Mathématiques II – Option scientifique

1 - Le sujet

Le problème étudiait quelques propriétés mathématiques du modèle de régression linéaire élémentaire, qui constitue l'instrument de base de l'économétrie.
L'économétrie mesure les phénomènes économiques quantifiables à l'aide des outils de la statistique descriptive et de la statistique mathématique.
Dans le modèle le plus élémentaire, on exprime une variable économique endogène en fonction d'une ou plusieurs variables exogènes sous la forme d'un ajustement linéaire multidimensionnel auquel on ajoute une composante (perturbation) aléatoire. Celle-ci est censée représenter d'une part toutes les erreurs de mesure et d'autre part tous les phénomènes n'entrant pas dans le modèle mais pouvant avoir (éventuellement) une influence sur la variable endogène à expliquer ; elle mesure en quelque sorte le «degré de notre ignorance» et confère ainsi au modèle un caractère aléatoire.
Le modèle va comporter un certain nombre de paramètres inconnus que l'on va pouvoir estimer à partir d'observations statistiques et d'hypothèses sur le modèle théorique retenu.
Les estimations ponctuelles et par intervalle permettent alors de porter un jugement sur la significativité des paramètres (cette question d'inférence statistique est bien entendu hors programme).
À partir du concept de matrice aléatoire, le sujet couvrait un certain nombre de points du programme d'algèbre linéaire et bilinéaire (méthode des moindres carrés, projecteurs orthogonaux, trace, etc.), de probabilités et de statistique mathématique (estimation des paramètres inconnus du modèle, biais et convergence des estimateurs retenus, matrice de variance-covariance, etc.).

2 - Les résultats obtenus

Sur l'ensemble des candidats de l'option scientifique ayant composé en mathématiques II, la moyenne s'établit à $9,65$ avec un écart-type assez élevé de $4,65$. Les premier et troisième quartiles sont égaux à $6,5$ et $12,5$ respectivement.

Le poids des différentes parties I, II A, II B et II C du problème dans la note finale a été respectivement de $20\%, 40\%, 30\%$ et 10%. Pour obtenir la note de 20, il fallait traiter correctement au moins les deux tiers du problème.

3 - Commentaires détaillés

Parmi les erreurs commises par un nombre non négligeable de candidats, on peut noter tout au long du problème le «non respect des règles du jeu» imposées par le sujet. C'est ainsi que la confusion entre variable aléatoire et vecteur aléatoire a généré des incohérences dans les calculs d'espérance ; on trouve ainsi :

- $E(Y) = E(X\alpha + U) = \alpha E(X) = \alpha X$, alors que α est un vecteur de $\mathbb{R}^k$.
- $E(Y^tY) = E(Y^2)$, alors que Y est un vecteur de $\mathbb{R}^n$.

De même, l'opérateur V est considéré comme la variance d'une variable aléatoire alors que l'énoncé précise qu'il s'agit de la matrice de variance-covariance d'un vecteur aléatoire (on lit par exemple $V(Y) = V(X\alpha + U) = V(X\alpha) + V(U)$).

On oublie également trop souvent de préciser que toute matrice symétrique à *coefficients réels* est diagonalisable.

Partie I

1 et 2 - En général, ces questions sont plutôt bien traitées par une grande majorité de candidats, excepté ceux qui égalisent une matrice à un de ses coefficients.

3 a) et b) - Si l'on fait abstraction des candidats qui confondent variable aléatoire et vecteur aléatoire, les formules demandées sont souvent démontrées.

3 c) - La première égalité est souvent bien établie par ceux qui l'abordent et il est à noter que peu de candidats travaillent en coordonnées, préférant une démonstration plus formelle en constatant notamment

que certaines des matrices considérées sont de taille 1, ce qui constitue une agréable surprise pour le correcteur. En revanche, la seconde partie de la question est très rarement menée à son terme.

Partie II

A - 1 a) Hormis les quelques candidats pour lesquels la confusion entre la matrice X de $\mathcal{M}_{n,k}(\mathbb{R})$ avec un vecteur de $\mathcal{M}_{n,1}(\mathbb{R})$ se traduit par ${}^tXX = ||X||^2$, ou bien qui ne vérifient pas que la matrice H est à coefficients réels et de taille convenable, cette question est très bien traitée.

A - 1 b) Très peu de candidats montrent que le noyau de h est réduit à $\{0\}$. Dans l'expression de la matrice H^{-1}, on voit trop so¡uvent X^{-1}, alors que X n'est pas une matrice carrée.

A - 2 Souvent, les candidats se contentent de réciter le cours (pas toujours avec toutes les hypothèses) alors qu'il fallait faire une démonstration («Montrer que... »). Toutefois, les correcteurs ont accordé des points à ceux qui citaient correctement le théorème du cours.

A - 3 a) Cette question est souvent bien résolue.

A - 3 b) On ne trouve pratiquement jamais que le rang de P est égal à k.

A - 4 a) La formule $Qu = \hat{u}$ est très rarement démontrée. On trouve souvent $Qu = y$ sans voir que $QX\alpha = X\alpha$.

A - 4 b) Cette question n'est résolue qu'à moitié en montrant que ${}^t\hat{u}\hat{u} = {}^tuQu$.

A - 5 a) Il est très rare de voir une démonstration complète (condition nécessaire et suffisante).

A - 5 b) Seule, la symétrie de la matrice ${}^tLL$ est vérifiée. La démonstration de la positivité est beaucoup moins fréquente.

B - 1 a) Il est très rare de trouver une justification correcte de la formule $V(Y) = \sigma^2 I_n$.

B - 2 b) La condition sur la matrice B est souvent trouvée, mais très fréquemment, on ne dépasse pas le calcul de ${}^tFX$.

La question B - 3 et la partie C ont été très peu abordées

BANQUE COMMUNE D'EPREUVES ECRITES 2005

CCIP : Mathématiques II - Option scientifique (épreuve n° 283)

ECOLES UTILISATRICES DE L'EPREUVE	*MOYENNES*	*ECARTS-TYPES*	*NBRE CANDIDATS*
E.M Lyon	9,42	4,40	2 680
E.S.C.P. - E.A.P.	10,01	4,29	2 320
H.E.C	**10,79**	**4,31**	**1 738**

ENSEMBLE	**9,37**	**4,40**	**2 758**

22/09/05 - DAC / CL