11 Test d'Homogénéité

vendredi 9 décembre 2022

1. Objectif de test d'homogénéité 同质性
   - -1- Valider l'homogénéité des deux échantillons
     1- Espérance, médiane
     2- Variance
     3- Probabilité de succès
     4- Loi
2. Deux type d'échantillons
   - -1- Échantillons indépendants
     1- Homogénéité des espérance
     2- Homogénéité des variances
     3- Homogénéité des proportions
     4- Homogénéité des lois
   - -2- Échantillons appariés $(\mathrm{X_i}, Y_i)$
3. Échantillons indépendants
   - -1- Homogénéité des espérance (Gaussien)
     1- Modélisation $X_i \sim N(\mu_X, \sigma_X^2), Y_j \sim N(\mu_Y, \sigma_Y^2), N = n + m$
     2- $V = \bar{X} - \bar{Y} \sim N\left(\mu_V = \mu_X - \mu_Y, \sigma_V^2 = \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}\right)$
     3- Test bilatéral
        1> $H_0 = \{\mu_X = \mu_Y\}\ contre\ H_1 = \{\mu_X \neq \mu_Y\}$
           a) $\mathrm{W} = \left\{|v| > u_{1-\frac{\alpha^*}{2}} \sigma_V\right\}$
     4- Test unilatéral
        1> $H_0 = \{\mu_X = \mu_Y\}\ contre\ H_1 = \{\mu_X > \mu_Y\}$
           a) Test UPP de RC : W = $\{v > u_{1-\alpha^*} \sigma_V\}$
        2> $H_0 = \{\mu_X = \mu_Y\}\ contre\ H_1 = \{\mu_X < \mu_Y\}$
           a) Test UPP de RC : W = $\{v < -u_{1-\alpha^*} \sigma_V\}$
     5- Variances connues

1> $\dfrac{\bar{X}-\bar{Y}}{\sqrt{\dfrac{\sigma_X^2}{n}+\dfrac{\sigma_Y^2}{m}}} \sim^{L} N(0,1)$

6- Variances égales mais inconnues (Test de Student)

1> $S^{*2} = \dfrac{1}{\mathrm{N}-2}\left(\sum (X_i-\bar{X})^2 + \sum (Y_j-\bar{Y})^2\right)$

2> $\Rightarrow \dfrac{(\mathrm{N}-2)S^{*2}}{\sigma^2} \sim \chi^2_{N-2}$

3> $\Rightarrow T = \dfrac{(\bar{X}-\bar{Y})}{S^{*}\sqrt{\dfrac{1}{n}+\dfrac{1}{m}}} \sim T_{N-2}$

a) RC, $\mathrm{W} = \left\{|t| > t_{N-2,1-\frac{\alpha^*}{2}}\right\}, H_1 = \{\mu_X \neq \mu_Y\}$

b) RC, $\mathrm{W} = \{t > t_{N-2,1-\alpha^*}\}, H_1 = \{\mu_X > \mu_Y\}$

c) RC, $\mathrm{W} = \{t < -t_{N-2,1-\alpha^*}\}, H_1 = \{\mu_X > \mu_Y\}$

7- Variances inconnues et différentes

1> $S_V^{*2} = \dfrac{S_X^{*2}}{n} + \dfrac{S_Y^{*2}}{m}\ estimateur\ de\ \sigma_V^2 = \dfrac{\sigma_X^2}{n}+\dfrac{\sigma_Y^2}{m}$

2> $v\dfrac{S_V^{*2}}{\sigma_V^2} = v\dfrac{\dfrac{S_X^{*2}}{n}+\dfrac{S_Y^{*2}}{m}}{\dfrac{\sigma_X^2}{n}+\dfrac{\sigma_Y^2}{m}} \sim_{approx} \chi_V^2$, où v

$$= \frac{\left(\dfrac{S_X^{*2}}{n}+\dfrac{S_Y^{*2}}{m}\right)^2}{\dfrac{1}{n-1}\left(\dfrac{S_X^2}{n}\right)^2+\dfrac{1}{m-1}\left(\dfrac{S_Y^2}{m}\right)^2}, v\ t.q. E\left(v\frac{S_V^{*2}}{\sigma_V^2}\right) = v, Var\left(v\frac{S_V^{*2}}{\sigma_V^2}\right) = 2v$$

3> $\dfrac{V}{\sqrt{S_V^{*2}}} \sim_{approx} T_v$

a) RC, $\mathrm{W} = \left\{\left|\dfrac{V}{\sqrt{S_V^{*2}}}\right| > t_{v,1-\frac{\alpha^*}{2}}\right\}, H_1 = \{\mu_X \neq \mu_Y\}$

b) RC, $\mathrm{W} = \left\{\dfrac{V}{\sqrt{S_V^{*2}}} > t_{v,1-\alpha^*}\right\}, H_1 = \{\mu_X > \mu_Y\}$

c) RC, $\mathrm{W} = \left\{\dfrac{V}{\sqrt{S_V^{*2}}} < -t_{v,1-\alpha^*}\right\}, H_1 = \{\mu_X < \mu_Y\}$

8- Cas de deux échantillons appariés (X_i, Y_i)

1> $D = \bar{X}-\bar{Y} \sim N\left(\mu_D = \mu_X-\mu_Y, \sigma_D^2 = \dfrac{\sigma_X^2}{n}+\dfrac{\sigma_Y^2}{m}\right)$

2> On appliquera donc dans ces cas un test de conformité sur l'échantillon

-2- Homogénéité des variances (Gaussien)

1- $\dfrac{(n-1)S^{*2}_X}{\sigma_X^2} \sim \chi^2_{n-1}\ and\ \dfrac{(m-1)S^{*2}_Y}{\sigma_Y^2} \sim \chi^2_{m-1}$

1> $\dfrac{\dfrac{S_X^{*2}}{\sigma_X^2}}{\dfrac{S_Y^{*2}}{\sigma_Y^2}} \sim F_{n-1,m-1}$

2- Test bilatéral

1> $H_0 = \{\sigma_X = \sigma_Y\}\ contre\ H_1 = \{\sigma_X \neq \sigma_Y\}$

3- Test unilatéral

1> $H_0 = \{\sigma_X = \sigma_Y\}\ contre\ H_1 = \{\sigma_X > \sigma_Y\}$

2> $H_0 = \{\sigma_X = \sigma_Y\}\ contre\ H_1 = \{\sigma_X < \sigma_Y\}$

4- $\dfrac{\dfrac{Z_1}{v_1}}{\dfrac{Z_2}{v_2}} \sim F_{v_1,v_2}, Z_1 \sim \chi^2_{v_1}, Z_2 \sim \chi^2_{v_2}$

5- $H_0, F = \dfrac{S_X^{*2}}{S_Y^{*2}} \sim F_{n-1,m-1}$

1> $\mathrm{RC}, \mathrm{W} = \left\{F < f_{n-1,m-1,\frac{\alpha^*}{2}}\right\} \cup \left\{F > f_{n-1,m-1,1-\frac{\alpha^*}{2}}\right\}, H_1 = \{\sigma_X \neq \sigma_Y\}$

2> $\mathrm{RC}, \mathrm{W} = \left\{\left\{F > f_{n-1,m-1,1-\alpha^*} = \dfrac{1}{f_{n-1,m-1,\alpha^*}}\right\}\right\}, H_1 = \{\sigma_X > \sigma_Y\}$

3> $\mathrm{RC}, \mathrm{W} = \left\{\left\{F < f_{n-1,m-1,\alpha^*}\right\}\right\}, H_1 = \{\sigma_X < \sigma_Y\}$

-3- Homogénéité des proportions

1- $X \sim B(n, p_1), Y \sim B(m, p_2)$

2- Test bilatéral

1> $H_0 = \{p_X = p_Y\}\ contre\ H_1 = \{p_X \neq p_Y\}$

3- Test unilatéral

1> $H_0 = \{p_X = p_Y\}\ contre\ H_1 = \{p_X > p_Y\}$

2> $H_0 = \{p_X = p_Y\}\ contre\ H_1 = \{p_X < p_Y\}$

4- $H_0 : p = p_1 = p_2, p\ inconnue$

1> $\dfrac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n} + \dfrac{1}{m}\right)}} \sim^L N(0,1)$

a) $\mathrm{RC}, \mathrm{W} = \left\{\left|\dfrac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n} + \dfrac{1}{m}\right)}}\right| > u_{1-\frac{\alpha^*}{2}}\right\}, H_1 = \{p_X \neq p_Y\}$

b) $\mathrm{RC}, \mathrm{W} = \left\{\dfrac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n} + \dfrac{1}{m}\right)}} > u_{1-\alpha^*}\right\}, H_1 = \{p_X > p_Y\}$

c) $\mathrm{RC}, \mathrm{W} = \left\{\dfrac{(\hat{p}_1 - \hat{p}_2)}{\sqrt{\hat{p}(1-\hat{p})\left(\dfrac{1}{n} + \dfrac{1}{m}\right)}} < -u_{1-\alpha^*}\right\}, H_1 = \{p_X < p_Y\}$

-4- Homogénéité des lois

1- $X_i \sim F_X$ *loi continue*, $y_j \sim F_Y$ *loi continue*

1> Définition : X est stochastiquement supérieur à Y si

a) $F_X(z) \leq F_Y(z), \forall z \in R$

2- Test bilatéral

1> $H_0 = \{F_X = F_Y\}$ *contre* $H_1 = \{F_X \neq F_Y\}$

3- Test unilatéral

1> $H_0 = \{F_X = F_Y\}$ *contre* $H_1 = \{F_X > F_Y\}$

a) ~~Test UPP de RC : W =~~ $\cancel{\{v > u_{1-\alpha^*}\sigma_V\}}$

2> $H_0 = \{F_X = F_Y\}$ *contre* $H_1 = \{F_X < F_Y\}$

a) ~~Test UPP de RC : W =~~ $\cancel{\{v > u_{1-\alpha^*}\sigma_V\}}$

4- Test de Wilcoxon

1> Somme des rangs des X_i après avoir ordonné la série des $N = m + n$ observations

2> $\{W_X \leq c_1\} \cup \{W_X \geq c_2\}, H_1 = \{F_X \neq F_Y\}$

3> $\{W_X \geq c\}, H_1 = \{F_X > F_Y\}$

4> $\{W_X \leq \tilde{c}\}, H_1 = \{F_X < F_Y\}$

5> $$\begin{cases} E_{H_0}(W_X) = \dfrac{n(n+m+1)}{2} \\ Var_{H_0}(W_X) = \dfrac{nm(n+m+1)}{12} \end{cases}$$

4. Échantillons appariés 配对 (X_i, Y_i)

-1- Modélisation gaussienne de la différence $X - Y$

1- $D_i = X_i - Y_i \sim N(\mu_X - \mu_Y, \sigma_D^2)$

2- $\Rightarrow \sqrt{n}\dfrac{\bar{D}}{\sqrt{{S_D^*}^2}} \sim T_{n-1} - -{S_D^*}^2 = \dfrac{\sum(D_i - \bar{D})^2}{n-1}$

1> $$W = \left\{\left|\sqrt{n}\frac{\bar{d}}{\sqrt{{S_D^*}^2}}\right| > t_{n-1,1-\frac{\alpha^*}{2}}\right\} - -H_1: \mu_D \neq 0$$

2> $$W = \left\{\sqrt{n}\frac{\bar{d}}{\sqrt{{S_D^*}^2}} > t_{n-1,1-\alpha^*}\right\} - -H_1: \mu_D > 0$$

3> $$W = \left\{\sqrt{n}\frac{\bar{d}}{\sqrt{{S_D^*}^2}} < -t_{n-1,1-\alpha^*}\right\} - -H_1: \mu_D < 0$$

-2- Loi quelconque (non gaussienne) mais continue de la différence X − Y

1- $D_i = X_i - Y_i$

2- Test sur la médiane de D, $m_D = D_{\left[\frac{n}{2}\right]}$

1> $\begin{cases} H_0: m_D = 0 \\ H_1: m_D > 0 \end{cases}$

2> Test du signe

1> Test du signe dépend uniquement du signe de chaque différence

2> On pose $Z = \sum 1_{D_i \geq 0} \sim B(n, \theta)$

i) $\Rightarrow \begin{cases} H_0: \theta = \frac{1}{2} \\ H_1: \theta > \frac{1}{2} \end{cases}$

3> Test de Wilcoxon signé

1> Test de Wilcoxon signé tient aussi compte de la valeur de chaque différence $|D_i|$

2> $W_+ = \sum_{i: D_i > 0} R_i Z_i \;--R_i \in \{1, \dots, n\}$

i) $\begin{cases} H_0: \theta = \frac{1}{2} \\ H_1: \theta > \frac{1}{2} \end{cases}$

$x_i - y_i$	3	2	0	2	4	3	-1	1	3	-2
$\lvert x_i - y_i \rvert$ ordonnées	0	$\lvert -1 \rvert$	1	$\lvert -2 \rvert$	2	2	3	3	3	4

ii)

signe de $(x_i - y_i) * \lvert x_i - y_i \rvert$	0	-1	1	-2	$\lvert 2$	2	3	3	3	4
r_i	1	2.5	2.5	5	5	5	8	8	8	10

$W_+ = 46.5$

3> RC

i) $\begin{cases} \left\{W_+ \leq k_{\frac{\alpha^*}{2}}\right\} \cup \left\{W_+ \geq \frac{n(n+1)}{2} - k_{\frac{\alpha^*}{2}}\right\} --H_1: m_D \neq 0 \\ \left\{W_+ \geq \frac{n(n+1)}{2} - k_{\alpha^*}\right\} --H_1: m_D > 0 \\ \{W_+ \leq k_{\alpha^*}\} --H_1: m_D < 0 \end{cases}$

$--k \sim N\left(E(W_+) = \frac{n(n+1)}{4}, Var(W_+) = \frac{n(n+1)(2n+1)}{24}\right)$