

CORRIGE PARTIEL BLANC DE STATISTIQUE

AES-L2STAT.

AVRIL2013

1 EXERCICE-1

C.A.	n_i	x_i	A_i	d_i	n_{icc}	$n_i x_i$	$n_i x_i^2$
[0;2.5[51	1,25	2,5	20,4	51	63,75	79,6875
[2.5;5[360	3,75	2,5	144	411	1350	5062,5
[5;10[482	7,5	5	96,4	893	3615	27112,5
[10;15[154	12,5	5	30,8	1047	1925	24062,5
[15;20[47	17,5	5	9,4	1094	822,5	14393,75
[20;25[28	22,5	5	5,6	1122	630	14175
	1122					8406,25	84885,94

- La classe modale est la classe de plus grande densité, la classe $[2, 5; 5[$.
- On calcule : $0.25 * 1122 = 280.5$; d'après les effectifs cumulés croissants, on voit que Q_1 est dans la classe $[2.5 ; 5 [$; on le détermine par interpolation linéaire : $\frac{411-51}{5-2.5} = \frac{280.5-51}{Q_1-2.5}$, soit : $Q_1 = \frac{280.5-51}{411-51} * 2.5 + 2.5 = 4.09$. 25% des entreprises ont un chiffre d'affaires inférieur ou égal à 4.09 millions d'euros.
- On trouve : $\bar{x} = \frac{\sum n_i x_i}{N} = \frac{8406.25}{1122} \simeq 7.49$ et $V(x) = \frac{1}{N} \sum n_i x_i^2 - (\bar{x})^2 \simeq \frac{84885.94}{1122} - \left(\frac{8406.25}{1122}\right)^2 \simeq 19.52$ et l'écart-type $\sigma(x) = \sqrt{V(x)} \sqrt{19.52} = 4.42$

2 EXERCICE-2

	X	Y
Moyennes	8,2095	13,4598
Variances	76,5818	263,2720
Ecart-types	8,7511	16,2257

1.

	x	y	xy
Argentine	0,98	1	0,98
Australie	1,18	1,98	2,34
Brésil	1,42	2,19	3,11
Royaume Uni	0,78	0,70	0,55
Canada	1,31	1,56	2,04
Chine	3,9	8,28	32,29
Rép.tchèque	22	39	858,00
Danemark	9,74	8,46	82,40
Zone euro	1,01	1,14	1,15
France	7,28	7,44	54,16
Allemagne	2,01	2,22	4,46
Hong Kong	4,21	7,81	32,88
Philippines	23,2	50,3	1166,96
Pologne	2,32	4,03	9,35
Russie	13,8	28,9	398,82
Afrique du Sud	3,82	8,13	31,06
Suède	9,45	10,28	97,15
Suisse	2,48	1,73	4,29
Taiwan	27,6	32,9	908,04
Thaïlande	21,7	45,5	987,35
			4677,37

$$Cov(x; y) = \frac{1}{n} \sum y_i x_i - \bar{x} \bar{y} = \frac{4677.37}{20} - 8.21 * 13.46 =$$

2. Covariance
123.36

3. La droite de régression de Y en X, $D_{Y/X}$, est : $y = 1.7491x - 0.8995$

4. Le coefficient de corrélation linéaire est : $r \approx 0,9433$. r est proche de 1 donc un ajustement affine est justifié.

5. Le cours du dollar en pesos impliqué par la PPA est : $x = \frac{21.9}{2.54}$, soit : $x \approx 8.62$

donc une estimation du taux de change du dollar en pesos est : $\hat{Y} = 1.7491 * 8.62 - 0.8995 = 14.18$.

6. Le coefficient de détermination est $R^2 \approx : R^2 = \frac{SCE}{SCT} = 0.8899$

R^2 donne le pourcentage de variation expliquée par le modèle. Ici, on explique 88.99% de la variation par le modèle.

7. Analyse de la variance

L'équation de l'analyse de la variance s'écrit : $SCT = SCE + SCR$, $SCT = nV(y) = 19 * 263.2720 = 5002.168$ et $SCE = R^2 SCT = 0.8899 * 5002.168 = 4451.429$ donc $SCR = SCT - SCE = 5002.168 - 4451.429 = 550.739$

3 EXERCICE-3

SEXE	Y	AGE				
	X	[15;25[[25;30[[30;50[n_{i+}	f_{i+}
	H	710	110	180	1000	0,7042
	F	240	60	120	420	0,2958
	n_{+j}	950	170	300	1420	
	f_{+j}	0,6690	0,1197	0,2113		1

1.

2. Cette fréquence partielle est donnée par : $f_{11} = \frac{710}{1420} = 0.5$
3. Il s'agit d'une fréquence conditionnelle : $f_{(i=3/j=2)} = \frac{120}{300} = 0.4$

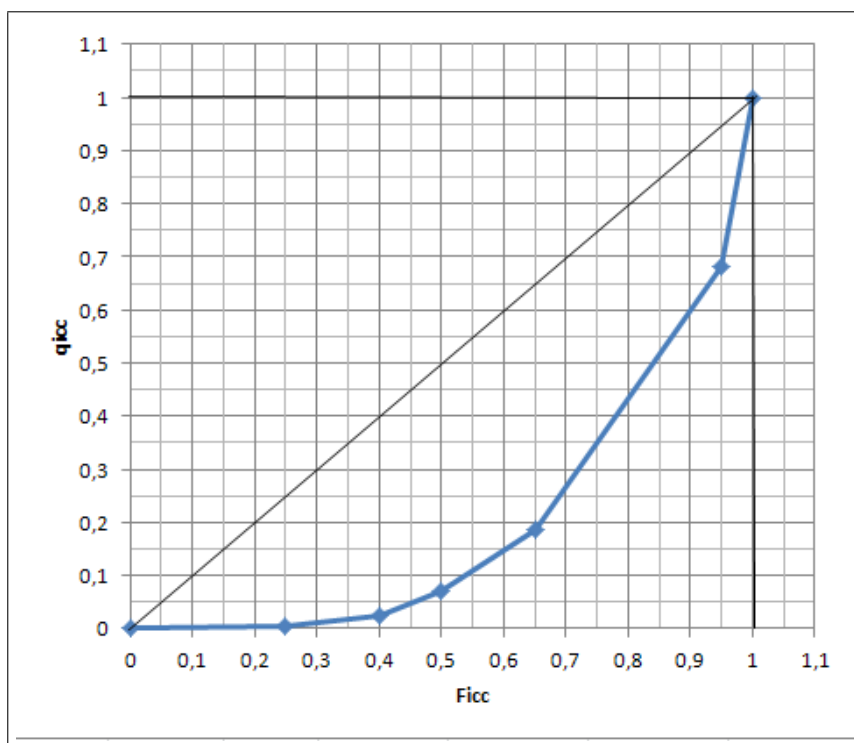
Y_j	f_{+j}	$f_{+j} * Y_j$
20	0,6690	13,38
27,5	0,1197	3,29
40	0,2113	8,45
		25,12

4.

4 EXERCICE-4

f_{icc}	q_{icc}	f_i	Aires
0,250	0,003	0,2500	0,0004
0,400	0,025	0,1500	0,0021
0,500	0,072	0,1000	0,0049
0,650	0,187	0,1500	0,0194
0,950	0,68	0,3000	0,1301
1,000	1	0,0500	0,0420
			0,1988

1.



2.

On obtient alors l'indice de Gini : $2 * \text{Aire de concentration}$ soit $I_G = 2 * (0.5 - 0.1988) = 0.6024$

Le coefficient de Gini est plus proche de 1 que de 0, la concentration élevée.