

Tests non-paramétriques

Myriam Maumy-Bertrand¹ & Marie Chion¹

¹IRMA, Université de Strasbourg
Strasbourg, France

Master 1^{re} Année
2019-2020

1^{ère} partie

Tests non paramétriques pour un ou deux échantillons

Tests non libres de distribution

Certains tests statistiques ne sont valables que sous certaines conditions concernant la distribution de la ou les variable(s).

Exemples

Le test de Student impose que les deux variables sont issues d'une distribution normale, l'analyse de la variance également.

Tests libres de distribution

D'autres tests au contraire sont valables indépendamment de toute distribution. Nous les appelons les tests « libres de distribution » (distribution-free tests).

Exemples

C'est le cas du test du Khi-deux, du test des signes, ou du test du coefficient de Spearman.

Tests paramétriques

Certains tests ont pour but de montrer une égalité sur certains paramètres : ce sont les tests paramétriques.

Exemples de paramètres

- 1 La moyenne (test de comparaison de deux moyennes ou plus),
- 2 la variance (test de comparaison de deux variances ou plus),
- 3 etc.

Tests non paramétriques

D'autres tests testent des hypothèses plus générales : ce sont les tests non paramétriques.

Exemples

- 1 Une égalité de lois,
- 2 l'indépendance entre deux variables qualitatives,
- 3 etc.

Une question naturelle : quel test choisir ?

Habituellement, les tests paramétriques sont plus puissants. Par conséquent, ils seront choisis plutôt que les tests non paramétriques.

De même les tests non libres sont généralement plus efficaces que les tests libres. Cependant, ils sont aussi plus contraignants, car il faut vérifier les conditions d'application qui sont plus nombreuses dans ce cas.

On choisira généralement un test libre ou non paramétrique lorsque

- 1 les conditions d'application du test ne sont pas vérifiées
- 2 ou il est impossible de vérifier ces conditions.

Soit un échantillon aléatoire (X_1, X_2, \dots, X_n) de loi parente une loi continue de fonction de répartition F_X dont la médiane est notée m_e et la moyenne μ .

Hypothèses testées

Le **test des signes** permet de tester l'hypothèse suivante :

$$\mathcal{H}_0 : m_e = 0 \text{ ou de façon équivalente } \mathbb{P}(X_i > 0) = 1/2$$

contre

$$\mathcal{H}_1 : m_e \neq 0 \text{ ou de façon équivalente } \mathbb{P}(X_i > 0) \neq 1/2.$$

Remarques

- 1 La formulation de ce test est bien sûr la formulation d'un test bilatéral. Nous pouvons envisager les deux tests unilatéraux correspondants qui s'écrivent :

$$\mathcal{H}_0 : \mathbb{P}(X_i > 0) = 1/2$$

contre

$$\mathcal{H}_1 : \mathbb{P}(X_i > 0) < 1/2.$$

Ou

$$\mathcal{H}_0 : \mathbb{P}(X_i > 0) = 1/2$$

contre

$$\mathcal{H}_1 : \mathbb{P}(X_i > 0) > 1/2.$$

Remarques

- ② Lorsque m_0 est un nombre réel, ce test permet de tester plus généralement l'hypothèse nulle

$$\mathcal{H}_0 : m_e = m_0$$

contre

$$\mathcal{H}_1 : m_e \neq m_0.$$

La formulation de ce test est bien sûr la formulation d'un test bilatéral.

Nous pouvons envisager les deux tests unilatéraux correspondants qui s'écrivent :

$$\mathcal{H}_0 : m_e = m_0$$

contre

$$\mathcal{H}_1 : m_e < m_0.$$

ou

$$\mathcal{H}_0 : m_e = m_0$$

contre

$$\mathcal{H}_1 : m_e > m_0.$$

Pour cela il suffit de considérer l'échantillon Y_1, \dots, Y_n avec $Y_i = X_i - m_0$ et nous sommes ramenés au test précédent.

Statistique du test

La statistique S_n^+ du test des signes de l'échantillon se définit par le nombre de variables aléatoires X_i , $1 \leq i \leq n$, qui prennent une valeur positive ou encore

$$S_n^+ = \sum_{i=1}^n \mathbf{1}_{\{X_i > 0\}}.$$

Remarque

La loi de la statistique S_n^+ ne dépend pas de la loi continue F_X .

Règle de décision et conclusion du test

- Premier cas** : La taille n est inférieure à 40. Pour un seuil donné α , nous cherchons, dans les tables de la loi binomiale, le plus grand nombre entier k_α tel que $\mathbb{P}_{(\mathcal{H}_0)} (S_n^+ \leq k_\alpha) \leq \alpha/2$. Alors nous décidons :

$$\left\{ \begin{array}{ll} \text{si } S_{n,\text{obs}}^+ \notin]k_\alpha; n - k_\alpha[& \mathcal{H}_1 \text{ est vraie,} \\ \text{si } S_{n,\text{obs}}^+ \in]k_\alpha; n - k_\alpha[& \mathcal{H}_0 \text{ est vraie.} \end{array} \right.$$

Règle de décision et conclusion du test (suite)

- **Second cas** : La taille n est supérieure ou égale à 40. La statistique S_n^+ suit approximativement une loi normale et nous utilisons alors la statistique suivante, en tenant compte de la correction de continuité :

$$Z_n = \frac{2S_n^+ + 1 - n}{\sqrt{n}}.$$

Règle de décision et conclusion du test (suite)

Pour un seuil donné α , la table de la loi d'une variable aléatoire Z normale centrée réduite nous fournit une valeur critique c telle que $\mathbb{P}_{(\mathcal{H}_0)}(-c < Z_n < c) = 1 - \alpha$. Alors nous décidons :

$$\begin{cases} \text{si } \mathbf{Z}_{n,\text{obs}} \notin]-c; +c[& (\mathcal{H}_1) \text{ est vraie,} \\ \text{si } \mathbf{Z}_{n,\text{obs}} \in]-c; +c[& (\mathcal{H}_0) \text{ est vraie.} \end{cases}$$

Remarques

- 1 Le niveau de signification réel du test est alors égal à $2\mathbb{P}(S_n^+ \leq k_\alpha)$ qui est généralement différent de α .
- 2 Dans le cas où il y a des ex æquo dans les données, le protocole ne change pas.

Méthode

Pour traiter ce problème la méthode recommandée est la suivante : les éliminer et se ramener à un jeu de données de taille n' , où n' est le nombre d'observations non nulles, puis le traiter comme ci-dessus.

Les prématurés

Il est admis que des prématurés nés avec une masse de 2,15 kg arrivent à une masse de 2,80 kg en un mois s'ils sont nourris avec du lait maternel. 12 nourrissons pesant approximativement 2,15 kg à la naissance ont été nourris avec un lait sensé remplacer le lait maternel. Les gains de masse en kilogrammes, notés x_i , ont été de

0,55; 0,62; 0,54; 0,58; 0,63; 0,64;
0,60; 0,62; 0,59; 0,67; 0,62; 0,61.

Est-il possible de conclure, au seuil de $\alpha = 5\%$, à une différence significative entre les effets du lait maternel et ceux du lait de remplacement relativement à la prise de masse des nourrissons ?

Les prématurés

Compte tenu du faible effectif de l'échantillon, nous décidons de tester l'hypothèse nulle :

$$\mathcal{H}_0 : m_e = m_0 = 2,80 \text{ kg}$$

contre

$$\mathcal{H}_1 : m_e \neq m_0 = 2,80 \text{ kg.}$$

Les prématurés : Première méthode : « à la main »

Nous allons d'abord transformer les données observées x_1, \dots, x_{12} en un échantillon y_1, \dots, y_{12} . Pour cela, nous calculons les y_i avec la formule suivante : $y_i = 2,15 + x_i - 2,80$. Nous avons donc la suite de données suivantes :

$-0,10; -0,03; -0,11; -0,07; -0,02; -0,01$
 $-0,05; -0,03; -0,06; +0,02; -0,03; -0,04.$

Nous en déduisons que $S_{12,obs}^+ = 1$. Il ne reste plus qu'à déterminer le plus grand nombre entier k_α tel que

$$\mathbb{P}_{(\mathcal{H}_0)} (S_{12}^+ \leq k_\alpha) \leq 0,025.$$

Les prématurés : Première méthode : « à la main »

Pour trouver ce nombre entier, nous allons utiliser le logiciel R.

Nous tapons donc la ligne de commande suivante :

```
pbinom(0:11, 11, 0.5)
```

```
0.0004882812 0.0058593750 0.0327148438
```

```
0.1132812500 0.2744140625 0.5000000000
```

```
0.7255859375 0.8867187500 0.9672851562
```

```
0.9941406250 0.9995117188 1.0000000000
```

Nous trouvons que ce nombre entier k_α est égal à 1.

Les prématurés : Première méthode : « à la main »

Par conséquent, comme $S_{12,obs}^+ \notin]1; 11[$, le test est significatif au seuil de $\alpha = 5\%$. Nous rejetons donc l'hypothèse nulle (\mathcal{H}_0) et nous décidons que l'hypothèse alternative (\mathcal{H}_1) est vraie. Il est donc possible de conclure, avec un risque d'erreur de première espèce de $\alpha = 5\%$, à une différence significative entre les effets du lait maternel et ceux du lait de remplacement relativement à la prise de poids des nourrissons.

Les prématurés : **Seconde méthode** : « avec R »

Nous tapons les deux lignes de commande suivante :

```
>poids<-c(0.55, 0.62, 0.54, 0.58, 0.63, 0.64,
0.60, 0.62, 0.59, 0.67, 0.62, 0.61)
>SIGN.test(poids,md=0.65)
One-sample Sign-Test
data: poids
s = 1, p-value = 0.006348
alternative hypothesis: true median is not
equal to 0.65
95 percent confidence interval:
0.5810636 0.6289364
sample estimates:
median of x
0.615
```


Soit un échantillon aléatoire (X_1, X_2, \dots, X_n) de loi parente une loi continue de fonction de répartition F_X dont la médiane est notée m_e et la moyenne μ .

Hypothèses testées

Le **test des rangs signés de Wilcoxon** permet de tester l'hypothèse suivante :

$\mathcal{H}_0 : F_X$ est symétrique par rapport à l'origine

contre

$\mathcal{H}_1 : F_X$ n'est pas symétrique par rapport à l'origine.

Ici l'origine c'est 0.

Remarques

- 1 Nous pouvons remplacer la valeur 0 dans les hypothèses ci-dessus par une valeur fixée à l'avance, comme 1, 2 ou π .
- 2 Si nous savons que F_X est symétrique, (pour le savoir, par exemple, tracer un histogramme) alors le test des rangs signés de Wilcoxon permet de tester :

$$\mathcal{H}_0 : \mu = 0$$

contre

$$\mathcal{H}_1 : \mu \neq 0.$$

Ce qui permet de s'intéresser à la moyenne μ de la loi. Nous rappelons, que dans le cas d'une loi symétrique, la moyenne et la médiane sont confondues.

Remarques

- ③ Nous souhaitons tester l'hypothèse nulle

$$\mathcal{H}_0 : \mu = \mu_0$$

contre

$$\mathcal{H}_1 : \mu \neq \mu_0.$$

Alors nous introduisons l'échantillon Y_1, \dots, Y_n avec $Y_i = X_i - \mu_0$.

- ④ La formulation de ce test est la formulation d'un test bilatéral. Nous pourrions envisager d'étudier les deux tests unilatéraux correspondants.

Soit (x_1, \dots, x_n) une réalisation de l'échantillon précédent. À chaque x_i nous attribuons le **rang** r_i^a qui correspond au rang de $|x_i|$ lorsque que les n réalisations sont classées par ordre croissant de leurs valeurs absolues. Le rang r_i^a est la réalisation d'une variable aléatoire R_i^a .

Remarque

La lettre a est là pour rappeler que nous travaillons sur les valeurs absolues des x_i .

Statistique du test

Nous déterminons alors la somme $W_{n,obs}^+$ des rangs r_i^a des seules observations strictement positives. La statistique W_n^+ des rangs signés de Wilcoxon est la variable aléatoire qui prend pour valeur la somme $W_{n,obs}^+$. Par conséquent, la statistique W_n^+ des rangs signés de Wilcoxon de l'échantillon se définit par :

$$W_n^+ = \sum_{\substack{1 \leq i \leq n \\ X_i > 0}} R_i^a.$$

Remarque

La loi de la statistique W_n^+ ne dépend pas de la loi continue F_X des variables aléatoires X_i .

Règle de décision et conclusion du test

- **Premier cas** : La taille n est inférieure à 15. Pour un seuil donné α , nous cherchons, dans les tables de la loi de Wilcoxon, le plus grand nombre entier w_α tel que $\mathbb{P}_{(\mathcal{H}_0)} (W_n^+ \leq w_\alpha) \leq \alpha/2$. Alors nous décidons :

$$\left\{ \begin{array}{l} \text{si } W_{n,\text{obs}}^+ \notin]w_\alpha; n(n+1)/2 - w_\alpha[\quad (\mathcal{H}_1) \text{ est vraie,} \\ \text{si } W_{n,\text{obs}}^+ \in]w_\alpha; n(n+1)/2 - w_\alpha[\quad (\mathcal{H}_0) \text{ est vraie.} \end{array} \right.$$

Règle de décision et conclusion du test (suite)

- **Deuxième cas** : La taille n est supérieure ou égale à 15. La statistique W_n^+ suit approximativement une loi normale et nous utilisons alors la statistique suivante, en tenant compte de la correction de continuité :

$$Z_n = \frac{2W_n^+ + 1 - \frac{n(n+1)}{2}}{\sqrt{\frac{n(n+1)(2n+1)}{6}}}.$$

Règle de décision et conclusion du test (fin)

Pour un seuil donné α , la table de la loi d'une variable aléatoire Z normale centrée réduite nous fournit une valeur critique c telle que $\mathbb{P}_{(\mathcal{H}_0)} [-c < Z_n < c] \leq 1 - \alpha$. Alors nous décidons :

$$\begin{cases} \text{si } \mathbf{Z}_{n,\text{obs}} \notin]-c; +c[& (\mathcal{H}_1) \text{ est vraie,} \\ \text{si } \mathbf{Z}_{n,\text{obs}} \in]-c; +c[& (\mathcal{H}_0) \text{ est vraie.} \end{cases}$$

Cette méthode est la plus utilisée, en particulier dans la plupart des logiciels statistiques.

Les observations x_1, \dots, x_n peuvent présenter des ex æquo et *a fortiori* leurs valeurs absolues.

Statistique du test

En associant à la variable X_i son rang moyen R_i^{a*} dans le classement des valeurs absolues et en sommant tous les rangs pour lesquels $X_i > 0$ nous obtenons la statistique :

$$W_n^{+*} = \sum_{\substack{1 \leq i \leq n \\ X_i > 0}} R_i^{a*}.$$

Remarque

Le symbole \star est là pour rappeler que nous sommes dans le cas où il y a des ex æquo. Les valeurs absolues observées $|x_1|, \dots, |x_n|$ sont ordonnées puis regroupées en classes d'ex æquo, C_0 pour la première classe qui est constituée des nombres $|x_i|$ nuls, s'il en existe, et C_j , $1 \leq j \leq h$ pour les autres nombres. Certaines classes C_j peuvent comporter un seul élément, si cet élément n'a pas d'ex æquo. Notons d_j le nombre d'ex æquo de la classe C_j . Nous avons

$$d_0 + \sum_{j=1}^h d_j = n.$$

Règle de décision et conclusion du test

- **Premier cas** : La taille n est inférieure à 15. Pour ces valeurs, les calculs « à la main » sont fastidieux. Mais il est à noter qu'il existe des logiciels qui traitent parfaitement ce cas.
- **Deuxième cas** : Même règle et même conclusion que dans le cas où il n'y a pas d'ex æquo en remplaçant W_n^+ par W_n^{+*} .

Remarques

- 1 Lorsque nous utilisons cette méthode des rangs moyens, nous ne pouvons pas utiliser les tables statistiques usuelles qui concernent la distribution de la statistique W_n^+ .
- 2 Par extension nous pourrions utiliser la procédure ci-dessus lorsque la loi F des variables aléatoires X_i est discrète.

Nous considérons deux variables aléatoires X et Y de lois continues, observées toutes les deux sur les mêmes unités d'un n -échantillon. Les observations se présentent alors sous la forme d'une suite de couples $(x_1, y_1), \dots, (x_n, y_n)$.

Hypothèses testées

Le **test de Wilcoxon** permet de tester l'hypothèse suivante :

$$\mathcal{H}_0 : \mathcal{L}(X) = \mathcal{L}(Y)$$

contre

$$\mathcal{H}_1 : \mathcal{L}(X) \neq \mathcal{L}(Y).$$

Remarque

Ce test suppose que la loi de la différence entre les deux variables étudiées X et Y est symétrique par rapport à 0.

Statistique du test

Pour obtenir la statistique du test notée W_n^+ en général, nous devons procéder à des calculs successifs :

- 1 Après avoir calculé les différences d_i , nous classons par ordre croissant les $|d_i|$ non nulles, c'est-à-dire les d_i sans tenir compte des signes.
- 2 Nous attribuons à chaque $|d_i|$ le rang correspondant.
- 3 Nous restituons ensuite à chaque rang le signe de la différence correspondante.
- 4 Enfin, nous calculons la somme W_n^+ des rangs positifs (P) et la somme W_n^- des rangs négatifs (M).

La somme W_n^+ des rangs positifs (P) permet de tester l'hypothèse nulle (\mathcal{H}_0).

Règle de décision et conclusion du test

- Premier cas** : La taille n est inférieure strictement à 15. Pour un seuil donné α ($= 5\% = 0,05$ en général), nous cherchons le plus grand nombre entier k_α tel que $\mathbb{P}_{(\mathcal{H}_0)} [W_n^+ \leq k_\alpha] \leq \alpha/2$. Alors nous décidons :

$$\left\{ \begin{array}{l} \text{si } W_{n,\text{obs}}^+ \notin]k_\alpha; n(n+1)/2 - k_\alpha[\quad (\mathcal{H}_1) \text{ est vraie,} \\ \text{si } W_{n,\text{obs}}^+ \in]k_\alpha; n(n+1)/2 - k_\alpha[\quad (\mathcal{H}_0) \text{ est vraie.} \end{array} \right.$$

Règle de décision et conclusion du test (suite et fin)

- **Second cas** : La taille n est supérieure ou égale à 15. Nous utilisons l'approximation normale avec correction de continuité :

$$\mathbb{P}_{(\mathcal{H}_0)} [W_n^+ \leq k] = \Phi \left(\frac{2k + 1 - \frac{n(n+1)}{2}}{\sqrt{\frac{n(n+1)(2n+1)}{6}}} \right)$$

où Φ est la fonction de répartition de la loi normale centrée réduite et k un nombre entier compris entre 0 et n .

Méthode

Il se traite de la même manière que pour la statistique du test des rangs signés de Wilcoxon.

Méthode

Il se traite de la même manière que pour la statistique du test des rangs signés de Wilcoxon.

Cas d'école

Un psychologue de l'enfance veut tester l'effet de l'assistance à l'école maternelle sur la compréhension sociale des enfants. Il estime cette compréhension à partir des réponses que les enfants donnent à une série de questions portant sur des images représentant diverses situations sociales. Chaque enfant obtient ainsi un score compris entre 0 et 100.

Le psychologue ne peut pas affirmer que les différences observées entre scores sont numériquement exactes (il ne peut pas dire qu'un score de 60 est le double d'un score de 30, ni que la différence entre 60 et 40 est exactement le double de la différence entre 40 et 30). Cependant, il pense que les scores sont suffisamment précis pour qu'il puisse les ranger selon les valeurs absolues de leurs différences deux à deux.

Cas d'école : suite

Pour tester l'effet de l'assistance à l'école maternelle sur la compréhension sociale des enfants, il utilise 8 paires de jumeaux. L'un des jumeaux est envoyé à l'école, alors que l'autre reste à la maison pendant un trimestre. L'affectation se fait au hasard. À la fin du trimestre, il estime la compréhension sociale de chacun des enfants.

Cas d'école : suite

Paires	Scores enfants scolarisés	Scores enfants non-scolarisés
<i>a</i>	82	63
<i>b</i>	69	42
<i>c</i>	73	74
<i>d</i>	43	37
<i>e</i>	58	51
<i>f</i>	56	43
<i>g</i>	76	80
<i>h</i>	65	62

Cas d'école : suite

Est-il possible de conclure, au seuil de $\alpha = 5\%$, à une différence significative entre la compréhension sociale des enfants restés à la maison et celle des enfants scolarisés ?

Compte tenu du faible effectif de l'échantillon apparié, nous décidons de tester l'hypothèse nulle

$$\mathcal{H}_0 : \mathcal{L}(X) = \mathcal{L}(Y)$$

contre

$$\mathcal{H}_1 : \mathcal{L}(X) \neq \mathcal{L}(Y).$$

Calcul de la statistique du test

- 1 Calculons les différences $|d_i|$:

19; 27; -1; 6; 7; 13; -4; 3.

- 2 Attribuons le rang correspondant à chaque $|d_i|$ et rendons le signe :

7; 8; -1; 4; 5; 6; -3; 2.

- 3 Calculons la statistique du test W_n^+ :

$$W_n^+ = 7 + 8 + 4 + 5 + 6 + 2 = 32.$$

Conclusion du test

- 4 Dans la table associée au test de Wilcoxon, nous lisons $k_\alpha = 4$.
- 5 Nous calculons l'intervalle correspondant : $]4; 32[$.
- 6 $W_{n,obs}^+ = 32 \notin]4; 32[$.
- 7 Comme la valeur de la statistique du test calculée sur l'échantillon n'appartient pas à l'intervalle, le test est significatif au seuil $\alpha = 5\%$. Nous décidons de rejeter \mathcal{H}_0 et nous décidons que \mathcal{H}_1 est vraie avec un risque de première espèce $\alpha = 5\%$.

Le **test de Mann-Whitney** a été introduit en 1947 indépendamment du **test de Wilcoxon de la somme des rangs** qui a été élaboré en 1945. Ces deux tests, d'une formulation différente, sont en fait équivalents. En fonction de l'outil informatique que vous utiliserez, la dénomination du test pourra être l'une des suivantes : Test de Mann-Whitney, Test de Wilcoxon de la somme des rangs ou encore Test de Mann-Whitney-Wilcoxon.

L'approche de Mann et Whitney paraît souvent plus facile à mettre en pratique. Si nous devons utiliser une table, il nous faudra déterminer quelle a été l'approche utilisée par le logiciel statistique et nous servir de l'une des tables appropriées.

Nous observons, de manière indépendante, une variable aléatoire X de loi continue, sur deux populations, ou sur une population divisée en deux sous-populations. Nous obtenons ainsi deux séries d'observations notées (x_1, \dots, x_{n_1}) pour la première et (y_1, \dots, y_{n_2}) pour la seconde. Nous notons \mathcal{L}_i la loi de la variable aléatoire X sur la (sous-)population d'ordre i .

Hypothèses testées

Le **test de Mann-Whitney** permet de tester l'hypothèse suivante :

$$\mathcal{H}_0 : \mathcal{L}_1(X) = \mathcal{L}_2(X)$$

contre

$$\mathcal{H}_1 : \mathcal{L}_1(X) \neq \mathcal{L}_2(X).$$

Statistique du test

Pour obtenir la statistique U_{n_1, n_2} du test de Mann-Whitney, en général, nous devons procéder à des calculs successifs :

- Nous classons par ordre croissant l'ensemble des observations des deux échantillons (x_1, \dots, x_{n_1}) et (y_1, \dots, y_{n_2}) de taille respective n_1 et n_2 .
- Nous affectons le rang correspondant.
- Nous effectuons les sommes des rangs (**rank sums**) pour chacun des deux échantillons, notées R_{n_1} et R_{n_2} .

Statistique du test (suite et fin)

- Nous en déduisons les quantités U_{n_1} et U_{n_2} qui se calculent ainsi :

$$U_{n_1} = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_{n_1}$$

et

$$U_{n_2} = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_{n_2} = n_1 n_2 - U_{n_1}.$$

Par conséquent, la statistique U_{n_1, n_2} du test de Mann-Whitney de l'échantillon se définit comme étant la plus petite des deux valeurs U_{n_1} et U_{n_2} . C'est cette statistique qu'il faut considérer car les tables sont construites autour de cette statistique.

Règle de décision et conclusion du test

- **Premier cas** : Les tailles n_1 ou n_2 sont inférieures ou égales à 20. Pour un seuil donné α ($= 5\% = 0,05$ en général), les tables de Mann-Whitney nous fournissent une valeur critique c . Alors nous décidons :

$$\left\{ \begin{array}{l} \text{si } U_{n_1, n_2, \text{obs}} \leq c \quad (\mathcal{H}_1) \text{ est vraie,} \\ \text{si } U_{n_1, n_2, \text{obs}} > c \quad (\mathcal{H}_0) \text{ est vraie.} \end{array} \right.$$

Règle de décision et conclusion du test (suite)

- **Second cas** : Les tailles n_1 et n_2 sont supérieures strictement à 20. La statistique U_{n_1, n_2} suit approximativement une loi normale et nous utilisons alors la statistique suivante, en tenant compte de la correction de continuité :

$$Z_{n_1, n_2} = \frac{2U_{n_1, n_2} + 1 - n_1 n_2}{\sqrt{\frac{(n_1 n_2)(n_1 + n_2 + 1)}{3}}}$$

Règle de décision et conclusion du test (fin)

Pour un seuil donné α , la table de la loi d'une variable aléatoire Z normale centrée réduite nous fournit une valeur critique c telle que $\mathbb{P}_{(\mathcal{H}_0)}(-c < Z_{n_1, n_2} < c) = 1 - \alpha$. Alors nous décidons :

$$\begin{cases} \text{si } Z_{n_1, n_2, \text{obs}} \notin]-c; +c[& (\mathcal{H}_1) \text{ est vraie,} \\ \text{si } Z_{n_1, n_2, \text{obs}} \in]-c; +c[& (\mathcal{H}_0) \text{ est vraie.} \end{cases}$$

Méthode

Cette méthode est la plus utilisée, en particulier dans la plupart des logiciels statistiques.

Les observations $x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}$ peuvent présenter des ex aequo. Les valeurs absolues observées

$x_1, \dots, x_{n_1}, y_1, \dots, y_{n_2}$ sont ordonnées puis regroupées en h classes d'ex aequo C_j , $1 \leq j \leq h$. Certaines classes C_j peuvent comporter un seul élément, si cet élément n'a pas d'ex aequo.

Méthode (suite et fin)

Notons d_j le nombre d'ex æquo de la classe C_j . Nous avons

$$\sum_{j=1}^h d_j = n_1 + n_2.$$

En associant à la variable X_i son rang moyen R_i^* dans ce classement et en sommant les rangs de tous les X_i , nous obtenons la statistique :

$$U_{n_1, n_2}^* = \sum_{i=1}^{n_1} R_i^*.$$

Remarques

- 1 Pour $n_1 \leq 15$ et $n_2 \leq 15$, les calculs « à la main » sont fastidieux. Mais vous devez savoir qu'il existe des logiciels qui traitent parfaitement ce cas.
- 2 Lorsque nous utilisons cette méthode des rangs moyens nous ne pouvons pas utiliser les tables statistiques usuelles qui concernent la distribution de la variable aléatoire U_{n_1, n_2} .
- 3 Par extension nous pourrions utiliser cette procédure lorsque les lois \mathcal{L}_1 et \mathcal{L}_2 des variables aléatoires X_i sont discrètes.

Les arbres

Dans deux types de forêts distincts, nous avons mesuré les hauteurs respectivement de 13 et 14 arbres choisis au hasard et indépendamment, dans le but de vérifier si les deux distributions des hauteurs des deux types de forêts sont ou ne sont pas égales.

Forêt 1	Forêt 1	Forêt 2	Forêt 2
23,4	24,4	22,5	22,9
24,6	24,9	23,7	24,0
25,0	26,2	24,3	24,5
26,3	26,5	25,3	26,0
26,6	26,8	26,1	26,4
27,0	27,6	26,7	26,9
27,7		27,4	28,5

Statistique du test

1

Forêt 1	Forêt 1	Forêt 2	Forêt 2
3	7	1	2
9	10	4	5
11	15	6	8
16	18	12	13
19	21	14	17
23	25	20	22
26		24	27

- 2 Calculons R_{n_1} et R_{n_2} : $R_{n_1} = 203$ et $R_{n_2} = 175$.
 Nous vérifions que $R_{n_1} + R_{n_2} = (27 * 28)/2 = 378$.

Statistique du test

3 Calculons

$$U_{n_1} = 13 * 14 + (13 * 14)/2 - 203 = 70$$

et

$$U_{n_2} = 13 * 14 - 70 = 112.$$

4 Calculons la statistique du test :

$$U_{n_1, n_2} = \min(U_{n_1}, U_{n_2}) = 70.$$

Conclusion du test

- 5 Dans la table associée au test de Mann-Whitney, nous lisons $c = 50$.
- 6 $U_{n_1, n_2, obs} = 70 > 50 = c$.
- 7 Comme la valeur de la statistique du test calculée sur l'échantillon est strictement supérieure à la valeur critique, le test n'est pas significatif au seuil $\alpha = 5\%$. Nous décidons de ne pas rejeter \mathcal{H}_0 . Le risque d'erreur associé à cette décision est un risque d'erreur de seconde espèce β . Nous ne pouvons pas l'évaluer dans ce cas.