CONCOURS INTERNE 2018

POUR LE RECRUTEMENT D'ADMINISTRATEURS STAGIAIRES DE L'INSEE

COMPOSITION DE MATHÉMATIQUES ET STATISTIQUES

(durée : 4 heures)

Cette composition comporte 8 pages :

L'épreuve est constituée de deux parties indépendantes, chacune comptant pour moitié dans la note finale.

Les exercices sont indépendants et sont tous à traiter.

L'usage de la calculatrice est interdit.

Vous composerez pour chacune de ces deux parties sur une copie séparée.

Institut National de la Statistique et des Études Économiques - 88 avenue Verdier – CS 70058 - 92541 MONTROUGE CEDEX - FRANCE - www.insee.fr
Tél. standard : 01.41.17.50.50 - N° SIRET : 120 027 016 00019 - Code APE : 84.11Z - Service Insee Contact : 09 72 72 4000 - (tarification "appel local")

Concours administrateur Insee interne 2018

L'épreuve est constituée de deux parties indépendantes, chacune comptant pour moitié dans la note finale.

Partie 1 : analyse-algèbre

Exercice 1

Dans tout l'exercice, n désigne un entier naturel supérieur ou égal à 1 et E_n l'espace vectoriel des polynômes de degré inférieur ou égal à n.

Pour tout k dans $[\![0,n]\!]$, on note $\binom{n}{k}$ le coefficient du binôme défini par : $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$

1. Montrer que, pour tout élément k de $[\![0,n]\!]$, il existe un unique polynôme $L_{n,k}$ de degré inférieur ou égal à n tel que :

$$\forall j \in [\![0,n]\!], \quad L_{n,k}(j) = \begin{cases} 1 \text{ si } j = k \\ 0 \text{ si } j \neq k \end{cases}$$

et que ce polynôme est donné par : $L_{n,k}(X) = (-1)^{n-k} \dfrac{\binom{n}{k}}{n!} \displaystyle\prod_{\substack{j=0 \\ j \neq k}}^{n} (X - j)$

2. Montrer que $\mathcal{B} = (L_{n,0}, L_{n,1}, \dots, L_{n,n})$ est une base de E_n.

3. On définit, pour $(P,Q) \in E^2$, $\langle P, Q \rangle = \sum\limits_{k=0}^{n} P(k)Q(k)$.

   a) Montrer que l'on définit ainsi un produit scalaire sur E_n. On notera $||.||$ la norme associée à ce produit scalaire.

   b) Que peut-on dire de $\mathcal{B}$ pour ce produit scalaire ?

4. Soit k un entier tel que $0 \leq k \leq n$.
   Déterminer un réel λ_k tel que $X^n + \lambda_k L_{n,k} \in E_{n-1}$.

5. Soit $P \in E_{n-1}^{\perp}$. Montrer que les coordonnées de P sur la base $\mathcal{B}$ sont $\alpha_0, \dots \alpha_n$, où pour tout k de $[\![0,n]\!]$:

$$\alpha_k = \frac{(-1)^{n-k}}{k!(n-k)!} \sum_{j=0}^{n} P(j) j^n$$

6. (a) Montrer que pour tout k de $[\![1,n]\!]$ on a : $L'_{n,k}(0) = (-1)^{k-1} \dfrac{\binom{n}{k}}{k}$.

   (b) Calculer $L'_{n,0}(0)$.

7. Montrer qu'il existe un unique polynôme A appartenant à E_n tel que :

$$\forall P \in E_n, \; < A, P >= P'(0)$$

et que ce polynôme est donné par : $A = \displaystyle\sum_{j=0}^{n} L'_{n,j}(0) L_{n,j}$

8. On désigne par H_n la somme partielle de la série harmonique : $H_n = \displaystyle\sum_{k=1}^{n} \frac{1}{k}$.

   Établir la formule suivante :

$$\forall P \in E_n, \; \sum_{k=1}^{n} (-1)^{k-1} \frac{\binom{n}{k}}{k} P(t+k) = P'(t) + H_n P(t)$$

9. En déduire enfin la formule sommatoire suivante : $\displaystyle\sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} \binom{n}{k} (n-k)^n = n^n (H_n - 1)$.

Exercice 2

1. Étudier la convergence de l'intégrale $\int_0^{+\infty} \frac{\sin^2 t}{t^2} dt$. On notera S sa valeur.
2. Montrer que : $S = \int_0^{+\infty} \frac{\sin t}{t} dt$.

On définit maintenant les fonctions :

$$\psi : x \geq 0 \to \int_0^{+\infty} e^{-tx} \frac{\sin t}{t} \, dt$$

$$\text{et : } \psi_n : x \geq 0 \to \int_0^{n} e^{-tx} \frac{\sin t}{t} \, dt \text{ pour } n \in \boldsymbol{N}^*.$$

3.
    a. Montrer que : $\forall x \geq 0,\ \forall x_0 \geq 0,\ \forall t \geq 0 : \left| e^{-tx} - e^{-tx_0} \right| \leq t \left| x - x_0 \right|$.
    b. En déduire que chaque fonction ψ_n est continue sur $\boldsymbol{R}^+$.

4.
    a. Vérifier que la fonction ψ est bien définie.
    b. Montrer que la suite de fonctions $\{\psi_n\}$ converge *uniformément* vers ψ, quand $n \to +\infty$, sur tout intervalle $[\rho, +\infty[$, avec $\rho > 0$.
    c. Qu'en déduit-on, en termes de propriétés de la fonction ψ ?
    d. Montrer que : $\forall x > 0 : \left| \psi(x) \right| \leq \frac{1}{x}$.

Des raisonnements analogues à ceux des questions 3 et 4 permettent de démontrer que :

- chaque fonction ψ_n est dérivable sur $\boldsymbol{R}^+$ et : $\psi'_n(x) = -\int_0^n e^{-tx} \sin t \, dt$.
- la suite de fonctions $\{\psi'_n\}$ converge *uniformément*, quand $n \to +\infty$, vers la fonction $h : x > 0 \to -\int_0^{+\infty} e^{-tx} \sin t \, dt$, sur tout intervalle $[\rho, +\infty[$, avec $\rho > 0$.

5. On fixe $\rho > 0$. En exprimant $\psi_n(x)$, pour $x \in [\rho, +\infty[$, sous forme d'une intégrale faisant intervenir la fonction ψ'_n, montrer que la fonction ψ est dérivable sur $[\rho, +\infty[$ et que : $\psi' = h$.
6. Calculer explicitement les valeurs $h(x)$ pour $x \geq \rho > 0$, au moyen des fonctions usuelles. En déduire la valeur de ψ sur $\boldsymbol{R}^{+*}$.

7. On cherche à montrer que ψ est continue en 0.

   a. Démontrer la relation :

   $$\forall x > 0 : \psi(x) - \psi(0) = x \int_0^{+\infty} [1 - \cos(\frac{u}{x})] \frac{1 - e^{-u}(1+u)}{u^2} \; du .$$

   b. En déduire qu'il existe une constante $C > 0$ telle que : $\forall x > 0 : |\psi(x) - \psi(0)| \leq C x$

   c. Conclure.

   d. En déduire la valeur de S.

■ ■ ■ ■ ■ ■

Partie 2 : probabilités-statistiques

Cette partie est constituée de deux exercices indépendants

Exercice 1

On note λ et θ deux réels vérifiant $\lambda > 0$, $\theta > 1$ et X une variable aléatoire qui suit la loi de Pareto de paramètre λ et θ, dont une densité est :

$$f_X(t) = \begin{cases} \dfrac{\lambda}{\theta}\left(\dfrac{\theta}{t}\right)^{\lambda+1} & \text{si } t > \theta \\ \\ 0 \text{ sinon} \end{cases}$$

On note :$X \hookrightarrow \mathcal{VP}(\lambda, \theta)$ On considère une suite $(X_n)_{n\in\mathbb{N}}$ de variables aléatoires indépendantes suivant toutes la loi $\mathcal{VP}(\lambda, \theta)$

1. Déterminer la fonction de répartition de X.
2. On pose, pour tout entier n supérieur où égal à 1 : $Z_n = \displaystyle\prod_{k=1}^{n} X_k$.
   (a) Établir, pour toute variable aléatoire Z l'inégalité suivante :
   $$\mathbb{P}\Big(\big[X_{n+1} > \frac{3}{2}\big]\Big) \leqslant \mathbb{P}\left(\left[|Z_{n+1} - Z| > \frac{1}{4}\right]\right) + \mathbb{P}\left(\left[|Z_n - Z| > \frac{1}{4}\right]\right)$$
   (b) Étudier la convergence en probabilité de la suite $(Z_n)_{n\in\mathbb{N}^*}$.

On admet pour la suite le résultat suivant :

Si (X_n) converge en loi vers la loi de la variable aléatoire X et que (T_n) converge en probabilité vers la constante c, alors (X_nT_n) converge en loi vers la loi de la variable cX

3. On suppose dans cette question que θ est connu et que λ est un paramètre inconnu que l'on souhaite estimer.

   On pose : $Y_i = \ln\left(\dfrac{X_i}{\theta}\right)$, $\overline{Y}_n = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} Y_k$ et $\hat{\lambda}_n = \dfrac{1}{\overline{Y}_n}$.

   (a) i. Déterminer la loi de Y_k.
   ii. En déduire les valeurs de $\mathbb{E}(\overline{Y}_n)$ et de $\mathrm{Var}(\overline{Y}_n)$.

   (b) Justifier les deux convergences suivantes :
   i.
   $$\lambda\sqrt{n}\left(\overline{Y}_n - \frac{1}{\lambda}\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$
   ii.
   $$\lambda\overline{Y}_n \xrightarrow{P} 1$$

   (c) Montrer que :
   $$\frac{\sqrt{n}(\hat{\lambda}_n - \lambda)}{\lambda} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$$

   (d) En approchant la loi de $\dfrac{\sqrt{n}(\hat{\lambda}_n - \lambda)}{\lambda}$ par $\mathcal{N}(0,1)$, déterminer, en fonction de n et de la valeur observée λ_0, un intervalle de confiance au risque 0.05 pour λ. (on rappelle que, en notant Φ la fonction de répartition de la loi normale centrée réduite, on a $\Phi(1,96) = 0,975$).

Exercice 2

1ère partie : préliminaires probabilistes

On considère deux variables aléatoires réelles *indépendantes* :

$$U \leadsto \mathcal{N}(0, \sigma^2) \quad (\sigma > 0)$$
$$X \leadsto \mathcal{B}(1, p) \text{ [loi de BERNOULLI] } (p \in \,]0, 1[).$$

a étant un paramètre réel, on définit la variable aléatoire : $Y = aX + U$.

1. Calculer $EY, VY, E(XY), V(XY)$.

2.
   a. Calculer la fonction de répartition de Y [on retient, par convention qu'elle s'écrit : $P\{Y \leq y\}$]. On l'exprimera en fonction de la fonction de répartition de la loi $\mathcal{N}(0, 1)$, notée H.
   b. En déduire la densité de la loi de Y.

3.
   a. Calculer la fonction de répartition de la loi conditionnelle de Y sachant $X = x$.
   b. Calculer la loi conditionnelle de X sachant $Y \leq y$, puis celle de X sachant $y < Y \leq y + \eta$ ($\eta > 0$).
   c. On note : $P\{X = x / Y = y\} = \lim_{\eta \to 0^+} P\{X = x / y < Y \leq y + \eta\}$.
   Calculer la loi $P\{X = 1 / Y = y\}$. *On montrera qu'elle s'exprime, dans le cas* $a > 0$, *au moyen de la fonction de répartition d'une loi logistique de la forme :* $\dfrac{1}{1 + A e^{-By}}$.

4. On pose : $Z = Y(1 - X)$. Calculer la fonction de répartition de Z (exprimée au moyen de la fonction H définie ci-dessus).

5. Lorsqu'on ne peut pas observer la variable aléatoire X, on cherche à en construire une approximation à partir de l'observation de la variable Y. Pour cela, on définit :

$$X^* = \left| \begin{array}{l} 1 \ si \ Y > c \\ 0 \ si \ Y \leq c \end{array} \right.,$$ où c est un réel donné. On suppose ici : $\boxed{a > 0}$.

Déterminer la valeur de c minimisant l'erreur quadratique de prévision : $E(X - X^*)^2$.

2ème partie : estimation

On considère maintenant une suite de couples de variables aléatoires $(X_i, U_i)_{i=1,\ldots,N}$, indépendants et dont les lois et les constructions sont identiques à celle du couple (X, U) de la première partie.
On pose également : $Y_i = aX_i + U_i$. Les observations disponibles sont les couples $(X_i, Y_i)_{i=1,\ldots,N}$.

6.
   a. Donner l'expression de l'estimateur des moindres carrés ordinaires de a, soit $\hat{a}_N$, en fonction de a et des X_i et des U_i.
   b. Étudier la convergence en probabilité de l'estimateur $\hat{a}_N$ quand $N \to +\infty$.
   c. Étudier la convergence en loi de $\sqrt{N}(\hat{a}_N - a)$ quand $N \to +\infty$.

7. Comment estimeriez-vous les paramètres p et σ^2 ? Étudier la convergence en probabilité des estimateurs proposés.