Corrigé contrôle continu de statistique : sujet B

I EXERCICE-1

x_i	n_i	n_{iCC}
0	25	25
1	54	79
2	23	102
3	67	169
4	19	188
5	14	202
6	2	204
7	1	205

L'effectif total est 205, il est impair, on calcule $\frac{n}{2} = 102.5$, la médiane correspond au terme de rang 103 ; on utilise les effectifs cumulés croissants.qui nous indiquent que ce terme vaut 3 ; 50% des familles ont eu 3 ou moins de 3 accidents.

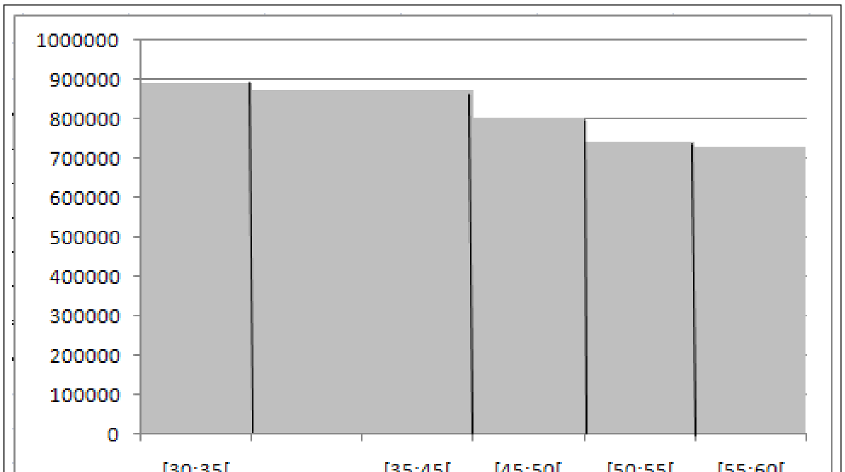
II EXERCICE-2$(3pts)$

1. $y = 0.87x$,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 : $\overline{y} = 0.87\overline{x} = 0.87 * 21.97 = 19.11mn$ et $\sigma(y) = 0.87\sigma(x) = 0.87 * 5.38 = 4.68$
2. Pour comparer la dispersion des caractères x et y, on calcule : $CV(y) = \dfrac{\sigma(y)}{\overline{y}} = \dfrac{0.87\sigma(x)}{0.87\overline{x}} = \dfrac{\sigma(x)}{\overline{x}} = CV(x) = \dfrac{5.38}{21.97} = 0.244\,9$ soit 24.49%.Les caractères $x$et y ont la même dispersion.

III EXERCICE-3

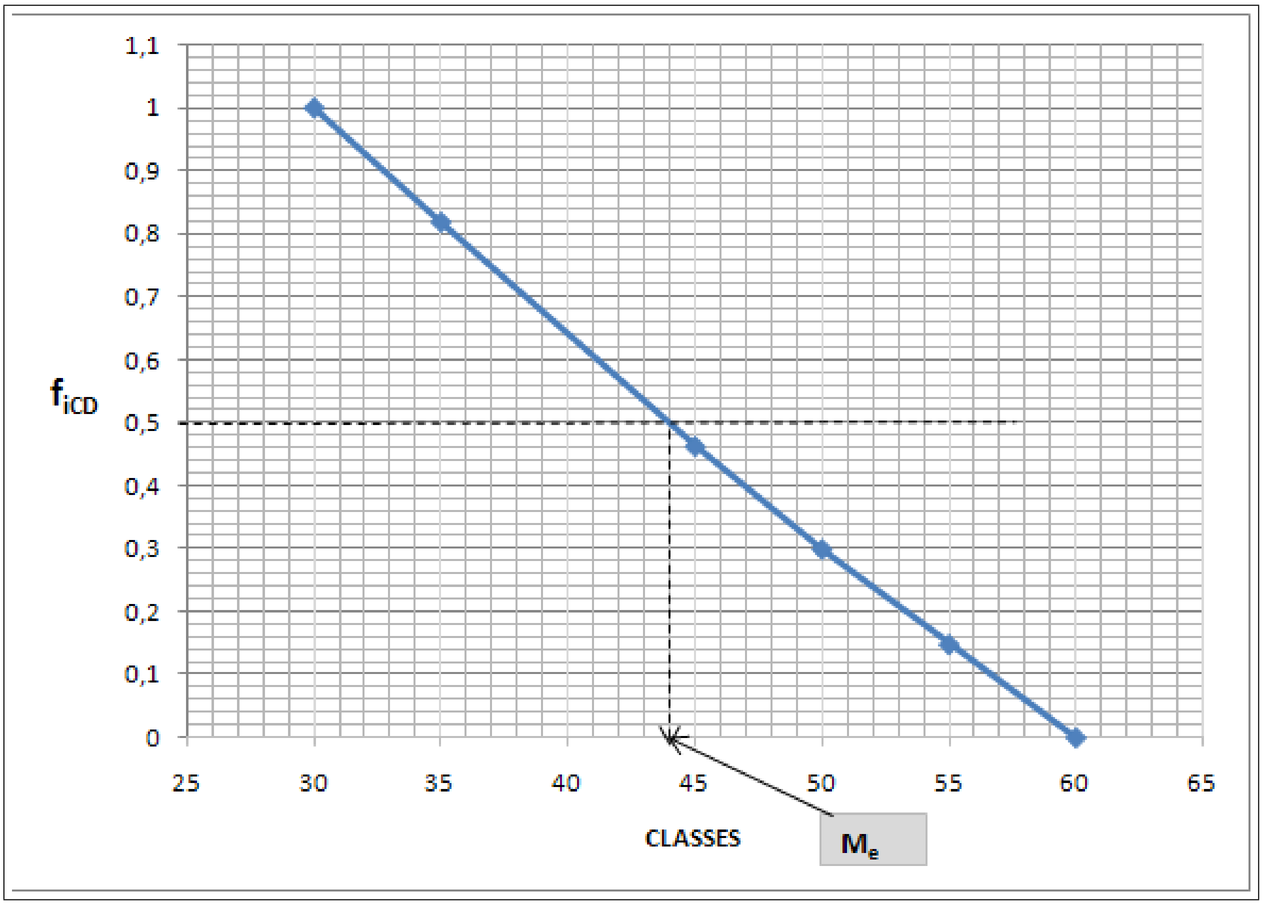
1. La représentation de cette série représentant les données relatives à un caractère quantitatif continu est un histogramme ; les classes étant d'amplitudes inégales, on utilise la densité, $d_i = \dfrac{n_i}{A_i}$ et les effectifs corrigés $n_{icor} = 5d_i$, 5 étant l'amplitude minimale de classe.

a_i	b_i	n_i	x_i	A_i	d_i	$n_i$cor	$n_i$cd	$f_i$cd	nx
30	35	890 605	32,5	5	178121	890605	4 909 873	1,0000	28944662,5
35	45	1 748 413	40	10	174841,3	874207	4 019 268	0,8186	69936520
45	50	801 683	47,5	5	160336,6	801683	2 270 855	0,4625	38079942,5
50	55	741 702	52,5	5	148340,4	741702	1 469 172	0,2992	38939355
55	60	727 470	57,5	5	145494,00	727470	727 470	0,1482	41829525
		4 909 873							217730005



2. La classe modale est celle de plus grande densité, c'est-à dire la classe $[30;35[$ et le mode est calculé en considérant les classes encadrant la classe modale, ce qui donne avec les notations du cours : $\begin{cases} x_1 = 30 \\ x_2 = 35 \end{cases}$, $\begin{cases} h = 890605 \\ h_1 = 0 \text{ et } h_2 = 874206.5 \end{cases}$
$\begin{cases} k_1 = h - h_1 = 890605 \\ k_2 = h - h_2 = 890605 - 874206.5 = 16399 \end{cases}$ et pour conclure :
$M_o = \dfrac{k_2 x_1 + k_1 x_2}{k_2 + k_1} = \dfrac{16399 * 30 + 890605 * 35}{16399 + 890605} = 34.91$; comme prévu, le mode est très proche de 35, car il est attiré par la classe de droite, de densité plus importante.

3. Le polygone des fréquences cumulées décroissantes



Ce graphique permet d'estimer la médiane à environ 44, en prenant l'intersection du polygone des fréquences cumulées décroissantes avec la droite horizontale : $y = 0.5$.

4. Calcul de Q_3 : on localise Q_3 dans la classe $[50;55[$ (la fréquence cumulée passe le seuil des 75%), puis on effectue une interpolation linéaire :
$\frac{0.8518-0.7008}{55-50} = \frac{0.75-0.7018}{Q_3-50}$ soit $Q_3 - 50 = 5\frac{0.75-0.7018}{0.8518-0.7008}$ soit $Q_3 = 50 + 5\frac{0.75-0.7018}{0.8518-0.7008} = 51.60$. Il y a donc 75% de la population d'île de France d'âge compris entre 30 et 60 ans qui avait moins de 51.60 ans en 2008.

5. La moyenne est donnée par : $\overline{x} = \dfrac{1}{n}\sum n_i x_i = 44.35$; $V(x) = \dfrac{1}{n}\sum n_i x_i^2 - \overline{x}^2 = 8.34^2 = 69.56$ et $\sigma(x) = \sqrt{V(x)} \simeq 8.34$

6. 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
[36;45[	9	174841,3	1573571,7
[45;50[			801 683
			2375254,7

, donc une estimation de 2375254.7.