

Le tableau ci-dessous donne, pour l'année 19xx, les tranches de revenus des ménages exprimées en FF. Le REVENU est ici une variable quantitative continue comportant 9 classes.

REVENU (en FF)	3250-3750	3750-4250	4250-4750	4750-5250	5250-5750	5750-6250	6250-6750	6750-7250	7250-7750
Effectifs ( $n_i$ )	28	20	14	11	8	7	5	4	3
Centres des classes ( $r_i$ )									
Amplitudes des classes ( $a_i$ )									
Fréquences ( $f_i$ )									
Fréquences cumulées ( $F_i$ )									

TRAVAIL DEMANDE :

1. Statistique descriptive

- a) Compléter le tableau ci-dessus, puis représenter la *fonction de répartition empirique* et l'*histogramme* associés à la variable continue REVENU. (3p)
- b) Déduire le *mode* de la variable continue REVENU. Dans quelle classe se situe la *médiane* ? Calculer, par interpolation linéaire, sa valeur. (2p)
- c) Calculer la *moyenne*, la *variance* et l'*écart-type empiriques* associés à la variables continue REVENU. Comparer les valeurs des 3 indicateurs de tendance centrale : *mode* , *médiane* et *moyenne*. Conclure. (3p)

2. Statistique inférentielle

On assimile chacune des *valeurs observées* sur l'échantillon *i.i.d.* Tiré, valeurs représentées dans le tableau ci-dessus, au *centre de sa classe*. On modélise le *revenu d'un ménage* choisi au hasard par une *variable aléatoire* que l'on note *R*. Cette variable aléatoire suit une loi de Pareto dont la densité est la suivante pour un revenu *r* :

$$f_R(r) = \begin{cases} \frac{ba^b}{r^{b+1}} & \text{si } r \geq a \\ 0 & \text{sinon} \end{cases} \text{ avec } a \text{ et } b > 0$$

- a) Donner l'expression de la fonction de vraisemblance  $a \rightarrow L(r_1, r_2, \dots, r_n; a)$  , puis représenter graphiquement la courbe associée. Déduire l'estimateur  $\hat{a}_{MV}$  du *max. de vraisemblance* du paramètre *a*. Montrer que, pour *b* fixé, l'estimation du *max. de vraisemblance* du paramètre *a* est 3500. Quelle est donc la signification d'un point de vue économique du paramètre *a* ? (3p)
- b) Déterminer, pour *a* fixé, l'estimateur  $\hat{b}_{MV}$  du *max. de vraisemblance* du paramètre *b*. Calculer sa réalisation à partir des résultats du sondage. (3p)
- c) Est-ce que l'estimateur  $\hat{b}_{MV}$  du paramètre *b* est efficace ? Sinon, l'estimateur  $\hat{u}(b)$  d'une fonction  $u(b)$  de *b* ? Calculer la variance de cet estimateur  $\hat{u}(b)$  . (3p)
- d) Calculer l'espérance et la variance de l'estimateur  $\hat{a}_{MV}$  . (Indication : Calculer la fonction de répartition de  $\hat{a}_{MV}$  ). Déterminer le risque statistique associé à cet estimateur. (3p)

**NOTA :** Le médian est noté sur 20 points.