Statistiques L2Bio, M.Boutahar

Tests d'hypothèses

Pour tous les tests qui suivent, T désigne la statistique de décision et $\mathcal{R}$ la zone de rejet, c'est à dire que si $T \in \mathcal{R}$ alors on rejette l'hypothèse nulle H_0 avec un risque d'erreur α.
$N(0,1)$: loi normale centrée réduite (loi de Gauss).
$\mathcal{X}^2(n)$: loi de Khi-deux à n degrés de liberté.
$t(n-1)$: loi de Student à $(n-1)$ degrés de liberté.
$F(n_1, n_2)$: loi de Fisher à (n_1, n_2) degrés de liberté.

1.Test d'indépendance de $\mathcal{X}^2$

$$\begin{cases} H_0: \ X \text{ est indépendante de } Y \text{ (hypothèse nulle)} \\ \text{contre} \\ H_1: X \text{ dépends de } Y \text{ (hypothèse alternative)} \end{cases}$$

$$T = \sum_{i=1}^{p}\sum_{j=1}^{r} \frac{\left(n_{i,j} - \frac{n_{i,\bullet}n_{\bullet,j}}{n}\right)^2}{\frac{n_{i,\bullet}n_{\bullet,j}}{n}},$$

$n_{i,\bullet} = \sum_{j=1}^{r} n_{i,j}, n_{\bullet,j} = \sum_{i=1}^{p} n_{i,j}, n = \sum_{i=1}^{p}\sum_{j=1}^{r} n_{i,j}$

$$\mathcal{R} = [x, +\infty[, P(\mathcal{X}^2((p-1)(r-1)) \leq x) = 1-\alpha.$$

2.Test de conformité d'une proportion

$$\begin{cases} H_0: \quad p \ = p_0 \\ \text{contre} \\ H_1: \quad p \ \neq p_0 \end{cases}$$

$f = \frac{n_1}{n}$: proportion des individus de la catégorie A, $v_0 = \frac{p_0(1-p_0)}{n}$,

$$T = \frac{f - p_0}{\sqrt{v_0}},$$

$$\mathcal{R} =]-\infty - u]\bigcup[u, +\infty[, P(N(0,1) \in [-u, u]) = 1-\alpha.$$

3. Test bilatéral de conformité d'une moyenne

$$\begin{cases} H_0: \quad m \ = m_0 \\ \text{contre} \\ H_1: \quad m \ \neq m_0 \end{cases}$$

• Test bilatéral de conformité d'une moyenne avec variance connue

$$T = \sqrt{n}\frac{\overline{X} - m_0}{\sigma}, \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

$$\mathcal{R} =]-\infty - u]\bigcup[u, +\infty[, P(N(0,1) \in [-u, u]) = 1-\alpha.$$

• Test bilatéral de conformité d'une moyenne avec variance inconnue

$$T = \sqrt{n}\frac{\overline{X} - m_0}{S'_n}, S'^2_n = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2,$$

$$\mathcal{R} =]-\infty, -t]\bigcup[t, +\infty[, P(t(n-1) \in [-t, t]) = 1-\alpha.$$

5. Test unilatéral de conformité d'une moyenne avec variance inconnue

$$\begin{cases} H_0: \quad m \ \leq m_0 \\ \text{contre} \\ H_1: \quad m \ > m_0 \end{cases}$$

$$T = \sqrt{n}\frac{\overline{X} - m_0}{S'_n}, \mathcal{R} = [t', +\infty[, P(t(n-1) \leq t') = 1-\alpha.$$

6. Test bilatéral de conformité d'une variance

$$\begin{cases} H_0: \sigma^2 = \sigma_0^2 \\ \text{contre} \\ H_1: \sigma^2 \neq \sigma_0^2 \end{cases}$$

$T = \frac{(n-1)S'^2_n}{\sigma_0^2}, \mathcal{R} = [0, a]\bigcup[b, +\infty[,$

$$P(\mathcal{X}^2(n-1) < a) = \alpha/2, P(\mathcal{X}^2(n-1) > b) = \alpha/2.$$

7. Test unilatéral de conformité d'une variance

$$\begin{cases} H_0: \sigma^2 \leq \sigma_0^2 \\ \text{contre} \\ H_1: \sigma^2 > \sigma_0^2 \end{cases}$$

$T = \frac{(n-1)S'^2_n}{\sigma_0^2}, \mathcal{R} = [b, +\infty[, P(\mathcal{X}^2(n-1) < b) = 1-\alpha.$

8. Test bilatéral de comparaison de deux proportions

$f_1 = \frac{k_1}{n_1}$ et $f_2 = \frac{k_2}{n_2}$ les proportions des individus de la catégorie A dans les échantillons E_1 et E_2.

$$\begin{cases} H_0: p_1 = p_2 = p_0 \\ \text{contre} \\ H_1: \text{ non } H_0 \end{cases}$$

$$T = \frac{f_1 - f_2}{\widehat{\sigma}_D}, \widehat{\sigma}_D^2 = \hat{p}_0(1-\hat{p}_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right),$$

$$\hat{p}_0 = \frac{n_1 f_1 + n_2 f_2}{n_1 + n_2},$$

$$\mathcal{R} =]-\infty, -u]\bigcup[u, +\infty[, P(N(0,1) \in [-u, u]) = 1-\alpha.$$

9. Test bilatéral de comparaison des moyennes de deux échantillons indépendants

$$\begin{cases} H_0: \quad m_1 = m_2 \\ \text{contre} \\ H_1: \quad m_1 \neq m_2 \end{cases}$$

• Cas σ_1^2 et σ_2^2 connues:

$$T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}, \quad \overline{X}_k = \frac{1}{n_k}\sum_{i=1}^{n_k} X_{k,i}, k = 1, 2, \qquad (1)$$

$\mathcal{R} =]-\infty, -u] \bigcup [u, +\infty, P(N(0,1) \in [-u, u]) = 1 - \alpha$.
• Cas σ_1^2 et σ_2^2 inconnues et les échantillons sont grands:

$n_1 \geq 30$ et $n_2 \geq 30$: On remplace dans (1) σ_1^2 (resp. σ_2^2) par les variances empiriques modifiées $S_1'^2 = \frac{1}{n_1-1}\sum_1^{n_1}(X_{1,i} - \overline{X}_1)^2$ (resp. $S_2'^2$).
• Cas σ_1^2 et σ_2^2 inconnues, les échantillons sont petits et $\sigma_1^2 = \sigma_2^2 = \sigma^2$:
$T = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{f(1,2)}}$,

$$f(1,2) = \frac{(n_1-1)S_1'^2 + (n_2-1)S_2'^2}{n_1+n_2-2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right),$$

$\mathcal{R} =]-\infty, -t] \bigcup [t, +\infty[, P(t(n_1+n_2-2) \in [-t, t]) = 1 - \alpha$,
• Test de Mann-Whitney Wilcoxon (cas σ_1^2 et σ_2^2 inconnues et différentes et les échantillons sont petits):

$$U_{1,2} = \sum_{i=1}^{n_1}\sum_{j=1}^{n_2} \mathbf{1}_{\{X_{1,i} \leq X_{2,j}\}},$$

on compte pour chaque $X_{1,i}, i = 1, ..., n_1$ le nombre de $X_{2,j}, j = 1, .., n_2$ qui lui sont supérieurs et on somme les résultats obtenus pour tous les $X_{1,i}$.

$$T = \frac{U_{1,2} - \frac{n_1 n_2}{2}}{\sqrt{n_1 n_2(n_1+n_2+1)/12}},$$

$\mathcal{R} =]-\infty, -u] \bigcup [u, +\infty[, P(N(0,1) \in [-u, u]) = 1 - \alpha$.

10. Test bilatéral de comparaison des moyennes de deux échantillons appariées

$T = \sqrt{n}\frac{\overline{D}}{S_D}, D_i = X_{1,i} - X_{2,i}$,
$S_D'^2 = \frac{1}{n-1}\sum_1^n (D_i - \overline{D})^2, \overline{D} = \frac{1}{n}\sum_{i=1}^n D_i$.
• Si $n \geq 30$, alors T est assimilée à une variable centrée réduite de Gauss.
• Si $n < 30$, alors T suit une loi de Student à $(n-1)$ degrés de liberté.

11. Test bilatéral de comparaison des variances de deux échantillons indépendants

$$\begin{cases} H_0: & \sigma_1^2 = \sigma_2^2 \\ \text{contre} & \\ H_1: & \sigma_1^2 \neq \sigma_2^2 \end{cases}$$

• Si $\frac{S_1'^2}{S_2'^2} > 1$ on pends $T = \frac{S_1'^2}{S_2'^2}$,
$\mathcal{R} = [f, +\infty[, P(F(n_1-1, n_2-1) > f) = \alpha/2$.
• Si $\frac{S_2'^2}{S_1'^2} > 1$ on prends $T = \frac{S_2'^2}{S_1'^2}$,
$\mathcal{R} = [f, +\infty[, P(F(n_2-1, n_1-1) > f) = \alpha/2$.

12. Test unilatéral de comparaison des variances de deux échantillons indépendants

$$\begin{cases} H_0: & \sigma_1^2 \leq \sigma_2^2 \\ \text{contre} & \\ H_1: & \sigma_1^2 > \sigma_2^2 \end{cases}$$

$T = \frac{S_1'^2}{S_2'^2}$,
$\mathcal{R} = [f, +\infty[, P(F(n_1-1, n_2-1) > f) = \alpha$.

13. Test de comparaison simultanée de plusieurs moyennes

$$\begin{aligned} S^2_{\text{intra}} &= \frac{1}{n-k}\sum_{i=1}^k \nu_i S_i'^2, \nu_i = n_i - 1, \\ S_i'^2 &= \frac{1}{n_i-1}\sum_{j=1}^{n_i}(X_{i,j} - \overline{X}_i)^2, \overline{X}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} X_{i,j}. \end{aligned}$$

$$\begin{aligned} S^2_{\text{inter}} &= \frac{1}{k-1}\sum_{i=1}^k n_i(\overline{X}_i - \overline{X})^2 , \\ \overline{X} &= \frac{1}{n}\sum_{i=1}^k\sum_{j=1}^{n_i} X_{i,j}, \quad n = \sum_{i=1}^k n_i. \end{aligned}$$

$$\begin{cases} H_0: & m_1 = m_2 = ... = m_k \\ \text{contre} & \\ H_1: & \exists i \neq j \text{ tels que } m_i \neq m_j \end{cases}$$

$T = \frac{S^2_{\text{inter}}}{S^2_{\text{intra}}}$, $\mathcal{R} = [f, +\infty[, P(F(k-1, n-k) > f) = \alpha$.

14. Test de comparaison simultanée de plusieurs variances (Test de Bartlett)

$$\begin{cases} H_0: & \sigma_1^2 = \sigma_2^2 = ... = \sigma_k^2 \\ \text{contre} & \\ H_1: & \exists i \neq j \text{ tels que } \sigma_i^2 \neq \sigma_j^2 \end{cases}$$

$$\begin{aligned} T &= \frac{1}{\lambda}\left\{(n-k)\ln S^2_{\text{intra}} - \sum_{j=1}^k \nu_i \ln S_i'^2\right\}, \\ \lambda &= 1 + \frac{1}{3(k-1)}\left\{\frac{1}{\nu_1} + \frac{1}{\nu_2} .... + \frac{1}{\nu_k} - \frac{1}{n-k}\right\}, \end{aligned}$$

$\nu_k = n_k - 1, \mathcal{R} = [\mathbf{x}, +\infty[, P(\mathcal{X}^2(k-1) \leq \mathbf{x}) = 1 - \alpha$.