Quelques tests...

Ségolen Geffray

Ecole doctorale Chimie/Biologie

Année 2008-2009

Des bonnes questions à se poser

- Qu'est-ce je veux (peux) faire ?
  - comparer une moyenne, une proportion, une variance à une valeur de référence ou une distribution à une loi de référence...
  - comparer des moyennes, des proportions, des variances, des distributions entre elles...
  - mettre en évidence des liaisons, des associations
- un, deux ou plusieurs échantillons ?
- dans le cas de plusieurs échantillons, sont-ils indépendants ou appariés ?
- petit(s) ou grand(s) échantillon(s) ?
- test paramétrique ou non-paramétrique ?

Tests de comparaison de moyennes

Comparer une moyenne à une valeur de référence :

- one sample z-test pour échantillon gaussien avec variance connue
- one-sample t-test pour échantillon gaussien avec variance inconnue (robuste à l'hypothèse de normalité)
- one-sample z-test pour grand échantillon avec variance inconnue

Tests de comparaison de moyennes (suite)

Comparer deux moyennes entre elles :

- **échantillons indépendants**
  - two-sample z-test pour échantillons gaussiens avec variances connues
  - two-sample t-test pour échantillons gaussiens homoscédastiques (robuste à l'hyp de normalité)
  - two-sample z-test pour grands échantillons avec variances inconnues
  - test d'Aspin-Welch $\simeq$ two-sample t-test pour échantillons gaussiens avec variances inconnues (robuste à l'hyp de normalité)
  - sinon test de comparaison de 2 distributions
- **échantillons appariés**
  - one-sample z-test pour échantillon gaussien avec variance connue sur la différence $D = X - Y$
  - one-sample t-test pour échantillon gaussien avec variance inconnue sur la différence $D = X - Y$ (robuste à l'hyp de normalité)
  - one-sample z-test pour grand échantillon avec variance inconnue sur la différence $D = X - Y$
  - sinon test d'ajustement de la loi de $D = X - Y$ à une loi de moyenne nulle

Tests de comparaison de moyennes (fin)

Comparer k moyennes ($k > 2$) entre elles :

- **échantillons indépendants**
  - échantillons gaussiens homoscédastiques : anova à un facteur
  - sinon test de Kruskal-Wallis
  - méthodes de comparaisons multiples pour identifier les échantillons qui diffèrent les uns des autres
- **échantillons appariés**
  - échantillons gaussiens homoscédastiques : anova sur mesures répétées

Tests de comparaison de variances

- **Comparer une variance à une valeur de référence** :
  - test avec une statistique du χ^2 pour échantillon gaussien
  - z-test pour grand échantillon
- **Comparer deux variances** :
  - **échantillons indépendants**
    - f-test pour échantillons gaussiens
    - z-test pour grands échantillons
    - sinon test de Levene, test log-anova...
  - **échantillons appariés**
    - test avec une statistique du χ^2 sur la différence pour échantillons gaussiens
    - one-sample z-test sur la différence pour grands échantillons
- **Comparer k variances ($k > 2$) entre elles** :
  - **échantillons indépendants**
    - test de Bartlett pour échantillons gaussiens
    - test L_1 de Neyman-Pearson
    - test C de Cochran
    - test log-anova
    - ...

Tests de comparaison de proportions

- **Comparer une proportion à une valeur de référence** :
  - z-test pour échantillon satisfaisant l'approximation normale
  - test basé sur l'approximation de Poisson pour échantillon satisfaisant l'approximation de Poisson
  - sinon test exact basé sur la loi binômiale
- **Comparer deux proportions** :
  - **échantillons indépendants**
    - z-test pour échantillons satisfaisant l'approximation normale
    - test exact de Fisher
    - test G du rapport de vraisemblance
  - **échantillons appariés**
    - test de Mac Neymar
- **Comparer k proportions** :
  - **échantillons indépendants**
    - comparer k proportions à k valeurs de référence : test du χ^2 à un facteur de classification
    - test G du rapport de vraisemblance

Tests de comparaison de distributions (forcément non-paramétriques !)

- **Comparer une distribution à une loi de référence** :
  - test d'ajustement de Kolmogorov-Smirnov
  - test du χ^2 d'ajustement
  - test G du rapport de vraisemblance
  - test de normalité de Shapiro-Wilk
- **Comparer deux distributions entre elles** :
  - **échantillons indépendants**
    - two-sample test de Kolmogorov-Smirnov
    - test U de Mann-Whitney $\simeq$ test de la somme des rangs de Wilcoxon
  - **échantillons appariés**
    - test de Wilcoxon des rangs signés de la différence
- **Comparer k distributions ($k > 2$)** :
  - **échantillons indépendants**
    - test de Kruskal-Wallis
  - **échantillons appariés**
    - test Q de Cochran
    - test de Friedman

Tests d'indépendance et de corrélation

- **Tester l'indépendance entre 2 variables** :
  - Test d'indépendance entre 2 variables qualitatives ou réparties en classes
    - test du χ^2 d'indépendance
    - test exact de Fisher
  - Test d'indépendance entre une variable qualitative et une variable quantitative : cf comparaison de 2 ou plusieurs échantillons
  - Test d'indépendance entre 2 variables quantitatives
- **Tester la corrélation de 2 variables quantitatives** :
  - données gaussiennes : coefficient R de corrélation de Bravais-Pearson
  - données non-gaussiennes : coefficient de corrélation de Spearman

Test paramétrique ou non-paramétrique ?

- un test est dit paramétrique s'il nécessite des hypothèses sur la distribution des variables de l' (des) échantillon(s). Un test est dit non-paramétrique s'il ne nécessite aucune hypothèse de distribution.
- Une méthode statistique est dite robuste si elle peut être encore employée lorsque les hypothèses relatives aux distributions des variables de l' (des) échantillon(s) peuvent ne pas être exactement remplies.
- Il est classique de considérer que les tests paramétriques précédents sont robustes(!)....
- 2 petits avantages de tests paramétriques : simplicité et petit gain de puissance par rapport aux tests non-paramétriques.
- le gros avantage des tests non-paramétriques : absence de conditions de validité donc pas d'angoisse quant à l'observance (parfois inconnue) des conditions de validité et à la robustesse (supposée) des tests paramétriques.

Echantillons indépendants ou appariés ?

- Pour améliorer la puissance d'un test, une méthode simple consiste à essayer de diminuer la variance de l'estimateur du paramètre à mesurer car moins il y a de "bruit" (i.e. moins de pertubations), plus on a de chances de détecter un phénomène particulier. C'est ce que l'on cherche à faire lorsqu'une mesure est effectuée à deux reprises sur le même sujet (avant et après traitement, par ex). On obtient alors deux échantillons : celui avec les premières mesures réalisées et celui avec les secondes mesures réalisées. Bien entendu, ces deux échantillons ne sont pas indépendants, ils sont appariés.
- Réaliser un appariement consiste à faire la différence de ces deux mesures afin d'éliminer une grande partie de la variabilité intra-individuelle.
- Cette stratégie peut se généraliser à un grand nombre de situations expérimentales, par ex, 15 patients diagnostiqués par 2 médecins.

Rappel : principe des tests statistiques

- On compare une moyenne (ou une variance ou un paramètre de distribution) à une valeur de référence sur la base des valeurs d'un échantillon.
- A l'issue de la comparaison, on décide si l'on accepte l'égalité proposée ou non.
- Pour déterminer si une telle assertion sur une caractéristique de la population doit être acceptée ou rejetée, on utilise un test statistique.
- Un test statistique spécifie les hypothèses en compétition et le risque associé à la décision prise.

Rappel : démarche pour tester l'adéquation à une valeur de référence

- Etape 1 : on formule l'**hypothèse nulle** H_0 sous la forme $H_0 : \theta = \theta_0$ où θ désigne une **caractéristique de la population** et où θ_0 désigne une **valeur de référence**.
- Etape 2 : on formule l'**hypothèse alternative** H_1 en contradiction avec H_0 sous la forme $H_1 : \theta \neq \theta_0$ ou $H_1 : \theta < \theta_0$ ou $H_1 : \theta > \theta_0$.
- Etape 3 : on fixe α le risque de $1^{\text{ère}}$ espèce accepté.
- Etape 4 : on utilise les données issues de l'échantillon pour tester les deux hypothèses en compétition H_0 contre H_1.

Prise de décision au moyen de la région de rejet

- On veut tester l'hypothèse nulle $H_0 : \theta = \theta_0$ contre l'hypothèse alternative $H_1 : \theta \neq \theta_0$ ou $H_1 : \theta < \theta_0$ ou $H_1 : \theta > \theta_0$.
- On fixe la valeur de α.
- On choisit une statistique de test T dont on connait la loi sous H_0. On en calcule une réalisation t **à partir des données**.
- On détermine une **région de rejet** $\mathcal{R}$ telle que :
  - Si $t \in \mathcal{R}$, alors on décide de rejeter H_0 (d'accepter H_1). Le risque de se tromper est inférieur ou égal à α.
  - Si $t \notin \mathcal{R}$, alors on décide d'accepter H_0.
- Lorsque H_1 est de la forme $H_1 : \theta \neq \theta_0$, le test est dit **bilatéral**. Lorsque H_1 est de la forme $H_1 : \theta > \theta_0$ ou $H_1 : \theta < \theta_0$, le test est dit **unilatéral**.

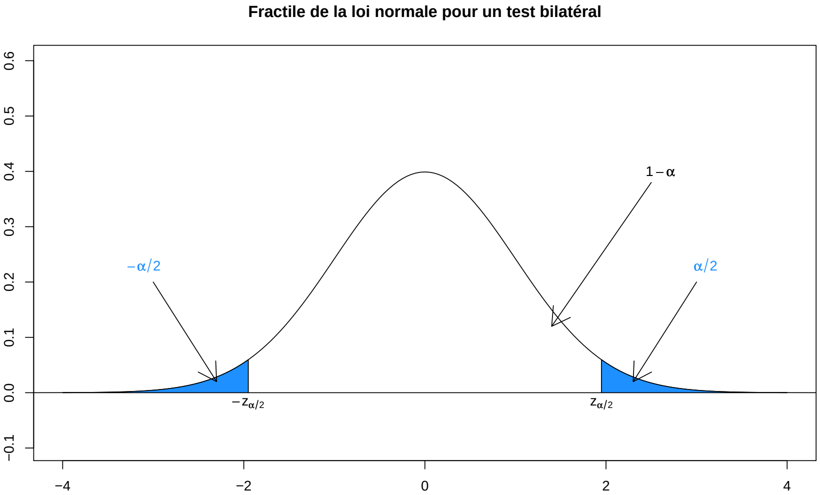
Prise de décision au moyen de la p-value

- On veut tester l'hypothèse nulle $H_0 : \theta = \theta_0$ contre l'hypothèse alternative $H_1 : \theta \neq \theta_0$ ou $H_1 : \theta < \theta_0$ ou $H_1 : \theta > \theta_0$.
- On fixe la valeur de α.
- On choisit une statistique de test T dont on connait la loi sous H_0. On en calcule une réalisation t à partir des données.
- On détermine la p-value p qui représente la probabilité d'obtenir des résultats au moins autant éloignés de H_0 que ceux obtenus pour t, et ce, alors que H_0 est vraie.
- Si $p \leq \alpha$, alors on décide de rejeter H_0 (d'accepter H_1). Le risque de se tromper est inférieur ou égal à α.
- Si $p > \alpha$, alors on décide de ne pas rejeter H_0.

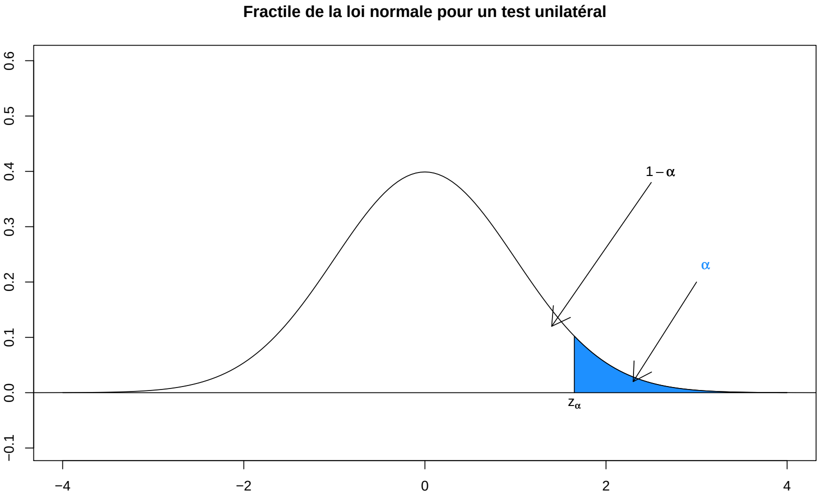
Fractiles de la loi normale et de la loi de Student

- Soit $\mathbf{z}_{\alpha/2}$ le fractile **bilatéral** de la loi $\mathcal{N}(0,1)$. Il est défini par $\mathbb{P}[|Z| > z_{\alpha/2}] = \alpha$ ou de manière équivalente par $\mathbb{P}[Z \leq z_{\alpha/2}] = 1 - \alpha/2$ pour une variable Z de loi $\mathcal{N}(0,1)$.
- Soit $\mathbf{z}_{\alpha}$ le fractile **unilatéral** de la loi $\mathcal{N}(0,1)$. Il est défini par $\mathbb{P}[Z > z_{\alpha}] = \alpha$ ou de manière équivalente par $\mathbb{P}[Z \leq z_{\alpha}] = 1 - \alpha$ pour une variable Z de loi $\mathcal{N}(0,1)$.
- Soit $t_{\alpha/2}$ le fractile **bilatéral** de la loi de Student à d degrés de liberté $T(d)$. Il est défini par $\mathbb{P}[|T| > t_{\alpha/2}] = \alpha$ ou de manière équivalente par $\mathbb{P}[T \leq t_{\alpha/2}] = 1 - \alpha/2$ pour une variable T de loi $T(d)$.
- Soit t_{α} le fractile **unilatéral** de $T(d)$. Il est défini par $\mathbb{P}[T > t_{\alpha}] = \alpha$ ou de manière équivalente par $\mathbb{P}[T \leq t_{\alpha}] = 1 - \alpha$ pour une variable T de loi $T(d)$.

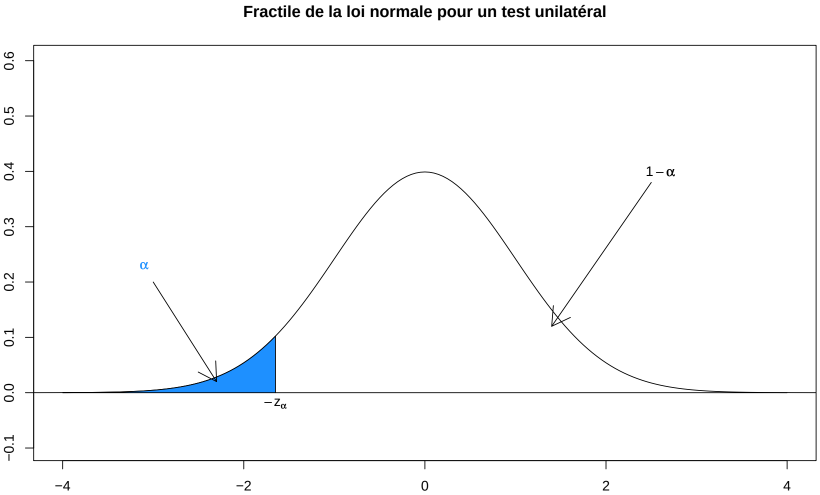
Loi normale : fractile bilatéral



Loi normale : fractile unilatéral supérieur



Loi normale : fractile unilatéral inférieur



1ère partie

Comparaisons de moyennes

Comparer une moyenne à une valeur de référence

- Soit $(X_1, ...X_n)$ un échantillon de taille n ce qui signifie que les variables $X_1, ...X_n$ sont indépendantes et de même loi qu'une variable parente notée X.
- On veut comparer $\mathbb{E}[X]$ à une valeur m_0.
- On note :

$$\overline{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \,,$$

$$S^2 = \frac{1}{n-1} \left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2 \right) .$$

Comparaison de $\mathbb{E}[X]$ à m_0 : one sample z-test pour échantillon gaussien avec variance connue

- **Conditions d'application** : les X_i suivent une loi $\mathcal{N}(m, \sigma^2)$ avec σ^2 **connue**.
- La statistique de test Z suit une loi $\mathcal{N}(0, 1)$ sous H_0 :

$$Z = \frac{\overline{X} - m_0}{\sqrt{\sigma^2/n}}.$$

- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = m_0$ contre $\mathbf{H_1}$: $\mathbb{E}[X] \neq m_0$, on calcule la réalisation z, puis on rejette H_0 si $|z| > z_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = m_0$ contre $\mathbf{H_1}$: $\mathbb{E}[X] > m_0$, on calcule la réalisation z, puis on rejette H_0 si $z > z_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = m_0$ contre $\mathbf{H_1}$: $\mathbb{E}[X] < m_0$, on calcule la réalisation z, puis on rejette H_0 si $z < -z_\alpha$, sinon on ne rejette pas H_0.

Comparaison de $\mathbb{E}[X]$ à m_0 : one sample t-test pour échantillon gaussien avec variance inconnue

- **Conditions d'application** : les X_i suivent une loi $\mathcal{N}(m, \sigma^2)$ avec σ^2 **inconnue** (test **robuste** par rapport aux écarts à la normalité).
- La statistique de test T suit une loi de Student $T(n-1)$ sous H_0 :

$$T = \frac{\overline{X} - m_0}{\sqrt{S^2/n}}.$$

- Soit $t_{\alpha/2}$ le fractile bilatéral de la loi $T(n-1)$ et soit t_α le fractile unilatéral de la loi $T(n-1)$.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = m_0$ contre $\mathbf{H_1}$: $\mathbb{E}[X] \neq m_0$, on calcule la réalisation t, puis on rejette H_0 si $|t| > t_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = m_0$ contre $\mathbf{H_1}$: $\mathbb{E}[X] > m_0$, on calcule la réalisation t, puis on rejette H_0 si $t > t_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = m_0$ contre $\mathbf{H_1}$: $\mathbb{E}[X] < m_0$, on calcule la réalisation t, puis on rejette H_0 si $t < -t_\alpha$, sinon on ne rejette pas H_0.

Comparaison de $\mathbb{E}[X]$ à m_0 : one sample z-test pour grand échantillon avec variance inconnue

- **Conditions d'application** : $n \geq 30$.
- La statistique de test Z suit approximativement une loi $\mathcal{N}(0,1)$ sous H_0 :
$$Z = \frac{\overline{X} - m_0}{\sqrt{S^2/n}}.$$
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = m_0$ contre $\mathbf{H_1}$: $\mathbb{E}[X] \neq m_0$, on calcule la réalisation z, puis on rejette H_0 si $|z| > z_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = m_0$ contre $\mathbf{H_1}$: $\mathbb{E}[X] > m_0$, on calcule la réalisation z, puis on rejette H_0 si $z > z_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = m_0$ contre $\mathbf{H_1}$: $\mathbb{E}[X] < m_0$, on calcule la réalisation z, puis on rejette H_0 si $z < -z_\alpha$, sinon on ne rejette pas H_0.

Exemple de comparaison de $\mathbb{E}[X]$ à m_0

- On ensemence 100 boîtes de Pétri avec $1cm^3$ de solution bactérienne de titre 8 bactéries/cm^3 à laquelle on a ajouté un traitement dont on espère qu'il modifie les conditions de culture. On observe que le nombre moyen de colonies sur les 100 boîtes de Pétri est $\overline{x} = 12$ avec $s^2 = 403$.
- On veut tester H_0 : "le traitement n'a eu aucun effet sur la culture" contre H_1 : "le traitement a eu un effet sur la culture". Autrement écrit, si l'on note X la variable représentant le nombre de bactéries survivant au traitement, on veut tester $H_0 : \mathbb{E}[X] = 8$ contre $H_1 : \mathbb{E}[X] \neq 8$.
- On choisit le risque de 1ère espèce $\alpha = 0.05$.
- $n = 100 \geq 30$ donc on effectue un one-sample z-test pour grand échantillon. La statistique de test est $Z = \frac{\overline{X} - m_0}{\sqrt{S^2/n}}$.
- On calcule la réalisation $z = 1.99$ et on détermine $z_{\alpha/2}$ par $\mathbb{P}[\mathcal{N}(0,1) \leq z_{\alpha/2}] = 1 - 0.05/2 = 0.975$, on lit dans la table $z_{\alpha/2} = 1.96$.
- On voit que $|z| > 1.96$ donc on rejette H_0.
- Conclusion : le traitement a eu un effet significatif sur la culture au risque 5%.

Exemple de comparaison de $\mathbb{E}[X]$ à m_0

- On considère que la valeur moyenne "normale" du taux de cholestérol est 2g/L. Dans une population de patients bien définie cliniquement, on a prélevé au hasard un échantillon de 150 sujets sur lesquels on a mesuré le taux de cholestérol X en g/L. Les résultats ont été les suivants : $\sum_{i=1}^{n} x_i = 273\text{g/L}$, $\sum_{i=1}^{n} x_i^2 = 512.12(\text{g/L})^2$. On veut tester si la population de patients en question présente un taux de cholestérol nominal.
- On teste $H_0 : m = 2\text{g/L}$ contre $H_1 : m \neq 2$ au risque $\alpha = 0.05$.
- $n = 150$ ($n \geq 30$) donc on peut utiliser un one-sample z-test. La statistique de test est $Z = \frac{\overline{X} - m_0}{\sqrt{S^2/n}}$.
- On calcule à partir de l'échantillon $\overline{x} = \frac{273}{150} = 1.82\text{g/L}$ et $s^2 = \frac{512.12 - 150 \times 1.82^2}{149} = 0.102(\text{g/L})^2$ puis $z = \frac{1.82 - 2.00}{\sqrt{\frac{0.102}{150}}} = -6.97$.
- Valeur critique : fractile bilatéral de la loi $\mathcal{N}(0, 1)$ correspondant au risque $\alpha = 5\%$: $z_{\alpha/2} = 1.96$.
- $|z| = 6.97 > 1.96$ donc on rejette H_0.
- Conclusion : dans cette population de patients, le taux moyen de cholestérol est significativement différent de la valeur de référence 2g/L au risque 5%.

Comparaison de 2 moyennes

- Soit $(X_1, ...X_{n_1})$ échantillon de taille n_1 et de variable parente X et soit $(Y_1, ...Y_{n_2})$ échantillon de taille n_2 et de variable parente Y.
- On veut comparer $\mathbb{E}[X]$ à $\mathbb{E}[Y]$.
- Lorsque $\text{Var}(X) = \text{Var}(Y)$, on dit qu'il y a **homoscédasticité**.
- On note :

$$\begin{aligned}
\overline{X} &= \frac{1}{n_1}\sum_{i=1}^{n_1} X_i \\
\overline{Y} &= \frac{1}{n_2}\sum_{i=1}^{n_2} Y_i \\
S_1^2 &= \frac{1}{n_1 - 1}\left(\sum_{i=1}^{n_1} X_i^2 - n_1\overline{X}^2\right) \\
S_2^2 &= \frac{1}{n_2 - 1}\left(\sum_{i=1}^{n_2} Y_i^2 - n_2\overline{Y}^2\right)
\end{aligned}$$

Comparaison de 2 moyennes : two-sample z-test pour échantillons gaussiens indépendants avec variances connues

- **Conditions d'application** : les X_i suivent une loi $\mathcal{N}(m_1, \sigma_1^2)$ avec σ_1^2 **connue** et les Y_i suivent une loi $\mathcal{N}(m_2, \sigma_2^2)$ avec σ_2^2 **connue**.
- La statistique de test Z suit une loi $\mathcal{N}(0, 1)$ sous H_0 :

$$Z = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] \neq \mathbb{E}[Y]$, on calcule la réalisation z, puis on rejette H_0 si $|z| > z_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] > \mathbb{E}[Y]$, on calcule la réalisation z, puis on rejette H_0 si $z > z_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] < \mathbb{E}[Y]$, on calcule la réalisation z, puis on rejette H_0 si $z < -z_\alpha$, sinon on ne rejette pas H_0.

Comparaison de 2 moyennes : two-sample t-test pour échantillons gaussiens homoscédastiques indépendants

- **Conditions d'application** : les X_i suivent une loi $\mathcal{N}(m_1, \sigma^2)$, les Y_i suivent une loi $\mathcal{N}(m_2, \sigma^2)$ où la variance commune σ^2 **inconnue**.
- La statistique de test T suit une loi de Student $T(n_1 + n_2 - 2)$ sous H_0 :

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S^2}{n_1} + \frac{S^2}{n_2}}}$$

avec

$$S^2 = \frac{\sum_{i=1}^{n_1}(X_i - \overline{X})^2 + \sum_{i=1}^{n_2}(Y_i - \overline{Y})^2}{n_1 + n_2 - 2}.$$

Autre expression utile de S^2 :

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{(n_1 - 1) + (n_2 - 1)}.$$

Comparaison de 2 moyennes : two-sample t-test pour échantillons gaussiens homoscédastiques indépendants (suite)

- Soit $t_{\alpha/2}$ le fractile bilatéral de la loi $T(n_1 + n_2 - 2)$ et soit t_α le fractile unilatéral de la loi $T(n_1 + n_2 - 2)$.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] \neq \mathbb{E}[Y]$, on calcule la réalisation t, puis on rejette H_0 si $|t| > t_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] > \mathbb{E}[Y]$, on calcule la réalisation t, puis on rejette H_0 si $t > t_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] < \mathbb{E}[Y]$, on calcule la réalisation t, puis on rejette H_0 si $t < -t_\alpha$, sinon on ne rejette pas H_0.

Exemple de comparaison de 2 moyennes : two-sample t-test pour échantillons indépendants

- On se demande si le taux sanguin est modifié chez les sujets atteints d'une maladie rare M. On a mesuré le taux sanguin chez 10 malades et on a trouvé $\overline{x} = 91$ avec $s_1^2 = 35$. On a aussi mesuré le taux sanguin chez 20 sujets sains et on a trouvé $\overline{y} = 91$ avec $s_2^2 = 27$.
- On veut tester H_0 : "le taux sanguin est identique chez les malades et les sujets sains" contre H_1 : "le taux sanguin est différent chez les malades et les sujets sains". Autrement dit, si l'on note X la variable représentant le taux sanguin d'un malade et si l'on note Y la variable représentant le taux sanguin d'un sujet sain, on veut tester $H_0 : \mathbb{E}[X] = \mathbb{E}[Y]$ contre $H_1 : \mathbb{E}[X] \neq \mathbb{E}[Y]$. On fixe le risque de 1ère espèce $\alpha = 0.05$.
- Les 2 échantillons sont indépendants, de petite taille $n_1 = 10$ et $n_2 = 20$, les variances sont inconnues donc on pense au two-sample t-test.

Exemple de comparaison de 2 moyennes (suite)

- L'usage du test de Student suppose que le taux sanguin est distribué normalement chez les malades et chez les témoins et que les variances de X et de Y sont égales. Avec les données dont on dispose, on ne peut pas vérifier ces hypothèses (voir comparaison de distributions et de variances). Mais on utilise quand même le test de Student car il est réputé robuste.
- On calcule $s^2 = \frac{19\times 27+9\times 35}{19+9} = 30$ puis $t = \frac{91-85}{\sqrt{\frac{30}{20}+\frac{30}{10}}}$.
- On détermine $t_{\alpha/2}$ par $\mathbb{P}[T(10+20-2) \leq t_{\alpha/2}] = 1 - 0.05/2 = 0.975$, on trouve $t_{\alpha/2} = 2.05$.
- $|t| = 2.83 > 2.05$ donc on rejette H_0 et on conclut que le taux sanguin est supérieur chez les malades au risque $\alpha = 0.05$.
- NB : dans la réalité, avec si peu de données, il faut s'assurer de la représentativité de l'échantillon par rapport à la population de référence qui doit être bien définie.

Test d'Aspin-Welch

- C'est un t-test pour échantillons gaussiens indépendants SANS hyp d'homoscédasticité.
- **Conditions d'application** : les X_i suivent une loi $\mathcal{N}(m_1, \sigma_1^2)$ avec σ_1^2 **inconnue**, les Y_i suivent une loi $\mathcal{N}(m_2, \sigma_2^2)$ avec σ_2^2 **inconnue**.
- La statistique de test T suit une loi de Student $T(\nu)$ sous H_0 :

$$T = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

où ν est l'entier le plus proche de

$$\frac{\left(S_1^2/n_1 + S_2^2/n_2\right)^2}{\left(S_1^2/n_1\right)^2 \times \frac{1}{n_1-1} + \left(S_2^2/n_2\right)^2 \times \frac{1}{n_2-1}}.$$

- Soit $t_{\alpha/2}$ le fractile bilatéral de la loi $T(\nu)$ et soit t_α le fractile unilatéral de la loi $T(\nu)$.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] \neq \mathbb{E}[Y]$, on calcule la réalisation t, puis on rejette H_0 si $|t| > t_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] > \mathbb{E}[Y]$, on calcule la réalisation t, puis on rejette H_0 si $t > t_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] < \mathbb{E}[Y]$, on calcule la réalisation t, puis on rejette H_0 si $t < -t_\alpha$, sinon on ne rejette pas H_0.

Comparaison de 2 moyennes : two-sample z-test pour grands échantillons indépendants avec variances inconnues

- **Conditions d'application** : $n_1 \geq 30$ et $n_2 \geq 30$.
- La statistique de test Z suit approximativement une loi $\mathcal{N}(0,1)$ sous H_0 :

$$Z = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}.$$

- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] \neq \mathbb{E}[Y]$, on calcule la réalisation z, puis on rejette H_0 si $|z| > z_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] > \mathbb{E}[Y]$, on calcule la réalisation z, puis on rejette H_0 si $z > z_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] < \mathbb{E}[Y]$, on calcule la réalisation z, puis on rejette H_0 si $z < -z_\alpha$, sinon on ne rejette pas H_0.

Exemple de comparaison de 2 moyennes : two-sample z-test pour grands échantillons indépendants

- Deux groupes de 100 souris ont reçu les traitements A et B resp. Après traitement, la durée de vie moyenne dans le groupe A était de $\overline{x} = 114$j avec $s_1^2 = 410$ tandis que dans le groupe B, la durée de vie moyenne était de $\overline{y} = 119$j avec $s_2^2 = 490$.
- On veut tester H_0 : "la durée de vie moyenne est identique dans les groupes A et B" contre H_1 : "la durée de vie moyenne est différente dans les groupes A et B". Autrement dit, si l'on note X la variable représentant la durée de vie d'une souris dans le groupe A et si l'on note Y la variable représentant la durée de vie dans le groupe B, on veut tester H_0 : $\mathbb{E}[X] = \mathbb{E}[Y]$ contre H_1 : $\mathbb{E}[X] \neq \mathbb{E}[Y]$. On fixe le risque à $\alpha = 0.05$.
- Les 2 échantillons sont indépendants, $n_1 = n_2 = n = 100$ donc on utilise un z-test avec variances estimées.
- On calcule $z = \frac{114-119}{\sqrt{\frac{410}{100}+\frac{490}{100}}} = -1.67$.
- On détermine $z_{\alpha/2}$ par $\mathbb{P}[\mathcal{N}(0,1) \leq z_{\alpha/2}] = 1 - 0.05/2 = 0.975$, on lit dans la table $z_{\alpha/2} = 1.96$.
- On voit que $|z| \leq 1.96$ donc on ne rejette pas H_0.

Exemple de comparaison de 2 moyennes : two-sample z-test pour grands échantillons indépendants

- On étudie l'efficacité d'un régime hypocholestérolémiant sur des animaux comparativement à des animaux témoins (non traités). On mesure le taux de cholestérol X dans les deux groupes et on obtient les résultats suivants :
  Groupe traité : $n_1 = 100$, $\sum_{i=1}^{n_1} x_{1,i} = 250$g/L et $\sum_{i=1}^{n_1} x_{1,i}^2 = 1025$(g/L)^2
  Groupe témoin : $n_2 = 100$, $\sum_{i=1}^{n_2} x_{2,i} = 400$g/L et $\sum_{i=1}^{n_2} x_{2,i}^2 = 2140$(g/L)^2
  On souhaite tester l'efficacité du régime hypocholestérolémiant au risque 5%.
- On teste $H_0 : m_1 = m_2$ contre $H_1 : m_1 < m_2$.
- Les effectifs sont $n_1 = 80$ et $n_2 = 90$ donc supérieurs à 30 donc on utilise un two-sample z-test.
- On calcule : $\overline{x}_1 == \frac{250}{100} = 2.50$g/L, $\overline{x}_2 = \frac{400}{100} = 4.00$g/L,
  $S_1^2 = \frac{1025-100\times 2.50^2}{99} = 4.04$g/L^2, $S_2^2 = \frac{2140-100\times 4.00^2}{99} = 5.45$g/L^2 puis

$$z = \frac{2.50 - 4.00}{\sqrt{\frac{4.04}{100} + \frac{5.45}{100}}} = -4.869$$

- Valeur critique : fractile unilatéral inférieur de la loi $\mathcal{N}(0,1)$ correspondant au risque $\alpha = 0.05$: $-z_\alpha = -1.65$.
- $z = -4.869 < -1.65$ donc on rejette H_0.
- Conclusion : au risque 5%, le traitement hypocholestérolémiant est efficace.

Comparaison de 2 moyennes : two-sample z-test pour échantillons gaussiens indépendants avec variances connues

- **Conditions d'application** : les X_i suivent une loi $\mathcal{N}(m_1, \sigma_1^2)$ avec σ_1^2 **connue** et les Y_i suivent une loi $\mathcal{N}(m_2, \sigma_2^2)$ avec σ_2^2 **connue**.
- La statistique de test Z suit une loi $\mathcal{N}(0, 1)$ sous H_0 :

$$Z = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] \neq \mathbb{E}[Y]$, on calcule la réalisation z, puis on rejette H_0 si $|z| > z_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] > \mathbb{E}[Y]$, on calcule la réalisation z, puis on rejette H_0 si $z > z_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1}$: $\mathbb{E}[X] < \mathbb{E}[Y]$, on calcule la réalisation z, puis on rejette H_0 si $z < -z_\alpha$, sinon on ne rejette pas H_0.

Comparaison de 2 moyennes : z-test pour grands échantillons appariés

- Soient $(X_1, ..., X_n)$ et $(Y_1, ..., Y_n)$ deux échantillons appariés (de même taille forcément !).
- On crée un unique échantillon $(D_1, ..., D_n)$ où $D_i = X_i - Y_i$.
- Tester l'hypothèse $H_0 : \mathbb{E}[X] = \mathbb{E}[Y]$ revient à tester $H_0 : \mathbb{E}[D] = 0$.
- **Conditions d'application** : $n \geq 30$.
- La statistique de test Z suit une loi $\mathcal{N}(0,1)$ sous H_0 :

$$Z = \frac{\overline{D}}{\sqrt{S^2/n}}$$

avec $\overline{D} = \frac{1}{n} \sum_{i=1}^{n} D_i$ et $S^2 = \frac{1}{n-1} \left(\sum_{i=1}^{n} D_i^2 - n\overline{D}^2 \right)$.

- Pour tester $\mathbf{H_0} : \mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1} : \mathbb{E}[X] \neq \mathbb{E}[Y]$, on calcule la réalisation z, puis on rejette H_0 si $|z| > z_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0} : \mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1} : \mathbb{E}[X] > \mathbb{E}[Y]$, on calcule la réalisation z, puis on rejette H_0 si $z > z_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0} : \mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1} : \mathbb{E}[X] < \mathbb{E}[Y]$, on calcule la réalisation z, puis on rejette H_0 si $z < -z_\alpha$, sinon on ne rejette pas H_0.

Exemple de comparaison de 2 moyennes : z-test pour grands échantillons appariés

- Un essai thérapeutique visant à évaluer l'efficacité de deux traitements a été réalisé en cross-over (i.e. chaque patient est pris comme son propre témoin et reçoit les deux médicaments dont l'ordre d'administration est tiré au sort). Les valeurs suivantes de taux de cholestérol X en g/L ont été obtenues après administration de chaque traitement :
  Après traitement A : $n_1 = n = 100$, $\sum_{i=1}^{n_1} x_{1,i} = 269$, $\sum_{i=1}^{n_1} x_{1,i}^2 = 1370$,
  Après traitement B : $n_2 = n = 100$, $\sum_{i=1}^{n_2} x_{2,i} = 302$, $\sum_{i=1}^{n_1} x_{2,i}^2 = 2899$,
  Différence : $\sum_{i=1}^{n}(x_{1,i} - x_{2,i}) = -33$, $\sum_{i=1}^{n}(x_{1,i} - x_{2,i})^2 = 194$.
  On veut tester l'efficacité du traitement au risque $\alpha = 0.05$ et au risque $\alpha = 0.01$.
- On teste $H_0 : m_1 = m_2$ contre $H_1 : m_1 \neq m_2$.

Exemple de comparaison de 2 moyennes : z-test pour grands échantillons appariés (suite)

- Les échantillons sont appariés donc on travaille sur l'échantillon des différences $d_i = x_{1,i} - x_{2,i}$ pour $i = 1, ..., 100$. On effectue le test de $H_0 : \mathbb{E}[D] = 0$ contre $H_1 : \mathbb{E}[D] \neq 0$ (comparaison d'une moyenne à une valeur de référence).
- L'effectif est $n = 100 > 30$ donc on utilise un one-sample z-test.
- On calcule $\overline{d} = \frac{-33}{100} = -0.33$, $S^2 = \frac{194-100\times(-0.33)^2}{99} = 1.85$ puis la réalisation de la statistique de test $z = \frac{-0.33-0}{\sqrt{1.85/100}} = -2.43$
- Valeur critique : fractile bilatéral de la loi $\mathcal{N}(0,1)$ correspondant au risque $\alpha = 0.05$: $z_{\alpha/2} = 1.96$, au risque $\alpha = 0.01$: $z_{\alpha/2} = 2.58$.
- $|z| = 2.43 > 1.96$ donc on rejette H_0 au risque 5% mais $|z| = 2.43 < 2.58$ donc on ne rejette pas H_0 au risque 1%.
- Conclusion : au risque 5%, il existe une différence significative entre A et B mais pas au risque 1%.

Comparaison de 2 moyennes : t-test pour échantillons gaussiens appariés

- Soient $(X_1, ..., X_n)$ et $(Y_1, ..., Y_n)$ deux échantillons appariés (de même taille forcément !).
- On crée un unique échantillon $(D_1, ..., D_n)$ où $D_i = X_i - Y_i$.
- Tester l'hyp $H_0 : \mathbb{E}[X] = \mathbb{E}[Y]$ revient à tester $H_0 : \mathbb{E}[D] = 0$.
- **Conditions d'application** : les D_i suivent une loi $\mathcal{N}(m, \sigma^2)$ où la variance σ^2 **inconnue**.
- La statistique de test T suit une loi de Student $T(n-1)$ sous H_0 :

$$T = \frac{\overline{D}}{\sqrt{S^2/n}}$$

avec $\overline{D} = \frac{1}{n}\sum_{i=1}^{n} D_i$ et $S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} D_i^2 - n\overline{D}^2\right)$.

- Soit $t_{\alpha/2}$ le fractile bilatéral de la loi $T(n-1)$ et soit t_α le fractile unilatéral de la loi $T(n-1)$.
- Pour tester $\mathbf{H_0} : \mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1} : \mathbb{E}[X] \neq \mathbb{E}[Y]$, on calcule la réalisation t, puis on rejette H_0 si $|t| > t_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0} : \mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1} : \mathbb{E}[X] > \mathbb{E}[Y]$, on calcule la réalisation t, puis on rejette H_0 si $t > t_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0} : \mathbb{E}[X] = \mathbb{E}[Y]$ contre $\mathbf{H_1} : \mathbb{E}[X] < \mathbb{E}[Y]$, on calcule la réalisation t, puis on rejette H_0 si $t < -t_\alpha$, sinon on ne rejette pas H_0.

Exemple de t-test sur échantillons appariés

- On souhaite évaluer l'action d'un anti-cancéreux sur une tumeur solide. Au moyen d'une technique d'imagerie, on mesure la taille (en mm) de la tumeur avant et après traitement. On obtient

patient	1	2	3	4	5	6	7
avant	20.4	25.4	25.6	25.6	26.6	28.6	28.7
après	21.7	26.3	26.8	28.1	26.2	27.3	29.5

patient	8	9	10	11	12
avant	29.0	29.8	30.5	30.9	31.1
après	32.0	30.9	32.3	32.3	31.7

Exemple de t-test sur échantillons appariés (suite)

- On note X la variable représentant la taille de la tumeur avant traitement et Y la variable représentant la taille de la tumeur après traitement. On veut tester $H_0 : \mathbb{E}[X] = \mathbb{E}[Y]$ contre $H_1 : \mathbb{E}[X] \neq \mathbb{E}[Y]$ à partir de 2 échantillons appariés. Cela revient à tester $H_0 : \mathbb{E}[D] = 0$ contre $H_1 : \mathbb{E}[D] \neq 0$ en notant $D = X - Y$.
- On construit un échantillon de taille $n = 12$ en effectuant les différences "avant"-"après". On obtient les valeurs : -1.3, -0.9, -1.2, -2.5, 0.4, 1.3, -0.8, -3.0, -1.1, -1.8, -1.4, -0.6.
- On peut montrer (test de Shapiro-Wilk) que l'échantillon des différences est gaussien. On effectue un t-test pour échantillon gaussien sur l'échantillon des différences.
- On calcule $\overline{d} = -1.075$, $s^2 = 1.326$ puis $t = -3.234$. On lit dans la table que $t_{\alpha/2}$, le fractile bilatéral de $T(11)$ défini par $\mathbb{P}[T(11) \leq t_{\alpha/2}] = 1 - 0.05/2 = 0.975$, vaut $t_{\alpha/2} = 2.201$ pour $\alpha = 0.05$. On rejette donc H_0 au risque $\alpha = 0.05$.

$2^{\text{ème}}$ partie

Comparaisons de variances

Comparer une variance à une valeur de référence

- Soit $(X_1, ... X_n)$ un échantillon de taille n de variable parente X. On veut comparer $\text{Var}(X)$ à une valeur σ_0^2. On teste l'hypothèse H_0 : $\text{Var}(X) = \sigma_0^2$.
- On note
$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right).$$
- On note α le risque de 1$^{\text{ère}}$ espèce.
- Soit $\mathbf{z}_{\alpha/\mathbf{2}}$ le fractile **bilatéral** de la loi $\mathcal{N}(0,1)$ défini par $\mathbb{P}[|Z| > z_{\alpha/2}] = \alpha$ ou de manière équivalente par $\mathbb{P}[Z \leq z_{\alpha/2}] = 1 - \alpha/2$ pour une variable Z de loi $\mathcal{N}(0,1)$.
- Soit $\mathbf{z}_{\alpha}$ le fractile **unilatéral** de la loi $\mathcal{N}(0,1)$ défini par $\mathbb{P}[Z > z_{\alpha}] = \alpha$ ou de manière équivalente par $\mathbb{P}[Z \leq z_{\alpha}] = 1 - \alpha$ pour une variable Z de loi $\mathcal{N}(0,1)$.

Comparer une variance à une valeur de référence : one-sample χ^2-test pour échantillon gaussien

- **Conditions d'application** : les X_i suivent une loi $\mathcal{N}(m, \sigma^2)$ avec σ^2 **inconnue**
- La statistique de test K suit une loi $\chi^2(n-1)$ sous H_0 :
$$K = (n-1)\frac{S^2}{\sigma_0^2}.$$
- Pour tester $\mathbf{H_0}$: $\mathrm{Var}(X) = \sigma_0^2$ contre $\mathbf{H_1}$: $\mathrm{Var}(X) \neq \sigma_0^2$, on calcule la réalisation k, puis on rejette H_0 si $k < k_1$ ou $k > k_2$ avec k_1 et k_2 déterminés par $\mathbb{P}[\chi^2(n-1) < k_1] = \alpha/2$ et $\mathbb{P}[\chi^2(n-1) > k_2] = \alpha/2$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathrm{Var}(X) = \sigma_0^2$ contre $\mathbf{H_1}$: $\mathrm{Var}(X) > \sigma_0^2$, on calcule la réalisation k, puis on rejette H_0 si $k > k_2$ avec k_2 déterminé par $\mathbb{P}[\chi^2(n-1) > k_2] = \alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathrm{Var}(X) = \sigma_0^2$ contre $\mathbf{H_1}$: $\mathrm{Var}(X) < \sigma_0^2$, on calcule la réalisation k, puis on rejette H_0 si $k < k_1$ avec k_1 déterminé par $\mathbb{P}[\chi^2(n-1) < k_1] = \alpha$, sinon on ne rejette pas H_0.

Comparer une variance à une valeur de référence : one-sample z-test pour grand échantillon

- **Conditions d'application** : $n \geq 30$.
- La statistique de test Z suit approximativement une loi $\mathcal{N}(0, 1)$ sous H_0 :

$$Z = \frac{S^2 - \sigma_0^2}{\sqrt{2\sigma_0^4/(n-1)}}.$$

- Pour tester $\mathbf{H_0}$: $\text{Var}(X) = \sigma_0^2$ contre $\mathbf{H_1}$: $\text{Var}(X) \neq \sigma_0^2$, on calcule la réalisation z, puis on rejette H_0 si $|z| > z_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\text{Var}(X) = \sigma_0^2$ contre $\mathbf{H_1}$: $\text{Var}(X) > \sigma_0^2$, on calcule la réalisation z, puis on rejette H_0 si $z > z_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\text{Var}(X) = \sigma_0^2$ contre $\mathbf{H_1}$: $\text{Var}(X) < \sigma_0^2$, on calcule la réalisation z, puis on rejette H_0 si $z < -z_\alpha$, sinon on ne rejette pas H_0.

Comparaison de 2 variances

- Soit $(X_1, ...X_{n_1})$ échantillon de taille n_1 et de variable parente X et soit $(Y_1, ...Y_{n_2})$ échantillon de taille n_2 et de variable parente Y.
- On veut comparer $\mathrm{Var}(X)$ à $\mathrm{Var}(Y)$ (test d'homoscédasticité). On teste l'hypothèse H_0 : $\mathrm{Var}(X) = \mathrm{Var}(Y)$.
- On note :

$$\begin{aligned}
S_1^2 &= \frac{1}{n_1 - 1}\left(\sum_{i=1}^{n_1} X_i^2 - n_1 \overline{X}^2\right) \\
S_2^2 &= \frac{1}{n_2 - 1}\left(\sum_{i=1}^{n_2} Y_i^2 - n_2 \overline{Y}^2\right) \\
S^2 &= \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}
\end{aligned}$$

Comparaison de 2 variances, two sample f-test pour échantillons gaussiens indépendants

- **Conditions d'applications** : les X_i et les Y_i sont de loi gaussienne (test réputé robuste à l'hypothèse de normalité).
- La statistique de test $F = S_1^2/S_2^2$ suit une loi de Fisher $F(n_1-1, n_2-1)$ sous H_0.
- Pour tester $\mathbf{H_0}$: $\text{Var}(X) = \text{Var}(Y)$ contre $\mathbf{H_1}$: $\text{Var}(X) \neq \text{Var}(Y)$,
  - lorsque $s_1^2 > s_2^2$, on calcule la réalisation $f' = s_1^2/s_2^2$, puis on rejette H_0 si $f' > f_{\alpha/2}$, sinon on ne rejette pas H_0, $f_{\alpha/2}$ étant déterminé par $\mathbb{P}[F(n_1-1, n_2-1) \leq f_{\alpha/2}] = 1-\alpha/2$,
  - lorsque $s_2^2 > s_1^2$, on calcule la réalisation $f' = s_2^2/s_1^2$, puis on rejette H_0 si $f' > f_{\alpha/2}$, sinon on ne rejette pas H_0, $f_{\alpha/2}$ étant déterminé par $\mathbb{P}[F(n_2-1, n_1-1) \leq f_{\alpha/2}] = 1-\alpha/2$.
- Pour tester $\mathbf{H_0}$: $\text{Var}(X) = \text{Var}(Y)$ contre $\mathbf{H_1}$: $\text{Var}(X) > \text{Var}(Y)$, on calcule $f' = s_1^2/s_2^2$, puis on rejette H_0 si $f > f_\alpha$ où f_α est déterminé par $\mathbb{P}[F(n_1-1, n_2-1) \leq f_\alpha] = 1-\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\text{Var}(X) = \text{Var}(Y)$ contre $\mathbf{H_1}$: $\text{Var}(X) < \text{Var}(Y)$, on calcule $f' = s_2^2/s_1^2$, puis on rejette H_0 si $1/f > f_\alpha$ où f_α est déterminé par $\mathbb{P}[F(n_2-1, n_1-1) \leq f_\alpha] = 1-\alpha$, sinon on ne rejette pas H_0.

Exemple de comparaison de 2 variances, two sample f-test pour échantillons gaussiens indépendants

- Pour fabriquer une solution contenant une certaine espèce chimique, on dispose de 2 procédés A et B. On se demande si les deux procédés fournissent des solutions dont le titre est bien homogène ou s'il existe une différence entre les deux procédés. On fabrique alors deux séries de solutions, l'une avec le procédé A, l'autre avec le procédé B, dont on vérifie le titre. L'homogénéité est quantifiée par la variance observée sur les résultats obtenus pour un même procédé (plus une méthode est homogène, plus la variance est petite).
- Notons σ_A^2 la variance du titre obtenu avec le procédé A et σ_B^2 la variance du titre obtenu avec le procédé B. On veut tester $H_0 : \sigma_A^2 = \sigma_B^2$ contre $H_1 : \sigma_A^2 \neq \sigma_B^2$.
- On réalise 20 aliquots de solution avec A et 25 avec B. On a donc $n_A = 20$ et $n_B = 25$. On trouve $s_A^2 = 37$ et $s_B^2 = 41$.
- On calcule $f' = 41/37 = 1.108$
- On lit dans la table la valeur critique pour une loi $F(24, 19)$: $f_{\alpha/2} = 2.452$ pour $\alpha = 0.05$. On a $f' < 2.452$ donc on ne rejette pas H_0.

Exemple de comparaison de 2 variances, two sample f-test pour échantillons gaussiens indépendants

- Pour fabriquer une solution contenant une certaine espèce chimique, on utilise un procédé A. Un labo vient proposer un procédé B dont il affirme qu'il est très fiable en terme d'homogénéité du titre obtenu. On se demande si le procédé B est plus homogène que le procédé A. On fabrique alors deux séries de solutions, l'une avec le procédé A, l'autre avec le procédé B dont on vérifie le titre. L'homogénéité est quantifiée par la variance observée sur les résultats obtenus pour un même procédé.
- Notons σ_A^2 la variance du titre obtenu avec le procédé A et σ_B^2 la variance du titre obtenu avec le procédé B. On veut tester $H_0 : \sigma_A^2 = \sigma_B^2$ contre $H_1 : \sigma_A^2 > \sigma_B^2$.
- On réalise 25 aliquots de solution avec A et 22 avec B. On a donc $n_A = 25$ et $n_B = 22$. On trouve $s_A^2 = 37$ et $s_B^2 = 35$.
- On calcule $f' = 37/35 = 1.057$
- On lit dans la table la valeur critique pour une loi $F(24, 21)$: $f_\alpha = 2.054$ pour $\alpha = 0.05$. On a $f' < 2.054$ donc on ne rejette pas H_0.

Comparaison de 2 variances, two sample z-test pour grands échantillons indépendants

- **Conditions d'applications** : $n_1 \geq 30$, $n_2 \geq 30$.
- La statistique de test Z suit approximativement une loi $\mathcal{N}(0,1)$ sous H_0 :
$$Z = \frac{S_1^2 - S_2^2}{S^2\sqrt{\frac{2}{n_1} + \frac{2}{n_2}}}.$$
- Pour tester $\mathbf{H_0}$: $\mathrm{Var}(X) = \mathrm{Var}(Y)$ contre $\mathbf{H_1}$: $\mathrm{Var}(X) \neq \mathrm{Var}(Y)$, on calcule la réalisation z, puis on rejette H_0 si $|z| > z_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathrm{Var}(X) = \mathrm{Var}(Y)$ contre $\mathbf{H_1}$: $\mathrm{Var}(X) > \mathrm{Var}(Y)$, on calcule la réalisation z, puis on rejette H_0 si $z > z_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $\mathrm{Var}(X) = \mathrm{Var}(Y)$ contre $\mathbf{H_1}$: $\mathrm{Var}(X) < \mathrm{Var}(Y)$, on calcule la réalisation z, puis on rejette H_0 si $z < -z_\alpha$, sinon on ne rejette pas H_0.

Exemple de comparaison de 2 variances, two sample z-test pour grands échantillons indépendants

- On dispose de 2 méthodes de dosage A et B. On se demande si l'une des deux méthodes est plus précise que l'autre. La précision est quantifiée par la variance observée sur les résultats obtenus pour une même substance à doser (plus une méthode est précise, plus la variance est petite).
- Notons σ_A^2 la variance d'un dosage réalisé avec la méthode A et σ_B^2 la variance d'un dosage réalisé avec la méthode B. On veut tester $H_0 : \sigma_A^2 = \sigma_B^2$ contre $H_1 : \sigma_A^2 \neq \sigma_B^2$.
- On réalise 200 aliquots d'une substance à doser dont 100 sont dosés avec A et 100 sont dosés avec B. On a donc $n_A = n_B = 100$. On trouve $s_A^2 = 37$ et $s_B^2 = 41$.
- On calcule $s^2 = 39$ et $$z = \frac{41 - 37}{39\sqrt{\frac{2}{100} + \frac{2}{100}}} = 0.50.$$
- On lit dans la table $z_{\alpha/2} = 1.96$ pour $\alpha = 0.05$. On a $z < 1.96$ donc on ne rejette pas H_0.

3$^{\text{ème}}$ partie

Comparaisons de proportions

Comparer une proportion à une valeur de référence : z-test

- On se demande si la proportion p de sujets présentant une certaine caractéristique diffère significativement d'une certaine valeur p_0. Pour cela, on dispose d'un échantillon de n sujets dans lequel la proportion $\widehat{p}$ présente la caractéristique en question.
- **Conditions d'application** : $np_0 > 5$ et $n(1 - p_0) > 5$.
- La statistique de test Z suit approximativement une loi $\mathcal{N}(0, 1)$ sous H_0 :
$$Z = \frac{\widehat{p} - p_0}{\sqrt{p_0(1 - p_0)/n}}.$$
- Pour tester $\mathbf{H_0}$: $p = p_0$ contre $\mathbf{H_1}$: $p \neq p_0$, on calcule la réalisation z, puis on rejette H_0 si $|z| > z_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $p = p_0$ contre $\mathbf{H_1}$: $p > p_0$, on calcule la réalisation z, puis on rejette H_0 si $z > z_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $\mathbf{H_0}$: $p = p_0$ contre $\mathbf{H_1}$: $p < p_0$, on calcule la réalisation z, puis on rejette H_0 si $z < -z_\alpha$, sinon on ne rejette pas H_0.

Exemple de comparaison d'une proportion à une valeur de référence : vaccin

- Sachant que 20% des personnes vaccinées par la formule F d'un vaccin présentent des allergies ou troubles secondaires, un laboratoire pharmaceutique propose une formule améliorée de ce vaccin Fa et espère diminuer le taux d'allergies et de troubles secondaires. Sur un échantillon de 400 personnes prises au hasard et ayant opté pour la formule Fa, on observe 60 cas d'allergies ou troubles secondaires. On veut tester au risque $\alpha = 0.05$ si la formule améliorée apporte réellement un bénéfice.
- On teste $H_0 : p = 0.20$ contre $H_1 : p < 0.20$.
- La réalisation de la statistique de test est

$$z = \frac{\frac{60}{400} - 0.20}{\sqrt{\frac{0.20 \times (1-0.20)}{400}}} = \frac{0.15 - 0.20}{\sqrt{0.2 \times 0.8/400}} = -2.5$$

- Valeur critique : fractile unilatéral inférieur de la loi $\mathcal{N}(0,1)$ pour $\alpha = 0.05$: $-z_\alpha = -1.65$, pour $\alpha = 0.01$: $-z_\alpha = -2.33$.
- $z < -1.645$ et $z < -2.33$ donc on rejette H_0 dans les deux cas.
- Conclusion : le taux d'allergies ou troubles secondaires est diminué par la formule Fa du vaccin au risque 5% et au risque 1%.

Exemple de comparaison d'un pourcentage à une valeur de référence : test de la 1ère loi de Mendel

- Soit un couple [A,a] de gènes où A est dominant. Le croisement d'hétérozygotes fournit les génotypes Aa, aA, aa et AA. Les génotypes Aa, aA et AA présentent le phénotype [A] tandis que le génotype aa présente le phénotype [a]. Lorsque le 1ère loi de Mendel s'applique, les proportions attendues de [A] et de [a] dans la descendance sont donc 3/4 de [A] et 1/4 de [a].
- Dans une descendance de 40 individus, on observe 27 individus [A].
- On veut tester si ces données permettraient de rejeter la loi de Mendel pour le caractère étudié avec un risque de 1ère espèce égal à 5%. Soit p la proportion d'individus présentant le phénotype [A]. On veut tester H_0 : $p = 3/4$ contre H_1 : $p \neq 3/4$.
- On calcule :
$$z = \frac{\frac{27}{40} - \frac{3}{4}}{\sqrt{\frac{3}{4} \times \frac{1}{4} \times \frac{1}{40}}} = -1.095.$$
- Valeur critique : fractile bilatéral de la loi $\mathcal{N}(0,1)$ pour $\alpha = 0.05$: $z_{\alpha/2} = 1.96$.
- $|z| = 1.095 < 1.96$ donc on ne rejette pas H_0.

Comparer une proportion à une valeur de référence : approximation poissonnienne

- Lorsque les conditions d'application ne sont pas vérifiées i.e. si $np_0 \leq 5$ et/ou $n(1-p_0) \leq 5$, l'approximation normale ne s'applique plus et il faut recourir à d'autres tests. Ceci arrive dans deux cas :
- **1er cas** : l'échantillon est grand mais p_0 est très petit (ne pas utiliser la loi normale !). Lorsque $p < 0.1$, on utilise le fait que la loi exacte de $n\widehat{p} = \sum_{i=1}^{n} X_i$ sous H_0 qui est une loi $\mathcal{B}(n, p_0)$ peut s'approximer par la loi $\mathcal{P}(np_0)$.
- Pour tester $H_0 : p = p_0$ contre $H_1 : p \neq p_0$, on détermine le plus grand k_1 tel que $\mathbb{P}[\mathcal{P}(np_0) \leq k_1] \leq \alpha/2$ et le plus petit k_2 tel que $\mathbb{P}[\mathcal{P}(np_0) \geq k_2] \leq \alpha/2$. On rejette H_0 lorsque $\widehat{p} < k_1/n$ ou $\widehat{p} > k_2/n$.
- Pour tester $H_0 : p = p_0$ contre $H_1 : p > p_0$, on détermine le plus grand k_1 tel que $\mathbb{P}[\mathcal{P}(np_0) \leq k_1] \leq \alpha$. On rejette H_0 si $\widehat{p} > k_1/n$.
- Pour tester $H_0 : p = p_0$ contre $H_1 : p < p_0$, on détermine le plus petit k_2 tel que $\mathbb{P}[\mathcal{P}(np_0) \geq k_2] \leq \alpha$. On rejette H_0 si $\widehat{p} < k_2/n$.

Exemple de comparaison d'une proportion à une valeur de référence par approximation poissonnienne

- On effectue une surveillance des effets indésirables d'un médicament. On considère qu'un risque inférieur à $1/1000$ de subir un effet indésirable est "acceptable".
- On a observé 2 évènements indésirables sur $n = 5000$ prescriptions.
- Soit p la proportion d'effets indésirables liés au médicament. On veut tester $H_0 : p = 1/1000$ contre $H_1 : p > 1/1000$ au risque de 1ère espèce $\alpha = 0.05$. Ici, on a : $np_0 = 5$ donc on n'utilise pas l'approximation normale mais comme $p < 0.1$, on utilise l'approximation poissonnienne $\mathcal{P}(5)$ de la loi $\mathcal{B}(5000, 0.001)$.
- On détermine le plus grand k_1 tel que $\mathbb{P}[\mathcal{P}(5) \leq k_1] \leq 0.05$, on lit dans la table $k_1 = 1$. On a $\widehat{p} = 2/5000 > 1/5000$ donc on rejette H_0.

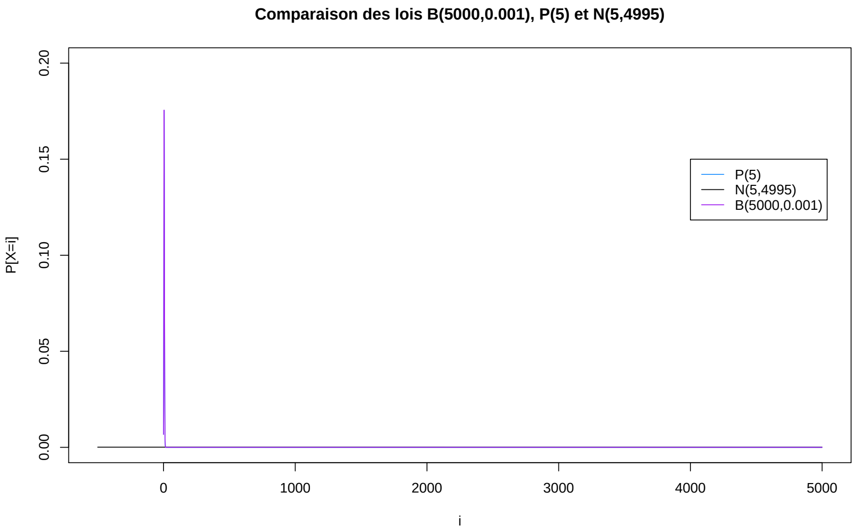
Exemple de comparaison d'une proportion à une valeur de référence par approximation poissonnienne (suite)

- NE PAS UTILISER l'approximation normale ! ! !
- Pour s'en convaincre, calculons la statistique de test du z-test :

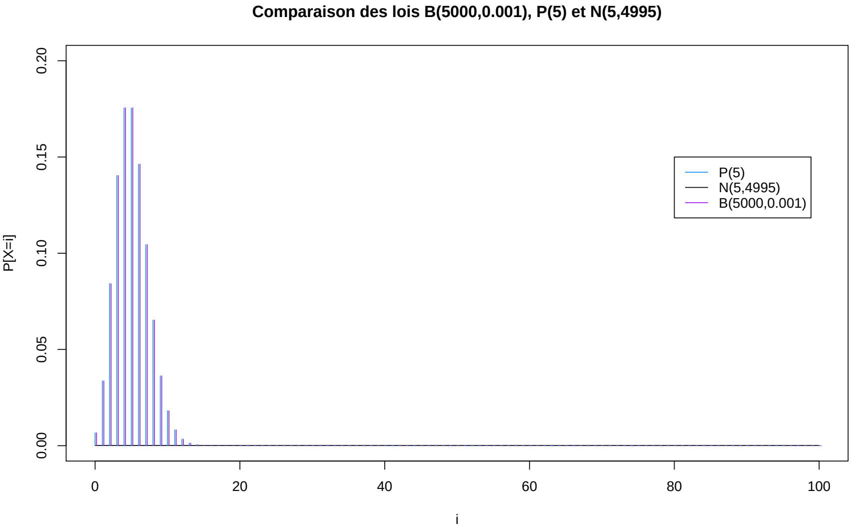
$$z = \frac{2/5000 - 0.001}{\sqrt{\frac{1}{1000} \times \frac{999}{1000} \times \frac{1}{5000}}} = -1.34.$$

- Le z-test rejette H_0 lorsque $z > z_\alpha$ pour $\alpha = 0.05$ i.e. $z_\alpha = 1.64$ donc ici on accepterait H_0 si on utilisait le z-test ! ! ! !
- Qui a raison ? le z-test ou l'approximation de Poisson ? ?

Attention aux conditions des approximations !



Attention aux conditions des approximations ! (suite)



Comparer une proportion à une valeur de référence : test exact

- **2ème cas** pour lequel l'approximation normale ne s'applique plus : l'échantillon est très petit et $np_0 \leq 5$ et/ou $n(1-p_0) \leq 5$. Dans ce cas, la solution est obtenue grâce à la loi exacte de $n\widehat{p}$ sous H_0 : $n\widehat{p} \sim \mathcal{B}(n, p_0)$.
- Pour tester $H_0 : p = p_0$ contre $H_1 : p \neq p_0$, on détermine le plus grand k_1 tel que $\mathbb{P}[\mathcal{B}(n, p_0) \leq k_1] \leq \alpha/2$ et le plus petit k_2 tel que $\mathbb{P}[\mathcal{B}(n, p_0) \geq k_2] \leq \alpha/2$. On rejette H_0 lorsque $\widehat{p} < k_1/n$ ou $\widehat{p} > k_2/n$.
- Pour tester $H_0 : p = p_0$ contre $H_1 : p > p_0$, on détermine le plus grand k_1 tel que $\mathbb{P}[\mathcal{B}(n, p_0) \leq k_1] \leq \alpha$. On rejette H_0 lorsque $\widehat{p} > k_1/n$.
- Pour tester $H_0 : p = p_0$ contre $H_1 : p < p_0$, on détermine le plus petit k_2 tel que $\mathbb{P}[\mathcal{B}(n, p_0) \geq k_2] \leq \alpha$. On rejette H_0 lorsque $\widehat{p} < k_2/n$.

Exemple de test exact pour comparer une proportion à une valeur de référence

- On réalise un essai préclinique sur l'action d'une nouvelle molécule contre une pathologie donnée.
- On observe 4 souris guéries sur $n = 10$ souris traitées.
- On veut tester $H_0 : p = 0.5$ contre $H_1 : p \neq 0.5$ au risque de 1ère espèce de 5%.
- On détermine le plus grand k_1 tel que $\mathbb{P}[\mathcal{B}(10, 0.5) \leq k_1] \leq 0.025$ et le plus petit k_2 tel que $\mathbb{P}[\mathcal{B}(10, 0.5) \geq k_2] \leq 0.025$.
- On lit dans la table de la f.r. de la loi binomiale :

$$\mathbb{P}[\mathcal{B}(10, 0.5) \leq 0] = 0.001$$

$$\mathbb{P}[\mathcal{B}(10, 0.5) \leq 1] = 0.011$$

donc $k_1 = 0$.

Exemple de test exact pour comparer une proportion à une valeur de référence (suite)

- On lit dans la table de la f.r. de la loi binomiale et on calcule :

$$\begin{aligned}\mathbb{P}[\mathcal{B}(10,0.5)\geq 10] &= 1-\mathbb{P}[\mathcal{B}(10,0.5)<10]\\ &= 1-\mathbb{P}[\mathcal{B}(10,0.5)\leq 9] = 0.001\end{aligned}$$

$$\begin{aligned}\mathbb{P}[\mathcal{B}(10,0.5)\geq 9] &= 1-\mathbb{P}[\mathcal{B}(10,0.5)<9]\\ &= 1-\mathbb{P}[\mathcal{B}(10,0.5)\leq 8] = 0.011\end{aligned}$$

$$\begin{aligned}\mathbb{P}[\mathcal{B}(10,0.5)\geq 8] &= 1-\mathbb{P}[\mathcal{B}(10,0.5)<8]\\ &= 1-\mathbb{P}[\mathcal{B}(10,0.5)\leq 7] = 0.055\end{aligned}$$

  donc $k_2 = 9$.
- La région de rejet est donc : $\{\frac{0}{10}, \frac{9}{10}, \frac{10}{10}\}$. Donc on ne rejette pas H_0.

Comparaison de deux proportions pour échantillons indépendants satisfaisant l'approximation normale

- **conditions d'applications** : $n_1 p_1 > 5$, $n_1(1-p_1) > 5$, $n_2 p_2 > 5$, $n_2(1-p_2) > 5$
- La statistique de test Z suit approximativement une loi $\mathcal{N}(0,1)$ sous H_0 :

$$Z = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}(1-\widehat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

où
$$\widehat{p} = \frac{n_1\widehat{p}_1 + n_2\widehat{p}_2}{n_1+n_2}.$$

- Pour tester $H_0 : p_1 = p_2$ contre $H_1 : p_1 \neq p_2$, on calcule la réalisation z. On rejette H_0 si $|z| > z_{\alpha/2}$, sinon on ne rejette pas H_0.
- Pour tester $H_0 : p_1 = p_2$ contre $H_1 : p_1 > p_2$, on calcule la réalisation z. On rejette H_0 si $z > z_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $H_0 : p_1 = p_2$ contre $H_1 : p_1 < p_2$, on calcule la réalisation z. On rejette H_0 si $z < -z_\alpha$, sinon on ne rejette pas H_0.

Exemple de comparaison de deux proportions

- Deux médicaments A et B en concurrence ont été administrés à deux groupes distincts de 50 personnes atteintes de la même maladie. Les patients sont suivis de façon absolument identique. Après 15 jours, on remarque que 44 personnes parmi les 50 traitées par A ont guéri alors qu'il y en a 40 guéries parmi les 50 personnes traitées par B. On veut tester au risque 5% si les deux médicaments sont significativement différents.
- On teste $H_0 : p_1 = p_2$ contre $H_1 : p_1 \neq p_2$.
- On calcule : $\overline{p}_1 = \frac{44}{50} = 0.88$, $\overline{p}_2 = \frac{40}{100} = 0.80$
- Les effectifs sont $n_1 = n_2 = 50 > 30$ donc on utilise un two-sample z-test avec
$$\overline{p} = \frac{50 \times 0.88 + 50 \times 0.80}{50 + 50} = 0.84$$
- La statistique de test vaut :
$$z = \frac{0.88 - 0.80}{\sqrt{2 \times \frac{0.84 \times (1 - 0.84)}{50}}} = 1.09$$
- Valeur critique : fractile bilatéral de la loi $\mathcal{N}(0,1)$ correspondant au risque $\alpha = 0.05$: $z_{\alpha/2} = 1.96$.
- $|z| < 1.96$ donc on ne rejette pas H_0 au risque 5%.
- Conclusion : les deux médicaments ne sont pas significativement différents en termes d'efficacité au risque 5%.

Comparaison de 2 proportions au moyen d'un test du χ^2

- On veut tester $H_0 : p_1 = p_2$ contre $H_1 : p_1 \neq p_2$ en notant p_1 la proportion d'observations de catégorie 1 dans l'échantillon 1 et p_2 la proportion d'observations de catégorie 1 dans l'échantilllon 2. Fixons le risque de 1ère espèce à α.
- Présentons les résultats obtenus lors de l'expérience sous la forme suivante.

	échantillon 1	échantillon 2	total
catégorie 1	a	b	$a+b$
catégorie 2	c	d	$c+d$
total	$a+c$	$b+d$	n

- Si $(a+c)(a+b)/n \geq 5$, $(a+c)(c+d)/n \geq 5$, $(b+d)(c+d)/n \geq 5$ et $(b+d)(a+b)/n \geq 5$, alors on peut utiliser la statistique K^2 qui suit sous H_0 une loi $\chi^2(1)$:

$$K^2 = \frac{n(ad - bc)^2}{(a+b)(a+c)(b+d)(c+d)}$$

  Il peut être souhaitable d'effectuer une correction dite de continuité ou correction de Yates :

$$K^2 = \frac{n(|ad - bc| - \frac{n}{2})^2}{(a+b)(a+c)(b+d)(c+d)}$$

- On calcule la réalisation k^2 à partir de l'échantillon et on rejette H_0 lorsque k^2 dépasse la valeur k_α définie par $\mathbb{P}[\chi^2(1) \leq k_\alpha] = 1 - \alpha$.

Test exact de Fisher pour comparer 2 proportions à partir d'échantillons indépendants

- Lorsque les effectifs des échantillons sont très faibles, il est souvent impossible de respecter les conditions d'application du test du χ^2. La méthode exacte de Fisher est alors utile. Présentons les résultats obtenus lors de l'expérience sous la forme suivante.

	échantillon 1	échantillon 2	total
catégorie 1	a	b	$a+b$
catégorie 2	c	d	$c+d$
total	$a+c$	$b+d$	n

- On veut tester $H_0 : p_1 = p_2$ contre $H_1 : p_1 > p_2$ ou contre $H_1 : p_1 < p_2$ en notant p_1 la proportion d'observations de catégorie 1 dans l'échantillon 1 et p_2 la proportion d'observations de catégorie 1 dans l'échantilllon 2.
- La méthode de Fisher est basée sur le calcul de la probabilité d'obtenir des échantillons aussi ou encore plus différents entre eux que ceux observés alors qu'il n'existe pas de différence en réalité. Cela nécessite de construire tous les tableaux de contingence présentant les mêmes totaux marginaux que ceux observés et affichant des différences encore plus marqués entre les échantillons (dans le même sens que celui observé).
- Notons p la probabilité obtenue par somme des probabilités de chacun des tableaux. Fixons le risque de 1ère espèce à α. Lorsque $p \leq \alpha$, on rejette H_0. Lorsque $p > \alpha$, on ne rejette pas H_0.

Exemple de test exact de Fischer

Un laboratoire veut développer un médicament destiné à soigner la grippe. Les chercheurs se demandent si le nouveau traitement conduit à une guérison plus rapide qu'avec le traitement classique. Un essai est conduit sur 19 patients qui reçoivent soit le nouveau traitement A soit un traitement classique B. Les résultats sont présentés dans le tableau suivant.

	traitement A	traitement B	total
guérison rapide (GR)	6	3	9
guérison "normale" (GN)	2	8	10
total	8	11	19

- On veut tester $H_0 : p_1 = p_2$ contre $H_1 : p_1 > p_2$ en notant p_1 la proportion de guérisons rapides avec le traitement A et p_2 la proportion de guérisons rapides avec le traitement B. On fixe $\alpha = 0.05$.
- Construisons les tableaux suivants :

cas n°1	A	B	total
GR	6	3	9
GN	2	8	10
total	8	11	19

cas n°2	A	B	total
GR	7	2	9
GN	1	9	10
total	8	11	19

cas n°3	A	B	total
GR	8	1	9
GN	0	10	10

Exemple de test exact de Fischer (suite)

- La probabilité d'obtenir le cas n°1 est donnée par

$$p_1 = \frac{(a+b)!(a+c)!(b+d)!(c+d)!}{n!a!b!c!d!}$$

  ce qui fait :

$$p_1 = \frac{9!10!8!11!}{19!6!3!2!8!} = 0.0500$$

  La probabilité d'obtenir le cas n°2 est donnée par :

$$p_2 = \frac{9!10!8!11!}{19!7!2!1!9!} = 0.0048$$

  La probabilité d'obtenir le cas n°3 est donnée par :

$$p_3 = \frac{9!10!8!11!}{19!8!1!0!10!} = 0.0001$$

- On en déduit $p = 0.0500 + 0.0048 + 0.0001 = 0.0549$
- $p > \alpha = 0.05$ donc on accepte H_0 à la limite du seuil critique.

Test de Mac Neymar pour comparer 2 proportions à partir d'échantillons appariés

- On veut tester $H_0 : p_1 = p_2$ en notant p_1 la proportion d'observations positives avant traitement et p_2 la proportion d'observations positives après traitement. Fixons le risque de 1ère espèce à α.
- Présentons les résultats obtenus lors de l'expérience sous la forme suivante.

avant	après	effectif
positif	positif	a
positif	négatif	b
négatif	positif	c
négatif	négatif	d

- La statistique de test Z suit une loi $\mathcal{N}(0,1)$ sous H_0 :

$$Z = \frac{b-c}{\sqrt{b+c}}$$

  - Pour tester $H_0 : p_1 = p_2$ contre $H_1 : p_1 \neq p_2$, on calcule la réalisation z. On rejette H_0 si $|z| > z_{\alpha/2}$, sinon on ne rejette pas H_0.

Test de Mac Neymar pour comparer 2 proportions à partir d'échantillons appariés (suite)

- (suite)
  - Pour tester $H_0 : p_1 = p_2$ contre $H_1 : p_1 > p_2$, on calcule la réalisation z. On rejette H_0 si $z > z_\alpha$, sinon on ne rejette pas H_0.
  - Pour tester $H_0 : p_1 = p_2$ contre $H_1 : p_1 < p_2$, on calcule la réalisation z. On rejette H_0 si $z < -z_\alpha$, sinon on ne rejette pas H_0.
- Pour tester $H_0 : p_1 = p_2$ contre $H_1 : p_1 \neq p_2$, on peut également souhaiter utiliser la statistique K^2 qui suit sous H_0 une loi $\chi^2(1)$:

$$K^2 = \frac{(b-c)^2}{(b+c)}.$$

Il peut être souhaitable d'effectuer une correction dite de continuité ou correction de Yates :

$$K^2 = \frac{(|b-c|-1)^2}{(b+c)}.$$

On calcule la réalisation k^2 à partir de l'échantillon et on rejette H_0 lorsque k^2 dépasse la valeur k_α définie par $\mathbb{P}[\chi^2(1) \leq k_\alpha] = 1 - \alpha$.

Exemple de test de Mac Neymar

Le diagnostic de sténose peut être posé grâce à l'angiographie effectuée sur des patients en position assise ou couchée. Pour tester la position la plus efficace, les résultats d'examen effectués sur 112 patients se plaignant de fourmillements dans le bras ont été résumés dans le tableau ci-dessous :

position assise	position couchée	nombre de patients
positif	positif	59
positif	négatif	8
négatif	positif	20
négatif	négatif	25

On se demande si une position est plus favorable que l'autre à la détection de la sténose à l'angiographie.

- On teste $H_0 : p_1 = p_2$ contre $H_1 : p_1 \neq p_2$ en notant p_1 la proportion de sténose détectées en posiion assise et p_1 la proportion de sténose détectées en posiion couchée au risque $\alpha = 0.05$. On effectue un test du χ^2 avec correction de Yates.
- La réalisation de la statistique de test vaut :

$$k^2 = \frac{(|8-20|-1)^2}{8+20} = 4.321$$

La valeur critique est le fractile unilatéral à 95% de la loi $\chi^2(1)$: $k_\alpha = 3.841$. On voit que $k^2 > 3.841$. On rejette donc H_0 au risque 5%. Les deux positions n'ont donc pas la même efficacité pour détecter une sténose.

Comparer k proportions à k valeurs de référence : test du χ^2 à un critère de classification

- Soit X une variable catégorielle ou quantitative discrète ou continue mais réparties en classes. Soit K le nombre de modalités (=classes) prises par X. Soit p_k la probabilité que X soit dans la $k^{\text{ème}}$ classe pour $k = 1, ...K$ (proportion théorique). Soit $p_{0,k}$ une proportion proposée par l'expérimentateur pour $k = 1, ...K$.
- On veut tester $\mathbf{H_0}$: $p_k = p_{0,k}$ pour $k = 1, ...K$ contre $\mathbf{H_1}$: $p_k \neq p_{0,k}$ pour au moins un k_0 dans $\{1, ...K\}$.
- Soit n_k pour $k = 1, ...K$ l'effectif de la $k^{\text{ème}}$ classe, la taille de l'échantillon est donc $n = \sum_{k=1}^{K} n_k$.
- **Conditions d'application** : pour $k = 1, \ldots, K$, $np_{0,k} \geq 5$.
- La statistique de test K^2 suit une loi $\chi^2(K-1)$ sous H_0 :

$$K^2 = \sum_{k=1}^{K} \frac{(n_k - np_{0,k})^2}{np_{0,k}}$$

- On calcule la réalisation k^2 à partir des données. On rejette H_0 au risque α lorsque $k^2 > k_\alpha$ où k_α est la valeur définie par $\mathbb{P}[\chi^2(K-1) > k_\alpha] = \alpha$.
- Le test du χ^2 à un critère de classification s'appelle également test du χ^2 d'ajustement ou test du χ^2 d'adéquation ou test du χ^2 de conformité.

Exemple de test du χ^2 à un critère de classification

- Considérons un dosage biologique. Les médecins analysent les résultats de ce dosage en disant que le dosage est faible, normal ou élevé. On veut tester au risque $\alpha = 0.05$ le fait que 90% des gens ont un dosage normal, 5% ont un dosage faible et 5% ont un dosage élevé. Pour cela, on sélectionne au hasard 100 sujets et on constate que, sur les 100 dosages, 76 sont normaux, 10 faibles et 14 forts.
- On veut tester H_0 : $\mathbb{P}$[avoir un dosage faible] $= 0.05$,
  $\mathbb{P}$[avoir un dosage normal] $= 0.90$ et
  $\mathbb{P}$[avoir un dosage élevé] $= 0.05$ contre H_1 : "au moins une de ces 3 égalités n'est pas satisfaite".
- On calcule
  $k^2 = \frac{(10-100\times0.05)^2}{100\times0.05} + \frac{(76-100\times0.9)^2}{100\times0.9} + \frac{(14-100\times0.05)^2}{100\times0.05} = 23.378.$
- Valeur critique : on détermine la valeur k_α définie par :
  $\mathbb{P}[\chi^2(3-1) > k_\alpha] = 0.05$. On lit dans la table du χ^2 que
  $k_\alpha = 5.991$.
- On voit que $k^2 > k_\alpha$ donc on rejette H_0.

$4^{\text{ème}}$ partie

Comparaisons de distributions

Comparer une distribution à une loi de référence : test d'ajustement de Kolmogorov-Smirnov

- Soit X une variable quantitative de loi F.
- On veut tester $H_0 : F = F_0$ contre $H_1 : F \neq F_0$ i.e. pour au moins un x dans $\mathbb{R}$, on a $F(x) \neq F_0(x)$.
- Soit $(X_1, \ldots, X_n)$ un échantillon de variable parente X. Soit F_n la fonction de répartition empirique de X, égale au point x à

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(X_i \leq x) = \frac{\text{nb d'observations inférieures ou égales à } x}{n}.$$

- On rejette H_0 au risque α lorsque la statistique de test

$$D = \sup_{x \in \mathbb{R}} |F_n(x) - F_0(x)|$$

dépasse une valeur d_α définie par $\mathbb{P}[Y_n > d_\alpha] = \alpha$ où la loi de la variable Y_n est tabulée selon α et n.
- Intuitivement, D mesure en chaque point la distance entre la loi empirique et la loi théorique proposée.
- NB : dans le cas d'un ajustement à une loi normale, on préfèrera le test de Shapiro-Wilk.

Exemple n°1 de test d'ajustement de Kolmogorov-Smirnov

- Un appareil de radiographie admet 5 réglages possibles en ce qui concerne le tirage, allant du plus clair au plus foncé. On veut tester l'hypothèse H_0 selon laquelle la lisibilité de la radiographie est la même pour les 5 tirages possibles, au risque $\alpha = 0.05$. Autrement dit, sous H_0, les préférences des médecins en ce qui concerne la lisibilité des radios est uniformément répartie sur les 5 tirages. Si l'on note F la loi théorique des préférences des médecins, on veut tester H_0 : $F = \mathcal{U}\{1, 2, 3, 4, 5\}$ contre H_1 : $F \neq \mathcal{U}\{1, 2, 3, 4, 5\}$.
- On demande à 10 médecins d'observer les 5 tirages différents d'une même radiographie. On obtient

tirage sélectionné	1	2	3	4	5
nb de médecins	0	1	0	5	4

- On calcule

tirage sélectionné	1	2	3	4	5
nb de médecins	0	1	0	5	4
F_n	$0/10$	$1/10$	$1/10$	$6/10$	$10/10$
F_0	$1/5$	$2/5$	$3/5$	$4/5$	$5/5$
$\lvert F_n - F_0 \rvert$	$2/10$	$3/10$	$5/10$	$2/10$	0

- On obtient la réalisation $d = 5/10 = 0.500$
- On lit dans la table que $\mathbb{P}[Y_n > d_\alpha] = 0.05$ pour $d_\alpha = 0.409$ avec $n = 10$.
- On voit que $d > d_\alpha$ donc on rejette H_0.

Exemple n°2 de test d'ajustement de Kolmogorov-Smirnov

- En médecine nucléaire, des sources de radionucléides sont injectées à des patients. Pour les pathologies thyroïdiennes, on utilise classiquement l'isotope ${}^{135}_{53}I$. L'élimination effective de la radioactivité incorporée s'effectue par combinaison de la décroissance radioactive du radionucléide et de l'élimination biologique propre à l'organe cible. On sait que l'élimination effective de l'iode ${}^{135}_{53}I$ obéit à une loi $\mathcal{E}(0.0912)$ lorsque le temps d'élimination est exprimé en jours.
- Un praticien développe l'utilisation de l'isotope ${}^{135}_{53}I$ pour un nouveau traitement d'une autre pathologie. Il veut tester au risque $\alpha = 0.05$ si la durée d'élimination suit les mêmes règles que lors de l'utilisation classique. Autrement dit, si l'on note F la loi théorique de la durée d'élimination après ce nouveau traitement, on veut tester $H_0 : F = \mathcal{E}(0.0912)$ contre $F \neq \mathcal{E}(0.0912)$.
- On suit 10 patients après ce nouveau traitement. On obtient les durées d'élimination suivantes : 3 4 4 5 6 6 11 18 22 36.
- On calcule

durée	3	4	5	6	11	18	22	36
F_n	$1/10$	$3/10$	$4/10$	$6/10$	$7/10$	$8/10$	$9/10$	$10/10$
F_0	0.213	0.274	0.330	0.381	0.585	0.763	0.828	0.944
$\lvert F_n - F_0\rvert$	0.113	0.026	0.070	0.219	0.115	0.037	0.072	0.056

- On obtient la réalisation $d = 0.219$
- On lit dans la table que $\mathbb{P}[Y_n > d_\alpha] = 0.05$ pour $d_\alpha = 0.409$ avec $n = 10$.
- On voit que $d < d_\alpha$ donc on ne rejette pas H_0.

Comparer deux distributions, échantillons indépendants : test two-sample de Kolmogorov-Smirnov

- Soit une variable X de f.r. F et soit Y une variable de f.r. G.
- On veut tester au risque α $\mathbf{H_0}$: $F = G$ contre $\mathbf{H_1}$: $F \neq G$ i.e. pour au moins un x dans $\mathbb{R}$, on a $F(x) \neq G(x)$.
- Soit $(X_1, \ldots, X_{n_1})$ échantillon de variable parente X et soit $(Y_1, \ldots, Y_{n_2})$ échantillon de variable parente Y. Soit F_{n_1} la f.r. empirique de X, définie au point x par

$$F_{n_1}(x) = \frac{1}{n_1} \sum_{i=1}^{n_1} I(X_i \leq x) = \frac{\text{nb d'observations de } X \text{ inférieures ou égales à } x}{n}$$

et soit G_{n_2} la f.r. empirique de Y, définie au point x par

$$G_{n_2}(x) = \frac{1}{n_2} \sum_{i=1}^{n_2} I(Y_i \leq x) = \frac{\text{nb d'observations de } Y \text{ inférieures ou égales à } x}{n}.$$

- On rejette H_0 au risque α lorsque la réalisation de la statistique de test

$$D = \sqrt{\frac{n_1 n_2}{n_1 + n_2}} \sup_{x \in \mathbb{R}} |F_{n_1}(x) - G_{n_2}(x)|$$

dépasse une valeur d_α. Pour $n_1 > 12$ et $n_2 > 12$, on considère que pour $\alpha = 0.05$, $d_\alpha = 1.36$.
- Intuitivement, D mesure en chaque point la distance entre la loi empirique de X et la loi empirique de Y.

Exemple de two-sample test de Kolmogorov-Smirnov

- On s'intéresse à l'effet d'un médicament sur les infections des souris par une larve.
- 16 souris sont infectées par le même nombre de larves, puis réparties au hasard entre 2 groupes égaux. Le premier groupe reçoit le traitement, pas le second. Au bout d'une semaine, toutes les souris sont sacrifiées et les nombres suivants de vers adultes ont été retrouvés dans les intestins :
  souris traitées : 44 47 49 53 57 60 62 67
  souris non traitées : 51 55 62 63 68 71 75 79
- On veut conclure sur l'éventuelle efficacité du traitement.
- On note X la variable représentant le nombre de vers chez une souris traitée et Y la variable représentant le nombre de vers chez une souris non traitée. Soit F la loi de X et soit G la loi de Y. On veut tester $H_0 : F = G$ contre $H_1 : F \neq G$ au risque $\alpha = 0.05$.

Exemple de two-sample test de Kolmogorov-Smirnov

nb de vers	44	47	49	51	53	55	57	60
G_n (SNT)	0	0	0	1/8	1/8	2/8	2/8	2/8
F_n (ST)	1/8	2/8	3/8	3/8	4/8	4/8	5/8	6/8
$G_n - F_n$	-1/8	-2/8	-3/8	-2/8	-3/8	-2/8	-3/8	-4/8

nb de vers	62	63	67	68	71	75	79
G_n (SNT)	3/8	4/8	4/8	5/8	6/8	7/8	8/8
F_n (ST)	7/8	7/8	8/8	8/8	8/8	8/8	8/8
$G_n - F_n$	-4/8	-3/8	-4/8	-3/8	-2/8	-1/8	0

- On obtient la réalisation $d = \sqrt{\frac{8\times 8}{8+8}} \times \frac{4}{8} = 1$. On ne rejette pas H_0.

Comparer deux distributions entre elles, échantillons indépendants : test U de Wilcoxon-Mann-Whitney

- Soit une variable X de f.r. F et soit Y une variable de f.r. G.
- On veut tester $H_0 : F = G$ contre $H_1 : F \neq G$ au risque α.
- Soit $(X_1, \ldots, X_{n_1})$ échantillon de variable parente X et soit $(Y_1, \ldots, Y_{n_2})$ échantillon de variable parente Y.
- On calcule $U = \min(U_1, U_2)$ avec

$$\begin{cases} U_1 = \text{somme des scores des éléments du plus petit échantillon} \\ U_2 = n_1 n_2 - U_1 \end{cases}$$

  où le score d'un élément e est défini comme étant le nombre d'éléments de l'autre échantillon qui sont inférieurs à e plus la moitié du nombre d'éléments de l'autre échantillon qui sont égaux à e.
- Lorsque n_1 ou $n_2 \geq 20$, la statistique de test Z suit une loi $\mathcal{N}(0, 1)$:

$$Z = \frac{U - n_1 n_2/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}}.$$

  On rejette H_0 si la réalisation z satisfait $|z| > z_{\alpha/2}$ où $z_{\alpha/2}$ est le fractile bilatéral de la loi $\mathcal{N}(0, 1)$.
- Si $n_1 < 20$ et $n_2 < 20$, la loi de U est tabulée en fonction de n et de α. On rejette H_0 si la réalisation u satisfait $u < u_{\alpha/2}$ où $u_{\alpha/2}$ est le fractile lu dans la table pour tests de Mann-Whitney bilatéraux (attention au sens inhabituel pour la région de rejet ! ! !).

Comparer deux distributions entre elles, échantillons indépendants : test U de Wilcoxon-Mann-Whitney (suite)

- NB : Intuitivement, si on mélange les deux séries de valeurs $(X_1, \ldots, X_{n_1})$ et $(Y_1, \ldots, Y_{n_2})$ et si on ordonne les $(n_1 + n_2)$ valeurs des 2 échantillons par ordre croissant, on doit obtenir un mélange homogène sous H_0.
- NB2 : on ne prend pas en compte les valeurs observées, seulement leur rang !
- NB3 : s'il y a beaucoup d'ex-aequo ou s'il y a peu d'ex-aequo mais que z est très proche de la zone de rejet, il est recommandé d'utiliser une formule plus exacte. Supposons qu'il y ait k groupes de valeurs (éventuellement réduits à un seul point) et que chaque groupe est constitué d'un nombre t_i ex-aequos. On calcule alors :
$$Z = \frac{S - n_1 n_2/2}{\sqrt{\sigma^2}}$$
avec
$$\sigma^2 = \frac{n_1 n_2}{n(n-1)} \left(\frac{n^3 - n}{12} - \frac{\sum_{i=1}^{k} (t_i^3 - t_i)}{12} \right).$$

Comparer deux distributions entre elles, échantillons indépendants : test de la somme des rangs de Wilcoxon

- Soit une variable X de f.r. F et soit Y une variable de f.r. G.
- On veut tester $H_0 : F = G$ contre $H_1 : F \neq G$ au risque α.
- Soit $(X_1, \ldots, X_{n_1})$ échantillon de variable parente X et soit $(Y_1, \ldots, Y_{n_2})$ échantillon de variable parente Y.
- On mélange les deux séries de valeurs et on ordonne les $(n_1 + n_2)$ valeurs des 2 échantillons par ordre croissant. Supposons que n_1 soit la taille du plus petit échantillon. On détermine $U = \min(U_1, U_2)$ avec

  $$\begin{cases} S = \text{somme des rangs obtenus par les éléments du plus petit échantillon} \\ U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - S \\ U_2 = n_1 n_2 - U_1 \end{cases}$$

  En cas d'ex-aequos, on attribue à chacun des ex-aequos le rang moyen obtenu : par ex, s'il y a 3 ex-aequos pour la 5ème place, on attribue à tous les 3 le rang $(5+6+7)/3 = 6$, s'il y en a 4, on attribue à tous les 4 le rang $(5+6+7+8)/4 = 6.5$.
- Lorsque $n_1 \geq 20$ ou $n_2 \geq 20$, la statistique de test Z suit approximativement une loi $\mathcal{N}(0,1)$ sous H_0 :

  $$Z = \frac{S - n_1(n_1 + n_2 + 1)/2}{\sqrt{n_1 n_2 (n_1 + n_2 + 1)/12}}.$$

  On rejette H_0 lorsque la réalisation z satisfait $|z| > z_{\alpha/2}$.

Comparer deux distributions entre elles, échantillons indépendants : test de la somme des rangs de Wilcoxon (suite)

- Lorsque $n_1 < 20$ et $n_2 < 20$, on rejette H_0 lorsque la réalisation s satisfait $s < s_{\alpha/2}$ pour une valeur $s_{\alpha/2}$ tabulée en fonction de n et de α.
- NB : s'il y a beaucoup d'ex-aequo ou s'il y a peu d'ex-aequo mais que z est très proche de la zone de rejet, il est recommandé d'utiliser une formule plus exacte pour la variance. Supposons qu'il y ait k groupes de valeurs (éventuellement réduits à un seul point) et que chaque groupe est constitué de t_i ex-aequos. On calcule alors :
$$Z = \frac{S - n_1(n_1 + n_2 + 1)/2}{\sqrt{\sigma^2}}$$
avec
$$\sigma^2 = \frac{n_1 n_2}{2}\left(n_1 + n_2 - \frac{\sum_{i=1}^{k} t_i(t_i^2 - 1)}{(n_1 + n_2)(n_1 + n_2 - 1)}\right)$$
- NB2 : le test de Wilcoxon est équivalent au test U de Mann-Whitney sur le plan de l'idée et sur le plan mathématique mais il est plus simple à mettre en oeuvre.
- NB3 : ne pas confondre avec le test de Wilcoxon de la somme des rangs signés qui est valide pour les données appariées.

Exemple de comparaison de distributions avec échantillons indépendants

Afin d'évaluer la superficie du domaine vital de l'ours noir, mâle ou femelle, 15 colliers émetteurs ont été posés sur 15 ours, permettant ainsi de suivre leurs déplacements. Les superficies suivantes ont été obtenues (en km^2) :
Mâles : 94 504 173 560 274 168
Femelles : 37 72 60 49 18 50 102 49 20
On se demande si le domaine vital des mâles est aussi étendu que celui des femelles ou non.

- Soit X la variable représentant la superficie du domaine vital des mâles de f.r. F et soit Y la variable représentant la superficie du domaine vital des femelles de f.r. G.
- On veut tester $H_0 : F = G$ contre $H_1 : F \neq G$ au risque $\alpha = 0.05$.
- Effectuons les calculs préliminaires au calcul de U selon les deux approches : voir transparent suivant
- Approche de Mann-Whitney : $U_1 = 1$, $U_2 = 6 \times 9 - 1 = 53$ donc $U = 1$.
- Approche de Wilcoxon : $U_1 = 6 \times 9 + \frac{6(6+1)}{2} - 74 = 1$, $U_2 = 6 \times 9 - 1 = 53$ donc $U = 1$.
- La valeur critique pour un test bilatéral au risque $\alpha = 0.05$ lue dans la table vaut 10. On rejette donc H_0 au risque $\alpha = 0.05$.

Exemple de comparaison de distributions avec échantillons indépendants (suite)

données	échantillons	scores	rangs
18	F	0	1
20	F	0	2
37	F	0	3
49	F	0	4.5
49	F	0	4.5
50	F	0	6
60	F	0	7
72	F	0	8
94	M	8	9
102	F	1	10
168	M	9	11
173	M	9	12
274	M	9	13
504	M	9	14
560	M	9	15

Exemple de test de Wilcoxon pour échantillons indépendants

- On veut comparer la perte de poids chez les hommes et chez les femmes après un régime de 3 mois. On dispose des observations suivantes (en kg) :
  hommes : -10 -5 -4 -4 -4 -4 -3 -3 -3 -2 -2 -2 -1 0 0 2 2 4 4 4
  femmes : -7 -6 -6 -5 -4 -2 -1 0 1 1

valeur	-10	-7	-6	-6	-5	-5	-4	-4	-4	-4	-4	-3
rang	1	2	3.5	3.5	5.5	5.5	9	9	9	9	9	13
rang F		2	3.5	3.5	5.5		9					

valeur	-3	-3	-2	-2	-2	-2	-1	-1	0	0
rang	13	13	16.5	16.5	16.5	16.5	19.5	19.5	22	22
rang F			16.5				19.5		22	

valeur	0	1	1	2	2	4	4	4
rang	22	24.5	24.5	26.5	26.5	29	29	29
rang F		24.5	24.5					

Exemple de test de Wilcoxon pour échantillons indépendants

- On a mélangé les deux échantillons, on a classé les valeurs par ordre croissant et déterminé les rangs.
- On calcule la somme des rangs du plus petit échantillon (celui des femmes), on trouve $s = 130.5$.
- On détermine alors

$$z = \frac{130.5 - 10(10 + 20 + 1)/2}{\sqrt{10 \times 20 \times (10 + 20 + 1)/12}} = -1.08.$$

Comme $|z| < z_{\alpha/2} = 1.96$ pour $\alpha = 0.05$, on ne rejette pas H_0.

Comparer deux distributions entre elles, échantillons appariés : test de la somme des rangs signés de Wilcoxon

- Soit une variable X de f.r. F et soit Y une variable de f.r. G.
- On veut tester $\mathbf{H_0}$: $F = G$ contre $\mathbf{H_1}$: $F \neq G$ au risque de 1ère espèce α.
- Soit $(X_1, \ldots, X_n)$ échantillon de variable parente X et soit $(Y_1, \ldots, Y_n)$ échantillon de variable parente Y, les 2 échantillons ont la même taille et sont appariés.
- On crée une série de variables D_i pour $i = 1, \ldots, n$ définies par $D_i = X_i - Y_i$.
- On ordonne les n valeurs des $|D_i|$ par ordre croissant. On fait la somme S des rangs qui portent le signe $-$. En cas d'ex-aequos, on attribue à chacun des ex-aequos le rang moyen obtenu.
- Pour $n < 60$, on rejette H_0 au risque α lorsque $s < s_{\alpha/2}$ où $s_{\alpha/2}$ est tabulé en fonction de n et de α.
- Pour $n \geq 60$, la statistique de test Z suit approximativement une loi $\mathcal{N}(0, 1)$ sous H_0 :

$$Z = \frac{S - n(n+1)/4}{\sqrt{n(n+1)(2n+1)/24}}.$$

On rejette H_0 au risque α lorsque $|z| \geq z_{\alpha/2}$.

Exemple de test de Wilcoxon pour échantillons appariés

- Des échantillons de 10 ml de crème sont prélévés dans 10 laiteries puis divisés en deux sous-échantillons de 10 fois 5 ml chacun. L'un est envoyé dans le laboratoire A et l'autre dans le laboratoire B pour compter les bactéries, les 2 laboratoires employant des méthodes différentes.
- Les résultats (en milliers de bactéries par ml) sont :
  labo A : 11.7 12.1 13.3 15.1 15.9 15.3 11.9 16.2 15.1 13.8
  labo B : 10.9 11.9 13.4 15.4 14.8 14.8 12.3 15.0 14.2 13.2
- On se demande si la différence de méthode influe sur le résultat. On effectue un test de Wilcoxon pour échantillons appariés au risque $\alpha = 0.05$.

Exemple de test de Wilcoxon pour échantillons appariés (suite)

- On construit l'échantillon des différences :

Z	0.8	0.2	-0.1	-0.3	1.1	0.5
rang($\|Z\|$)	7	2	1	3	9	5
rang des < 0			1	3		

Z	-0.4	1.2	0.9	0.6
rang($\|Z\|$)	4	10	8	6
rang des < 0	4			

- On obtient $s = 1 + 3 + 4 = 8$ puis on lit dans la table que $s_{\alpha/2} = 8$. On ne rejette pas H_0 (à la limite du seuil critique).

$5^{\text{ème}}$ partie

Tests d'association, de liaison, d'indépendance

Test d'indépendance entre 2 variables qualitatives

- Contexte : sur une même personne, on mesure 2 quantités X et Y. La variable X est qualitative ou discrétisée en K_1 modalités tandis que la variable Y est qualitative ou discrétisée en K_2 modalités. On cherche à montrer que ces variables sont dépendantes (sont liées).
- Exemples concrets :
  - Etat rénal (absence/présence d'insuffisance rénale) et état hépatique (absence/présence d'insuffisance hépatique) sont liés.
  - Perte de connaissance après traumatisme (oui/non) et survie à un mois (oui/non) sont liées.
  - Couleur des yeux (bleu/marrons/noirs) et couleur des cheveux (bruns/chatains/roux/blonds) sont liées.
- On effectue un test du χ^2 d'indépendance, appelé aussi χ^2 d'association ou χ^2 à deux critères de classification.

Test du χ^2 d'indépendance entre 2 variables qualitatives

- On veut tester H_0 : "les variables X et Y sont indépendantes" contre H_1 : "les variables X et Y sont liées".
- Les données sont rangées sous forme de table de contingence :

	col 1	col 2	col 3	...	col K_2	total
ligne 1	n_{11}	n_{12}	n_{13}	...	n_{1K_2}	$n_{1.}$
ligne 2	n_{21}	n_{22}	n_{23}	...	n_{2K_2}	$n_{2.}$
...	...	...	...	...	...	...
ligne K_1	n_{K_11}	n_{K_12}	n_{K_13}	...	$n_{K_1K_2}$	$n_{K_1.}$
total	$n_{.1}$	$n_{.2}$	$n_{.3}$	...	$n_{.K_2}$	$n_{..}$

avec n_{ij} représentant l'effectif des individus présentant la modalité i de la variable X ainsi que la modalité Y.

Test du χ^2 d'indépendance entre 2 variables qualitatives

- **Conditions d'application** : $n_{i.}n_{.j}/n_{..} \geq 5$ pour $i = 1, \ldots, K_1$ et $j = 1, \ldots, K_2$.
- On rejette H_0 avec un risque de 1ère espèce de α lorsque

$$K^2 = \sum_{i=1}^{K_1} \sum_{j=1}^{K_2} \frac{(n_{ij} - n_{i.}n_{.j}/n_{..})^2}{n_{i.}n_{.j}/n_{..}}$$

dépasse une certaine valeur k_0 déterminée par

$$\mathbb{P}[\chi^2((K_1 - 1)(K_2 - 1)) > k_0] = \alpha.$$

Exemple de test d'indépendance du χ^2

- On veut savoir si le temps écoulé depuis la vaccination contre la petite vérole a ou non une influence sur le degré de gravité de la maladie lorsqu'elle apparaît. Les patients sont divisés en trois catégories selon la gravité de leur maladie : légère (L), moyenne (M) ou grave (G) et en trois autres quant à la durée écoulé depuis la vaccination : moins de 10 ans (A), entre 10 et 25 ans (B) et plus de 25 ans (C).
- Les résultats d'une étude portant sur $n = 1574$ malades sont :

	A	B	C	total
G	1	42	230	273
M	6	114	347	467
L	23	301	510	834
total	30	457	1087	1574

Exemple de test d'indépendance du χ^2 (suite)

- On teste H_0 : "le temps écoulé depuis la vaccination contre la petite vérole et le degré de gravité de la maladie lorsqu'elle apparaît sont indépendants" contre H_1 : "le temps écoulé depuis la vaccination contre la petite vérole et le degré de gravité de la maladie lorsqu'elle apparaît sont liés". On fixe le risque de 1ère espèce $\alpha = 0.05$.
- Le tableau des effectifs attendus sous H_0 (*expected*) est :

	A	B	C	total
G	5.2	79.3	188.5	273
M	8.9	135.6	322.5	467
L	15.9	242.1	576.0	834
total	30	457	1087	1574

- On calcule $K^2 = 61.4$ et on lit dans la table $\mathbb{P}[\chi^2((3-1)(3-1)) > k_0] = 0.05$ pour $k_0 = 9.5$ donc on rejette H_0 au risque $\alpha = 0.05$.

Test d'indépendance entre 1 variable qualitative et 1 variable quantitative

- Soit X une variable qualitative ou discrétisée en K modalités et soit Y une variable quantitative de loi notée $\mathcal{L}(Y)$, par ex, X=stade de dépistage du cancer du poumon (parmi $K = 3$ stades : "précoce", "moyen", "avancé") et Y=durée de vie en j.
- Lorsque X et Y sont indépendantes, les lois $\mathcal{L}(Y|X = k)$ pour $k = 1, \dots, K$ sont toutes identiques entre elles et identiques à $\mathcal{L}(Y)$.
- On crée K groupes d'individus comparables à tous points de vue, sauf en ce qui concerne la valeur de X : X vaut k pour tous les individus du groupe k. On note Y_k la variable quantitative mesurée dans le groupe k.
- On compare ensuite les distributions, ou les espérances (ou autre) des variables Y_k pour $k = 1, \dots, K$.
- Ainsi, l'hypothèse H_0 : X et Y sont indépendantes peut se reformuler en $H_0 : \mathcal{L}(Y_k) = \mathcal{L}(Y)$ pour $k = 1, \dots, K$ ou en H_0 : $\mathbb{E}[Y_k] = \mathbb{E}[Y]$ pour $k = 1, \dots, K$.