# LSTAT 2040 - TP 7 : Solutions

Tests d'hypothèse

## Exercice 1

(a) On rappelle que

$$\Lambda_n = \frac{\sup_{\lambda \geq \lambda_0} L_n(\lambda)}{\sup_\lambda L_n(\lambda)}.$$

On calcule pour tout $\lambda > 0$,

$$L_n(\lambda) = \prod_{i=1}^n \frac{\lambda^{X_i}}{X_i!} \exp(-\lambda) = \frac{\lambda^{\sum_{i=1}^n X_i}}{\prod_{i=1}^n X_i!} \exp(-n\lambda),$$

d'où l'on déduit

$$\ell_n(\lambda) = -n\lambda + \sum_{i=1}^n X_i \log(\lambda) - \sum_{i=1}^n \log(X_i!).$$

Dès lors, l'équation

$$\partial_\lambda \ell_n(\lambda) = -n + \frac{\sum_{i=1}^n X_i}{\lambda} = 0$$

n'admet qu'une unique solution donnée par

$$\lambda = \overline{X}_n.$$

Puisque pour tout $\lambda > 0$,

$$\partial_\lambda^2 \ell_n(\lambda) = -\frac{\sum_{i=1}^n X_i}{\lambda^2} < 0,$$

on déduit que le MLE de $\lambda$ est donné par

$$\hat{\lambda}_{\text{MLE}} := \overline{X}_n.$$

On déduit également de la stricte concavité de la vraisemblance que le MLE contraint à la région $H_0 = \{\lambda \geq \lambda_0\}$ vaut $\hat{\lambda}_{\text{MLE}}$ si ce dernier est plus grand que $\lambda_0$ et vaut $\lambda_0$ sinon. On déduit que

$$\begin{aligned}
\Lambda_n &= \begin{cases} \frac{L_n(\lambda_0)}{L_n(\hat{\lambda}_{\text{MLE}})} & \text{si } \hat{\lambda}_{\text{MLE}} \leq \lambda_0 \\ 1 & \text{sinon} \end{cases} \\
&= \begin{cases} \exp(T_n - n\lambda_0) \left(\frac{n\lambda_0}{T_n}\right)^{T_n} & \text{si } T_n \leq n\lambda_0 \\ 1 & \text{si } T_n > n\lambda_0, \end{cases}
\end{aligned}$$

où l'on a posé $T_n := \sum_{i=1}^n X_i$.

(b) Soit $c \in (0, 1)$. Définissons $g : (0, n\lambda_0] \to (0, 1]$ par $g(t) = \exp(t - n\lambda_0)(n\lambda_0/t)^t$. Alors, on a que la région de rejet est donnée par

$$\{\Lambda_n \leq c\} = \{g(T_n) \leq c\}.$$

Montrons que la fonction $g$ est croissante. Pour cela, puisque le logarithme est une transformation croissante, on peut simplement montrer que la composée $\log \circ g$ est croissante. On calcule pour tout $t \in (0, n\lambda_0]$,

$$\partial_t \log(g(t)) = \partial_t \left\{ t - n\lambda_0 + t \log\left(\frac{n\lambda_0}{t}\right) \right\} = \log\left(\frac{n\lambda_0}{t}\right) \geq 0.$$

On déduit que $g$ est croissante et donc que la région de rejet vaut

$$\{g(T_n) \leq c\} = \{T_n \leq a\}, \qquad a \in (0, n\lambda_0].$$

Soit $\alpha \in (0,1)$. Le test sera de niveau $\alpha$ si

$$\sup_{\lambda \geq \lambda_0} P\left(T_n \leq a(\alpha)|\lambda\right) \leq \alpha,$$

i.e., la probabilité de rejeter $H_0$ si le vrai paramètre est dans $H_0$ est au maximum de $\alpha$. Soit $\lambda \geq \lambda_0$. Puisque $T_n|\lambda \sim \mathrm{Poi}(n\lambda)$, on a pour tout $x \in [0,\infty)$,

$$P\left(T_n \leq x|\lambda\right) \leq P\left(T_n \leq x|\lambda_0\right),$$

d'où l'on déduit que

$$\sup_{\lambda \geq \lambda_0} P\left(T_n \leq a(\alpha)|\lambda\right) = P\left(T_n \leq a(\alpha)|\lambda_0\right).$$

Dès lors, si l'on choisit $a(\alpha)$ comme étant le quantile d'ordre $\alpha$ d'une Poisson $\mathrm{Poi}(n\lambda_0)$, on est sûr que le test sera de niveau $\alpha$.

(c) Pour de telles valeurs, on a $a(0.05) = 39$ (obtenu via la commande `qpois(0.05, lambda = 25*2)` sur `R`). Dès lors, $t_n = 38$ appartient à la région de rejet et l'on rejette $H_0$.

(d) Avec les nouvelles hypothèses, on aurait obtenu que la statistique de test du LRT exact est donnée par

$$\Lambda_n = g(T_n),$$

pour n'importe quelle valeur de la statistique $T_n$. Ainsi, puisque la fonction $g$ est décroissante sur la région $[n\lambda_0, \infty)$, ce qui se voit directement de notre calcul de dérivée au point (a), on aurait eu que la région de rejet s'exprime sous la forme

$$\{\Lambda_n \leq c\} = \{g(T_n) \leq c\} = \{T_n \leq a\} \cup \{T_n \geq b\},$$

pour $0 < a < n\lambda_0 < b < \infty$. Dès lors, pour obtenir un test de niveau $\alpha \in (0,1)$, on aurait du prendre $a(\alpha)$ et $b(\alpha)$ comme les quantiles d'ordre $\alpha/2$ et $1-\alpha/2$ d'une Poisson $\mathrm{Poi}(n\lambda_0)$. Avec les données du point (c), on aurait alors trouvé $a(0.05) = 37$ et $b(0.05) = 64$, et on aurait pas pu conclure au rejet de l'hypothèse $H_0$ (bien que les données sous-jacentes à l'obtention de la statistique observée $t_n = 38$ aient été simulées avec un paramètre dans la région $H_1$).

## Exercice 2

(a) On rappelle que

$$\Lambda_n = \frac{\sup_{\pi \leq \pi_0} L_n(\pi)}{\sup_{\pi} L_n(\pi)}.$$

On calcule pour tout $\pi \in (0,1)$,

$$L_n(\pi) = \prod_{i=1}^n \pi^{X_i}(1-\pi)^{1-X_i} = \pi^{\sum_{i=1}^n X_i}(1-\pi)^{n-\sum_{i=1}^n X_i},$$

d'où l'on déduit

$$\ell_n(\pi) = (\sum_{i=1}^n X_i)\log(\pi) + (n - \sum_{i=1}^n X_i)\log(1-\pi).$$

Dès lors, l'équation

$$\partial_\pi \ell_n(\pi) = \frac{\sum_{i=1}^n X_i}{\pi} - \frac{n - \sum_{i=1}^n X_i}{1-\pi} = 0$$

n'admet qu'une unique solution donnée par

$$\pi = \overline{X}_n.$$

Puisque pour tout $\pi \in (0,1)$

$$\partial_\pi^2 \ell_n(\pi) = -\frac{\sum_{i=1}^n X_i}{\pi^2} - \frac{n - \sum_{i=1}^n X_i}{(1-\pi)^2} < 0,$$

on déduit que le MLE de $\pi$ est donné par

$$\hat{\pi}_{\text{MLE}} := \overline{X}_n.$$

On déduit également de la stricte concavité de la vraisemblance que le MLE contraint à la région $H_0 = \{\pi \leq \lambda_0\}$ vaut $\hat{\pi}_{\text{MLE}}$ si ce dernier est plus petit que $\pi_0$ et vaut $\pi_0$ sinon. On déduit que

$$\begin{aligned}\Lambda_n &= \begin{cases} \frac{L_n(\pi_0)}{L_n(\hat{\pi}_{\text{MLE}})} & \text{si } \hat{\pi}_{\text{MLE}} \geq \pi_0 \\ 1 & \text{sinon} \end{cases} \\ &= \begin{cases} \left(\frac{n\pi_0}{T_n}\right)^{T_n} \left(\frac{n-n\pi_0}{n-T_n}\right)^{n-T_n} & \text{si } T_n \geq n\pi_0 \\ 1 & \text{si } T_n < n\pi_0, \end{cases}\end{aligned}$$

où l'on a posé $T_n := \sum_{i=1}^n X_i$.

(b) Il suffit de montrer que la fonction $g : [n\pi_0, n) \to (0,1]$ définie par $g(t) := (n\pi_0/t)^t[(n-n\pi_0)/(n-t)]^{n-t}$ est décroissante. Pour cela, comme pour l'exercice précédent, montrons que la composition $\log \circ g$ est décroissante. On calcule pour tout $t \in [n\pi_0, n)$,

$$\partial_t \log(g(t)) = \partial_t \left\{ t \log\left(\frac{n\pi_0}{t}\right) + (n-t)\log\left(\frac{n-n\pi_0}{n-t}\right)\right\} = \log\left(\frac{n\pi_0}{t}\right) - \log\left(\frac{n-n\pi_0}{n-t}\right) \leq 0.$$

(c) Afin de réaliser le test au niveau $\alpha = 0.05$, il faut trouver une région de rejet telle que la probabilité de commettre une erreur de type I est au maximum $\alpha$. Il suit des calculs précédents que pour $c \in (0,1)$, la région de rejet peut s'écrire

$$\{\Lambda_n \leq c\} = \{T_n \geq b\},$$

pour un certain $b \in (n\pi_0, n)$. Soit $\pi \in H_0 = \{\pi \leq \pi_0\}$. Alors, on a $T_n|\pi \sim \text{Bin}(n,\pi)$ et on peut montrer, comme à l'exercice précédent,

$$\sup_{\pi \leq \pi_0} P\left(T_n \geq b|\pi\right) = P\left(T_n \geq b|\pi_0\right),$$

pour n'importe quel $b \in (n\pi_0, n)$. Ainsi, si l'on prend $b = b(\alpha)$ comme étant le quantile d'ordre $\alpha = 0.95$ d'une Binomiale $\text{Bin}(n,\pi_0)$, on est certain que le test est de niveau alpha. On trouve

$$b(0.05) = \texttt{qbinom(0.95, size = 50, prob = 0.5)} = 31,$$

et comme $t_n < 31$, on ne rejette pas $H_0$ pour un tel choix de $\alpha$.

(d) La région critique associée à ce nouveau test aurait été déterminée par

$$\tilde{b}(0.05) = \texttt{qbinom(0.95, size = 100, prob = 0.5)} = 58.$$

Puisque maintenant on a $t_n = 59$, on aurait rejeté $H_0$, bien que la probabilité de succès estimée sur les données est restée identique, à savoir $\hat{\pi} = 0.59$. Cela provient du fait que, comme la taille d'échantillon $n$ a augmenté, le test a gagné en puissance.

## Exercice 3

(a) Par souci de clarté, notons $\theta_0 = \theta_X = \theta_Y$ la valeur commune du paramètre sous $H_0$. La statistique de test $\Lambda_{n,m}$ est donnée par

$$\Lambda_{n,m} = \frac{\sup_{\theta_0} L_{n,m}(\theta_0,\theta_0)}{\sup_{\theta_X,\theta_Y} L_{n,m}(\theta_X,\theta_Y)},$$

où $L_{n,m}$ est la vraisemblance jointe de l'échantillon joint $(X_1, \ldots, X_n, Y_1, \ldots, Y_m)$ donnée par

$$L_{n,m}(\theta_X, \theta_Y) = \theta_X^{-n} \exp\left(-\frac{\sum_{i=1}^n X_i}{\theta_X}\right) \times \theta_Y^{-m} \exp\left(-\frac{\sum_{j=1}^m Y_j}{\theta_Y}\right).$$

On montre assez facilement, en utilisant des arguments similaires à ceux utilisés précédemment, que les MLE non contraints au dénominateur sont donnés par

$$\hat{\theta}_X = \overline{X}_n \qquad \text{et} \qquad \hat{\theta}_Y = \overline{Y}_m,$$

et que le MLE contraint du numérateur est donné par

$$\hat{\theta}_0 = \frac{\sum_{i=1}^n X_i + \sum_{j=1}^m Y_j}{n+m},$$

qui, intuitivement, est l'estimateur que l'on s'attend à obtenir sous $H_0$ puisque sous cette hypothèse, les deux exponentielles ont le même paramètre et on peut estimer ce paramètre commun en joignant toutes les données. Ainsi, on obtient

$$\begin{aligned}
\Lambda_{n,m} &= \frac{\left(\frac{\sum_{i=1}^n X_i + \sum_{j=1}^m Y_j}{n+m}\right)^{-(n+m)} \exp\left(-(n+m)\right)}{\left(\frac{\sum_{i=1}^n X_i}{n}\right)^{-n} \left(\frac{\sum_{j=1}^m Y_j}{m}\right)^{-m} \exp\left(-(n+m)\right)} \\
&= \frac{(n+m)^{n+m}}{n^n m^m} \frac{\left(\sum_{i=1}^n X_i\right)^n \left(\sum_{j=1}^m Y_j\right)^m}{\left(\sum_{i=1}^n X_i + \sum_{j=1}^m Y_j\right)^{n+m}} \\
&= \frac{(n+m)^{n+m}}{n^n m^m} T_{n,m}^n \left(1 - T_{n,m}\right)^m.
\end{aligned}$$

(b) Définissons la fonction $g : t \in [0,1] \mapsto g(t) := \frac{(n+m)^{n+m}}{n^n m^m} t^n (1-t)^m$ qui est telle que

$$\Lambda_{n,m} = g\left(T_{n,m}\right).$$

Soit $c \in (0,1)$. On cherche à réécrire la région de rejet $\{\Lambda_{n,m} \leq c\}$ comme un région de rejet sur $T_{n,m}$. Pour cela, notons que la fonction $g$ satisfait $g(0) = g(1) = 0$, que l'unique solution de l'équation de premier ordre $g'(t_0) = 0$ est donnée par

$$t_0 := \frac{n}{n+m} \in (0,1),$$

et que $g''(t_0) < 0$. Ainsi, on déduit que

$$\{\Lambda_{n,m} \leq c\} = \{g\left(T_{n,m}\right) \leq c\} = \{T_{n,m} \leq a\} \cup \{T_{n,m} \geq b\},$$

pour $0 < a < b < 1$. Soit $\alpha \in (0,1)$. On chercher $a = a(\alpha)$ et $b = b(\alpha)$ tels que

$$P\left(\{T_{n,m} \leq a\} \cup \{T_{n,m} \geq b\} \,\middle|\, H_0\right) \leq P\left(T_{n,m} \leq a | H_0\right) + P\left(T_{n,m} \geq b | H_0\right) \leq \alpha$$

Comme $T_{n,m} | H_0 \sim \text{Beta}(n,m)$, si l'on souhaite répartir les probabilités d'erreur de type I de manière symétrique, on prendra respectivement comme valeurs de $a(\alpha)$ et $b(\alpha)$ les quantiles d'ordre $\alpha/2$ et $1-\alpha/2$ d'une distribution Beta $\text{Beta}(n,m)$.

(c) Pour de telles valeurs, on a

$$\begin{aligned}
a(0.05) &= \texttt{qbeta(0.025, shape1 = 35, shape2 = 40)} = 0.356 \\
b(0.05) &= \texttt{qbeta(0.975, shape1 = 35, shape2 = 40)} = 0.579,
\end{aligned}$$

et l'on voit que $t_{n,m} > b(0.05)$, ce qui nous permet de conclure au rejet de $H_0$.

## Exercice 4

(a) La statistique de test $\Lambda_n$ est donnée par

$$\Lambda_n = \frac{\sup_\nu L_n(1,\nu)}{\sup_{\theta,\nu} L_n(\theta,\nu)}.$$

En faisant un graphique de la vraisemblance $L_n(\theta,\nu) = \theta^n \nu^{n\theta} \left(\prod_{i=1}^n X_i\right)^{-(\theta+1)} \mathrm{I}(\nu \leq X_{(1)})$, on voit facilement que peu importe la valeur de $\theta$, le paramètre $\nu$ qui maximise la vraisemblance est

$$\hat{\nu} := X_{(1)}.$$

Dès lors, la log-vraisemblance profilée en $\theta$ est donnée par

$$\ell_n^p(\theta) = n\log(\theta) + n\theta \log\left(X_{(1)}\right) - (\theta+1)\sum_{i=1}^n \log(X_i).$$

En résolvant la condition de premier ordre et en observant que pour tout $\theta$, on a $\partial_\theta^2 \ell_n^p(\theta) < 0$, on trouve que la valeur de $\theta$ qui maximise cette expression est donnée par

$$\hat{\theta} := \frac{n}{\sum_{i=1}^n \log(X_i/X_{(1)})} = \frac{n}{T_n}.$$

Ainsi, on déduit

$$\begin{aligned}\Lambda_n &= \frac{T_n^n}{n^n} \frac{\left(\prod_{i=1}^n X_i\right)^{(n-T_n)/T_n}}{X_{(1)}^{n(n-T_n)/T_n}} \\ &= \frac{T_n^n}{n^n} \exp\left(-\frac{T_n - n}{T_n} T_n\right) \\ &= \frac{T_n^n}{n^n} \exp\left(-T_n + n\right).\end{aligned}$$

(b) Soit la fonction $g : t \in [0,\infty) \mapsto g(t) := t^n/n \times \exp(-t+n)$ telle que $\Lambda_n = g(T_n)$. Notons que $g$ satisfait $g(0) = \lim_{t\to\infty} g(t) = 0$, que l'unique solution strictement positive de l'équation de premier ordre $g'(t_0) = 0$ est donnée par $t_0 = n$ et que $g''(t_0) < 0$. Dès lors, on déduit que pour tout $c \in (0,1)$,

$$\{\Lambda_n \leq c\} = \{g(T_n) \leq c\} = \{T_n \leq a\} \cup \{T_n \geq b\},$$

où $0 < a < b < \infty$. Soit $\alpha \in (0,1)$. Afin d'obtenir un test de niveau $\alpha$, il suffit (voir exercices précédents) de choisir $a = a(\alpha)$ et $b = b(\alpha)$ tels que

$$P\left(T_n \leq a | H_0\right) + P\left(T_n \geq b | H_0\right) \leq \alpha.$$

Ainsi, puisque $T_n | H_0 \sim \text{Gamma}(n-1,1)$, un choix valide est de prendre respectivement pour $a$ et $b$ les quantiles d'ordre $\alpha/2$ et $1-\alpha/2$ de la distribution Gamma $\text{Gamma}(n-1,1)$.

(c) Pour de telles valeurs, on a

$$\begin{aligned}a(0.05) &= \texttt{qgamma(0.025, shape = 25 - 1, rate = 1)} = 15.377 \\ b(0.05) &= \texttt{qgamma(0.975, shape = 25 - 1, rate = 1)} = 34.511,\end{aligned}$$

et l'on voit que $t_n > b(0.05)$, ce qui nous permet de conclure au rejet de $H_0$.

## Exercice 5

(a) On rappelle que les statistiques des tests proposés sont données, respectivement, par

$$\begin{aligned}W_n &= I_n(\hat{\theta})(\hat{\theta} - \theta_0)^2 \\ R_n &= \frac{S_n(\theta_0)^2}{I_n(\theta_0)} \\ T_n &= 2\left(\ell_n(\hat{\theta}) - \ell_n(\theta_0)\right),\end{aligned}$$

où $\hat{\theta}$ désigne le MLE de $\theta$, $S_n(\theta) = \ell'_n(\theta)$ désigne le score et $I_n = -\operatorname{E}[S'_n(\theta)] = -\operatorname{E}[\ell''_n(\theta)]$ désigne l'information de Fisher de l'échantillon. Puisque toutes les quantités utiles au calcul de ces statistiques ont déjà été calculées précédemment pour la distribution $\text{Beta}(\theta, 2)$, nous nous contentons de les rappeler ici :

$$\begin{aligned}
\ell_n(\theta) &= n\left(\log(\theta) + \log(\theta+1)\right) + (\theta - 1)\sum_{i=1}^n \log(X_i) + \sum_{i=1}^n \log(1 - X_i) \\
\hat{\theta} &= \frac{-2 - \overline{\log(X)}_n - \sqrt{4 + \overline{\log(X)}_n^2}}{2\overline{\log(X)}_n} \\
S_n(\theta) &= n\frac{2\theta+1}{\theta(\theta+1)} + \sum_{i=1}^n \log(X_i) \\
I_n(\theta) &= n\frac{2\theta^2 + 2\theta + 1}{\theta^2(\theta+1)^2}.
\end{aligned}$$

(b) Puisque la distribution $\text{Beta}(\theta, 2)$ satisfait les hypothèses de normalité du MLE, on sait que sous $H_0$,

$$W_n, R_n, T_n \xrightarrow{d} \chi_1^2,$$

quand $n \to \infty$. Dès lors, pour un niveau $\alpha \in (0, 1)$ donné, les régions critiques des tests asymptotiques de niveau $\alpha$ basés sur ces statistiques sont toutes identiques et sont données par

$$\{x > 0 : x > \chi^2_{1;1-\alpha}\},$$

où $\chi^2_{1;1-\alpha}$ désigne le quantile d'ordre $1-\alpha$ d'une distribution Chi-carrée $\chi_1^2$. Avec les données proposées, on calcule

$$\begin{aligned}
w_n &= 7.248 \\
r_n &= 7.239 \\
t_n &= 8.377,
\end{aligned}$$

et $\chi^2_{1;0.95} = 3.84$, ainsi, les trois tests mènent au rejet de l'hypothèse nulle. Cela n'est pas étonnant puisque les trois tests sont asymptotiquement équivalents, i.e., leur décision sera identique avec probabilité 1 quand $n \to \infty$.

## Exercice 6

(a) Comme pour l'exercice précédent, nous nous contentons de rappeler les quantités utiles au calcul des statistiques des trois tests proposés :

$$\begin{aligned}
\ell_n(\theta) &= n\left(\log(2) - \log(\theta)\right) + \sum_{i=1}^n \log(X_i) - \frac{1}{\theta}\sum_{i=1}^n X_i^2 \\
\hat{\theta} &= \overline{X^2}_n \\
S_n(\theta) &= -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^n X_i^2 \\
I_n(\theta) &= \frac{n}{\theta^2},
\end{aligned}$$

où il est utile de noter, pour le calcul de l'information de Fisher, que $X^2 \sim \text{Exp}(\theta)$.

(b) Puisque la distribution considérée satisfait les hypothèses de normalité du MLE, on sait que sous $H_0$,

$$W_n, R_n, T_n \xrightarrow{d} \chi_1^2,$$

quand $n \to \infty$. Dès lors, pour un niveau $\alpha \in (0,1)$ donné, les régions critiques des tests asymptotiques de niveau $\alpha$ basés sur ces statistiques sont toutes identiques et sont données par

$$\{x > 0 : x > \chi^2_{1;1-\alpha}\},$$

où $\chi^2_{1;1-\alpha}$ désigne le quantile d'ordre $1-\alpha$ d'une distribution Chi-carrée $\chi^2_1$. Avec les données proposées, on calcule

$$\begin{aligned} w_n &= 0.0106 \\ r_n &= 0.0104 \\ t_n &= 0.0105, \end{aligned}$$

et $\chi^2_{1;0.95} = 3.84$, ainsi, aucun des trois tests ne mène au rejet de l'hypothèse nulle. Cela n'est pas étonnant puisque les trois tests sont asymptotiquement équivalents, i.e., leur décision sera identique avec probabilité 1 quand $n \to \infty$. En particulier, le vecteur de données ayant mené à l'observation $\sum_{i=1}^n x_i^2 = 98.98$ a été généré à partir de la distribution avec $\theta = 1$.

## Exercice 7

(a) On rappelle que pour, tester de telles hypothèses, les statistiques des trois tests asymptotiques sont respectivement données par

$$\begin{aligned} W_n &= \left((\hat\theta_X, \hat\theta_Y) - (\theta_X^0, \theta_Y^0)\right)^T I_n(\hat\theta_X, \hat\theta_Y) \left((\hat\theta_X, \hat\theta_Y) - (\theta_X^0, \theta_Y^0)\right) \\ R_n &= S_n(\theta_X^0, \theta_Y^0)^T I_n^{-1}(\theta_X^0, \theta_Y^0) S_n(\theta_X^0, \theta_Y^0) \\ T_n &= 2\left(\ell_n(\hat\theta_X, \hat\theta_Y) - \ell_n(\theta_X^0, \theta_Y^0)\right) \end{aligned}$$

Afin de calculer ces dernières, on vérifie aisément que

$$\begin{aligned} \ell_n(\theta_X, \theta_Y) &= -n\left(\log(\theta_X) + \log(\theta_Y)\right) - \frac{\sum_{i=1}^n X_i}{\theta_X} - \frac{\sum_{j=1}^n Y_j}{\theta_Y} \\ S_n(\theta_X, \theta_Y) &= n\left(\frac{\overline{X}_n - \theta_X}{\theta_X^2}, \frac{\overline{Y}_n - \theta_Y}{\theta_Y^2}\right)^T \\ (\hat\theta_X, \hat\theta_Y) &= \left(\overline{X}_n, \overline{Y}_n\right) \\ I_n(\theta_X, \theta_Y) &= n\begin{pmatrix} 1/\theta_X^2 & 0 \\ 0 & 1/\theta_Y^2 \end{pmatrix} \end{aligned}$$

Dès lors, quelques manipulations algébriques permettent de montrer que

$$\begin{aligned} W_n &= n\left\{\frac{(\overline{X}_n - \theta_X^0)^2}{(\overline{X}_n)^2} + \frac{(\overline{Y}_n - \theta_Y^0)^2}{(\overline{Y}_n)^2}\right\} \\ R_n &= n\left\{\frac{(\overline{X}_n - \theta_X^0)^2}{(\theta_X^0)^2} + \frac{(\overline{Y}_n - \theta_Y^0)^2}{(\theta_Y^0)^2}\right\} \\ T_n &= 2n\left\{\log\left(\frac{\theta_X^0}{\overline{X}_n}\right) + \log\left(\frac{\theta_Y^0}{\overline{Y}_n}\right) + \frac{\overline{X}_n}{\theta_X^0} + \frac{\overline{Y}_n}{\theta_Y^0} - 2\right\} \end{aligned}$$

Soit $\alpha \in (0,1)$. Comme chacune de ces statistiques de test converge en distribution vers une distribution Chi-carrée $\chi^2_2$, la région de rejet est identique pour chaque test et vaut

$$\{x : x > \chi^2_{2;1-\alpha}\}$$

si l'on souhaite effectuer le test au niveau $\alpha$.

(b) L'avantage d'un tel test par rapport à deux tests indépendants sur $\theta_X$ et $\theta_Y$ est qu'il n'y a pas de correction de région de rejet (ou, de manière équivalente, de p-valeur) à effectuer pour garder le même niveau $\alpha = 0.025$. Par contre, évidemment, le test proposé ici, s'il rejette $H_0$, ne permet pas de savoir de quel paramètre provient le rejet.

## Exercice 8

Notons $X_{ij}$ pour $i = 1, \dots, 200$ et $j = 1, \dots, 3$, la variable aléatoire qui vaut 1 si l'individu $i$ dans le quartier $j$ est pour le candidat A et qui vaut 0 sinon. Alors, on a $X_{ij} \sim \mathrm{Be}(\pi_j)$ pour un certain $\pi_j \in (0, 1)$ et l'on souhaite effectuer le test

$$H_0 : \pi_1 = \pi_2 = \pi_3 \qquad \text{contre} \qquad H_1 : \exists j_1 \neq j_2, \quad \pi_{j_1} \neq \pi_{j_2}.$$

Ces hypothèses peuvent se réécrire sous la forme

$$H_0 : A(\pi_1, \pi_2, \pi_3)^T = (0, 0)^T \qquad \text{contre} \qquad H_1 : A(\pi_1, \pi_2, \pi_3)^T \neq (0, 0)^T,$$

où $A : \mathbb{R}^3 \to \mathbb{R}^2$ est l'opérateur linéaire défini par la matrice

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{pmatrix}.$$

En particulier, il s'agit d'un test (asymptotique) sur une hypothèse composée. Par conséquent, les statistiques des deux tests demandés sont données par

$$\begin{aligned} W_n &= \left(A(\hat{\pi}_1, \hat{\pi}_2, \hat{\pi}_3)^T\right)^T \left(A I_n^{-1}(\hat{\pi}_1, \hat{\pi}_2, \hat{\pi}_3) A^T\right)^{-1} \left(A(\hat{\pi}_1, \hat{\pi}_2, \hat{\pi}_3)^T\right) \\ T_n &= 2\left(\ell_n(\hat{\pi}_1, \hat{\pi}_2, \hat{\pi}_3) - \ell_n(\tilde{\pi}_{1;0}, \tilde{\pi}_{2;0}, \tilde{\pi}_{3;0})\right), \end{aligned}$$

où l'on vérifie facilement que

$$\begin{aligned} (\hat{\pi}_1, \hat{\pi}_2, \hat{\pi}_3) &= \left(\overline{X_1}, \overline{X_2}, \overline{X_3}\right) \\ (\tilde{\pi}_{1;0}, \tilde{\pi}_{2;0}, \tilde{\pi}_{3;0}) &= \left(\overline{X}, \overline{X}, \overline{X}\right) \\ \ell_n(\pi_1, \pi_2, \pi_3) &= \sum_{j=1}^{3} \left\{ \left(\sum_{i=1}^{200} X_{ij}\right) \log(\pi_j) + \left(200 - \sum_{i=1}^{200} X_{ij}\right) \log(1 - \pi_j) \right\} \\ I_n(\pi_1, \pi_2, \pi_3) &= n \begin{pmatrix} \frac{1}{\pi_1(1-\pi_1)} & 0 & 0 \\ 0 & \frac{1}{\pi_2(1-\pi_2)} & 0 \\ 0 & 0 & \frac{1}{\pi_3(1-\pi_3)} \end{pmatrix} \end{aligned}$$

où l'on a noté, pour $j = 1, \dots, 3$, $\overline{X_j} := (200)^{-1} \sum_{i=1}^{200} X_{ij}$ et $\overline{X} := (600)^{-1} \sum_{j=1}^{3} \sum_{i=1}^{200} X_{ij}$. Quelques manipulations algébriques conduisent alors aux expressions suivantes pour les statistiques de test :

$$\begin{aligned} W_n &= 200 \frac{(\overline{X_1} - \overline{X_2})^2 \left(\overline{X_2}(1 - \overline{X_2}) + \overline{X_3}(1 - \overline{X_3})\right) + 2(\overline{X_1} - \overline{X_2})(\overline{X_2} - \overline{X_3})\overline{X_2}(1 - \overline{X_2})}{\overline{X_1}(1 - \overline{X_1})\overline{X_2}(1 - \overline{X_2}) + \overline{X_1}(1 - \overline{X_1})\overline{X_3}(1 - \overline{X_3}) + \overline{X_2}(1 - \overline{X_2})\overline{X_3}(1 - \overline{X_3})} \\ &\quad + n \frac{(\overline{X_2} - \overline{X_3})^2 \left(\overline{X_1}(1 - \overline{X_1}) + \overline{X_2}(1 - \overline{X_2})\right)}{\overline{X_1}(1 - \overline{X_1})\overline{X_2}(1 - \overline{X_2}) + \overline{X_1}(1 - \overline{X_1})\overline{X_3}(1 - \overline{X_3}) + \overline{X_2}(1 - \overline{X_2})\overline{X_3}(1 - \overline{X_3})} \\ T_n &= 2 \times 200 \sum_{j=1}^{3} \left\{ \overline{X_j} \log\left(\frac{\overline{X_j}}{\overline{X}}\right) + \left(1 - \overline{X_j}\right) \log\left(\frac{1 - \overline{X_j}}{1 - \overline{X}}\right) \right\} \end{aligned}$$

Puisque la distribution des données satisfait les hypothèses de normalité du MLE, on sait que sous $H_0$,

$$W_n, T_n \xrightarrow{d} \chi^2_2,$$

quand $n \to \infty$. Dès lors, au niveau $\alpha = 0.05$, on doit comparer les valeurs observées des statistiques de test, données par

$$\begin{aligned} w_n &= 10.253 \\ t_n &= 10.535, \end{aligned}$$

au quantile $\chi_{2;0.95} = 5.99$ et l'on constate que $H_0$, l'hypothèse d'égalité des proportions de votants en faveur du candidat A entre les trois régions est rejetée, peu importe le test utilisé.

## Exercice 9

Notons que les hypothèses du test peuvent être réécrites comme

$$H_0 : A(\mu_X, \mu_Y, \sigma)^T = (\mu_X^0, \mu_Y^0)^T \qquad \text{contre} \qquad H_1 : A(\mu_X, \mu_Y, \sigma)^T \neq (\mu_X^0, \mu_Y^0)^T,$$

où $A : \mathbb{R}^3 \to \mathbb{R}^2$ est l'opérateur linéaire défini par la matrice

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

En particulier, il s'agit d'un test (asymptotique) sur une hypothèse composée. Par conséquent, les statistiques des deux tests demandés sont données par

$$\begin{aligned} W_n &= \left(A(\hat{\mu}_X, \hat{\mu}_Y, \hat{\sigma})^T - (\mu_X^0, \mu_Y^0)^T\right)^T \left(A I_n^{-1}(\hat{\lambda}_X, \hat{\lambda}_Y) A^T\right)^{-1} \left(A(\hat{\mu}_X, \hat{\mu}_Y, \hat{\sigma})^T - (\mu_X^0, \mu_Y^0)^T\right) \\ T_n &= 2\left(\ell_n(\hat{\mu}_X, \hat{\mu}_Y, \hat{\sigma})) - \ell_n(\tilde{\mu}_{X;0}, \tilde{\mu}_{Y;0}, \tilde{\sigma}_0)\right), \end{aligned}$$

où l'on vérifie facilement que

$$\begin{aligned} (\hat{\mu}_X, \hat{\mu}_Y, \hat{\sigma}) &= \left(\overline{X}_n, \overline{Y}_n, \sqrt{\frac{\sum_{i=1}^n (X_i - \overline{X}_n)^2 + \sum_{j=1}^n (Y_j - \overline{Y}_n)^2}{2n}}\right) \\ (\tilde{\mu}_{X;0}, \tilde{\mu}_{Y;0}, \tilde{\sigma}_0) &= \left(\mu_X^0, \mu_Y^0, \sqrt{\frac{\sum_{i=1}^n (X_i - \mu_X^0)^2 + \sum_{j=1}^n (Y_j - \mu_Y^0)^2}{2n}}\right) \\ \ell_n(\mu_X, \mu_Y, \sigma) &= -n\left(\log(2\pi) + 2\log(\sigma)\right) - \frac{1}{2\sigma^2}\left(\sum_{i=1}^n (X_i - \mu_X)^2 + \sum_{j=1}^n (Y_j - \mu_Y)^2\right) \\ I_n(\mu_X, \mu_Y, \sigma) &= n \begin{pmatrix} \frac{1}{\sigma^2} & 0 & 0 \\ 0 & \frac{1}{\sigma^2} & 0 \\ 0 & 0 & \frac{1}{\sigma^4} \end{pmatrix} \end{aligned}$$

Quelques manipulations algébriques conduisent alors aux expressions suivantes pour les statistiques de test :

$$\begin{aligned} W_n &= 2n^2 \frac{(\overline{X}_n - \mu_X^0)^2 + (\overline{Y}_n - \mu_Y^0)^2}{\sum_{i=1}^n (X_i - \overline{X}_n)^2 + \sum_{j=1}^n (Y_j - \overline{Y}_n)^2} \\ T_n &= 2n \log\left(\frac{\sum_{i=1}^n (X_i - \mu_X^0)^2 + \sum_{j=1}^n (Y_j - \mu_Y^0)^2}{\sum_{i=1}^n (X_i - \overline{X}_n)^2 + \sum_{j=1}^n (Y_j - \overline{Y}_n)^2}\right) \end{aligned}$$

Puisque la distribution considérée satisfait les hypothèses de normalité du MLE, on sait que sous $H_0$,

$$W_n, T_n \xrightarrow{d} \chi_2^2,$$

quand $n \to \infty$. Dès lors, pour un niveau $\alpha \in (0, 1)$ donné, les régions critiques des tests asymptotiques de niveau $\alpha$ basés sur ces statistiques sont toutes deux identiques et sont données par

$$\{x > 0 : x > \chi^2_{2;1-\alpha}\},$$

où $\chi^2_{2;1-\alpha}$ désigne le quantile d'ordre $1 - \alpha$ d'une distribution Chi-carrée $\chi_2^2$.

## Exercice 10

(a) Notons que les hypothèses du test peuvent être réécrites comme

$$H_0 : A(\lambda_X, \lambda_Y)^T = 0 \qquad \text{contre} \qquad H_1 : A(\lambda_X, \lambda_Y)^T \neq 0,$$

où $A : \mathbb{R}^2 \to \mathbb{R}^1$ est l'opérateur linéaire défini par la matrice

$$A = (2, -1).$$

En particulier, il s'agit d'un test (asymptotique) sur une hypothèse composée. Par conséquent, les statistiques des trois tests demandés sont données par

$$\begin{aligned}
W_n &= \left(A(\hat{\lambda}_X, \hat{\lambda}_Y)^T\right)^T \left(A I_n^{-1}(\hat{\lambda}_X, \hat{\lambda}_Y) A^T\right)^{-1} \left(A(\hat{\lambda}_X, \hat{\lambda}_Y)^T\right) \\
R_n &= S_n^T(\tilde{\lambda}_{X;0}, \tilde{\lambda}_{Y;0}) I_n^{-1}(\tilde{\lambda}_{X;0}, \tilde{\lambda}_{Y;0}) S_n(\tilde{\lambda}_{X;0}, \tilde{\lambda}_{Y;0}) \\
T_n &= 2\left(\ell_n(\hat{\lambda}_X, \hat{\lambda}_Y) - \ell_n(\tilde{\lambda}_{X;0}, \tilde{\lambda}_{Y;0})\right),
\end{aligned}$$

où l'on vérifie facilement que

$$\begin{aligned}
(\hat{\lambda}_X, \hat{\lambda}_Y) &= (\overline{X}_n, \overline{Y}_n) \\
(\tilde{\lambda}_{X;0}, \tilde{\lambda}_{Y;0}) &= \left(\frac{\overline{X}_n + \overline{Y}_n}{3}, \frac{2(\overline{X}_n + \overline{Y}_n)}{3}\right) \\
\ell_n(\lambda_X, \lambda_Y) &= -n\left(\lambda_X + \lambda_Y\right) + \left(\sum_{i=1}^n X_i\right)\log(\lambda_X) + \left(\sum_{j=1}^n Y_j\right)\log(\lambda_Y) - \left(\sum_{i=1}^n \log(X_i!) + \sum_{j=1}^n \log(Y_j!)\right) \\
S_n(\lambda_X, \lambda_Y) &= \left(\frac{\overline{X}_n}{\lambda_X} - 1, \frac{\overline{Y}_n}{\lambda_Y} - 1\right)^T \\
I_n(\lambda_X, \lambda_Y) &= n\begin{pmatrix} \frac{1}{\lambda_X} & 0 \\ 0 & \frac{1}{\lambda_Y} \end{pmatrix}
\end{aligned}$$

Quelques manipulations algébriques conduisent alors aux expressions suivantes pour les statistiques de test :

$$\begin{aligned}
W_n &= n\frac{(2\overline{X}_n - \overline{Y}_n)^2}{4\overline{X}_n + \overline{Y}_n} \\
R_n &= n\frac{(2\overline{X}_n - \overline{Y}_n)^2}{2(\overline{X}_n + \overline{Y}_n)} \\
T_n &= 2n\left\{\overline{X}_n \log\left(\frac{3\overline{X}_n}{\overline{X}_n + \overline{Y}_n}\right) + \overline{Y}_n \log\left(\frac{3\overline{Y}_n}{2(\overline{X}_n + \overline{Y}_n)}\right)\right\}
\end{aligned}$$

(b) Puisque la distribution considérée satisfait les hypothèses de normalité du MLE, on sait que sous $H_0$,

$$W_n, R_n, T_n \xrightarrow{d} \chi_1^2,$$

quand $n \to \infty$. Dès lors, pour un niveau $\alpha \in (0, 1)$ donné, les régions critiques des tests asymptotiques de niveau $\alpha$ basés sur ces statistiques sont toutes identiques et sont données par

$$\{x > 0 : x > \chi^2_{1;1-\alpha}\},$$

où $\chi^2_{1;1-\alpha}$ désigne le quantile d'ordre $1 - \alpha$ d'une distribution Chi-carrée $\chi_1^2$. Avec les données proposées, on calcule

$$\begin{aligned}
w_n &= 2.95 \\
r_n &= 2.77 \\
t_n &= 2.83,
\end{aligned}$$

et $\chi^2_{1;0.95} = 3.84$, ainsi, aucun des trois tests ne mène au rejet de l'hypothèse nulle. Cela n'est pas étonnant puisque les trois tests sont asymptotiquement équivalents, i.e., leur décision sera identique avec probabilité 1 quand $n \to \infty$.