Faculté des Sc. Exactes et Sc. de la Nature et la Vie
Département de Mathématiques
Master 1: 2023/2024
Module: tests statistiques

Solution de l'examen

Exercice 1. La loi exponentielle de paramètre $\lambda > 0$ est définie par $\mathbf{P}(X \geq x) = \exp(-\lambda x)$, $x > 0$. Au niveau de signification 5%, et sur la base d'un échantillon, de la v.a. X, de taille 100 et de moyenne $\overline{x} = 0.457$, on s'intéresse à tester les deux hypothèses suivantes: $H_0 : \lambda = 2$ contre $H_1 : \lambda \neq 2$. Que peut-on en conclure?

* * * * * * *

Exercice 2. Etant donné un échantillon de taille 16, d'une v.a. $X \rightsquigarrow \mathcal{N}(0, \sigma^2)$, d'un écart-type $\widetilde{s} = 0.828$. Au niveau de signification 5%, on veut tester les deux hypothèses suivantes: $H_0 : \sigma^2 = 1$ contre $H_1 : \sigma^2 \neq 1$. Que peut-on en conclure?

* * * * * * * * *

Solution de l'exercice 1(10pts). Tout d'abort on note que la densité de probabilité associée à la v.a. X est $f_\lambda(x) = \lambda \exp(-\lambda x)$. Le rapport de vraisemblance généralisé (ou maximal) associé à ce test est:

$$\begin{aligned}
R_1 &= \frac{\sup_{\lambda>0} L(x_1, ..., x_n; \lambda)}{L(x_1, ..., x_n; 2)} \\
&= \frac{L(x_1, ..., x_n; \lambda_{EMV})}{L(x_1, ..., x_n; 2)} = \frac{L(x_1, ..., x_n; 1/\overline{x})}{L(x_1, ..., x_n; 2)} \\
&= \frac{\prod_{i=1}^{n} \frac{1}{\overline{x}} \exp\left(-\frac{1}{\overline{x}} x_i\right)}{\prod_{i=1}^{n} 2\exp(-2x_i)} = \frac{\frac{1}{(\overline{x})^n} \exp\left(-\frac{1}{\overline{x}} \sum_{i=1}^n x_i\right)}{2^n \exp\left(-2\sum_{i=1}^n x_i\right)},
\end{aligned}$$

avec $n = 100$. Nous avons $\sum_{i=1}^n x_i = n\overline{x}$, donc

$$R_1 = \frac{1}{(2\overline{x})^n} \frac{\exp(-n)}{\exp(-2n\overline{x})} = \left(e^{-1} \frac{1}{2\overline{x}} \exp(2\overline{x})\right)^n.$$

La région critique de ce test est:

$$W = \left\{(x_1, ..., x_{100}) \in (\mathbb{R}_+ / \{0\})^{100} : \left(e^{-1} \frac{1}{2\overline{x}} \exp(2\overline{x})\right)^{100} > c\right\},$$

où

$$\mathbf{P}_{\lambda=2}\left(\left(e^{-1} \frac{1}{2\overline{x}} \exp(2\overline{x})\right)^{100} > c\right) = 0.05.$$

Ceci peut être simplifiée en

$$W = \left\{(x_1, ..., x_{100}) \in (\mathbb{R}_+ / \{0\})^{100} : \frac{1}{2\overline{x}} \exp 2\overline{x} > k\right\},$$

pour une certaine $k > 0$, où

$$\mathbf{P}_{\lambda=2}\left(\frac{1}{2\overline{X}} \exp 2\overline{X} > k\right) = 0.05.$$

Nous avons déjà noté en cours que

$$\forall t > 0, \ \frac{1}{t} \exp t > k \Longleftrightarrow \exists\, (0 < \eta_1 < \eta_2) \text{ telles que } t > \eta_1 \text{ ou } t < \eta_2.$$

Ce qui implique (vue la continuité de la v.a.) que

$$W=\left\{(x_1,...,x_{100})\in(\mathbb{R}_+/\{0\})^{100}:\ \overline{x}\geq c_1 \text{ ou } \overline{x}\leq c_2\right\},$$

où

$$\mathbf{P}_{\lambda=2}\left(\overline{X}\geq c_1\right)=\mathbf{P}_{\lambda=2}\left(\overline{X}\leq c_2\right)=0.05/2=0.025.$$

Comme $n=100>30$, on peut alors appliquer le théorème centrale limite. En effet, noter que $\mathbf{E}[X]=1/\lambda$ et $\mathbf{Var}[X]=1/\lambda^2$ donc sous $H_0:\lambda=2$, on a $\mathbf{E}[X]=1/2$ et $\mathbf{Var}[X]=1/4$, ainsi

$$\mathbf{P}_{\lambda=2}\left(\sqrt{100}\frac{\overline{X}-1/2}{1/2}\geq\rho_1\right)\simeq\mathbf{P}(Z\geq\rho_1)=0.025,$$

où

$$\rho_1:=\sqrt{100}\frac{c_1-1/2}{1/2} \text{ et } Z\rightsquigarrow\mathcal{N}(0,1).$$

Il claire que $\rho_1=\Phi^{-1}(1-0.025)=\Phi^{-1}(0.975)\simeq1.96$, ce qui implique que

$$\sqrt{100}\frac{c_1-1/2}{1/2}=1.96\Longleftrightarrow c_1=0.598.$$

De même

$$\mathbf{P}_{\lambda=2}\left(\sqrt{100}\frac{\overline{X}-1/2}{1/2}\leq\rho_2\right)\simeq\mathbf{P}(Z\leq\rho_2)=0.025,$$

où

$$\rho_2:=\sqrt{100}\frac{c_2-1/2}{1/2}.$$

Il claire que $\rho_2=\Phi^{-1}(0.025)=-\Phi^{-1}(1-0.025)=-\Phi^{-1}(0.975)\simeq-1.96$, ce qui implique que

$$\sqrt{100}\frac{c_2-1/2}{1/2}=-1.96\Longleftrightarrow c_2=0.402.$$

Finalement, la région critique est

$$W=\left\{(x_1,...,x_{100})\in(\mathbb{R}_+/\{0\})^{100}:\overline{x}\geq0.598 \text{ ou } \overline{x}\leq0.402\right\}.$$

La moyenne des données observées est égale à $\overline{x}=0.457$ qui est ni ≥0.598 ni ≤0.402, donc en garde H_0, c'est à dire $\lambda=2$.

* * * * **

Solution de l'exercice 2 (10pts). Il s'agit du test de la variance

$$\begin{cases} H_0: & \sigma^2=1 \\ H_1: & \sigma^2\neq1 \end{cases},$$

d'une moyenne "connue $\mu=0$". Nous avons aussi annoncé au cours la région critique de ce test est donnée par

$$W=\left\{(x_1,...,x_{16})\in\mathbb{R}_+^{16}\mid\frac{16v^2}{1}\geq k_1 \text{ ou } \frac{16v^2}{1}\leq k_2\right\},$$

où $v^2=\frac{1}{16}\sum_{i=1}^{16}(x_i-0)^2$, et k_1 et k_2 sont données par

$$\mathbf{P}\left(\chi^2_{16}\geq k_1\right)=\mathbf{P}\left(\chi^2_{16}\leq k_1\right)=0.05/2=0.025,$$

où χ^2_{16} désigne la variable de Chi-deux (ou de Pearson) à 16 degrés de liberté. De la table statistique on obtient $k_1=28.8454$ et $k_2=6.9077$ ainsi

$$W=\left\{(x_1,...,x_{16})\in\mathbb{R}_+^{16}\mid16v^2\geq28.8454 \text{ ou } 16v^2\leq6.9077\right\}.$$

La fonction statistique du test correspondante est

$$\delta(x_1,...,x_{16})=\begin{cases}1 & \text{si } v^2\geq1.8028 \text{ ou } v^2\leq0.43173\\ 0 & \text{sinon}\end{cases}.$$

On note que comme $\mu=0$, et vu "la nécessité" on peut admettre que sont estimateur $\overline{X}$ est "presque nulle" ansi

$$v^2_{obs}=\frac{1}{16}\sum_{i=1}^{16}(x_i-0)^2\simeq\widetilde{s}^2_{obs}=(0.828)^2=0.68558.$$

La valeur du $v^2_{obs}=0.68558$ est ni ≥1.8028 ni ≤0.43173, alors on garde H_0; c'est à dire $\sigma^2=1$.