

## RAPPORT DE L'ÉPREUVE ÉCRITE DE MODÉLISATION MATHÉMATIQUE ET INFORMATIQUE

### Contexte scientifique du sujet de l'épreuve

L'épreuve de modélisation mathématique et informatique visait à guider les candidat(e)s dans la mise en place de différentes méthodes de modélisation mathématique issues du domaine des statistiques et de l'analyse numérique pour aborder le problème de la modélisation et de la prédiction d'observations dépendant du temps. Plus précisément, nous avons proposé dans ce sujet aux candidat(e)s de découvrir le domaine des séries temporelles ou chronologiques qui consistent à modéliser des observations dépendant du temps (ici l'évolution du niveau du lac Huron en fonction du temps) comme les réalisations d'un processus stochastique. Parmi les ouvrages de référence sur ce sujet, on peut citer [1] et [2].

Dans cette épreuve, différentes notions du programme de mathématiques de classe préparatoires ont été abordées : probabilités, algèbre linéaire et analyse. Cette épreuve contient également une partie importante d'informatique. Il est à noter que l'épreuve contenait un certain nombre de questions ouvertes permettant d'évaluer les capacités d'initiative des candidat(e)s en matière de modélisation et d'interprétation des résultats.

Nous reviendrons sur la manière dont les différentes questions ont été traitées, les principales erreurs et les points positifs relevés par l'équipe de correction. Nous proposerons à ce propos quelques pistes pour le traitement des questions les plus délicates.

### Commentaires généraux sur le traitement du sujet

Le sujet abordait des thématiques variées : statistiques, probabilités, algèbre linéaire, notions d'analyse numérique, algorithmique et programmation. La majorité des candidat(e)s a réparti son travail de façon plutôt équilibrée entre les questions théoriques et l'informatique. Parmi les points positifs l'équipe de correction a pu noter la quasi absence de copies vides, l'informatique en particulier permettant, même aux candidat(e)s les plus fragiles, d'assurer un minimum de points. Les connaissances et les méthodes classiques de programmation sont à présent bien maîtrisés, les candidat(e)s n'hésitant pas à aller chercher jusqu'au bout de la partie des questions à traiter.

Toujours dans le domaine des satisfactions, on note un réel effort concernant la rédaction. Les dérivées partielles pour la fonction de trois variables dans la partie 1 ont été trouvées pour la grande majorité des candidats, prouvant leur capacité à adapter leurs connaissances concernant les fonctions de deux variables.

Parmi les points négatifs, on peut s'étonner de la très mauvaise connaissance du théorème du rang en algèbre (partie 1). Ce théorème est un incontournable en classe préparatoire. Les étudiants donnent en majorité les énoncés erronés suivants :  $\text{rg}(A) + \dim(\text{Ker}(A)) = \dim(A)$  ou  $= np$  ou  $= \text{rg}(A) \dots$  Toujours en algèbre, on peut lire de nombreuses fois l'erreur grave :  ${}^tAAu = 0$  est équivalent à  ${}^tA = 0$  ou  $Au = 0$ . En analyse, la confusion : continue implique dérivable a été lue malheureusement à plusieurs reprises.

# Commentaires détaillés et pistes de résolution

## PARTIE 1

- 1) a) Un schéma peut être utile à l'explication pour mettre en avant la parabole et la notion d'approximation. "Minimiser  $F(a, b, c)$  revient à chercher  $a, b, c$  tels que  $F(a, b, c)$  soit minimal" est une explication peu convaincante "droite de régression linéaire" lancée sans plus de commentaires non plus. Encore plus obscur : "l'ajustement du polynôme de degré 2 est beaucoup plus linéaire que la droite".
- 1) b) Une fonction de plusieurs variables n'est pas dérivable. Trop souvent on rencontre continue donc dérivable.
- 2) Calculs des dérivées partielles et produit scalaire satisfaisants.
- 3) Beaucoup d'erreurs sur cette question, la plupart des candidat(e)s affirment que " $Au=0$  et  $A$  non nul implique  $u=0$ " ou que "une somme de termes est nulle si tous les facteurs sont nuls" ou encore " $\langle u, v \rangle = 0$  implique  $u = 0$  ou  $v = 0$ ".
- 4) a) L'identification de la double inclusion est déjà une difficulté. L'intervention de la norme pour prouver la nullité d'un vecteur n'est pas un réflexe. Il est beaucoup trop fréquent de voir ' $AAu = 0$  et  $A \neq 0$  donc  $Au = 0$ . Enfin le  $\ker(A) = 0$  est encore bien fréquent.
- 4) b) Le théorème du rang appliqué à une matrice reste une difficulté importante. On rappelle aussi que  $\dim(A)$  n'a pas de sens.
- 5) Il convient d'être vigilant avec un produit matriciel non commutatif. Par ailleurs, beaucoup de candidats inversent des matrices rectangulaires.

## PARTIE 2

- 2) c) Le principe de la double affectation ou de l'échange de variables semble bien acquis en majorité. Attention aux fonctions qui doivent modifier l'entrée à la différence de celles qui doivent renvoyer quelque chose. Le module "copy" n'est pas au programme encore moins l'instruction "deepcopy". En cas d'utilisation, il convient de bien connaître la syntaxe surtout si on peut s'en passer.
- 3) Plusieurs copies renvoient la valeur maximale au lieu de renvoyer l'indice.
- 4) a) L'algorithme du pivot de Gauss est bien identifié néanmoins la réalisation pratique sur une matrice de taille 3 a conduit à beaucoup d'erreurs de calculs.
- 5) a) Beaucoup d'erreurs dans les "range".
- 6) a) Deux boucles imbriquées suivies d'un "if  $j == i + n$ " même si syntaxiquement correct peut se simplifier aisément en une seule boucle.

## PARTIE 3

- 1) Beaucoup de candidat(e)s appliquent une recette de cuisine sans explication. Dire au moins EDL à coefficients constants par exemple et parler de la structure des solutions. Par ailleurs, utiliser la méthode de variations de la constante est bien légitime mais ne pas négliger les solutions évidentes. Enfin, la recherche de la valeur de la constante d'intégration doit se faire, via la condition initiale, sur la solution complète pas sur la solution homogène!
- 2) La disjonction de cas est bien maîtrisée en général.
- 3) Lorsque la solution d'une EDL est donnée dans l'énoncé, on peut toujours vérifier qu'elle fonctionne en l'injectant dans l'EDL.
- 4) Question peu abordée. Si elle l'était, c'était par le principe de superposition et non par récurrence.

- 5) Le changement d'indice n'a été trouvé que dans les meilleures copies. Les étudiant(e)s se sont souvent contenté(e)s d'analyser l'intervalle de sommation sans regarder le terme  $u_j$  qui devient  $u_{j-1}$ .
- 6) On attendait une interprétation du type suivant pour cette question : l'équation précédente montre que le niveau du lac à l'instant  $t_i(x(t_i))$  est une somme pondérée des précipitations des années précédentes  $u_{j-1}$ , avec une pondération décroissante ( $\Phi$ ) suivant l'éloignement dans le temps ( $\cdot$ ) des précipitations considérées.

#### PARTIE 4

- 1) a) Erreur vue plusieurs fois :  $Cov(Z_n, Z_n) = E(Z_n^2) - E(Z_n)^2 = E(Z_n)^2 - E(Z_n)^2 = 0$  car les  $Z_n$  sont indépendants ...
- 1) b) Les questions de synthèse ne doivent pas être manquées mais nécessitent de bien dresser le bilan des différents points importants.
- 3) b) Le  $X_n$  de la question 3 n'est pas le même que dans la question 2.
- 3) e) Parmi les candidats qui ont obtenu l'expression, peu font le lien avec l'équation (3). On attendait une réponse du type suivant : on retrouve la somme pondérée de la partie 3 avec des coefficients qui sont ceux d'une série géométrique de raison strictement plus petite que 1 en valeur absolue.

#### PARTIE 5

- 1) L'interprétation du modèle a été très décevante. Il s'agit d'une compétence importante à ne pas négliger.
- 2) a) Lorsqu'on manipule une double somme, il est important de bien utiliser des indices de sommation différents.
- 2) c) Il s'agit d'une question de comptabilité; la somme  $\sum_{p=1}^n p$  est bien connue mais le manque de lucidité de fin d'épreuve a conduit plusieurs candidats à manquer  $\sum_{p=1}^{n-1} p$ .
- 2) d) Très peu traité. Les contrôles avec domination bien propres sont très rares.
- 3) On attendait dans cette question une réponse du type suivant : on prédit le niveau de l'année  $n$  comme  $\hat{\mu} + \varphi \hat{X}_{n-1}$  où  $X_{n-1}$  est le niveau du lac l'année  $n - 1$ .

## Références

- [1] P. J. Brockwell and R. A. Davis. *Time Series : Theory and Methods*. Springer Verlag, New York, second edition edition, 1991.
- [2] P. J. Brockwell and R. A. Davis. *Introduction to time series and forecasting*. SpringerVerlag Inc, Berlin; New York, 2002.