Statistiques décisionnelles : TD 4

Afin de vous entraîner pour le CC du jeudi 24 février, je vous encourage fortement à tout faire à la main et à la calculatrice, puis à vérifier votre résultat sur R.

Exercice 1. *On observe le vecteur suivant*

$$0.67 \quad 8.09 \quad 2.31 \quad 3.10 \quad 2.06 \quad 0.70 \quad 2.80 \quad 1.68 \quad 4.79 \quad 8.38$$

Ces observations sont-elles compatibles avec une loi du $\chi^2(4)$ *(pour un niveau* 10%*) ? Quel test appliquez-vous ? On s'aidera des valeurs ci-dessous pour la fonction de répartition de la loi* $\chi^2(4)$.

t	0.67	0.70	1.68	2.06	2.31	2.80	3.10	4.79	8.09	8.38
$\mathbb{P}(\chi^2 \le t)$	0.04	0.05	0.2	0.27	0.32	0.40	0.45	0.69	0.91	0.92

Exercice 2. *On a acheté un dé truqué, spécialement conçu pour gagner au jeu* 421. *Le vendeur nous assure que les chiffre* 1, 2 *et* 4 *ont deux fois plus de chances d'être tirés que les autres. Pour vérifier ses dires, on a effectué* 120 *lancers de ce dé et noté les résultats dans le tableau suivant :*

Résultat du dé	1	2	3	4	5	6
Nombre d'observations	28	24	13	33	11	11

1. *À quel vecteur de probabilités les dires du vendeur correspondent-ils ?*
2. *Tester si l'affirmation du vendeur est cohérente avec les résultats observés, au niveau* 10%.

Exercice 3. *On regarde le nombre de tentatives de passage du permis de conduire dans une auto-école avant le premier succès. L'auto-école affiche sur son site qu'à chaque essai, ses candidats ont* 75% *de chances d'obtenir le permis* B.

1. *Quelle est la loi du nombre de passages avant l'obtention du permis ?*
2. *On veut appliquer un test du* χ^2 *pour valider l'adéquation à cette loi. On a relevé le nombre de passages de* 2000 *candidats et candidates. A partir de quelle valeur* k *a-t-on* $2000\mathbb{P}(Z = k) < 5$ *si* Z *est une variable aléatoire de loi* $\mathcal{G}(0.75)$ *?*
3. *Proposer un découpage en classes pour effectuer un test du* χ^2 *pour l'adéquation à la loi proposée. Les effectifs sont donnés par le tableau ci-dessous (mais le découpage n'est peut-être pas le bon !)*

Nombre de passages	1	2	3	4	5	6
Effectif	1488	387	90	25	7	3

4. *À quel vecteur de probabilités les dires du vendeur correspondent-ils ?*
5. *Tester si l'affirmation du vendeur est cohérente avec les résultats observés, au niveau* 10%.

Exercice 4. *Le tableau suivant donne la répartition du nombre d'accidents par jour observés sur une durée de* 100 *jours à un carrefour donné :*

Nombre d'accidents	*0*	*1*	*2*	*3*	*4*	*5*
Nombre de jours correspondants	*19*	*37*	*22*	*14*	*6*	*2*

1. *Qu'est-ce qui permet de penser que le nombre d'accidents par jours à ce carrefour suit une loi de Poisson ?*
2. *Quel est l'estimateur du maximum de vraissemblance pour le paramètre d'un loi de Poisson ?*
3. *Tester l'ajustement à une loi de Poisson au niveau* 5%.

Exercice 5 (χ^2 pour les lois continues)**.** *Il est possible d'adapter le test du* χ^2 *pour tester l'adéquation à une loi continue, et c'est le but de cet exercice dans le cas particulier de lois uniformes sur* $[0,1]$*. Vous pouvez aussi appliquer directement ce qui a été fait en cours avec le découpage suggéré.*

On dispose de 100 *réalisations indépendantes* X_i *de loi inconnue, et on souhaite tester si elles proviennent d'une loi uniforme sur* $[0,1]$*. Malheureusement, il vous est interdit d'utiliser le test de Kolmogorov-Smirnov. Vous disposez uniquement du tableau récapitulatif ci-dessous.*

Intervalles	$[0,0.2[$	$[0.2,0.4[$	$[0.4,0.6[$	$[0.6,0.8[$	$[0.8,1]$
Nombre de variables	*19*	*20*	*16*	*23*	*22*

1. *Si on numérote les intervalles du tableau de* 1 *à* 5 *et qu'on construit les variables*
$$Y_i = \begin{cases} 1 \text{ si } X_i \in [0,0.2[\\ 2 \text{ si } X_i \in [0.2,0.4[\\ \dots \\ 5 \text{ si } X_i \in [0.8,1]. \end{cases}$$
Quelle est la loi des Y_i *sous l'hypothèse* H_0 *: "les* X_i *sont de loi uniforme sur* $[0,1]$ *?*
2. *En déduire un test asymptotique de niveau* α *pour l'adéquation des* X_i *à une loi uniforme sur* $[0,1]$*".*
3. *Existe-t-il d'autres lois de probabilités sur les* X_i *qui mènent à la même loi sur les* Y_i *? Si oui, construisez en une.*
4. *Qu'en déduisez-vous sur l'erreur de seconde espèce de ce test ?*

Exercice 6. *On dispose de* 10 *résultats de simulation de la loi uniforme sur* $[0,1]$ *:*

$$0,134 \quad 0,628 \quad 0,789 \quad 0,905 \quad 0,250 \quad 0,563 \quad 0,790 \quad 0,470 \quad 0,724 \quad 0,569.$$

Étudiez si cet échantillon conduit à rejeter l'hypothèse nulle selon laquelle le tirage a bien eu lieu selon la loi uniforme sur $[0,1]$*. Quel test choisissez-vous ? et pourquoi ?*

Exercice 7. *On dispose de l'échantillon de* 50 *variables suivant, les données ont été classées par ordre croissant pour en faciliter le traitement. Les ex-aequo proviennent d'arrondis.*

0.02	0.02	0.07	0.07	0.07	0.09	0.11	0.13	0.16	0.16
0.2	0.2	0.21	0.25	0.29	0.32	0.35	0.35	0.37	0.37
0.38	0.4	0.4	0.41	0.47	0.49	0.54	0.55	0.59	0.6
0.68	0.68	0.7	0.73	0.75	0.79	0.79	0.8	0.81	0.81
0.86	0.89	1.03	1.04	1.04	1.09	1.12	1.21	1.33	1.42

1. *Tester l'adéquation de ces données à une loi uniforme sur* $[0, 1.5]$ *à l'aide d'un test du* χ^2*, au niveau* 90%*. On proposera un découpage en* 5 *intervalles de même longueur.*
2. *Tester l'adéquation de ces données à la loi dont la densité est*

$$f(x) = \cos(x) 1_{x \in [0, \pi/2]}$$

à l'aide d'un test du χ^2*, au niveau* 90%*. On pourra calculer la fonction de répartition de cette fonction et déterminer les probabilités des intervalles construits précédemment.*