



PRÉPASIMPLE

prepasimple.fr

HEC / ESCP

Mathématiques 2

1995

Sujet de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Sujet

Partie I. La loi de Pareto

Toutes les variables aléatoires considérées sont réelles. Dans tout le problème, x_0 , C et α désignent des réels vérifiant $\alpha > 0$ et $x_0 + C > 0$.

On dit que X suit une loi de Pareto de paramètres α, x_0, C si elle est à valeurs dans $[x_0, +\infty[$ et admet pour densité

$$f(x) = \begin{cases} 0 & \text{si } x < x_0, \\ \frac{\alpha(x_0 + C)^\alpha}{(x + C)^{\alpha+1}} & \text{si } x \geq x_0. \end{cases}$$

On note cette loi $VP(\alpha, x_0, C)$, loi de Pareto à trois paramètres.

Lorsque $C = 0$, on utilise la notation $VP(\alpha, x_0)$ pour la loi à deux paramètres. Dans ce cas, $x_0 > 0$ et sa densité est

$$f(x) = \begin{cases} 0 & \text{si } x < x_0, \\ \frac{\alpha x_0^\alpha}{x^{\alpha+1}} & \text{si } x \geq x_0. \end{cases}$$

A. Quelques résultats probabilistes

1. Soit $X \sim VP(\alpha, x_0, C)$.

1.a) Vérifier que f est bien une fonction de densité.

1.b) Déterminer la fonction de répartition de X .

2. Soit $X \sim VP(\alpha, x_0)$.

2.a) Déterminer les valeurs de α pour lesquelles X admet une espérance et la calculer dans ce cas.

2.b) Déterminer les valeurs de α pour lesquelles X admet une variance et la calculer dans ce cas.

3. Soit $X \sim VP(\alpha, x_0)$ et $\lambda > 0$. Déterminer la fonction de répartition de $Y = \lambda X$. Quelle loi reconnaît-on ?

4.a) Soit $X \sim VP(\alpha, x_0)$ et $\mu \in \mathbb{R}$. Montrer que $U = X + \mu$ suit une loi de Pareto à trois paramètres que l'on déterminera.

4.b) Réciproquement, soit $Z \sim VP(\alpha, x_0, C)$. Quelle est la loi de $V = Z + C$?

4.c) Soit $Z \sim VP(\alpha, x_0, C)$. Dédurre des questions 2 et 4.b) les valeurs de α pour lesquelles Z admet une espérance et la déterminer dans ce cas. Déterminer les valeurs de α pour lesquelles Z admet une variance et la calculer dans ce cas.

5. Soit W de loi exponentielle de paramètre $\beta > 0$, admettant la densité

$$g(x) = \begin{cases} 0 & \text{si } x < 0, \\ \beta e^{-\beta x} & \text{si } x \geq 0. \end{cases}$$

Soit $k > 1$ et $x_0 > 0$. Déterminer la fonction de répartition de $T = x_0 k^W$. Quelle loi reconnaît-on ?

6. Soit $X \sim VP(\alpha, x_0)$. Déterminer la loi de $\sqrt{X}$.

B. Propriété caractéristique de la loi de Pareto

Soit X de densité f , admettant une espérance, et soit $x \in \mathbb{R}$ tel que $\int_x^{+\infty} f(t) dt \neq 0$.

On appelle moyenne de X sur $[x, +\infty[$ le réel

$$M_X(x) = \frac{\int_x^{+\infty} t f(t) dt}{\int_x^{+\infty} f(t) dt}.$$

1. Montrer que $M_X(x) > x$.
2. On suppose que $X \sim VP(\alpha, x_0)$ avec $\alpha > 1$. Calculer $M_X(x)$ pour $x \geq x_0$.
3. Réciproquement, soit $x_0 > 0$ et soit X à valeurs dans $[x_0, +\infty[$, de densité f continue et strictement positive sur $[x_0, +\infty[$, admettant une espérance. On suppose qu'il existe $k > 1$ tel que $M_X(x) = kx$ pour tout $x \geq x_0$.

On se propose d'établir que X suit une loi de Pareto à deux paramètres. Pour $x \geq x_0$, on pose

$$G(x) = \int_x^{+\infty} f(t) dt, \quad H(x) = \int_x^{+\infty} t f(t) dt.$$

3.a) Montrer que G et H sont dérivables sur $[x_0, +\infty[$, au sens de la dérivée à droite en x_0 , et calculer leurs dérivées.

3.b) Prouver que

$$G(x) = \frac{1-k}{k} x G'(x) \quad (x \geq x_0).$$

3.c) Pour $x \geq x_0$, on pose $I(x) = x^{k/(k-1)} G(x)$. Calculer $I'(x)$ et en déduire la valeur de $G(x)$.

3.d) En déduire que X suit $VP(k/(k-1), x_0)$.

4. Soit $x_0 \in \mathbb{R}$ et soit X à valeurs dans $[x_0, +\infty[$, de densité f continue et strictement positive sur $[x_0, +\infty[$, admettant une espérance.

4.a) Soit $\lambda \in \mathbb{R}$ et $Y = X + \lambda$. Prouver que

$$M_Y(y) = M_X(y - \lambda) + \lambda,$$

pour les valeurs de y où ces moyennes sont définies.

4.b) On suppose que $X \sim VP(\alpha, x_0, C)$ avec $\alpha > 1$. Exprimer $M_X(x)$ en fonction de x pour $x \geq x_0$.

4.c) Réciproquement, on suppose qu'il existe $h, k \in \mathbb{R}$, avec $k > 1$ et $x_0 + h/(k-1) > 0$, tels que

$$M_X(x) = kx + h \quad (x \geq x_0).$$

Soit $Y = X + h/(k-1)$. Calculer $M_Y(y)$ pour $y \geq x_0 + h/(k-1)$. En déduire la loi de Y , puis montrer que

$$X \sim VP\left(\frac{k}{k-1}, x_0, \frac{h}{k-1}\right).$$

5. Un exemple statistique : la répartition des revenus

Des statistiques provenant de la Direction générale des impôts donnent, pour l'année 1988, la répartition des revenus d'environ 25 000 000 de contribuables.

On note x le niveau de revenu en KF, milliers de francs ; $1 - F(x)$ la proportion de contribuables ayant un revenu strictement supérieur à x ; et $M(x)$ le revenu moyen des contribuables ayant un revenu strictement supérieur à x .

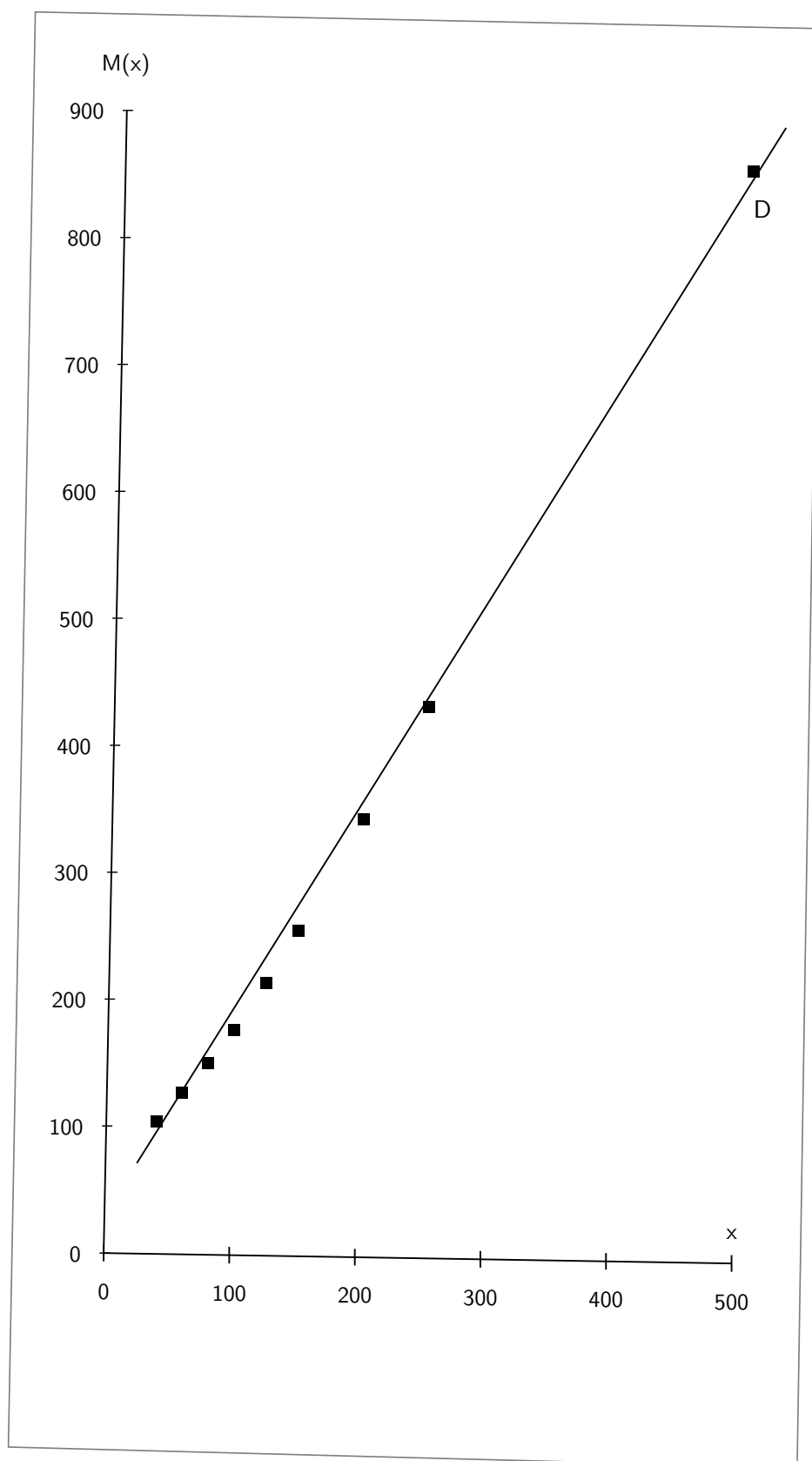
On dispose du tableau suivant :

x	$1 - F(x)$	$M(x)$
500	0,006	860
250	0,027	434
200	0,046	345
150	0,095	256

x	$1 - F(x)$	$M(x)$
125	0,145	215
100	0,226	177
80	0,323	151
60	0,458	127
40	0,658	104

Les graphiques ci-dessous représentent ces données et ne sont pas à reproduire sur la copie.

5.1) Le graphique 1 représente les points d'abscisse x et d'ordonnée $M(x)$ pour les valeurs du tableau. Une étude statistique permet d'estimer que ce nuage peut être modélisé par une droite D , figurant sur le graphique.



Revenus 1988 - Graphique n°1

5.1.a) Lire sur le graphique le coefficient directeur de D .

5.1.b) Expliquer pourquoi on peut modéliser la distribution des revenus par une loi de Pareto à trois paramètres $VP(\alpha, x_0, C)$.

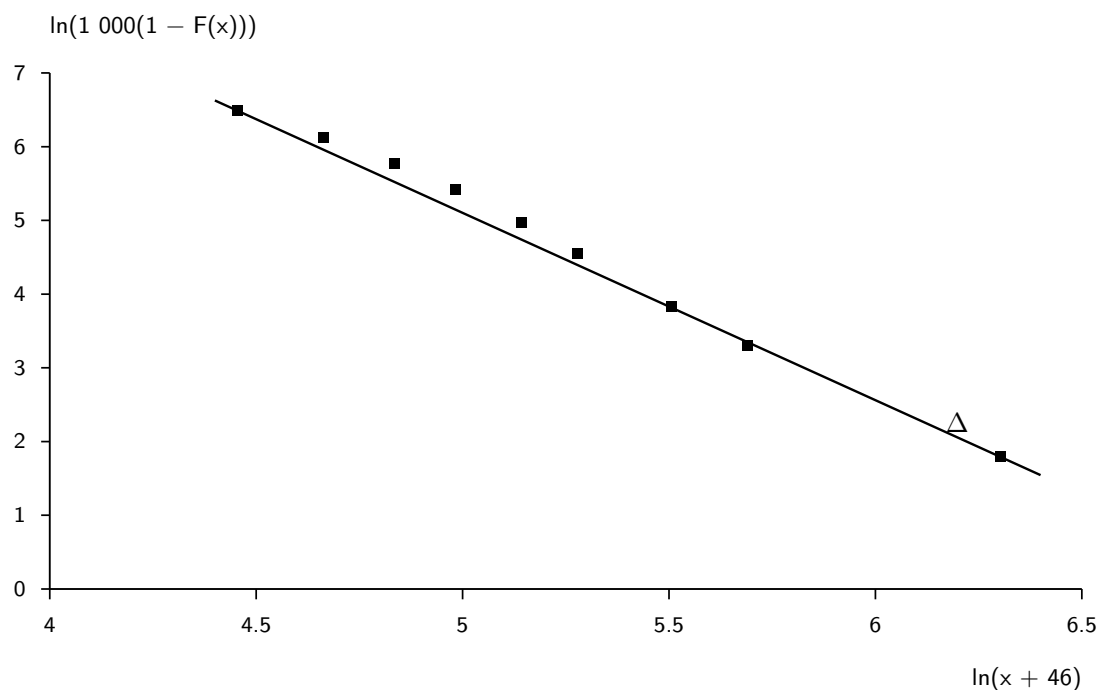
5.1.c) Donner, d'après le graphique, une valeur approchée de α et de C .

5.1.d) Sachant que le revenu moyen de tous les contribuables est $\overline{M} = 75$ KF, donner à l'aide du graphique une valeur approchée de x_0 .

5.2) Pour construire le graphique 2, on retient $C_{gr} = 46$ KF, valeur approchée compatible avec la lecture du graphique 1. Pour chaque ligne du tableau, on calcule les coordonnées

$$u = \ln(x + C_{gr}), \quad v = \ln(1\,000(1 - F(x))).$$

On rappelle que $\ln$ désigne le logarithme népérien. On définit Δ comme la droite passant par les deux points associés aux revenus $x = 40$ et $x = 500$.



Revenus 1988 - Graphique n°2

5.2.a) Déterminer le coefficient directeur de Δ à partir des coordonnées de ces deux points.

5.2.b) Expliquer pourquoi l'allure de ce graphique est cohérente avec la modélisation de la distribution des revenus par une loi de Pareto à trois paramètres.

5.2.c) Retrouver ainsi une valeur approchée de α et de x_0 , en utilisant $C \simeq C_{\text{gr}}$.

Partie II. Courbe de concentration et inégalité des revenus

On désigne par F_x la proportion des individus d'une population ayant un revenu inférieur ou égal à x , et par Q_x le quotient de la masse des revenus de ces mêmes individus par la masse totale des revenus de la population.

On appelle courbe de concentration la représentation graphique de la fonction donnant Q_x en fonction de F_x . Dans la suite, X représente le revenu d'un individu de cette population.

A. Courbe de concentration et indice de Gini

On suppose $x_0 > 0$ et X à valeurs dans $[x_0, +\infty[$, de densité f continue et strictement positive sur $[x_0, +\infty[$, possédant une espérance $E(X)$.

Pour $x \geq x_0$, on pose

$$F(x) = \int_{x_0}^x f(t) dt, \quad Q(x) = \frac{1}{E(X)} \int_{x_0}^x t f(t) dt.$$

1.a) Montrer que la restriction de F à $[x_0, +\infty[$ admet une application réciproque notée F^{-1} . Quel est le domaine de définition de F^{-1} ?

1.b) On pose $\mathcal{C} = Q \circ F^{-1}$. Montrer que $\mathcal{C}$ se prolonge en une fonction continue strictement croissante de $[0, 1]$ dans lui-même. Déterminer $\mathcal{C}(0)$ et $\mathcal{C}(1)$.

Ainsi, la courbe de concentration de X existe ; c'est la courbe représentative de $\mathcal{C}$ dans un repère orthonormé du plan.

2.a) Prouver que $\mathcal{C}$ est dérivable sur $[0, 1[$, au sens de la dérivée à droite en 0, et que

$$\mathcal{C}'(t) = \frac{F^{-1}(t)}{E(X)} \quad (t \in [0, 1]).$$

2.b) En déduire sans calcul que $\mathcal{C}$ est convexe et que la courbe de concentration est située en dessous de la première bissectrice.

2.c) Déterminer $\mathcal{C}'(0)$ et $\lim_{t \rightarrow 1^-} \mathcal{C}'(t)$.

2.d) Tracer l'allure de la courbe de concentration en précisant les tangentes aux deux extrémités.

On appelle indice d'inégalité de Gini de X le réel $I(X)$ égal à deux fois l'aire située entre la courbe de concentration et la première bissectrice :

$$I(X) = 2 \int_0^1 (t - \mathcal{C}(t)) dt.$$

On estime que plus $I(X)$ est grand, plus l'inégalité des revenus est grande.

B. Application : comparaison de quelques procédures d'imposition des revenus

Dans toute cette partie, $X \sim VP(\alpha, x_0)$ avec $\alpha > 1$.

1.a) Calculer $Q(x)$ pour $x \geq x_0$.

1.b) Prouver que

$$\mathcal{C}(t) = 1 - (1 - t)^{(\alpha-1)/\alpha} \quad (t \in [0, 1]).$$

1.c) Déterminer α si les 30 % d'individus ayant les plus hauts revenus se partagent 60 % de la masse des revenus. Dans la suite, α n'est plus supposé égal à cette valeur.

1.d) Déterminer l'indice d'inégalité de Gini $I(X)$.

2. On prélève sur tous les revenus un impôt proportionnel. Le revenu disponible après imposition est $Y = (1 - \lambda)X$, avec $\lambda \in]0, 1[$.

2.a) Calculer $I(Y)$. On pourra utiliser I.A.3.

2.b) Quel est l'effet de cette imposition sur l'inégalité des revenus?

3. Le revenu disponible après une imposition progressive est $T = h\sqrt{X}$, avec $h > 0$.

3.a) À l'aide de la partie I, déterminer la loi de T et calculer $I(T)$.

3.b) Quel est l'effet de cette imposition sur l'inégalité des revenus?

4. On prélève sur tous les revenus un impôt constant $a \in]0, x_0[$. Le revenu disponible après imposition est $Z = X - a$.

4.a) Déterminer la loi de Z .

4.b) On note $\mathcal{C}_Z$ la fonction dont le graphe est la courbe de concentration de Z . Montrer que

$$t - \mathcal{C}_Z(t) = \frac{E(X)}{E(X) - a}(t - \mathcal{C}(t)) \quad (t \in [0, 1]).$$

4.c) En déduire $I(Z)$.

4.d) Quel est l'effet de cette imposition sur l'inégalité des revenus?