L2 AES

Corrigé contrôle continu de statistique : SUJET B

EXERCICE-1

Xi	n _i	n _{iCd}	f _{iCd}	nx
102	65	335	1	6630
103	80	270	0,806	8240
104	56	190	0,567	5824
105	48	134	0,4	5040
106	34	86	0,257	3604
107	23	52	0,155	2461
108	17	29	0,087	1836
109	7	12	0,036	763
110	3	5	0,015	330
111	2	2	0,006	222
	335			34950

- 1. Il s'agit d'un caractère quantitatif discret.
- 2. Les fréquences cumulées décroissantes sont dans le tableau. On lit que 15.5% des individus ont plus de 107 ans donc 100 15.5 =84.5% ont moins de 107ans. On lit directement que 25.70% des individus ont un âge supérieur ou égal à 106 ans donc plus que 105
- 3. n = 335 ; c'est un nombre impair, il y a donc une médiane qui est l'observation de rang 168, car $\frac{335}{2} = 167.5$; les effectifs cumulés décroissants nous montrent que la médiane est 104 ans ; il y a 50% des individus qui ont moins de 104 ans.

4.
$$\overline{x} = \frac{\sum n_i x_i}{n} = \frac{34950}{335} = 104.33$$
 ans.

II EXERCICE-2

1. Les classes étant d'amplitudes inégales, on utilise la densité, $d_i = \frac{n_i}{A_i}$ et les effectifs corrigés $n_{icor} = 5d_i$, 5 étant l'amplitude minimale de classe

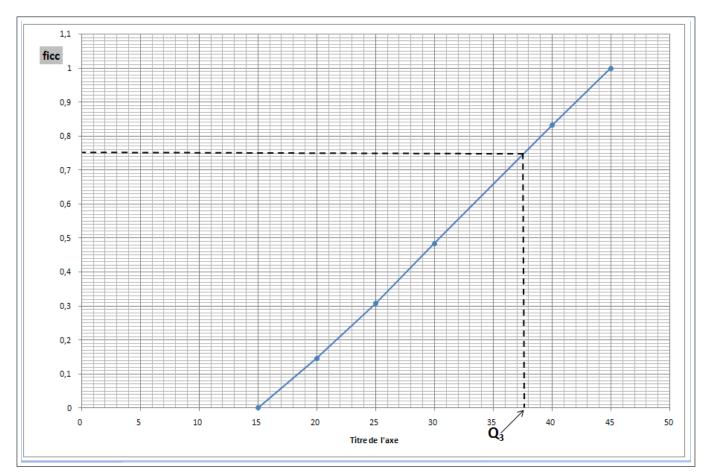
a _i	b _i	n _i	Xi	Ai	d _i	n _i cor	n _i cc	f _i cc	n _i xi
15	20	748 828	17,5	5	149765,6	748828	748 828	0,1463	13104490
20	25	822 939	22,5	5	164587,8	822939	1 571 767	0,3071	18516127,5
25	30	906 967	27,5	5	181393,4	906967	2 478 734	0,4843	24941592,5
30	40	1 780 325	35	10	178032,5	890162,5	4 259 059	0,8322	62311375
40	45	858 693	42,5	5	171738,6	858693	5 117 752	1,0000	36494452,5
		5 117 752							155368038

La classe modale est celle de plus grande densité, c'est-à dire la classe [25; 30] et le mode est calculé en considérant les classes encadrant la classe modale, ce qui donne avec les notations du cours : $\left\{ \begin{array}{l} x_1 = 25 \\ x_2 = 30 \end{array} \right., \left\{ \begin{array}{l} h = 906967 \\ h_1 = 822939 \text{ et } h_2 = 890162.5 \end{array} \right.$

$$\begin{cases} k_1 = h - h_1 = 906967 - 822939 = 84028 \\ k_2 = h - h_2 = 906967 - 890162.5 = 16805 \end{cases} : : : et pour conclure :$$

 $\begin{cases} k_1 = h - h_1 = 906967 - 822939 = 84028 \\ k_2 = h - h_2 = 906967 - 890162.5 = 16805 \end{cases} : \dots \text{ et pour conclure} : \\ M_o = \frac{k_2 x_1 + k_1 x_2}{k_2 + k_1} = \frac{16805 * 25 + 84028 * 30}{16805 + 84028} = 29.17 \text{ ; comme prévu, le mode est très proche de } 30, \text{ car il est attiré par la classe de droite, de densité plus importante.} \end{cases}$

2. Le polygone des fréquences cumulées décroissantes



Ce graphique permet d'estimer le troisième quartile à environ 37.6, en prenant l'intersection du polygone des fréquences cumulées décroissantes avec la droite horizontale : y = 0.75.

3. Calcul de Q_3 : on localise la médiane dans la classe [30; 40] (la fréquence cumulée passe le seuil des 75%), puis on effectue une interpolation linéaire:

 $\frac{0.8322-0.4843}{40-30} = \frac{0.75-0.4843}{Q_3-30} \text{ soit } Q_3 - 30 = 10 \\ \frac{0.75-0.4843}{0.8322-0.4843} \text{ soit } Q_3 = 30 + 10 \\ \frac{0.75-0.4843}{0.8322-0.4843} = 37.64 \\ \text{; Il y a donc } 75\% \text{ de la population d'île de France d'âge compris entre } 15 \text{ et } 45 \text{ ans qui avait moins de } 37.64 \text{ ans en } 2008.$

4. La moyenne est donnée par : $\overline{x} = \frac{1}{n} \sum n_i x_i = \frac{155368038}{5117752} = 30.36$; $V(x) = \frac{1}{n} \sum n_i x_i^2 - \overline{x}^2 = 8.23^2 = 67.80$ et $\sigma(x) = \sqrt{V(x)} \simeq 8.23$

5. Pour les intervalles ne correspondant pas à une classe entière, on multiplie la densité correspondante par l'amplitude de l'intervalle, d'après la formule : $n_i = A_i * d_i$.

 $\overline{x} + 1.5\sigma(x) = 30.36 + 8.23 = 38.59$

 $\overline{x} - 1.5\sigma(x) = 30.36 - 8.23 = 22.13$; on doit donc estimer les effectifs correspondant aux intervalles : [22.13; 25[, [25; 30[, [30; 38.59]]]])

Pour les intervalles ne correspondant pas à une classe entière, on multiplie la densité correspondante par l'amplitude de l'intervalle, d'après la formule : $n_i = A_i * d_i$.

2

	Amplitude	densité	Effectif estimé
[22,13;25[2,87	164587,8	472366,99
[25;30[906967,00
[30;38,59[8,59	1529299,18	
Effe	2908633,16		

ce qui donne un nombre total de : 2908633 individus.

III EXERCICE-3(5 pts)

L2 AES Corrigé contrôle continu de statistique : SUJET B

- 1. On se ramène à la loi normale centrée réduite en standardisant : on pose $Z=\frac{X-90}{15}$, on sait alors que Z suit la loi normale centrée réduite : $E\left(Z\right)=0$ et $\sigma\left(Z\right)=1$; on peut alors utiliser la table de la fonction de répartition de la loi $\mathcal{N}\left(0;1\right)$, et en utilisant la fonction de répartition de la loi $\mathcal{N}\left(0;1\right)$, ce qui donne : $P\left(X<4.25\right)=P\left(Z<\frac{4.25-4}{0.16}\right)=P\left(Z<1.56\right)\simeq \boxed{0.9406}$.
- 2. $P(3.80 \le X \le 4.20) = P(\frac{3.8-4}{0.16} \le Z \le \frac{4.20-4}{0.16}) = 2F(1.25) 1 \simeq 2 * 0.89435 1 = 0.7887$
- 3. $P(X \ge 3.80) = 1 P(X \le 3.80) = 1 P(Z < \frac{3.8 4}{0.16}) = 1 F(1.25) = 1 0.8944 = 0.1056$