MATHEMATIQUES 2S (épreuve n°283)

ANNEE 2012

Epreuve conçue par CCIP

Voie Scientifique

	NBRE CANDIDATS	MOYENNES	ECARTS-TYPE
RESULTATS GLOBAUX	2 946	10,37	4,44

VOIES PREPARATOIRES			
Scientifique	2 946	10,37	4,44

ECOLES UTILISATRICES			
HEC Paris	2 231	11,44	4,20
ESSEC	2 261	11,43	4,20
ESCP-EUROPE	2 470	11,11	4,25
EMLYON Business School	2 871	10,43	4,43

Le sujet

Le problème de cette année avait pour objet l'étude de quelques propriétés du modèle de régression linéaire élémentaire, notamment l'estimation des paramètres inconnus du modèle.

Ce sujet faisait appel à de larges connaissances du programme d'analyse (optimisation sous contraintes linéaires), de statistique descriptive univariée et bivariée, de probabilités (convergence en probabilité, loi normale, loi gamma, vecteurs gaussiens), de statistique mathématique (estimation et prévision) et d'algèbre linéaire et bilinéaire (méthode des moindres carrés, projecteurs orthogonaux, orthodiagonalisabilité).

La partie I, très proche du cours, était consacrée à la mise en évidence de résultats statistiques et algébriques.

A partir de la description et de l'écriture du modèle, la partie II se proposait de découvrir les propriétés des estimateurs des paramètres de la composante déterministe et de la fluctuation aléatoire du modèle.

Enfin, dans la partie III, l'hypothèse classique de normalité et d'indépendance des perturbations aléatoires permettait de trouver la loi de probabilité de la somme des carrés estimés du modèle ainsi qu'un intervalle de prévision de la valeur inconnue de la variable « à expliquer » correspondant à une unité statistique supplémentaire dans l'échantillon considéré.

Les résultats statistiques

Plus d'un tiers des candidats obtiennent une note supérieure à 12 et environ 12% de l'ensemble des candidats se voient attribuer une note supérieure à 16 ; enfin, 3% de candidats, soit une centaine, se situent entre 19 et 20, et parmi ceux-ci, 45 obtiennent la note maximale de 20.

Le barème de notation accordait aux trois parties du problème les poids respectifs de 30%, 19% et 51%. Les meilleures copies réalisent les 2/3 du problème, c'est-à-dire les parties I et II et quelques questions de la partie III.

Commentaires

Les remarques générales qui ressortent de l'examen des copies sont les suivantes :

1.a) Cette question facile n'a été traitée que par la moitié des candidats et a révélé de grandes lacunes logiques. Ainsi, beaucoup pensent que si les x_i ne sont pas tous égaux, ils sont alors tous distincts.

1.b)c) Les résultats étant fournis, les candidats y parviennent à peu près tous, mais plus ou moins laborieusement.
2.a) Question facilement résolue par deux candidats sur trois. Les autres ont trouvé un rang égal à 1, ou à n, voire $2n$.

2.b) Un candidat sur trois avance des critères d'inversibilité d'une matrice totalement faux : ainsi, la matrice est inversible « car elle est symétrique » ou « car elle est symétrique et ne possède pas de 0 sur sa diagonale » ou encore « car elle est symétrique et que ses coefficients sont strictement positifs ». La plupart de ceux qui ont donné une justification satisfaisante ont utilisé la caractérisation à l'aide du déterminant (au programme à l'ordre 2).

3.a) Bien que l'énoncé incitait à donner une preuve explicite du résultat, on peut considérer qu'il s'agissait d'une question de cours et qu'un théorème correctement énoncé et intégré au contexte du problème constituait une réponse satisfaisante ; encore, fallait-il préciser les hypothèses du théorème. Le total des points de cette question n'a été accordé qu'aux (rares) candidats maîtrisant vraiment le « problème des moindres carrés » et son interprétation en termes de projection orthogonale.

3.b) Question plutôt bien traitée par la moitié des candidats.

3.c)d)e) Quelques candidats connaissent la formule donnant la matrice K, mais il convenait de la justifier. Peu de candidats voient que G est une matrice de projecteur.

4.a) Seuls, quelques rares candidats ne savent pas ce qu'est un estimateur sans biais.

4.b) Le calcul de la variance de A_n a été fait correctement, en signalant l'indépendance des Y_i. Dans le calcul de la variance de B_n, beaucoup trouvent un résultat correct sans que leur réponse soit acceptable car elle se base sur l'indépendance de Y_n et A_n alors que ces deux variables aléatoires sont seulement non corrélées (l'indépendance ne s'applique que dans le cas gaussien).

4.c) Le résultat n'étant pas donné, une bonne réponse était fort appréciée, mais la même erreur que dans la question précédente amenait souvent une pollution dommageable. Elle n'a pas été sanctionnée trop lourdement afin d'éviter une « double peine ».

5. Question facile pour les candidats (nombreux) qui connaissent l'inégalité de Bienaymé-Tchebychev.

6.a) Cette question (triviale) n'a été traitée que par trois candidats sur quatre.

6.b) Très peu de candidats ont mené à bien ce calcul assez difficile. D'autres ont tenté d'abuser le correcteur en masquant des erreurs dans le cœur de leur calcul ; rappelons que cette pratique est fort peu appréciée par les correcteurs…..

6.c) Beaucoup de résultats erronés pour ce calcul assez simple mais dont le résultat n'était pas fourni.

7.a) Beaucoup d'erreurs dans le calcul de la variance d'une combinaison linéaire des variables aléatoires Y_i.

7.b) La normalité du vecteur (A_n, B_n) a rarement été établie.
8.a)b) Très peu de calculs corrects de la variance d'une combinaison linéaire des T_i. L'intérêt de l'orthogonalité de S ne pouvait être perçu qu'après un calcul exact de variance, rarement rencontré.

9.b) Il s'agissait de justifier l'application du théorème spectral en prouvant que G est symétrique réelle. Signaler que G est la matrice d'un projecteur orthogonal dans une base orthonormale était suffisant.

9.c) Très peu de succès pour ce calcul classique.

9.e) Question facile pour ceux qui connaissent les lois gamma. Très peu de réponses satisfaisantes.

10.a) Le succès n'a pas été au rendez-vous pour les candidats (nombreux) qui confondent extremum libre et extremum sous contrainte. On pouvait aussi répondre à la question, sans même avoir à vérifier que le point donné est critique, en démontrant directement et algébriquement que la valeur en ce point était la plus petite possible, sous les contraintes imposées.

Les questions suivantes ne sont quasiment jamais abordées, sauf par une poignée de très bons candidats.