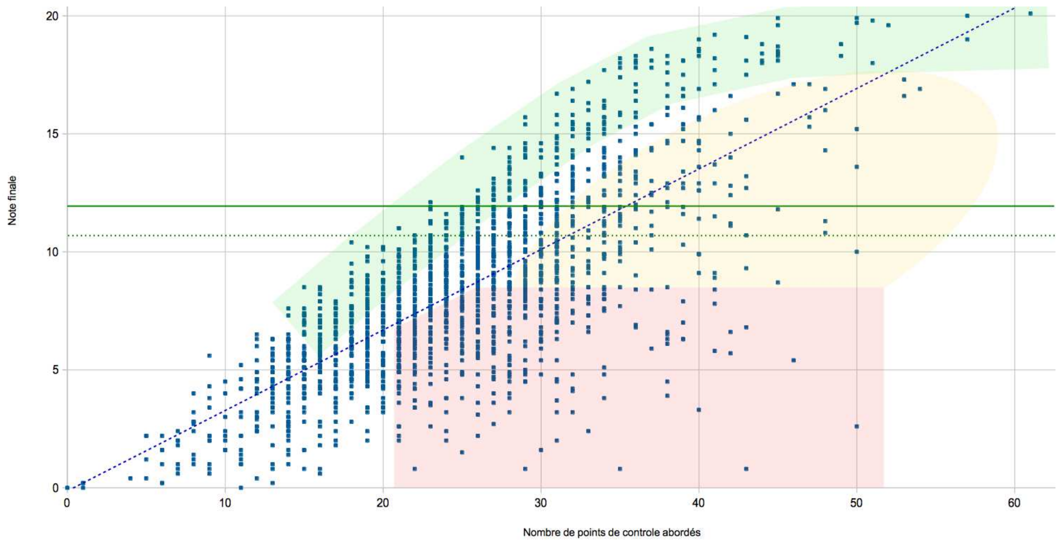


Fig. 4 – Corrélation entre le nombre de questions traitées et la note finale.

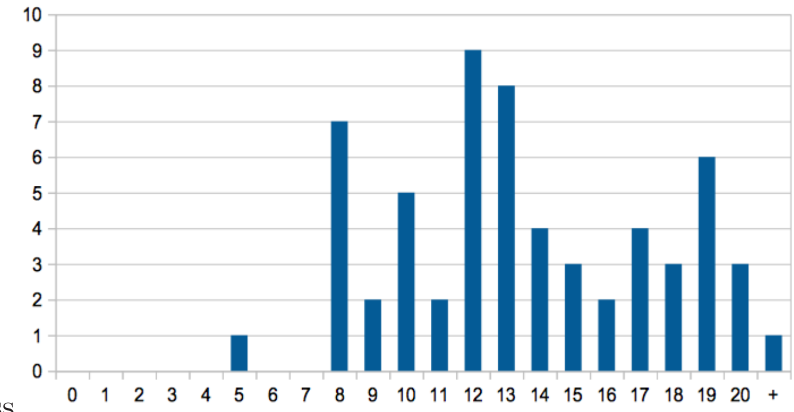
Examen détaillé des questions

Première partie

Question 1a Ces deux égalités ensemblistes ainsi que l'inclusion sont des conséquences immédiates de la stricte croissance (presque sûre) de la suite $(S_n)_{n\geq 0}$ et du fait que $S_0 = 0$. Si la plupart des candidats ont remarqué cette propriété, il s'agissait ensuite de rédiger proprement la démonstration de la première égalité par double inclusion. Un nombre non négligeable de candidats rédige l'une des inclusions et conclut immédiatement à l'égalité ensembliste sans même mentionner l'autre inclusion ce qui leur coûtait les points prévus pour cette question.

Notons qu'en toute rigueur, les variables aléatoires X_k peuvent prendre des valeurs dans $\mathbb{R}\backslash\{x_i | p_i > 0\}$, mais que cet événement est de probabilité nulle. On peut donc, soit signaler qu'on peut remplacer les variables X_k par des variables de même loi, à valeurs dans $\{x_i | p_i > 0\}$, soit signaler que les identités demandées ne sont vraies que presque sûrement, ce qui ne changeait rien à leur utilisation dans les questions suivantes, ces événements étant ensuite considérés aux travers de leurs probabilités. Cependant, comme le programme précise que *tout développement sur les événements négligeables ou presque sûrs est hors programme*, les concepteurs du sujet, les inspecteurs généraux ayant testé le sujet et les correcteurs sont unanimes pour ne pas pénaliser un candidat qui considère que chaque variable X_k ne peut prendre que les valeurs $\{x_i | p_i > 0\}$.

Nous avons cependant décidé d'accorder un bonus aux candidats qui ont utilisé (à bon escient) des identités « presque sûres », ici ou ailleurs. Seuls 60 candidats (3% des présents) ont identifié la subtilité et donc bénéficié de ce bonus.



ss

Fig. 5 – Histogramme des notes des candidats ayant utilisé l'expression « presque sûrement » a bon escient.

Il est intéressant de remarquer que, même si l'utilisation correcte de l'expression « presque sûrement » est un gage de qualité, elle ne suffit pas pour traiter correctement un problème de probabilités...

Question 1b Cette question pouvait se voir comme une conséquence de la loi faible des grands nombres. Au vu de sa place en début d'épreuve et le commentaire du programme officiel qui indique explicitement que les étudiants doivent savoir retrouver l'inégalité $\mathbb{P}\left(|S_n/n - m| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$, il fallait considérer cette question comme une question de cours et donc redémontrer l'inégalité à partir de l'inégalité de Bienaymé-Tchebychev.

Question 2 Cette question a dans la majorité des cas été traitée soit par le théorème de sommation par paquets (théorème de Fubini pour les familles sommables), soit par une somme téléscopique utilisant le fait que Y est supposée à valeurs entières. La sommabilité dans le premier cas et la convergence vers 0 du reste de la série à termes positifs dans le second sont toutes deux une conséquence du fait que Y admet une espérance.

Cette question, ainsi que les questions $4b$, 5, $6b$ et $8c$ pointent un manquement du programme. Toutes ces questions peuvent bien entendu se traiter dans le cadre du programme officiel, mais elles gagneraient à être traitées de manière plus naturelle à l'aide du théorème de sommation par paquets pour les familles à termes positifs. Il nous semble que l'absence de ce résultat conduit à des difficultés inutiles pour justifier des convergences dans le cadre des probabilités.

Question 3a Bien que cette question vienne juste après la question 2, on ne pouvait l'utiliser puisque les variables considérées n'étaient pas à valeurs entières. On pouvait éventuellement utiliser des parties entières (et être très rigoureux dans les inégalités utilisées ensuite), mais c'est à l'inégalité de Markov qu'il fallait penser en premier lieu. Certains candidats ont fait appel à l'inégalité de Cauchy-Schwarz. Il faut, dans ce cas, justifier que les variables sont de variance finie. La deuxième partie de la question relevait du théorème des coalitions et utilisait donc l'indépendance des variables aléatoires X_k.

Question 3b Cette question se déduisait immédiatement de la précédente une fois démontré que l'espérance de X est strictement plus petite que 1. La deuxième partie de la question revenait alors à identifier une série géométrique dont on calculait la somme.

Question 3c Cette question représentait la première véritable difficulté de l'épreuve. L'inégalité $\mathbb{P}(A \cap B) \leq \mathbb{P}(A)\mathbb{P}(B)$ est fausse en général. Il suffit de considérer un évènement A de probabilité $1/2$ et de prendre $A = B$.

La notion de probabilité conditionnelle ne semble pas maîtrisée au-delà des systèmes complets d'évènements. En particulier, Il n'est pas évident de comparer des probabilités conditionnelles et des probabilités inconditionnelles en utilisant des inclusions d'évènements. Plus généralement, alors que les candidats semblent bien savoir ce qu'ils ont le droit de faire avec des variables aléatoires indépendantes (cf question $3a$), ils ne savent pas, pour la plupart, identifier quels évènements mettant en jeu ces variables aléatoires sont indépendants. Ainsi, les évènements $(S_{n-1} < x \leq S_n)$ et $(N(0,\ell) \geq k)$ n'ont aucune raison d'être indépendants, même si pour de nombreux candidats, le fait qu'ils ne dépendent pas des mêmes variables parmi n, x et k suffit à les rendre indépendants...

Deuxième partie

Question 4a Cette question ne posait aucune difficulté et a été réussie par presque tous les candidats, même les plus faibles.

Question 4b Les examinateurs sont agréablement surpris par le bon taux de réussite pour cette question. On notera cependant que les évènements $N(x-K,x) > k$ et $x-K < S_k < x$ ne sont pas la même chose, contrairement à ce qui a pu être écrit dans certaines copies. Par ailleurs, la linéarité de l'espérance ne peut servir de justification à l'interversion de l'espérance avec une somme infinie.

Question 4c Cette question a été très bien réussie, la plupart des candidats ayant constaté que puisque g est à support compact dans $[0, K]$, g est majorée sur $\mathbb{R}$ par $||g||_\infty \mathbf{1}_{[0,K]}$.

Question 4d Certaines copies confondent le support et l'image d'une fonction. En particulier, une fonction positive n'a, a priori, pas de raison d'avoir un support positif. Une autre erreur récurrente est de penser que puisque g est à support dans $[0, K]$, f l'est aussi ou encore que f est à support compact.

Question 5 Il s'agissait d'appliquer la formule de transfert. Relevons que pour un certain nombre de candidats, toutes les espérances sont, d'après la formule de transfert, des sommes sur l'ensemble des entiers naturels, que les variables aléatoires soient à valeurs entières ou non. Il a aussi été vu des expressions du type « espérance de l'espérance de » qui formellement permettait d'arriver à la formule demandée mais qui n'a pas réellement de sens dans le contexte étudié.

Question 6a Il était possible d'utiliser la question 5 en l'appliquant au bon couple de variables aléatoires indépendantes, à savoir (S_n, X_{n+1}). Comme à la question 3*c*, on observe des confusions autour de la notion de variables aléatoires indépendantes. Les variables X et S_n ne sont pas a priori indépendantes, même si X_{n+1} suit la même loi que X...

Question 6b Il fallait bien entendu justifier d'une manière ou d'une autre l'interversion limite - somme infinie, ce qui pouvait par exemple s'obtenir en utilisant le fait que la somme des p_i vaut 1.

Question 7a Cette question se traitait de façon tout à fait similaire à la question 6*a* et nous avons pu y retrouver le même type d'erreurs.

Question 7b Cette question était la seconde question réellement difficile du sujet après la 3*c* et n'a été réussi que par les meilleurs candidats. Il s'agissait de majorer $h(x)$ en valeur absolue par $||h||_\infty \mathbb{P}(S_n \leq x)$ puis d'utiliser la question 3*b*. Une erreur courante était d'écrire que (S_n) tend vers $+\infty$ puisqu'elle est strictement croissante (ce qui laisse songeur quant aux connaissances de base sur les suites réelles de certains candidats) puis de passer à la limite sans justification sous l'espérance. Plus subtil, le résultat de la question 3*b* ou le fait que l'espérance de S_n (qui vaut nm) tend vers $+\infty$ ne permet pas de justifier que S_n tend vers $+\infty$, même presque sûrement. Ce qui est vrai mais hors programme est qu'une sous-suite de (S_n) tend alors vers l'infini presque sûrement.

Question 7c La fonction g ou la fonction nulle ne sont pas les uniques fonctions recherchées. Il s'agit bien entendu de f obtenue à la question $4a$.

Question 8a Étonnamment, de nombreux candidats oublient de montrer le second point, à savoir que l'ensemble Λ_X est inclus dans $\mathbb{R}_+$.

Question 8b C'est une application directe de la formule de transfert.

Question 8c Il s'agissait de passer à la limite dans la formule précédente puis de justifier l'interversion des deux sommes infinies. La seconde formule pouvait s'obtenir par une application directe du résultat de la question $4b$.

Question 9a Il s'agit de la première des trois questions permettant de vérifier les compétences en analyse des candidats. Cette question, dont la principale difficulté tenait à sa rédaction, n'a été réussie que par 11% des candidats. Une erreur courante est de ne pas se placer sur un intervalle compact quelconque, mais de majorer directement la somme des q_i indexée par les indices i tels que $y_i \in [x - K, x]$ en tout point x de $\mathbb{R}$ et ce sans se soucier de l'uniformité en x du majorant.

Question 9b La continuité uniforme sur tout segment n'entraîne pas la continuité uniforme sur $\mathbb{R}$. Par ailleurs, la fonction f n'est pas à support compact, sa continuité sur $\mathbb{R}$ qui est une conséquence directe de la convergence uniforme sur tout compact prouvée à la question précédente ne suffit donc pas à obtenir la continuité uniforme de f sur tout $\mathbb{R}$. Pour cette question on pouvait revenir à la définition de la continuité uniforme et utiliser la borne démontrée en $3c$.

Question 9c Cette question mettait en jeu plusieurs théorèmes classiques de l'analyse réelle et des séries de fonctions. Le caractère borné de g' est une conséquence du fait que g étant à support compact, le support de g' l'est aussi et est inclus dans celui de g. Le caractère borné de f' devait se démontrer directement en utilisant à nouveau la question $3c$. L'argument affirmant que g est lipschitzienne (ce qui est affirmé sans jamais être démontré) donc que g' est bornée n'est pas valable. Les hypothèses permettant d'intervertir la dérivation et la somme infinie ne sont bien rédigées que chez 12% des candidats. Sur ce type de questions, la faiblesse en analyse des candidats est préoccupante.

Troisième partie

Question 10a Cette question élémentaire a été réussie par près de 80% des candidats dont de nombreux étant passés directement de la question 5 à celle-ci en ne traitant pas l'essentiel de la seconde partie du problème.

Question 10b Il était inutile de chercher des exemples compliqués pour répondre à cette question, mais il fallait prendre garde à bien prendre des sous-ensembles de $\mathbb{R}_+^*$ et non simplement de $\mathbb{R}$. Par ailleurs, le nombre $r(\Lambda)$ mesure la distance minimale entre deux éléments de Λ et non pas le plus petit élément de Λ comme semblent le penser certains candidats. En particulier $r([1,+\infty[)$ vaut 0 et non pas 1...

Question 11a Cette question est une application directe de la définition de la borne inférieure. Il est inutile ici de distinguer dans la rédaction le cas où la borne inférieure est atteinte de celui où elle ne l'est pas.

Question 11b Cette question a été réussie par 20% des candidats. Pour y répondre, il suffisait de réaliser que si un réel se trouve dans un intervalle de longueur d, il est à une distance inférieure à $\frac{d}{2}$ de l'une des deux bornes de l'intervalle.

Question 11c Il était important de remarquer que les intervalles de la forme $[na+kd, na+(k+1)d]$ pour $k \leq n$ ne recouvrent pas nécessairement tout $\mathbb{R}^+$ mais la réunion contient $[n_0a, +\infty)$ ce qui est suffisant pour conclure.

Question 12a Cette question plus difficile n'a été réussie que par moins de 10% des candidats. Le résultat démontré dans cette question ne permet pas d'affirmer qu'un ensemble Λ tel que $r(\Lambda) = 0$ est dense dans $\mathbb{R}_+$ mais seulement qu'il est, en un certain sens, asymptotiquement dense ce qui est suffisant pour démontrer le résultat de la question suivante.

Question 12b Dans cette question, il fallait revenir à la définition de la continuité uniforme de f et utiliser la question 12a pour η un ε-module d'uniforme continuité pour f. Tout comme à la question 9b, nous constatons là encore que la plupart des candidats maîtrisent mal la notion de continuité uniforme.

Hormis les excellents candidats et ceux tentant d'aller grappiller les points des questions faciles de la partie 4, la plupart des bons candidats se sont arrêtés à cette question et ne sont pas allés plus loin dans le problème.

Quatrième partie

Cette quatrième et dernière partie a été très peu abordée.

Question 13a Cette question pouvait s'aborder par double inégalité en utilisant le cas d'égalité dans une inégalité de convexité qui apparaissait en écrivant $h(0)$ comme étant $h(-S_n)$ et en utilisant la formule de transfert.

Question 13b Dans cette question, il fallait proprement utiliser le caractère indépendant et identiquement distribué des variables aléatoires X_k. L'hypothèse faite en début de partie couplée au résultat de la question 11d permettait d'obtenir que $r(\Lambda_X) = 0$.

Question 13c Il s'agissait de vérifier que l'on pouvait appliquer la question 12b. Cette question a été abordée par des candidats qui souvent ont admis les résultats des questions précédentes.

Question 13d Cette question n'a que très peu été abordée. En coupant en deux la somme donnant la valeur $\mathbb{E}(h(x-S_n))$ à l'aide d'un réel à partir duquel $h(-x)$ est proche de $h(0)$ à ε près et en choisissant n suffisamment grand pour contrôler la première somme à l'aide de la question 3a, on montre que pour tout x réel fixé, $h(-x) = h(0)$.

Question 14a et 14b Ces questions très simples se résolvaient en utilisant le théorème de la limite monotone puis la caractérisation séquentielle de la borne supérieure.

Question 14c L'une des questions les plus difficiles du sujet. On pouvait se ramener au résultat de la question 13d en vérifiant que ξ vérifiait bien toutes les hypothèses de la question 13. L'un des points clés de la démonstration est d'utiliser que pour k assez grand, $t + t_k$ sort du support de g' et qu'ainsi ξ_k vérifie la même équation que h avec le terme en $g'(t + t_k)$ valant 0. Il est faux de croire qu'un simple passage à la limite lorsque k tend vers l'infini permet d'obtenir que $\xi = c$. En effet, le résultat de la question 14b ne permet pas de conclure à l'existence d'une limite pour $(f'(t + t_k))_{k\in\mathbb{N}}$ qui serait égale à c, ceci à cause de la translation par $-t$ fixé.

Question 14d Là encore, il s'agit d'une question difficile, nécessitant d'utiliser soit l'inégalité des accroissements finis soit le théorème fondamental du calcul intégral.

Question 14e et 14f Ces questions élémentaires ont été l'objet de tentatives de grappillage et ne valaient donc que très peu de points.

Très peu de candidats sont parvenus jusque là et ont eu le temps d'aborder les dernières questions du problème.

La fonction g_0 pouvait présenter une infinité de discontinuités sur un intervalle borné donc son intégrale n'avait pas de sens dans le cadre du programme qui se restreint à l'intégrale de Riemann des fonctions continues par morceaux. Le sujet admettait donc l'existence de l'intégrale. Pourtant, la fonction g_0 est bien entendue Riemann-intégrable et son intégrale relève donc de la même construction que celle enseignée en MPSI. Rappelons que l'hypothèse restrictive du programme de continuité par morceaux des fonctions que l'on intègre en Classes Préparatoires aux Grandes Écoles ne simplifie en rien la construction de l'intégrale et donc la notion de fonction intégrable, mais permet tout juste d'avoir un théorème simple d'existence des primitives. Ce dernier résultat est disjoint de la notion d'intégrabilité d'une fonction mais donne un outil permettant de calculer explicitement certaines intégrales.

D'autre part, dans ce cas précis, si on pense à l'intégrale comme à « l'aire sous la courbe » et que la courbe est constituée de segments horizontaux (éventuellement en nombre dénombrable), l'aire est une série à termes positifs (la somme des aires des rectangles) donc aura toujours un sens dans $[0, +\infty]$ et comme le graphe est contenu dans le rectangle $[0, K] \times [0, \|g\|_\infty]$ qui est d'aire finie, c'est bien un nombre réel ; admettre l'existence de l'intégrale, comme le demande sujet, est un pas certes, mais qu'il est légitime d'attendre des meilleurs candidats.

Enfin, nous rappelons qu'un candidat mal à l'aise avec les hypothèses du sujet peut toujours, s'il le souhaite, poser des hypothèses restrictives et admettre humblement qu'il ne sait pas traiter le cas général. Un candidat ayant une attitude humble devant une difficulté réelle et conscient des limites de ses connaissances sera presque automatiquement récompensé par les correcteurs.

Question 15a Nous concédons que l'unicité de μ était en défaut : sur $[0, +\infty[$, c'est une propriété alors que sur $]-\infty, 0[$, elle n'a pas lieu d'être. Cependant, comme la suite du sujet ne s'intéresse qu'aux valeurs de μ sur $[0, +\infty[$ puisqu'on intègre μ contre une fonction g à support compact $\mathbb{R}_+$, cette nuance était sans incidence. Un candidat qui traite précisément la question 15a devait vérifier les supports des fonctions et aurait alors clairement vu la subtilité. De même, en 15a, on pouvait démontrer l'égalité demandée avec un signe $+$ ou un signe $-$; le signe proposé par l'énoncé nécessitait une restriction de support qu'un candidat rigoureux devait dégager. De fait, les meilleurs candidats ont vu la difficulté et ont été récompensés. Ceux qui ont abordé sérieusement la question mais sans se préoccuper des supports n'ont cependant pas été pénalisés.

Question 15b Il s'agit d'utiliser la continuité de $s \mapsto \mu(s-a) - \mu(s)$ pour tout réel positif a fixé puis d'appliquer la question 15a pour une fonction test g bien choisie.

Question 16a La fonction caractéristique de $[0, +\infty[$ étant à support dans $\mathbb{R}_+$, il suffit de vérifier qu'elle satisfait à l'équation (E) dont la solution est unique.

Question 16b D'après la partie 2, il est possible de définir Lg_0 et l'hypothèse de continuité par morceaux ne jouant aucun rôle dans les questions précédentes, il est légitime d'appliquer la propriété $(\mathcal{P})$ avec g_0 puis d'utiliser 15b et 16a pour arriver au résultat.

Question 17 L'existence de la limite est une conséquence de la question 12b, sa valeur se déduit de 16b.

Question 18 En appliquant la question 17 à la fonction caractéristique de l'intervalle $[0, \ell]$ et en utilisant la question 4b, on obtient le résultat énoncé au début de ce rapport. Bien entendu ce résultat n'est plus vrai si la marche aléatoire $(S_n)_{n\in\mathbb{N}}$ est à valeur dans $d\mathbb{Z}$ dès que $\ell < d$.