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

Soultion de l'Interrogation 1

Soit $(X_1, ..., X_{35})$ un échantillon d'une population X normale centrée de variance σ^2.

On s'intéresse à tester les deux hypothèses suivantes:

$$\begin{cases} H_0: & \sigma^2 \geq 4 \\ H_1: & \sigma^2 < 4 \end{cases}.$$

1- Montrer que le rapport de vraisemblance associé à l'échantillon ci-dessus est décroisant par rapport à une statistique $t = t(x_1, ..., x_{35})$ qu'on demande à la déterminer.
2-Construire le test uniformément le plus puissant, au niveau de signification $\alpha = 0.05$, associé à cette statistique.
3-Déterminer le risque du premier espèce et le risque de deuxième espèce.
4. Déterminer la dimension de ce test.
5. Tracer le graphe de la fonction puissance.
Remarque: $\mathbf{E}[Z] = 0 = \mathbf{E}[Z^3]$, $\mathbf{E}[Z^2] = 1$, et $\mathbf{E}[Z^4] = 3$, où $Z \rightsquigarrow N(0,1)$.

* * * * * * * * * * *

Solution. **1)** Nous allons appliquer la méthode du test de *rapport de vraisemblance monotone (croissant ou décroissant)*. Soit $\sigma_1 > \sigma_2$ et écrivons

$$\frac{L_{\sigma_1}}{L_{\sigma_2}} = \frac{\prod_{i=1}^{35} \frac{1}{\sqrt{2\pi\sigma_1^2}} \exp\left\{-\frac{1}{2}\left(\frac{x_i}{\sigma_1}\right)^2\right\}}{\prod_{i=1}^{35} \frac{1}{\sqrt{2\pi\sigma_2^2}} \exp\left\{-\frac{1}{2}\left(\frac{x_i}{\sigma_2}\right)^2\right\}} = \left(\frac{\sigma_2}{\sigma_1}\right)^{35} \exp\left\{\frac{1}{2}\left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}\right)\sum_{i=1}^{35} x_i^2\right\}.$$

On pose $t := -\sum_{i=1}^{35} x_i^2$, $b := \frac{1}{2}\left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}\right)$ et $d := \left(\frac{\sigma_2}{\sigma_1}\right)^{35} > 0$. Il est claire que la fonction $t \to \frac{L_{\sigma_1}}{L_{\sigma_2}}(t) = d\exp(-bt)$ est décroissante en t, (car $\sigma_1 > \sigma_2$ implique $b > 0$). Alors la loi de la v.a, X possède un rapport de vraisemblance décroissant en $t = -\sum_{i=1}^{35} x_i^2$.

2) En appliquant la proposition 3, le test upp est

$$\delta = \begin{cases} 1 & \text{si } -\sum_{i=1}^{35} x_i^2 \geq c \\ 0 & \text{si } -\sum_{i=1}^{35} x_i^2 < c \end{cases}$$

ce qui est équivalent à

$$\delta = \begin{cases} 1 & \text{si } \sum_{i=1}^{35} x_i^2 \leq k \\ 0 & \text{si } \sum_{i=1}^{35} x_i^2 > k \end{cases}$$

où $k = -c$. La constante k est définie par $\mathbf{P}_{\sigma^2=4}\left(\sum_{i=1}^{35} X_i^2 \leq k\right) = 0.05$. Notons que $X_i \rightsquigarrow \mathcal{N}(0,4)$ alors la v.a. $\sum_{i=1}^{35}(X_i/2)^2 \rightsquigarrow \chi^2_{35}$, ce pendant la table statistique 3 offre que des quantiles pour des degrés de liberté inferieurs 30.

Comme $n=35>30$, on peut appliquer le théorème centrale limite (TCL). Ecrivons

$$\begin{aligned}0.05&=\mathbf{P}_{\sigma^2=4}\left(\sum_{i=1}^{35}X_i^2\leq k\right)\\&=\mathbf{P}_{\sigma^2=4}\left(\sqrt{35}\frac{\frac{1}{35}\sum_{i=1}^{35}X_i^2-\mathbf{E}\left[X_1^2\right]}{\sqrt{\mathbf{Var}\left[X_1^2\right]}}\leq\sqrt{35}\frac{\frac{k}{35}-\mathbf{E}\left[X_1^2\right]}{\sqrt{\mathbf{Var}\left[X_1^2\right]}}\right).\end{aligned}$$

Nous avons $\mathbf{E}\left[X_1\right]=0$ donc $\mathbf{E}\left[X_1^2\right]=\mathbf{Var}\left[X_1^2\right]=4$. D'autre par

$$\begin{aligned}\mathbf{Var}\left[X_1^2\right]&=\mathbf{E}\left[X_1^4\right]-\left(\mathbf{E}\left[X_1^2\right]\right)^2\\&=\mathbf{E}\left[X_1^4\right]-16.\end{aligned}$$

Comme $X_1\rightsquigarrow\mathcal{N}(0,4)$ alors $X_1=2Z$ où $Z\rightsquigarrow N(0,1)$, ainsi $\mathbf{E}\left[X_1^4\right]=\mathbf{E}\left[2Z\right]^4=2^4\mathbf{E}\left[Z^4\right]=16\times\mathbf{E}\left[Z^4\right]=16\times3=48$, donc $\mathbf{Var}\left[X_1^2\right]=48-16=32$. Par conséquent

$$\begin{aligned}0.05&=\mathbf{P}_{\sigma^2=4}\left(\sum_{i=1}^{35}X_i^2\leq k\right)\\&=\mathbf{P}_{\sigma^2=4}\left(\sqrt{35}\frac{\frac{1}{35}\sum_{i=1}^{35}X_i^2-4}{\sqrt{32}}\leq\sqrt{35}\frac{\frac{k}{35}-4}{\sqrt{32}}\right).\end{aligned}$$

D'après le TCL on a

$$Z^*=:\sqrt{35}\frac{\frac{1}{35}\sum_{i=1}^{35}X_i^2-4}{\sqrt{32}}\text{ suit approximativement la loi }\mathcal{N}(0,1).$$

Le fait que $\mathbf{P}\left(Z^*\leq\sqrt{35}\frac{\frac{k}{35}-4}{\sqrt{32}}\right)=0.05$ implique que $\sqrt{35}\frac{\frac{k}{35}-4}{\sqrt{32}}=\Phi^{-1}(0.05)=-\Phi^{-1}(1-0.05)=-\Phi^{-1}(0.95)=-1.64$, ce qui implique que $k=85.115$. Ainsi le test uniformément le plus puissant, au niveau de signification $\alpha=0.05$ est

$$\delta^{(TCL)}(x_1,...,x_{35})=\begin{cases}1 & \text{si }\sum_{i=1}^{35}x_i^2\leq 85.115\\ 0 & \text{si }\sum_{i=1}^{35}x_i^2>85.115\end{cases}$$

On peut aussi appliquer l'approximation de Fisher de la loi de khi-deux qui ce trouve à la fin de la page 3 des tables statistiques:

$$\sqrt{2\chi_\nu^2}-\sqrt{2v-1}\rightsquigarrow\mathcal{N}(0,1),\text{ pour }\nu>30.$$

Ecrivons

$$\begin{aligned}0.05&=\mathbf{P}_{\sigma^2=4}\left(\sum_{i=1}^{35}X_i^2\leq k\right)=\mathbf{P}\left(\sum_{i=1}^{35}\left(X_i/2\right)^2\leq k/4\right)\\&=\mathbf{P}\left(\chi_{35}^2\leq k/4\right)\\&=\mathbf{P}\left(\sqrt{2\chi_{35}^2}-\sqrt{2\times35-1}\leq\sqrt{2\left(k/4\right)}-\sqrt{2\times35-1}\right)\\&\simeq\mathbf{P}\left(Z\leq\sqrt{k/2}-\sqrt{69}\right),\end{aligned}$$

donc $\sqrt{k/2}-\sqrt{69}=-1.64 \Rightarrow k=88.888$. Ainsi le test uniformément le plus puissant, au niveau de signification $\alpha=0.05$ est

$$\delta^{(Fisher)}(x_1,...,x_{35})=\begin{cases} 1 & \text{si } \sum_{i=1}^{35} x_i^2 \leq 88.888 \\ 0 & \text{si } \sum_{i=1}^{35} x_i^2 > 88.888 \end{cases}$$

On note que, pour les petits échantillons, l'approximation de Fisher est plus précise que le TCL. Pour que cette dernière atteint la précision de Fisher on doit avoir $n>100$. Par la suite nous allons considérer le test $\delta^{(Fisher)}(x_1,...,x_{35})$.

3) Le risque de première espèce et celui de deuxième espèce sont respectivement définies par

$$\alpha(\sigma):=\mathbf{P}_{\sigma^2}\left(\sum_{i=1}^{35} X_i^2 \leq 88.888\right), \text{ pour } \sigma^2 \geq 4,$$

et

$$\beta(\sigma):=1-\mathbf{P}_{\sigma^2}\left(\sum_{i=1}^{35} X_i^2 \leq 88.888\right), \text{ pour } \sigma^2 < 4.$$

4) La dimension du test est définie par $\sup_{\sigma^2\geq 4}\alpha(\sigma)$ qui est égale au seuil de signification $\alpha=0.05$.

5) Nous traçons le graphe au TD.