Université Mohamed Khider, Biskra
Faculté des Sc. Exactes et Sc. de la Nature et la Vie
Département de Mathématiques
Master 1: 2021/2022

Solution de l'exercice 5 de la Série N°3

Exercice 5. On veut tester l'égalité des variances de deux populations $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ et $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ au niveau de signification $\alpha = 0.05$. Un échantillon de taille 16 de X_1 donne une variance empirique $\widetilde{s}_{1,obs}^2 = 7.62$ et un échantillon de taille 12 de X_2 donne une variance empirique $\widetilde{s}_{2,obs}^2 = 3.96$. Conclure.

****

Solution. Il s'agit ici d'un test (bilatéral) de comparaison entre deux variances:

$$\begin{cases} H_0: & \sigma_1^2 = \sigma_2^2 \\ H_0: & \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

Nous avons $\alpha = 0.05$, $n_1 = 16$, $n_2 = 12$, $\widetilde{s}_{1,obs}^2 = 7.62$, $\widetilde{s}_{2,obs}^2 = 3.96$. La statistique de test (variable de décision) à utiliser est:

$$\frac{\frac{16\widetilde{S}_1^2/\sigma_1^2}{15}}{\frac{12\widetilde{S}_2^2/\sigma_2^2}{11}} \rightsquigarrow F(15, 11), \quad \left(\text{sous } H_0 : \sigma_1^2 = \sigma_2^2\right), \tag{1}$$

où $F(15, 11)$ désigne la loi de Fisher à $(15, 11)$ degrés de liberté. A un seuil de signification $\alpha = 5\%$, la région critique associée est

$$W = \left\{ \left(x_1^{(1)}, ..., x_1^{(16)}; x_2^{(1)}, ..., x_2^{(12)} \right) \in \mathbb{R}^{28} \mid \right.$$
$$\left. \frac{\widetilde{s}_1^2}{\widetilde{s}_2^2} \geq \frac{12 \times 15}{16 \times 11} c_1 \text{ ou } \frac{\widetilde{s}_1^2}{\widetilde{s}_2^2} \leq \frac{12 \times 15}{16 \times 11} c_2 \right\},$$

où c_1 et c_2 sont deux constantes telles que

$$\mathbf{P}\left(F(15, 11) \geq c_1\right) = \mathbf{P}\left(F(15, 11) \leq c_2\right) = \alpha/2 = 0.025.$$

Donc c_1 est le quantile d'ordre $1 - 0.025 = 0.975$ et c_2 est le quantile d'ordre 0.025 de la loi Fisher à $(15, 11)$ degrés de liberté. De la table statistique de Fisher on obtient $c_1 = 3.40$. Pour trouver la valeur de c_2, on utilise la formule suivante:

$$F(\nu_1, \nu_2) \rightsquigarrow \frac{1}{F(\nu_2, \nu_1)}.$$

Donc

$$\mathbf{P}\left(F(15, 11) \leq c_2\right) = 0.025 \Longleftrightarrow \mathbf{P}\left(F(11, 15) \geq 1/c_2\right) = 0.025.$$

De la table statistique on obtient $1/c_2 =$qui donne $c_2 = 0.33$. La région critique est donc

$$W = \left\{ \left(x_1^{(1)}, ..., x_1^{(16)}; x_2^{(1)}, ..., x_2^{(12)} \right) \in \mathbb{R}^{28} \mid \right.$$
$$\left. \frac{\widetilde{s}_1^2}{\widetilde{s}_2^2} \geq 3.40 \text{ ou } \frac{\widetilde{s}_1^2}{\widetilde{s}_2^2} \leq 0.33 \right\},$$

Nous avons

$$\frac{\widetilde{s}_{1,obs}^2}{\widetilde{s}_{2,obs}^2} = \frac{7.62}{3.96} = 1.92.$$

Cette valeur est ni ≥ 3.40 ni ≤ 0.33, donc on accepte l'égalité des deux variances. La p-value dans notre cas est égale à

$$\begin{aligned} p - value &= 2 \min \left\{ \mathbf{P}\left(\frac{\widetilde{S}_1^2}{\widetilde{S}_2^2} \geq 1.92 \right), \mathbf{P}\left(\frac{\widetilde{S}_1^2}{\widetilde{S}_2^2} \leq 1.92 \right) \right\} \\ &= 2 \min \left\{ \mathbf{P}\left(\frac{\frac{16\widetilde{S}_1^2/\sigma_1^2}{15}}{\frac{12\widetilde{S}_2^2/\sigma_2^2}{11}} \geq \frac{\frac{16/\sigma_1^2}{15}}{\frac{12/\sigma_2^2}{11}} 1.92 \right), \mathbf{P}\left(\frac{\frac{16\widetilde{S}_1^2/\sigma_1^2}{15}}{\frac{12\widetilde{S}_2^2/\sigma_2^2}{11}} \leq \frac{\frac{16/\sigma_1^2}{15}}{\frac{12/\sigma_2^2}{11}} 1.92 \right) \right\} \\ &= 2 \min \left\{ \mathbf{P}\left(F(15, 11) \geq \frac{\frac{16}{15}}{\frac{12}{11}} 1.92 \right), \mathbf{P}\left(F(15, 11) \leq \frac{\frac{16}{15}}{\frac{12}{11}} 1.92 \right) \right\}, \end{aligned}$$

car sous H_0 on a $\sigma_1^2 = \sigma_2^2$. Donc

$$p - value = 2 \min \{ \mathbf{P} \left(F \left(15, 11 \right) \geq 1.87 \right), \mathbf{P} \left(F \left(15, 11 \right) \leq 1.87 \right) \}.$$

En utilisant le langage R on obient $\mathbf{P} \left(F \left(15, 11 \right) \leq 1.87 \right) = 0.85$, ainsi

$$p - value = 2 \min \{ 1 - 0.85, 0.85 \} = 0.3 > 0.05,$$

donc, effet l'égalités des deux variances est encore une fois confirmée.