Mathématiques II – Option économique

1. Le sujet

L'objet du problème était l'étude des propriétés d'un estimateur (celui obtenu par la méthode du maximum de vraisemblance) du paramètre p d'une loi géométrique, ainsi que la détermination d'un intervalle de confiance de p.

Le sujet faisait appel à de larges parties du programme de mathématiques générales et de probabilités (séries numériques, intégrales, équivalents, existence et calculs d'espérance, etc.), même si sa problématique était de nature statistique. Ainsi, s'agissait-il essentiellement de déterminer le degré de distorsion (biais) et la convergence de l'estimateur étudié par des méthodes d'analyse du programme et retrouver ainsi certaines propriétés (bien connues des statisticiens) de la méthode du maximum de vraisemblance.

2. Les résultats obtenus

Dans l'ensemble des 1910 candidats de l'option économique, la moyenne obtenue est de $8,15$ avec un écart-type très élevé qui atteint 4. Il est vraisemblable que cette épreuve a permis d'effectuer un bon classement des candidats et a pu distinguer les meilleurs d'entre eux. Ainsi, 25% des candidats obtiennent une note supérieure à 14 et 25% une note inférieure à 4 ; il y a environ 5% de copies dont la note est très voisine de 0.

Dans le barème de notation, les poids respectifs des six parties du problème ont été les suivantes : partie I : 14%, partie II : 6%, partie III : 29%, partie IV : 20%, partie V : 21%, partie VI : 10%.
Compte tenu de la relative « longueur » du problème, la note de 20 était accordée à tout candidat qui résolvait correctement au moins la moitié du sujet.

3. Commentaires détaillés

En général, la présentation des copies est très convenable, les écritures sont lisibles et les résultats sont mis en évidence.

Partie I

1. Une proportion importante de candidats tente une récurrence sur r, ce qui les conduit soit à une impasse, soit à écrire des formules erronées. Les récurrences correctement rédigées, sur l'entier n, ne mentionnent quasiment jamais la possibilité pour r d'être égal à $n+1$.

2. a) Beaucoup de candidats ne parviennent pas à établir l'égalité demandée et les changements d'indices induisent de grosses erreurs de calcul dans de nombreuses copies.

2. b) Cette question est bien traitée lorsqu'elle est abordée. Cependant, on affirme souvent que $n! \sim n^n$.

3. a) Une majorité de candidats traite cette question mais la justification de l'existence des limites proposées est rarissime : on écrit la somme infinie et le résultat, puis on invoque le cours.

3. b) Le raisonnement par récurrence, lorsqu'il est abordé, est très mal conduit, ce qui constitue une nouveauté par rapport aux années précédentes. Ainsi, dans un grand nombre de copies, on suppose le résultat vrai au rang r et on remplace r par $r+1$ pour obtenir le résultat au rang $r+1$.

Partie II

1. Cette question est correctement résolue par les candidats qui l'ont abordée.

2. Dans la majorité des copies, on trouve $0 < \dfrac{1}{1-t} < 1$, alors que $0 < t < 1$.

3. Les réponses sont parfois embrouillées. Quelques candidats évoquent la convergence de l'intégrale pour en déduire celle de la série, car ils confondent la convergence de la suite $\left(\int_0^x \frac{t^n}{1-t}dt\right)$ avec une convergence d'intégrale.

Partie III

1. Beaucoup de candidats ont confondu « calculer » et « donner ». Il fallait ici utiliser les résultats de la partie I.

2. a) Cette question est rarement bien traitée. On trouve ainsi $Y(\Omega) = \mathbb{N}^*$ ou $1/\mathbb{N}^*$ ou $[0, 1]$, ainsi que $P(Y = x) = \frac{1}{P(X = x)}$.

2. b) Les candidats qui ont vu le lien avec la partie II ont traité correctement cette question. En revanche, on trouve souvent $E(Y) = \frac{1}{E(X)}$.

2. c) Cette question a très rarement été abordée.

3. On trouve beaucoup d'erreurs dans les trois sous-parties de cette question. Ainsi, « S_2 suit une loi géométrique de paramètre $2p$ » ou bien encore des notions de linéarité fantaisistes, telles que « $P(X_1 + X_2 = k) = P(X_1 = k) + P(X_2 = k)$ ».

4. a) Cette question est bien traitée en général même si on oublie parfois d'invoquer l'indépendance pour calculer $V(S_n)$.

4. b) La loi de S_n est très peu souvent abordée et rarement bien traitée.

5. On trouve les mêmes erreurs que pour Y (question 2. a) et l'existence des moments de Y_n n'est quasiment jamais abordée.

Partie IV

1. La notion d'estimateur est rarement perçue ; en revanche, la propriété de « sans biais » est bien connue.

2. a) Cette question est peu abordée. On trouve des tentatives d'inversion de sommes et intégrales. En général, seule la première partie de la question est traitée.

2. b) Le changement de variable n'est jamais justifié mais l'intégration par parties est bien faite.

3. Lorsque la sous-question a) est abordée, elle est correctement traitée. Les autres sous-questions donnent lieu à beaucoup d'erreurs. Ainsi, de nombreux candidats affirment que $p/q < 1$; pour obtenir la négligeabilité de certains termes, les questions d'inégalités dans les intégrales sont mal maîtrisées même dans les bonnes copies.

Partie V

L'équivalent de la question 1. a) est rarement obtenu. Les questions 2 et 3 sont peu abordées.

Partie VI

Cette partie a été assez souvent abordée mais excepté quelques candidats, on se contente de « grappiller » des points ça et là (voir la question 2. a). Parmi les erreurs importantes, on peut noter des différences d'équivalents et même $o(1/n) - o(1/n) = 0$. On pense qu'il faut choisir p et q ($p = \frac{2}{3\sqrt{3}}$ et $q = 1$) dont la somme ne vaut pas 1.