CONTROLE CONTINU DE STATISTIQUES DESCRIPTIVES
L1 AES
- Correction -

EXERCICE 1 (8,5 pts)

Les deux parties sont indépendantes et peuvent être traitées séparément.

Une étude sur le budget consacré aux vacances d'été auprès de ménages Bretons a donné les résultats suivants.

Budget	Fréquence cumulée
	0
[800-1000[	
	0,08
[1000-1400[	
	0,18
[1400-1600[	
	0,34
[1600-y[	
	0,64
[y-2400[	
	0,73
[2400-x[	
	1

PARTIE 1

1/ Certaines données sont manquantes. Calculer la borne manquante x sachant que l'étendue de la série est égale à 3200 (0,5pt).

On sait que l'étendue est égale au maximum-le minimum. Ainsi :
$3200 = Max - Min = Max - 800$ et donc $Max = 4000$

2/ Calculer la borne manquante y dans les deux cas suivants :

a- Le budget moyen est égal à 1995 euros (1,5 pts).

Si le budget moyen est égal à 1995 euros :

$$\bar{x} = 1995 = \frac{1}{N}\sum_i n_i x_i = \sum_i f_i x_i$$

Il faut donc au préalable calculer les fréquences à partir des fréquences cumulées dans le tableau précédent.

Budget	Fréquence cumulée	Fréquences
	0	
[800-1000[		0,08
	0,08	
[1000-1400[		0,1
	0,18	
[1400-1600[		0,16
	0,34	
[1600-y[		0,3
	0,64	
[y-2400[		0,09
	0,73	
[2400-x[		0,27
	1	

Ainsi :

$$\begin{aligned}\bar{x} &= 1995\\ &= \sum_i f_i x_i = 0,08*900+0,1*1200+0,16*1500+0,3*\left(\tfrac{1600+y}{2}\right)+0,09*\left(\tfrac{y+2400}{2}\right)+0,27*3200\\ &= 1644+0,195y\end{aligned}$$

et on trouve : $y = 1800$

b- Le budget médian est égal à 1920 euros (1,5 pts).

Il faut raisonner par interpolation linéaire sur l'intervalle [1600-y[. On pose le rapport des distances suivant :

$$\frac{1920-1600}{y-1600} = \frac{0,5-0,34}{0,64-0,34}$$

et on trouve : $y = 2200$

PARTIE 2

Considérons maintenant que la borne manquante y est égale à 2000 euros.

3/ Donner une représentation graphique de la distribution des budgets " vacances " (1,5 pts).

Pour donner une représentation graphique correcte de la distribution (histogramme), il faut au préalable corriger les fréquences puisque l'amplitude est différente selon les intervalles.

Budget	F	f	A	f'
	0			
[800-1000[		0,08	200	0,08
	0,08			
[1000-1400[		0,1	400	0,05
	0,18			
[1400-1600[		0,16	200	0,16
	0,34			
[1600-y[		0,3	400	0,15
	0,64			
[y-2400[		0,09	400	0,045
	0,73			
[2400-x[		0,27	1600	0,03
	1			

4/Calculer et interpréter le budget moyen et médian. Que pouvez-vous conclure de la comparaison entre ces deux valeurs (1,5 pts)?

Pour calculer le budget moyen, il fallait à nouveau utiliser les fréquences :

$$\begin{aligned}\bar{x} &= \sum_i f_i x_i \\ &= 0,08*900 + 0,1*1200 + 0,16*1500 + 0,3*1800 + 0,09*2200 + 0,27*3200 \\ &= 2034\end{aligned}$$

Pour calculer le budget médian, il faut à nouveau raisonner par interpolation linéaire sur l'intervalle [1600-2000[. On pose le rapport des distances suivant :

$$\frac{m-1600}{2000-1600} = \frac{0,5-0,34}{0,64-0,34}$$

et on trouve : $M = 1813$

5/ Sachant que $\sum_i n_i x_i^2 = 4741200000$ et V(x)=604044, retrouver les effectifs correspondant à chacune des tranches de budgets ainsi que l'effectif total des ménages enquêtés (2 pts).

La variance est égale à 604 044 :

$$V(x) = 604044 = \frac{1}{N}\sum_i n_i x_i^2 - \bar{x}^2$$

On sait que $\sum_i n_i x_i^2 = 4741200000$.

Ainsi :

$$V(x) = 604044 = \frac{1}{N}4741200000 - 2034^2$$

On trouve que N=1000. Cela nous permet de calculer les effectifs :

Budget	**n**
[800-1000[	80
[1000-1400[	100
[1400-1600[	160
[1600-y[	300
[y-2400[	90
[2400-x[	270

EXERCICE 2 (11,5 pts)

La répartition des salaries mensuels d'une entreprise est donnée par le tableau suivant :

Salaire	**nb. de salariés**
[1000-1400[	100
[1400-1800[	150
[1800-2200[	40
[2200-3000[	10

1 / Calculer le salaire moyen ainsi que l'écart-type de cette distribution (1,5pts). Commenter.

Salaire moyen :

$$\bar{x} = \frac{1}{N}\sum_i n_i x_i = \frac{1}{30}(100 * 1200 + 150 * 1600 + 40 * 2000 + 10 * 2600) = 1553,33$$

Ecart-type : racine de la variance.

$$V(x) = \frac{1}{N}\sum_i n_i x_i^2 - \bar{x}^2 = \frac{1}{30}(100 * 1200^2 + 150 * 1600^2 + 40 * 2000^2 + 10 * 2600^2) - 1553,33^2 = 105832,58$$

L'écart-type est égal à 325,3.

2/ Tracer la boîte à moustache de cette série. Commenter. (2 pt)

Pour tracer la boîte à moustache, il faut au préalable calculer les valeurs suivantes par interpolation linéaire :
Min=1000
Max=3000
Q1=1303
Q3=1736
D1=1121
D9=2015
Mé=1536

3/ Calculer les fréquences et les fréquences cumulées de la distribution de la masse salariale (masse salariale détenue par chaque catégorie d'individus) (1,5 pts).

Salaire	**nb. de salariés**	$n_i * x_i$	**f**$(n_i x_i)$	F$(n_i x_i)$
				0
[1000-1400[	100	120000	0,26	
				0,26
[1400-1800[	150	240000	0,52	
				0,78
[1800-2200[	40	80000	0,17	
				0,95
[2200-3000[	10	26000	0,06	
	300	**466000**		1

4/ A partir du tableau précédent, calculer la médiale. Que vous indique la comparaison entre la médiane et la médiale (1,5 pts)?

La médiale est la médiane de la nouvelle série calculée. Comme la médiane, elle se calcule par interpolation linéaire. Ici, la médiale est égale à 1588.

Le salaire médian est égal à 1536, cela signifie que 50% des salariés gagnent moins de 1536. La médiale implique que 50% de la masse salariale est versée aux salariés gagnant moins que 1588. Ainsi, 50% des salariés gagnent moins de 50% de la masse salariale.

5/ Tracez la courbe de Lorenz. Que remarquez-vous ? (1,5 pt)

Pour tracer la courbe de Lorenz il faut placer en abcisse la fréquence cumulée de la série "classique" et en ordonnée la fréquence cumulée de la la série $n_i x_i$.

Lorsque l'on trace cette courbe, le centre d'intérêt est la distance entre la première bissectrice et cette courbe. Ainsi, plus l'aire comprise entre les deux est importante, plus il y a des inégalités (ou plus la concentration est importante).

6/ Après avoir rappelé ce que mesure l'indice de Gini et comment il se calcule, calculer et interpréter l'indice de Gini. (2,5 pts)

L'indice de Gini mesure le pourcentage d'inégalités réalisées sur les 100% possibles. ***(Se rapporter au cours pour la construction de l'indice)***

Salaire	**nb. de salariés**	**F(xi)**	F(nixi)	Trapèzes=(b+B)*h/2
		0	0	
[1000-1400[	100			**0,0429**
		0,33	0,26	
[1400-1800[	150			**0,2575**
		0,83	0,78	
[1800-2200[	40			**0,11115**
		0,96	0,95	
[2200-3000[	10			**0,0388**
		1	1	

Ici, la somme des trapèzes est égale à 0,45035, ainsi, l'aire de concentration est égale à 0,05 et l'indice de Gini est égal à 0,1.

7/ Supposons que, sur une période de 5 ans, le salaire moyen augmente de 20% :

a - Quelle est l'évolution annuelle moyenne du salaire moyen ? (0,5pt)

Pour calculer l'évolution annuelle moyenne : si le salaire a augmenté de 20%, cela implique que $\bar{S}_5 = 1,2\bar{S}_0$. Un taux identique appliqué pendant 5 ans veut dire $\bar{S}_5 = (1+\bar{g})^5\bar{S}_0$.
Ainsi :

$$\bar{g} = (1,2)^{(1/5)} - 1 = 0,037$$

Le taux de croissance annuel moyen est de 3,7%.

b - Au bout de combien d'années le salaire moyen aura-t-il doublé ? (0,5pt)

On cherche t tel que $S_t = 2 * S_0$ et on sait que $S_t = (1+\bar{g})^t * S_0$.
Ainsi, $(1,037)^t = 2 \Leftrightarrow t \ln 1,037 = \ln 2$. Après calcul, on trouve $t = 19$. Il faudra donc 19 ans à un taux de croissance annuel moyen de 3,7% pour doubler le salaire moyen.