STATISTIQUE COMPUTATIONNELLE – 7

TESTS D'HYPOTHÈSES SUR LES MOYENNES

7.1

Généralités sur les tests d'hypothèses

On considère un échantillon de taille n représenté par les v.a. $X_1,\ldots,X_n$, qui sont considérées i.i.d., issu d'une population d'effectif beaucoup plus grand.et soient $x_1,\ldots,x_n$ les observations d'un échantillon. On s'intéresse à une grandeur statistique particulière de cette population $\theta \in \Theta$. On suppose que $\Theta = \Theta_0 \cup \Theta_1$ avec $\Theta_1 = \Theta - \Theta_0$. Selon les valeurs des mesures, nous pouvons formuler les hypothèses suivantes :

– $H_0 : \theta \in \Theta_0$
  qui est *l'hypothèse nulle*. L'hypothèse contraire
– $H_1 : \theta \in \Theta_1$
  est appelée *hypothèse alternative*.

Afin d'examiner la validité de ces hypothèses on élabore des tests d'hypothèses. L'application de ce test peut conduire à des erreurs répertoriée en deux catégories :

– Erreur de 1e espèce.- Rejet de l'hypothèse H_0 bien qu'elle soit vraie. (Rejet à tort de l'hypothèse H_0 ou, ce qui revient au même, acceptation à tort de l'hypothèse H_1).
– Erreur de 2e espèce.- Acceptation de l'hypothèse H_0 bien qu'elle soit fausse. (Rejet à tort de l'hypothèse H_1 ou, ce qui revient au même, acceptation à tort de l'hypothèse H_0).

Réalité / Décision	H_0 **vraie**	H_1 **vraie**
Accepter H_0	Acceptation à raison de l'hypothèse nulle Prob. $1-\alpha$	Acceptation à tort de l'hypothèse nulle Prob. β
Refuser H_0	Refus à tort de l'hypothèse nulle Prob. α	Refus à raison de l'hypothèse nulle Prob. $1-\beta$

TABLE 7.1 – Décisions d'un test d'hypothèses

On suppose que l'erreur de 1e espèce se réalise avec une probabilité α que l'on appelle *risque de 1e espèce*. Par conséquent la probabilité d'accepter à raison H_0 est égale à $1-\alpha$. De même on suppose que l'erreur de 2e espèce se réalise avec une probabilité β qui est le risque *de 2e espèce*. Par conséquent la probabilité de refuser à raison H_0 est égale à $1-\beta$. Cette situation est répertoriée au tableau 7.1

En général les deux erreurs ne conduisent pas à des conséquences de même nature. De ce fait nous ne pouvons pas traiter de façon symétrique les deux hypothèses H_0 et H_1. Lors de l'élaboration d'un test on prendra comme hypothèse nulle celle dont le rejet à tort a les conséquences les plus pénalisantes.

7.2 Test simple de la moyenne avec une valeur de référence

Supposons que d'une population, de taille potentiellement infinie, on extrait un échantillon de taille n. Les mesures effectuées sur cet échantillon sont représentés par les variables aléatoires $X_1,\ldots,X_n$. Pour ces variables aléatoires nous faisons l'hypothèse qu'elles sont indépendantes et identiquement distribuées avec moyenne arithmétique $\overline{X} = \frac{1}{n}\sum_{k=1}^{n} X_k$.
Le paramètre que nous voulons estimer est la moyenne $\theta = \mu$ de la population totale et nous utiliserons $\overline{X}$ comme valeur de référence pour le test l'estimation. On prend comme hypothèse nulle

$H_0 : \theta = \theta_0$

et comme hypothèse alternative

$H_1 : \theta \neq \theta_0$

Parfois on cherche à détecter si $\theta = \theta_1$, avec $\theta_1 \neq \theta_0$, qui est l'hypothèse de détection :

$H_D : \theta = \theta_1$

Notons, enfin, que selon les valeurs que θ puisse prendre, nous avons un test simple ou un test double.

On distingue deux cas :
– les variances de deux populations sont connues ;
– les variances de deux populations sont inconnues
La procédure du test est la suivante.

1°.- $H_0 : \theta = \theta_0$; $H_1 : \theta < \theta_0$ (ou $\theta > \theta_0$) ; $H_D : \theta = \theta_1$

2°.- On fixe les probabilités α et β.

3°.- Calcul de la taille de l'échantillon :

3.1.- Calcul de la valeur U_α. Nous avons

$$P\left(\frac{\theta_0 - \overline{X}}{\sigma/\sqrt{n}} > U_\alpha\right) = \alpha$$

avec

$$U_\alpha = \begin{cases} Z_\alpha \sim \mathcal{N}(0,1), & \text{loi normale si la variance } \sigma^2 \text{ de la population est connue} \\ T_{n-1,\alpha} \sim t_{n-1,\alpha} & \text{loi de Student si la variance } \sigma^2 \text{ de la population est inconnue} \end{cases}$$

Par lecture des tables de la distribution normale (resp. de Student) on établit la valeur de Z_α (resp. $T_{n-1,\alpha}$).

3.2.- Calcul de la valeur U_β. Nous avons

$$P\left(\frac{\overline{X} - \theta_1}{\sigma/\sqrt{n}} > U_\beta\right) = \beta$$

avec

$$U_\beta = \begin{cases} Z_\beta \sim \mathcal{N}(0,1), & \text{loi normale sii la variance } \sigma^2 \text{ de la population est connue} \\ T_{n-1,\beta} \sim t_{n-1,\beta} & \text{loi de Student si la variance } \sigma^2 \text{ de la population est inconnue} \end{cases}$$

Par lecture des tables de la distribution normale (resp. de Student) on établit la valeur de Z_β (resp. $T_{n-1,\beta}$).

3.3.- Calcul de la taille minimale n_C de l'échantillon :

$$n_C = \frac{\left(U_\alpha + U_\beta\right)^2 \Sigma^2}{(\theta_0 - \theta_1)_D^2}$$

avec

$$\Sigma^2 = \begin{cases} \sigma^2, & \text{si la variance } \sigma^2 \text{ de la population est connue} \\ s^2 = \frac{1}{n-1} \sum_{k=1}^{n} \left(X_k - \overline{X}\right)^2 & \text{si la variance } \sigma^2 \text{ de la population est inconnue} \end{cases}$$

Si $n_C > n$, où n la taille de l'échantillon, il faut recommencer l'échantillonnage avec une valeur de n plus grande.

N.B. L'indice D à la différence $(\theta_0 - \theta_1)_D^2$ indique que nous devons travailler sous l'hypothèse de détection $\theta_0 \neq \theta_1$, sinon on risque de faire une division par 0.

4°.- Calcul de la valeur du critère C :

$$C = \pm\frac{1}{2}\left(U_\beta - U_\alpha\right)\sqrt{\frac{\Sigma^2}{n}} + \frac{(\theta_0+\theta_1)_D}{2}$$

On prend le signe "+" si $H_1 : \theta < \theta_0$, on prend le signe "-" si $H_1 : \theta > \theta_0$

5°.- Décision :

Si $H_1 : \theta > \theta_0$, alors l'hypothèse H_0 est acceptée si $\overline{X} \leq C$.

Si $H_1 : \theta < \theta_0$, alors l'hypothèse H_0 est acceptée si $\overline{X} \geq C$.

7.3 Test double de la moyenne avec une valeur de référence

Comme précédement, on distingue deux cas :

– les variances de deux populations sont connues ;

– les variances de deux populations sont inconnues

La procédure du test est la suivante.

1°.- $H_0 : \theta = \theta_0$; $H_1 : \theta \neq \theta_0$; $H_D : \mid \theta_1 - \theta_0 \mid = d$

2°.- On fixe les probabilités α et β.

3°.- Calcul de la taille de l'échantillon :

3.1.- Calcul de la valeur $U_{\alpha/2}$. Nous avons

$$P\left(\frac{\theta_0 - \overline{X}}{\sigma/\sqrt{n}} > U_{\alpha/2}\right) = \frac{\alpha}{2}$$

avec

$$U_{\alpha/2} = \begin{cases} Z_{\alpha/2} \sim \mathcal{N}(0,1), & \text{loi normale si la variance } \sigma^2 \text{ de la population est connue} \\ T_{n-1,\alpha/2} \sim t_{n-1,\alpha/2} & \text{loi de Student si la variance } \sigma^2 \text{ de la population est inconnue} \end{cases}$$

Par lecture des tables de la distribution normale (resp. de Student) on établit la valeur de $Z_{\alpha/2}$ (resp. $T_{n-1,\alpha/2}$).

3.2.- Calcul de la valeur U_β. Nous avons

$$P\left(\frac{\overline{X} - \theta_1}{\sigma/\sqrt{n}} > U_\beta\right) = \beta$$

avec

$$U_\beta = \begin{cases} Z_\beta \sim \mathcal{N}(0,1), & \text{loi normale si la variance } \sigma^2 \text{ de la population est connue} \\ T_{n-1,\beta} \sim t_{n-1,\beta} & \text{loi de Student si la variance } \sigma^2 \text{ de la population est inconnue} \end{cases}$$

Par lecture des tables de la distribution normale (resp. de Student) on établit la valeur de Z_β (resp. $T_{n-1,\beta}$).

3.3.- Calcul de la taille minimale n_C de l'échantillon :

$$n_C = \frac{\left(U_{\alpha/2} + U_\beta\right)^2 \Sigma^2}{(\theta_0 - \theta_1)^2}$$

Si $n_C > n$, où n la taille de l'échantillon, il faut recommencer l'échantillonnage avec une valeur de n plus grande.

4°.- Calcul des valeurs des critères C_1 et C_2 :

$$C_1 = \frac{1}{2}\left(Z_\beta - Z_{\alpha/2}\right)\sqrt{\frac{\sigma^2}{n}} + \frac{(\theta_0 + \theta_1)_D}{2} \quad ; \quad C_2 = 2\theta_0 - C_1$$

5°.- Décision :

Si $\overline{X} \in [C_1, C_2]$, alors l'hypothèse H_0 est acceptée.

7.4

Comparaison simple entre les moyennes

Nous avons deux séries de mesures représentées par les deux séries des v.a. $X_1, \ldots, X_n$ et $Y_1, \ldots, Y_m$. Ces deux séries de mesures sont issues

- soit de la même population et, par conséquent, nous voulons tester l'homogénéïté de la population ;
- soit de deux populations différentes et, par conséquent, nous voulons tester l'équivalence de deux populations.

On distingue trois cas :

- les variances de deux populations sont connues ;
- les variances de deux populations sont inconnues et considérées comme étant égales, et
- les variances de deux populations sont inconnues et non égales.

La procédure du test est la suivante.

1°.- $H_0 : \theta_X - \theta_Y = 0$; $H_1 : \theta_X - \theta_Y > 0$; $H_D : \theta_X - \theta_Y = d$ ou $\theta_X - \theta_Y > 0$ dans le cas des variances inconnues et non égales.

2°.- On fixe les probabilités α et β.

3°.- Calcul de la taille de l'échantillon :

3.1.- Calcul de la valeur U_α. Nous avons

$$P\left(\frac{\overline{X} - \overline{Y} - (\theta_X - \theta_Y)}{\sigma_W} > U_\alpha\right) = \alpha$$

avec

$$U_\alpha = \begin{cases} Z_\alpha \sim \mathcal{N}(0,1) & \text{si les variances } \sigma_X^2, \sigma_Y^2 \text{ des populations connues} \\ T_{\alpha,\nu} \sim t_{\alpha,\nu} & \text{si les variances sont inconnues et égales } (\nu = n+m-2) \\ T_{\alpha,\nu} \sim t_{\alpha,\nu} & \text{si les variances sont inconnues et non égales } \left(\nu = \dfrac{\left(\frac{s_X^2}{n} + \frac{s_Y^2}{m}\right)^2}{\frac{\left(\frac{s_X^2}{n}\right)^2}{n+1} + \frac{\left(\frac{s_Y^2}{m}\right)^2}{m+1}} - 2 \right) \\ & \text{N.B. Si } \nu > n+m-2, \text{il faut recommencer avec des échantillons plus grands.} \end{cases}$$

et

$$\sigma_W^2 = \begin{cases} \dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m} & \text{si les variances } \sigma_X^2, \sigma_Y^2 \text{ des populations connues} \\ \dfrac{(n-1)s_X^2 + (m-1)s_Y^2}{n+m-2} \left(\dfrac{1}{n} + \dfrac{1}{m}\right) & \text{si les variances sont inconnues et égales} \\ \dfrac{s_X^2}{n} + \dfrac{s_Y^2}{m} & \text{si les variances sont inconnues et non égales} \end{cases}$$

Dans les formules précédentes, on a posé

$$s_X^2 = \frac{1}{n-1} \sum_{k=1}^{n} \left(X_k - \overline{X}\right)^2, s_Y^2 = \frac{1}{n-1} \sum_{k=1}^{n} \left(Y_k - \overline{Y}\right)^2$$

Par lecture des tables de la distribution normale (resp. de Student) on établit la valeur de Z_α (resp. de $T_{\alpha,\nu}$).

3.2.- Calcul de la valeur U_β. Nous avons

$$P\left(\frac{(\theta_X - \theta_Y)_D - \left(\overline{X} - \overline{Y}\right)}{\sigma_W} > U_\beta\right) = \beta$$

avec

$$U_\beta = \begin{cases} Z_\beta \sim \mathcal{N}(0,1) & \text{si les variances } \sigma_X^2, \sigma_Y^2 \text{ des populations connues} \\ T_{\beta,\nu} \sim t_{\beta,\nu} & \text{si les variances sont inconnues et égales} \end{cases}$$

Par lecture des tables de la distribution normale (resp. de Student) on établit la valeur de Z_β (resp. de $T_{\beta,\nu}$).

3.3.- Calcul de la taille minimale n_C de l'échantillon :

$$n_C = \begin{cases} \frac{\left(Z_\alpha + Z_\beta\right)^2}{(\theta_X - \theta_Y)_D^2} \left(\sigma_X^2 + \sigma_Y^2\right) & \text{si les variances } \sigma_X^2, \sigma_Y^2 \text{ des populations connues} \\ \frac{\left(T_{\alpha,\nu} + T_{\beta,\nu}\right)^2}{(\theta_X - \theta_Y)_D^2} \left(s_X^2 + s_Y^2\right) & \text{si les variances sont inconnues et égales} \end{cases}$$

Si $n_C > n$, alors il faut recommencer avec un échantillon plus grand.

4°.- Calcul de la valeur du critère C :

$$C = \begin{cases} \frac{1}{2}\left(Z_\alpha - Z_\beta\right)\sqrt{\frac{\sigma_X^2}{n}+\frac{\sigma_Y^2}{m}}+\frac{(\theta_X-\theta_Y)_D}{2} & \text{si les variances } \sigma_X^2, \sigma_Y^2 \text{ des populations connues} \\ \frac{1}{2}\left(T_{\alpha,\nu} - T_{\beta,\nu}\right)\sqrt{\frac{(n-1)s_X^2+(m-1)s_Y^2}{n+m-2}}\cdot\sqrt{\frac{1}{n}+\frac{1}{m}}+\frac{(\theta_X-\theta_Y)_D}{2} & \text{si les variances sont inconnues et égales} \\ T_{\alpha;\nu}\sqrt{\frac{s_X^2}{n}+\frac{s_Y^2}{m}} & \text{si les variances sont inconnues et non égales} \end{cases}$$

5°.- Décision :

Si $\overline{X}-\overline{Y} \leq C$, alors l'hypothèse H_0 est acceptée.

7.5

Comparaison double entre les moyennes

Comme précédemment on distingue trois cas :

– les variances de deux populations sont connues ;
– les variances de deux populations sont inconnues et considérées comme étant égales, et
– les variances de deux populations sont inconnues et non égales.

La procédure du test est la suivante.

1°.- $H_0 : \theta_X - \theta_Y = 0$; $H_1 : \theta_X - \theta_Y \neq 0$; $H_D : \theta_X - \theta_Y = d$ ou $\theta_X - \theta_Y > 0$ dans le cas des variances inconnues et non égales.

2°.- On fixe les probabilités α et β.

3°.- Calcul de la taille de l'échantillon :

3.1.- Calcul de la valeur $U_{\alpha/2}$. Nous avons

$$P\left(\frac{\overline{X}-\overline{Y}-(\theta_X-\theta_Y)}{\sigma_W} > U_{\alpha/2}\right) = \frac{\alpha}{2}\ ;\ Z_{\alpha/2} \sim \mathcal{N}(0,1)$$

avec

$$U_\alpha = \begin{cases} Z_{\alpha/2} \sim \mathcal{N}(0,1) & \text{si les variances } \sigma_X^2, \sigma_Y^2 \text{ des populations connues} \\ T_{\alpha/2,\nu} \sim t_{\alpha/2,\nu} & \text{si les variances sont inconnues et égales } (\nu = n+m-2) \\ T_{\alpha/2,\nu} \sim t_{\alpha/2,\nu} & \text{si les variances sont inconnues et non égales } \left(\nu = \dfrac{\left(\frac{s_X^2}{n}+\frac{s_Y^2}{m}\right)^2}{\frac{\left(\frac{s_X^2}{n}\right)^2}{n+1}+\frac{\left(\frac{s_Y^2}{m}\right)^2}{m+1}} - 2\right) \\ & \text{N.B. Si } \nu > n+m-2, \text{ il faut recommencer avec des échantillons plus grands.} \end{cases}$$

et

$$\sigma_W^2 = \begin{cases} \dfrac{\sigma_X^2}{n} + \dfrac{\sigma_Y^2}{m} & \text{si les variances } \sigma_X^2, \sigma_Y^2 \text{ des populations connues} \\ \dfrac{(n-1)\, s_X^2 + (m-1)\, s_Y^2}{n+m-2} \left(\dfrac{1}{n} + \dfrac{1}{m} \right) & \text{si les variances sont inconnues et égales} \\ \dfrac{s_X^2}{n} + \dfrac{s_Y^2}{m} & \text{si les variances sont inconnues et non égales} \end{cases}$$

Par lecture des tables de la distribution normale on établit la valeur de Z_α.

3.2.- Calcul de la valeur U_β. Nous avons

$$P\left(\frac{(\theta_X - \theta_Y)_D - \left(\overline{X} - \overline{Y}\right)}{\sigma_W} > U_\beta \right) = \beta$$

avec

$$U_\beta = \begin{cases} Z_\beta \sim \mathcal{N}(0,1) & \text{si les variances } \sigma_X^2, \sigma_Y^2 \text{ des populations connues} \\ T_{\beta,\nu} \sim t_{\beta,\nu} & \text{si les variances sont inconnues et égales} \end{cases}$$

Par lecture des tables de la distribution normale (resp. de Student) on établit la valeur de Z_β (resp. de $T_{\beta,\nu}$).

3.3.- Calcul de la taille minimale n_C de l'échantillon :Calcul de la taille minimale n_C de l'échantillon :

$$n_C = \begin{cases} \frac{\left(Z_{\alpha/2} + Z_\beta\right)^2}{(\theta_X - \theta_Y)_D^2} \left(\sigma_X^2 + \sigma_Y^2\right) & \text{si les variances } \sigma_X^2, \sigma_Y^2 \text{ des populations connues} \\ \frac{\left(T_{\alpha/2,\nu} + T_{\beta,\nu}\right)^2}{(\theta_X - \theta_Y)_D^2} \left(s_X^2 + s_Y^2\right) & \text{si les variances sont inconnues et égales} \end{cases}$$

Si $n_C > n$, alors il faut recommencer avec un échantillon plus grand.Si $n_C > n$, alors il faut recommencer avec un échantillon plus grand.

4°.- Calcul de la valeur du critère C :

$$C = \begin{cases} \frac{1}{2}\left(Z_\alpha - Z_\beta\right) \sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} + \frac{(\theta_X - \theta_Y)_D}{2} & \text{si les variances } \sigma_X^2, \sigma_Y^2 \text{ des populations connues} \\ \frac{1}{2}\left(T_{\alpha,\nu} - T_{\beta,\nu}\right) \sqrt{\frac{(n-1)s_X^2 + (m-1)s_Y^2}{n+m-2}} \cdot \sqrt{\frac{1}{n} + \frac{1}{m}} + \frac{(\theta_X - \theta_Y)_D}{2} & \text{si les variances sont inconnues et égales} \\ T_{\alpha/2;\nu} \sqrt{\dfrac{s_X^2}{n} + \dfrac{s_Y^2}{m}} & \text{si les variances sont inconnues et non égales} \end{cases}$$

5°.- Décision :

Si $\overline{X} - \overline{Y} \in [-C, C]$, alors l'hypothèse H_0 est acceptée.

7.6
Démarche pour effectuer un test

D'une façon générale les démarches à suivre pour réaliser n'importe quel test sont les suivantes :

(1) Définir bien les hypothèses

(2) Écrire les équations qui déterminent α et β. (On disposez de deux équations)
Attention : Dans un test il y a quatre paramètres α, β, C, n. On aura deux qui sont définis et les deux autres on doit les calculer.

(3) Définir bien la v.a (estimateur naturel) qu'on doit utiliser pour le test, sa loi, ses paramètres. Pour calculer les probabilités on utilise les tables de la loi Normale, $\mathcal{X}^2$, Student et Fisher.

(4) Dessiner les courbes qui représentent les hypothèses (distributions) à tester, les risques α , , β , C. On indique la partie de la demi-droite où on doit refuser ou accepter H_0

Ce corrigé concerne tous les exercices du TD n°10.

Les formules utilisées ci-dessous peuvent être utilisées dans tous les tests si la loi suivie par la population est la loi Normale ou si on considère un grand échantillon (par T.C.L on arrive à la loi Normale).

Dans ces algorithmes on suppose connaître :

(1) α et β ce qui nous oblige à calculer la taille nécessaire et C. (Souvent un effectif est proposé, il suffit de constater si il est correct ou non)

(2) dans le cas des tests bilatéraux, pour le calcul de la region critique, le point C_1 est l'extrémité gauche de l'intervalle.

Remarque : On peut utiliser une partie seulement de ces algorithmes dans le cas où β n'est pas fixé.

7.7
Exercices

EXERCICE 7.1 *Un fabricant produit des piles avec durée de vie* 80 *heures. Suite à des réclamations il veut vérifier si la qualité a baissée ou non. Dans le cas où la moyenne de la durée de vie est inférieure à* 75 *heures, il sera obligé de baisser le prix de vente..*

Mettez au point un test avec $\alpha = 0.025$, $\beta = 0.05$. *On suppose* $\sigma = 6.44$.

Application : On a calculé $\bar{x} = 78$ *heures. Conclusion ?*

EXERCICE 7.2 *Supposons maintenant que le fabricant vend trois qualités de piles* A , B *et* C. *Leur durée de vie est, respectivement,* 75 , 80 *et* 85. *Un client souhaite vérifier si les piles reçues appartiennent à la qualité moyenne.*

Mettez au point un test avec : $\alpha = 0.05$, $\beta = 0.05$ *et* $\sigma = 6.44$.

Application : On a calculé $\bar{x} = 83$ *heures. Conclusion ?*

EXERCICE 7.3 *Le même fabricant veut solder ses piles de seconde choix qui ont une durée de vie de* 60 *heures. Il veut bien vérifier que les lots n'appartiennent pas à la qualité* 70 *heures avec probabilité de* 95%. *Mettez au point un test avec* $\alpha = 0.05$ *et en disposant de deux séries de mesures :*

- *Échantillon de taille* 11, *avec* $s = 10$ *et* $\bar{x} = 65$ *;*
- *Échantillon de taille* 21, *avec* $s = 11$ *et* $\bar{x} = 64$.

EXERCICE 7.4 *Un autre client voudrait avoir de piles dont la moyenne de la durée de vie est de* 100 *heures avec une confiance de* 95%. *Le client ne veut pas accepter de piles avec une durée de vie inférieure à* 90 *heures, (qualité mauvaise) et le fabricant décide à ne pas livrer à son client des piles avec une durée de vie supérieure à* 110 *heures, (qualité supérieure).* $\beta = 0.025$.

Mettez au point un test avec une série de mesures effectuées sur un échantillon de taille 11, *avec* $s = 5$ *et* $\bar{x} = 94$.

EXERCICE 7.5 *Deux groupes A et B, de* 30 *élèves chacun, sont notés sur une matière de* 0100. *La moyenne des notes est de* 80 *pour le premier groupe et de* 72 *pour le second. D'autre part on sait que l'écart-type des notes est, toujours, de* 5 *pour le groupe A et de* 10 *pour le groupe B. Nous dirons qu'il y a une différence significative entre les deux groupes si la différence des moyennes dépasse* 10 *points.*

Si nous prenons $\alpha = \beta = 0.01$, *que peut-on en conclure ?*

EXERCICE 7.6 *Pour deux types d'essence nous faisons des mesures du nombre d'octanes.*

Pour le premier type nous avons les mesures suivantes : $81, 84, 79, 76, 82, 85, 88, 84, 80, 79, 82, 81$ *et pour le second :* $76, 74, 78, 79, 80, 79, 82, 76, 81, 79, 82, 78$.

Le deuxième type d'essence étant moins cher que le premier, on décide d'en acheter à condition que les deux types ont en moyenne le même nombre d'octanes avec une confiance de 99%. *Toutefois si le nombre moyen d'octanes du premier type est supérieur ou égal d'au moins cinq octanes du nombre moyen du second type, avec une confiance de* 95%, *on achètera de l'essence du premier type.*

Selon vous, que doit-on faire, si on suppose que les variances de deux types sont égales ?

EXERCICE 7.7 *Un traitement spécifique modifie la durée de vie d'un appareil. Nous possédons deux séries de mesures concernant la durée de vie en mois des appareils : une pour des appareils traités et qui sont les suivantes :*

$31, 34, 29, 26, 32, 35, 38, 34, 30, 29, 32, 31$

et une pour des appareils non traités :

$26, 24, 28, 29, 30, 29, 32, 26, 31, 29, 32, 28$

On voudrait savoir si le traitement spécifique modifie la durée de vie moyenne des appareils traités avec confiance de 99%.

Que doit-on en conclure si on suppose que les variances de deux populations sont égales et que $\beta = 0.1$ *?*

EXERCICE 7.8 *Décider, sur la base des données ci-après, si la valeur d'une résistance* y *est aussi élevée que la valeur d'une résistance étalon* x *avec* $\alpha = 0.05$.

$s_x^2 = 10.02$, $\bar{x} = 32.25\ \Omega \cdot cm \times 10^6$, $n_x = 18$

$s_y^2 = 6.05$, $\bar{y} = 28.50\ \Omega \cdot cm \times 10^6$, $n_y = 18$

EXERCICE 7.9 *Décider, sur la base des données ci-après issues de deux séries de mesures, si, avec un niveau de confiance de* 95%, *il y a une différence significative entre ces deux séries de mesures.* $s_x^2 = 0.0001$, $\bar{x} = 0.070$, $n_x = 15$.

$s_y^2 = 0.0002$, $\bar{y} = 0.081$, $n_y = 15$

Exercices à faire chez soi.

EXERCICE 7.10 *Une machine produit des billes de roulement de diamètre fixe. Si elle fonctionne normalement il y a une proportion de 5% de billes défectueuses. Si elle est dereglée la proportion de billes défectueuses est de 10%.*

Soit X la variable aléatoire qui représente le nombre de billes défectueuses.

(1) *Le fabricant teste des lots de 500 billes. Il ne livre pas à ses clients si le nombre de billes défectueuses X est* $X \geq 50$.
   (a) *Déterminer le niveau et la puissance du test.*
   (b) *En déduire si un client doit acheter ou non un lot qui a subi un tel test.*
(2) *On procède au test précédent sur le même nombre de billes et on pose comme niveau du test* $\alpha = 0.01$.
   (a) *Évaluer la région critique du test et calculer sa puissance.*
   (b) *Le client décide à procéder à son propre test avec la même valeur du niveau. Évaluer la région critique de ce test.*
   (c) *Lors d'une vérification on a trouvé* $X = 35$. *Le client doit-il accepter ce lot ?*
(3) *Le client et le fabricant décident de procéder comme suit : Chaque partie fera son propre test sur le même échantillon de taille* n *et avec le même niveau* $\alpha = 0.01$.
   (a) *Si on veut que les deux tests donnent les mêmes résultats quelle doit être la valeur de* n *?*
   (b) *Dans ce cas quelles sont les régions critiques pour les deux parties ?*

EXERCICE 7.11 *Nous voulons tester un nouveau médicament T contre la fièvre, en le comparant avec un ancien médicament C. La comparaison concerne la durée moyenne entre la prise du médicament et la manifestation des premiers symptômes de son action (baisse de la temperature). Cette durée moyenne est de* 1 *heure pour le médicament C. Nous déciderons que T est meilleur si la réduction de la durée n'est pas inférieure à* 0.5 *heure. On cherche à construire le test en supposant que les écarts-types des tests concernant T et C ont la même valeur, égale à* 0.12 *h. Si nous prenons* $\alpha = \beta = 0.01$ *et si pour un test de chacun des médicaments sur* 10 *malades nous avons eu comme valeur de la durée moyenne* 0.9*h pour C et* 0.6 *h pour* T, *que peut-on en conclure ?*