

Licence mention Mathématiques et mention Informatique parcours MIAGE - Semestre 3
Statistique et Probabilités

Estimation, intervalle de confiance, test statistique (suite)
Cas d'une ou de deux moyennes, d'une ou de deux variances

1. Introduction

On s'intéresse à l'étude d'un **caractère** (quantitatif ou qualitatif) des N individus d'une **population**. Pour chacun des individus de la population, le caractère peut a priori prendre des valeurs aléatoirement différentes. Ainsi, le caractère peut être représenté par une **variable aléatoire** X .

Lorsque le caractère est **quantitatif** (taille des individus,...), X sera une variable aléatoire égale aux valeurs du caractère ; on supposera en général que X est une variable aléatoire d'**espérance mathématique** (**moyenne**) μ , d'**écart-type** σ , et éventuellement de **loi normale**.

Lorsqu'on n'a pas accès à l'ensemble de la population, on procède à un **échantillonnage**, i.e. au choix de n individus dans la population, sur lesquels on observe la valeur x du caractère X . On aura ainsi un échantillon (X_1, X_2, \dots, X_n) est un **échantillon de taille n de X** ; pour tout $i = 1, \dots, n$, la variable aléatoire X_i correspond aux valeurs du caractère du i -ème individu obtenu par échantillonage, et aura donc la **même loi de probabilité que X** . De plus, l'échantillonnage étant non-exhaustif (tirages avec remise), les variables aléatoires X_i sont indépendantes.

Exemple introductif sur la moyenne

On considère un groupe de quatre enfants, Alexis, Benjamin, Cyril et David, d'âges respectifs 12, 13, 14 et 15 ans. Lorsqu'on choisit un enfant au hasard dans le groupe, on peut considérer :

- X , âge de l'enfant, variable aléatoire de loi uniforme sur $\{12, 13, 14, 15\}$:

$$P(X = 12) = \dots = P(X = 15) = \frac{1}{4}, \text{ de moyenne } \mu = 13,5 \text{ et d'écart-type } \sigma = \sqrt{1.25} \approx 1.118 ;$$

Cherchons à retrouver ou à approcher ces résultats à partir d'échantillons non-exhaustifs (avec remise) de taille $n = 3$. Il y en a $4^3 = 64$, ils forment un univers Ω' , ensemble des résultats possibles de l'expérience aléatoire "choisir un échantillon". On peut munir Ω' de la tribu des événements $\mathcal{A}' = \mathcal{P}(\Omega')$ et de l'équiprobabilité P' sur (Ω', \mathcal{A}') . A chacun des résultats (échantillons) ω , on peut associer la moyenne $\bar{X}(\omega) = \bar{x}$ des âges de l'échantillon. On obtient les résultats présentés dans le tableau page 2.

On définit ainsi une variable aléatoire \bar{X} , dont on peut obtenir la loi de probabilité :

x_i	12,00	12,33	12,67	13,00	13,33	13,67	14,00	14,33	14,67	15,00
$P(\bar{X} = x_i)$	1/64	3/64	6/64	10/64	12/64	12/64	10/64	6/64	3/64	1/64

On peut alors calculer :

$$- E(\bar{X}) = \sum x_i P(\bar{X} = x_i) = 13,5 : \text{on remarque que } E(\bar{X}) = \mu = E(X).$$

$$- Var(\bar{X}) = \sum x_i^2 P(\bar{X} = x_i) - (E(\bar{X}))^2 = \frac{5}{12} : \text{on remarque que } Var(\bar{X}) = \frac{\sigma^2}{n} = \frac{Var(X)}{n}.$$

2. Estimateur - Estimation

2.1. Moyenne et variance d'échantillon

Considérons un caractère quantitatif représenté par une variable aléatoire X d'espérance mathématique μ , d'écart-type σ , et un échantillon (X_1, X_2, \dots, X_n) de taille n de X .

Pour chaque échantillonnage on peut calculer la moyenne observée du caractère $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$,

$\bar{x}' = \frac{1}{n} \sum_{i=1}^n x'_i$, ... Ces moyennes observées peuvent être considérées comme les valeurs observées de la

variable aléatoire $\bar{X} = \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$, **moyenne d'échantillon**.

ω	\bar{x}	ω	\bar{x}	ω	\bar{x}	ω	\bar{x}
(A,A,A)	12	(B,A,A)	12,33	(C,A,A)	12,67	(D,A,A)	13
(A,A,B)	12,33	(B,A,B)	12,67	(C,A,B)	13	(D,A,B)	13,33
(A,A,C)	12,67	(B,A,C)	13	(C,A,C)	13,33	(D,A,C)	13,67
(A,A,D)	13	(B,A,D)	13,33	(C,A,D)	13,67	(D,A,D)	14
(A,B,A)	12,33	(B,B,A)	12,67	(C,B,A)	13	(D,B,A)	13,33
(A,B,B)	12,67	(B,B,B)	13	(C,B,B)	13,33	(D,B,B)	13,67
(A,B,C)	13	(B,B,C)	13,33	(C,B,C)	13,67	(D,B,C)	14
(A,B,D)	13,33	(B,B,D)	13,67	(C,B,D)	14	(D,B,D)	14,33
(A,C,A)	12,67	(B,C,A)	13	(C,C,A)	13,33	(D,C,A)	13,67
(A,C,B)	13	(B,C,B)	13,33	(C,C,B)	13,67	(D,C,B)	14
(A,C,C)	13,33	(B,C,C)	13,67	(C,C,C)	14	(D,C,C)	14,33
(A,C,D)	13,67	(B,C,D)	14	(C,C,D)	14,33	(D,C,D)	14,67
(A,D,A)	13	(B,D,A)	13,33	(C,D,A)	13,67	(D,D,A)	14
(A,D,B)	13,33	(B,D,B)	13,67	(C,D,B)	14	(D,D,B)	14,33
(A,D,C)	13,67	(B,D,C)	14	(C,D,C)	14,33	(D,D,C)	14,67
(A,D,D)	14	(B,D,D)	14,33	(C,D,D)	14,67	(D,D,D)	15

On démontre que $E(\bar{X}) = \mu$ (on dit que \bar{X} est un **estimateur sans biais** de μ) et $Var(\bar{X}) = \frac{\sigma^2}{n}$.

Pour une observation (x_1, x_2, \dots, x_n) de l'échantillon, on dit que $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \hat{\mu}$ est une **estimation ponctuelle** de μ .

De même, on considère la **variance d'échantillon** $S^2 = S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2$.

On a alors $E(S^2) = \frac{n-1}{n} \sigma^2$ et S^2 est un **estimateur** avec biais de σ^2 . On considère alors la **variance corrigée d'échantillon** $S_c^2 = \frac{n}{n-1} S^2$: on a alors $E(S_c^2) = \sigma^2$ et S_c^2 est un **estimateur sans biais** de σ^2 .

Pour une observation (x_1, x_2, \dots, x_n) de l'échantillon, une **estimation ponctuelle** de σ^2 est

$$s_c^2 = \frac{n-1}{n} s^2 = \hat{\sigma}^2 \text{ avec } s^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2.$$

2.2. Loi de probabilité des estimateurs

Cas d'un petit échantillon gaussien : $n \leq 30$ et X de loi normale $\mathcal{N}(\mu; \sigma)$

Si σ est connu (cas peu utile en pratique), alors $U = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ suit la loi normale $\mathcal{N}(0; 1)$.

Si σ est inconnu, alors $T = \frac{\bar{X} - \mu}{\frac{S_c}{\sqrt{n}}}$ suit la loi de Student à $n-1$ degrés de liberté.

Cas d'un grand échantillon : $n > 30$ (et X de loi quelconque)

Dans ce cas, $U = \frac{\bar{X} - \mu}{\frac{S_c}{\sqrt{n}}}$ suit approximativement la loi normale $\mathcal{N}(0; 1)$.

Cas d'un échantillon gaussien : X de loi normale $\mathcal{N}(\mu; \sigma)$

Dans ce cas, $Y^2 = \frac{n-1}{\sigma^2} S_c^2$ suit la loi de khi deux à $n-1$ degrés de liberté.

3. Intervalle de confiance

3.1. Pour une moyenne μ

Considérons un caractère quantitatif représenté par une variable aléatoire X d'espérance mathématique μ , d'écart-type σ , et un échantillon (X_1, X_2, \dots, X_n) de taille n de X . La moyenne d'échantillon est $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ et la variance corrigée d'échantillon est $S_c^2 = \frac{n}{n-1} S^2$, avec $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2$.

3.1.1. Cas d'un petit échantillon gaussien : $n \leq 30$ et X de loi normale $\mathcal{N}(\mu; \sigma)$

3.1.1.1. Cas σ connu (peu utile en pratique)

On sait que $U = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ suit la loi normale $\mathcal{N}(0; 1)$.

On fixe une valeur $\alpha \in]0, 1[$. On peut trouver un réel u_α tel que $P(-u_\alpha < U < u_\alpha) = 1 - \alpha$ (voir table 2).

$$\begin{aligned} \text{On a alors } 1 - \alpha &= P\left(-u_\alpha < \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} < u_\alpha\right) = P\left(-\frac{\sigma}{\sqrt{n}} u_\alpha < \bar{X} - \mu < \frac{\sigma}{\sqrt{n}} u_\alpha\right) \\ &= P\left(-\bar{X} - \frac{\sigma}{\sqrt{n}} u_\alpha < -\mu < -\bar{X} + \frac{\sigma}{\sqrt{n}} u_\alpha\right) = P\left(\bar{X} - \frac{\sigma}{\sqrt{n}} u_\alpha < \mu < \bar{X} + \frac{\sigma}{\sqrt{n}} u_\alpha\right). \end{aligned}$$

On dit que $I_\mu = \left[\bar{X} - \frac{\sigma}{\sqrt{n}} u_\alpha, \bar{X} + \frac{\sigma}{\sqrt{n}} u_\alpha \right]$ est un **intervalle de confiance de μ au niveau $1 - \alpha$** (ou au seuil α). En pratique, on a une observation \bar{x} de \bar{X} , d'où une observation de cet intervalle :

$$i_\mu = \left[\bar{x} - \frac{\sigma}{\sqrt{n}} u_\alpha, \bar{x} + \frac{\sigma}{\sqrt{n}} u_\alpha \right].$$

3.1.1.2. Cas σ inconnu

On sait que $T = \frac{\bar{X} - \mu}{\frac{S_c}{\sqrt{n}}}$ suit la loi de Student à $n - 1$ degrés de liberté.

On détermine alors le réel t_α tel que $P(-t_\alpha < T < t_\alpha) = 1 - \alpha$ (table 3). On en déduit un **intervalle de confiance de μ au niveau $1 - \alpha$** :

$$i_\mu = \left[\bar{x} - \frac{S_c}{\sqrt{n}} t_\alpha, \bar{x} + \frac{S_c}{\sqrt{n}} t_\alpha \right].$$

3.1.2. Cas d'un grand échantillon : $n > 30$

On sait que $U = \frac{\bar{X} - \mu}{\frac{S_c}{\sqrt{n}}}$ suit approximativement la loi normale $\mathcal{N}(0; 1)$. On procède alors comme au

3.1.1.1. en remplaçant σ par s_c et on obtient un **intervalle de confiance approché de μ au niveau $1 - \alpha$** :

$$i_\mu = \left[\bar{x} - \frac{s_c}{\sqrt{n}} u_\alpha, \bar{x} + \frac{s_c}{\sqrt{n}} u_\alpha \right]$$

3.2. Pour une variance σ^2

Considérons un caractère quantitatif représenté par une variable aléatoire X de loi normale $\mathcal{N}(\mu; \sigma)$, et un échantillon (X_1, X_2, \dots, X_n) de taille n de X . La moyenne d'échantillon est \bar{X} et la variance corrigée d'échantillon est $S_c^2 = \frac{n}{n-1} S^2$.

Alors $Y^2 = \frac{n-1}{\sigma^2} S_c^2$ suit la loi de khi deux à $n - 1$ degrés de liberté.

On détermine alors les réels a_α et b_α tels que $P(Y^2 \geq a_\alpha) = 1 - \frac{\alpha}{2}$ et $P(Y^2 \geq b_\alpha) = \frac{\alpha}{2}$ (table 4).

On a alors

$$1 - \alpha = P\left(a_\alpha < \frac{n-1}{\sigma^2} S_c^2 < b_\alpha\right) = P\left(\frac{1}{b_\alpha} < \frac{\sigma^2}{(n-1)S_c^2} < \frac{1}{a_\alpha}\right) = P\left(\frac{(n-1)S_c^2}{b_\alpha} < \sigma^2 < \frac{(n-1)S_c^2}{a_\alpha}\right).$$

On en déduit un **intervalle de confiance de la variance σ^2 au niveau $1 - \alpha$** :

$$i_{\sigma^2} = \left[\frac{n-1}{b_\alpha} s_c^2, \frac{n-1}{a_\alpha} s_c^2 \right].$$

3.3. Exemple

Un échantillon de 30 enfants d'une ville donnée à fourni les tailles suivantss (en cm) :

70	85	93	99	101	105	110	121	138	166
74	85	93	99	102	106	110	125	140	180
79	87	94	99	102	107	114	128	147	180

On peut considérer la situation suivante.

Population : les enfants de la ville considérée.

Caractère : la taille, variable aléatoire de moyenne μ et de variance σ^2 (écart-type σ).

Echantillon (X_1, X_2, \dots, X_n) de taille $n = 30$ de X .

Observation de l'échantillon : $(x_1, x_2, \dots, x_n) = (70, 74, 79, \dots, 180)$.

1) Estimateurs

a) de la moyenne μ : $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$; b) de la variance σ^2 : $S_c^2 = \frac{n}{n-1} S^2$, avec $S^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2$.

2) Estimations ponctuelles

- a) de la moyenne μ : $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{30} \times 3339 = 111,3$;
 b) de la variance σ^2 : $s^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - \bar{x}^2 = \frac{1}{30} \times 395347 - (111,3)^2 \simeq 790,54$;
 et donc $s_c^2 = \frac{n}{n-1} s^2 = \frac{30}{29} \times 790,54 \simeq 817,80$;
 c) de l'écart-type σ : $s_c = \sqrt{s_c^2} \simeq 28,60$.

3) Intervalle de confiance

a) de la moyenne μ

1^{ère} méthode : X supposée de loi normale $\mathcal{N}(\mu; \sigma)$, σ inconnu

$$T = \frac{\bar{X} - \mu}{\frac{S_c}{\sqrt{n}}} \text{ suit la loi de Student à } n-1 \text{ degrés de liberté.}$$

On détermine le réel t_α tel que $P(-t_\alpha < T < t_\alpha) = 1 - \alpha$ (table 3).

$$\text{On en déduit un intervalle de confiance de } \mu \text{ au niveau } 1 - \alpha : i_\mu = \left[\bar{x} - \frac{s_c}{\sqrt{n}} t_\alpha ; \bar{x} + \frac{s_c}{\sqrt{n}} t_\alpha \right].$$

On a $n = 30$ et $n-1 = 29$. Pour $\alpha = 0,05$ (i.e. 5%), on a $t_\alpha = 2,045$,
 et donc un intervalle de confiance de μ au niveau 0,95 (i.e. 95%) est $i_\mu = [100,6 ; 122,0]$.

2^{ème} méthode : grand échantillon ($n > 30$, pas d'hypothèse sur la loi de X)

$$U = \frac{\bar{X} - \mu}{\frac{S_c}{\sqrt{n}}} \text{ suit approximativement la loi normale } \mathcal{N}(0; 1).$$

On détermine le réel u_α tel que $P(-u_\alpha < U < u_\alpha) = 1 - \alpha$.

$$\text{On en déduit un intervalle de confiance de } \mu \text{ au niveau } 1 - \alpha : i_\mu = \left[\bar{x} - \frac{s_c}{\sqrt{n}} u_\alpha ; \bar{x} + \frac{s_c}{\sqrt{n}} u_\alpha \right].$$

On a $n = 30$. Pour $\alpha = 0,05$ (i.e. 5%), on a $u_\alpha = 1,96$,
 et donc un intervalle de confiance de μ au niveau 0,95 (i.e. 95%) est $i_\mu = [101,1 ; 121,5]$.

b) de la variance σ^2 : X supposée de loi normale $\mathcal{N}(\mu; \sigma)$

$$Y^2 = \frac{n-1}{\sigma^2} S_c^2 \text{ suit la loi de chi deux à } n-1 \text{ degrés de liberté.}$$

On détermine alors les réels a_α et b_α tels que $P(Y^2 \geq a_\alpha) = 1 - \frac{\alpha}{2}$ et $P(Y^2 \geq b_\alpha) = \frac{\alpha}{2}$ (table 4).

$$\text{On en déduit un intervalle de confiance de } \sigma^2 \text{ au niveau } 1 - \alpha : i_{\sigma^2} = \left[\frac{n-1}{b_\alpha} s_c^2 ; \frac{n-1}{a_\alpha} s_c^2 \right].$$

On a $n = 30$ et $n-1 = 29$. Pour $\alpha = 0,05$ (i.e. 5%), on a $a_\alpha = 16,05$ et $b_\alpha = 45,72$,
 et donc un intervalle de confiance de σ^2 au niveau 0,95 (i.e. 95%) est $i_{\sigma^2} = [518,73 ; 1477,65]$.

4. Test de conformité

4.1. Pour une moyenne μ

Considérons un caractère quantitatif représenté par une variable aléatoire X d'espérance mathématique μ , d'écart-type σ , et un échantillon (X_1, X_2, \dots, X_n) de taille n de X . La moyenne d'échantillon est $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ et la variance corrigée d'échantillon est $S_c^2 = \frac{n}{n-1} S^2$, avec $S^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2$.

4.1.1. Cas d'un petit échantillon gaussien : $n \leq 30$ et X de loi normale $\mathcal{N}(\mu; \sigma)$

4.1.1.1. Cas σ connu (exemple introductif)

Il s'agit de faire un choix entre plusieurs hypothèses possibles sur μ sans disposer d'informations suffisantes pour que ce choix soit sûr. On met en avant deux hypothèses privilégiées : *l'hypothèse nulle* H_0 et *l'hypothèse alternative* H_1 . Par exemple, on testera $H_0 : \mu = \mu_0$ contre $H_1 : \mu \neq \mu_0$, avec μ_0 fixé arbitrairement. On veut savoir si l'on doit rejeter H_0 ou pas.

Test (bilatéral) de $H_0 : \mu = \mu_0$ contre $H_1 : \mu \neq \mu_0$.

On utilise alors une variable aléatoire dont on connaît la loi de probabilité lorsque H_0 est vraie. Par exemple $U = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$, car lorsque H_0 est vraie, on sait que $U = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ suit la loi $\mathcal{N}(0; 1)$.

On fixe une valeur $\alpha \in]0, 1[$. En général, on prend α petit, le plus souvent 0,05, 0,01, 0,001. On peut trouver un réel u_α tel que $P(-u_\alpha < U < u_\alpha) = 1 - \alpha$. Ce réel u_α peut être trouvé dans la table 2.

On est donc amené à comparer la moyenne \bar{X} de l'échantillon à la moyenne théorique $\mu = \mu_0$. L'hypothèse H_0 signifiera que les différences observées sont seulement dues aux fluctuations d'échantillonnage (i.e. ne sont pas significatives).

On ne rejette pas H_0 si les différences observées ne sont pas significatives, c'est-à-dire si U est "petite", ce que l'on peut traduire par $-u_\alpha < U < u_\alpha$, c'est-à-dire $|U| < u_\alpha$.

On rejette donc H_0 si les différences observées sont significatives, ce que l'on peut traduire par $U > u_\alpha$ ou $U < -u_\alpha$, c'est-à-dire $|U| > u_\alpha$. Par construction de u_α , on a $P(U > u_\alpha) = P(U < -u_\alpha) = \frac{\alpha}{2}$, soit encore $P(|U| > u_\alpha) = \alpha$, i.e. $P(U \notin]-u_\alpha, u_\alpha[) = \alpha$.

En pratique, on calcule $u = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$ et on décide

- de rejeter H_0 si $u \notin]-u_\alpha, u_\alpha[$, car si H_0 était vraie, l'événement $U \notin]-u_\alpha, u_\alpha[$ aurait une probabilité faible de se réaliser ; on pourra dire que la valeur observée \bar{x} n'est pas conforme à la valeur théorique μ_0 mais on ne pourra pas donner de valeur acceptable de μ ;

- de ne pas rejeter H_0 si $u \in]-u_\alpha, u_\alpha[$, car si H_0 était vraie, l'événement $U \in]-u_\alpha, u_\alpha[$ aurait une probabilité forte de se réaliser ; on pourra dire que la valeur observée \bar{x} est conforme à la valeur théorique μ_0 et que la valeur μ_0 ne peut être rejeter. Attention : d'autres valeurs μ'_0, μ''_0, \dots peuvent également convenir.

Erreurs de décision.

Lorsqu'on rejette H_0 alors que H_0 est vraie, on commet une erreur. On a donc une probabilité α de se tromper : α est appelée **erreur de première espèce**. En effet, lorsque H_0 est vraie, on a $P(U \notin]-u_\alpha, u_\alpha[) = \alpha$.

Lorsque l'on ne rejette pas H_0 alors que H_0 est fausse, on commet une erreur. On a une probabilité β de se tromper : β est appelée **erreur de deuxième espèce**. Cette erreur est difficilement calculable. La plupart du temps, on ne connaît pas la loi de U lorsque H_0 est fausse. La valeur $1 - \beta$ est appelée la **puissance du test**.

Test (unilatéral) de $H_0 : \mu = \mu_0$ contre $H_1 : \mu > \mu_0$.

On détermine u'_α tel que $P(U > u'_\alpha) = \alpha$, i.e. $P(U < u'_\alpha) = 1 - \alpha$, i.e. $u'_\alpha = \phi^{-1}(1 - \alpha) = u_{2\alpha}$, et on décide que :

- si $u < u'_\alpha$, alors on ne peut rejeter H_0 ;
- si $u \geq u'_\alpha$, alors on rejette H_0 avec une probabilité α de se tromper.

4.1.1.2. Cas σ inconnu

On sait que $T = \frac{\bar{X} - \mu}{\frac{s_c}{\sqrt{n}}}$ suit la loi de Student à $n - 1$ degrés de liberté.

On détermine alors le réel t_α tel que $P(-t_\alpha < T < t_\alpha) = 1 - \alpha$ (table 3).

Test (bilatéral) de $H_0 : \mu = \mu_0$ contre $H_1 : \mu \neq \mu_0$.

On calcule $t = \frac{\bar{x} - \mu_0}{\frac{s_c}{\sqrt{n}}}$. On détermine t_α tel que $P(-t_\alpha < T < t_\alpha) = 1 - \alpha$ et on décide que :

- si $t \in]-t_\alpha, t_\alpha[$, alors on ne peut rejeter H_0 ;
- si $t \notin]-t_\alpha, t_\alpha[$, alors on rejette H_0 avec une probabilité α de se tromper.

Test (unilatéral) de $H_0 : \mu = \mu_0$ contre $H_1 : \mu > \mu_0$.

On détermine t'_α tel que $P(T < t'_\alpha) = 1 - \alpha$, i.e. $t'_\alpha = t_{2\alpha}$, et on décide que :

- si $t < t'_\alpha$, alors on ne peut rejeter H_0 ;
- si $t \geq t'_\alpha$, alors on rejette H_0 avec une probabilité α de se tromper.

Test (unilatéral) de $H_0 : \mu = \mu_0$ contre $H_1 : \mu < \mu_0$.

On détermine t''_α tel que $P(T \geq t''_\alpha) = 1 - \alpha$, i.e. $t''_\alpha = t_{2-2\alpha} = -t_{2\alpha}$, et on décide que :

- si $t > t''_\alpha$, alors on ne peut rejeter H_0 ;
- si $t \leq t''_\alpha$, alors on rejette H_0 avec une probabilité α de se tromper.

4.1.2. Cas d'un grand échantillon : $n > 30$

On sait que $U = \frac{\bar{X} - \mu}{\frac{s_c}{\sqrt{n}}}$ suit approximativement la loi normale $\mathcal{N}(0; 1)$.

Test (bilatéral) de $H_0 : \mu = \mu_0$ contre $H_1 : \mu \neq \mu_0$.

On calcule $u = \frac{\bar{x} - \mu_0}{\frac{s_c}{\sqrt{n}}}$. On détermine u_α tel que $P(-u_\alpha < U < u_\alpha) = 1 - \alpha$, et on décide que :

- si $u \in]-u_\alpha, u_\alpha[$, alors on ne peut rejeter H_0 ;
- si $u \notin]-u_\alpha, u_\alpha[$, alors on rejette H_0 avec une probabilité α de se tromper.

Test (unilatéral) de $H_0 : \mu = \mu_0$ contre $H_1 : \mu > \mu_0$.

On détermine u'_α tel que $P(U < u'_\alpha) = 1 - \alpha$, i.e. $u'_\alpha = \phi^{-1}(1 - \alpha) = u_{2\alpha}$, et on décide que :

- si $u < u'_\alpha$, alors on ne peut rejeter H_0 ;
- si $u \geq u'_\alpha$, alors on rejette H_0 avec une probabilité α de se tromper.

Test (unilatéral) de $H_0 : \mu = \mu_0$ contre $H_1 : \mu < \mu_0$.

On détermine u''_α tel que $P(U \geq u''_\alpha) = 1 - \alpha$, i.e. $u''_\alpha = \phi^{-1}(\alpha) = u_{2-2\alpha} = -u_{2\alpha}$, et on décide que :

- si $u > u''_\alpha$, alors on ne peut rejeter H_0 ;
- si $u \leq u''_\alpha$, alors on rejette H_0 avec une probabilité α de se tromper.

4.1.3. Exemple

Dans une usine du secteur de l'agroalimentaire, une machine à embouteiller est alimentée par un réservoir d'eau et par une file d'approvisionnement en bouteilles vides. Pour contrôler le bon fonctionnement de la machine, on veut construire un test d'hypothèse bilatéral qui sera mis en oeuvre toutes les heures.

Pour une production d'une heure, on suppose que la variable aléatoire X qui à toute bouteille, prise au hasard dans cette production, associe le volume d'eau (en litres) qu'elle contient, est une variable aléatoire d'espérance μ et d'écart-type σ inconnus. On considère que la machine est bien réglée lorsque le volume d'eau moyen dans une bouteille est $\mu = 1,5 \text{ l}$.

On a prélevé un échantillon de 100 bouteilles, et on a obtenu un volume d'eau moyen de $1,495 \text{ l}$ et un écart-type corrigé de $0,01$. Peut-on conclure, au risque 5% , que la machine est bien réglée ?

On peut considérer la situation suivante.

Population : bouteilles produites

Variable X : volume d'eau, variable aléatoire de moyenne μ et d'écart-type $\sigma = 0,01$.

Echantillon $E = (X_1, X_2, \dots, X_n)$ de taille $n = 100$ de X .

Observation de l'échantillon : $e = (x_1, x_2, \dots, x_n)$.

Estimateurs $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ de μ et $S_c^2 = \frac{n}{n-1} S^2$ de σ^2 , avec $S^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2$.

Estimations ponctuelles : $\bar{x} = 1,495$ et $s_c \simeq 0,01$.

On a $n = 100 > 30$ donc un grand échantillon.

On effectue un test (bilatéral) de $H_0 : \mu = \mu_0$ contre $H_1 : \mu \neq \mu_0$.

On sait que $U = \frac{\bar{X} - \mu}{\frac{S_c}{\sqrt{n}}}$ suit approximativement la loi normale $\mathcal{N}(0; 1)$.

On calcule $u = \frac{\bar{x} - \mu_0}{\frac{s_c}{\sqrt{n}}} = \frac{1,495 - 1,5}{\frac{0,01}{\sqrt{100}}} = -5$.

On détermine u_α tel que $P(-u_\alpha < U < u_\alpha) = 1 - \alpha$ (table 2) : pour $\alpha = 0,05$, on trouve $u_\alpha = 1,96$.

Comme $u \notin]-u_\alpha, u_\alpha[$, on rejette H_0 avec une probabilité α de se tromper : la machine n'est pas bien réglée.

4.2. Pour une variance σ^2

Considérons un caractère quantitatif représenté par une variable aléatoire X de loi normale $\mathcal{N}(\mu; \sigma)$, et un échantillon (X_1, X_2, \dots, X_n) de taille n de X . La moyenne d'échantillon est \bar{X} et la variance corrigée d'échantillon est $S_c^2 = \frac{n}{n-1} S^2$.

Alors $Y^2 = \frac{n-1}{\sigma^2} S_c^2$ suit la loi de khi deux à $n-1$ degrés de liberté.

Test (bilatéral) de $H_0 : \sigma^2 = \sigma_0^2$ contre $H_1 : \sigma^2 \neq \sigma_0^2$.

On calcule $y^2 = \frac{n-1}{\sigma_0^2} s_c^2$. On détermine a_α et b_α tels que $P(Y^2 \geq a_\alpha) = 1 - \frac{\alpha}{2}$ et $P(Y^2 \geq b_\alpha) = \frac{\alpha}{2}$ (table 4). On a donc $P(Y^2 \in]a_\alpha, b_\alpha[) = P(a_\alpha < Y^2 < b_\alpha) = 1 - \alpha$ et $P(Y^2 \notin]a_\alpha, b_\alpha[) = \alpha$. On décide que :

- si $y^2 \in]a_\alpha, b_\alpha[$, alors on ne peut rejeter H_0 ;
- si $y^2 \notin]a_\alpha, b_\alpha[$, alors on rejette H_0 avec une probabilité α de se tromper.

Test (unilatéral) de $H_0 : \sigma^2 = \sigma_0^2$ contre $H_1 : \sigma^2 > \sigma_0^2$.

On détermine b'_α tel que $P(Y^2 \geq b'_\alpha) = \alpha$, i.e. $b'_\alpha = b_{2\alpha}$ et on décide que :

- si $y^2 < b'_\alpha$, alors on ne peut rejeter H_0 ;
- si $y^2 \geq b'_\alpha$, alors on rejette H_0 avec une probabilité α de se tromper.

Test (unilatéral) de $H_0 : \sigma^2 = \sigma_0^2$ contre $H_1 : \sigma^2 < \sigma_0^2$.

On détermine a''_α tel que $P(Y^2 \geq a''_\alpha) = 1 - \alpha$, i.e. $a''_\alpha = a_{2\alpha}$ et on décide que :

- si $y^2 > a''_\alpha$, alors on ne peut rejeter H_0 ;
- si $y^2 \leq a''_\alpha$, alors on rejette H_0 avec une probabilité α de se tromper.

5. Test d'homogénéité

Dans deux populations P_1 et P_2 , on étudie un même caractère. On cherche à comparer les deux populations quant à ce caractère, et donc à savoir si elles sont homogènes ou pas.

5.1. Comparaison de deux variances

Soient X_1 et X_2 des variables aléatoires représentant le caractère dans chaque population, de moyennes respectives μ_1 et μ_2 , d'écart-types respectifs σ_1 et σ_2 . De P_1 et P_2 on extrait un échantillon

$E_1 = (X_{1,1}, X_{1,2}, \dots, X_{1,n_1})$ de taille n_1 de X_1 et un échantillon $E_2 = (X_{2,1}, X_{2,2}, \dots, X_{2,n_2})$ de taille n_2 de X_2 .

Les moyennes d'échantillon sont alors $\bar{X}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1,i}$ et $\bar{X}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2,i}$, et les variances corrigées d'échantillon $S_{c,1}^2 = \frac{n_1}{n_1 - 1} S_1^2$ et $S_{c,2}^2 = \frac{n_2}{n_2 - 1} S_2^2$, avec $S_1^2 = \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1,i}^2 - \bar{X}_1^2$ et $S_2^2 = \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2,i}^2 - \bar{X}_2^2$.

5.1.1. Cas d'échantillons indépendants

Les échantillons E_1 et E_2 sont supposés indépendants. On suppose de plus que X_1 et X_2 suivent les lois normales $\mathcal{N}(\mu_1; \sigma_1)$ et $\mathcal{N}(\mu_2; \sigma_2)$.

Test de $H_0 : \sigma_1^2 = \sigma_2^2$ contre $H_1 : \sigma_1^2 \neq \sigma_2^2$.

Sous l'hypothèse H_0 , $F = \frac{S_{c,1}^2}{S_{c,2}^2}$ suit la loi de Snédécor à $(n_1 - 1, n_2 - 1)$ degrés de liberté.

On calcule $f = \frac{S_{c,1}^2}{S_{c,2}^2}$. Si nécessaire, on permute les échantillons de sorte que $f \geq 1$. On détermine f_α tel que

$P(F \geq f_\alpha) = \frac{\alpha}{2}$ (table 5 ou 6), et on décide que :

- si $f < f_\alpha$, alors on ne peut rejeter H_0 ;

- si $f \geq f_\alpha$, alors on rejette H_0 avec une probabilité α de se tromper.

5.1.2. Cas d'échantillons appariés

Deux échantillons E_1 et E_2 sont dits *appariés* lorsque chaque observation $x_{1,i}$ de E_1 est associée à une valeur $x_{2,i}$ de E_2 (appariés = associés par paires). C'est par exemple le cas lorsque E_1 et E_2 proviennent d'un même groupe de malades avant et après traitement. Deux échantillons appariés ont donc la même taille $n_1 = n_2 = n$.

On suppose que E_1 et E_2 sont appariés et que X_1 et X_2 suivent les lois normales $\mathcal{N}(\mu_1; \sigma_1)$ et $\mathcal{N}(\mu_2; \sigma_2)$.

Test de $H_0 : \sigma_1^2 = \sigma_2^2$ contre $H_1 : \sigma_1^2 \neq \sigma_2^2$.

Sous l'hypothèse H_0 , $T = \frac{(S_{c,1}^2 - S_{c,2}^2) \sqrt{n-2}}{2 \sqrt{S_{c,1}^2 S_{c,2}^2 - \frac{1}{(n-1)^2} \sum_{i=1}^n (X_{1,i} - \bar{X}_1)(X_{2,i} - \bar{X}_2)}}$ suit la loi de Student à $n - 2$

degrés de liberté. On calcule $t = \frac{(s_{c,1}^2 - s_{c,2}^2) \sqrt{n-2}}{2 \sqrt{s_{c,1}^2 s_{c,2}^2 - \frac{1}{(n-1)^2} \sum_{i=1}^n (x_{1,i} - \bar{x}_1)(x_{2,i} - \bar{x}_2)}}$. On détermine t_α tel que

$P(-t_\alpha < T < t_\alpha) = 1 - \alpha$ (table 3), et on décide que :

- si $t \in]-t_\alpha, t_\alpha[$, alors on ne peut rejeter H_0 ;

- si $t \notin]-t_\alpha, t_\alpha[$, alors on rejette H_0 avec une probabilité α de se tromper.

5.2. Comparaison de deux moyennes

5.2.1. Cas de grands échantillons indépendants

On suppose que $n_1 > 30$ et $n_2 > 30$, et que les échantillons E_1 et E_2 sont indépendants.

Test (bilatéral) de $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$.

Sous l'hypothèse H_0 , $U = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)}}$ suit approximativement la loi normale $\mathcal{N}(0; 1)$. Le résultat

reste valable si on remplace σ_1^2 et σ_2^2 par leurs estimations $s_{c,1}^2$ et $s_{c,2}^2$. On calcule $u = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{s_{c,1}^2}{n_1} + \frac{s_{c,2}^2}{n_2}\right)}}$. On

détermine u_α tel que $P(-u_\alpha < U < u_\alpha) = 1 - \alpha$, i.e. $u_\alpha = \phi^{-1}\left(1 - \frac{\alpha}{2}\right)$ (table 2) et on décide que :

- si $u \in]-u_\alpha, u_\alpha[$, alors on ne peut rejeter H_0 ;
- si $u \notin]-u_\alpha, u_\alpha[$, alors on rejette H_0 avec une probabilité α de se tromper.

Test (unilatéral) de $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 > \mu_2$.

On détermine u'_α tel que $P(U < u'_\alpha) = 1 - \alpha$, i.e. $u'_\alpha = \phi^{-1}(1 - \alpha) = u_{2\alpha}$, et on décide que :

- si $u < u'_\alpha$, alors on ne peut rejeter H_0 ;
- si $u \geq u'_\alpha$, alors on rejette H_0 avec une probabilité α de se tromper.

Test (unilatéral) de $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 < \mu_2$.

On détermine u''_α tel que $P(U \geq u''_\alpha) = 1 - \alpha$, i.e. $u''_\alpha = \phi^{-1}(\alpha) = u_{2-2\alpha} = -u_{2\alpha}$, et on décide que :

- si $u > u''_\alpha$, alors on ne peut rejeter H_0 ;
- si $u \leq u''_\alpha$, alors on rejette H_0 avec une probabilité α de se tromper.

5.2.2. Cas de petits échantillons indépendants extraits de populations gaussiennes

On suppose que $n_1 \leq 30$ ou $n_2 \leq 30$, et que les échantillons E_1 et E_2 sont indépendants. On suppose de plus que X_1 et X_2 suivent les lois normales $\mathcal{N}(\mu_1; \sigma_1)$ et $\mathcal{N}(\mu_2; \sigma_2)$, et que $\sigma_1 = \sigma_2 = \sigma$.

Test (bilatéral) de $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$.

Sous l'hypothèse H_0 , $T = \frac{\bar{X}_1 - \bar{X}_2}{\sigma \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ suit approximativement la loi de Student à $n_1 + n_2 - 2$

degrés de liberté. Comme on ne connaît pas $\sigma_1 = \sigma_2$, on doit d'abord tester l'égalité des variances $\sigma_1^2 = \sigma_2^2$ (paragraphe 2.1.1.). Si cette hypothèse est retenue, alors cette valeur commune σ^2 peut être estimée par

$$s_{c,1,2}^2 = \frac{(n_1 - 1)s_{c,1}^2 + (n_2 - 1)s_{c,2}^2}{n_1 + n_2 - 2}$$

On calcule $t = \frac{\bar{X}_1 - \bar{X}_2}{s_{c,1,2} \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$. On détermine t_α tel que $P(-t_\alpha < T < t_\alpha) = 1 - \alpha$ (table 3) et on

décide que

- si $t \in]-t_\alpha, t_\alpha[$, alors on ne peut rejeter H_0 ;
- si $t \notin]-t_\alpha, t_\alpha[$, alors on rejette H_0 avec une probabilité α de se tromper.

Test (unilatéral) de $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 > \mu_2$.

On détermine t'_α tel que $P(T < t'_\alpha) = 1 - \alpha$, i.e. $t'_\alpha = t_{2\alpha}$, et on décide que :

- si $t < t'_\alpha$, alors on ne peut rejeter H_0 ;
- si $t \geq t'_\alpha$, alors on rejette H_0 avec une probabilité α de se tromper.

Test (unilatéral) de $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 < \mu_2$.

On détermine t''_α tel que $P(T \geq t''_\alpha) = 1 - \alpha$, i.e. $t''_\alpha = t_{2-2\alpha} = -t_{2\alpha}$, et on décide que :

- si $t > t''_\alpha$, alors on ne peut rejeter H_0 ;
- si $t \leq t''_\alpha$, alors on rejette H_0 avec une probabilité α de se tromper.

5.2.3. Cas de petits échantillons indépendants : test de Mann et Whitney

Non traité ici.

5.2.4. Cas de grands échantillons appariés

On suppose que $n_1 = n_2 = n > 30$, et que les échantillons E_1 et E_2 sont appariés.

On considère la variable aléatoire $D = X_1 - X_2$, dont un échantillon est (D_1, D_2, \dots, D_n) , avec $D_i = X_{1,i} - X_{2,i}$. Les moyenne et variance corrigée d'échantillon sont alors $\bar{D} = \frac{1}{n} \sum_{i=1}^n D_i$ et $S_{c,d}^2 = \frac{n}{n-1} S_d^2$, avec $S_d^2 = \frac{1}{n} \sum_{i=1}^{n_1} D_i^2 - \bar{D}^2$. Désignons par $\mu = \mu_1 - \mu_2$ la moyenne de D .

Puisque $n > 30$, $U = \frac{\bar{D} - \mu}{\frac{S_{c,d}}{\sqrt{n}}}$ suit approximativement la loi normale $\mathcal{N}(0; 1)$.

Test (bilatéral) de $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$.

Ce test est équivalent au test (bilatéral) de $H_0 : \mu = 0$ contre $H_1 : \mu \neq 0$ (paragraphe 4.1.2.).

Test (unilatéral) de $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 > \mu_2$.

Ce test est équivalent au test (unilatéral) de $H_0 : \mu = 0$ contre $H_1 : \mu > 0$ (paragraphe 4.1.2.).

Test (unilatéral) de $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 < \mu_2$.

Ce test est équivalent au test (unilatéral) de $H_0 : \mu = 0$ contre $H_1 : \mu < 0$ (paragraphe 4.1.2.).

5.2.5. Cas de petits échantillons appariés extraits de populations gaussiennes

On suppose que $n_1 = n_2 = n \leq 30$, que les échantillons E_1 et E_2 sont appariés et que X_1 et X_2 suivent les lois normales $\mathcal{N}(\mu_1; \sigma_1^2)$ et $\mathcal{N}(\mu_2; \sigma_2^2)$.

Les notations sont les mêmes que dans le paragraphe 5.2.4. Dans ce cas, $T = \frac{\bar{D} - \mu}{\frac{S_{c,d}}{\sqrt{n}}}$ suit la loi de Student

à $n - 1$ degrés de liberté. On adapte alors les résultats ci-dessus (paragraphes 5.2.4. et 4.1.1.2.).

5.2.6. Cas de petits échantillons appariés : test de Wilcoxon

Non traité ici.

5.3. Exemples

5.3.1. Comparaison de deux moyennes (1)

Dans un article de la revue "Biometrika", le biologiste Latter donne la longueur (en mm) des oeufs de Coucou trouvés dans les nids de deux espèces d'oiseaux :

- dans des nids de petite taille (Roitelet) : $\begin{cases} 19,8 & 22,1 & 21,5 & 20,9 & 22,0 & 21,0 & 22,3 & 21,0 \\ 20,3 & 20,9 & 22,0 & 22,0 & 20,8 & 21,2 & 21,0 \end{cases}$
- dans des nids de taille plus grande (Fauvette) : $\begin{cases} 22,0 & 23,9 & 20,9 & 23,8 & 25,0 & 24,0 & 23,8 \\ 21,7 & 22,8 & 23,1 & 23,5 & 23,0 & 23,0 & 23,1 \end{cases}$

On se demande si le Coucou adapte la taille de ses oeufs à la taille du nid.

On peut considérer la situation suivante.

Population 1 : oeufs de Coucou dans des nids de Roitelet.

Variable X_1 : la longueur, variable aléatoire de moyenne μ_1 et de variance σ_1^2 .

Echantillon $E_1 = (X_{1,1}, X_{1,2}, \dots, X_{1,n_1})$ de taille $n_1 = 15$ de X_1 .

Observation de l'échantillon : $e_1 = (x_{1,1}, x_{1,2}, \dots, x_{1,n_1}) = (19,8, 22,1, \dots, 21,0)$.

Estimateurs $\bar{X}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1,i}$ de μ_1 et $S_{c,1}^2 = \frac{n_1}{n_1 - 1} S_1^2$ de σ_1^2 , avec $S_1^2 = \frac{1}{n_1} \sum_{i=1}^{n_1} X_{1,i}^2 - \bar{X}_1^2$.

Population 2 : oeufs de Coucou dans des nids de Fauvette.

Variable X_2 : la longueur, variable aléatoire de moyenne μ_2 et de variance σ_2^2 .

Echantillon $E_2 = (X_{2,1}, X_{2,2}, \dots, X_{2,n_2})$ de taille $n_2 = 14$ de X_2 .

Observation de l'échantillon : $e_2 = (x_{2,1}, x_{2,2}, \dots, x_{2,n_2}) = (22,0, 23,9, \dots, 23,1)$.

Estimateurs $\bar{X}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2,i}$ de μ_2 et $S_{c,2}^2 = \frac{n_2}{n_2 - 1} S_2^2$ de σ_2^2 , avec $S_2^2 = \frac{1}{n_2} \sum_{i=1}^{n_2} X_{2,i}^2 - \bar{X}_2^2$.

1) Estimations ponctuelles

- a) observation sur l'échantillon e_1 de taille $n_1 = 15$: $\bar{x}_1 \approx 21,25$ et $s_{c,1}^2 \approx 0,516$;
- b) observation sur l'échantillon e_2 de taille $n_2 = 14$: $\bar{x}_2 \approx 23,11$ et $s_{c,2}^2 \approx 1,101$.

2) Test de $H_0 : \sigma_1^2 = \sigma_2^2$ contre $H_1 : \sigma_1^2 \neq \sigma_2^2$.

Les échantillons E_1 et E_2 sont indépendants et on suppose que X_1 et X_2 suivent les lois normales $\mathcal{N}(\mu_1; \sigma_1)$ et $\mathcal{N}(\mu_2; \sigma_2)$.

Sous l'hypothèse H_0 , $F = \frac{S_{c,1}^2}{S_{c,2}^2}$ suit la loi de Snédécor à $(n_1 - 1, n_2 - 1)$ degrés de liberté.

Comme $f = \frac{s_{c,1}^2}{s_{c,2}^2} < 1$, on permute les échantillons.

Sous l'hypothèse H_0 , $F' = \frac{S_{c,2}^2}{S_{c,1}^2}$ suit la loi de Snédécor à $(n_2 - 1, n_1 - 1) = (13, 14)$ degrés de liberté.

On calcule $f' = \frac{s_{c,2}^2}{s_{c,1}^2} \approx 2,14$. On détermine f_α tel que $P(F' \geq f_\alpha) = \frac{\alpha}{2}$ (table 5 ou 6) : pour $\alpha = 0,05$,

on trouve f_α compris entre 2,95 et 3,15 (table 5).

Comme $f' < f_\alpha$, on ne peut rejeter H_0 et les variances des deux populations ne sont pas différentes significativement au risque 5%. Pour cette décision de non-rejet, on ne connaît pas la probabilité de se tromper (erreur de deuxième espèce).

3) Test (bilatéral) de $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 \neq \mu_2$.

On a $n_1 \leq 30$ ou $n_2 \leq 30$, et les échantillons E_1 et E_2 sont indépendants. On suppose que X_1 et X_2 suivent les lois normales $\mathcal{N}(\mu_1; \sigma_1)$ et $\mathcal{N}(\mu_2; \sigma_2)$. On est alors dans le cas de petits échantillons gaussiens indépendants. D'après le test précédent, on peut admettre $\sigma_1 = \sigma_2 = \sigma$.

Sous l'hypothèse H_0 , $T = \frac{\bar{X}_1 - \bar{X}_2}{\sigma \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$ suit approximativement la loi de Student à

$n_1 + n_2 - 2 = 27$ degrés de liberté. Comme on a retenu $\sigma_1^2 = \sigma_2^2$, cette valeur commune σ^2 peut être estimée par $s_{c,1,2}^2 = \frac{(n_1 - 1)s_{c,1}^2 + (n_2 - 1)s_{c,2}^2}{n_1 + n_2 - 2} \approx 0,798$, et en remplaçant σ par $s_{c,1,2}^2$ dans T , on ne modifie pas la loi approchée de T .

On calcule alors $t = \frac{\bar{x}_1 - \bar{x}_2}{s_{c,1,2} \sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \approx -5,61$.

On détermine t_α tel que $P(-t_\alpha < T < t_\alpha) = 1 - \alpha$ (table 3) : pour $\alpha = 0,05$, on trouve $t_\alpha = 2,052$.

Comme $t \notin]-t_\alpha, t_\alpha[$, on rejette H_0 avec une probabilité $\alpha = 0,05$ de se tromper. La taille moyenne des oeufs de Coucou sont différentes dans les nids de Roitelet et de Fauvettes.

Comme on observe $\bar{x}_1 < \bar{x}_2$, on aurait pu faire le test unilatéral de $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 < \mu_2$.

On détermine t''_α tel que $P(U \geq t''_\alpha) = 1 - \alpha$, i.e. $t''_\alpha = t_{2-\alpha} = -t_{2\alpha}$: pour $\alpha = 0,05$, on trouve $t''_\alpha = -1,703$.

Comme $t \leq t''_\alpha$, on rejette H_0 avec une probabilité α de se tromper. La taille moyenne des oeufs de Coucou dans les nids de Roitelet est inférieure à celle dans les nids de Fauvettes.

Ainsi, on peut conclure que le Coucou adapte la grosseur de ses oeufs à la taille du nid. (Il s'agit d'un phénomène de mimétisme qui permet aux oeufs de Coucou de passer plus facilement inaperçus.)

5.3.2. Comparaison de deux moyennes (2)

Chez un groupe de 10 malades, on expérimente les effets d'un traitement destiné à diminuer la pression artérielle. On observe les résultats suivants (valeur de la tension artérielle systolique en cm Hg) :

sujet n°	1	2	3	4	5	6	7	8	9	10
avant traitement	15	18	17	20	21	18	17	15	19	16
après traitement	12	16	17	18	17	15	18	14	16	18

On se demande si le traitement a une action significative. On peut considérer la situation suivante.

Population 1 : malades avant traitement.

Variable X_1 : la tension, variable aléatoire de moyenne μ_1 et de variance σ_1^2 .

Echantillon $E_1 = (X_{1,1}, X_{1,2}, \dots, X_{1,n_1})$ de taille $n_1 = 10$ de X_1 .

Observation de l'échantillon : $e_1 = (x_{1,1}, x_{1,2}, \dots, x_{1,n_1}) = (15, 18, \dots, 16)$.

Population 2 : malades après traitement.

Variable X_2 : la tension, variable aléatoire de moyenne μ_2 et de variance σ_2^2 .

Echantillon $E_2 = (X_{2,1}, X_{2,2}, \dots, X_{2,n_2})$ de taille $n_2 = 10$ de X_2 .

Observation de l'échantillon : $e_2 = (x_{2,1}, x_{2,2}, \dots, x_{2,n_2}) = (12, 16, \dots, 18)$.

On a $n_1 = n_2 = n = 10 \leq 30$ et les échantillons E_1 et E_2 sont appariés. On suppose que X_1 et X_2 suivent les lois normales $\mathcal{N}(\mu_1; \sigma_1)$ et $\mathcal{N}(\mu_2; \sigma_2)$. On a donc de petits échantillons appariés extraits de populations gaussiennes.

On considère la variable aléatoire $D = X_1 - X_2$, dont un échantillon est (D_1, D_2, \dots, D_n) , avec

$D_i = X_{1,i} - X_{2,i}$. Les moyenne et variance corrigée d'échantillon sont alors $\bar{D} = \frac{1}{n} \sum_{i=1}^n D_i$ et $S_{c,d}^2 = \frac{n}{n-1} S_d^2$,

avec $S_d^2 = \frac{1}{n} \sum_{i=1}^{n_1} D_i^2 - \bar{D}^2$. Désignons par $\mu = \mu_1 - \mu_2$ la moyenne de D .

A partir de l'observation de l'échantillon $(d_1, d_2, \dots, d_n) = (3, 2, 0, 2, 4, 3, -1, 1, 3, -2)$, on obtient les estimations $\bar{d} = 1,5$ et $s_{c,d}^2 \simeq 1,96$.

Test (unilatéral) de $H_0 : \mu_1 = \mu_2$ contre $H_1 : \mu_1 > \mu_2$.

Ce test est équivalent au test (unilatéral) de $H_0 : \mu = 0$ contre $H_1 : \mu > 0$ (test de conformité).

Sous l'hypothèse H_0 , on sait que $T = \frac{\bar{D}}{S_{c,d}} \sqrt{n}$ suit la loi de Student à $n-1$ degrés de liberté.

On calcule $t = \frac{\bar{d}}{S_{c,d}} \sqrt{n} \simeq 2,42$.

On détermine t'_α tel que $P(T < t'_\alpha) = 1 - \alpha$, i.e. $t'_\alpha = t_{2\alpha}$ (table 3) : pour $\alpha = 0,05$, on trouve $t'_\alpha = 1,833$.

Comme $t \geq t'_\alpha$, alors on rejette H_0 avec une probabilité α de se tromper. On conclut que la tension a diminué après le traitement et donc que ce dernier a une action significative.

6. Exercices

Exercice 1. Une usine fabrique des pièces métalliques. Le client réceptionne sa commande. Dans le lot reçu, il préleve un échantillon de 20 billes choisies au hasard et avec remise, et mesure les diamètres suivants :

24,7	24,9	25,0	25,0	25,1	25,1	25,1	25,2	25,3	25,4
24,8	24,9	25,0	25,0	25,1	25,1	25,2	25,3	25,3	25,5

- 1) Préciser la population et le caractère étudiés.
- 2) Expliquer pourquoi on peut considérer que ce caractère est une variable aléatoire. Préciser en particulier l'expérience aléatoire et l'espace probabilisé permettant cette modélisation.
- 3) Préciser la taille d'échantillon, le(s) estimateur(s) mis en jeu et leur loi.
- 4) Donner une estimation ponctuelle de la moyenne et de la variance du diamètre.

Exercice 2. On admet que le taux de cholestérol chez une femme suit une loi normale $\mathcal{N}(\mu; \sigma)$. Sur un échantillon de 10 femmes, on a obtenu les taux de cholestérol (en g/l) suivants :

3,0	1,8	2,1	2,7	1,4	1,9	2,2	2,5	1,7	2,0
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

- 1) Déterminer une estimation ponctuelle de la moyenne et de l'écart-type du taux.
- 2) Déterminer un intervalle de confiance pour la moyenne du taux au seuil 1%.
- 3) Déterminer un intervalle de confiance pour l'écart-type du taux au seuil 5%.

Exercice 3. Dans la fabrication de comprimés effervescents, il est prévu que chaque comprimé doit contenir 1625 mg de bicarbonate de sodium. Afin de contrôler la fabrication de ces médicaments, on a prélevé un échantillon de 150 comprimés, et on a mesuré la quantité de bicarbonate de sodium pour chacun d'eux. On a obtenu les résultats suivants :

Classes	[1610; 1615]]1615; 1620]]1620; 1625]]1625; 1630]]1630; 1635]
Effectifs	7	8	42	75	18

- 1) Déterminer une estimation ponctuelle de la moyenne et de l'écart-type de la quantité de bicarbonate de sodium.
- 2) Déterminer un intervalle de confiance au seuil 5% de la moyenne de la quantité de bicarbonate de sodium.
- 3) Quelle devrait-être la taille n de l'échantillon pour connaître la quantité moyenne de bicarbonate de sodium à 1 mg près ?

Exercice 4. On a mesuré, avant et après une course de 400 mètres, le pouls (en battements par minute) de 7 étudiants suivants un cours d'éducation physique :

Avant	74	87	77	99	103	81	60
Après	83	96	99	110	130	95	74

On suppose que l'accroissement du pouls est une variable aléatoire de loi normale $\mathcal{N}(\mu; \sigma)$.

Déterminer un intervalle de confiance au niveau 95 % pour la moyenne de l'accroissement du pouls.

Exercice 5. Le temps (exprimé en minutes) mis par une machine A pour fabriquer une pièce suit une loi Normale $\mathcal{N}(48,5)$. La machine A tombant en panne, on fabrique la même pièce avec une machine B. On suppose que le temps de fabrication suit encore une loi Normale de même écart-type. Pour un échantillon de 25 pièces, on a obtenu un temps moyen de fabrication de 51 min.

La machine B a-t-elle les mêmes performances que la machine A ?

Exercice 6. On suppose que chez les femmes non malades, la teneur en hémoglobine du sang (en g pour 100 mL) est une variable aléatoire de loi normale de moyenne 14,5 et d'écart-type 1,1. Sur un échantillon de 20 femmes, on trouve une teneur moyenne en hémoglobine de 13,8 et un écart-type corrigé de 1,2.

Au risque de 5%, peut-on conclure que la population de femmes dont est extrait cet échantillon présente une teneur en hémoglobine normale ? trop faible ?

Exercice 7. Le volume d'une pipette d'un type donné suit une loi normale $\mathcal{N}(\mu; \sigma)$. Le fabricant annonce un écart-type $\sigma = 0,2\mu\text{l}$. Pour le vérifier, on pipette 20 fois un liquide. On observe une moyenne de $10\mu\text{l}$ et un écart-type de $0,4\mu\text{l}$.

Tester l'affirmation du fabricant.

Exercice 8. Des dosages de l'acide aspartique total de l'urine, en mg/24h, ont été effectués sur deux groupes d'adultes à régime alimentaire normal. Les dosages ont donné les résultats suivants :

	Hommes	Femmes
Effectif du groupe	53	48
Moyenne	91,13	112,9
Ecart-type	30,49	48,99

1) Préciser la(les) population(s) et le(s) caractère(s) étudié(s), ainsi que la(les) taille(s) d'échantillon. Indiquer le(s) estimation(s) mises en jeu dans la suite.

2) Peut-on considérer, au risque 5%, que les deux populations étudiées ont le même dosage d'acide aspartique ? Justifier votre réponse.

3) Si les deux groupes avaient les mêmes moyennes et écart-types que dans le tableau ci-dessus, mais n'étaient constitués que de 12 hommes et 10 femmes, chacun, pourrait-on encore appliquer la méthode suivie au 2) ? En cas de réponse négative, indiquer la démarche à suivre.

Exercice 9. Sur un groupe de 10 malades, on expérimente les effets d'un traitement destiné à diminuer la pression artérielle. On observe les résultats suivants (valeurs de la tension artérielle systolique en cm Hg) :

Sujet n°	1	2	3	4	5	6	7	8	9	10
avant traitement	15	18	17	20	21	18	17	15	19	16
après traitement	12	16	17	18	17	15	18	14	16	18

Le traitement a-t-il une action significative, au risque de 5% ? On supposera les populations gaussiennes.

Exercice 10. La durée de gestation humaine est en moyenne de 40,5 semaines.

1) Dans une maternité, on a noté l'âge gestationnel de 100 nouveaux-nés successifs. On a observé une moyenne de 38,5 semaines et un écart-type de 5 semaines. On pense que cette maternité est spécialisée dans les accouchements prématurés.

Tester cette hypothèse au risque 5%.

2) Dans cette même maternité, les mères des 100 nouveaux-nés suivants ont reçu un traitement inhibant les contractions utérines. Pour ces nouveaux-nés, on a observé une moyenne de 39,5 semaines et un écart-type de 4 semaines.

Tester l'égalité des moyennes des durées de gestation des 2 groupes au risque 2%.

Exercice 11. D'après examen de mars 2011

1) Un fabricant de téléviseurs achète un certain composant électronique à un fournisseur A. L'accord entre le fabricant et le fournisseur stipule que les composants doivent avoir une durée de vie au moins égale à 600 heures.

Le fabricant reçoit un lot important de composants et (ayant des doutes) décide de tester la durée de vie (en heures) sur un échantillon de 16 composants choisis au hasard dans le lot. Il obtient les résultats suivants :

565	620	570	525	605	590	590	560
590	575	625	560	550	625	570	515

a) Préciser la population et le caractère étudié. Préciser la taille d'échantillon, le(s) estimateur(s) mis en jeu et leur loi. Préciser les hypothèses éventuelles à faire sur la variable étudiée pour connaître la loi des estimateurs.

b) Donner une estimation ponctuelle de la moyenne et de l'écart-type de la durée de vie d'un composant.

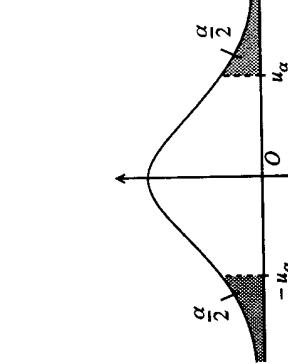
c) Donner un intervalle de confiance au niveau 95% de la durée de vie moyenne d'un composant.

Peut-on en déduire que l'accord entre le fabricant et le fournisseur n'est pas respecté ? Expliquer.

d) Effectuer un test statistique au risque 5% pour savoir si l'accord est respecté ou pas.

2) Le fabricant a réalisé un autre test de durée de vie sur un échantillon de 16 composants prélevés au hasard dans un lot d'un autre fournisseur B annonçant une durée de vie plus longue que celle de A. Les résultats obtenus sur la durée de vie de ces 16 composants donnent une moyenne de 600 h et un écart-type corrigé de 30 h.

Effectuer le(s) test(s) statistique(s) adéquat(s) pour savoir si on peut considérer, au risque 5%, que les composants du fournisseur B ont une durée de vie plus longue que ceux du fournisseur A.



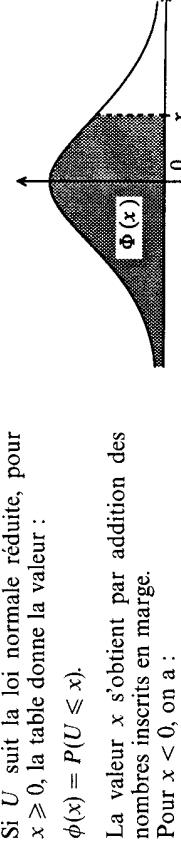
Si U est une variable aléatoire qui suit la loi normale réduite, la table donne, pour α choisi, la valeur u_α telle que :

$$P(|U| \geq u_\alpha) = \alpha.$$

La valeur α s'obtient par addition des nombres inscrits en marge.

α	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0	∞	2,576	2,326	2,170	2,054	1,960	1,881	1,812	1,751
0,1	0,1	1,645	1,598	1,555	1,514	1,476	1,440	1,405	1,372	1,341
0,2	0,2	1,282	1,254	1,227	1,200	1,175	1,150	1,126	1,103	1,080
0,3	0,3	1,036	1,015	0,994	0,974	0,954	0,935	0,915	0,896	0,878
0,4	0,4	0,842	0,824	0,806	0,789	0,772	0,755	0,739	0,722	0,706
0,5	0,5	0,674	0,659	0,643	0,628	0,613	0,598	0,583	0,568	0,553
0,6	0,6	0,524	0,510	0,496	0,482	0,468	0,454	0,440	0,426	0,412
0,7	0,7	0,385	0,372	0,358	0,345	0,332	0,319	0,305	0,292	0,279
0,8	0,8	0,253	0,240	0,228	0,215	0,202	0,189	0,176	0,164	0,151
0,9	0,9	0,126	0,113	0,100	0,088	0,075	0,063	0,050	0,038	0,025

TABLE 2
Loi normale réduite
(table de l'écart réduit)



Si U suit la loi normale réduite, pour $x \geq 0$, la table donne la valeur :

$$\phi(x) = P(U \leq x).$$

La valeur x s'obtient par addition des nombres inscrits en marge.

Pour $x < 0$, on a :

$$\phi(x) = 1 - \phi(-x).$$

x	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7390	0,7424	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9982	0,9983	0,9984	0,9985	0,9985	0,9986	0,9986

TABLE 1
Fonction de répartition
de la loi normale réduite

Si U suit la loi normale réduite, pour $x \geq 0$, la table donne la valeur :

$$\phi(x) = P(U \leq x).$$

La valeur x s'obtient par addition des nombres inscrits en marge.

Pour $x < 0$, on a :

$$\phi(x) = 1 - \phi(-x).$$

TABLE 3
Lois de Student

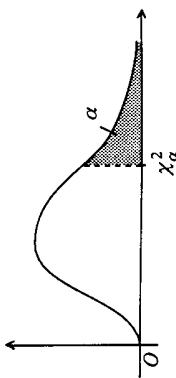
Si T est une variable aléatoire qui suit la loi de Student à v degrés de liberté, la table donne, pour α choisi, le nombre t_α tel que $P(|T| \geq t_\alpha) = \alpha$.

α	0,90	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001
v									
1	0,158	1,000	1,963	3,078	6,314	12,706	31,82	63,657	636,619
2	0,142	0,816	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,765	1,250	1,638	2,353	3,182	4,541	5,841	12,924
4	0,134	0,741	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,727	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,718	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,711	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,706	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,703	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,700	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,697	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,695	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,694	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,692	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,691	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,690	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,689	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,688	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,688	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,687	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,686	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,686	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,685	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,685	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,684	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,684	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,684	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,683	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,683	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,683	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,681	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,126	0,679	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,677	1,041	1,289	1,658	1,980	2,358	2,617	3,373

TABLE 4
Lois de Pearson
ou lois du χ^2

Si Y^2 est une variable aléatoire qui suit la loi du Y^2 à v degrés de liberté, la table donne, pour α choisi, le nombre χ_α^2 tel que $P(Y^2 \geq \chi_\alpha^2) = \alpha$.

α	0,99	0,99	0,975	0,95	0,90	0,10	0,05	0,025	0,01	0,001
v										
1	0,0002	0,001	0,004	0,016	0,04	0,21	0,78	2,71	3,84	6,63
2	0,002	0,05	0,10	0,21	0,58	4,61	5,99	7,38	9,21	10,83
3	0,12	0,22	0,35	0,48	0,71	1,06	7,78	9,49	11,14	13,82
4	0,30	0,55	0,83	1,15	1,61	2,20	10,64	12,59	15,09	16,27
5	0,87	1,24	1,64	2,20	2,83	12,02	14,07	16,01	18,47	20,52
6	1,24	1,69	2,17	2,73	3,49	13,36	15,51	17,53	20,09	22,46
7	1,24	1,69	2,17	2,73	3,47	14,88	16,92	19,02	21,67	24,32
8	1,24	1,69	2,17	2,73	3,47	15,99	18,31	20,48	23,21	29,59



Lorsque le degré de liberté v est tel que $v > 30$, la variable aléatoire :

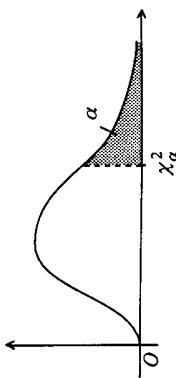
$$U = \sqrt{2Y^2 - \sqrt{2v-1}}$$

suit à peu près la loi normale réduite.

TABLE 3
Lois de Student

Si T est une variable aléatoire qui suit la loi de Student à v degrés de liberté, la table donne, pour α choisi, le nombre t_α tel que $P(|T| \geq t_\alpha) = \alpha$.

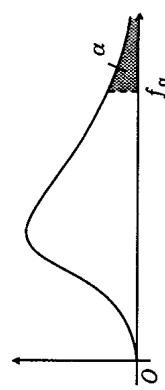
α	0,90	0,50	0,30	0,20	0,10	0,05	0,02	0,01	0,001
v									
1	0,158	1,000	1,963	3,078	6,314	12,706	31,82	63,657	636,619
2	0,142	0,816	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,765	1,250	1,638	2,353	3,182	4,541	5,841	12,924
4	0,134	0,741	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,727	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,718	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,711	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,706	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,703	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,700	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,697	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,695	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,694	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,692	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,691	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,690	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,689	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,688	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,688	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,687	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,686	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,686	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,685	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,685	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,684	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,684	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,684	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,683	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,683	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,683	1,055	1,310	1,697	2,042	2,457	2,750	3,646
40	0,126	0,681	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,126	0,679	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,677	1,041	1,289	1,658	1,980	2,358	2,617	3,373



Lorsque le degré de liberté est infini, il s'agit du nombre t_α correspondant à la loi normale centrée réduite (cf. table 2).

TABLE 5
Lois de Snédécor ($\alpha = 0,025$)

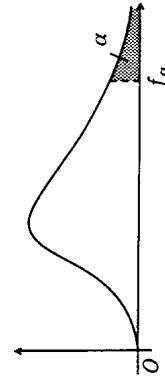
Si F est une variable aléatoire qui suit la loi de Snédécor à (v_1, v_2) degrés de liberté, la table donne le nombre f_α tel que $P(F \geq f_\alpha) = \alpha = 0,025$.



$v_1 \backslash v_2$	1	2	3	4	5	6	8	10	15	20	30	∞
1	648	800	864	900	922	937	957	969	985	993	1001	1018
2	38,5	39,0	39,2	39,3	39,4	39,5	39,4	39,5	39,5	39,5	39,5	39,5
3	17,4	16,0	15,4	15,1	14,9	14,7	14,5	14,4	14,3	14,2	14,1	13,9
4	12,2	10,6	9,98	9,60	9,36	9,20	8,98	8,84	8,66	8,56	8,46	8,26
5	10,0	8,43	7,76	7,39	7,15	6,98	6,76	6,62	6,43	6,33	6,23	6,02
6	8,81	7,26	6,60	6,23	5,99	5,82	5,60	5,46	5,27	5,17	5,07	4,85
7	8,07	6,54	5,89	5,52	5,29	5,12	4,90	4,76	4,57	4,47	4,36	4,14
8	7,57	6,06	5,42	5,05	4,82	4,65	4,43	4,30	4,10	4,00	3,89	3,67
9	7,21	5,71	5,08	4,72	4,48	4,32	4,10	3,96	3,77	3,67	3,56	3,33
10	6,94	5,46	4,83	4,47	4,24	4,07	3,85	3,72	3,52	3,42	3,31	3,08
11	6,72	5,26	4,63	4,28	4,04	3,88	3,66	3,53	3,33	3,23	3,12	2,88
12	6,55	5,10	4,47	4,12	3,89	3,73	3,51	3,37	3,18	3,07	2,96	2,72
13	6,41	4,97	4,35	4,00	3,77	3,60	3,39	3,25	3,05	2,95	2,84	2,60
14	6,30	4,86	4,24	3,89	3,66	3,50	3,29	3,15	2,95	2,84	2,73	2,49
15	6,20	4,76	4,15	3,80	3,58	3,41	3,20	3,06	2,86	2,76	2,64	2,40
16	6,12	4,69	4,08	3,73	3,50	3,34	3,12	2,99	2,79	2,68	2,57	2,32
17	6,04	4,62	4,01	3,66	3,44	3,28	3,06	2,92	2,72	2,62	2,50	2,25
18	5,98	4,56	3,95	3,61	3,38	3,22	3,01	2,87	2,67	2,56	2,44	2,19
19	5,92	4,51	3,90	3,56	3,33	3,17	2,96	2,82	2,62	2,51	2,39	2,13
20	5,87	4,46	3,86	3,51	3,29	3,13	2,91	2,77	2,57	2,46	2,35	2,09
22	5,79	4,38	3,78	3,44	3,22	3,05	2,84	2,70	2,50	2,39	2,27	2,00
24	5,72	4,32	3,72	3,38	3,15	2,99	2,78	2,64	2,44	2,33	2,21	1,94
26	5,66	4,27	3,67	3,33	3,10	2,94	2,73	2,59	2,39	2,28	2,16	1,88
28	5,61	4,22	3,63	3,29	3,06	2,90	2,69	2,55	2,34	2,23	2,11	1,84
30	5,57	4,18	3,59	3,25	3,03	2,87	2,65	2,51	2,31	2,20	2,07	1,79
40	5,42	4,05	3,46	3,13	2,90	2,74	2,53	2,39	2,18	2,07	1,94	1,64
50	5,34	3,98	3,39	3,06	2,83	2,67	2,46	2,32	2,11	1,99	1,87	1,55
60	5,29	3,93	3,34	3,01	2,79	2,63	2,41	2,27	2,06	1,94	1,82	1,48
80	5,22	3,86	3,28	2,95	2,73	2,57	2,36	2,21	2,00	1,88	1,75	1,40
100	5,18	3,83	3,25	2,92	2,70	2,54	2,32	2,18	1,97	1,85	1,71	1,35
∞	5,02	3,69	3,12	2,79	2,57	2,41	2,19	2,05	1,83	1,71	1,57	1,00

TABLE 6
Lois de Snédécor ($\alpha = 0,05$)

Si F est une variable aléatoire qui suit la loi de Snédécor à (v_1, v_2) degrés de liberté, la table donne le nombre f_α tel que $P(F \geq f_\alpha) = \alpha = 0,05$.



$v_1 \backslash v_2$	1	2	3	4	5	6	8	10	15	20	30	∞
1	161	200	216	225	230	234	239	242	246	248	250	254
2	18,5	19,0	19,2	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,85	8,79	8,70	8,66	8,62	8,53
4	7,71	6,94	6,59	6,39	6,16	6,04	5,96	5,86	5,80	5,75	5,63	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,74	4,62	4,56	4,50	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,06	3,94	3,87	3,81	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,64	3,51	3,44	3,38	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,35	3,22	3,15	3,08	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,14	3,01	2,94	2,86	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,98	2,85	2,77	2,70	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	2,95	2,85	2,72	2,65	2,57	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,85	2,75	2,62	2,54	2,47	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,77	2,67	2,53	2,46	2,38	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,60	2,46	2,39	2,31	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,64	2,54	2,40	2,33	2,25	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,49	2,35	2,28	2,19	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,55	2,45	2,31	2,23	2,15	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,41	2,27	2,19	2,11	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,48	2,38	2,23	2,16	2,07	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,35	2,20	2,12	2,04	1,84