Université Batna 2 **3^{em} Licence SAD**
Département de Mathématiques
Année Universitaire 2020/2021

Exercices corrigés du TD

Exercise 1 (Ad'équation à la loi de poisson) *On souhaite étudier la circulation en un point fixe d'une autoroute en comptant, pendant deux heures, le nombre de voitures passant par minute devant un observateur. Le tableau suivant résume les données obtenues :*

débit en voitures par minute	*Fréquences observées n_i*
00	04
01	09
02	24
03	25
04	22
05	18
06	06
07	05
08	03
09	02
10	01
11	01
Total	120

Tester l'adéquation de la loi empirique observée à une loi théorique simple au seuil $\alpha = 0,10$?

Solution Soit X une variable aléatoire représentant le trafic des voitures par minute. Le nombre moyen $\bar{x}$ de voitures par minute est égal à $3,70$ et la variance empirique $s^2 = 4,37$. Les deux valeurs sont proches à 4.

On peut penser ajuster la loi de X à une loi de Poisson de paramètre $\lambda = 4$.

On obtient les probabilités $p_i = P(X = x)$, $x \geq 0$ en lisant la table de Poisson $P(4)$. Par exemple pour $X = 3$, $p_4 = P(X = 3) = 0,1954$, $np_4 = 120 \times 0,1954 = 23,448$ arrondi à 23 voitures

débite n voitures par minute	Fréquences observées n_i	Fréquences théoriques n_ip_i
00	04	02
01	09	09
02	24	18
03	25	23
04	22	23
05	18	19
06	06	13
7	05	07
8	03	03
9	02	02
10	01	01
11	01	00
Total	120	120

On constate que certains effectifs n_ip_i sont inférieurs à 5. On regroupe les deux premières classes et aussi les quatre dernières classes, on obtient :

débite n voitures par minute	Fréquences observées n_i	Fréquences théoriques n_ip_i
$[00, 02[$	13	11
02	24	18
03	25	23
04	22	23
05	18	19
06	06	13
7	05	07
$[8, 12[$	07	06
Total	120	120

La valeur de

$$\begin{aligned}\chi^2_{calc} &= \frac{(13-11)^2}{11} + \frac{(24-18)^2}{18} + \frac{(25-23)^2}{23} + \frac{(22-23)^2}{23} \\ &\quad + \frac{(18-19)^2}{19} + \frac{(6-13)^2}{13} + \frac{(5-7)^2}{7} + \frac{(7-6)^2}{6} \\ &= 7,141\end{aligned}$$

On compare cette valeur de $7,141$ à la valeur du fractile de la loi de khi-deux à 7 *d.d.l* (car on a 8 classes) d'ordre 0.90, lue dans la table de khi-deux, qui est égal à $12,017$. On est donc dans la région d'acceptation de H_0 et par conséquent on ne rejette pas l'ajustement à la loi de $\mathcal{P}(4)$

Exercise 2 *(Ad'équation à la loi normale) Le satellite Landsat a mesuré la lumière, dans le proche infrarouge, réfléchie par des zones urbanisées. Voici des points de relevées thermographiques*

mesures	71	72	73	74	75	77	78	79	80	81	82	84
effectifs observées	1	1	1	1	1	1	1	4	3	3	4	6
mesures	85	86	87	88	90	91	94					
effectifs observés	6	2	1	1	1	1	1					

On veut tester si cette série peut provenir d'une loi normale.

On a besoin d'estimer deux paramètres, la moyenne (par la moyenne empirique) et l'écart-type (par l'écart-type empirique)

$$\bar{x}_{40} = 82.075 \text{ et } s_{40} = 4.98$$

On dispose d'un échantillon de taille 40. Le nombre de classes sera donc 8. On les choisit de telle sorte que chacune d'elle ait la probabilitè $\frac{1}{8}$, de telle sorte que les efiectifs théoriques soient tous égaux à 5.

La partition pour la loi normale $\mathcal{N}(0,1)$ est obtenue en lisant les tables

$$-1.15 \quad -0.67 \quad -0.32 \quad 0 \quad 0.32 \quad 067 \quad 1.15$$

ce qui donne pour la loi $\mathcal{N}(82.075, 4.98)$

$$76.35 \quad 78.72 \quad 80.49 \quad 82.075 \quad 83.66 \quad 85.43 \quad 87.8$$

On obtient donc le tableau

classe	n_{th}	n_{obs}
$]-\infty, 76.3]$	5	5
$[76.3, 78.7]$	5	2
$[78.7, 80.5]$	5	7
$[80.5, 82.7]$	5	7
$[82.7, 83.7]$	5	0
$[83.7, 85.4]$	5	12
$[85.4, 87.8]$	5	3
$[87.8, +\infty]$	5	4

La distance du chi-deux vaut alors
$\chi^2_{cal} = \frac{(5-5)^2}{5} + \frac{(2-5)^2}{5} + \frac{(7-5)^2}{5} + \frac{(7-5)^2}{5} + \frac{(0-5)^2}{5} + \frac{(12-5)^2}{5} + \frac{(3-5)^2}{5} + \frac{(4-5)^2}{5} = 19.2$

On a une valeur observée ici beaucoup plus grande, donc on rejette H_0 : on décide que la loi observée n'est pas une loi normale.

Exercise 3 *On a testé un échantillon de* 5 *appareils et noté leurs durées de vie en heures :*

Appareil	1	2	3	4	5
Durée de vie	133	169	8	122	58

On voudrait savoir si la durée de vie suit une loi de probabilité exponentielle On dispose de $n = 5$ *observations.*

On estime le paramètre λ *de la loi exponentielle par la moyenne empirique* $\bar{X}$ *de l'échantillon car* $\bar{X}$ *est un estimateur de*

$$E(X) = \frac{1}{\lambda}.$$

On trouve $\bar{X} = 98$ *et donc on fera les calculs avec* $\lambda = \frac{1}{98}$*. La fonction de répartition de la loi exponentielle est donnée par la formule :*

$$F(x) = 1 - e^{-\lambda x}$$

Voici comment il faut disposer les calculs :

i	1	2	3	4	5
X_i	8	58	122	133	169
$F(X_i)$	0.078	0.447	0.712	0.743	0.822
$\frac{i}{n}$	0.2	0.4	0.6	0.8	1.0
$\left\lvert F(X_i) - \frac{i}{n} \right\rvert$	0.122	0.047	0.112	0.057	0.178
$\frac{i-1}{n}$	0.0	0.2	0.4	0.6	0.8
$\left\lvert F(X_i) - \frac{i-1}{n} \right\rvert$	0.078	0.247	0.312	0.143	0.022

La distance de Kolmogorov-Smirnov est le plus grand des écarts en valeur absolue. On trouve ici

$$D = \max\left\{\left|F(X_i) - \frac{i}{n}\right|, \left|F(X_i) - \frac{i-1}{n}\right|\right\} = 0,312.$$

La table de Kolmogorov-Smirnov pour $n = 5$ *au seuil* $\alpha = 0,05$ *donne la valeur critique* $0,565$. *Puisque* $0,312 < 0,565$, *on accepte l'hypothèse* H_0

Exercise 4 (Application : adéquation à une famille de lois binomiales)
Nombre de garçons dans une fratrie. Pour 10000 *fratries de quatre enfants exactement, on a relevé le nombre de garçons :*

Nombre de garçons	0	1	2	3	4
Effectifs	572	2329	3758	2632	709

On modélise les naissances de la manière suivante :
– *les naissances sont indépendantes;*
– *chaque naissance correspond à la naissance d'un garçon avec probabilité* p, *ou d'une fille avec probabilité* $1 - p$.
Testons

$$H_0 : X \sim \mathcal{B}(4, p),\ \text{où } 0 < p < 1$$

contre

$$H_1 : X \text{ ne suit pas une loi } \mathcal{B}(4, p),\ 0 < p < 1$$

Sous H_0, *l'EMV de* p *est* $\hat{p}_n = \dfrac{\bar{X}_n}{4}$, *qui vaut ici* $0,514425$*:*
avec

$$\bar{X}_n = \frac{\sum_{i=0}^{4} n_i x_i}{n} = 2,0577 \text{ et } n = 10000$$

Nombre de garçons	0	1	2	3	4	*Total*
Probabilités théoriques p_i	$0,0556$	$0,2356$	$0,3744$	$0,2644$	$0,07$	1
Effectifs théoriques np_i	$555,9$	$2355,9$	$3743,8$	$2644,1$	$700,3$	10000
Effectifs observés n_i	572	2329	3758	2632	709	10000
Écarts $\dfrac{(n_i - np_i)^2}{np_i}$	0.4641	0.3064	0.0542	0.0556	0.1079	$\chi^2 \approx 0,99$

Sous H_0,

$$\chi^2_{calc} = \sum_{i=0}^{4} \frac{(n_i - np_i)^2}{np_i} = 0.99$$

suit approximativement la loi $\chi^2_{\upsilon,\alpha}$ $(\upsilon = 3)$ *donc on rejette* H_0 *au niveau asymptotique de* 5% *lorsque* $\chi^2_{calc} \geq \chi^2_{Tab} = 7.81$
On observe ici $\chi^2_{calc} \approx 0,99$ *donc on ne rejette pas* H_0