TP n° 8 : Adéquation à une loi de probabilité

Exercice 1. On lance 100 fois un dé à 6 faces numérotées de 1 à 6. On obtient les résultats suivants :

Numéro	1	2	3	4	5	6
Nombre de fois	18	23	19	12	11	15

Peut-on affirmer que le dé est truqué ? *(on fera une analyse graphique convenable, puis un test statistique adapté au risque* 5%*).*

Exercice 2. Une enquête effectuée auprès du comptoir de 116 coopératives agricoles a permis d'étudier l'arrivée dans le temps des usagers de ces coopératives. Pendant une heure, on a les résultats suivants :

Nombre d'usagers arrivés	0	1	2	3	4	5
Nombre de coopératives	28	31	27	19	7	4

Soit X la *var* égale au nombre d'usager arrivés en une heure. Peut-on affirmer, au risque 5%, que X suit la loi de Poisson $\mathcal{P}(2)$?

Exercice 3. Le fabriquant d'un certain type d'appareil affirme que :

- la durée de vie moyenne d'un appareil est de 1500 heures,
- la durée de vie d'un appareil peut être modélisée par une *var* suivant une loi exponentielle.

Afin de tester cette affirmation, on mesure la durée de vie en heures de 10 de ces appareils pris au hasard. Les résultats sont :

576	617	1718	698	3335	512	2024	3259	3146	754

Peut-on dire, au risque 5%, que le fabriquant a tort ?

Exercice 4. Les données considérées sont les observations d'un caractère X que l'on représente comme une *var*. Elles sont disponibles ici :

`https://chesneau.users.lmno.cnrs.fr/observations.txt`

1. Mettre les données dans un vecteur `x`.
2. Représenter l'histogramme des fréquences, le boxplot et le QQ plot associés à `x`. Est-que X semble suivre une loi normale ?
3. Apporter une réponse plus tranchée avec un test statistique adapté.

Exercice 5. La pression artérielle systolique est la pression maximale du sang dans les artères au moment de la contraction du cœur. Celle-ci a été mesurée pour 29 individus de différents âges. Ainsi, pour chacun d'entre eux, on dispose :

- de leur pression systolique en mmHg,
- de leur âge en années.

On modélise ces deux variables comme des *var* Y et X_1. Le jeu de données "pression" est disponible ici :

`https://chesneau.users.lmno.cnrs.fr/pression.txt`

1. Mettre le jeu de données sous la forme d'une data frame `w`, puis attacher les noms des colonnes.
2. Peut-on affirmer, au risque 5%, que Y ne suit pas une loi normale ?
3. Peut-on affirmer, au risque 5%, que X_1 ne suit pas une loi normale ?
4. Reproduire et comprendre l'enjeu des commandes suivantes :

```
par(mfrow = c(1, 2))
qqnorm(scale(X1))
abline(0, 1, col = "red")
qqnorm(scale(Y))
abline(0, 1, col = "blue")
```

Exercice 6. Soient X et Y deux *var* indépendantes. Illustrer les résultats ci-dessous avec la commande `qqplot`.

- Caractérisation de la loi du chi-deux $\chi^2(2)$: Si $X \sim \mathcal{N}(0,1)$ et $Y \sim \mathcal{N}(0,1)$, alors

$$X^2 + Y^2 \sim \chi^2(2).$$

- Caractérisation de la loi de Student $\mathcal{T}(\nu)$: Si $X \sim \mathcal{N}(0,1)$ et $Y \sim \chi^2(\nu)$, alors

$$\frac{X}{\sqrt{\frac{Y}{\nu}}} \sim \mathcal{T}(\nu).$$

  Prendre $\nu = 3.9$.

- Caractérisation de la loi de Fisher $\mathcal{F}(\nu_1, \nu_2)$: Si $X \sim \chi^2(\nu_1)$ et $Y \sim \chi^2(\nu_2)$, alors

$$\frac{\frac{X}{\nu_1}}{\frac{Y}{\nu_2}} = \frac{\nu_2 X}{\nu_1 Y} \sim \mathcal{F}(\nu_1, \nu_2).$$

  Prendre $(\nu_1, \nu_2) = (2.1, 8.3)$.