LSTAT 2040 - TP 7

Tests d'hypothèse

Exercice 1

Soit $X_1, \ldots, X_n$ un échantillon iid de distribution commune $\text{Poi}(\lambda)$ où $\lambda > 0$. On souhaite réaliser un test du rapport de vraisemblance (LRT) exact pour les hypothèses

$$H_0 : \lambda \geq \lambda_0 \qquad \text{contre} \qquad H_1 : \lambda < \lambda_0,$$

où $\lambda_0 > 0$.

(a) Calculer la statistique Λ_n du LRT.

(b) Déduire du point précédent que rejeter H_0 pour $\Lambda_n \leq c$, où $c \in (0,1)$, revient à rejeter H_0 si $T_n := \sum_{i=1}^n X_i$ satisfait

$$T_n \leq a, \qquad 0 < a < n\lambda_0 < \infty.$$

Trouver la valeur de a pour que le test soit de niveau $\alpha \in (0,1)$.

(c) Effectuer le test pour $\alpha = 0.05$ et $\lambda_0 = 2$, en supposant que $n = 25$ et que $t_n :=$ valeur observée de $T_n = 38$.

(d) Si les hypothèses avaient été modifiées en $\tilde{H}_0 : \lambda = \lambda_0$ et $\tilde{H}_1 : \lambda \neq \lambda_0$, que devient la statistique du LRT et la région de rejet associée ? Que donne le test dans ce cas ?

Exercice 2

Soit $X_1, \ldots, X_n$ un échantillon iid de distribution commune $\text{Be}(\pi)$ où $\pi \in (0,1)$. On souhaite réaliser un LRT exact pour les hypothèses

$$H_0 : \pi \leq \pi_0 \qquad \text{contre} \qquad H_1 : \pi > \pi_0,$$

où $\pi_0 \in (0,1)$.

(a) Calculer la statistique de test Λ_n.

(b) Montrer que Λ_n peut être exprimée comme

$$\Lambda_n = \begin{cases} g(T_n), & \text{si } T_n \geq n\pi_0 \\ 1, & \text{si } T_n < n\pi_0 \end{cases},$$

où $T_n := \sum_{i=1}^n X_i$ et $g : [n\pi_0, n) \to (0,1]$ est une certaine fonction décroissante.

(c) Effectuer le test pour $\alpha = 0.05$ et $\pi_0 = 0.5$, en supposant que $n = 50$ et que $t_n :=$ valeur observée de $T_n = 29.5$.

(d) Le résultat du test effectué au point (c) aurait-il été différent si on avait eu $n = 100$ avec la même valeur de probabilité pour $\hat{\pi} = T_n/n$, c'est-à-dire, $t_n = 59$?

Exercice 3

Soient $X_1, \ldots, X_n$ et $Y_1, \ldots, Y_m$ deux échantillons iid et mutuellement indépendants provenant de distributions exponentielles de paramètre $\theta_X > 0$ et $\theta_Y > 0$, respectivement. On souhaite réaliser un LRT exact pour les hypothèses

$$H_0 : \theta_X = \theta_Y \qquad \text{contre} \qquad H_1 : \theta_X \neq \theta_Y.$$

(a) Montrer que la statistique de test $\Lambda_{n,m}$ peut être exprimée en fonction de la statistique

$$T_{n,m} := \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n X_i + \sum_{j=1}^m Y_j}.$$

(b) Déduire de votre calcul au point (a) que rejeter H_0 pour $\Lambda_{n,m} \leq c$, où $c \in (0,1)$, revient à rejeter H_0 si

$$T_{n,m} \leq a \qquad \text{ou} \qquad T_{n,m} \geq b, \qquad 0 < a < b < 1.$$

En utilisant le fait que sous H_0, on a $T_{n,m} \sim \mathrm{Beta}(n,m)$, trouver les valeurs de a et b pour que le test soit de niveau $\alpha \in (0,1)$.

(c) Effectuer le test avec $\alpha = 0.05$ pour des données telles que $t_{n,m} = 0.59$ où $n = 35$ et $m = 40$.

Exercice 4

Soit $X_1, \ldots, X_n$ un échantillon iid dont la distribution commune est une Pareto à deux paramètres avec fonction de densité donnée par

$$f(x;\theta,\nu) = \frac{\theta\nu^\theta}{x^{\theta+1}} \mathrm{I}(x \geq \nu),$$

où $\theta, \nu > 0$. On souhaite réaliser un LRT exact pour les hypothèses

$$H_0 : \theta = 1 \qquad \text{contre} \qquad H_1 : \theta \neq 1.$$

(a) Montrer que la statistique de test Λ_n peut être exprimée en fonction de la statistique

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

(b) Déduire de votre calcul au point (a) que rejeter H_0 pour $\Lambda_n \leq c$, où $c \in (0,1)$, revient à rejeter H_0 si

$$T_n \leq a \qquad \text{ou} \qquad T_n \geq b, \qquad 0 < a < b < \infty.$$

En utilisant le fait que sous H_0, on a $T_n \sim \mathrm{Gamma}(n-1,1)$, trouver les valeurs de a et b pour que le test soit de niveau $\alpha \in (0,1)$.

(c) Effectuer le test avec $\alpha = 0.05$ pour des données telles que $t_n = 55$ où $n = 25$.

Exercice 5

Soit $X_1, \ldots, X_n$ un échantillon iid de distribution commune $\mathrm{Beta}(\theta, 2)$ où $\theta > 0$. On souhaite effectuer un test asymptotique pour les hypothèses

$$H_0 : \theta = \theta_0 \qquad \text{contre} \qquad H_1 : \theta \neq \theta_0,$$

où $\theta_0 > 0$.

(a) Calculer les statistiques des tests asymptotiques de Wald, du Score et du LRT.

(b) Effectuer les différents tests avec $\alpha = 0.05$ et $\theta_0 = 2$ pour des données telles que $\sum_{i=1}^n \log(x_i) = -67.165$ où $n = 100$. Tous les tests donnent-ils la même conclusion ?

Exercice 6

Soit $X_1, \ldots, X_n$ un échantillon iid dont la densité de la distribution commune est donnée par

$$f(x;\theta) = \frac{2x}{\theta} \exp(-x^2/\theta) \mathrm{I}(x > 0),$$

où $\theta > 0$. On souhaite effectuer un test asymptotique pour les hypothèses

$$H_0 : \theta = \theta_0 \qquad \text{contre} \qquad H_1 : \theta \neq \theta_0,$$

où $\theta_0 > 0$.

(a) Calculer les statistiques des tests asymptotiques de Wald, du Score et du LRT.

(b) Effectuer les différents tests avec $\alpha = 0.05$ et $\theta_0 = 1$ pour des données telles que $\sum_{i=1}^n x_i^2 = 98.98$ où $n = 100$. Tous les tests donnent-ils la même conclusion ?

Exercice 7

Soient $X_1, \ldots, X_n$ et $Y_1, \ldots, Y_n$ deux échantillons iid et mutuellement indépendants provenant de distributions exponentielles de paramètre $\theta_X > 0$ et $\theta_Y > 0$, respectivement. On souhaite effectuer un test asymptotique pour les hypothèses

$$H_0 : (\theta_X, \theta_Y) = (\theta_X^0, \theta_Y^0) \qquad \text{contre} \qquad H_1 : (\theta_X, \theta_Y) \neq (\theta_X^0, \theta_Y^0),$$

où $\theta_X^0, \theta_Y^0 > 0$.

(a) Calculer les statistiques des tests asymptotiques de Wald, du Score et du LRT et donner les régions de rejet associées.

(b) Donner un avantage des tests basés sur les statistiques calculées en (a) pour tester l'hypothèse H_0 par rapport à la réalisation de deux tests indépendants pour tester $H_{0,X} : \theta_X = \theta_X^0$ et $H_{0,Y} : \theta_Y = \theta_Y^0$ si tous ces tests sont réalisés avec un même niveau $\alpha = 0.05$.

Exercice 8

Une étude sur les opinions des électeurs est menée dans trois quartiers ayant des politiques différentes afin de comparer la proportion des électeurs étant pour un certain candidat A. On interroge 200 électeurs dans chaque quartier et on obtient les résultats suivants :

Quartier	1	2	3	**Total**
Pour le candidat A	76	53	48	177
Pas pour le candidat A	124	147	152	423
Total	200	200	200	600

En supposant que les résultats des sondages dans les trois quartiers sont indépendants, réaliser les tests asymptotiques de Wald et du LRT afin de tester l'égalité des proportions d'électeurs en faveur du candidat A dans les trois quartiers. Utiliser $\alpha = 0.05$ pour effectuer votre test. De plus, calculer les p-valeurs asymptotiques de ces deux tests.

Exercice 9

Soient $X_1, \ldots, X_n$ et $Y_1, \ldots, Y_n$ deux échantillons iid et mutuellement indépendants provenant de distributions $\mathrm{N}(\mu_X, \sigma^2)$ et $\mathrm{N}(\mu_Y, \sigma^2)$ respectivement, avec $\mu_X, \mu_Y \in \mathbb{R}$. Trouver le test de Wald et le LRT asymptotique pour tester les hypothèses

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

où $\mu_X^0, \mu_Y^0 \in \mathbb{R}$ au niveau $\alpha \in (0, 1)$.

Exercice 10

Soient $X_1, \ldots, X_n$ et $Y_1, \ldots, Y_n$ deux échantillons iid et mutuellement indépendants provenant de distributions de Poisson de paramètre $\lambda_X > 0$ et $\lambda_Y > 0$, respectivement. On souhaite effectuer un test asymptotique pour les hypothèses

$$H_0 : \lambda_Y = 2\lambda_X \qquad \text{contre} \qquad H_1 : \lambda_Y \neq 2\lambda_X.$$

(a) Calculer les statistiques des tests asymptotiques de Wald, du Score et du LRT.

(b) Effectuer les différents tests avec $\alpha = 0.05$ pour des données telles que $\sum_{i=1}^n x_i = 102$ et $\sum_{i=1}^n y_i = 248$ où $n = 100$. Tous les tests donnent-ils la même conclusion ?