Université de Carthage
Ecole Supérieure de la Statistique
et de l'Analyse de l'Information à Tunis

Année Universitaire
2014-2015
Première Année

Statistique Descriptive : Corrigé du Devoir surveillé octobre 2014

Enseignante : Mme Héla Ouaili Mallek   Durée :1h30

Corrigé de l'exercice 1 *:*

1. $\mathcal{T} = \prod_{i=1}^{4} (1+\tau_i) - 1.$

2. $\tau = \sqrt[4]{1+\mathcal{T}} - 1 = \sqrt[4]{\prod_{i=1}^{4} (1+\tau_i)} - 1$

3. $\tau_{34} = \sqrt{(1+\tau_3)(1+\tau_4)} - 1$

   $\mathcal{T} = 0 \iff \sqrt[4]{\prod_{i=1}^{4} (1+\tau_i)} = 1 \iff \prod_{i=1}^{4} (1+\tau_i) = 1 \iff (1+\tau_{34})^2 = \dfrac{1}{(1+\tau_1)(1+\tau_2)}$

   $\iff \tau_{34} = \frac{1}{\sqrt{(1+\tau_1)(1+\tau_2)}} - 1$

Corrigé de l'exercice 2 *:*

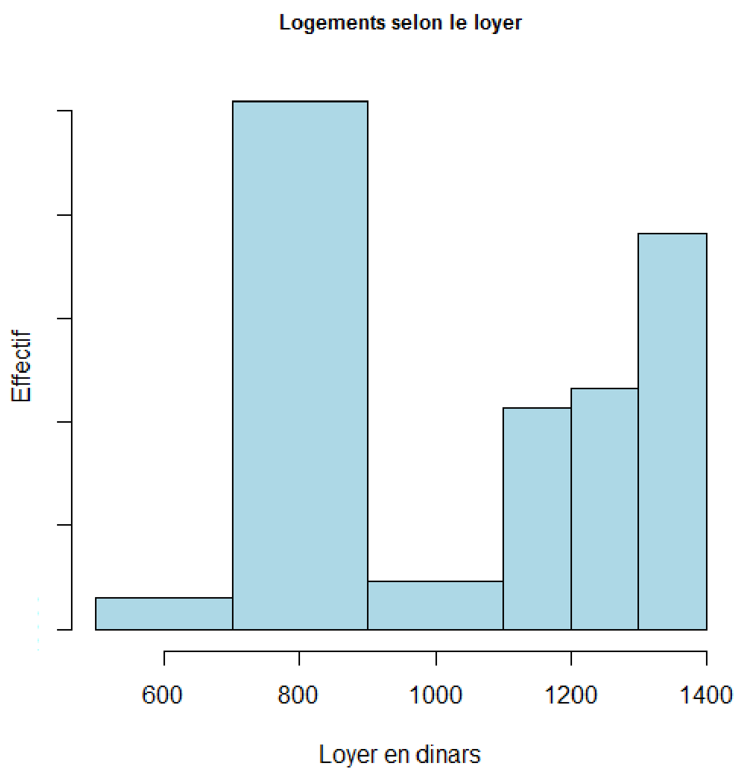
Tableau des calculs préliminaires :

$\sum n_i c_i = 4291 \times 10^2 \quad et \sum n_i c_i^2 = 45325 \times 10^4$

1. *Amplitude de référence :* 100.

Loyer	$[500;700[$	$[700;900[$	$[900;1100[$	$[1100;1200[$	$[1200;1300[$	$[1300;1400[$
Effectif	13	219	20	46	50	82
f_i	0.030	0.509	0.047	0.107	0.116	0.191
$F(e_i)$	0.030	0.539	0.586	0.693	0.809	1
n_i^C	6.5	109.5	10	46	50	82

Histogramme :



2. *La distribution présente 2 modes. On a donc une première sous-population de logements à loyer "bas" et une seconde à loyers élevés.*

3. *Pour la classe* $[700; 900[$, *malgré une très forte concentration (effectif moyen par unité d'intervalle égal à* $\frac{219}{200} = 1.095$*) alors que la concentration la plus élevée parmi les classes restantes de même amplitude est largement en dessous (effectif moyen par unité d'intervalle égal à* $\frac{20}{200} = 0.1$*), il a été décidé de prendre une amplitude de classe égale à* 200*. Il aurait été plus judicieux de découper la classe en deux classes d'amplitude* 100*, permettant ainsi de préserver l'information sans pour autant perdre en termes de synthèse.*

4. $P\left[X \leq 550\right] = P\left[X < 550\right] = F\left(550\right).$

   $$F\left(550\right) = F\left(500\right) + \left(F\left(700\right) - F\left(500\right)\right) * \frac{550-500}{700-500}$$

   $$F\left(550\right) = 0.030 * \frac{550-500}{700-500} = 0.0075$$

   Seulement 0.75% *des logements sont loués à moins de* 550 *dinars.*

   $P\left[X \geq 1150\right] = 1 - F\left(1150\right)$

   $$F\left(1150\right) = F\left(1100\right) + \left(F\left(1200\right) - F\left(1100\right)\right) * \frac{1150-1100}{1200-1100}$$

   $$F\left(1150\right) = 0.586 + \left(0.693 - 0.586\right) * \frac{1150-1100}{1200-1100} = 0.6395$$

$P\,[X \geq 1150] = 0.360$

Ainsi 36% *des logements ont un loyer supérieur à* 1150 *dinars.*

5. *La classe modale est* $[700; 900[$.

   $$M_o = 700 + 200 * \frac{109.5 - 6.5}{(109.5 - 6.5) + (109.5 - 10)} = 801.73 \ \textit{dinars}$$

   La classe médiane est $[700; 900[$.

   $$M_e = 700 + 200 * \frac{0.5 - 0.030}{0.539 \ - 0.030} = 884.68 \ \textit{dinars.}$$

6. *D'après la relation empirique de Pearson, on a*

   $$\overline{x} \simeq \frac{3M_e - M_o}{2} \Longrightarrow \frac{3 * 884.68 - 801.73}{2} = 926.16 \ \textit{dinars.}$$

7. $\overline{x} = \frac{1}{n} n_i c_i = \frac{1}{430} * 4291 \times 10^2 = 997.91 \ \textit{dinars.}$

   La valeur calculée est très différente de celle approximée car la distribution n'est ni unimodale ni peu asymétrique.

8. $s_X = \sqrt{\frac{1}{n} \sum n_i c_i^2 - \overline{x}^2} = \sqrt{\frac{1}{430} * 45325 \times 10^4 - (997.91)^2} = 241.\,34 \ \textit{dinars.}$