

Banque BCPST Inter-ENS / ENPC - Session 2022

Rapport sur l'épreuve écrite de mathématiques

Épreuve comptant uniquement pour l'admission

Ecoles concernées : ENS (Paris) - ENS de Lyon - ENS de Paris-Saclay - ENPC

Coefficients (en pourcentage du total d'admission) :

ENS (Paris) : 11,3%

ENS de Lyon : 6,6%

ENS Paris-Saclay : 6,2%

ENPC : 20,0%

Membre du jury : Grégoire Nadin

Le sujet étudiait deux modèles d'épidémiologie, en deux parties distinctes. Le premier modèle était probabiliste, de type Galton-Watson, le second déterministe, de type SIR. Ce sujet abordait ainsi plusieurs parties du programmes des classes préparatoires BCPST : probabilités (formule des probabilités totales, indépendance), analyse réelle, polynômes, équations différentielles. Le choix de coller à l'actualité scientifique, la crise sanitaire ayant mis sur le devant de la scène la modélisation mathématique en épidémiologie, a également permis d'évaluer l'interprétation que faisait les candidats des modèles et des résultats obtenus.

Les notes obtenues par les candidats admissibles sont comprises entre 1 et 20. La moyenne est de 10,22, l'écart-type de 3,69. On peut estimer que la seule mobilisation des connaissances issues du lycée (et revues en classes préparatoires) permettait d'obtenir environ 7 points. En répondant aux questions relevant d'une application immédiate du cours de classes préparatoires, les candidats pouvaient atteindre 12 points. Pour aller au-delà il fallait savoir appliquer le cours dans des situations plus complexes.

Aux vues du coefficient de cette épreuve, on ne peut donc qu'encourager les candidats à ne pas négliger les mathématiques, d'autant que, si la plupart des candidats retenus sur liste complémentaire ont de bonnes notes, plusieurs ont des notes sous la moyenne. Enfin, on ne peut que rappeler ici l'importance que prennent les mathématiques aujourd'hui en biologie, que ce soit en termes de statistiques, d'exploitation de données, ou encore de modélisation. En conséquence, les candidats seront forcément amenés à les utiliser plus tard dans leur carrière professionnelle, autant donc profiter de ses études pour se doter de bases correctes.

Voici quelques commentaires aux réponses apportées par les candidats.

1.a) Cette question a été bien traitée par la quasi-totalité des candidats.

1.b) Quelques candidats n'ont pas réussi à calculer $\phi(0)$.

2.a) La plupart des candidats ayant abordé cette question ont tenté d'appliquer la formule des probabilités totales en indexant par l'évènement $\{Z_{n+1} = j\}$ plutôt que

$\{Z_n = j\}$, et beaucoup se sont ainsi perdus en route. Certains candidats se sont par exemple mélangés les pinceaux en traitant des doubles sommes avec le même indice.

2.b) Pour obtenir tous les points à cette question, il fallait à la fois exploiter l'indépendance des variables aléatoires $(X_n)_n$ et le fait qu'elles sont de même loi. La récurrence a été traitée par presque tous les candidats (mais rapportait peu de points).

2.c) De nombreux candidats ont fait l'erreur de sortir Z_n de l'espérance, ce qui est interdit puisqu'il s'agit d'une variable aléatoire. Le plus simple était de dériver la relation obtenue à la question précédente.

3.a) Question bien traitée par quasiment tous les candidats.

3.b) Cette question a été bien traitée par une poignée de candidats et était sans doute trop difficile. L'indication devait être utilisée pour calculer ξ mais au lieu de ça a été utilisée par bien des candidats pour démontrer l'inégalité, qui était en fait quasiment immédiate par inclusion.

3.c) Plusieurs candidats n'ont pas réussi à calculer $\phi_n(0)$, souvent parce qu'ils sortaient la variable aléatoire Z_n de la probabilité. Le fait que ξ est point fixe de ϕ se prouvait par un simple argument de continuité mais a été démontré par très peu de candidats.

3.d) Question bien traitée par les candidats qui l'abordaient.

3.e) Il suffisait de rassembler les questions précédentes, mais les candidats qui n'avaient pas réussi à calculer $\phi_n(0)$ ne pouvaient pas la traiter.

4.a) Cette question a en général été mal traitée. Une somme de termes positifs n'est strictement positive que si au moins un de ses termes est strictement positif. Il fallait donc évoquer l'hypothèse $p_0 + p_1 < 1$. Quand on ne savait pas démontrer ce résultat, il valait mieux se borner à expliquer pourquoi $\phi''(t) \geq 0$ pour grappiller quelques points, plutôt que d'affirmer sans le montrer que $\phi''(t) > 0$.

4.b) Il fallait appliquer un théorème des accroissements finis ou une variante. Un argument de type tangente pouvait fonctionner à condition d'expliquer le signe strict.

4.c) Question bien traitée par les candidats qui l'ont abordée.

4.d) Question difficile abordée seulement par une poignée de candidats.

5) Il suffisait d'appliquer le binôme de Newton. Certains candidats ont peut-être été perdus par une faute de frappe dans le sujet : il restait un C_n^k au lieu de $\binom{n}{k}$. Ceci dit, les candidats doivent savoir que $\binom{n}{k} = 0$ si $k > n$. Le calcul de ξ relevait d'un calcul de racines de polynôme d'ordre 2 et a malgré tout fait l'objet d'erreurs. Pour avoir tous les points, il fallait vérifier le signe des racines obtenues et sélectionner la plus petite positive, ce qui n'a été fait que par un petit nombre de candidats.

6.a) Cette question n'a été traitée que par une poignée de candidats. Il fallait utiliser le triangle de Pascal.

6.b) Très peu de candidats ont traité cette question. Il fallait appliquer la question précédente en changeant d'indice dans la somme. Quand les calculs sont compliqués, il est pertinent de vérifier le résultat obtenu sur certains cas, par exemple ici en vérifiant qu'on retrouve bien $\phi(1) = 1$.

6.c) Plusieurs candidats ont bien factorisé les polynômes, certains ont résolu l'équation d'ordre 2 qui en découlait, mais beaucoup ont abandonné en route.

7) Cette question a été traitée par presque tous les candidats. Plus globalement, le modèle SIR semble avoir été traité dans la plupart des classes préparatoires, ce dont

on ne peut que se féliciter. Cependant, le sujet proposait une certaine façon d'étudier celui-ci, basée sur le programme, et les candidats avaient parfois tendance à vouloir se raccrocher à l'approche qu'ils avaient dû voir en cours, plutôt que de s'adapter au sujet. Pour en revenir à la question 7, pour avoir tous les points il fallait par exemple expliquer en deux mots pourquoi $S'(t)$ est proportionnel à $S(t)$ ainsi qu'à $I(t)$, pas seulement expliquer pourquoi il y a un signe $-$ devant βSI etc.

8.a) De nombreuses erreurs à cette question pourtant basique. Les équations différentielles d'ordre 1 à coefficients non-constants sont pourtant bien au programme. L'erreur standard consistait à écrire $S(t) = S_0 \exp\left(-t \int_0^t \beta I(\tau) d\tau\right)$.

8.b) Souvent la même erreur qu'au a).

8.c) Les candidats ont souvent fait la confusion entre hypothèse de modélisation et preuve mathématique. Il fallait ici démontrer ce résultat (en sommant les trois lignes du système différentiel).

9.a) Il fallait utiliser la décroissance de S . Plusieurs candidats ont tenté de démontrer que I était croissante à partir d'un certain rang. Rappelons ici que ça n'est pas parce que la dérivée d'une fonction tend vers une limite positive que la fonction sera croissante à partir d'un certain rang (contre-exemple: $1/x$). Ça n'est vrai que si la limite est strictement positive. On attendait pour la contradiction une preuve rigoureuse rarement pourvue (par exemple $R'(t) \geq \gamma I_0$ donc $R(t) \geq \gamma I_0 t \rightarrow +\infty$ car $I_0 > 0$ ce qui est impossible car R est bornée).

9.b) Question peu traitée. Elle se démontre comme décrit à la question précédente.

9.c) Il fallait ici étudier les variations de S et I en fonction des valeurs de S et I . Beaucoup de candidats ont tracé un tableau de variations sans fournir de preuve, sans doute sur la base de leurs souvenirs de cours, souvent avec des erreurs.

10.a) La double implication (« si et seulement si ») a souvent été oubliée.

10.b) Question bizarrement très peu traitée. Les critères $m > 1$ et $\beta S_0/\gamma > 1$ sont pourtant très proches. La quantité m représente le nombre moyen de personnes contaminées par chaque malade. C'est aussi le cas de $\beta S_0/\gamma$, le fameux R_0 dont on a tant entendu parler dans les médias pendant l'épidémie.

11) Question facile, bien résolue quand elle était traitée.

12.a) Question très facile.

12.b) Cette question nécessitait des capacités d'abstraction puisqu'il fallait considérer S_∞ comme une fonction de I_0 . Elle a peu été traitée.

13.a) Souvent de petites erreurs de signe ou d'oublis de termes dans cette question quand elle était traitée.

13.b) Il s'agissait ici de confronter les résultats obtenus aux données de la population française. Peu de candidats ont traité cette question. Souvent les erreurs de la question précédente donnait des résultats extravagants (population négative, population d'infectés supérieure à la population totale...). Quel dommage que les candidats ne se soient pas saisis de ces aberrations pour s'interroger et corriger la formule obtenue au 13.a)!

14.a) Plusieurs candidats ont tenté de grappiller quelques points à cette question. L'interprétation était souvent pertinente mais aussi souvent extravagante.

14.b) Question ouverte : plusieurs candidats donnaient une condition juste sur la base d'une simple heuristique, quelques uns ont pris le temps de fournir quelques éléments de preuve.