MESURES ET ANALYSES STATISTIQUES DE DONNÉES

Probabilités

Master Génie des Systèmes Industriels, mentions ACCIE et RIM

Université du Littoral - Côte d'Opale, La Citadelle

Laurent SMOCH

(smoch@lmpa.univ-littoral.fr)

Septembre 2012

Laboratoire de Mathématiques Pures et Appliquées Joseph Liouville
Université du Littoral, zone universitaire de la Mi-Voix, bâtiment H. Poincarré
50, rue F. Buisson, BP 699, F-62228 Calais cedex

Table des matières

Chapitre 5

Lois de probabilités continues usuelles

5.1 Loi et variable uniformes

5.1.1 Définition

On dit que la loi de probabilité d'une variable aléatoire réelle est **uniforme** sur un segment $[a;b]$, avec $0 \leq a < b$, si sa densité de probabilité f est définie par

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{pour } x \in [a;b] \\ 0 & \text{pour } x < a \text{ ou } x > b \end{cases}$$

On note alors $X \rightsquigarrow \mathcal{U}([a;b])$. f admet la représentation graphique de la Figure 5.1.

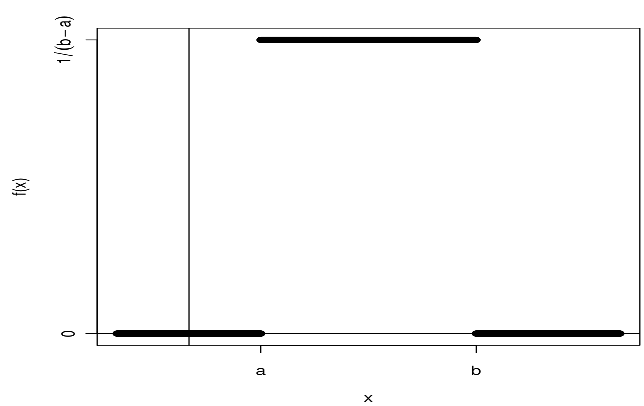

FIGURE 5.1

On a bien une densité de probabilité puisque

- $f(x) \geq 0 \ \forall x \in \mathbb{R}$,
- f est continue sur $]-\infty; a[\cup]a; b[\cup]b; +\infty[$,
- $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = \int_{-\infty}^{a} f(t)dt + \int_{a}^{b} f(t)dt + \int_{b}^{+\infty} f(t)dt = 0 + 1 + 0 = 1.$

5.1.2 Fonction de répartition

1. On sait que $F(x) = p(\{X \leq x\}) = \displaystyle\int_{-\infty}^{x} f(t)dt$ donc si $X \rightsquigarrow \mathcal{U}([a;b])$,

$$F(x) = \begin{cases} 0 & \text{pour } x < a \\ \dfrac{x-a}{b-a} & \text{pour } a \leq x \leq b \\ 1 & \text{pour } x > b \end{cases}$$

<u>Preuve</u> : On distingue trois cas :
- si $x < a$, $F(x) = \displaystyle\int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} 0dt = 0$,
- si $a \leq x \leq b$, $F(x) = \displaystyle\int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{a} 0dt + \int_{a}^{x} \frac{1}{b-a}dt = \frac{x-a}{b-a}$,
- si $x > b$, $F(x) = \displaystyle\int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{a} 0dt + \int_{a}^{b} \frac{1}{b-a}dt + \int_{b}^{x} 0dt = 0 + 1 + 0 = 1$.

■

2. Représentation graphique

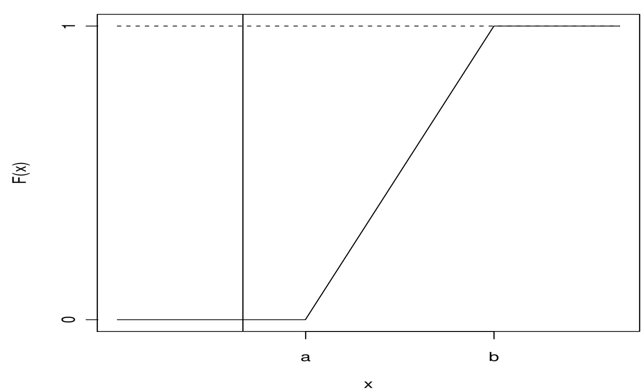

FIGURE 5.2

5.1.3 Moments

Si $X \rightsquigarrow \mathcal{U}([a;b])$ alors
- $E(X) = \displaystyle\int_{-\infty}^{+\infty} tf(t)dt = \int_{a}^{b} \frac{t}{b-a}dt = \left[\frac{t^2}{2(b-a)}\right]_a^b = \frac{a+b}{2}$.
- $V(X) = \dfrac{a^2+ab+b^2}{3} - \left(\dfrac{a+b}{2}\right)^2 = \dfrac{(b-a)^2}{12}$.

  En effet, on a $E(X^2) = \displaystyle\int_{a}^{b} \frac{t^2}{b-a}dt = \left[\frac{t^3}{3(b-a)}\right]_a^b = \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}$.

Exemple 5.1.1 On considère la fonction f définie par :

$$f(x) = \begin{cases} 0 & \text{pour } x < 0 \text{ ou } x > 1 \\ 1 & \text{pour } 0 \leq x \leq 1 \end{cases}.$$

Il apparaît en intégrant que

$$F(x) = \begin{cases} 0 & \text{pour } x < 0 \\ x & \text{pour } 0 \leq x \leq 1 \\ 1 & \text{pour } x > 1 \end{cases}$$

On trouve $E(X) = \dfrac{1}{2}$ et $V(X) = \dfrac{1}{12}$.

Remarque 5.1.1 Dans l'exemple précédent, comme dans toute variable aléatoire absolument continue, on a $p(\{X = 0\}) = 0$. En effet, $\forall a \in \mathbb{R}$, $\displaystyle\int_{a}^{a} f(t)dt = 0$.

5.2 La loi exponentielle

5.2.1 Définition

Une variable aléatoire X suit une loi **exponentielle** de paramètre λ ($\lambda \in \mathbb{R}^{+\star}$) si X est une variable aléatoire absolument continue dont la densité de probabilité est définie par

$$f(x) = \begin{cases} 0 & \text{pour } x < 0 \\ \lambda e^{-\lambda x} & \text{pour } x \geq 0 \end{cases}$$

On note alors $X \rightsquigarrow \mathcal{E}(\lambda)$. La fonction f admet la représentation graphique de la Figure 5.3.

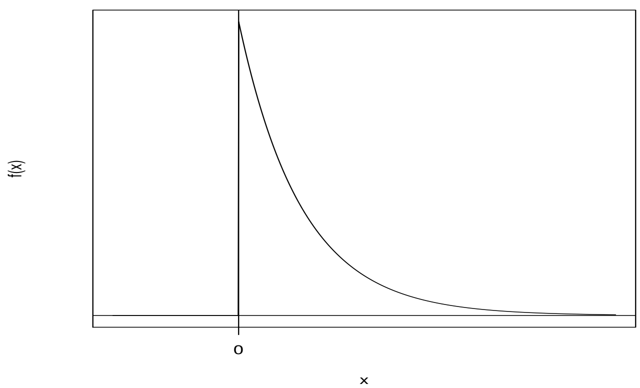

FIGURE 5.3

On a bien une densité de probabilité puisque
- $f(x) \geq 0 \ \forall x \in \mathbb{R}$,
- f est continue sur $\mathbb{R}^+$ et $\mathbb{R}^{-\star}$,
- $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = \int_{-\infty}^{0} 0dt + \int_{0}^{+\infty} \lambda e^{-\lambda t}dt$. Or, si $A > 0$, $\displaystyle\int_0^A \lambda e^{-\lambda t}dt = [-e^{\lambda t}]_0^A = 1 - e^{-\lambda A} \to 1$ quand $A \to +\infty$ donc $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = 1$.

La loi exponentielle peut être considérée comme l'équivalent en continu de la loi géométrique dans le cas discret. En effet, elle modélise un temps d'attente du premier succès dans un processus de Poisson.

5.2.2 Fonction de répartition

Si $X \rightsquigarrow \mathcal{E}(\lambda)$, on a

$$F(x) = \begin{cases} 0 & \text{pour } x < 0 \\ 1 - \lambda e^{-\lambda x} & \text{pour } x \geq 0 \end{cases}$$

dont la représentation graphique est donnée à la Figure 5.4.

5.2.3 Les moments

Si $X \rightsquigarrow \mathcal{E}(\lambda)$,

1. $E(X) = \displaystyle\int_{-\infty}^{+\infty} tf(t)dt = \int_0^{+\infty} \lambda te^{-\lambda t}dt = \frac{1}{\lambda}$ à l'aide d'une intégration par parties.
2. $V(X) = \dfrac{1}{\lambda^2}$ et $\sigma(X) = \dfrac{1}{\lambda}$.

   En effet, on sait que $V(X) = E(X^2) - (E(X))^2$. Comme $E(X^2) = \displaystyle\int_0^{+\infty} \lambda t^2 e^{-\lambda t}dt$, à l'aide de deux intégrations par parties, on obtient $E(X^2) = \dfrac{2}{\lambda^2}$.

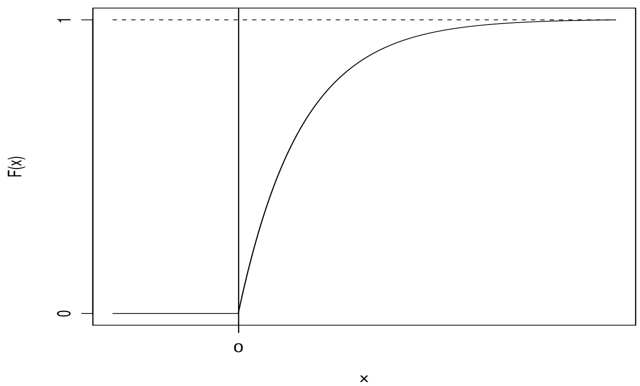

FIGURE 5.4

5.3 La loi de Laplace-Gauss ou loi normale

5.3.1 Définition

On appelle **variable aléatoire normale** ou **gaussienne** toute variable aléatoire absolument continue dont la densité de probabilité f est définie par

$$\boxed{f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}}$$

m étant une constante réelle, σ une constante réelle strictement positive. On utilise la notation suivante

$$\boxed{X \rightsquigarrow \mathcal{N}(m, \sigma)}$$

Remarque 5.3.1 On admettra que f est bien une densité de probabilité (la difficulté étant de montrer que $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = 1$).

5.3.2 Représentation graphique

La courbe représentative de f est donnée par la Figure 5.5.

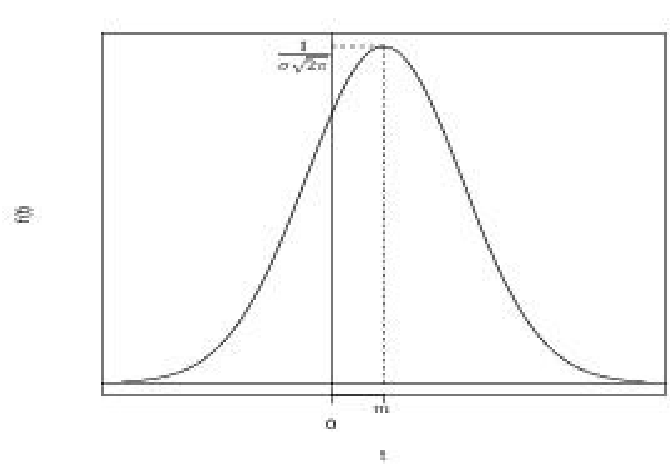

FIGURE 5.5

Remarque 5.3.2

- La courbe, dite courbe en cloche, a un axe de symétrie qui est la droite d'équation $x = m$.
- La densité f a un maximum atteint pour $x = m$ valant $\dfrac{1}{\sigma\sqrt{2\pi}}$.
- La courbe est d'autant plus pointue que σ est petit.

Rappels : Pour déterminer le(s) **point(s) d'inflexion** d'une fonction, on calcule sa dérivée seconde et on détermine le signe de cette dernière. Si le signe change pour une abscisse particulière, la fonction y admet un point d'inflexion.
Déterminons le(s) point(s) d'inflexion de f. Le calcul de la dérivée première donne :

$$f'(x) = \frac{1}{\sigma\sqrt{2\pi}}\frac{-2}{2\sigma^2}(x-m)e^{-\dfrac{(x-m)^2}{2\sigma^2}} = -\frac{1}{\sigma^3\sqrt{2\pi}}(x-m)e^{-\dfrac{(x-m)^2}{2\sigma^2}}.$$

Le calcul de la dérivée seconde donne :

$$f''(x) = -\frac{1}{\sigma^3\sqrt{2\pi}}\left[1-\frac{(x-m)^2}{\sigma^2}\right]e^{-\dfrac{(x-m)^2}{2\sigma^2}} = \frac{1}{\sigma^5\sqrt{2\pi}}[(x-m-\sigma)(x-m+\sigma)]e^{-\dfrac{(x-m)^2}{2\sigma^2}}.$$

On a le tableau de signes suivant :

x	$-\infty$		$m-\sigma$		$m+\sigma$		$+\infty$
$f''(x)$		$+$	|	$-$	|	$+$	

Par conséquent, la courbe admet deux points d'inflexion pour $x = m+\sigma$ et $x = m-\sigma$.

Exemple 5.3.1 Les variables normales sont très fréquentes, par exemple la variable aléatoire réelle "poids" d'un français adulte, la variable aléatoire "quotient intellectuel" d'une population donnée.

5.3.3 Les moments

L'espérance et la variance d'une variable normale sont respectivement données par :

$$\boxed{E(X) = m} \quad \text{et} \quad \boxed{V(X) = \sigma^2}$$

5.3.4 Variable normale centrée réduite

Si $m = 0$ et $\sigma = 1$, la variable normale est appelée **variable normale centrée réduite** et est notée $\mathcal{Z}$ ou Γ, on note alors

$$\boxed{\mathcal{Z} \rightsquigarrow \mathcal{N}(0,1)}$$

La courbe représentative de la fonction g est donnée par la Figure 5.6. Sa densité de probabilité est la fonction g définie par $g(x) = \dfrac{1}{\sqrt{2\pi}}e^{-\dfrac{x^2}{2}}$.

- Cette fonction g est paire, la courbe a un axe de symétrie qui est la droite des ordonnées. En $x = 0$, la fonction g vaut $g(0) = \dfrac{1}{\sqrt{2\pi}}$.
- Les points d'inflexion de la fonction g se trouvent en $x = -1$ et $x = 1$.
- En général, les valeurs de $g(x) = \dfrac{1}{\sqrt{2\pi}}e^{-\dfrac{x^2}{2}}$ sont données à l'aide d'une table pour $x \geq 0$.

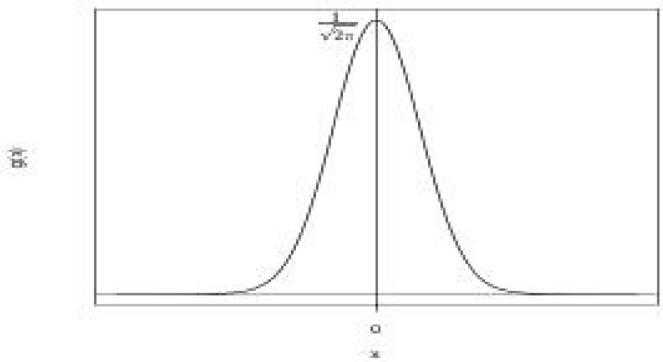

FIGURE 5.6

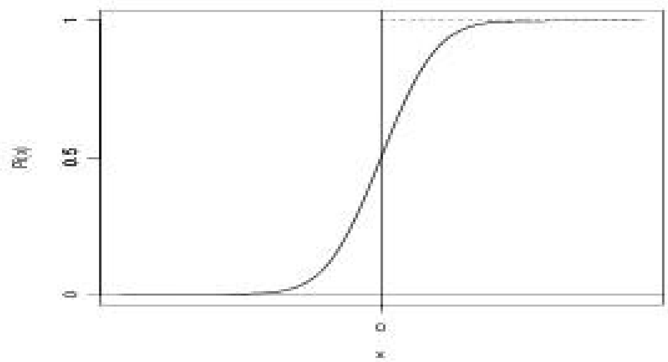

FIGURE 5.7

5.3.5 Fonction de répartition

Si on note Π la fonction de répartition de la loi normale centrée réduite $\mathcal{Z}$ associée à X,

$$\Pi(z) = \int_{-\infty}^{z} g(x)dx = p(\{\mathcal{Z} \leq z\})$$

cette fonction est représentée graphiquement à la Figure 5.7. Il existe des tables qui donnent la valeur de $\Pi(z)$ pour $z \geq 0$ (voir annexe A). $\Pi(z)$ désigne l'aire du domaine plan en jaune (voir Figure 5.8).

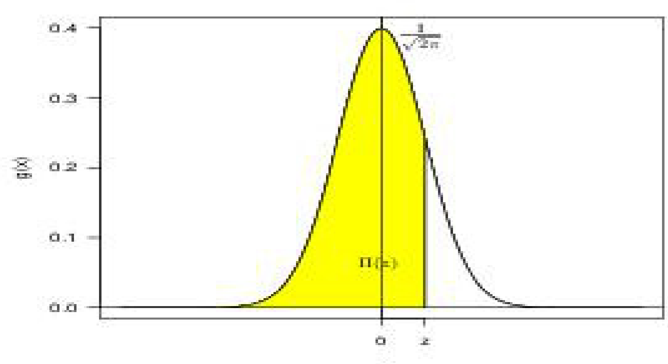

FIGURE 5.8

Exemple 5.3.2 $\Pi(0) = 0,5$ ce qui est évident puisqu'on considère exactement la moitié de l'aire totale (qui vaut 1). On peut également trouver à l'aide de la table $\Pi(1) = 0,8413$.

– Pour $z > 0$ on a la relation $\boxed{\Pi(z) + \Pi(-z) = 1}$

On peut observer cette propriété à l'aide de la Figure 5.9.

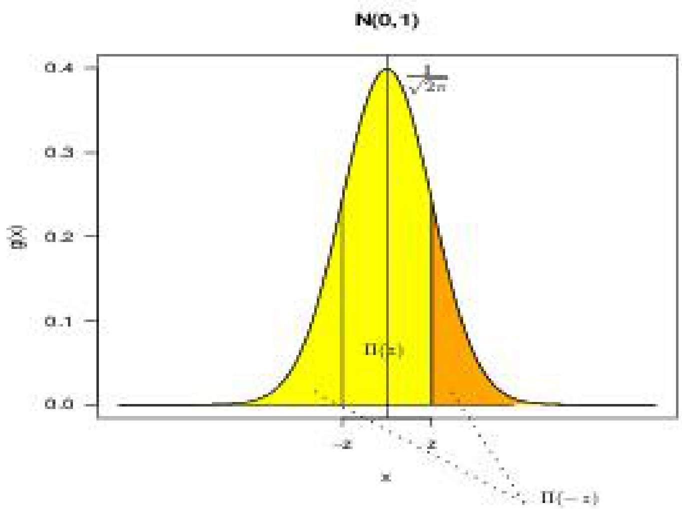

FIGURE 5.9

Exemple 5.3.3 $p(\{\mathcal{Z} \leq -1\}) = \Pi(-1) = 1 - \Pi(1) = 1 - 0,8413 = 0,1587$.

– Soit $z \in \mathbb{R}$, on a la relation $\boxed{p(\{\mathcal{Z} > z\}) = 1 - p(\{\mathcal{Z} \leq z\}) = 1 - \Pi(z)}$

On peut observer cette propriété à l'aide de la Figure 5.10.

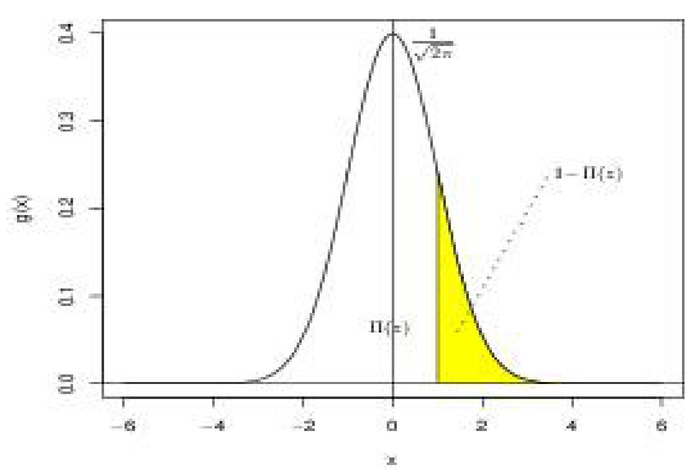

FIGURE 5.10

Exemple 5.3.4 $p(\{Z > 1\}) = 1 - \Pi(1) = 0,1587$.

– Soient $a, b \in \mathbb{R}$ vérifiant $a < b$ alors $\boxed{p(\{a \leq \mathcal{Z} \leq b\}) = \Pi(b) - \Pi(a)}$

On peut observer cette propriété à l'aide de la Figure 5.11.
Exemple 5.3.5 $p(\{1 \leq \mathcal{Z} \leq 2\}) = \Pi(2) - \Pi(1) = 0,9772 - 0,8413 = 0,1359$.

– Soit $z > 0$ alors $\boxed{p(\{-z \leq \mathcal{Z} \leq z\}) = \Pi(z) - \Pi(-z) = 2\Pi(z) - 1}$

On peut observer cette propriété à l'aide de la Figure 5.12.
Exemple 5.3.6 $p(\{-1 \leq \mathcal{Z} \leq 1\}) = 2\Pi(1) - 1 = 2 \times 0,8413 - 1 = 0,6826$.

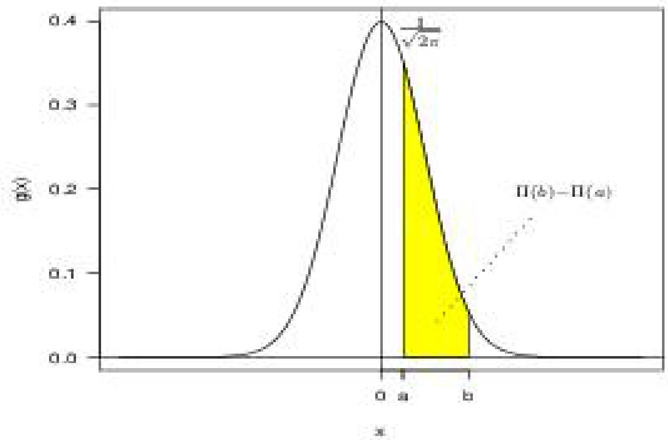

FIGURE 5.11

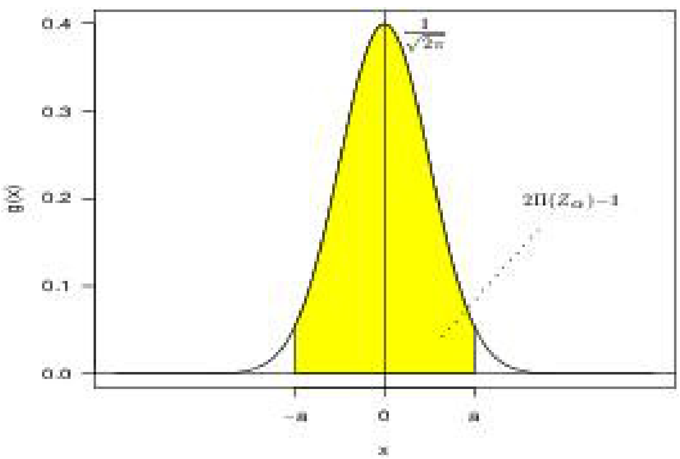

FIGURE 5.12

5.3.6 Table de l'écart réduit

Soit un intervalle centré en 0 de probabilité $1-\alpha$, on note $-Z_\alpha$ et Z_α ses bornes. Alors

$$p(\{-Z_\alpha \leq z \leq Z_\alpha\}) = 1-\alpha = 2\Pi(Z_\alpha)-1$$

ou encore

$$\boxed{\Pi(Z_\alpha) = 1 - \frac{\alpha}{2}}$$

Par conséquent,

$$\boxed{p(\{z > Z_\alpha\}) = p(\{z < -Z_\alpha\}) = \frac{\alpha}{2}}$$

Il existe une table (ne figurant pas dans les annexes) qui pour α fixé donne Z_α, ce qui permet d'obtenir les bornes d'un intervalle centré en 0 dont la probabilité est connue.

Exemple 5.3.7 Pour $\alpha = 0,03$ on obtient $Z_\alpha = 2,170090$ ce qui signifie

- $p(\{-2,17090 \leq Z \leq 2,17090\}) = 0,97$
- $p(\{Z > 2,17090\}) = 0,015$
- $p(\{Z < -2.17090\}) = 0,015$

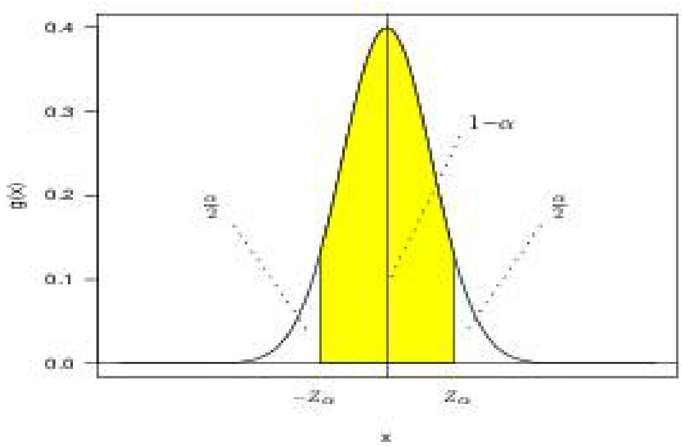

FIGURE 5.13

5.3.7 Exemples

1. Soient $X \rightsquigarrow \mathcal{N}(m = 30, \sigma = 3)$ et sa variable Z centrée réduite associée vérifiant $Z = \dfrac{X - 30}{3} \rightsquigarrow \mathcal{N}(0, 1)$.
   Déterminons les probabilités $p(\{X = 28\})$, $p(\{X \leq 33\})$, $p(\{X \leq 27\})$, $p(\{27 \leq X \leq 33\})$ et $p(\{X > 33\})$.

   - $p(\{X = 28\}) = 0$.
   - $X \leq 33 \Leftrightarrow Z \leq \dfrac{33 - 30}{3} = 1$ donc $p(\{X \leq 33\}) = p(\{Z \leq 1\}) = \Pi(1) = 0,8413$.
   - $\{X \leq 27\} = \left\{Z \leq \dfrac{27 - 30}{3} = -1\right\}$ donc $p(\{X \leq 27\}) = p(\{Z \leq -1\}) = \Pi(-1) = 1 - \Pi(1) = 0,1587$.
   - $\{27 \leq X \leq 33\} = \{-1 \leq Z \leq 1\}$ par conséquent $p(\{27 \leq X \leq 33\}) = p(\{-1 \leq Z \leq 1\}) = \Pi(1) - \Pi(-1) = 2\Pi(1) - 1$ ce qui donne $p(\{27 \leq X \leq 33\}) = 2 \times 0,8413 - 1 = 0,6826$.
   - $p(\{X > 33\}) = 1 - p(\{X \leq 33\}) = 1 - \Pi(1) = 0,1587$.

2. Soit une variable aléatoire réelle dont on sait qu'elle suit une loi normale. On sait de plus que $p(\{X \leq 3\}) = 0,5517$ et $p(\{X > 7\}) = 0,0166$.
   Déterminons les paramètres m et σ de la loi normale que suit X.

   On a $X \rightsquigarrow \mathcal{N}(m, \sigma) \Leftrightarrow Z = \dfrac{X - m}{\sigma} \rightsquigarrow \mathcal{N}(0, 1)$ donc

   - $X \leq 3 \Leftrightarrow Z \leq \dfrac{3 - m}{\sigma}$ et $p(\{X \leq 3\}) = p\left(\left\{Z \leq \dfrac{3 - m}{\sigma}\right\}\right) = 0,5517$. Alors $\Pi\left(\dfrac{3 - m}{\sigma}\right) = 0,5517$ or $\Pi(0,13) = 0,5517$ donc $\dfrac{3 - m}{\sigma} = 0,13$.
   - $X > 7 \Leftrightarrow Z > \dfrac{7 - m}{\sigma}$ et $p(\{X \leq 7\}) = 0,9834 = p\left(\left\{Z \leq \dfrac{7 - m}{\sigma}\right\}\right)$. Ainsi $\Pi\left(\dfrac{7 - m}{\sigma}\right) = 0,9834$ or $\Pi(2,13) = 0,9834$ donc $\dfrac{7 - m}{\sigma} = 2,13$.

   Afin de déterminer m et σ, on résout le système

$$\left\{\begin{array}{lcl} 7 - m & = & 2,13\sigma \\ 3 - m & = & 0,13\sigma \end{array}\right. \Leftrightarrow \left\{\begin{array}{lcl} 4 & = & 2\sigma \\ m & = & 3 - 0,13\sigma \end{array}\right. \Leftrightarrow \left\{\begin{array}{lcl} \sigma & = & 2 \\ m & = & 3 - 0,26 = 2,74 \end{array}\right. .$$

   Conclusion, $X \rightsquigarrow \mathcal{N}(m = 2,74; \sigma = 2)$.

5.3.8 Remarques

On sait que $X \rightsquigarrow \mathcal{N}(m,\sigma) \Leftrightarrow Z = \dfrac{X-m}{\sigma} \rightsquigarrow \mathcal{N}(0,1)$ donc :

- $\{m-\sigma \leq X \leq m+\sigma\} = \{-1 \leq Z \leq 1\}$ et $p(\{m-\sigma \leq X \leq m+\sigma\}) = \Pi(1) - \Pi(-1) = 2\Pi(1)-1 = 0,6826$. On peut alors affirmer que $68,26\%$ de la population étudiée appartient à l'intervalle $[m-\sigma; m+\sigma]$.
- $\{m-2\sigma \leq X \leq m+2\sigma\} = \{-2 \leq Z \leq 2\}$ et $p(\{m-2\sigma \leq X \leq m+2\sigma\}) = 2\Pi(2)-1 = 0,9544$. On peut alors affirmer que $95,44\%$ de la population étudiée appartient à l'intervalle $[m-2\sigma; m+2\sigma]$.
- $\{m-3\sigma \leq X \leq m+3\sigma\} = \{-3 \leq Z \leq 3\}$ et $p(\{m-3\sigma \leq X \leq m+3\sigma\}) = 2\Pi(3)-1 = 0,9973$. On peut alors affirmer que $99,73\%$ de la population étudiée appartient à l'intervalle $[m-3\sigma; m+3\sigma]$.

5.3.9 Relation entre la fonction de répartition et la densité de probabilité des loi normale et loi normale centrée réduite

– Soit $X \rightsquigarrow \mathcal{N}(m,\sigma)$ de densité f. La fonction de répartition F est définie par $F(x) = p(\{X \leq x\})$ et vérifie alors la relation

$$\boxed{F' = f}$$

– Soit $Z \rightsquigarrow \mathcal{N}(0,1)$ de fonction de répartition Π et de densité g alors

$$\boxed{\Pi' = g}$$

– On a l'égalité $\{X \leq x\} = \left\{Z \leq \dfrac{x-m}{\sigma}\right\}$ ainsi que la relation $F(x) = \Pi\left(\dfrac{x-m}{\sigma}\right)$. Par dérivation,

$$F'(x) = \frac{1}{\sigma}\Pi'\left(\frac{x-m}{\sigma}\right) \Leftrightarrow f(x) = \frac{1}{\sigma}g\left(\frac{x-m}{\sigma}\right),$$

ce qui permet l'utilisation de la table de densité de probabilité de la loi $\mathcal{N}(0,1)$ pour calculer les valeurs de la densité de probabilité de $\mathcal{N}(m,\sigma)$.

5.3.10 Propriétés

Soit $X \rightsquigarrow \mathcal{N}(m,\sigma)$ et k une constante. On a les résultats suivants :

– la variable kX suit une loi normale $\mathcal{N}(km, |k|\sigma)$,

– la variable $k+X$ suit une loi normale $\mathcal{N}(k+m,\sigma)$.

5.3.11 Somme de deux variables normales indépendantes

Soient deux variables $X_1 \rightsquigarrow \mathcal{N}(m_1,\sigma_1)$ et $X_2 \rightsquigarrow \mathcal{N}(m_2,\sigma_2)$ indépendantes. Alors,

$$\boxed{X_1 + X_2 \rightsquigarrow \mathcal{N}(m_1+m_2, \sqrt{\sigma_1^2+\sigma_2^2})}$$

et

$$\boxed{X_1 - X_2 \rightsquigarrow \mathcal{N}(m_1-m_2, \sqrt{\sigma_1^2+\sigma_2^2})}$$

Preuve : Ces propriétés sont obtenues en utilisant les résultats suivants

$$E(X_1+X_2) = E(X_1)+E(X_2) \;\; ; \;\; E(X_1-X_2) = E(X_1)-E(X_2)$$
$$\sigma(X_1+X_2) = \sqrt{\sigma^2(X_1)+\sigma^2(X_2)} \;\; ; \;\; \sigma(X_1-X_2) = \sqrt{\sigma^2(X_1)+\sigma^2(X_2)}.$$

■

Plus généralement, soient n variables aléatoires indépendantes deux à deux telles que $X_i \rightsquigarrow \mathcal{N}(m_i,\sigma_i)$, $\forall i \in \{1,2,\ldots,n\}$, alors

$$\boxed{\text{la variable } X = \sum_{i=1}^{n} a_i X_i \text{ suit une loi normale } \mathcal{N}(m, \sigma)}$$

de moyenne $m = \sum_{i=1}^{n} a_i m_i$ et d'écart-type $\sigma = \sqrt{\sum_{i=1}^{n} a_i^2 \sigma_i^2}$.

Remarque 5.3.3 Si les X_i suivent la même loi $\mathcal{N}(m, \sigma)$,

- la variable $X = X_1 + X_2 + \ldots + X_n$ suit une loi normale de moyenne nm et d'écart-type $\sqrt{\sigma^2 + \sigma^2 + \ldots + \sigma^2} = \sqrt{n\sigma^2} = \sigma\sqrt{n}$.
- La variable $Y = \dfrac{X_1 + X_2 + \ldots + X_n}{n}$ suit une loi normale de moyenne $\dfrac{nm}{n} = m$, d'écart-type $\dfrac{\sigma}{\sqrt{n}}$.

5.3.12 Approximation d'une loi binomiale par une loi normale

Soit une loi binomiale $\mathcal{B}(n, p)$ de moyenne $m = np$, d'écart-type $\sigma = \sqrt{npq}$.
On montre que l'on peut approximer la loi binomiale $\mathcal{B}(n, p)$ par la loi normale $\mathcal{N}(m = np, \sigma = \sqrt{npq})$ si $n \geq 15$, p et q étant non voisins de 0. Dans la pratique, l'approximation est admise si $n \geq 20$, $np \geq 10$, $nq \geq 10$.

Exemple 5.3.8 Soit $X \rightsquigarrow \mathcal{B}(n = 100; p = 0,4)$ avec $E(X) = 40$ et $\sigma(X) = \sqrt{40 \times 0,6} = \sqrt{24}$. On a dans ce cas

$$\begin{cases} n = 100 \geq 20 \\ np = 40 \geq 10 \\ nq = 60 > 10 \\ npq = 24 \end{cases}$$

ce qui implique que $X \rightsquigarrow \mathcal{N}(m = 40, \sigma = \sqrt{24})$. Calculons $p(\{X < 38\})$.

- On a $p(\{X = 38\})_{\mathcal{B}} = p(\{37,5 \leq X \leq 38,5\})_{\mathcal{N}}$. Posons $Z = \dfrac{X - 40}{\sqrt{24}}$ alors
$$\{37,5 \leq X \leq 38,5\} = \left\{\frac{-2,5}{\sqrt{24}} \leq Z \leq \frac{-1,5}{\sqrt{24}}\right\} = \{0,510 \leq Z \leq 0,306\}.$$
Ainsi, $p(\{X = 38\})_{\mathcal{B}} = \Pi(-0,306) - \Pi(-0,510) = \Pi(-0,510) - \Pi(-0,306) = 0,6950 - 0,6202 = 0,0748$.
Montrons que $\pi(0,306) = 0,6202$ à l'aide de l'interpolation linéaire : on a $\pi(0,30) = 0,6179$ et $\pi(0,31) = 0,6217$ et le tableau suivant

$0,6179$	x	$0,6217$
$0,30$	$0,306$	$0,31$

$$\Leftrightarrow \frac{x - 0,6179}{0,6217 - 0,6179} = \frac{0,306 - 0,30}{0,31 - 0,30} \Leftrightarrow x = 0,6179 + 0,0038 \times \frac{0,006}{0,01} = 0,6202.$$

- Ensuite, $p(\{X > 38\})_{\mathcal{B}} = p(\{X > 38,5\})_{\mathcal{N}}$ et $\{X > 38,5\} = \left\{Z > \dfrac{38,5 - 40}{\sqrt{24}}\right\}$. Par conséquent,
$p(\{X > 38\})_{\mathcal{B}} = p(\{X > 38,5\})_{\mathcal{N}} = 1 - p(\{X \leq 38,5\})_{\mathcal{N}}$ ou encore
$p(\{X > 38\})_{\mathcal{B}} = 1 - \Pi(-0,306) = \Pi(0,306) = 0,6202$.
- Enfin, $p(\{X \leq 38\})_{\mathcal{B}} = p\left(\{X \leq 38,5\}\right)_{\mathcal{N}}$ et $\{X \leq 38,5\} = \left\{Z \leq \dfrac{40 - 38,5}{\sqrt{24}}\right\}$. Par conséquent,
$p(\{X \leq 38\})_{\mathcal{B}} = p\left(\left\{Z \leq \dfrac{-1,5}{\sqrt{24}}\right\}\right)_{\mathcal{B}} = \Pi(-0,306)$.

Finalement, $p(\{X \leq 38\})_{\mathcal{B}} = 1 - \Pi(0,306) = 1 - 0,6202 = 0,3797$.

5.3.13 Résumé sur les approximations de lois

- $\boxed{\mathcal{H}(N,p,n) \sim \mathcal{B}(n,p)}$ pour $N > 10n$,
- $\boxed{\mathcal{B}(n,p) \sim \mathcal{P}(\lambda = np)}$ pour $n \geq 30$, $p \leq 0,1$ et $np \leq 10$,
- $\boxed{\mathcal{B}(n,p) \sim \mathcal{N}(m = np, \sigma = \sqrt{npq})}$ avec $\left\{\begin{array}{l} n \geq 20 \\ 0,1 < p < 0,9 \end{array}\right.$ ou $\left\{\begin{array}{l} np \geq 10 \\ nq \geq 10 \end{array}\right.$ ou $npq > 3$.
- $\boxed{\mathcal{P}(\lambda = np) \sim \mathcal{N}(m = np, \sigma = \sqrt{npq})}$ pour $np \geq 10$.

5.4 Loi et variable du χ^2 (Khi-deux) de Pearson

5.4.1 La distribution du χ^2

1. On considère n variables indépendantes d'une loi normale centrée réduite T_1, T_2,...,T_n. La quantité

$$T_1^2 + T_2^2 + \ldots + T_n^2 = \sum_{i=1}^{n} T_i^2$$

est une variable aléatoire dont la distribution est celle d'un χ^2 **à** n **degrés de liberté** de moyenne et variance respectives,

$$\boxed{E(\chi_n^2) = n} \quad \text{et} \quad \boxed{V(\chi_n^2) = 2n}$$

Lorsque n augmente, la densité f d'une loi du χ^2 ressemble de plus en plus à la densité d'une loi normale (voir la Figure 5.14) : La variable χ^2 est tabulée en fonction du nombre n de degrés de liberté.

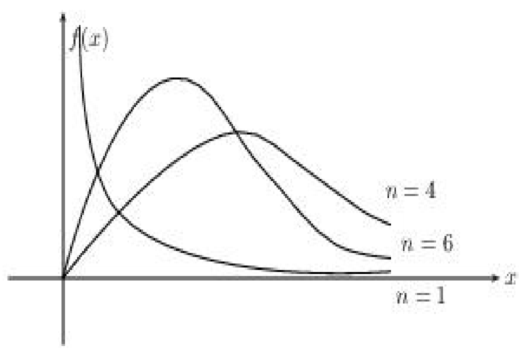

FIGURE 5.14

La table (voir annexe B) donne pour différentes valeurs de α, la valeur de x telle que :

$$P(\{\chi_n^2 < x\}) = 1 - \alpha$$

2. Graphiquement, cette valeur est égale à la surface grisée de la Figure 5.15 :

Exemple 5.4.1 Calculer $p(\{\chi_{10}^2 > 20,5\})$. On récupère à l'aide de la table, la probabilité $p(\{\chi_{10}^2 < 20,5\}) = 0,975$. Par conséquent, la probabilité recherchée $p(\{\chi_{10}^2 > 20,5\})$ est égale à $1-0,975 = 0,025$.

Remarque 5.4.1 Attention, d'autres tables donnent la probabilité α, en fonction du nombre de degrés de liberté ν pour qu'une variable aléatoire X suivant une loi de χ_ν^2 soit supérieure ou égale à une valeur donnée x : $\alpha = p(\{X \geq x\})$.

On a la propriété suivante :

$$\boxed{\chi_m^2 + \chi_n^2 = \chi_{m+n}^2}$$

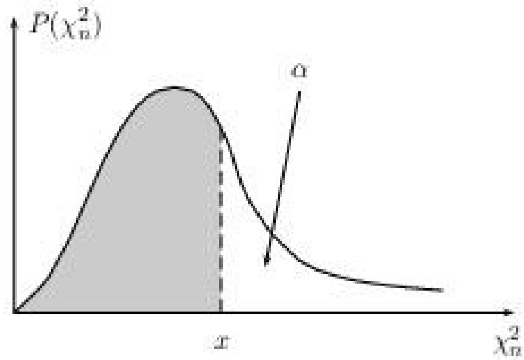

FIGURE 5.15

Ce χ^2 admet

– une moyenne $E(\chi^2_{m+n}) = m + n$,

– une variance $\sigma^2(\chi^2_{m+n}) = 2(m + n)$

et ceci par application directe du théorème sur l'addition de variables aléatoires indépendantes.

5.4.2 Les données du problème

Certains tests ont pour objet de tirer des conclusions relatives à la valeur des paramètres (moyenne, fréquence, variance) d'une ou plusieurs populations, sur la base d'informations partielles fournies par un ou plusieurs échantillons.
La même démarche peut être appliquée pour porter un "jugement" sur les caractéristiques encore plus générales de la population : la forme même de distribution du caractère étudié, la validité de sa représentation à l'aide de telle ou telle loi de probabilité particulière, les relations éventuelles entre plusieurs variables.

Concrètement, on dispose d'une distribution statistique empirique se présentant sous la forme d'une table d'effectifs ou de fréquences du caractère étudié. On désire savoir si ces effectifs ou ces fréquences sont compatibles avec une distribution théorique déterminée telle que la loi binomiale, la loi de Poisson, la loi normale ou toute autre loi de probabilité. Il s'agit en d'autres termes d'apprécier l'adéquation d'une distribution théorique particulière, en tant que représentation d'un phénomène concret observé (série empirique).

La démarche consiste donc à tester l'hypothèse selon laquelle notre échantillon serait tiré d'une population régie par une certaine loi de probabilité.
Il est évident que, même si le phénomène étudié suit effectivement la loi de probabilité dont on teste l'adéquation, les fréquences expérimentales (ou empiriques) observées sur un échantillon particulier différeront nécessairement peu ou prou des probabilités (fréquences que l'on devrait théoriquement observer selon la loi en question).

La problématique du test revient en définitive à savoir si les différences constatées entre la distribution expérimentale et la distribution théorique supposée sont explicables par l'aléa lié à la constitution de l'échantillon ou si elles sont trop importantes pour être imputables au seul hasard. En ce cas, c'est l'hypothèse de travail avancée sur la nature de la distribution qui devrait être mise en cause.

5.4.3 Ajustement d'une distribution observée à une distribution théorique

Construction du test

1. Les hypothèses du test sont les suivantes :
   - H_0 : X suit la loi théorique L,

- H_1 : X ne suit pas L.

2. La variable observée est :
   - soit discrète et prend k valeurs x_1, x_2,...,x_k
   - soit continue et classée en k classes $[a_0; a_1[$, $[a_1; a_2[$,...,$[a_{k-1}; a_k[$ de centres respectifs x_1, x_2,...,x_{k-1}, x_k.
3. Les N observations de l'échantillon sont réparties sur les k valeurs de X (si X est discrète) ou sur les k classes de X (si X est continue). On a les tableaux suivants :

x_i	n_i
x_1	n_1
x_2	n_2
$\vdots$	$\vdots$
x_k	n_k

Classes	Centres x_i	Effectifs n_i
$[a_0; a_1[$	x_1	n_1
$[a_1; a_2[$	x_2	n_2
$\vdots$	$\vdots$	$\vdots$
$[a_{k-1}; a_k[$	x_k	n_k

   avec $N = \sum_{i=1}^{k} n_i = n_1 + n_2 + \ldots + n_k$.

4. Sous H_0 on note p_i la probabilité dite théorique définie par
   - $p_i = p(\{X = x_i / X \rightsquigarrow L\})$ si X est discrète,
   - $p_i = p(\{X \in [a_{i-1}; a_i[/ X \rightsquigarrow L\})$ si X est continue.

   $e_i = Np_i$ est l'effectif théorique de la i-ième classe de X.

5. L'indicateur d'écart entre les distributions observées et théoriques est

$$\boxed{\sum_{i=1}^{k} \frac{(n_i - Np_i)^2}{Np_i}} \qquad (1)$$

   dit χ^2 **observé ou calculé**. Cet écart suit pour N suffisamment grand une loi du χ^2_ν d'où le nom du test.

   Intuitivement, on comprend que cette grandeur statistique traduise l'écart entre l'échantillon et la loi conjecturée.
   Si l'ajustement était parfait, cette expression du χ^2 serait nulle, les effectifs empiriques coïncidant exactement avec les effectifs théoriques.
   En revanche, plus grands sont les écarts entre les effectifs observés et les effectifs théoriques $(n_i - e_i)$ et plus forte sera la valeur du χ^2.

   En outre, comme la quantité (1) ne peut pas être négative, le test d'ajustement est nécessairement un test unilatéral droit.

   Le paramètre ν indiçant χ^2_ν définit le **nombre de degrés de liberté**. C'est le nom donné au nombre d'observations linéairement indépendantes qui apparaissent dans une somme de carrés. Autrement dit, c'est le nombre d'observations aléatoires indépendantes à qui l'on soustrait le nombre de contraintes imposées à ces observations.

   Le nombre ν de degrés de liberté est égal à
   - si les paramètres de la loi d'ajustement L sont donnés,

$$\boxed{\nu = k - 1}$$

En effet, aucun paramètre n'est à estimer puisque la loi d'ajustement est totalement spécifiée. Le χ^2 est constitué de k écarts $(n_i - e_i)$. Les écarts sont reliés par la contrainte

$$\sum(n_i - e_i) = \sum(n_i - Np_i) = \sum n_i - N\sum p_i = N - N = 0$$

En d'autres termes, lorsqu'on connaît la valeur de $k-1$ écarts, on peut en déduire la valeur du dernier qui n'est donc pas "libre" de varier de manière aléatoire,

– si la loi d'ajustement L comporte r paramètres inconnus,

$$\boxed{\nu = k - r - 1}$$

On impose de ce fait autant de contraintes supplémentaires entre les observations, diminuant d'autant le nombre de degrés de liberté.

Remarque 5.4.2 *Le nombre d'observations par classes ne doit pas être faible, Np_i doit être **supérieur** à 5, $\forall i = 1, 2, \ldots, k$. Dans le cas contraire, on regroupe deux ou plusieurs classes adjacentes de façon à réaliser cette condition. On tient compte de ce regroupement pour le nombre de degrés de liberté.*

6. Pour un risque de première espèce α, la région critique est définie pour

$$\boxed{\sum_{i=1}^{k} \frac{(n_i - Np_i)^2}{Np_i} \geq \chi^2_{\nu,1-\alpha}}$$

d'où la règle de décision :

- χ^2 observé $< \chi^2_{\nu,1-\alpha}$, on décide H_0 et $X \rightsquigarrow L$.
- χ^2 observé $\geq \chi^2_{\nu,1-\alpha}$, on décide H_1 et X ne suit pas la loi L.

Exemple 5.4.2 Loi uniforme.
Une statistique relative aux résultats du concours d'entrée à une grande école fait ressortir les répartitions des candidats et des admis selon la profession des parents.

Profession des candidats	Nombre de candidats	Nombre d'admis
① Fonctionnaires et assimilés	2244	180
② Commerce, industrie	988	89
③ Professions libérales	575	48
④ Propriétaires rentiers	423	37
⑤ Propriétaires agricoles	287	13
⑥ Artisans, petits commerçants	210	18
⑦ Banque, assurance	209	17
Total	4936	402

Question : Tester l'hypothèse (risque $\alpha = 0,05$) selon laquelle la profession des parents n'a pas d'influence sur l'accès à cette grande école.

Il s'agit du test d'ajustement d'une distribution théorique, on considère les hypothèses :

- H_0 : la profession des parents n'a pas d'influence sur l'accès à cette grande école, la proportion des admis est constante pour toutes les professions soit $p = \dfrac{402}{4936} \simeq 0,0814$.
- H_1 : la profession des parents influe sur l'accès à cette grande école.

Sous H_0, le nombre d'admis pour la i-ième profession est pN_i.

i	N_i	n_i effectif observé	N_ip effectif théorique	$\dfrac{(n_i - N_ip)^2}{N_ip}$
1	2244	180	$\dfrac{2244 \times 402}{4936} \simeq 182,76$	$0,0416$
2	988	89	$\dfrac{988 \times 402}{4936} \simeq 80,47$	$0,9042$
3	575	48	$\dfrac{575 \times 402}{4936} \simeq 46,83$	$0,0293$
4	423	37	$\dfrac{423 \times 402}{4936} \simeq 34,45$	$0,1887$
5	287	13	$\dfrac{287 \times 402}{4936} \simeq 23,37$	$4,6050$
6	210	18	$\dfrac{210 \times 402}{4936} \simeq 17,10$	$0,0471$
7	209	17	$\dfrac{209 \times 402}{4936} \simeq 17,02$	$\simeq 0$
Total	4936	402	402	$5,8181$

Le χ^2 observé vaut $5,8181$. Le nombre de degrés de liberté est $7 - 1 = 6$. La table de l'annexe B fournit $\chi^2_{6;0,95} = 12,59$ donc χ^2 observé $< \chi^2_{6/0,95}$. On choisit H_0, ce qui signifie que la profession des parents n'a pas d'influence sur l'accès à cette grande école.

Exemple 5.4.3 *Loi binomiale.*
Supposons qu'on ait recueilli 300 boîtes contenant chacune trois ampoules. Dans chaque boîte, on compte le nombre d'ampoules défectueuses. On obtient les résultats suivants :

Nombre d'ampoules défectueuses x_i	Nombre de boîtes observées n_i
0	190
1	95
2	10
3	5
Total	300

Pour chaque ampoule testée, on peut observer deux états différents : l'ampoule est défectueuse ou non. Le nombre X d'ampoules défectueuses par boîte suit une loi binomiale de paramètres $n = 3$ et p.
Dans la distribution observée, le nombre d'ampoules défectueuses est de

$$0 \times 190 + 1 \times 95 + 2 \times 10 + 3 \times 5 = 130$$

soit 130 ampoules défectueuses sur un total de 900 ampoules. La proportion d'ampoules défectueuses est alors de $\dfrac{130}{900} \simeq 0,144$. Prenons $p = 0,15$ et réalisons alors le test suivant : soit X le nombre d'ampoules défectueuses par boîte

- $H_0 : X \rightsquigarrow \mathcal{B}(3; 0,15)$.
- $H_1 : X$ ne suit pas cette loi binomiale.

On détermine les probabilités théoriques :

- $p_0 = p(\{X = 0/X \rightsquigarrow \mathcal{B}\}) = (0,85)^3 \simeq 0,6141$
- $p_1 = p(\{X = 1/X \rightsquigarrow \mathcal{B}\}) = C_3^1(0,15)(0,85)^2 \simeq 0,3251$
- $p_2 = p(\{X = 2/X \rightsquigarrow \mathcal{B}\}) = C_3^2(0,15)^2(0,85) \simeq 0,0574$
- $p_3 = p(\{X = 3/X \rightsquigarrow \mathcal{B}\}) = (0,15)^3 \simeq 0,0034$

On a le tableau

x_i	effectif observé n_i	p_i	effectif théorique Np_i
0	190	$0,6141$	$184,23$
1	95	$0,3251$	$97,53$
2	10	$0,0574$	$17,22$
3	5	$0,0034$	$1,02$
Total	$N = 300$	1	300

L'effectif théorique de la quatrième classe est faible : $1,02 < 5$. On effectue un regroupement de classes, les classes "2" et "3".

x_i	n_i	Np_i	$\frac{(n_i - Np_i)^2}{Np_i}$
0	190	$184,23$	$0,18071$
1	95	$97,53$	$0,06563$
2 ou 3	15	$18,24$	$0,57553$
Total	300	300	$0,82187$

Après le regroupement, le nombre de classes est 3, le nombre de degrés de liberté est $3 - 1 = 2$.
Au risque $\alpha = 0,01$ on a $\chi^2_{2;0,99} = 9,21$. Donc

$$\chi^2 \text{ observé} = 0,82187 < \chi^2_{2;0,99}.$$

On ne rejette pas H_0 au profit de H_1. On considère que le nombre d'ampoules défectueuses par boîte suit une loi binomiale de paramètres $n = 3$ et $p = 0,15$ au risque $\alpha = 0,01$.

Exemple 5.4.4 *Loi normale.*
On suppose que le rendement X (quintaux par hectares d'une parcelle de blé) suit une loi normale $\mathcal{N}(m, \sigma)$.
L'observation du rendement de 1000 parcelles a donné les résultats suivants :

Rendement	$[0;10[$	$[10;20[$	$[20;30[$	$[30;40[$	$[40;50[$	$[50;60[$	$[60;70[$	$[70;80[$	$[80;90[$
Nombre de parcelles	5	6	40	168	288	277	165	49	2

1. Déterminer la moyenne arithmétique et l'écart-type de la distribution observée.
   - $\overline{x} = \dfrac{\displaystyle\sum_i n_i x_i}{N} = 49,76$
   - $\sigma'^2 = \dfrac{\displaystyle\sum_i n_i x_i^2}{N} - \overline{x}^2 = 164,5424$ donc $\sigma' \simeq 12,827$.

2. Vérifier pour un test du χ^2 avec un risque de $0,05$ si l'ajustement de la distribution observée à une loi normale $\mathcal{N}(n = 50, \sigma = 13)$ est acceptable.

   Les hypothèses du test du χ^2 sont les suivantes :
   - $H_0 : X \rightsquigarrow \mathcal{N}(50, 13)$
   - $H_1 : X$ ne suit pas $\mathcal{N}(50, 13)$

   On désigne par $[a_0; a_1[$, $[a_1; a_2[$,...,$[a_8; a_9[$ les classes et par x_1, x_2,...,x_9 les centres de ces classes. Sous H_0, $X \rightsquigarrow \mathcal{N}(50, 13)$ et $Z = \dfrac{X - 50}{13} \rightsquigarrow \mathcal{N}(0, 1)$, donc $p_i = p(\{X \in [a_{i-1}; a_i[\}) = \Pi(z_i) - \Pi(z_{i-1})$ avec $z_i = \dfrac{a_i - 50}{13}$ et $z_{i-1} = \dfrac{a_{i-1} - 50}{13}$. L'effectif théorique de la i-ème classe est $1000p_i$ et

$$\sum_i \frac{(n_i - Np_i)^2}{Np_i} \rightsquigarrow \chi^2_\nu$$

Classe $[x_{i-1}; x_i[$	n_i	z_i	$\Pi(z_i)$	p_i	Np_i	Np_i corrigé	n_i corrigé	$\frac{(n_i - Np_i)^2}{Np_i}$
$[0; 10[$	5	$-3,0769$	$0,001$	$0,0009$	$0,9$	$10,4$	11	$0,0346$
$[10; 20[$	6	$-2,3077$	$0,0105$	$0,0095$	$9,5$			
$[20; 30[$	40	$-1,5385$	$0,0620$	$0,0515$	$51,5$	$51,5$	40	$2,568$
$[30; 40[$	168	$-0,7692$	$0,2209$	$0,1589$	$158,9$	$158,9$	168	$0,5211$
$[40; 50[$	288	0	$0,5$	$0,2791$	$279,1$	$279,1$	288	$0,283$
$[50; 60[$	277	$0,7692$	$0,7791$	$0,2791$	$279,1$	$279,1$	277	$0,0158$
$[60; 70[$	165	$1,5385$	$0,9380$	$0,1589$	$158,9$	$158,9$	165	$0,234$
$[70; 80[$	49	$2,3077$	$0,9895$	$0,0515$	$51,5$	$51,5$	49	$0,1214$
$[80; 90[$	2	$3,0769$	$0,9990$	$0,0095$	$9,5$	$9,5$	2	$5,9211$
Total	1000	–	–	1	1000	1000	1000	$9,7$

   On effectue le regroupement des deux premières classes car $Np_i < 5$. Le χ^2 observé vaut $9,7$. Après le regroupement, il reste 8 classes, les deux paramètres de la loi normale sont donnés, le nombre de degrés de liberté est $\nu = (9 - 1) - 1 = 7$. À l'aide de la table, on obtient $\chi^2_{7;0,95} = 14,07$. Ainsi,

$$\chi^2 \text{ observé} < \chi^2_{7;0,95}.$$

   On choisit H_0, l'ajustement de la distribution observée à une loi normale $\mathcal{N}(50, 13)$ est acceptable.

Exemple 5.4.5 *Loi de Poisson.*
Souvent, lorsqu'on envisage une modèle pour un phénomène qu'on étudie, on ne spécifie pas complètement la loi qu'on considère. Supposons qu'on s'intéresse au nombre de voitures se présentant par minute à un poste de péage sur une autoroute. On peut se demander si cette variable aléatoire peut être modélisée par une loi de Poisson. On souhaite donc tester l'hypothèse fondamentale

$$H_0 : X \rightsquigarrow \mathcal{P}(\lambda)$$

contre l'hypothèse alternative

$$H_1 : X \text{ ne suit pas } \mathcal{P}(\lambda).$$

On ne précise pas la valeur du paramètre λ. On peut toutefois l'estimer à partir des données disponibles mais dans ce cas, $r = 1$. Le nombre de degrés sera alors $\nu = k - r - 1 = k - 2$.

On effectue 200 comptages au péage.

x_i	0	1	2	3	4	5	6	7	8	≥ 9	Total
n_i	6	15	40	42	37	30	10	12	8	0	200
$n_i x_i$	0	15	80	126	148	150	60	84	64	0	727

où x_i est le nombre de voitures par minute lors de la i-ième l'observation et n_i est l'effectif correspondant. Par exemple, $x_1 = 0$ et $n_1 = 6$ c'est-à-dire que lors de 6 observations, il y a 0 voiture.
La moyenne arithmétique de cette distribution observée est

$$\overline{x} = \frac{\sum_i n_i x_i}{\sum_i n_i} = \frac{727}{200} = 3,635 \simeq 3,5$$

On peut tester l'hypothèse

$$H_0 : X \rightsquigarrow \mathcal{P}(\lambda = 3,5).$$

x_i	n_i	p_i	Np_i	Np_i corrigé	n_i corrigé	$\frac{(n_i - Np_i)^2}{Np_i}$
0	6	0,0302	6,04	6,04	6	0,00026
1	15	0,1057	21,14	21,14	15	1,78333
2	40	0,1850	37	37	40	0,24324
3	42	0,2158	43,16	43,16	42	0,03118
4	37	0,1888	37,76	37,76	37	0,01530
5	30	0,1322	26,44	26,44	30	0,47933
6	10	0.0771	15,42	15,42	10	1,90508
7	12	0,0385	7,7	7,7	12	2,40130
8	8	0,0169	3,38	5,34	8	1,32502
≥ 9	0	0,0098	1,96			
Total	200	1	200	200	200	8,18404

On a $p_i = p(\{X = x_i / X \rightsquigarrow \mathcal{P}(3,5)\})$ donc

- $p_0 = p(\{X = 0 / X \rightsquigarrow \mathcal{P}(3,5)\}) = e^{-3,5} \simeq 0,0302$ et
- $p_1 = p(\{X = 1 / X \rightsquigarrow \mathcal{P}(3,5)\}) = e^{-3,5} 3,5 \simeq 0,1057$.

On a effectué le regroupement des deux dernières classes car l'effectif théorique y est inférieur à 5. Après ce regroupement, le nombre de classes est de 9. Le nombre de degrés de liberté est $9 - 1 - 1 = 7$. Au risque $\alpha = 0,01$, $\chi^2_{7;0,99} = 18,48$ donc

$$\chi^2 \text{ observé} = 8,18404 < \chi^2_{7;0,99}.$$

On ne rejette pas l'hypothèse H_0 et $X \rightsquigarrow \mathcal{P}(\lambda = 3,5)$ au risque $\alpha = 0,01$.

5.5 Loi de Student-Fischer

5.5.1 Définition

La **loi de Student** est une loi continue qui comme la loi du χ^2 dépend d'un seul paramètre qu'on appellera également degré de liberté et qu'on note ν ($\nu \in \mathbb{N}^\star$). La variable X distribuée selon cette loi qu'on note

$$\boxed{X \rightsquigarrow t_\alpha}$$

prend toutes ses valeurs dans $\mathbb{R}$. Si $Y \rightsquigarrow \mathcal{N}(0,1)$ et $Z \rightsquigarrow \chi^2_\nu$, Y et Z étant indépendantes, la variable $X = \dfrac{Y}{\sqrt{\dfrac{Z}{\nu}}}$ suit une loi de Student à ν degrés de liberté.

On dit qu'une variable aléatoire réelle à densité X a une loi de probabilité de Student à ν degrés de liberté (n entier > 0) si, et seulement si, sa densité de probabilité est donnée par la formule :

$$f_\nu(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu}\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)\left(1+\frac{x^2}{\nu}\right)^{\frac{\nu+1}{2}}}.$$

Dans cette formule, Γ est la fonction Gamma d'Euler définie, lorsque la partie réelle de x est positive, par :

$$\Gamma(x) = \int_0^{+\infty} e^u u^{x-1} du.$$

La loi de Student à ν degrés de liberté est la loi de probabilité du quotient d'une variable normale centrée réduite par la racine carrée de la somme des carrés de ν variables normales centrées réduites indépendantes entre elles et indépendantes de la première variable.

Pour $\nu = 1$, la loi de Student s'appelle **loi de Cauchy**, ou **loi de Lorentz**. C'est la loi du rapport de deux variables normales centrées réduites indépendantes.

5.5.2 Les courbes

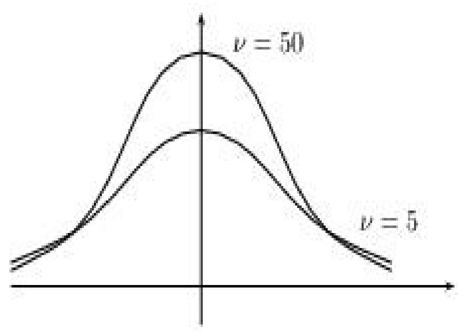

FIGURE 5.16

La courbe est unimodale, centrée, symétrique et plus plate que la courbe d'une loi normale. Lorsque le nombre de degrés de liberté augmente, la loi de Student tend vers la loi normale $\mathcal{N}(0, 1)$ (voir Figure 5.16).

5.5.3 Les moments

Soit $X \rightsquigarrow t_\nu$, on a

$$\boxed{E(X) = 0} \text{ et } \boxed{V(X) = \frac{\nu}{\nu - 2}} \text{ pour } \nu > 2.$$

Remarque 5.5.1

- Lorsque l'espérance existe, elle est nulle, puisque la loi est symétrique autour de 0.
- Lorsque $\nu = 1$ ou $\nu = 2$, la variance n'est pas déterminée.
- Lorsque ν tend vers l'infini, la variance tend vers 1.

5.5.4 Les tables

Soit $X \rightsquigarrow t_\nu$. Il existe une table (voir annexe C1) qui fournit les valeurs $t_{\nu,1-\alpha}$ pour ν et α donnés, telles que

$$p(\{X < t_{\nu,1-\alpha}\}) = 1 - \alpha.$$

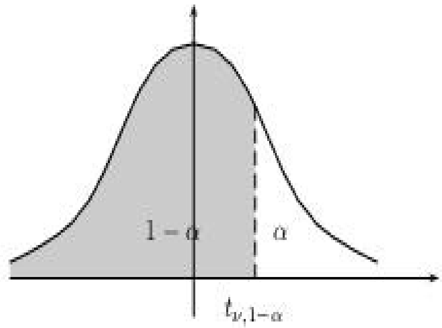

FIGURE 5.17

Graphiquement, cette probabilité est donnée par la surface grisée de la Figure 5.17 :

Exemple 5.5.1

- $\nu = 10$, $\alpha = 0,1$ et $X \rightsquigarrow t_{10}$ donc $p(\{X < 1,372\}) = 0,9$ et $t_{10;0,9} = 1,372$.
- $\nu = 20$, $\alpha = 0,05$ et $X \rightsquigarrow t_{20}$ donc $p(\{X < 1,725\}) = 0,95$ et $t_{20;0,95} = 1,725$.

Il existe une autre table (voir annexe C2) qui fournit pour ν et α donnés la valeur $t_{\nu,\alpha}$ telle que

$$p(\{-t_{\nu,\alpha} < X < t_{\nu,\alpha}\}) = 1 - \alpha.$$

Graphiquement, cette probabilité est donnée par la surface grisée de la Figure 5.18 :

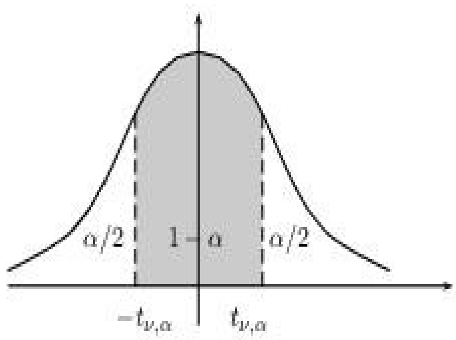

FIGURE 5.18

On remarque alors que $p(\{X < -t_{\nu,\alpha}\}) = p(\{X > t_{\nu,\alpha}\}) = \dfrac{1-(1-\alpha)}{2} = \dfrac{\alpha}{2}$.

Exemple 5.5.2 $\nu = 12$, $\alpha = 0,4$ et $t_{12;0,4} = 0,873$ donc $p(\{-0,873 < X < 0,873\}) = 0,6$ et $p(\{X < -0,873\}) = p(\{X > 0,873\}) = 0,2$.

5.6 Loi de Fischer-Snedecor

5.6.1 Définition

1. La **loi de Fischer-Snedecor** est une loi continue dépendant de deux paramètres notés ν_1 et ν_2, entiers naturels non nuls. La variable X distribuée selon cette loi prend toutes ses valeurs dans $\mathbb{R}^{+\star}$ ou dans $\mathbb{R}^+$.

   Si $Y \rightsquigarrow \chi^2_{\nu_1}$ et $Z \rightsquigarrow \chi^2_{\nu_2}$, Y et Z étant indépendantes, la variable $X = \dfrac{\frac{Y}{\nu_1}}{\frac{Z}{\nu_2}}$ suit une loi de Fischer-Snedecor. On note

$$\boxed{X \rightsquigarrow F_{(\nu_1,\nu_2)}}$$

La loi F de Fischer-Snedecor à (ν_1, ν_2) degrés de liberté est la loi de probabilité du rapport de deux variables de khi-deux indépendantes divisées par leurs nombres de degrés de liberté (ν_1 pour le numérateur, ν_2 pour le dénominateur).
Pour $\nu_1 = 1$, la loi F de Fischer-Snedecor à $(1,\nu_2)$ degrés de liberté est la loi de probabilité du carré d'une variable de Student à ν_2 degrés de liberté.

2. La densité de probabilité est, par définition :

$$f_{(\nu_1,\nu_2)}(x) = \nu_1^{\frac{\nu_1}{2}} \nu_2^{\frac{\nu_2}{2}} \frac{\Gamma\left(\dfrac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\dfrac{\nu_1}{2}\right)\Gamma\left(\dfrac{\nu_2}{2}\right)} \frac{x^{\frac{\nu_1}{2}-1}}{(\nu_1 x+\nu_2)^{\frac{\nu_1+\nu_2}{2}}} \text{ pour } x > 0,\ \nu_1 \text{ et } \nu_2 \in \mathbb{N}^\star.$$

Dans cette formule, Γ est la fonction Gamma d'Euler définie, lorsque la partie réelle de x est positive, par :

$$\Gamma(x) = \int_0^{+\infty} e^u u^{x-1} du.$$

La fonction $f_{(\nu_1,\nu_2)}$ est bien une densité de probabilité sur $]0; +\infty[$, car :

– ses valeurs sont positives,
– la fonction est intégrable et son intégrale est donnée par :

$$\int_0^{+\infty} f_{(\nu_1,\nu_2)}(x)dx = \nu_1^{\frac{\nu_1}{2}} \nu_2^{\frac{\nu_2}{2}} \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \int_0^{+\infty} \frac{x^{\frac{\nu_1}{2}-1}}{(\nu_1 x+\nu_2)^{\frac{\nu_1+\nu_2}{2}}} dx.$$

Pour calculer l'intégrale $I = \displaystyle\int_0^{+\infty} \frac{x^{\frac{\nu_1}{2}-1}}{(\nu_1 x+\nu_2)^{\frac{\nu_1+\nu_2}{2}}} dx$, on pose $t = \dfrac{\nu_1 x}{\nu_1 x + \nu_2} \Rightarrow dx = \dfrac{n_2}{n_1} \dfrac{dt}{1-t^2}$. De plus, $\nu_1 x + \nu_2 = \nu_2 \times \dfrac{1}{1-t}$ ce qui implique que lorsque $x = 0$, $t = 0$ et lorsque x tend vers l'infini, t tend vers 1. Par conséquent,

$$I = \int_0^1 \left(\frac{\nu_2}{\nu_1} \frac{t}{1-t}\right)^{\frac{\nu_1}{2}-1} \left(\frac{1-t}{\nu_2}\right)^{\frac{\nu_1+\nu_2}{2}} \frac{\nu_2}{\nu_1} \frac{dt}{(1-t)^2} = \nu_1^{-\frac{\nu_1}{2}} \nu_2^{-\frac{\nu_2}{2}} \int_0^1 t^{\frac{\nu_1}{2}-1} (1-t)^{\frac{\nu_2}{2}-1} dt.$$

Dans l'intégrale, on reconnaît la fonction Beta d'Euler définie, lorsque les parties réelles de x et de y sont positives, par :

$$B(x,y) = \int_0^1 u^{x-1}(1-u)^{y-1} du = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

donc $\displaystyle\int_0^1 t^{\frac{\nu_1}{2}-1}(1-t)^{\frac{\nu_2}{2}-1} dt = B\left(\frac{\nu_1}{2}, \frac{\nu_2}{2}\right)$ ce qui implique que

$$\int_0^{+\infty} f_{(\nu_1,\nu_2)}(x)dx = \nu_1^{\frac{\nu_1}{2}} \nu_2^{\frac{\nu_2}{2}} \frac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)} \nu_1^{-\frac{\nu_1}{2}} \nu_2^{-\frac{\nu_2}{2}} \frac{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)} = 1.$$

L'intégrale de $f_{(\nu_1,\nu_2)}$ est bien égale à 1, ce qui montre que $f_{(\nu_1,\nu_2)}$ est bien une densité de probabilité.

3. Si $X \rightsquigarrow F(\nu_1, \nu_2)$ la variable $\dfrac{1}{X} \rightsquigarrow F(\nu_2, \nu_1)$ donc $F(\nu_1, \nu_2, 1-\alpha) = \dfrac{1}{F(\nu_2, \nu_1, \alpha)}$.

5.6.2 Les courbes

On a représenté ci-dessus (Figure 5.19) la loi F de Fischer-Snedecor pour diverses valeurs de ν_1 et de ν_2.

5.6.3 Les moments

Soit $X \rightsquigarrow F(\nu_1, \nu_2)$.

– Pour $\nu_2 > 2$, l'espérance est définie par

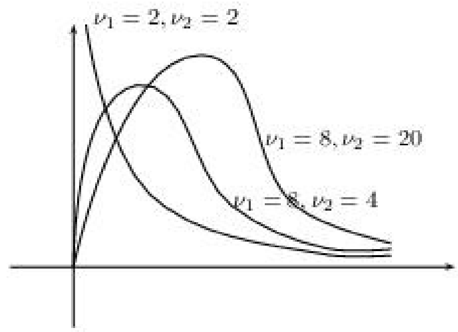

FIGURE 5.19

$$\boxed{E(X) = \frac{\nu_2}{\nu_2 - 2}}$$

Remarque 5.6.1 Pour $\nu_2 \leq 2$, l'espérance n'est pas déterminée.

– Pour $\nu_2 > 4$, la variance est définie par

$$\boxed{V(X) = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - 2)^2(\nu_2 - 4)}}$$

Remarque 5.6.2 Pour $\nu_2 \leq 4$, la variance n'est pas déterminée.

5.6.4 Les tables

Soit $X \rightsquigarrow F(\nu_1, \nu_2)$. La table (voir annexe D) fournit, pour $\alpha = 0,025$, pour ν_1 et ν_2 donnés, les valeurs $F_{\nu_1,\nu_2,1-\alpha}$ telles que $p(\{X < F_{\nu_1,\nu_2,1-\alpha}\}) = 1 - \alpha$. Cette table sert à la comparaison des variances de deux populations à partir de deux échantillons.

5.7 Exercices

Exercice 57 Une entreprise de transport a un parc total de 150 camions. On désigne par X la variable aléatoire qui à chaque camion choisi au hasard dans le parc, associe la distance qu'il a parcourue dans une journée (les distances sont mesurées en kilomètres). Un étude statistique permet d'admettre que cette variable aléatoire X suit une loi normale de moyenne 120 et d'écart-type 14.
Déterminer à 10^{-4} près la probabilité qu'un camion parcourt un jour donné une distance comprise entre 110 et 130 kilomètres (utiliser éventuellement une interpolation affine).

Exercice 58

1. *Statistique* - Avant d'accepter un contrat de livraison de véhicules, une société d'équipements automobiles établit une statistique de production journalière sur 100 jours. Le nombre de véhicules équipés journellement se répartit comme suit :

Production journalière de véhicules équipés	Nombre de jours	Production journalière de véhicules équipés	Nombre de jours
95	1	102	14
96	3	103	9
97	6	104	8
98	8	105	6
99	10	106	2
100	13	107	2
101	18	Total	100

Déterminer la valeur moyenne de la production journalière et une valeur approchée à 10^{-2} près de l'écart-type de cette production.

2. *Probabilités* - La production exigée par le contrat est de 100 véhicules équipés au moins par jour, pendant 100 jours de travail consécutif. À chaque journée on associe le nombre de véhicules équipés que l'on suppose indépendant du nombre obtenu chacun des autres jours. On définit ainsi une variable aléatoire X. On admet que la variable aléatoire discrète X peut être approchée par la loi normale de paramètres $m = 101$ et $\sigma = 2.59$. On note Y une variable aléatoire suivant la loi $\mathcal{N}(101; 2,59)$.
Calculer la probabilité de l'événement "le contrat est rempli", c'est-à-dire $p(\{Y \geq 99,5\})$.

Exercice 59 On jette 10 fois de suite une pièce de monnaie bien équilibrée en notant chaque fois le résultat, ce qui constitue une partie.

1. On note X la variable aléatoire qui à chaque partie associe le nombre de "face" obtenu.
   (a) Justifier que la loi de probabilité suivie par la variable X est une loi binomiale (on précisera les paramètres de cette loi).
   (b) Calculer la probabilité de l'événement E :
   "le nombre de 'face' est compris entre 3 et 6 (bornes incluses)".
2. On décide d'approcher la loi de variable aléatoire discrète X par la loi normale de paramètres m et σ.
   (a) Expliquer pourquoi on prend $m = 5$ et $\sigma = \sqrt{2,5}$.
   (b) On considère une variable aléatoire Y suivant une la loi $\mathcal{N}(5; \sqrt{2,5})$. En utilisant cette approximation, calculer la probabilité de l'événement :
   "le nombre de 'face' est compris entre 3 et 6 (bornes incluses)"
   c'est-à-dire $p\{(2,5 \leq Y \leq 6,5\})$.

Exercice 60 Une entreprise fabrique des imprimantes de modèle PRINT et constate que le nombre de commandes journalières définit une variable aléatoire Y dont la loi peut être approchée par la loi normale de paramètres $m = 80$ et $\sigma = 60$. On désigne par Z la variable aléatoire qui, à chaque mois de 25 jours ouvrables, associe le nombre d'unités du modèle PRINT demandé. Il y a indépendance entre les commandes journalières.

1. Montrer que la loi de Z peut être approchée par la loi normale $\mathcal{N}(2000, 300)$.
   L'entreprise a en stock, au début du mois, 2300 unités. Quelle est la probabilité qu'elle ne puisse satisfaire à la demande ?
2. On veut que la probabilité qu'elle ne puisse satisfaire à la demande soit inférieure à $0,05$. Quel doit être le nombre minimal d'unités que l'entreprise doit stocker en début de mois ?

Exercice 61 La variable aléatoire X suit une loi normale $\mathcal{N}(20, 5)$. Calculer

1. $p(\{X \leq 28\})$
2. $p(\{X \geq 28\})$
3. $p(\{X \geq 12\})$
4. $p(\{X \leq 12\})$
5. $p(\{12 \leq X \leq 28\})$

Exercice 62 Pour mesurer l'impact d'un régime amaigrissant, un club a choisi au hasard un échantillon de 5 individus avant le régime, et un échantillon de 5 autres individus après. Les masses corporelles se présentent ainsi :

Avant	84	92	72	91	84
Après	81	88	74	81	90

1. Déterminer un intervalle de confiance à 95% pour :
   (a) la masse corporelle moyenne avant le régime,
   (b) la masse corporelle moyenne après le régime,
   (c) la perte moyenne de masse corporelle durant le régime.
2. Tout compte fait, on a décidé qu'il aurait peut-être été plus adapté de peser les mêmes individus avant et après le régime. On a obtenu :

Individu	1	2	3	4	5
Avant	84	92	72	91	84
Après	81	88	74	81	90

   Sur la base de cet échantillon, déterminer un intervalle de confiance à 95% pour la perte moyenne de masse corporelle durant le régime. Conclusion ?

Exercice 63 Un laboratoire veut fabriquer des pilules se composant de deux substances A et B. Pour chaque pilule de la fabrication, on considère les masses a et b respectivement des 2 substances A et B qui la constituent. On désigne par X et Y respectivement les variables aléatoires qui associent à chaque pilule la masse a et la masse b des substances de cette pilule. On suppose que ces variables sont indépendantes et suivent des lois normales de moyennes respectives $m_X = 8,55$ mg et $m_Y = 5,20$ mg et de même écart-type $\sigma_X = \sigma_Y = 0,05$ mg.

1. Déterminer les probabilités $p(\{8,45 \leq X \leq 8,70\})$ et $p(\{5,07 \leq Y \leq 5,33\})$.
2. Les normes imposées pour la fabrication sont les suivantes : $8,45 \leq a \leq 8,70$ et $5,07 \leq b \leq 5,33$
   (a) Calculer le pourcentage de pilules qui seront hors normes à la sortie de la chaîne de fabrication.
   (b) En déduire que le procédé de fabrication ne peut être retenu si on veut que le pourcentage de pilules défectueuses ne dépasse pas 3%. On modifie alors la fabrication de la substance B. La moyenne de Y ne change pas mais son écart-type est modifié. Trouver la valeur minimum de ce nouvel écart-type pour que le pourcentage de pièces défecteuses soit inférieur à 3%.
3. (a) Déterminer la moyenne et l'écart-type de la variable aléatoire S qui associe à chaque pilule sa masse totale, les variables X et Y gardant leurs caractéristiques de la question 1.
   (b) On admet que S est encore une variable aléatoire normale dont les paramètres sont ceux calculés précédemment. Calculer $p(\{13,6 \leq S \leq 13,8\})$.
4. On assure le conditionnement des pilules par boîtes de 100 unités. Une boîte est constituée à partir d'un tirage au hasard dans un stock assez grand pour qu'on puisse estimer que les tirages successifs se font avec remises. On désigne par Z la variable aléatoire qui, à chaque boîte associe le nombre de pilules hors normes au sens de la question 2.(a). On pourra prendre pour probabilité p d'une pilule hors-norme $p = 0,01$.
   (a) Dans ces conditions, montrer que Z est une variable binomiale dont on précisera les paramètres.
   (b) Dire pourquoi on peut approcher cette variable par une loi de Poisson. En utilisant cette loi, donner une valeur approximative de $p(\{Z \geq 5\})$.
5. On désigne par U la variable aléatoire qui à chaque boîte associe le nombre de pilules dont la masse totale est supérieure à $13,8$. Là aussi, on peut supposer que U est une variable binomiale de paramètres n et p.
   (a) Calculer p.
   (b) Dire pourquoi on peut approcher U par une variable normale. À l'aide de cette approximation, donner une valeur approchée de $p(\{U \in \{70, 71, \ldots, 85\}\})$.

ANNEXE A - Fonction de répartition de la loi normale centrée réduite $\mathcal{N}(0,1)$.

Cette table donne $\Pi(x) = p(\{X \leq x\})$ pour $X \rightsquigarrow \mathcal{N}(0,1)$:

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	0,7389	0,7422	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	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

ANNEXE B - Probabilités individuelles de la loi du χ^2_ν,

Cette table donne les valeurs (quantiles) $\chi^2_{\nu,1-\alpha}$ telles que $p(\{\chi^2_\nu < \chi^2_{\nu,1-\alpha}\}) = 1-\alpha$:

ν \ $1-\alpha$	$0,005$	$0,010$	$0,025$	$0,050$	$0,100$	$0,900$	$0,950$	$0,975$	$0,990$	$0,995$
1	$0,0000393$	$0,000157$	$0,000982$	$0,00393$	$0,0158$	$2,71$	$3,84$	$5,02$	$6,63$	$7,88$
2	$0,0100$	$0,0201$	$0,0506$	$0,103$	$0,211$	$4,61$	$5,99$	$7,38$	$9,21$	$10,60$
3	$0,072$	$0,115$	$0,216$	$0,352$	$0,584$	$6,25$	$7,81$	$9,35$	$11,34$	$12,84$
4	$0,207$	$0,297$	$0,484$	$0,711$	$1,064$	$7,78$	$9,49$	$11,14$	$13,28$	$14,86$
5	$0,412$	$0,554$	$0,831$	$1,145$	$1,61$	$9,24$	$11,07$	$12,83$	$15,09$	$16,75$
6	$0,676$	$0,872$	$1,24$	$1,64$	$2,20$	$10,64$	$12,59$	$14,45$	$16,81$	$18,55$
7	$0,989$	$1,24$	$1,69$	$2,17$	$2,83$	$12,02$	$14,07$	$16,01$	$18,48$	$20,28$
8	$1,34$	$1,65$	$2,18$	$2,73$	$3,49$	$13,36$	$15,51$	$17,53$	$20,09$	$21,96$
9	$1,73$	$2,09$	$2,70$	$3,33$	$4,17$	$14,68$	$16,92$	$19,02$	$21,67$	$23,59$
10	$2,16$	$2,56$	$3,25$	$3,94$	$4,87$	$15,99$	$18,31$	$20,48$	$23,21$	$25,19$
11	$2,60$	$3,05$	$3,82$	$4,57$	$5,58$	$17,28$	$19,68$	$21,92$	$24,73$	$26,76$
12	$3,07$	$3,57$	$4,40$	$5,23$	$6,30$	$18,55$	$21,03$	$23,34$	$26,22$	$28,30$
13	$3,57$	$4,11$	$5,01$	$5,89$	$7,04$	$19,81$	$22,36$	$24,74$	$27,69$	$29,82$
14	$4,07$	$4,66$	$5,63$	$6,57$	$7,79$	$21,06$	$23,68$	$26,12$	$29,14$	$31,32$
15	$4,60$	$5,23$	$6,26$	$7,26$	$8,55$	$22,31$	$25,00$	$27,49$	$30,58$	$32,80$
16	$5,14$	$5,81$	$6,91$	$7,96$	$9,31$	$23,54$	$26,30$	$28,85$	$32,00$	$34,27$
17	$5,70$	$6,41$	$7,56$	$8,67$	$10,09$	$24,77$	$27,59$	$30,19$	$33,41$	$35,72$
18	$6,26$	$7,01$	$8,23$	$9,39$	$10,86$	$25,99$	$28,87$	$31,53$	$34,81$	$37,16$
19	$6,84$	$7,63$	$8,91$	$10,12$	$11,65$	$27,20$	$30,14$	$32,85$	$36,19$	$38,58$
20	$7,43$	$8,26$	$9,59$	$10,85$	$12,44$	$28,41$	$31,41$	$34,17$	$37,57$	$40,00$
21	$8,03$	$8,90$	$10,28$	$11,59$	$13,24$	$29,62$	$32,67$	$35,48$	$38,93$	$41,40$
22	$8,64$	$9,54$	$10,98$	$12,34$	$14,04$	$30,81$	$33,92$	$36,78$	$40,29$	$42,80$
23	$9,26$	$10,20$	$11,69$	$13,09$	$14,85$	$32,01$	$35,17$	$38,08$	$41,64$	$44,18$
24	$9,89$	$10,86$	$12,40$	$13,85$	$15,66$	$33,20$	$36,42$	$39,36$	$42,98$	$45,56$
25	$10,52$	$11,52$	$13,12$	$14,61$	$16,47$	$34,38$	$37,65$	$40,65$	$44,31$	$46,93$
26	$11,16$	$12,20$	$13,84$	$15,38$	$17,29$	$35,56$	$38,89$	$41,92$	$45,64$	$48,29$
27	$11,81$	$12,88$	$14,57$	$16,15$	$18,11$	$36,74$	$40,11$	$43,19$	$46,96$	$49,64$
28	$12,46$	$13,56$	$15,31$	$16,93$	$18,94$	$37,92$	$41,34$	$44,46$	$48,28$	$50,99$
29	$13,12$	$14,26$	$16,05$	$17,71$	$19,77$	$39,09$	$42,56$	$45,72$	$49,59$	$52,34$
30	$13,79$	$14,95$	$16,79$	$18,49$	$20,60$	$40,26$	$43,77$	$46,98$	$50,89$	$53,67$
40	$20,71$	$22,16$	$24,43$	$26,51$	$29,05$	$51,81$	$55,76$	$59,34$	$63,69$	$66,77$
50	$27,99$	$29,71$	$32,36$	$34,76$	$37,69$	$63,17$	$67,50$	$71,42$	$76,15$	$79,49$
60	$35,53$	$37,48$	$40,48$	$43,19$	$46,46$	$74,40$	$79,08$	$83,30$	$88,38$	$91,95$
70	$43,28$	$45,44$	$48,76$	$51,74$	$55,33$	$85,53$	$90,53$	$95,02$	$100,4$	$104,2$
80	$51,17$	$53,54$	$57,15$	$60,39$	$64,28$	$96,58$	$101,9$	$106,6$	$112,3$	$116,3$
90	$59,20$	$61,75$	$65,65$	$69,13$	$73,29$	$107,6$	$113,1$	$118,1$	$124,1$	$128,3$
100	$67,33$	$70,06$	$74,22$	$77,93$	$82,36$	$118,5$	$124,3$	$129,6$	$135,8$	$140,2$

ANNEXE C1 - Probabilités individuelles et cumulées de la loi de Student-Fischer $t_{\nu,\alpha}$,

Cette table donne les valeurs (quantiles) $t_{\nu,1-\alpha}$ telles que $p(\{-t_\nu < t_{\nu,1-\alpha}\}) = 1-\alpha$:

ν \ $1-\alpha$	0,55	0,60	0,65	0,70	0,75	0,80	0,85	0,90	0,95	0,975	0,99	0,995	0,9995
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	635,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,924
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,150
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,659
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,648
40	0,126	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,126	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,126	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,960	2,326	2,576	3,291

ANNEXE C2 - Probabilités individuelles et cumulées de la loi de Student-Fischer $t_{\nu,\alpha}$.

Cette table donne les valeurs $t_{\nu,\alpha}$ telles que $p(\{t_{\nu,\alpha} < t_\nu < +t_{\nu,\alpha}\}) = 1 - \alpha$:

ν \ α	0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,02	0,01	0,001
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657	636,619
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925	31,598
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841	12,929
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604	8,610
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032	6,869
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707	5,959
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499	5,408
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,387	1,860	2,306	2,896	3,355	5,041
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250	4,781
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169	4,587
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106	4,437
12	0,128	0,259	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055	4,318
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012	4,221
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977	4,140
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947	4,073
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,745	2,120	2,583	2,921	4,015
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898	3,965
18	0,l27	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878	3,922
19	0,l27	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861	3,883
20	0,l27	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845	3,850
21	0,l27	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831	3,819
22	0,l27	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819	3,792
23	0,l27	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807	3,767
24	0,l27	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797	3,745
25	0,l27	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787	3,725
26	0,l27	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779	3,707
27	0,l27	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771	3,690
28	0,l27	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763	3,674
29	0,l27	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756	3,649
30	0,l27	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750	3,656
40	0,l26	0,255	0,388	0,529	0,681	0,851	1,050	1,303	1,684	2,021	2,423	2,704	3,551
80	0,l26	0,254	0,387	0,527	0,679	0,848	1,046	1,296	1,671	2,000	2,390	2,660	3,460
120	0,l26	0,254	0,386	0,526	0,677	0,845	1,041	1,289	1,658	1,980	2,358	2,617	3,373
∞	0,l26	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,940	2,326	2,576	3,291

ANNEXE D - la loi de Fischer-Snedecor,

Cette table donne, pour $\alpha = 0,025$, pour ν_1 et ν_2 donnés, les valeurs $F_{\nu_1,\nu_2,1-\alpha}$ telles que $p(\{X < F_{\nu_1,\nu_2,1-\alpha}\}) = 1 - \alpha$,

ν_2 \ ν_1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	647,8	799,5	864,2	899,6	921,8	937,1	948,2	956,7	963,3	968,6	976,7	984,9	993,1	997,2	1001	1006	1010	1014	1018
2	38,51	39,00	39,17	39,25	39,30	39,33	39,36	39,37	39,39	39,40	39,41	39,43	39,45	39,46	39,46	39,47	39,48	39,49	39,50
3	17,44	16,04	15,44	15,10	14,88	14,73	14,62	14,54	14,47	14,42	14,34	14,25	14,17	14,12	14,08	14,04	13,99	13,95	13,90
4	12,22	10,65	9,98	9,60	9,36	9,20	9,07	8,98	8,90	8,84	8,75	8,66	8,56	8,51	8,46	8,41	8,36	8,31	8,26
5	10,01	8,43	7,76	7,39	7,15	6,98	6,85	6,76	6,68	6,62	6,52	6,43	6,33	6,28	6,23	6,18	6,12	6,07	6,02
6	8,81	7,26	6,60	6,23	5,99	5,82	5,70	5,60	5,52	5,46	5,37	5,27	5,17	5,12	5,07	5,01	4,96	4,90	4,85
7	8,07	6,54	5,89	5,52	5,29	5,12	4,99	4,90	4,82	4,76	4,67	4,57	4,47	4,42	4,36	4,31	4,25	4,20	4,14
8	7,57	6,06	5,42	5,05	4,82	4,65	4,53	4,43	4,36	4,30	4,20	4,10	4,00	3,95	3,89	3,84	3,78	3,73	3,67
9	7,21	5,71	5,08	4,72	4,48	4,32	4,20	4,10	4,03	3,96	3,87	3,77	3,67	3,61	3,56	3,51	3,45	3,39	3,33
10	6,94	5,46	4,83	4,47	4,24	4,07	3,95	3,85	3,78	3,72	3,62	3,52	3,42	3,37	3,31	3,26	3,20	3,14	3,08
11	6,72	5,26	4,63	4,28	4,04	3,88	3,76	3,66	3,59	3,53	3,43	3,33	3,23	3,17	3,12	3,06	3,00	2,94	2,88
12	6,55	5,10	4,47	4,12	3,89	3,73	3,61	3,51	3,44	3,37	3,28	3,18	3,07	3,02	2,96	2,91	2,85	2,79	2,72
13	6,41	4,97	4,35	4,00	3,77	3,60	3,48	3,39	3,31	3,25	3,15	3,05	2,95	2,89	2,84	2,78	2,72	2,66	2,60
14	6,30	4,86	4,24	3,89	3,66	3,50	3,38	3,29	3,21	3,15	3,05	2,95	2,84	2,79	2,73	2,67	2,61	2,55	2,49
15	6,20	4,77	4,15	3,80	3,58	3,41	3,29	3,20	3,12	3,06	2,96	2,86	2,76	2,70	2,64	2,59	2,52	2,46	2,40
16	6,12	4,69	4,08	3,73	3,50	3,34	3,22	3,12	3,05	2,99	2,89	2,79	2,68	2,63	2,57	2,51	2,45	2,38	2,32
17	6,04	4,62	4,01	3,66	3,44	3,28	3,16	3,06	2,98	2,92	2,82	2,72	2,62	2,56	2,50	2,44	2,38	2,32	2,25
18	5,98	4,56	3,95	3,61	3,38	3,22	3,10	3,01	2,93	2,87	2,77	2,67	2,56	2,50	2,44	2,38	2,32	2,26	2,19
19	5,90	4,51	3,90	3,56	3,33	3,17	3,05	2,96	2,88	2,82	2,72	2,62	2,51	2,45	2,39	2,33	2,27	2,20	2,13
20	5,87	4,46	3,86	3,51	3,29	3,13	3,01	2,91	2,84	2,77	2,68	2,57	2,46	2,41	2,35	2,29	2,22	2,16	2,09
21	5,83	4,42	3,82	3,48	3,25	3,09	2,97	2,87	2,80	2,73	2,64	2,53	2,42	2,37	2,31	2,25	2,18	2,11	2,04
22	5,79	4,38	3,78	3,44	3,22	3,05	2,93	2,84	2,76	2,70	2,60	2,50	2,39	2,33	2,27	2,21	2,14	2,08	2,00
23	5,75	4,35	3,75	3,41	3,18	3,02	2,90	2,81	2,73	2,67	2,57	2,47	2,36	2,30	2,24	2,18	2,11	2,04	1,97
24	5,72	4,32	3,72	3,38	3,15	2,99	2,87	2,78	2,70	2,64	2,54	2,44	2,33	2,27	2,21	2,15	2,08	2,01	1,94
25	5,69	4,29	3,69	3,35	3,13	2,97	2,85	2,75	2,68	2,61	2,51	2,41	2,30	2,24	2,18	2,12	2,05	1,98	1,91
26	5,66	4,27	3,67	3,33	3,10	2,94	2,82	2,73	2,65	2,59	2,49	2,39	2,28	2,22	2,16	2,09	2,03	1,95	1,88
27	5,63	4,24	3,65	3,31	3,08	2,92	2,80	2,71	2,63	2,57	2,47	2,36	2,25	2,19	2,13	2,07	2,00	1,93	1,85
28	5,61	4,22	3,63	3,29	3,06	2,90	2,78	2,69	2,61	2,55	2,45	2,34	2,23	2,17	2,11	2,05	1,98	1,91	1,83
29	5,59	4,20	3,61	3,27	3,04	2,88	2,76	2,67	2,59	2,53	2,43	2,32	2,21	2,15	2,09	2,03	1,96	1,89	1,81
30	5,57	4,18	3,59	3,25	3,03	2,87	2,75	2,65	2,57	2,51	2,41	2,31	2,20	2,14	2,07	2,01	1,94	1,87	1,79
40	5,42	4,05	3,46	3,13	2,90	2,74	2,62	2,53	2,45	2,39	2,29	2,18	2,07	2,01	1,94	1,88	1,80	1,72	1,64
60	5,29	3,93	3,34	3,01	2,79	2,63	2,51	2,41	2,33	2,27	2,17	2,06	1,94	1,88	1,82	1,74	1,67	1,58	1,48
120	5,15	3,80	3,23	2,89	2,67	2,52	2,39	2,30	2,22	2,16	2,05	1,94	1,82	1,76	1,69	1,61	1,53	1,43	1,31
∞	5,02	3,69	3,12	2,79	2,57	2,41	2,29	2,19	2,11	2,05	1,94	1,83	1,71	1,64	1,57	1,48	1,39	1,27	1,00