Corrige contrôle continu A

L2-AES-Statistique

Avril 2012

1 EXERCICE-1

x_i	n_i	n_{iCC}
0	25	25
1	53	78
2	58	136
3	37	173
4	19	192
5	5	197
6	2	199
7	1	200

1. Il s'agit d'un caractère quantitatif discret. Le mode est 2 car c'est la modalité de pus grand effectif.

2. L'effectif total est 200, il est pair, on calcule $\frac{n}{2} = 100$, et on détermine l'intervalle médian constitué par les deux termes centraux, c'est à dire de rangs respectifs 100 et 101 ; on utilise les effectifs cumulés croissants.qui nous indiquent que les deux termes centraux sont égaux à 2 ; la médiane étant leur moyenne arithmétique, on a : $M_e = \frac{2+2}{2} = 2$; 50% des familles ont eu 2 ou moins de 2 accidents.

2 EXERCICE-2

a_i	b_i	x_i	n_i	n_ix_i	$n_ix_i^2$
5	10	7.5	7	52.5	393.75
10	15	12.5	58	725	9062.5
15	25	20	242	4840	96800
25	30	27.5	181	4977.5	136881.3
30	35	32.5	12	390	12675
			500	10985	255812.5

1. Le tableau ci-dessus permet de calculer la moyenne : $\overline{x} = \dfrac{1}{n}\sum n_ix_i = \frac{10985}{500} = 21.97$; $V(x) = \dfrac{1}{n}\sum n_ix_i^2 - \overline{x}^2 = \frac{255812.5}{500} - 21.97^2 = 28.94$ et $\sigma(x) = \sqrt{28.94} = 5.38$.

2. $y = 0.82x$,ce qui donne avec les formules du cours : si $y = ax + b$, $\overline{y} = a\overline{x} + b$, $V(y) = a^2 V(x)$ et $\sigma(y) = |a|\,\sigma(x)$,soit ici : $\overline{y} = 0.82\overline{x} = 0.82 * 21.97 = 18.02\ mn$ et $\sigma(y) = 0.82\sigma(x) = 0.82 * \sqrt{28.94} = 4.41mn$.

3. Pour comparer la dispersion des caractères x et y, on calcule : $CV(y) = \dfrac{\sigma(y)}{\overline{y}} = \dfrac{0.82\sigma(x)}{0.82\overline{x}} = \dfrac{\sigma(x)}{\overline{x}} = CV(x) = \dfrac{\sqrt{28.94}}{21.97} = 0.244\,9$,soit 424.49%.Les caractères x et y ont la même dispersion.

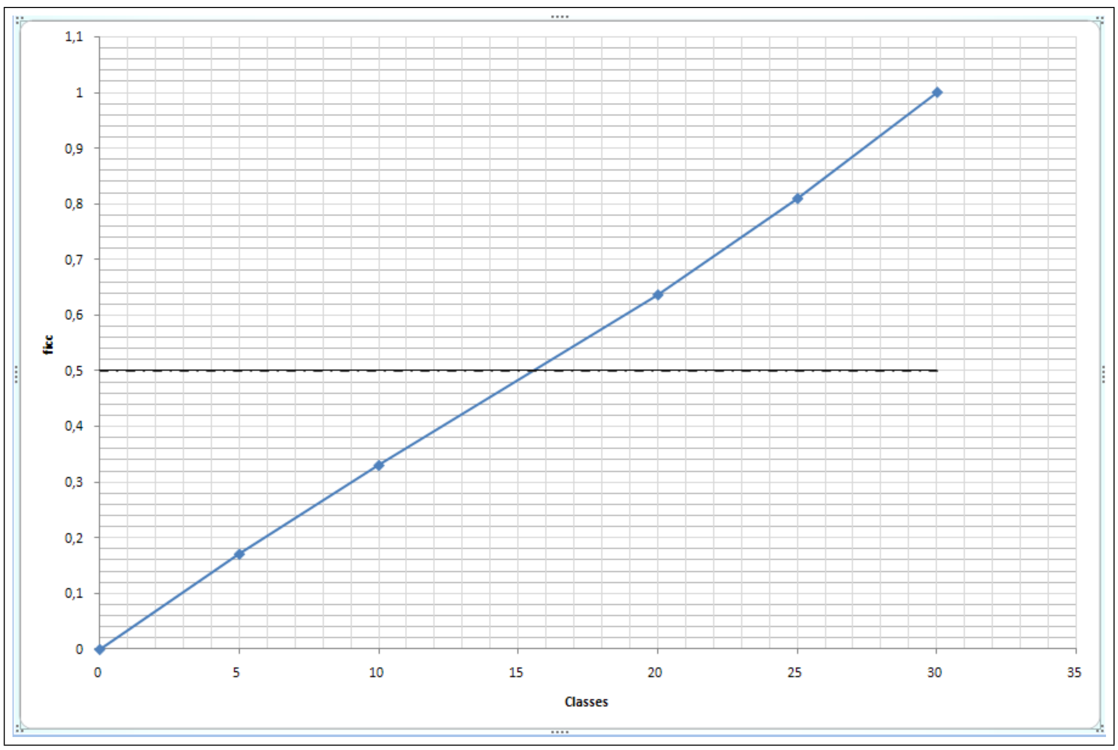
3 EXERCICE-3

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

a_i	b_i	n_i	x_i	A_i	d_i	$n_i$cor	$n_i$cc	$f_i$cc	nx	nx²
0	5	813 754	2,5	5	162750,8	813754	813754	0,1711	2034385	5085962,5
5	10	757 668	7,5	5	151533,6	757668	1571422	0,3303	5682510	42618825
10	20	1 456 075	15	10	145607,5	728037,5	3027497	0,6364	21841125	327616875
20	25	822 939	22,5	5	164587,8	822939	3850436	0,8094	18516128	416612869
25	30	906 967	27,5	5	181393,40	906967	4757403	1,0000	24941593	685893794
		4 757 403							73015740	1477828325

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 : $\begin{cases} x_1 = 25 \\ x_2 = 30 \end{cases}$, $\begin{cases} h = 906967 \\ h_1 = 822939 \text{ et } h_2 = 0 \end{cases}$ $\begin{cases} k_1 = h - h_1 = 906967 - 822939 = 84028 \\ k_2 = h - h_2 = 906967 \end{cases}$: . et pour conclure :
$M_o = \dfrac{k_2 x_1 + k_1 x_2}{k_2 + k_1} = \dfrac{25 * 906967 + 84028 * 30}{906967 + 84028} = 25.42$; comme prévu, le mode est très proche de 25, car il est attiré par la classe de gauche, de densité plus importante. L'âge le plus fréquent est estimé à 25.42 ans.

2. Le polygone des fréquences cumulées croissantes



Ce graphique permet d'estimer la médiane à environ 15.5, en prenant l'intersection du polygone des effectifs cumulés croissants avec la droite horizontale : $y = 0.5$.

3. Calcul de Q_1 : on localise Q_1 dans la classe $[5;10[$ (la fréquence cumulée passe le seuil des 25%), puis on effectue une interpolation linéaire :
$\frac{0.3303-0.1711}{10-5} = \frac{0.25-0.1711}{Q_1-5}$ soit $Q_1 - 5 = 5\frac{0.25-0.1711}{0.3303-0.1711}$ soit $Q_1 = 5 + 5\frac{0.25-0.1711}{0.3303-0.1711} = 7.48$. Il y a donc 25% de la

population d'île de France de moins de 30 ans qui avait moins de 7.48 ans en 2008.

4. La moyenne est donnée par : $\overline{x} = \dfrac{1}{n}\sum n_i x_i = 15.35$; $V(x) = \dfrac{1}{n}\sum n_i x_i^2 - \overline{x}^2 = 75.08$ et $\sigma(x) = \sqrt{V(x)} \simeq 8.66$
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 : $n_i = A_i * d_i$.

	Amplitude	densité	Effectif estimé
[9;10[	1	151533.6	151533.6
[10;20[			1 456 075
			1607608.6

ce qui donne une proportion de : $\frac{1607608.6}{4757403} = 0.337\,9$, soit 33.79%.

4 EXERCICE-4

a_i	b_i	n_i	x_i	f_i	$f_i cc$	$n_i xi$	q_i	$q_i cc$
0	2	222279	1	0.5870	0.5870	222279	0.1614	0.1614
2	5	142380	3.5	0.3760	0.9630	498330	0.3617	0.5231
5	30	8331	17.5	0.0220	0.9850	145792.5	0.1058	0.6289
30	150	5680	90	0.0150	1.0000	511200	0.3711	1.0000
		378670				1377601.5	1