

BCPST 2015 — Mathématiques

# MODÉLISATION MATHÉMATIQUE ET INFORMATIQUE

*Commentaires*

Le jury était cette année composé de Mesdames Odile Barka, Zoé Faget, Émilie Lebarbier, Céline Lévy-Leduc, Armèle Robin et de Messieurs Jean-Luc Fargier, Marc Jalard, Camille Male, Hachemi Nadir, Robert Thai, Mathieu Leroy-Lerêtre, Nicolas Guillaud (rapporteur) et Émeric Bouin (expert).

<b>1</b>	<b>Commentaires d'ordre général</b>	<b>1</b>
<b>2</b>	<b>Statistiques</b>	<b>1</b>
<b>3</b>	<b>Commentaires question par question</b>	<b>2</b>
3.1	Premier exercice . . . . .	2
3.2	Deuxième exercice . . . . .	4
3.3	Troisième exercice . . . . .	5

Le sujet comportait trois exercices indépendants ; le deuxième, et c'était la nouveauté, était un exercice d'informatique. Sur les trois heures et trente minutes que durait l'épreuve, les temps envisagés par les concepteurs sur chaque exercice étaient respectivement 1h30, 0h30 puis 1h30 : les points ont été répartis, dans le barème, proportionnellement à ces durées ; de sorte à ne pas avantager ou désavantager un candidat ayant abordé les exercices dans le désordre.

## 1 Commentaires d'ordre général

La première chose qui frappe lors de la correction est un écart de niveau important, entre les (nombreuses) très bonnes copies, avec des candidats maîtrisant les aspects techniques et calculatoires, ayant une bonne connaissance du cours, organisant correctement les démonstrations ; et des copies (là aussi, malheureusement, assez – trop – nombreuses) particulièrement faibles, de candidats qui manifestement ne connaissent pas les définitions, accumulent les erreurs de raisonnement dans les démonstrations, et passent sans même les amorcer les questions calculatoires, ce qui indique peut-être qu'ils ont peur de perdre du temps à faire la question se sachant très peu à l'aise, soit qu'ils ignorent le calcul à faire en lien avec la question posée.

Avec le recul, il apparaît assez clairement que bon nombre de candidats ont connaissance (soulignons que c'est un point tout à fait *positif*) des rapports des années antérieures, et ont été poussés, au cours de leur formation, à prêter la plus grande attention aux points justement soulignés ; tandis que d'autres ignorent les attentes du jury. Et c'est là, même s'il ne concerne qu'un nombre limité de candidats, un point *décevant* : un certain nombre de remarques, déjà formulées dans les rapports de l'an passé, sont à nouveau à faire ; elles sont sources d'une perte de points importante pour les candidats qui n'en tiennent pas compte. Avant d'y revenir en détail, question par question, citons déjà l'absence de rédaction des preuves par récurrence (« une récurrence facile montre que » : lorsqu'une question se résume à cela, ces mots, suivis d'une assertion donnée dans le sujet, ne montrent aucune production de la part du candidat, et ne sauraient lui rapporter des points !), les erreurs dans le calcul avec les fractions numériques (d'autant plus impardonnables que les calculatrices étaient ici autorisées. . .) et enfin, la phrase « prenons par exemple », qui ne constitue aucunement une *démonstration*.

Les copies sont bien présentées ; on encourage les très rares candidats ayant une écriture peu lisible à aérer leur rédaction, notamment les calculs, l'écriture des matrices, etc., afin de lever toute ambiguïté sur la lecture d'un chiffre, d'une lettre ou sur la position d'un coefficient. Dans quelques copies, la numérotation des questions n'est par moments pas respectée ou est erronée ; nous pressons le candidat à apporter le plus grand soin dans cette numérotation. Rappelons en outre que la couleur rouge (et les nuances proches) sont réservées à la correction, et qu'il convient de ne pas l'utiliser. Enfin un candidat a écrit les commentaires de ses programmes, dans le deuxième exercice, au crayon à papier : mieux vaudrait user d'un stylo à l'avenir.

L'orthographe et la grammaire laissent à désirer sur de nombreuses copies. Il ne s'agit évidemment pas d'une dictée, mais une copie dans laquelle la moitié des verbes et adjectifs sont mal accordés fait en général mauvais effet. Plus sérieusement, des erreurs de syntaxe, parfois rencontrées, peuvent changer totalement le sens d'une phrase. En particulier dans les démonstrations, le mot *si* est utilisé abusivement ; il n'a pas le même sens que le mot *puisque*.

Dans la suite, nous allons essentiellement nous concentrer sur les points mal traités par les candidats ; il ne faut en aucun cas oublier que beaucoup de choses sont *bien* faites dans les copies. Le but de ces pages, sans chercher à relever toutes les erreurs rencontrées, est de mettre l'accent sur les points *les plus facilement* améliorables (et ils sont nombreux), notamment sur les aspects méthodologiques ; ces remarques se veulent constructives et espèrent voir une progression, qualitative et quantitative, des copies les plus faibles (quantitative voulant dire que ces dernières soient davantage remplies, pas plus nombreuses). Cela suppose évidemment que les futurs candidats aient eu connaissance de ces remarques.

## 2 Statistiques

### 3 Commentaires question par question

#### 3.1 Premier exercice

Le premier exercice regroupait un certain nombre de questions classiques d’algèbre linéaire, souvent calculatoires et plusieurs très proches du cours, ainsi que quelques questions plus difficiles.

**Question 1.** Aucune justification n’était attendue dans cette question, les réponses

$$n_{s,1} + n_{s,2} + \dots + n_{s,p}$$

et

$$\sum_{i=1}^p n_{s,i}$$

sont justes l’une comme l’autre ; mais dans la deuxième les bornes ont été parfois mal écrites :  $i = 0$  jusqu’à  $p$  est une réponse non seulement fautive, mais qui n’a en plus aucun sens puisque la quantité  $n_{s,0}$  n’est pas définie dans le sujet.

**Question 2.1.** Question globalement bien traitée, et plus souvent juste que la précédente. De manière étonnante, des candidats qui se trompent sur les indices dans la question précédente les ont écrit justes dans celle-ci. Quelques erreurs de raisonnement ont été rencontrées, et menaient à des formules faisant intervenir les quantités  $P_i$ . Ceci aurait dû alerter les candidats, car le sujet demandait une formule en fonction des  $n_{s,i}$  et des  $F_i$  et ne mentionnait aucune autre grandeur.

**Question 2.2.** Voici une question ouverte, dont les réponses l’ont été encore davantage. De très nombreux candidats ont purement ignoré la question, d’autres ont essayé d’écrire le nom d’un théorème au hasard, sans la moindre explication mais avec beaucoup d’imagination. On n’écrit pas « le théorème de la loi faible des grands nombres ». L’explication attendue, idéalement, et effectivement rencontrée une fois, était la suivante : pour chacun des  $n_{s,i-1}$  individus de la classe  $C_i$  à l’instant  $s$ , on effectue une expérience de Bernoulli donnant le résultat 1 s’il survit (avec la probabilité  $P_{i-1}$ ) et 0 sinon. Si  $X_k$  est le résultat de l’expérience pour le  $k$ -ième individu, alors on a

$$n_{s+1,i} = X_1 + \dots + X_{n_{s,i-1}}.$$

Lorsque les expériences sont *indépendantes*, la Loi Faible des Grands Nombres permet d’affirmer qu’au moins pour les grandes valeurs de  $n_{s,i-1}$  la somme ci-dessus est peu différente de  $n_{s,i-1} \times \mathbf{E}(X_1) = n_{s,i-1} \times P_{i-1}$ . Bien entendu, des explications bien plus succinctes étaient acceptées, mais le point crucial, à savoir l’indépendance, n’a été que très rarement rencontré.

**Question 2.3.** Cette question a vu beaucoup de mauvaises réponses, certaines très surprenantes comme

$$M = \begin{pmatrix} F_1 & F_2 & \dots & \dots & F_p \\ P_1 & 0 & \dots & \dots & 0 \\ 0 & P_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & P_p \end{pmatrix} \quad \text{ou} \quad M = \begin{pmatrix} F_1 & F_2 & \dots & \dots & F_p \\ P_1 & 0 & \dots & \dots & 0 \\ 0 & P_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & P_p & 0 \end{pmatrix}.$$

On ne saurait inviter plus énergiquement les candidats à considérer attentivement ce qu’ils écrivent : une matrice carrée est attendue par l’énoncé, et la réponse de gauche ci-dessus ne l’est pas ; quant à la réponse de droite, rencontrée *plusieurs* fois, elle est tout simplement de forme incohérente.

Certains candidats ont essayé de deviner le résultat à partir de la question suivante, c’est un bon réflexe, mais la (mauvaise) réponse

$$M = \begin{pmatrix} F_1 & F_2 & \dots & \dots & F_p \\ P_1 & 0 & \dots & \dots & 0 \\ 0 & P_2 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & P_{p-1} & 0 \end{pmatrix}$$

aurait dû, une nouvelle fois, alerter ceux qui l'ont proposée : elle ne fait pas intervenir  $P_p$  qui pourtant apparaît dans les formules données par l'énoncé.

**Question 3.1.** Question globalement bien traitée. Quelques candidats se sont lancés dans une démonstration par récurrence, alors qu'il suffisait d'appliquer le résultat de la question précédente.

**Question 3.2.** C'est ici qu'il fallait *rédigé* une récurrence. Répétons ce qui a été dit en préambule : on peut dire qu'un résultat est la conséquence d'une récurrence triviale, si ce résultat est donné par le candidat lui-même et s'il n'est qu'une portion de raisonnement dans une question plus complexe. Ici la démonstration constitue la totalité de la question, le résultat étant fourni par l'énoncé : un candidat qui écrit « c'est évident » ou une réponse analogue n'apporte donc rien et n'obtient logiquement *aucun* point.

En revanche, soulignons que les nombreux candidats qui ont rédigé la récurrence l'ont très bien fait. Quelques-uns (rares) démarrent (est-ce une précaution ?) à  $s = 1$ , alors que le cas  $s = 0$  ne pose pas le moindre problème ; d'ailleurs l'énoncé demande d'établir la formule pour tout entier naturel, et zéro (au moins en France) en est un.

**Question 3.3.1.** Quatre choses sont à dire sur cette (longue) question. La première est qu'elle est en fait constituée de plusieurs questions : manifestement certains candidats ont oublié (n'ont pas lu ?) la dernière : « la matrice  $A$  est-elle diagonalisable ? ».

La deuxième chose concerne l'utilisation de la calculatrice : celle-ci donne (même une calculatrice non formelle) les valeurs propres (au moins approchées, mais celle qui manquait était  $-1,5$ ) d'une matrice. On peut donc tout à fait écrire « La calculatrice donne  $-1,5$  comme troisième valeur propre, vérifions que c'est le cas » et procéder comme pour les deux qui sont données par l'énoncé. C'est peut-être plus rapide et plus simple que de faire l'échelonnement avec un paramètre  $\lambda$  inconnu. Le fait est que certains candidats ont obtenu, après échelonnement, un polynôme  $P(\lambda)$ , de degré trois, et n'ont pas su, à partir de ses deux premières racines, trouver la troisième (appelons-là ici  $\lambda'$ ). Il fallait faire une identification, en partant de l'expression factorisée

$$(\lambda - 2)(\lambda + 0,5)(\lambda - \lambda'),$$

pour la trouver.

La troisième chose concerne la définition des valeurs propres et les conditions caractéristiques de la diagonalisabilité : on en trouve une grande variété, par exemple (tout ce qui suit est faux) : «  $\lambda$  est valeur propre de  $A$  si et seulement si  $A - \lambda I$  est inversible », «  $A$  est inversible donc diagonalisable », «  $A$  est diagonalisable car elle possède trois valeurs propres et est de rang trois ». Dans cette dernière phrase, le mot *rang* est parfois remplacé par *dimension*, sans qu'il soit toujours clair que ce mot ait signifié, dans la tête du candidat au moment où il l'a écrit, « rang » ou bien « taille ». Lorsque le bon mot est utilisé, il faut encore accoler *distinctes* à *valeurs propres* pour avoir les points.

La quatrième et dernière chose est *à peu près* très positive : la question a été bien traitée, et par une grande majorité ; même si quelques erreurs de calcul ont parfois été commises. Le *à peu près* est dû à certains candidats qui ont tout simplement passé la question sans essayer même de l'amorcer : est-ce par un manque de confiance envers les questions calculatoires ? Ou bien est-ce par une ignorance pure et simple de la méthode à employer ? Une telle question devrait pourtant mettre les élèves les plus mal à l'aise en confiance : elle ne demande pratiquement aucune réflexion ; c'est pour ainsi dire une question de cours.

**Question 3.3.2.** Dans cette question, on peut utiliser les calculs faits dans la précédente : d'ailleurs certains candidats ont traité les deux questions d'un seul coup.

Lorsqu'une famille est constituée d'un unique vecteur, on attend du candidat qu'il précise (au moins une fois) que celui-ci n'est pas nul, avant d'affirmer que c'est une base. En effet,  $\vec{0}$  est un vecteur du sous-espace propre de  $A$  associé à la valeur propre 2, mais n'en est pas pour autant une base.

Beaucoup moins anecdotique : la démonstration du fait que  $(V_1 ; V_2 ; V_3)$  est une base a vu beaucoup d'arguments erronés : une concaténation de familles libres (même distinctes) n'en donne en général pas une ! Pour démontrer qu'une famille est une base, on montre en effet qu'elle est libre et on conclut en disant qu'en dimension *finie*  $d$ , une famille d'exactly  $d$  vecteurs est une base si et seulement si elle est libre (ou bien si et seulement si elle est génératrice mais ici c'était de peu d'utilité). Dans quelques copies, un joyeux mélange de ces mots aboutit à des phrases dépourvues de sens : « : la famille est libre et de plus génératrice car il s'agit d'un endomorphisme fini de  $\mathcal{M}_{31}(\mathbf{R})$  vers  $\mathcal{M}_{31}(\mathbf{R})$  ».

**Question 3.4.1.** Dans cette question, on peut démontrer la formule directement par récurrence, ou plus simplement démontrer par récurrence que si  $Ax = \lambda x$ , alors  $A^s x = \lambda^s x$ . Quoi qu'il arrive, il y a à nouveau une récurrence à rédiger. Les démonstrations à base de points de suspension comme

$$A^s x = A^{s-1}(Ax) = A^{s-1}(\lambda x) = \lambda A^{s-1}x = \dots = \lambda^s x$$

sont à éviter, même si leur présence a été valorisée par rapport aux copies qui passent le point sous silence.

**Question 3.4.2.** Cette question n'est pas très compliquée mais demande d'être rédigée avec précision (et elle l'a souvent été, cela dit). On attend du candidat qu'il définisse un vecteur  $\varepsilon_s$  satisfaisant la relation de l'énoncé, et qu'il justifie convenablement que ses coefficients tendent vers 0.

Concernant ce dernier point, un argument erroné est apparu à de bien trop nombreuses reprises : « si une fraction  $p/q$  est telle que  $|q| > 1$ , alors  $(p/q)^s \rightarrow 0$  lorsque  $s$  tend vers l'infini » (répétons que cette proposition est fautive!). Ce n'est donc pas  $q = 2$  qu'il fallait considérer ici, mais bien les rapports de valeurs propres  $\lambda_2/\lambda_1$  et  $\lambda_3/\lambda_1$ , qui sont dans l'intervalle  $] -1; 1[$  assurant ainsi la convergence de la suite géométrique. Attention aussi à ne pas oublier les valeurs absolues :  $r < 1$  n'implique pas  $r^s \rightarrow 0$ .

**Question 3.4.3.** Cette dernière question a été peu abordée, mais la plupart du temps parce que le résultat de la question précédente manquait. On pouvait pourtant faire sans : si l'on note

$$V_1 = \begin{pmatrix} v_{11} \\ v_{21} \\ v_{31} \end{pmatrix} \quad \text{et} \quad \varepsilon_s = \begin{pmatrix} \varepsilon_{1s} \\ \varepsilon_{2s} \\ \varepsilon_{3s} \end{pmatrix},$$

alors

$$\frac{n_{s,1}}{n_{s,3}} = \frac{\lambda_1^s (c_1 v_{11} + \varepsilon_{1s})}{\lambda_1^s (c_1 v_{31} + \varepsilon_{3s})} = \frac{v_{11}}{v_{31}} \times \frac{c_1 + \varepsilon_{1s}/v_{11}}{c_1 + \varepsilon_{3s}/v_{31}} \xrightarrow{s \rightarrow \infty} \frac{v_{11}}{v_{31}}.$$

Ceci est possible dès lors que  $v_{11}$  et  $v_{31}$  sont non nuls ; on n'a donc (hormis pour cette dernière justification) pas vraiment besoin de ce qui précède (sauf des coefficients de  $V_1$ ). Le même calcul est à refaire pour le second rapport.

## 3.2 Deuxième exercice

Attardons-nous sur cet exercice, puisque c'est la nouveauté. Pour cette année de transition, et de manière très exceptionnelle, on a laissé les candidats choisir le langage dans lequel ils rédigeraient leurs programmes. Seul Python sera utilisé à l'avenir.

Concernant le choix du langage, il était clairement écrit dans le préambule d'indiquer en début de copie celui qui était retenu : cela a été fait sur la moitié des copies environ, et soulève une question : les candidats lisent-ils attentivement l'énoncé ? Le doute (déjà mentionné dans l'exercice précédent, et qui le sera encore dans l'exercice suivant) finit par ne plus en être un...

On constate, dans un nombre significatif de copies, que lorsque cet exercice d'informatique a été très bien traité, le reste est d'un niveau assez faible ; alors qu'inversement plusieurs très bonnes copies du point de vue mathématique ont purement et simplement fait l'impasse sur ce deuxième exercice. Ce constat est bien regrettable, car il montre que les candidats ne sont pas du tout sensibles à la grande articulation qu'il y a entre les mathématiques (ou plus généralement les sciences), et l'informatique appliquée à la résolution des problèmes qui se posent dans ces disciplines.

Des candidats semblent manquer cruellement de pratique de la programmation : cela se manifeste par des programmes formulés dans un langage plus proche des mathématiques que d'un langage de programmation : traits de fractions à l'horizontale, exposants notés comme tels, affectations du type  $\|v\| = \dots$ , etc.

Enfin, en annexe était donnée une liste de fonctions que le candidat pouvait librement utiliser. Concernant Python, la classe `matrix` (dont l'utilisation est la plus proche de la formulation mathématique) était clairement indiquée, mais certains candidats ont choisi d'utiliser une autre représentation : la classe `array` (dans de rares cas pratique, mais également moins intuitive, par exemple pour écrire le produit de deux matrices), ou bien (pire encore) les listes de listes. Pour ces dernières, le produit matriciel n'est pas défini du tout, et de très rares candidats sont allés jusqu'à ré-écrire la fonction l'effectuant. C'est doublement problématique : d'un point de vue pratique, la ré-écriture de cette fonction était plus compliquée que ce qui

était à proprement parler demandé par l'exercice ; et d'un point de vue conceptuel, on est en droit de se demander si l'étudiant qui reprogramme le produit matriciel avant d'écrire ses codes a saisi l'utilité de l'outil informatique pour faire des mathématiques. . .

Pour les épreuves à venir, insistons sur les bibliothèques : les élèves n'ont pas de connaissance à avoir sur leur utilisation ; mais lorsqu'on leur donne une documentation, avec une liste de fonctions et leurs descriptions, on attend d'eux qu'ils les utilisent. En aucun cas le but de l'épreuve ne saurait être de tout reprogrammer soi-même ; et par ailleurs, le candidat doit être capable de se plier aux conventions choisies par l'énoncé, comme il le ferait dans un problème de mathématiques : si une fonction s'appelle  $f$  dans le sujet, le candidat n'a pas à la nommer  $g$  dans sa copie.

**Question 1.** Cette question est presque une recherche de maximum, il faut parcourir tous les coefficients de la matrice et ceci se fait en imbriquant deux boucles `for`. Peut-être par analogie avec une recherche de maximum, dans un grand nombre de copies, les valeurs absolues ont été oubliées.

**Question 2.** Dans cette question on utilise la fonction de la question précédente, on ne la ré-écrit pas ; de la même manière qu'on a le droit d'utiliser, dans un problème de mathématiques, le résultat d'une question pour traiter la suivante.

**Question 3.** La notion de norme, en toute généralité, n'est pas au programme de BCPST. Beaucoup de candidats, dans cette question, ont (pour une raison qu'il n'est pas facile de deviner) normalisé la quantité seulement tout à la fin : ils ont pour ainsi dire calculé  $w = A^n v_0$  puis ont renvoyé  $w/\|w\|$ . Mathématiquement, c'est juste mais cela utilise l'homogénéité de la norme, et demande donc une justification. Numériquement, c'est catastrophique : calculer les puissances d'une matrice sans normaliser à chaque étape mène à des coefficients dont l'ordre de grandeur peut-être gigantesque. Cela peut résulter en des erreurs de calculs, ou éventuellement en un *dépassement de capacité*.

**Question 4.** Voilà la question la moins réussie de cet exercice. Toutes les réponses ont été données avec égale fréquence ; dans beaucoup de copies la question n'a même pas été abordée (alors que les trois autres l'ont été). Les candidats doivent être entraînés à ce type d'exercice, qu'on pourrait retrouver dans les sessions à venir.

### 3.3 Troisième exercice

Ce dernier exercice était un exercice d'analyse plus que de probabilités, beaucoup de questions demandant de calculer des intégrales, parfois sur des intervalles non compacts. Les définitions du cours de probabilités étaient quand même nécessaires (et ont souvent fait défaut), et ce dès la première question.

**Question 1.** Il y a trois axiomes à vérifier pour démontrer que  $\ell_{\alpha;r}$  est une densité de probabilité : sa continuité par morceaux sur  $\mathbf{R}$ , le fait qu'elle est positive, et enfin que son intégrale sur  $] -\infty ; +\infty[$  converge et vaut 1. Les trois points ont été vérifiés de manière totalement aléatoire par les candidats, ce qui montre qu'ils connaissent mal la définition.

Concernant chacun des points : le plus délicat à vérifier est le calcul de l'intégrale, qui demande de fixer d'abord des bornes finies  $A$  et  $B$ , puis de faire tendre celle de gauche vers 0 et celle de droite vers l'infini. Beaucoup de candidats n'ont pas vu que pour certaines valeurs de  $\alpha$ , la fonction n'était pas continue sur  $\mathbf{R}$ , avec une limite infinie en  $0^+$ .

Enfin, le *si et seulement si*, indiquant une équivalence, a posé de nombreux problèmes de rédaction, les candidats semblant pour bon nombre d'entre eux mal à l'aise avec la logique élémentaire, ou en tous cas ayant du mal à retranscrire convenablement les arguments à l'écrit.

**Question 2.1.** Pour la fonction de répartition, on a à nouveau besoin de calculer une intégrale sur un intervalle non compact. Certains candidats dérivent la densité (qui rappelons-le n'est pas continue lorsque  $\alpha \in ]0 ; 1[$ ) pour obtenir le résultat, ce qui évidemment ne donne rien de bon ; mais ce qui est remarquablement incohérent c'est que dans une question ultérieure, les mêmes vont à nouveau dériver, cette fois-ci la fonction de répartition, pour obtenir la densité.

Dans quelques copies, on trouve

$$\int_{-\infty}^t \alpha \frac{s^{\alpha-1}}{r^\alpha} ds$$

ce qui n'a évidemment aucun sens puisque  $\ell_{\alpha; r}(s)$  n'a l'expression utilisée ci-dessus comme intégrande que lorsque  $s \in ]0; r[$ ; et même lorsque le calcul est bien rédigé, la réponse est peu rigoureuse : l'expression de  $F_X$  s'obtient en distinguant trois cas, on ne peut pas conclure (et encadrer)  $F_X(t) = (t/r)^\alpha$  sans rien préciser d'autre : cette expression ne saurait quoi qu'il arrive être valable sur  $\mathbf{R}$  tout entier, vues les propriétés d'une fonction de répartition.

**Question 2.2.** Que vaut  $\mathbf{P}(X \leq 0)$  ? Eh bien elle vaut 0, ce qu'on obtient avec la formule trouvée pour  $F_X$  ou bien en remarquant que le support de  $X$  est contenu dans  $]0; +\infty[$  d'après la fonction de densité.

**Question 2.3.** Dans cette question, la conclusion est souvent juste, mais on remarque le même manque de rigueur dans les écritures intermédiaires : la densité de  $W$  n'est pas

$$f_W(t) = 1 - e^{-\alpha t},$$

mais bien

$$f_W(t) = \begin{cases} 1 - e^{-\alpha t} & \text{si } t \geq 0 \\ 0 & \text{sinon.} \end{cases}$$

Toujours concernant la rédaction, on ne peut passer directement de la fonction de répartition à la densité : ayant obtenu  $F_W$ , on constate que celle-ci est continûment dérivable par morceaux sur  $\mathbf{R}$ . On en déduit *alors* que  $W$  est une variable à densité, et que la densité s'obtient en dérivant (chaque morceau de)  $F_W$ .

**Question 3.1.** Il faut ici faire un calcul et ne surtout pas appliquer un faux argument de transformation affine, menant dans quelques copies à la densité négative (sic.)  $-f_{U_2}$ . Encore une fois, répétons aux candidats de prendre garde au sens de ce qu'ils écrivent, un signe  $-$  devant une densité devrait faire trembler leur poignet au moment où la pointe de leur stylo s'apprête à le tracer.

**Question 3.2.** Voici la première d'un doublet de questions calculatoires. Ceux qui l'ont abordée sérieusement l'ont en général bien menée, et obtiennent le bon résultat. Mais bien trop souvent la rédaction s'arrête après une ré-écriture de la formule rappelée dans l'énoncé, certains candidats n'allant même pas jusqu'à substituer les expressions de  $f_{U_1}$  et  $f_{-U_2}$ . À titre de comparaison, les calculs de ce problème d'analyse sont globalement bien moins maîtrisés que les questions calculatoires d'algèbre linéaire du premier exercice.

**Question 3.3.** Le point clé de cette question était d'écrire

$$-\ln\left(\frac{sX}{rY}\right) = \ln\left(\frac{Y}{s}\right) - \ln\left(\frac{X}{r}\right),$$

ce qui demande de connaître les propriétés algébriques du logarithme. Celles-ci, pourtant vues en classe de Terminale, ne sont toujours pas maîtrisées par un nombre *non négligeable* de candidats.

**Question 4.1.** La fin du sujet dissimulait deux questions faciles, que tous les candidats n'ont malheureusement pas pensé à aller voir. Dans cette première sous-question, il fallait faire le calcul

$$\mathbf{P}_M(\mathbf{R}) = 1 - \frac{4}{4+2} \left(\frac{400}{800}\right)^2 = \frac{5}{6}$$

(car lorsque  $M$  est réalisé, c'est-à-dire lorsque le sujet est malade, on a  $\alpha = 4$ ,  $\beta = 2$ ,  $r = 400$  et  $s = 800$ , ce qui en particulier indique qu'on est dans le cas  $r \leq s$  pour appliquer la formule de la question précédente), et

$$\mathbf{P}_{\overline{M}}(\mathbf{R}) = \frac{3}{2+3} \left(\frac{50}{100}\right)^2 = \frac{3}{20}$$

dans le cas où  $M$  n'est pas réalisé (l'individu n'est pas malade).

Arrêtons-nous sur deux remarques : la première c'est que la calculatrice était *autorisée* et qu'il est donc absolument impardonnable de faire une erreur de calcul sur ces fractions (et il y en a eu). La seconde, c'est

que sur le nombre malgré tout important de candidats ayant abordé la question, beaucoup n'ont pas réussi à substituer les valeurs des quatre paramètres ; certains n'ont même pas su comment remplacer la probabilité conditionnelle  $\mathbf{P}_M(\mathbf{R})$ . Il est indéniable que ce qui a posé problème à certains, c'est la compréhension de l'énoncé, et le lien entre les phrases et les calculs à faire. Il paraît *essentiel* de travailler sur ce point : faire le lien entre une situation concrète et une mise en équation est à la base de l'activité scientifique.

**Question 4.2.** On doit ici utiliser deux formules. D'abord, la définition des probabilités conditionnelles :

$$\mathbf{P}_R(M) = \frac{\mathbf{P}(M \cap R)}{\mathbf{P}(R)} = \frac{\mathbf{P}(M)}{\mathbf{P}(R)} \times \frac{\mathbf{P}(M \cap R)}{\mathbf{P}(M)} = \frac{\mathbf{P}(M)}{\mathbf{P}(R)} \times \mathbf{P}_M(R),$$

et la formule des probabilités totales pour calculer

$$\mathbf{P}(R) = \mathbf{P}(R \cap M) + \mathbf{P}(R \cap \bar{M}) = \mathbf{P}_M(R) \times \mathbf{P}(M) + \mathbf{P}_{\bar{M}}(R) \times \mathbf{P}(\bar{M}) = \frac{5}{6} \times \frac{1}{40} + \frac{3}{20} \times \frac{39}{40} = \frac{401}{2400}.$$

On en déduit que

$$\mathbf{P}_R(M) = \frac{\frac{1}{40}}{\frac{401}{2400}} \times \frac{5}{6} \simeq 0,12.$$

On peut aussi appliquer directement l'une des écritures du théorème de Bayes, à savoir

$$\mathbf{P}_R(M) = \frac{\mathbf{P}(M) \times \mathbf{P}_M(R)}{\mathbf{P}(M) \times \mathbf{P}_M(R) + \mathbf{P}(\bar{M}) \times \mathbf{P}_{\bar{M}}(R)},$$

à condition de ne pas écrire n'importe quoi au dénominateur (on a trouvé, parfois, le dénominateur  $\mathbf{P}_M(R) + \mathbf{P}_{\bar{M}}(R)$ ). Certains candidats ont opté pour cette alternative, et sont parvenus au bon résultat, alors qu'ils n'ont pas su faire la question précédente : et pourtant, ce faisant, ils ont donc trouvé (à leur insu) les réponses de celle-ci.

On a fait le calcul ci-dessus pour commenter la valeur trouvée (c'était la dernière question du sujet, là aussi quelques candidats n'ont pas vu qu'il y avait deux questions en une) : le test détecte peut-être cinq malades sur six (c'est le résultat de la question précédente), mais lorsqu'il est positif, il s'agit dans 88 % des cas d'un individu qui n'est en fait pas malade : ce test est inutilisable. Un nombre finalement assez restreint de candidats ont osé tirer cette conclusion ; parmi les autres, certains ont été jusqu'à écrire qu'ils « avaient dû faire une erreur quelque part ». Il faut avoir confiance en ses résultats, sans compter qu'ici, le calcul est suffisamment court pour pouvoir le relire et trouver si une erreur s'y est glissée.