République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ MOHAMED KHIDER, BISKRA

FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la VIE

DÉPARTEMENT DE MATHÉMATIQUES

Tests statistiques

(Cours destiné aux étudiants de 1ère année Master)

réalisé par :

Djamel Meraghni et Abdelhakim Necir

(Professeurs à l'Universite Mohamed Khider, Biskra)

Préface

Ce polycopié est le fruit de plus de dix années d'efforts consentis par Messieurs Djamel Meraghni (2010-2018) et Abdelhakim Necir (2019-2024), professeurs à l'université Mohamed Khider de Biskra. Il représente un recueil de cours, travaux dirigés et travaux pratiques administrés aux étudiants de première année Master, au département de Mathématiques, sur le thème des tests d'hypothèses statistiques. Ces derniers sont d'une importance capitale des points de vue théorique et pratique. Avec l'estimation statistique, les tests sont un outil incontournable dans la modélisation statistique.

Introduction

Lors de l'estimation d'un (ou plusieurs) paramètre, on utilise des résultats échantillonaux afin d'approximer la valeur exacte inconnue du paramètre sans aucune idée préalable sur celle-ci. Mais dans la majorité des applications, on peut avoir une certaine idée sur le paramètre de telle sorte qu'on puisse formuler une hypothèse concernant sa vraie valeur (toujours inconnue). Les résultats échantillonaux permettent alors de confirmer ou d'infirmer cette hypothèse.

L'hypothèse peut aussi être relative à la distribution de la population mère (binomiale, Poisson, normale,...) ou à la relation pouvant exister entre deux ou plusieurs populations (égalité de paramètres, indépendances,...).

Un test statistique est une procédure de décision permettant de trancher entre deux hypothèses après l'observation d'un échantillon. Les deux hypothèses sont généralement appelées hypothsèse nulle et hypothèse alternative.

Soient Θ_0 et Θ_1 deux sous ensembles disjoints de Θ. On veut savoir si la valeur du paramètre θ est dans Θ_0 où Θ_1, alors on réalise le test dont les hypothèses sont :

$$\begin{cases} H_0 : & \theta \in \Theta_0 \\ H_1 : & \theta \in \Theta_1 \end{cases}$$

Puisque les valeurs de θ qui n'appartiennent pas à $\Theta_0 \cup \Theta_1$ ne sont pas envisagées, on peut admettre que $\Theta = \Theta_0 \cup \Theta_1$.

Les tests sont classés selon les hypothèses formulées ou en d'autres termes selon le but à atteindre. On distingue différentes catégories de tests, parmi lesquelles :

Test de conformité :

Ils servent à vérifier si un paramètre (moyenne, variance, proportion,...) d'une population de distribution, connue est égal à une certaine valeur théorique appelée valeur hypothétique du paramètre.

Tests de comparaison (ou égalité ou homogénéité) :

Ils servent à vérifier si les paramètres de deux (ou plusieurs) populations de distributions, connues sont égaux. Par exemple, test d'égalité des moyennes.

Tests d'ajustement :

Ils servent à vérifier si un échantillon provient d'une population de distribution donnée.

Tests d'indépendance :

Ils servent à vérifier si deux ou plusieurs populations sont indépendantes.

Tests paramétriques et tests nonparamétriques :

Un test est dit paramétrique si la population mère est de distribution connue ; l'objet du test est alors de vérifier si certaines hypothèses relatives à un ou plusieurs paramètres de cette distribution. Les tests de conformité et de comparaison sont des tests paramétriques.

Chapitre 1

Hypothèses simples et hypothèses multiples ou composites

Pour les tests paramétriques on distingue : hypothèses simples et hypothèses multiples.

Une hypothèse H est dite simple si elle de type "$\theta = \theta_0$" où $\theta_0 \in \Theta$.

Une hypothèse H est dite multiple si elle est du type "$\theta \in A$" où A est une partie de Θ ayant plus d'un élément.

1.0.1 Erreurs et risques :

Les deux hypothèses H_0 et H_1 sont telles que une et une seule vraie. Lors de la prise de décision qui aboutera à choisir H_0 ou H_1 quatre situations peuvent être envisagées :

- accepter H_0 et elle est vrais.
- rejeter H_0 et elle fausse.
- accepter H_0 alors qu'elle est fausse
- rejeter H_0 alors qu'elle est vraie.

Dans les deux premiers cas, la décision prise est bonne mais les deux derniers cas, elle est erronée.

L'erreur qui consiste à rejeter une hypothèse vrais appelée *erreur de première espèce* et sa probabilité appelée *risque de première espèce.* On le note α.

L'erreur commise en acceptant une hypothèse fausse est appelée *erreur de deuxième espèce* et sa probabilité appelée *risque de deuxième espèce.* On le note β.

On a donc :

$$\alpha = \mathbf{P}\left[\text{rejeter } H_0 \mid H_0 \text{ est vraie}\right] = \mathbf{P}\left(H_1 \mid H_0\right),$$

et

$$\beta = \mathbf{P}\left[\text{accepter } H_0 \mid H_0 \text{ est fausse}\right] = \mathbf{P}\left(H_0 \mid H_1\right).$$

Ceci est résumé dans le tableau Tab.1 et la figure Fig.1 suivants :

		Vérité	
		H_0	H_1
Décision	H_0	$1-\alpha$	β
	H_1	α	$1-\beta$

Tab.1

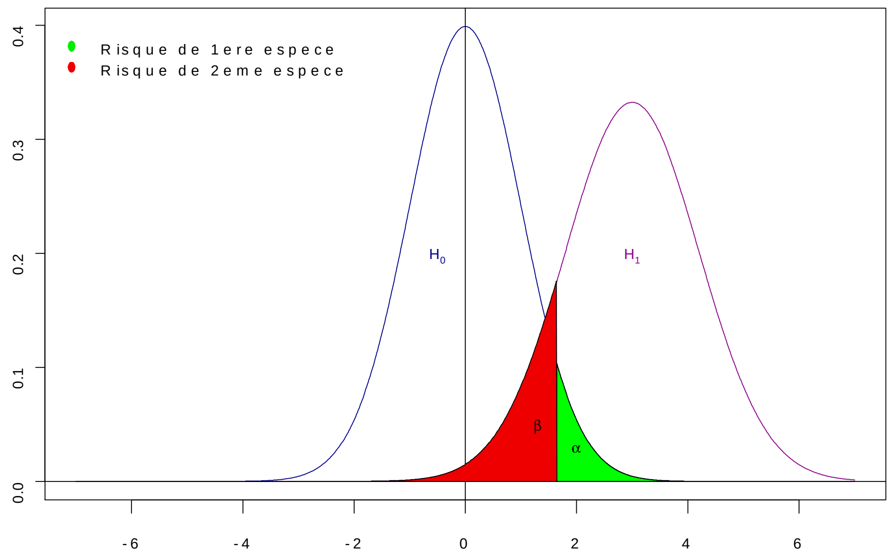

Fig.1

On remarque à la vue du graphique, que plus la différence entre les deux populations est faible, plus le risque β est important, donc plus il sera difficile dans ces conditions de conclure à une différence significative. Cette aptitude à pouvoir conclure en faveur d'une différence significative est appelée puissance du test et est représentée par la quantité $1-\beta$.

En pratique, on fixe une *limite supérieure* généralement égale à **0.10 ; 0.05 ; 0.01**, au risque de première espèce. Cette limite est appelée niveau ou *seuil de signification* du test. L'idéal est de construire un test pour lequel les valeurs de α et β sont toutes en sens petites. Ceci n'est pas possible car α et β en sens contraires. Si on diminue α (ou augmente $1-\alpha$) on est conduit à n'abandonner H_0 que très rarement et donc à l'accepter presque tout le temps (même à tort) d'où on augmente β.

Comme majorant, un niveau de signification *n'est pas unique.*

Exemple 2 : Une certaine personne se présente chez son médecin qui doit décider si celle-ci est atteinte d'une certaine maladie ou non. Le médecin désigne par H_0

l'hypothèse selon laquelle la personne est saine et par H_1 l'hypothèse selon laquelle la personne est atteinte. Le médecin dispose de données statistiques (résultats d'examens médicaux) qui l'aideront à prendre la décision. Si le médecin conclut que la personne est atteinte alors qu'elle est saine, il commet une erreur de première espèce. Et si au contraire il commet une erreur de deuxième espèce. Il est évident que les effets de ces deux erreurs sont fondamentalement différents.

Puissance :

La puissance d'un test est la probabilité de rejeter l'hypothèse nulle H_0 quand l'alternative H_1 est vraie. On la note par

$$\pi := \mathbf{P}\left[\text{rejeter } H_0 \mid H_1 \text{ est vraie}\right] = 1 - \beta.$$

Lorsque H_1 est composite, la puissance est variable sur Θ_1. De même lorsque H_0 est composite, le risque de première espèce est variable sur Θ_0. On définit alors une fonction sur l'ensemble Θ qu'on appelle fonction puissance

$$\pi\left(\theta\right) := \mathbf{P}_\theta\left[\text{rejeter } H_0\right], \ \theta \in \Theta.$$

- Si $\theta \in \Theta_0$, $\pi\left(\theta\right) = \alpha\left(\theta\right)$ c'est le risque de première espèce.
- Si $\theta \in \Theta_1$, $\pi\left(\theta\right) = 1 - \beta\left(\theta\right)$ c'est la puissance du test.

Variable de décision

La statistique qui apporte le plus de renseignement sur le problème posé est appelée *variable de décision* ou *statistique du test*. La loi de probabilité doit être différente selon que H_0 ou H_1, sinon elle ne servait à rien.

Région de rejet et région critique

La *région de rejet* d'un test est l'ensemble des points $(X_1, ..., X_n)$ de $\mathbb{R}^n$ pour lequel l'hypothèse nulle H_0 est écartée au profit de l'hypothèse alternative H_1. On appelle aussi *région critique* du test et on la note généralement par W. Elle est définie par la relation :

$$\mathbf{P}\left(W \mid H_0\right) = \alpha \tag{1.1}$$

Le complémentaire de la région critique est appelée région d'acceptation du test. Elle est notée par $\overline{W}$ et est définie par :

$$\mathbf{P}\left(\overline{W} \mid H_0\right) = 1 - \alpha$$

Remarque 1 :

1. L'indicatrice de W est appelée fonction critique du test. On note par

$$\delta\left(x_1, ..., x_n\right) := \mathbf{1}_W = \begin{cases} 1 & \text{si } \left(x_1, ..., x_n\right) \in W \\ 0 & \text{si } \left(x_1, ..., x_n\right) \notin W \end{cases}.$$

2. Le construction d'un test est en fait la détermination de la région critique W. D'où en vertu de la relation (1), la nécessité de connaitre la loi de probabilité de la variable de décision sous l'hypothèse H_0.
3. Puisque la puissance est $\pi = \mathbf{P}\left[W \mid H_1\right]$ alors, pour son calcul, il est nécessaire de connaitre la loi de probabilité de la variable de décision sous l'hypothèse H_1.

4. On a

$$\pi(\theta) = \mathbf{P}(W) = \mathbf{P}[\delta(X_1, ..., X_n) = 1]$$
$$= \mathbf{E}[\delta(X_1, ..., X_n)].$$

5. Certains auteurs s'intéressent au plus petit niveau de signification. Ils s'appellent dimension du test. C'est le risque de première espèce maximum

$$\alpha := \sup_{\theta \in \Theta_0} \alpha(\theta) = \sup_{\theta \in \Theta_0} \pi(\theta).$$

Exemple 3 :

Soit $X \rightsquigarrow U(0, \theta)$, $\theta > 0$. On désire tester les deux hypothèses :

$$\begin{cases} H_0 : & 3 \leq \theta \leq 4 \\ H_1 : & \theta < 3 \text{ ou } \theta > 4. \end{cases}$$

On a $\Theta =]0, +\infty[$, $\Theta_0 = [3, 4]$ et $\Theta_0 =]0, 3[\cup]4, +\infty[$. On suppose que la région d'acceptation du test est

$$\overline{W} = \{(x_1, ..., x_n) \in \mathbb{R}^n : \ 2.9 \leq \max\{x_1, ..., x_n\} \leq 4\}.$$

On rappelle que $Y_n := \max\{x_1, ..., x_n\}$ est l'estimateur de maximum de vraisemblance de θ et que sa fonction de répartition est définie sur $\mathbb{R}$ par :

$$F_{Y_n}(x) := \begin{cases} 0 & \text{si } x < 0 \\ (x/\theta)^n & \text{si } 0 \leq x \leq \theta \\ 1 & \text{si } x > \theta \end{cases}$$

La fonction puissance du test est définie sur $]0,+\infty[$ par

$$\begin{aligned}\pi(\theta) &= \mathbf{P}(W) = \mathbf{P}(Y_n < 2.9) + \mathbf{P}(Y_n > 4) \\ &= F_{Y_n}(2.9) + 1 - F_{Y_n}(4).\end{aligned}$$

– Si $\theta > 4$ alors $\pi(\theta) = (2.9/\theta)^n + 1 - (4/\theta)^n$

– Si $3 \leq \theta \leq 4$ alors $\pi(\theta) = (2.9/\theta)^n$

– Si $\theta < 3$ alors $\pi(\theta) = \begin{cases} 1 & \text{si } \theta < 2.9 \\ (2.9/\theta)^n & \text{si } 2.9 \leq \theta < 3 \end{cases}$

D'où

$$\pi(\theta) = \begin{cases} 1 & \text{si } \theta < 2.9 \\ (2.9/\theta)^n & \text{si } 2.9 \leq \theta < 4 \\ (2.9/\theta)^n + 1 - (4/\theta)^n & \text{si } \theta > 4 \end{cases}$$

La dimension du test est

$$\alpha = \sup_{3 \leq \theta \leq 4} \pi(\theta) = \pi(3) = (2.9/3)^n .$$

Pour un échantillon de taille 68 on trouve $\alpha \simeq 0.10$.

Chapitre 2

Test uniformément le plus puissant

Soit à tester l'hypothèse $H_0 : \theta \in \Theta_0$ contre l'hypothèse $H_1 : \theta \in \Theta_1$ où Θ_1 est composite.

Pour trancher entre H_0 et H_1, à niveau de signification α fixé, on doit trouver un test pour lequel la puissance est la plus grande possible. Quand il existe, un tel test est appelé upp. Sa puissance est supérieure à celle de tout autre test (de même niveau de signification α).

Un test δ^*, de niveau de signification α, est upp si et seulement si :

$$\pi(\theta, \delta^*) \geq \pi(\theta, \delta)$$

pour tout $\theta \in \Theta_1$ et pour tout autre test δ (de même niveau de signification α).

La région critique et la variable de décision correspondantes sont dites région critique et variable de décision optimals.

Dans le cas où H_1 est simple, on parle de test le plus puissant (pp).

Test sans biais :

Un test δ, niveau de signification α, est dit sans biais si $1-\beta \geq \alpha$; en d'autres termes

$$\pi(\theta) \geq \alpha \text{ pour tout } \theta \in \Theta_1.$$

Tests convergents ou consistants :

Un test est dit convergent si sa puissance tend vers 1 :

$$1-\beta \rightarrow 1, \text{ quand } n \rightarrow \infty.$$

Tests entre hypothèses simples :

Soit X une variable aléatoire de densité de probabilité f_θ où θ est un paramètre inconnu.

Il s'agit de tester

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta = \theta_1. \end{cases} \tag{2.1}$$

Soit

$$L_\theta(x_1, ..., x_n) := \prod_{i=1}^{n} f_\theta(x_i),$$

la fonction de vraisemblance (ou bien la densité) de l' échantillon $(X_1, ..., X_n)$ de X.

On pose

$$L_i(x_1, ..., x_n) := L_{\theta_i}(x_1, ..., x_n), \ i = 0, 1.$$

Le rapport

$$\frac{L_1(x_1, ..., x_n)}{L_0(x_1, ..., x_n)},$$

est appelé rapport de vraisemblance.

Lemme de Neyman-Pearson (cas continu)

On suppose ici que la v.a X est continue. Un test δ_k est le test qui rejette H_0 au niveau de signification α, si et seulement si le rapport de vraisemblance est au moins k, où $k = k(\alpha) \geq 0$.

Si δ est un autre test tel que $\alpha(\delta) \leq \alpha(\delta_k)$, alors

$$\pi(\theta; \delta_k) \geq \pi(\theta; \delta),$$

c'est à dire δ_k est le plus puissant. En d'autres termes, la région critique optimale est

$$W_k := \left\{ (x_1, ..., x_n) \in \mathbb{R}^n : \frac{L_1(x_1, ..., x_n)}{L_0(x_1, ..., x_n)} \geq k \right\},$$

où k est telle que $\mathbf{P}_0(W_k) = \alpha$, où $\mathbf{P}_i(A) := P(A \mid H_i)$, $i = 0, 1$.

Le test δ_k est alors défini comme suit :

$$\delta_k(x_1, ..., x_n) = \begin{cases} 1 & \text{si } \dfrac{L_1(x_1, ..., x_n)}{L_0(x_1, ..., x_n)} \geq k \\ 0 & \text{si } \dfrac{L_1(x_1, ..., x_n)}{L_0(x_1, ..., x_n)} < k \end{cases}$$

où est la solution de l'équation

$$\mathbf{P}_0 \left\{ \frac{L_1(X_1, ..., X_n)}{L_0(X_1, ..., X_n)} \geq k \right\} = \alpha.$$

Autrement dit

$$\alpha(\delta_k) := \mathbf{E}_0[\delta_k(X_1, ..., X_n)] = \mathbf{P}_0(W) = \alpha,$$

où $\mathbf{E}_i[X] := \mathbf{E}[X \mid H_i]$, $i = 0, 1$.

Exemple 4 : Soit X une population normale d'espérance inconnue μ et de variance 1. On veut tester l'hypothèse $H_0 : \mu = 0$ contre l'hypothèse $H_1 : \mu = 1$. Pour cela en prélève un échantillon de taille 9. Quel est le test le plus puissant au niveau de signification 0.05?

Solution : Le rapport de vraisemblance

$$\frac{L_1(x_1,...,x_n)}{L_0(x_1,...,x_n)} = \frac{\prod_{i=1}^{9} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x_i-1)^2\right\}}{\prod_{i=1}^{9} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}{x_i}^2\right\}}$$
$$= \exp\left\{9\left(\overline{x}-\frac{1}{2}\right)\right\},$$

où $\overline{x} = \sum_{i=1}^{9} x_i/9$ est la moyenne empirique. La région de rejet (critique) est donc

$$W = \left\{(x_1,...,x_n) \in \mathbb{R}^9 : \exp\left\{9\left(\overline{x}-\frac{1}{2}\right)\right\} \geq k\right\}$$
$$= \left\{(x_1,...,x_n) \in \mathbb{R}^9 : \overline{x} \geq \frac{1}{2} + \frac{\log k}{9} =: k'\right\},$$

où la constante k' est telle que :

$$0.05 = P(W \mid \mu = 0) = P_0(W) = P_0\left(\overline{X} \geq k'\right).$$

Or sous H_0, $\overline{X} \rightsquigarrow \mathcal{N}(0,1/9)$, ainsi $Z := 3\overline{X} \rightsquigarrow \mathcal{N}(0,1)$, d'où $P(Z \geq 3k') = 0.05$, en d'autres termes $\Phi(3k') = 0.95$ ou encore $k' = \frac{1}{3}\Phi^{-1}(0.95)$, où $\Phi^{-1}(\alpha)$ désigne la fonction de quantile d'ordre $0 < \alpha < 1$. De la table statistique de la loi normale (Gauss) en tire : $\Phi^{-1}(0.95) = 1.64$, par conséquent $k' = 1.64/3 = 0.54$. La région critique optimale est donc :

$$W = \left\{(x_1,...,x_n) \in \mathbb{R}^9 : \overline{x} \geq 0.54\right\}.$$

Le test optimal (le plus puissant) est par conséquent :

$$\delta\left(x_1,...,x_n\right)=\begin{cases}1 & \text{si } \overline{x}\geq 0.54\\ \\ 0 & \text{si } \overline{x}<0.54\end{cases} \tag{2.2}$$

Calculant la puissance $1-\beta$ du test δ :

$$\begin{aligned}1-\beta &= \mathbf{P}\left(W\mid\mu=1\right)=\mathbf{P}\left(\overline{X}\geq 0.54\mid\mu=1\right)\\ &= 1-\mathbf{P}_1\left(\overline{X}<0.54\right)\\ &= 1-\mathbf{P}_1\left(3\left(\overline{X}-1\right)<3\left(0.54-1\right)\right)\\ &= 1-\mathbf{P}\left(Z^*<-1.38\right),\ \text{où } Z^*\rightsquigarrow\mathcal{N}\left(0,1\right)\\ &= \mathbf{P}\left(Z^*<1.38\right)=0.91.\end{aligned}$$

Ainsi le risque de deuxième espèce est $\beta=1-0.91=0.09$.

Lemme de Neyman-Pearson (cas discret) :

Il arrive que pour certaines valeurs de α, il n'existe pas de constante k vérifiant l'équation $\alpha\left(\delta_k\right)=k$; ce qui ce passe surtout dans le cas où X est discrète. Dans ce cas, on modifie δ_k en définissant le δ_k^* (qui sera le plus puissant) :

$$\delta_k^*\left(x_1,...,x_n\right)=\begin{cases}1 & \text{si } \dfrac{L_1\left(x_1,...,x_n\right)}{L_0\left(x_1,...,x_n\right)}>k\\ p & \text{si } \dfrac{L_1\left(x_1,...,x_n\right)}{L_0\left(x_1,...,x_n\right)}=k\\ 0 & \text{si } \dfrac{L_1\left(x_1,...,x_n\right)}{L_0\left(x_1,...,x_n\right)}<k\end{cases}$$

où k et $0<p<1$ sont deux constantes définies, sous H_0, par l'équation :

$$\alpha\left(\delta_k^*\right)=\mathbf{P}_0\left\{\frac{L_1\left(X_1,...,X_n\right)}{L_0\left(X_1,...,X_n\right)}>k\right\}+p\mathbf{P}_0\left\{\frac{L_1\left(X_1,...,X_n\right)}{L_0\left(X_1,...,X_n\right)}=k\right\}=\alpha.$$

On note que δ_k est appelé test du rapport de vraisemblance et δ_k^* est appelé test du rapport de vraisemblance *randomisé*.

Exemple 5 : On prélève un échantillon de taille 8, d'une population X de Poisson de paramètre $\lambda > 0$, pour tester l'hypothèse $H_0 : \lambda = 1$ contre $H_0 : \lambda = 2$ au niveau de signification $\alpha = 0.1$. Déterminer le test le plus puissant et quelle est sa puissance ?

Solution : Rappelons que la loi de Poisson, de paramètre $\lambda > 0$, est définie par sa fonction de masse :

$$\mathbf{P}_\lambda (X = x) = \frac{\lambda^x}{x!} e^{-\lambda}, \ x = 0, 1, 2, ...$$

et sa fonction de répartition

$$\mathbf{P} (X \leq x) = \sum_{k=0}^{x} \mathbf{P}_\lambda (X = k) = e^{-\lambda} \sum_{k=0}^{x} \frac{\lambda^k}{k!}, \ x = 0, 1, 2, ...$$

Le rapport de vraisemblance qui correspond à ce test est :

$$\frac{L_1 (x_1, ..., x_8)}{L_0 (x_1, ..., x_8)} = \frac{\prod_{i=1}^{8} \mathbf{P}_{\lambda=2} (X = x_i)}{\prod_{i=1}^{8} \mathbf{P}_{\lambda=1} (X = x_i)} = \frac{\prod_{i=1}^{8} \frac{2^{x_i}}{x_i!} e^{-2}}{\prod_{i=1}^{8} \frac{1^{x_i}}{x_i!} e^{-1}} = e^{-8} 2^s,$$

où $s := x_1 + ... + x_8$. Ainsi le test statistique le plus puissant est :

$$\delta = \begin{cases} 1 & \text{si } e^{-8} 2^s > k \\ p & \text{si } e^{-8} 2^s = k \\ 0 & \text{si } e^{-8} 2^s < k \end{cases}$$

où k et $0 < p < 1$ sont deux constantes définies sous H_0 par l'équation :

$$\mathbf{P}_0 \left(e^{-8} 2^S > k \right) + p \mathbf{P}_0 \left(e^{-8} 2^S = k \right) = 0.1,$$

où $S := X_1 + ... + X_8$. En d'autres termes

$$\delta = \begin{cases} 1 & \text{si } s > c \\ p & \text{si } s = c \\ 0 & \text{si } s < c \end{cases}$$

où c et $0 < p < 1$ sont deux constantes définies sous H_0 par l'équation :

$$\mathbf{P}_0 (S > c) + p\mathbf{P}_0 (S = c) = 0.1.$$

Cette équation peut être réécrite comme suit :

$$\mathbf{P}_0 (S \leq c) = 0.90 + p\mathbf{P}_0 (S = c) . \tag{2.3}$$

Ce que signifie que c est le quantile d'ordre $\alpha_c := 0.90 + p\mathbf{P}_0 (S = c)$, en d'autres termes $c = F_S^{-1} (\alpha) := \inf \{x : F_S (x) \geq \alpha_c\}$, où $F_S (x) := \mathbf{P}_0 (S \leq x)$ et Sous H_0, l'échantillon de taille 8 provient d'une loi de Poisson de paramètre $\lambda = 1$, donc S est suit aussi une loi de Poisson de paramètre $8\lambda = 8$. En remarque que la quantité $0 < p\mathbf{P}_0 (S = c) < 1$, donc $0.90 < \alpha_c < 1$. De la table statistique de la loi de Poisson, on remarque que "la petite valeur de x" vérifiant $F_S (x) > 0.90$ est $x = 12$, qui correspond à $\mathbf{P}_0 (S \leq 12) = 0.93$, donc $c = 12$. On en déduit que

$$\begin{aligned} \mathbf{P}_0 (S = 12) &= \mathbf{P}_0 (S \leq 12) - \mathbf{P}_0 (S \leq 11) \\ &= 0.93 - 0.88 = 0.05, \end{aligned}$$

ce qui implique, de l'équation (2.3) , que

$$p = \frac{0.93 - 0.90}{0.05} = 0.60 = 60\%.$$

En d'autres termes si $s = 12$ on rejette H_0 avec une probabilité de 60%, et si $S > 12$ en rejette H_0 à 99.90%. Ainsi le test le plus puissant est :

$$\delta = \delta(x_1, ..., x_8) = \begin{cases} 1 & \text{si } s > 12 \\ 0.60 & \text{si } s = 12 \\ 0 & \text{si } s < 12 \end{cases}$$

Il est clair que

$$\begin{aligned} &\mathbf{E}_0 [\delta(X_1, ..., X_8)] \\ &= 1 \times \mathbf{P}_0 (S > 12) + 0.60 \times \mathbf{P}_0 (S = 12) + 0 \times \mathbf{P}_0 (S < 12) \\ &= 1 - 0.93 + 0.60 \times 0.05 = 0.1 = \alpha. \end{aligned}$$

La p-valeur (the p-value en anglais)

D'après le Lemme de Neyman-Pearson, ci-dessus. la statistique de décision (test) entre deux hypothèses simples, est basée sur le rapport de vraisemblance

$$T = T(X_1, ..., X_n) := \frac{L_1(X_1, .., X_n)}{L_0(X_1, .., X_n)} \text{ (ne dépend pas de } \theta).$$

Pour un seuil (niveau) α fixé, la région critique est définie par $W = \{T \geq k_\alpha\}$, où k_α est telle que $\mathbf{P}_0(T \geq k_\alpha) = \alpha$. Etant donné une observation $(x_1, ..., x_n)$ de l'échantillon $(X_1, .., X_n)$, nous disposons donc d'une valeur observée de T, notée

$$T_{obs} := T(x_1, ..., x_n)$$

Nous allons affaire à une valeur cruciale appelée la *p-valeur* (en anglais the *p-value*) définie par la probabilité

$$p := \mathbf{P}_0(T \geq T_{obs}).$$

Cette probabilité est fréquemment utilisée comme une règle de décision pour les tests statistiques implémentés dans les logiciels de programmations, tels que le R, Matlab, S, SAS,... Les résultats des tests statistiques sont exprimés en général par le T_{obs} et la p-valeur.

Supposons que

$$\mathbf{P}_0 \left(T \geq T_{obs} \right) < \alpha = \mathbf{P}_0 \left(T \geq k \right)$$

se qui implique $T_{obs} > k$, et par conséquent $(x_1, ..., x_n) \in W$ conduisant à la rejection de l'hypothèse nulle H_0. Le cas contraire conduit évidement à garder l'hypothèse nulle H_0 (voir la Fig.2).

A titre d'application, considérons l'Exemple 1 ci-dessus en prenant l'échantillon de taille 9 issu d'une loi normale $\mathcal{N}(0, 1)$:

$$-1.64; 0.13; 0.04; 1.29; 1.83; 0.49; 1.45; -1.08; -0.23$$

On veut tester l'hypothèse $H_0 : \mu = 0$ contre l'hypothèse $H_1 : \mu = 1$. La statistique de décision dans cette exemple est $T = \overline{X}$ et

$$\begin{aligned} T_{obs} &= \frac{1}{9} \left(-1.64 + 0.13 + 0.04 + 1.29 + 1.83 + 0.49 + 1.45 - 1.08 - 0.23 \right) \\ &= 0.25. \end{aligned}$$

La p-valeur qui correspond à cet échantillon (observé) est

$$\begin{aligned} p &= \mathbf{P}_0 \left(T \geq 0.25 \right) = \mathbf{P} \left(Z \geq 3 \left(0.25 \right) \right) \\ &= 1 - \mathbf{P} \left(Z \leq 0.75 \right) = 1 - \Phi \left(0.75 \right) = 1 - 0.77 = 0.23. \end{aligned}$$

Nous avons $p = 0.23 > 0.05 = \alpha$, ceci nous conduit à ne pas rejeter H_0. En utilisant la règle du test δ donnée dans (2.2) en remarque que $T_{obs} = \overline{x} = 0.25 < 0.54$ donc

$\delta = 0$, ce qui conduit aussi à garder H_0.

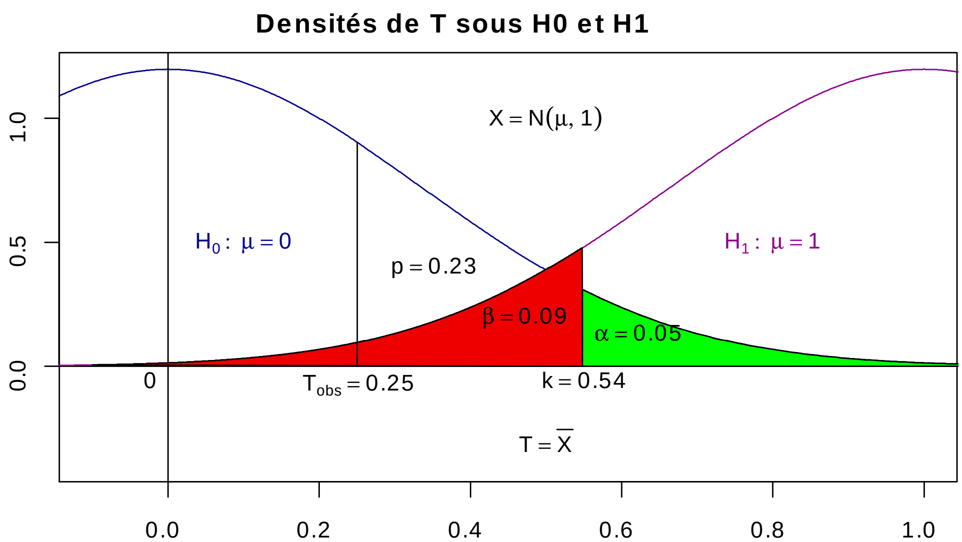

Fig. 2

Preuve du Lemme de Neyman-Pearson :

Considérons le cas continu. Le cas discret se traite de la même façon sauf il faut changer les signes $\int$ par les signes $\sum$. Nous allons montrer que, pour tout test statistique δ qui correspond au hypothèses (2.1) ci-dessus, tel que $\alpha(\delta) \leq \alpha(\delta_k)$, on a $\pi(\theta_1; \delta_k) \geq \pi(\theta_1; \delta)$. En effet, pour $(x_1, ..., x_n) \in \mathbb{R}^n$, nous examinons deux cas :

– Si $(x_1, ..., x_n) \in W_k$ alors $\delta_k := \delta_k(x_1, ..., x_n) = 1$ et $L_1 - kL_0 \geq 0$. Comme $\delta \leq 1$, alors

$$\delta \times [L_1 - kL_0] \leq 1 \times [L_1 - kL_0] = \delta_k \times [L_1 - kL_0].$$

– Si $(x_1, ..., x_n) \notin W_k$ alors $\delta_k(x_1, ..., x_n) = 0$ et $L_1 - kL_0 < 0$. Comme $\delta \geq 0$, alors

$$\delta \times [L_1 - kL_0] \leq 0 \times [L_1 - kL_0] = \delta_k \times [L_1 - kL_0].$$

D'où pour tout $(x_1, ..., x_n) \in \mathbb{R}^n$:

$$\delta \times [L_1 - kL_0] \leq \delta_k \times [L_1 - kL_0].$$

Ceci implique que

$$\int \delta \times [L_1 - kL_0]\, dx_1...dx_n \leq \int \delta_k \times [L_1 - kL_0]\, dx_1...dx_n.$$

En d'autres termes

$$\int \delta \times L_1 dx_1...dx_n - k \int \delta \times L_0 dx_1...dx_n$$
$$\leq \int \delta_k \times L_1 dx_1...dx_n - k \int \delta_k \times L_0 dx_1...dx_n,$$

et par conséquent

$$\mathbf{E}_1[\delta] - \mathbf{E}_1[\delta_k] \leq k[\mathbf{E}_0[\delta] - \mathbf{E}_0[\delta_k]].$$

Or

$$\mathbf{E}_0[\delta] = \pi(\theta_0; \delta) = \alpha(\delta)$$

et

$$\mathbf{E}_0[\delta_k] = \pi(\theta_0; \delta_k) = \alpha(\delta_k),$$

et puisque $\alpha(\delta) \leq \alpha(\delta_k) = \alpha$, alors

$$\mathbf{E}_1[\delta] - \mathbf{E}_1[\delta_k] \leq 0,$$

ainsi $\pi(\theta_1; \delta) \leq \pi(\theta_1; \delta_k)$.

Montrons maintenant que δ_k c'est le seul test le PP**.** En effet, soit δ_k et δ deux tests les PP au niveau de signification α, on a donc $\pi(\theta_1; \delta_k) \geq \pi(\theta_1; \delta)$ et $\pi(\theta_1; \delta) \geq$

$\pi(\theta_1; \delta_k)$, c'est à dire $\pi(\theta_1; \delta) = \pi(\theta_1; \delta_k)$. Reprenons l'intégrale

$$\int (\delta - \delta_k)(L_1 - kL_0)\, dx_1 ... dx_n$$
$$= \mathbf{E}_1[\delta] - k\mathbf{E}_0[\delta] - \mathbf{E}_1[\delta_k] + k\mathbf{E}_0[\delta_k].$$

On a noté que cette intégrale est positive donc $\mathbf{E}_0[\delta] - k\mathbf{E}_0[\delta_k] \leq 0$. Or on a vu que $L_1 - kL_0$ est différent de 0, cela implique $\delta_k = \delta$. Donc les tests coïncident (à p près qui reste à déterminer).

Proposition 1. Le test du rapport de vraisemblance est sans biais.

Preuve. Montrons que $1 - \beta \geq \alpha$. Nous distinguons deux cas :

– Si $k \geq 1$:

$$1 - \beta = \mathbf{P}_1(W) = \int_W L_1 dx_1 ... dx_n \geq k \int_W L_0 dx_1 ... dx_n$$
$$\geq \int_W L_0 dx_1 ... dx_n = \mathbf{P}_0(W) = \alpha.$$

– Si $k < 1$:

$$\beta = \mathbf{P}_1\left(\overline{W}\right) = \int_{\overline{W}} L_1 dx_1 ... dx_n < k \int_{\overline{W}} L_0 dx_1 ... dx_n$$
$$< \int_W L_0 dx_1 ... dx_n = \mathbf{P}_0\left(\overline{W}\right) = 1 - \alpha.$$

Donc les deux cas on a $1 - \beta \geq \alpha$. Le test est donc sans biais.

Tests d'une hypothèse simple contre une hypothèse alternative composite

La région critique du test du rapport de vraisemblance de l'hypothèse $H_0 : \theta = \theta_0$ contre l'hypothèse alternative H_1 ne dépend pas de façon explicite de H_1. Donc cette région est la même pour n'importe quel $\theta > \theta_0$ où $\theta < \theta_0$.

Proposition 2. Le test du rapport de vraisemblance défini, au paragraphe précédente, par le Lemme de Neyman-Pearson est uniformément le plus puissant pour l'hypothèse $H_0 : \theta = \theta_0$ contre l'hypothèse alternative $H_1 : \theta > \theta_0$ et pour l'hypothèse $H_0 : \theta = \theta_0$ contre l'hypothèse alternative $H_1 : \theta < \theta_0$.

Remarque 2 :

1. Une hypothèse H est dite *bilatérale* si elle est de la forme $\theta \neq \theta_0$. Elle est dite *unilatérale à gauche si elle est de la forme* $\theta < \theta_0$. Elle est dite *unilatérale à droite si elle est de la forme* $\theta > \theta_0$.

2. Il n'existe pas en général de test upp pour $H_0 : \theta = \theta_0$ contre l'hypothèse alternative $H_1 : \theta \neq \theta_0$; car s'il en existait il devrait être upp pour sous-alternatives $H_1' : \theta > \theta_0$ et $H_1'' : \theta < \theta_0$. Or les tests de H_0 contre H_1' et H_1'' sont d'après, la proposition, upp mais ils sont (évidement) différents l'un de l'autre. Nous allons voir par la suite que les upp's peux êtres déterminer pour une certaine classe de distributions.

Chapitre 3

Tests entres deux hypothèses composites

a) **Alternative unilatérale** *:*

$$\begin{cases} H_0: & \theta \leq \theta_0 \\ H_1: & \theta > \theta_0 \end{cases} \quad \text{ou} \quad \begin{matrix} H_0: & \theta \geq \theta_0 \\ H_1: & \theta < \theta_0 \end{matrix}$$

Définition 2. On dit que la distribution d'une variable aléatoire X possède **un rapport de vraisemblance monotone** s'il existe une statistique $T = T(X_1, ..., X_n)$ telle que pour tout $\theta_1 > \theta_2$ le rapport de vraisemblance

$$\frac{L_1(x_1, ..., x_n)}{L_2(x_1, ..., x_n)}$$

est une fonction monotone en $t = T(x_1, ..., x_n)$.

Exemple 6 : Soit $X \rightsquigarrow$Bernoulli(p), $0 < p < 1$, définie par sa masse

$$\mathbf{P}(X = x) = p^x (1-p)^{1-x}, \; x = 0, 1.$$

Soit $0 < p_1, p_2 < 1$, telles que $p_1 > p_2$. Nous avons

$$\frac{L_1}{L_2} = \frac{\prod_{i=1}^{n} p_1^{x_i} (1-p_1)^{1-x_i}}{\prod_{i=1}^{n} p_2^{x_i} (1-p_2)^{1-x_i}} = \frac{p_1^t (1-p_1)^{n-t}}{p_2^t (1-p_2)^{n-t}},\ t = \sum_{i=1}^{n} x_i.$$

En d'autres termes

$$\frac{L_1}{L_2} = \left(\frac{1-p_1}{1-p_2}\right)^n \left(\frac{p_1(1-p_2)}{p_2(1-p_1)}\right)^t.$$

Il est clair que $a := \frac{1-p_1}{1-p_2} > 0$, $b := \frac{p_1(1-p_2)}{p_2(1-p_1)} > 1$, donc la fonction $t \to a^n \times b^t$ est une fonction croissante en t. Donc la distribution de X, possède un rapport de vraisemblance croissant.

Proposition 3. Si la distribution de X possède **un rapport de vraisemblance monotone** pour une statistique T, alors le test upp δ que correspond aux deux hypothèses $H_0 : \theta \leq \theta_0$ contre $H_1 : \theta > \theta_0$, au niveau de signification α, est définit comme suit :

– (a) Si L_1/L_2 est **croissant** par rapport à t :

$$\delta = \begin{cases} 1 & \text{si } t > c \\ p & \text{si } t = c \\ 0 & \text{si } t < c \end{cases},$$

où les constantes c et p sont telles que $\mathbf{P}_{\theta=\theta_0}(T > c) + p\mathbf{P}_{\theta=\theta_0}(T = c) = \alpha$.

– (b) Si L_1/L_2 est **décroissant** par rapport à t :

$$\delta = \begin{cases} 1 & \text{si } t < c \\ p & \text{si } t = c \\ 0 & \text{si } t > c \end{cases},$$

où les constantes c et p sont telles que $\mathbf{P}_{\theta=\theta_0}(T < c) + p\mathbf{P}_{\theta=\theta_0}(T = c) = \alpha$.

Proposition 3 (bis). Si la distribution de X possède **un rapport de vraisemblance monotone** pour une statistique T, alors le test upp δ que correspond aux deux hypothèses $H_0 : \theta \geq \theta_0$ contre $H_1 : \theta < \theta_0$, au niveau de signification α, est définit comme suit :

– (a) Si L_1/L_2 est **croissant** par rapport à t :

$$\delta = \begin{cases} 1 & \text{si } t < c \\ p & \text{si } t = c \\ 0 & \text{si } t > c \end{cases},$$

où les constantes c et p sont telles que $\mathbf{P}_{\theta=\theta_0}(T < c) + p\mathbf{P}_{\theta=\theta_0}(T = c) = \alpha$.

– (b) Si L_1/L_2 est **décroissant** par rapport à t :

$$\delta = \begin{cases} 1 & \text{si } t > c \\ p & \text{si } t = c \\ 0 & \text{si } t < c \end{cases},$$

où les constantes c et p sont telles que $\mathbf{P}_{\theta=\theta_0}(T > c) + p\mathbf{P}_{\theta=\theta_0}(T = c) = \alpha$.

Lorsque X est continue, on prendra $p = 1$.

Remarque. On note que si le rapport L_1/L_2 est croissant en t alors décroissant en $-t$. Donc l'utilisation de l'assertion (a) (pour t) ou l'assertion (b) (pour $-t$) mène au même résultat.

Exemple 7. Soit $X \rightsquigarrow \mathcal{N}(\mu, 1)$ et

$$\begin{cases} H_0 : & \mu \leq 0 \\ H_1 : & \mu > 0 \end{cases}$$

avec $\alpha = 0.1$, $n = 65$. Pour $\mu_1 > \mu_2$, on écrit

$$\begin{aligned}
\frac{L_1}{L_2} &= \exp\left\{n\left(\mu_1 - \mu_2\right)\left[\overline{x} - \frac{1}{2}\left(\mu_1 + \mu_2\right)\right]\right\} \\
&= \exp\left\{at + b\right\},
\end{aligned}$$

où $t = \overline{x}$, $a := n\left(\mu_1 - \mu_2\right) > 0$ et $b := a - \frac{1}{2}\left(\mu_1 + \mu_2\right)$. Il est clair que la fonction $t \rightarrow \exp\left\{at + b\right\}$ est une fonction croissante en t. Donc la distribution de X possède un rapport de vraisemblance croissant. Comme X est continue, alors d'après la Proposition 3, le test upp le plus puissant est défini par

$$\delta = \begin{cases} 1 & \text{si } \overline{x} \geq c \\ 0 & \text{si } \overline{x} < c \end{cases}$$

avec $\mathbf{P}_{\mu=0}\left(\overline{X} \geq c\right) = 0.1$, ce qui implique $c = 1.28/\sqrt{65} = 0.16$. On a donc

$$\delta = \begin{cases} 1 & \text{si } \overline{x} \geq 0.16 \\ 0 & \text{si } \overline{x} < 0.16 \end{cases}$$

La fonction puissance du test δ est définie par

$$\pi\left(\mu\right) = \begin{cases} \alpha\left(\mu\right) & \text{si } \mu \leq 0 \\ 1 - \beta\left(u\right) & \text{si } \mu > 0 \end{cases}.$$

En d'autres termes

$$\begin{aligned}
\pi(\mu) &= \mathbf{P}\left(\overline{X} \geq 0.16 \mid \mu \in \mathbb{R}\right) \\
&= 1 - \mathbf{P}\left(\sqrt{65}\left(\overline{X} - \mu\right) < 1.28 - \sqrt{65}\mu \mid \mu \in \mathbb{R}\right) \\
&= 1 - \mathbf{P}\left(Z < 1.28 - \sqrt{65}\mu \mid \mu \in \mathbb{R}\right) \\
&= 1 - \Phi\left(1.28 - \sqrt{65}\mu\right), \ \mu \in \mathbb{R}. \\
&= \Phi\left(\sqrt{65}\mu - 1.28\right)
\end{aligned}$$

La fonction puissance π étant croissante sur $\mathbb{R}$, on trouve bien le fait que

$$\begin{aligned}
\alpha &= \sup_{\mu \leq 0} \pi(\mu) = \pi(0) \\
&= \Phi(-1.28) = 1 - \Phi(1.28) \\
&= 1 - 0.899 = 0.1,
\end{aligned}$$

qui correspond en effet au seuil (niveau) de signification du test. Le graphe de la fonction puissance $\mu \to \pi(\mu) = \Phi\left(\sqrt{65}\mu - 1.28\right)$, $\mu \in \mathbb{R}$ est donné par la figure Fig.2

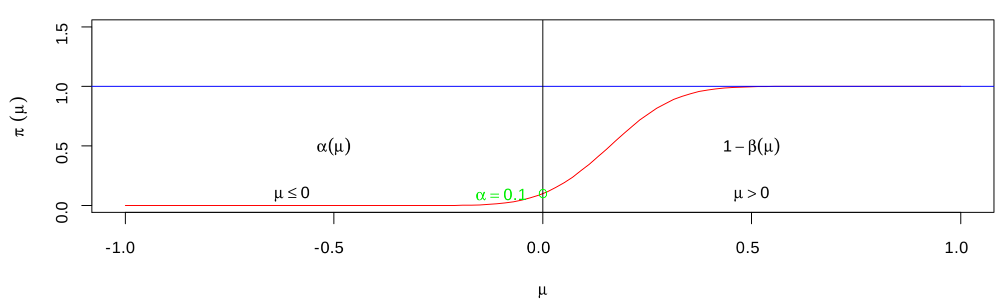

Fig.3

Voici les codes R de la figure Fig.3 :

```
f<-function(x){pnorm(sqrt(64)*x-1.28)}

x<-seq(-1,1,length=100)

plot(x,f(x),type="l",col="red",ylim=c(0,1.5),

xlab=expression(mu),ylab=expression(pi~(mu)))

abline(h=1,col="blue")

abline(v=0)

points(0,0.1,col="green")

text(-0.1,0.1,expression(alpha==0.1),col="green2")

text(-0.5,0.5,expression(alpha(mu)))

text(0.5,0.5,expression(1-beta(mu)))

text(-0.6,0.1,expression(mu<=0))

text(0.5,0.1,expression(mu>0))
```

Exemple 7 (bis). Si nous considérons le test suivant :

$$\begin{cases} H_0: & \mu \geq 0 \\ H_1: & \mu < 0 \end{cases}$$

alors d'après la Proposition 3 (bis), le test upp est

$$\delta = \begin{cases} 1 & \text{si } \overline{x} \leq c \\ 0 & \text{si } \overline{x} > c \end{cases}$$

avec $\mathbf{P}_0\left(\overline{X} \leq c\right) = 0.1$, ce qui implique $\sqrt{65}c = -1.28$ donc $c = -0.16$. On a donc

$$\delta = \begin{cases} 1 & \text{si } \overline{x} \leq -0.16 \\ \\ 0 & \text{si } \overline{x} > -0.16 \end{cases}$$

La fonction puissance du test δ est définie par

$$\pi\left(\mu\right) = \begin{cases} \alpha\left(\mu\right) & \text{pour } \mu \geq 0 \\ \\ 1 - \beta\left(u\right) & \text{pour } \mu < 0 \end{cases}.$$

En d'autre termes

$$\begin{aligned} \pi\left(\mu\right) &= \mathbf{P}\left(\overline{X} \leq -0.16 \mid \mu \in \mathbb{R}\right) \\ &= \mathbf{P}\left(\sqrt{65}\left(\overline{X} - \mu\right) < -1.28 - \sqrt{65}\mu \mid \mu \in \mathbb{R}\right) \\ &= \mathbf{P}\left(Z < -\sqrt{65}\mu - 1.28 \mid \mu \in \mathbb{R}\right) \\ &= \Phi\left(-\sqrt{65}\mu - 1.28\right),\ \ \mu \in \mathbb{R}. \end{aligned}$$

La fonction puissance π étant décroissante sur $\mathbb{R}$, on trouve bien le fait que

$$\begin{aligned} \alpha &= \sup_{\mu \geq 0} \pi\left(\mu\right) = \pi\left(0\right) \\ &= \Phi\left(-1.28\right) = 1 - \Phi\left(1.28\right) \\ &= 1 - 0.899 = 0.1, \end{aligned}$$

qui correspond en effet au seuil (niveau) de signification du test. Le graphe de la fonction puissance $\mu \to \pi\left(\mu\right) = \Phi\left(-1.28 - \sqrt{65}\mu\right)$, $\mu \in \mathbb{R}$ est donné par la figure

Fig.4

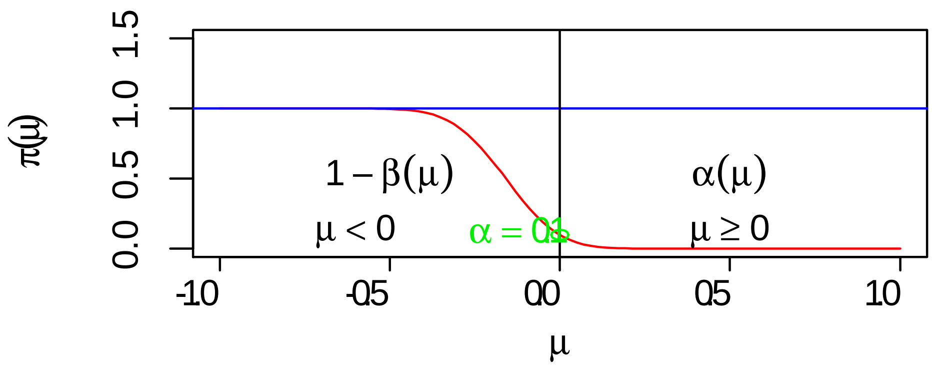

Fig.4

Voici les codes R de la figure Fig.4 :

```
f<-function(x){pnorm(-sqrt(64)*x-1.28)}

x<-seq(-1,1,length=100)

plot(x,f(x),type="l",col="red",ylim=c(0,1.5),

xlab=expression(mu),ylab=expression(pi~(mu)))

abline(h=1,col="blue")

abline(v=0)

points(0,0.1,col="green")

text(-0.1,0.1,expression(alpha==0.1),col="green2")

text(-0.5,0.5,expression(1-beta(mu)))

text(0.5,0.5,expression(alpha(mu)))

text(-0.6,0.1,expression(mu<0))

text(0.5,0.1,expression(mu>=0))
```

Preuve de la proposition 3.

Supposons que X est continue. D'après le paragraphe 3, le test qui rejette $H_0 : \theta = \theta_0$ au profit de $H_1 : \theta = \theta_1$ quand $L_1/L_0 \geq k$ est upp au niveau de signification α. Or L_1/L_0 est croissant en T d'où $L_1/L_0 \geq k \iff T \geq c$, avec c est k sont convenablement liés et

$$\alpha = \mathbf{P}_0 \left(L_1/L_0 \geq k \right) = \mathbf{P}_0 \left(T \geq c \right).$$

Donc le test qui rejette $H_0 : \theta = \theta_0$ au profit de $H_1 : \theta > \theta_1$ quand $T \geq k$ est upp au niveau de signification α. En d'autres termes dans ce cas l'hypothèse H_1 ($\theta = \theta_1$ où $\theta > \theta_0$) n'a pas d'influence sur la décision. On désigne maintenant par $\mathcal{C} := \{\delta \mid \pi(\theta_0, \delta) \leq \alpha\}$ la classe des tests ayant un risque de deuxième espèce inferieur ou égale à α. Il est clair que le test de Neyman-Pearson δ_k est tel que $\pi(\theta_0, \delta_k) = \alpha$ donc $\delta_k \in \mathcal{C}$. D'après qui précède, on peut dire que δ_k esr upp dans la classe $\mathcal{C}$. Soit $\mathcal{C}' := \{\delta \mid \pi(\theta, \delta) \leq \alpha, \text{ pour } \theta \leq \theta_0\}$, alors $\mathcal{C}' \subset \mathcal{C}$. En effet, si $\delta \in \mathcal{C}' \iff \pi(\theta, \delta) \leq \alpha$, pour tout $\theta \leq \theta_0 \Longrightarrow \pi(\theta_0, \delta) \leq \alpha \iff \delta \in \mathcal{C}$. D'après la première remarque ci-dessous, la fonction puissance du test δ_k est croissante (par rapport à θ), d'où

$$\theta \leq \theta_0 \Longrightarrow \pi(\theta, \delta_k) \leq \pi(\theta_0, \delta_k) = \alpha.$$

Donc $\delta_k \in \mathcal{C}'$ et par conséquent δ_k est upp dans la classe $\mathcal{C}'$.

Remarque 4 :

1) La fonction puissance du test δ_k $[\pi(\theta, \delta_k, \ \theta \in \Xi)]$ est une fonction croissante par rapport à θ dans Ξ, c'est à dire, $\theta' \leq \theta'' \Longrightarrow \pi(\theta', \delta_k) \leq \pi(\theta'', \delta_k)$

2) Pour $\theta < \theta_0$, le test δ_k minimise la fonction $\pi(\theta)$; dans le sens que si δ est un autre test de niveau de signification α alors pour tout $\theta < \theta_0$, $\pi(\theta, \delta_k) \leq \pi(\theta, \delta)$.

Preuve de la remarque 4.

1) On montre que si $\theta' \leq \theta'' \Longrightarrow \pi(\theta', \delta_k) \leq \pi(\theta'', \delta_k)$, pout tout $\theta', \theta'' \in \Xi$. On pose $\gamma := \pi(\theta', \delta_k)$ et $\mathcal{C} := \{\delta \mid \pi(\theta', \delta) = \gamma\}$, donc $\delta_k \in \mathcal{C}$. D'après la première partie de la démonstration ci-dessus, on peut dire que δ_k maximise $\pi(\theta'', \delta)$ pour $\delta \in \mathcal{C}$ (δ_k est upp dans $\mathcal{C}$). Soit $\widetilde{\delta}$ un test tel que $\pi\left(\theta, \widetilde{\delta}\right) = \gamma$, pour tout $\theta \in \Theta$, alors $\widetilde{\delta} \in \mathcal{C}$. Donc

$$\pi(\theta'', \delta_k) \geq \pi\left(\theta'', \widetilde{\delta}\right) = \pi\left(\theta', \widetilde{\delta}\right) = \gamma = \pi(\theta', \delta_k).$$

2) Pour tester $H_0' : \theta \geq \theta_0$ contre $H_1' : \theta < \theta_0$, le test upp est $\delta_k^* := 1- \delta_k$ au niveau de signification $1 - \alpha$. La puissance δ_k^* est maximale pour $\theta < \theta_0$. La puissance du test δ_k^* est :

$$\pi(\theta, \delta_k^*) = \mathbf{E}_\theta[\delta_k^*] = 1 - \mathbf{E}_\theta[\delta_k] = 1 - \pi(\theta, \delta_k), \text{ pour } \theta < \theta_{0.}$$

Par conséquent, δ_k minimise $\pi(\theta)$ pour $\theta < \theta_0$.

b) **Hypothèse nulle bilatérale** *:*

$$\begin{cases} H_0 : & \theta \leq \theta_1 \text{ ou } \theta \geq \theta_2 \\ H_1 : & \theta_1 < \theta < \theta_2 \end{cases}$$

Proposition 4.

Soit X une population de densité :

$$f_X(x; \theta) = \exp\{a(x)u(\theta) + b(x) + \vartheta(\theta)\}, \tag{3.1}$$

telle que la fonction $u(\theta)$ est monotone. Alors le test

$$\delta = \begin{cases} 1 & \text{si } c_1 < t < c_2 \\ p_i & \text{si } t = c_i;\ i = 1, 2 \\ 0 & \text{si } t < c_1 \text{ ou } t > c_2 \end{cases}$$

est le upp pour les hypothèses ci-dessus au niveau de signification α, où $t = \sum_{i=1}^{n} a(x_i)$ et c_i et p_i vérifient les conditions

$$\mathbf{E}_{\theta_1}[\delta] = \mathbf{E}_{\theta_2}[\delta] = \alpha,$$

en d'autres termes

$$\pi(\theta_1, \delta) = \pi(\theta_2, \delta) = \alpha.$$

Remarque 5 :

1) Lorsque X est continue on prend $p_1 = p_2 = 1$.

2) δ maximise la fonction puissance $\pi(\theta)$ à l'intérieur de l'intervalle $]\theta_1, \theta_2[$ et la minimise à l'extérieur.

3) $\pi(\theta, \delta)$ a un maximum en un point $\theta_0 \in]\theta_1, \theta_2[$ et décroit strictement à gauche et à droite de θ_0.

4) La proposition si dessus est valable pour le test de :

$$\begin{cases} H_0 : & \theta < \theta_1 \text{ ou } \theta > \theta_2 \\ H_1 : & \theta_1 \leq \theta \leq \theta_2 \end{cases}$$

5) Pour la preuve de la proposition voir le livre de Borovkov, page 320.

Exemple 8 :

Soit $X \rightsquigarrow \mathcal{N}(\mu, 1)$. On peut vérifier que la densité appartient à la famille (3.1), avec $a(x) = x$ et $u(\mu) = \mu$ (croissante sur $\mathbb{R}$). Nous avons affaire au test suivant :

$$\begin{cases} H_0 : & \mu \leq 1 \text{ ou } \mu \geq 2 \\ H_1 : & 1 < \mu < 2 \end{cases}$$

avec $n = 10$ et $\alpha = 0.1$. Suivant la proposition ci-dessus la statistique du test est $T = \sum_{i=1}^{10} X_i$ et le test upp est

$$\delta = \begin{cases} 1 & \text{si } c_1 \leq \sum_{i=1}^{10} x_i \leq c_2 \\ 0 & \text{si } \sum_{i=1}^{10} x_i < c_1 \text{ ou } \sum_{i=1}^{10} x_i > c_2 \end{cases}$$

où c_1 et c_2 sont solutions du systèmes :

$$\pi_{\mu=1}(\delta) = 0.1 \text{ et } \pi_{\mu=2}(\delta) = 0.1.$$

En d'autres termes

$$\begin{cases} \mathbf{P}_{\mu=1}\left(c_1 \leq \sum_{i=1}^{10} X_i \leq c_2\right) = 0.1 \\ \mathbf{P}_{\mu=2}\left(c_1 \leq \sum_{i=1}^{10} X_i \leq c_2\right) = 0.1 \end{cases}$$

En standardisant l'échantillon, on obtient

$$\begin{cases} \mathbf{P}_{\mu=1}\left(\frac{c_1-10}{\sqrt{10}} \leq \frac{\sum_{i=1}^{10} X_i-10}{\sqrt{10}} \leq \frac{c_2-10}{\sqrt{10}}\right) = 0.1 \\ \mathbf{P}_{\mu=2}\left(\frac{c_1-20}{\sqrt{10}} \leq \frac{\sum_{i=1}^{10} X_i-20}{\sqrt{10}} \leq \frac{c_2-20}{\sqrt{10}}\right) = 0.1 \end{cases}$$

Autrement dit

$$\begin{cases} \mathbf{P}\left(\frac{c_1-10}{\sqrt{10}} \leq Z \leq \frac{c_2-10}{\sqrt{10}}\right) = 0.1 \\ \mathbf{P}\left(\frac{c_1-20}{\sqrt{10}} \leq Z \leq \frac{c_2-20}{\sqrt{10}}\right) = 0.1 \end{cases}$$

où $Z \rightsquigarrow \mathcal{N}(0,1)$. Autrement dit

$$\begin{cases} \Phi\left(\frac{c_2-10}{\sqrt{10}}\right) - \Phi\left(\frac{c_1-10}{\sqrt{10}}\right) = 0.1 \\ \Phi\left(\frac{c_2-20}{\sqrt{10}}\right) - \Phi\left(\frac{c_1-20}{\sqrt{10}}\right) = 0.1 \end{cases}$$

La solution (numérique) de ce système nous donne :

$$c_1 = 13.67,\ c_2 = 16.32.$$

Voici le code R avec la package "nleqslv" qui résout ce système :

```
require(nleqslv)

h <-function(x){

y<-numeric(2)

y[1]<-pnorm((x[2]-10)/sqrt(10))-pnorm((x[1]-10)/sqrt(10))-0.1

y[2]<-pnorm((x[2]-20)/sqrt(10))-pnorm((x[1]-20)/sqrt(10))-0.1

y}

xstart<-c(1,2)

nleqslv(xstart,h,control=list(btol=.01))
```

Ainsi le test upp est :

$$\delta = \begin{cases} 1 & \text{si } 13.67 \leq \sum_{i=1}^{10} x_i \leq 16.32 \\ 0 & \text{si } \sum_{i=1}^{10} x_i < 13.67 \text{ ou } \sum_{i=1}^{10} x_i > 16.32 \end{cases}$$

La fonction puissance du test est

$$\begin{aligned}
\pi(\mu) &= \mathbf{P}_\mu\left(13.67 \leq \sum_{i=1}^{10} X_i \leq 16.32\right),\ \mu \in \mathbb{R} \\
&= \mathbf{P}_\mu\left(\frac{13.67 - 10\mu}{\sqrt{10}} \leq \frac{\sum_{i=1}^{10} X_i - 10\mu}{\sqrt{10}} \leq \frac{16.32 - 10\mu}{\sqrt{10}}\right) \\
&= \mathbf{P}\left(\frac{13.67 - 10\mu}{\sqrt{10}} \leq Z \leq \frac{16.32 - 10\mu}{\sqrt{10}}\right) \\
&= \Phi\left(\frac{16.32 - 10\mu}{\sqrt{10}}\right) - \Phi\left(\frac{13.67 - 10\mu}{\sqrt{10}}\right),\ \mu \in \mathbb{R}.
\end{aligned}$$

b) **Alternative Bilatérale** *:*

$$\begin{cases} H_0: & \theta_1 \leq \theta \leq \theta_2 \\ H_1: & \theta < \theta_1 \text{ ou } \theta > \theta_2 \end{cases},\ \text{avec } \theta_1 < \theta_2$$

Proposition 5.

Soit X une population de densité :

$$f_X(x;\theta) = \exp\{a(x)u(\theta) + b(x) + \vartheta(\theta)\}, \tag{3.2}$$

telle que la fonction $u(\theta)$ est monotone. Alors le test

$$\delta = \begin{cases} 1 & \text{si } t < c_1 \text{ ou } t > c_2 \\ p_i & \text{si } t = c_i;\ i = 1,2 \\ 0 & \text{si } c_1 < t < c_2 \end{cases}$$

est le upp pour les hypothèses ci-dessus au niveau de signification α, où $t = \sum_{i=1}^{n} a(x_i)$ et c_i et p_i vérifient les conditions

$$\mathbf{E}_{\theta_1}[\delta] = \mathbf{E}_{\theta_2}[\delta] = \alpha,$$

en d'autres termes

$$\pi(\theta_1, \delta) = \pi(\theta_2, \delta) = \alpha.$$

Autrement dit

$$\mathbf{P}_{\theta_i}(T < c_1) + p_1\mathbf{P}_{\theta_i}(T = c_1) + p_2\mathbf{P}_{\theta_i}(T = c_2) + \mathbf{P}_{\theta_i}(T > c_2) = \alpha, \ i = 1, 2.$$

Cas particulier : $\theta_1 = \theta_2 = \theta_0$. En d'autres termes nous avons affaire au test suivant :

$$\begin{cases} H_0 : & \theta = \theta_0 \\ H_1 : & \theta \neq \theta_0 \end{cases}$$

Dans ce cas, la fonction puissance $\pi(\theta)$ est minimale en θ_0. Les constantes c_i et p_i vérifient donc $\pi'(\theta_0) = 0$ en plus de la condition $\mathbf{E}_{\theta_0}[\delta] = \alpha$. Par conséquent, les constantes c_i et p_i se déterminent à partir de :

$$\begin{cases} \mathbf{P}_{\theta_0}(T < c_1) + p_1\mathbf{P}_{\theta_0}(T = c_1) + p_2\mathbf{P}_{\theta_0}(T = c_2) + \mathbf{P}_{\theta_0}(T > c_2) = \alpha \\ \pi'(\theta_0) = \left.\dfrac{d}{d\theta}\pi(\theta)\right|_{\theta=\theta_0} = 0 \end{cases}$$

Exemple 9

Soit $X \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$, oú $\sigma^2 < \infty$ est connue. Nous avons affaire au test suivant :

$$\begin{cases} H_0 : & \mu = \mu_0 \\ H_1 : & \mu \neq \mu_0 \end{cases}$$

et ce en se basant sur un échantillon de taille n et un seuil α. Le test upp sera de la forme

$$\delta = \delta(x_1, ..., x_n) = \begin{cases} 1 & \text{si } \sum_{i=1}^n x_i \leq c_1 \text{ ou} \sum_{i=1}^n x_i \geq c_2 \\ 0 & \text{si } c_1 < \sum_{i=1}^n x_i < c_2, \end{cases}$$

où $c_1 \neq c_2$ sont deux constantes à déteminer. En d'autres termes

$$\delta = \begin{cases} 1 & \text{si } \overline{x} \leq k_1 \text{ ou } \overline{x} \geq k_2 \\ 0 & \quad \text{si } k_1 < \overline{x} < k_2 \end{cases}$$

où $k_i := c_i/n$, $i = 1, 2$. La fonction puissance de ce test est

$$\begin{aligned} \pi(\mu; \delta) &= \mathbf{E}_\mu[\delta] = \mathbf{P}_\mu\left(\overline{X} \leq k_1 \text{ ou } \overline{X} \geq k_2\right), \text{ pour } \mu \in \mathbb{R} \\ &= \mathbf{P}_\mu\left(\overline{X} \leq k_1\right) + \mathbf{P}_\mu\left(\overline{X} \geq k_2\right) \\ &= \mathbf{P}_\mu\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{k_1 - \mu}{\sigma/\sqrt{n}}\right) + \mathbf{P}_\mu\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \geq \frac{k_2 - \mu}{\sigma/\sqrt{n}}\right) \\ &= \Phi\left(\frac{k_1 - \mu}{\sigma/\sqrt{n}}\right) + 1 - \Phi\left(\frac{k_2 - \mu}{\sigma/\sqrt{n}}\right). \end{aligned}$$

En passant à la dérivée par rapport à μ, on trouve

$$\pi'(\mu; \delta) = -\frac{1}{\sigma/\sqrt{n}}\varphi\left(\frac{k_1 - \mu}{\sigma/\sqrt{n}}\right) + \frac{1}{\sigma/\sqrt{n}}\varphi\left(\frac{k_2 - \mu}{\sigma/\sqrt{n}}\right), \text{ avec } \varphi = \Phi',$$

d'où

$$\pi'(\mu_0; \delta) = \frac{1}{\sigma/\sqrt{n}}\frac{1}{\sqrt{2\pi}}\left[\exp\left\{-\frac{(k_2 - \mu_0)^2}{2\sigma^2/n}\right\} - \exp\left\{-\frac{(k_1 - \mu_0)^2}{2\sigma^2/n}\right\}\right].$$

Nous avons $\pi'(\mu_0; \delta) = 0$, ce qui est équivalent à $(k_1 - \mu_0)^2 = (k_2 - \mu_0)^2$, ainsi

$$k_1 - \mu_0 = -(k_2 - \mu_0), \tag{3.3}$$

car $k_1 \neq k_2$. D'autres part $\mathbf{E}_{\mu=\mu_0}[\delta] = \alpha$, donc

$$\Phi\left(\frac{-(k_2 - \mu)}{\sigma/\sqrt{n}}\right) + 1 - \Phi\left(\frac{k_2 - \mu_0}{\sigma/\sqrt{n}}\right) = \alpha.$$

On sait que $\Phi(-x) = 1 - \Phi(x)$, alors

$$1 - \Phi\left(\frac{k_2 - \mu_0}{\sigma/\sqrt{n}}\right) + 1 - \Phi\left(\frac{k_2 - \mu_0}{\sigma/\sqrt{n}}\right) = \alpha.$$

ce qui implique que

$$\Phi\left(\frac{k_2 - \mu_0}{\sigma/\sqrt{n}}\right) = 1 - \alpha/2.$$

En d'autres termes $\frac{k_2-\mu_0}{\sigma/\sqrt{n}} = \xi_{(1-\alpha/2)}$, où $\xi_{(1-\alpha/2)} := \Phi^{-1}(1-\alpha/2)$ (le quantile d'ordre $1-\alpha/2$ de la loi normale centrée-réduite), ce qui donne $k_2 = \mu_0 + \sigma\xi_{(1-\alpha/2)}/\sqrt{n}$, Grace à l'égalité (3.3), on obient aussi $k_1 = \mu_0 - \sigma\xi_{(1-\alpha/2)}/\sqrt{n}$. Ainsi, le test upp δ est donc

$$\delta = \begin{cases} 1 & \text{si } \overline{x} \leq \mu_0 - \sigma\xi_{(1-\alpha/2)}/\sqrt{n} \text{ ou } \overline{x} \geq \mu_0 + \sigma\xi_{(1-\alpha/2)}/\sqrt{n} \\ 0 & \text{si } \mu_0 - \sigma\xi_{(1-\alpha/2)}/\sqrt{n} < \overline{x} < \mu_0 - \sigma\xi_{(1-\alpha/2)}/\sqrt{n}, \end{cases}$$

autrement dit

$$\delta = \begin{cases} 1 & \text{si } \frac{|\overline{x}-\mu_0|}{\sigma/\sqrt{k}} \geq \xi_{(1-\alpha/2)} \\ 0 & \text{si } \frac{|\overline{x}-\mu_0|}{\sigma/\sqrt{k}} < \xi_{(1-\alpha/2)}. \end{cases}$$

Application : Soit $X \rightsquigarrow \mathcal{N}(\mu, 1)$. Supposons que nous avons affaire au test suivant :

$$\begin{cases} H_0 : & \mu = 0 \\ H_1 : & \mu \neq 0 \end{cases}$$

avec $n = 9$ et $\alpha = 0.05$. Nous avons

$$\xi_{1-\alpha/2} = \Phi^{-1}\left(1 - \frac{0.05}{2}\right) = \Phi^{-1}(0.975) = 1.96.$$

Donc le test δ upp est de la forme

$$\delta = \begin{cases} 1 & \text{si } 3\left|\overline{x}\right| \geq 1.96 \\ 0 & \text{si } 3\left|\overline{x}\right| < 1.96. \end{cases}$$

ou simplement

$$\delta = \begin{cases} 1 & \text{si } \left|\overline{x}\right| \geq 0.65 \\ 0 & \text{si } \left|\overline{x}\right| < 0.65. \end{cases}$$

Exemple 10

Soit $X \rightsquigarrow \mathcal{N}(\mu, 1)$. Nous avons affaire au test suivant :

$$\begin{cases} H_0 : & 9 \leq \mu \leq 10 \\ H_1 : & \mu < 9 \text{ ou } \mu > 10 \end{cases}$$

avec $n = 16$ et $\alpha = 0.05$. Montrer que

$$\delta = \begin{cases} 1 & \text{si } \overline{x} \leq 8.59 \text{ ou } \overline{x} \geq 10.41 \\ 0 & \text{si } 8.59 < \overline{x} < 10.41. \end{cases}$$

En effet, de l'exemple précédent on conclut que le test upp est

$$\delta = \delta(x_1, ..., x_n) = \begin{cases} 1 & \text{si } \overline{x} \leq k_1 \text{ ou } \overline{x} \geq k_2 \\ 0 & \text{si } k_1 < \overline{x} < k_2 \end{cases}$$

Les constantes k_1 et k_2 solution du système

$$\begin{cases} \mathbf{P}_9\left(\overline{X} \leq k_1\right) + \mathbf{P}_9\left(\overline{X} > k_2\right) = 0.05 \\ \mathbf{P}_{10}\left(\overline{X} \leq c_1\right) + \mathbf{P}_{10}\left(\overline{X} > c_2\right) = 0.05 \end{cases}$$

Ceci peut être réécrit comme suit

$$\begin{cases} \mathbf{P}_9\left(4\left(\overline{X}-9\right)<4\left(k_1-9\right)\right)+\mathbf{P}_9\left(4\left(\overline{X}-9\right)>4\left(k_2-9\right)\right)=0.05 \\ \mathbf{P}_{10}\left(4\left(\overline{X}-10\right)<4\left(k_1-10\right)\right)+\mathbf{P}_{10}\left(4\left(\overline{X}-10\right)>4\left(k_2-10\right)\right)=0.05 \end{cases}$$

En d'autres termes

$$\begin{cases} \mathbf{P}_9\left(4\left(\overline{X}-9\right)\leq 4\left(k_2-9\right)\right)-\mathbf{P}_9\left(4\left(\overline{X}-9\right)<4\left(k_1-9\right)\right)=1-0.05=0.95 \\ \mathbf{P}_{10}\left(4\left(\overline{X}-10\right)\leq 4\left(k_1-10\right)\right)-\mathbf{P}_{10}\left(4\left(\overline{X}-10\right)<4\left(k_2-10\right)\right)=1-0.05=0.95 \end{cases}$$

Ainsi nous obtenons

$$\begin{cases} \Phi\left(4\left(k_2-9\right)\right)-\Phi\left(4\left(k_1-9\right)\right)=0.95 \\ \Phi\left(4\left(k_2-10\right)\right)-\Phi\left(4\left(k_1-10\right)\right)=0.95 \end{cases}$$

où Φ désigne la fonction de répartition de la loi normale centrée-réduite. La résolution de ce système donne $k_1 = 8.59$ et $k_2 = 10.41$. Voici un programme en langage R pour résoudre ce système :

```
-----------------------------------------------------------------------
require(nleqslv)
h<-function(x){
y<-numeric(2)
y[1]<-pnorm(4*(x[2]-9))-pnorm(4*(x[1]-9))-0.95
y[2]<-pnorm(4*(x[2]-10))-pnorm(4*(x[1]-10))-0.95
y}
xstart<-c(9,10)
nleqslv(xstart,h,control=list(btol=.01))
-----------------------------------------------------------------------
```

La fonction puissance est

$$\begin{aligned}\pi(\mu) &= \mathbf{E}_\mu\left[\delta(X_1,...,X_{16})\right]\\ &= \mathbf{P}_\mu\left(\overline{X}\leq 8.59 \text{ ou } \overline{X}\geq 10.41\right)\\ &= \mathbf{P}_\mu\left(\overline{X}\leq 8.59\right)+\mathbf{P}_\mu\left(\ \overline{X}\geq 10.41\right),\ \mu\in\mathbb{R}.\end{aligned}$$

En d'autre termes

$$\pi(\mu)=\Phi(4(8.59-\mu))+1-\Phi(4(10.41-\mu)),\ \mu\in\mathbb{R}.$$

Le graphe de la fonction puissance est donné par la figure Fig.5.

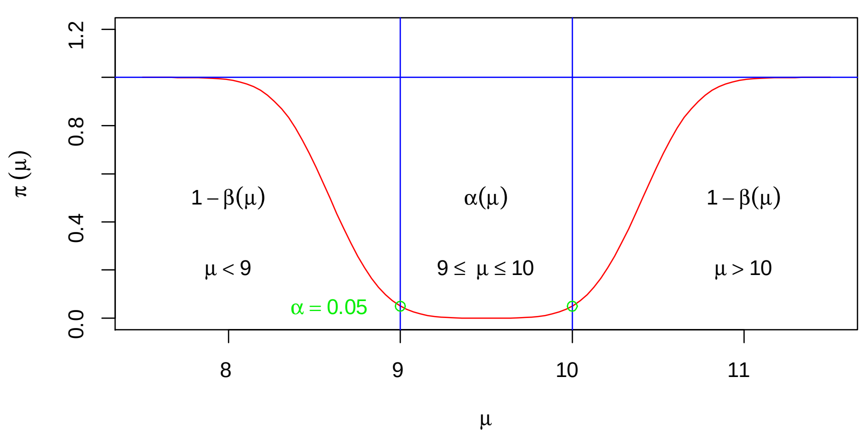

Fig.5

Voici le code R de cette figure :

```
f<-function(x){pnorm(4*(8.59-x))+1-pnorm(4*(10.41-x))}

x<-seq(7.5,11.5,length=100)
```

```
plot(x,f(x),type="l",col="red",ylim=c(0,1.2),
 xlab=expression(mu),ylab=expression(pi~(mu)))
abline(h=1,col="blue")
abline(v=9,col="blue")
abline(v=10,col="blue")
points(9,0.05,col="green2",)
points(10,0.05,col="green2")
text(8.59,0.05,expression(alpha==0.05),col="green2")
text(11,0.5,expression(1-beta(mu)))
text(8,0.5,expression(1-beta(mu)))
text(9.5,0.5,expression(alpha(mu)))
text(8,0.2,expression(mu<9))
text(11,0.2,expression(mu>10))
text(9.5,0.2,expression(paste(9<=~mu~<=10)))
```

Chapitre 4

Rapport de vraisemblance généralisé

Ce test est utile surtout là où les méthodes précédentes sont échoué. Il s'applique aussi le cas où le paramètre θ est vectoriel : $\theta = (\theta_1, ..., \theta_p) \in \Theta \subset \mathbb{R}^p$. Les resutlats exposés ci-dessous s'appliquent aussi bien aux variables aléatoires continues que disctrètes.

Rapport de vraisemblance généralisé

On appelle rapport de vraisemblance généralisé la quantité :

$$R = R\left(x_1, ..., x_n\right) = \frac{\sup_{\theta \in \Theta_1} L\left(x_1, ..., x_n; \theta\right)}{\sup_{\theta \in \Theta_0} L\left(x_1, ..., x_n; \theta\right)}.$$

Définition. On appelle test du rapport de vraisemblance généralisé de $H_0 : \theta \in \Theta_0$ contre $H_1 : \theta \in \Theta_1$, au niveau de signification α, le test δ dont la région critique est :

$$W := \left\{ \left(x_1, ..., x_n\right) \in \mathbb{R}^n : \ R > c \right\}.$$

La constante c se détermine par la relation :

$$\sup_{\theta\in\Theta_0} \mathbf{P}_\theta \left(R\left(X_1, ..., X_n \right) > c \right) = \alpha.$$

Remarque.

1) Un test équivalent, de même appellation, est souvent envisagé. Il est basé sur le rapport

$$R_1 := \frac{\sup_{\theta\in\Theta} L\left(x_1, ..., x_n; \theta\right)}{\sup_{\theta\in\Theta_0} L\left(x_1, ..., x_n; \theta\right)} = \frac{L\left(x_1, ..., x_n; \widehat{\theta}_{emv}\right)}{\sup_{\theta\in\Theta_0} L\left(x_1, ..., x_n; \theta\right)},$$

où $\widehat{\theta}_{emv}$ désigne l'estimateur de maximum de vraisemblance (EMV) de θ dans Θ. Il est bien clair que

$$L\left(x_1, ..., x_n; \widehat{\theta}\right) = \max\left\{ \sup_{\theta\in\Theta_0} L\left(x_1, ..., x_n; \theta\right), \sup_{\theta\in\Theta_1} L\left(x_1, ..., x_n; \theta\right) \right\},$$

ainsi $R_1 = \max\{1, R\}$.

2) Si l'hypothèse H_0 est simple $\Theta_0 = \{\theta_0\}$ et $\Theta_1 = \Theta - \{\theta_0\}$, alors pour les $f(x;\theta)$ continues par rapport à θ on a $R_1 = R$.

Exemple 11 : Test de la moyenne d'une population normale de variance inconnue

Soit $X \rightsquigarrow \mathcal{N}\left(\mu, \sigma^2\right)$, ici $\theta = \left(\mu, \sigma^2\right) \in \mathbb{R} \times \mathbb{R}_+^* =: \Theta$. La fonction de vraisemblance est

$$L(\theta) := L\left(x_1, ..., x_n, \theta\right) = \left(2\pi\sigma^2\right)^{-n/2} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} \left(x_i - \mu\right)^2 \right\}.$$

1) Test bilatéral :

$$\begin{cases} H_0 : & \mu = \mu_0 \\ H_1 : & \mu \neq \mu_0 \end{cases}$$

Dans ce cas $\Theta_0 = \{(\mu, \sigma^2) : \ \mu = \mu_0\}$ et $\Theta_1 = \{(\mu, \sigma^2) : \ \mu \neq \mu_0\}$. Nous avons

$$\sup_{\theta \in \Theta} L(\theta) = L\left(\widehat{\theta}\right),$$

où $\widehat{\theta} = (\widehat{\mu}, \widehat{\sigma}^2) = \left(\overline{X}, \widetilde{S}^2\right)$ est le EMV de θ dans Θ, où

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

On peut vérifier facilement que

$$\sup_{\theta \in \Theta_0} L(\theta) = L\left(\widehat{\theta}_0\right),$$

où $\widehat{\theta}_0 = (\mu_0, \widehat{\sigma}_0^2)$ est l' EMV dans Θ_0, avec $\widehat{\sigma}_0^2 = n^{-1}\sum_{i=1}^{n}(X_i - \mu_0)^2$. Ainsi le rapport de vraisemblance est

$$R_1(X_1, ..., X_n) = \left(\frac{\widehat{\sigma}_0^2}{\widehat{\sigma}^2}\right)^{n/2}.$$

Le test rejette H_0 pour les plus grandes valeurs de R_1, c'est à dire

$$W = \{(x_1, ..., x_n) \in \mathbb{R}^n : R_1(x_1, ..., x_n) > k\},$$

où k est telle que

$$\sup_{(\mu, \sigma^2) \in \Theta_0} \mathbf{P}(R_1(X_1, ..., X_n) > k) = \alpha.$$

Dans la suite nous allons déterminer la loi de la statistique du test $R_1 = (\widehat{\sigma}_0^2 / \widehat{\sigma}^2)^{n/2}$. On peut vérifier facilement que

$$\widehat{\sigma}_0^2 = \widehat{\sigma}^2 + \left(\overline{X} - \mu_0\right)^2. \tag{4.1}$$

En effet

$$\begin{aligned}
\widehat{\sigma}_0^2-\widehat{\sigma}^2 &= \frac{1}{n}\sum_{i=1}^{n}\left[(X_i-\mu_0)^2-\left(X_i-\overline{X}\right)^2\right] \\
&= \frac{1}{n}\sum_{i=1}^{n}\left[\left(X_i-\mu_0-\left(X_i-\overline{X}\right)\right)\left(X_i-\mu_0+\left(X_i-\overline{X}\right)\right)\right] \\
&= \left(\overline{X}-\mu_0\right)\frac{1}{n}\sum_{i=1}^{n}\left(2X_i-\mu_0-\overline{X}\right) \\
&= \left(\overline{X}-\mu_0\right)\left(2\overline{X}-\mu_0-\overline{X}\right)=\left(\overline{X}-\mu_0\right)^2,
\end{aligned}$$

ce qu'il fallait démontrer. Revenons maintenant a notre problème en remarquant que l'équation (4.1) implique que

$$\frac{\widehat{\sigma}_0^2}{\widehat{\sigma}^2}=1+\frac{\left(\overline{X}-\mu_0\right)^2}{\widehat{\sigma}^2}=1+\frac{\left(\overline{X}-\mu_0\right)^2}{\widetilde{S}^2}.$$

En posant

$$T=\frac{\overline{X}-\mu_0}{\widetilde{S}/\sqrt{n-1}},$$

nous obtenons

$$\left(\frac{\widehat{\sigma}_0^2}{\widehat{\sigma}^2}\right)^{n/2}=\left(1+\frac{T^2}{n-1}\right)^{n/2}. \tag{4.2}$$

Donc les deux événements $\{R_1>k\}$ et $\{|T|>c\}$ sont équivalents, pour une certaine $c=c(k)$. La constante c doit être obtenue par la relation

$$\sup_{(\mu,\sigma^2)\in\Theta_0}\mathbf{P}_{(\mu,\sigma^2)}\left(|T|>c\right)=\alpha.$$

On note que la variable aléatoire T est une Student à $(n-1)$ degré de liberté :

$$T:=\frac{\overline{X}-\mu_0}{\widetilde{S}/\sqrt{n-1}}\rightsquigarrow t_{n-1}.$$

La v.a est T est standardisée, et que et sa la loi est bien déterminée et ne dépend plus de Θ_0, donc le supremum n'as plus de rôle à jouer en fin. Ainsi

$$\mathbf{P}\left(|T| > c\right) = \alpha,$$

ce qui est équivalent à

$$\mathbf{P}\left(T < -c\right) + \mathbf{P}\left(T > c\right) = \alpha.$$

Puisque la distribution de T est symétrique alors

$$\mathbf{P}\left(T < -c\right) = \mathbf{P}\left(T > c\right).$$

D'où c est telle que

$$\mathbf{P}\left(T > c\right) = \alpha/2.$$

La valeur critique c est donc le quantile d'ordre $(1 - \alpha/2)$ de t_{n-1} (Student à $n - 1$ degré de liberté) qu'on le note par $t_{1-\alpha/2}$. En conclusion, la région critique du test est

$$W := \left\{(x_1, ..., x_n) \in \mathbb{R}^n : \frac{|\overline{x} - \mu_0|}{\widetilde{s}/\sqrt{n-1}} > t_{1-\alpha/2}\right\}.$$

Application numérique :

On désigne par X la longueur (en mm) de fibres métalliques produites par une certaine unité industrielle, et on suppose que X est de distribution normale d'espérance μ et de variance σ^2 inconnues. On veut tester l'hypothèse $H_0 : \mu = 5.2$ contre $H_0 : \mu \neq 5.2$ au niveau de signification $\alpha = 0.05$. Pour cela on observe 15 fibres et on obtient une moyenne empirique $\overline{x} = 5.4$ et une variance empirique $\widetilde{s}^2 = 0.17$.

Solution :

De la table statistique des quantiles de la loi de Student on trouve

$$t_{1-\alpha/2} = t_{1-\frac{0.05}{2}} = t_{0.975} = 2.145.$$

D'autre part on a

$$\frac{|\overline{x}-\mu_0|}{\widetilde{s}/\sqrt{n-1}} = \frac{|5.4-5.2|}{\sqrt{0.17}/\sqrt{15-1}} = 1.815.$$

Comme $1.815 < 2.145$ on ne peut pas rejeter H_0 (on la garde). Nous pouvons aussi arriver au même résultat en utilisant la p-value. En effet nous avons

$$\text{p-value} = \mathbf{P}\left(|T| \geq T_{obs}\right),$$

où $T_{obs} = 1.815$. Donc

$$\begin{aligned}\text{p-value} &= \mathbf{P}\left(|T| \geq 1.815\right) = 2\mathbf{P}\left(T \geq 1.815\right) \\ &= 2\left(1-\mathbf{P}\left(T \leq 1.815\right)\right).\end{aligned}$$

De la table de loi de Student on obtient $\mathbf{P}\left(T \leq 1.815\right) = 0.95$. Ainsi

$$\text{p-value} = 2\left(1-0.95\right) = 0.1 > \alpha = 0.05,$$

Donc on ne rejette pas $H_0 : \mu = 5.2$ on la garde.

2) **Test unilatéral :**

$$\begin{cases} H_0 : & \mu \leq \mu_0 \\ H_1 : & \mu > \mu_0 \end{cases}$$

Dans ce cas $\Theta_0 = \{(\mu,\sigma^2) \mid \mu \leq \mu_0\}$, $\Theta_1 = \{(\mu,\sigma^2) \mid \mu > \mu_0\}$ et $\Theta = \Theta_0 \cup \Theta_1$. La région critique

$$W := \left\{(x_1,...,x_n) \in \mathbb{R}^n : \frac{\overline{x}-\mu_0}{\widetilde{s}/\sqrt{n-1}} > t_{1-\alpha}\right\},$$

où $t_{1-\alpha}$ est le quantile d'ordre $(1-\alpha)$ de t_{n-1} (Student à $n-1$ degré de liberté). En effet, nous avons

$$\sup_{(\mu,\sigma^2)\in\Theta_i} L\left(X_1,...,X_n;\mu,\sigma^2\right) = L\left(X_1,...,X_n;\widehat{\mu}_i,\widehat{\sigma}_i^2\right),\ i=0,1,$$

où $(\widehat{\mu}_i,\widehat{\sigma}_i^2)$ est le EMV dans H_i, $i=0,1$. Le rapport des vraisemblances maximal est donc

$$R = R\left(X_1,...,X_n\right) = \frac{L\left(X_1,...,X_n;\widehat{\mu}_1,\widehat{\sigma}_1^2\right)}{L\left(X_1,...,X_n;\widehat{\mu}_0,\widehat{\sigma}_0^2\right)}.$$

On sait que la fonction de vraisemblance $(\mu,\sigma^2)\to L\left(X_1,...,X_n;\mu,\sigma^2\right)$ atteint sont maximum global, i.e. $(\mu,\sigma^2)\in\Theta$, en $\widehat{\mu}=\overline{X}$ et $\widehat{\sigma}^2=\widetilde{S}^2$. Nous distinguons ici deux cas : $\mu_0\leq\overline{X}$ et $\mu_0\geq\overline{X}.$

1) Pour le cas $\mu_0\leq\overline{X}$, (voir Fig. 6), nous avons

$$\sup_{\mu\leq\mu_0} L\left(X_1,...,X_n;\mu,\sigma^2\right) = L\left(X_1,...,X_n;\widehat{\mu}_0,\widehat{\sigma}_0^2\right),$$

où

$$\widehat{\mu}_0=\mu_0 \text{ et } \widehat{\sigma}_0^2=\frac{1}{n}\sum_{i=1}^{n}\left(X_i-\mu_0\right)^2.$$

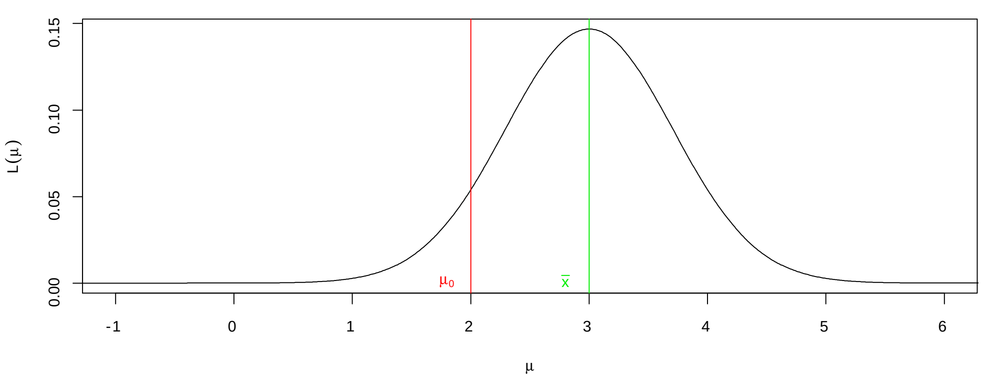

Fig.6

Tandis que

$$\sup_{\mu>\mu_0} L\left(X_1, ..., X_n; \mu, \sigma^2\right) = L\left(X_1, ..., X_n; \widehat{\mu}_1, \widehat{\sigma}_1^2\right),$$

où

$$\widehat{\mu}_1 = \overline{X} \text{ et } \widehat{\sigma}_1^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2.$$

Ainsi le rapport des vraisemblances maximal est

$$R = \left(\frac{\widehat{\sigma}_0^2}{\widehat{\sigma}_1^2}\right)^{n/2}.$$

On sait que la région critique du test rapport de vraisemblance maximal est de la forme

$$W = \{(x_1, ..., x_n) \in \mathbb{R}^n \mid R\left(x_1, ..., x_n\right) > k\},$$

où k est telle que

$$\sup_{(\mu,\sigma^2)\in\Theta_0} \mathbf{P}_{(\mu,\sigma^2)}\left\{R\left(X_1, ..., X_n\right) > k\right\} = \alpha.$$

Il nous reste donc à déterminer la constante k qui dépend de α. Comme $\overline{X} \geq \mu_0$, alors

$$T = \frac{\overline{X} - \mu_0}{\widetilde{S}/\sqrt{n-1}} \geq 0.$$

Nous avons établie à l'equation (4.2) ci-dessus que

$$R = \left(1 + \frac{T^2}{n-1}\right)^{n/2}.$$

Le fait que $R\left(x_1, ..., x_n\right) > k$ est équivalent à

$$\left(1 + \frac{T^2}{n-1}\right)^{n/2} > k \Longleftrightarrow T^2 > k' \Longleftrightarrow T > c \text{ (car } T > 0).$$

Ainsi

$$\begin{aligned}\alpha &= \sup_{(\mu,\sigma^2)\in\Theta_0} \mathbf{P}_{(\mu,\sigma^2)}\left\{R\left(X_1,...,X_n\right) > k\right\} \\ &= \sup_{(\mu,\sigma^2)\in\Theta_0} \mathbf{P}_{(\mu,\sigma^2)}\left(T > c\right) = \mathbf{P}\left(T > c\right),\end{aligned}$$

car sans paramétrage on a $T \rightsquigarrow t_{n-1}$ (Student à $n-1$ degré de liberté), ainsi $k = t_{1-\alpha}$ est le quantile d'ordre $1-\alpha$ de t_{n-1}. Notre région critique est alors

$$W := \left\{(x_1,...,x_n) \in \mathbb{R}^n : \ \frac{\overline{x}-\mu_0}{\widetilde{s}/\sqrt{n-1}} > t_{1-\alpha}\right\}.$$

2) Considérons maintenant le cas $\mu_0 \geq \overline{X}$, (voir Fig. 7). Dans cette situation nous avons

$$\sup_{\mu\leq\mu_0} L\left(X_1,...,X_n;\mu,\sigma^2\right) = L\left(X_1,...,X_n;\widehat{\mu}_0,\widehat{\sigma}_0^2\right),$$

où

$$\widehat{\mu}_0 = \overline{X} \text{ et } \widehat{\sigma}_0^2 = \widetilde{S}^2.$$

Tandis que

$$\sup_{\mu>\mu_0} L\left(X_1,...,X_n;\mu,\sigma^2\right) = L\left(X_1,...,X_n;\widehat{\mu}_1,\widehat{\sigma}_1^2\right),$$

où

$$\widehat{\mu}_1 = \mu_0 \text{ et } \widehat{\sigma}_1^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i-\mu_0\right)^2.$$

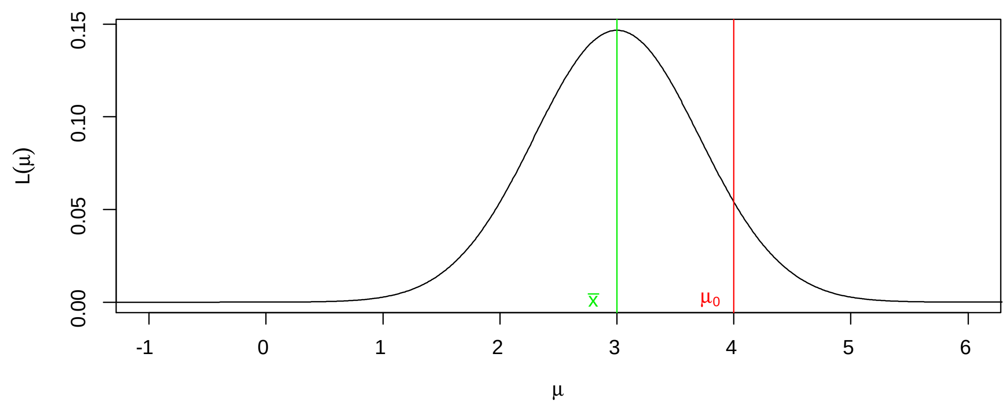

Fig. 7

Comme $\overline{X} \leq \mu_0$, donc

$$T = \frac{\overline{X} - \mu_0}{\widetilde{S}^2/\sqrt{n-1}} \leq 0.$$

La méthode utilisée pour avoir l'équation (4.1) , nous mène aussi à obtenir l'équation suivante :

$$1/R = \left(1 + \frac{T^2}{n-1}\right)^{n/2}.$$

Donc le fait que $R\left(x_1, ..., x_n\right) > k$ est équivalent à

$$\left(1 + \frac{T^2}{n-1}\right)^{n/2} < k^{-1} \Longleftrightarrow T^2 < k' \Longleftrightarrow T > c \text{ (car } T < 0),$$

et ceci nous conduit au même résultat que celui du premier cas : $\mu_0 \leq \overline{X}$.

Application : On reprend l'application numérique de l'exemple précédent pour tester :

$$\begin{cases} H_0 : & \mu \leq 5.2 \\ H_1 : & \mu > 5.2 \end{cases}$$

De la table de Student on obtient $t_{1-0.05} \equiv t_{0.95} = 1.761$. D'où l'hypothèse H_0 est à rejeter (car $1.815 > 1.761$).

Remarque :

1) La région critique du test des hypothèses

$$\begin{cases} H_0 : & \mu \geq \mu_0 \\ H_1 : & \mu < \mu_0 \end{cases}$$

et

$$W := \left\{ (x_1, ..., x_n) \in \mathbb{R}^n : \frac{\overline{x} - \mu_0}{\widetilde{s}/\sqrt{n-1}} < t_\alpha \right\}.$$

2) La fonction puissance de ces tests est basée sur la distribution de Student non centrée (voir Bickel p. 212).

Propriété asymptotique :

Théorème 1 : Soit à tester l'hypothèse

$$\begin{cases} H_0 : & \theta = \theta_0 \\ H_1 : & \theta \neq \theta_0. \end{cases}$$

Sous l'hypothèse nulle H_0, on a la distribution asymptotique suivante :

$$2 \log R \xrightarrow{\mathcal{D}} \chi_p^2, \text{ quand } n \to \infty.$$

La région critique du test est $W = \{(x_1, ..., x_n) \in \mathbb{R}^n : 2 \log R > c\}$, où c est le quantile d'ordre $1 - \alpha$ de χ_p^2.

Preuve : cas où $p = 1$.

Le développement en série Taylor de $\log R$ nous donne :

$$\log \frac{1}{R} = \log L(\theta_0) - \log L\left(\widehat{\theta}\right)$$
$$= \left(\theta_0 - \widehat{\theta}\right) \frac{\partial L\left(\widehat{\theta}\right)}{\partial \theta} + \frac{1}{2}\left(\theta_0 - \widehat{\theta}\right)^2 \frac{\partial^2 L(\theta^*)}{\partial \theta^2},$$

où θ^* est entre θ_0 est $\widehat{\theta}$. Nous avons $\widehat{\theta} \xrightarrow{ps} \theta$, ce qui implique que sous H_0 : $\widehat{\theta} \xrightarrow{ps} \theta$. D'un autre coté $\partial L\left(\widehat{\theta}\right)/\partial\theta = 0$ car $\widehat{\theta}$ est EMV de θ, d'où

$$-2\log R \sim \left(\theta_0 - \widehat{\theta}\right)^2 \frac{\partial^2 L(\theta_0)}{\partial \theta^2} = \left[\sqrt{n}\left(\theta_0 - \widehat{\theta}\right)\right]^2 \frac{1}{n}\frac{\partial^2 L(\theta_0)}{\partial \theta^2}$$

En utilisant la loi forte des grands nombres, on écrit :

$$\frac{1}{n}\frac{\partial^2 L(\theta_0)}{\partial \theta^2} = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2 f(x_i;\theta_0)}{\partial \theta^2} \xrightarrow{ps} \mathbf{E}\left[\frac{\partial^2 f(X;\theta_0)}{\partial \theta^2}\right].$$

On note que sous H_0, on a $\mathbf{E}\left[\frac{\partial^2 f(X;\theta_0)}{\partial \theta^2}\right] = -I(\theta_0)$ (la quantité d'information de Fisher), d'où

$$-2\log R \sim -\left[\sqrt{n}\left(\widehat{\theta} - \theta_0\right)\right]^2 I(\theta_0).$$

Rappelons que d'après la normalité asymptotique de $\widehat{\theta}$ on a

$$\sqrt{n}\left(\widehat{\theta} - \theta_0\right) I(\theta_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0,1), \text{ quand } n \to \infty,$$

ainsi $2\log R \sim \chi_1^2$.

Théorème 2 : Le test du rapport de vraisemblance maximale est convergent :

$$1 - \beta \to 1, \text{ quand } n \to 1.$$

Remarque. Lorsque les hypothèses sont composites, on utilisera R_1 et on a les mêmes résultats.

Exemple : Soit $X \rightsquigarrow \mathcal{P}(\lambda)$ (la loi de Poisson de paramètre $\lambda > 0$). On s'intéresse au test suivant :

$$\begin{cases} H_0: & \lambda = \lambda_0 \\ H_1: & \lambda \neq \lambda_0. \end{cases}$$

La fonction de vraisemblance est

$$L(x_1, ..., x_n; \lambda) = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!} = e^{-n\lambda} \frac{\lambda^{n\overline{x}}}{\prod_{i=1}^n x_i!}$$

Les estimateurs de maximum de vraisemblance de λ sous H_0 et H_1 sont respectivement $\widehat{\lambda}_0 = \lambda_0$ et $\widehat{\lambda} = \overline{x}$, ainsi le rapport des vraisemblances maximal est

$$R = e^{-(\overline{x}-\lambda_0)} \left(\frac{\overline{x}}{\lambda_0}\right)^{n\overline{x}}.$$

Dans le but d'effectuer un développement de Taylor de $2\log R$, au voisinage de λ_0, nous écrivons

$$\begin{aligned} 2\log R &= -2n(\overline{x} - \lambda_0) - 2n\overline{x} \log \frac{\lambda_0}{\overline{x}} \\ &= 2n(\lambda_0 - \overline{x}) - 2n\overline{x} \log\left(1 + \frac{\lambda_0 - \overline{x}}{\overline{x}}\right). \end{aligned}$$

On note que pour un $\lambda > 0$ quelconque, son estimateur de maximum de vraisemblance est $\widehat{\lambda} = \overline{X}$ qui converge en probabilité vers λ (estimateur consistent). Donc, sous H_0, $\widehat{\lambda} \xrightarrow{p} \lambda_0$. En d'autres termes pour n grand $\overline{x}$ est au voisinage de λ_0 avec qu'une probabilité qui tend vers 1. Comme $\log(1+u) = u - \frac{1}{2}u^2 + O(u^3)$, alors

$$\log\left(1 + \frac{\lambda_0 - \overline{x}}{\overline{x}}\right) = \frac{\lambda_0 - \overline{x}}{\overline{x}} - \frac{1}{2}\left(\frac{\lambda_0 - \overline{x}}{\overline{x}}\right)^2 + O_p(\lambda_0 - \overline{x})^3.$$

Ainsi

$$\begin{aligned}
&2\log R \\
&= -2n\left(\overline{x}-\lambda_0\right) - 2n\overline{x}\left[\frac{\lambda_0-\overline{x}}{\overline{x}} - \frac{1}{2}\left(\frac{\lambda_0-\overline{x}}{\overline{x}}\right)^2\right] + O_p\left[n\left(\lambda_0-\overline{x}\right)^3\right] \\
&= n\frac{\left(\lambda_0-\overline{x}\right)^2}{\overline{x}} + +O_p\left[n\left(\lambda_0-\overline{x}\right)^3\right].
\end{aligned}$$

Rappelons aussi que d'après le théorème central limite $\sqrt{n}\left(\overline{X}-\lambda_0\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0,\lambda_0^2\right)$, quand $n\to\infty$. Ceci implique que $\sqrt{n}\left(\overline{X}-\lambda_0\right)$ est asymptotiquement bornée (en probabilité), en d'autres termes $\sqrt{n}\left(\overline{X}-\lambda_0\right) = O_p(1)$. Ce qui implique que $n\left(\overline{X}-\lambda_0\right)^2 = O_p(1)$. Comme $\overline{X} \xrightarrow{p} \lambda_0$, alors $n\left(\lambda_0-\overline{x}\right)^3 \xrightarrow{p} 0$; ceci est notée $n\left(\lambda_0-\overline{x}\right)^3 = o_p(1)$.

$$2\log R = n\frac{\left(\lambda_0-\overline{x}\right)^2}{\overline{x}} + o_p(1).$$

Encore une fois le fait que

$$\sqrt{n}\left(\overline{X}-\lambda_0\right)/\lambda_0 \xrightarrow{\mathcal{D}} \mathcal{N}(0,1),$$

et $\overline{X} \xrightarrow{p} \lambda_0$, on a $\left(\sqrt{n}\left(\overline{X}-\lambda_0\right)/\lambda_0\right)^2 \xrightarrow{\mathcal{D}} \chi_1^2$ (un qui-deux à 1 degré de liberté). Par définition $o_p(1)$ est un terme qui tend vers zéro en probabilité, donc

$$2\log R \xrightarrow{\mathcal{D}} \chi_1^2, \text{ quand } n\to\infty,$$

aussi. La région critique du test est alors

$$W = \left\{(x_1,...,x_n)\in\mathbb{N}^n \mid n\frac{\left(\lambda_0-\overline{x}\right)^2}{\overline{x}} > c\right\},$$

où c est le quantile d'ordre $(1-\alpha)$ de χ_1^2.

Application numérique : $\lambda_0 = 1$, $n = 50$, la moyenne empirique observée $\overline{x}_{obs} =$

1.2 et $\alpha = 0.05$. De la table statistique de la loi de Pearson (ou de qui-deux χ_1^2), on a $c = 3.84$, or

$$50\frac{(\lambda_0 - \overline{x}_{obs})^2}{\overline{x}_{obs}} = 50\frac{(1-1.2)^2}{1.2} = 1.66 < 3.84.$$

Donc on accepte l'hypothèse H_0 qui dit que $\lambda_0 = 1$.

3) **Test de la variance d'une population** $\mathcal{N}(\mu, \sigma^2)$:

-Le cas bilatéral : Soit à tester

$$\begin{cases} H_0 : & \sigma^2 = \sigma_0^2 \\ H_1 : & \sigma^2 \neq \sigma_0^2. \end{cases}$$

Nous traitons deux cas : μ connue et μ inconnue. Le rapport de vraisemblance maximale est

$$R = \frac{\sup\limits_{(\mu,\sigma^2):\sigma^2\neq\sigma_0^2} L(x_1, ..., x_n; \sigma^2)}{\sup\limits_{(\mu,\sigma^2):\sigma^2=\sigma_0^2} L(x_1, ..., x_n; \sigma^2)} = \frac{\sup\limits_{\sigma^2\neq\sigma_0^2} L(x_1, ..., x_n; \mu, \sigma^2)}{L(x_1, ..., x_n; \mu, \sigma_0^2)}.$$

a) μ connue : Dans ce cas nous avons

$$R = \frac{L(x_1, ..., x_n; \mu, \widehat{\sigma}^2)}{L(x_1, ..., x_n; \mu, \sigma_0^2)},$$

où $\widehat{\sigma}^2 = n^{-1}\sum_{i=1}^n (\mu - x_i)^2$ est l'EMV de σ^2. Explicitement le rapport R se simplifie en

$$\left(\frac{\sigma_0^2}{\widehat{\sigma}^2}\right)^{n/2} \frac{\exp -\frac{1}{2}\sum_{i=1}^n \left(\frac{x_i-\mu}{\widehat{\sigma}}\right)^2}{\exp -\frac{1}{2}\sum_{i=1}^n \left(\frac{x_i-\mu}{\sigma_0}\right)^2}$$
$$= \left(\frac{\sigma_0}{\widehat{\sigma}^2}\right)^{n/2} \exp -\frac{1}{2}\left(\frac{1}{\widehat{\sigma}^2} - \frac{1}{\sigma_0^2}\right)\sum_{i=1}^n (x_i - \mu)^2.$$

Comme $\sum_{i=1}^{n}(x_i-\mu)^2 = n\widehat{\sigma}^2$, alors

$$R = \left(\frac{\sigma_0^2}{\widehat{\sigma}^2}\right)^{n/2} \exp\frac{n}{2}\left(\frac{\widehat{\sigma}^2}{\sigma_0^2}-1\right)$$
$$= \left[\frac{\sigma_0^2}{\widehat{\sigma}^2}\exp\left(\frac{\widehat{\sigma}^2}{\sigma_0^2}-1\right)\right]^{n/2}.$$

Le région critique est

$$W = \left\{(x_1,...,x_n)\in\mathbb{R}^n \mid \left[\frac{\sigma_0^2}{\widehat{\sigma}^2}\exp\left(\frac{\widehat{\sigma}^2}{\sigma_0^2}-1\right)\right]^{n/2} \geq c'\right\}$$
$$= \left\{(x_1,...,x_n)\in\mathbb{R}^n \mid \frac{\sigma_0^2}{\widehat{\sigma}^2}\exp\frac{\widehat{\sigma}^2}{\sigma_0^2} \geq c\right\}.$$

Posons $x = \widehat{\sigma}^2/\sigma_0^2$ et $g(x) := x^{-1}\exp x$. Cette fonction décroit sur $0 < x < 1$ puis elle croit sur $x > 1$ et elle atteint son minimum en $x = 1$ dans $x > 0$. On rappel que sous H_1 on a $\sigma^2 \neq \sigma_0^2$, en d'autres termes $x > 1$ ou $x < 1$. En vertu de a continuité et la monotonie (la bijection) de la fonction g sur chaque intervalle, on peu dire que $g(x) \geq c$ est équivalent à l'existence de deux constantes $0 < \rho_1 < 1$ et $\rho_2 > 1$, telle que $x \leq \rho_1$ ou $x \geq \rho_2$. Ainsi la région critique est

$$W = \left\{(x_1,...,x_n)\in\mathbb{R}^n \mid \frac{\widehat{\sigma}^2}{\sigma_0^2} \leq \rho_1 \text{ ou } \frac{\widehat{\sigma}^2}{\sigma_0^2} \geq \rho_2\right\}.$$

Par commodité on note $V^2 := \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$, $v^2 := \frac{1}{n}\sum_{i=1}^{n}(x_i-\mu)^2$, et $k_i/n = \rho_i$. Donc

$$W = \left\{(x_1,...,x_n)\in\mathbb{R}^n \mid \frac{nv^2}{\sigma_0^2} \leq k_1 \text{ ou } \frac{nv^2}{\sigma_0^2} \geq k_2\right\},$$

où k_1 et k_2 sont des solutions de l'équation

$$\mathbf{P}_{\sigma^2=\sigma_0^2}\left(\frac{nV^2}{\sigma_0^2} \leq k_1 \text{ ou } \frac{nV^2}{\sigma_0^2} \geq k_2\right) = \alpha.$$

En d'autres ternes

$$\mathbf{P}_{\sigma^2=\sigma_0^2}\left(\frac{nV^2}{\sigma_0^2}\leq k_1\right)+\mathbf{P}_{\sigma^2=\sigma_0^2}\left(\frac{nV^2}{\sigma_0^2}\geq k_2\right)=\alpha.$$

Comme solution de cette équation, on prend

$$\mathbf{P}_{\sigma^2=\sigma_0^2}\left(\frac{nV^2}{\sigma_0^2}\leq k_1\right)=\mathbf{P}_{\sigma^2=\sigma_0^2}\left(\frac{nV^2}{\sigma_0^2}\geq k_2\right)=\alpha/2.$$

Rappelons que $\sum_{i=1}^n\left(\frac{x_i-\mu}{\sigma_0}\right)^2 \rightsquigarrow \chi_n^2$ c'est à dire $nV^2/\sigma_0^2 \rightsquigarrow \chi_n^2$, ainsi

$$\mathbf{P}\left(\chi_n^2\leq k_1\right)=\mathbf{P}\left(\chi_n^2\geq k_2\right)=\alpha/2.$$

b) μ inconnue : en utilisant les arguments précédents, nous montrons aussi que la statistique de test (ou la variable de décision) dans ce cas est :

$$\frac{n\widetilde{S}^2}{\sigma_0^2}\rightsquigarrow\chi_{n-1}^2,$$

où $\widetilde{S}^2:=\frac{1}{n}\sum_{i=1}^n\left(X_i-\overline{X}\right)^2$. La région critique est donc

$$W=\left\{(x_1,...,x_n)\in\mathbb{R}^n \mid n\widetilde{s}^2/\sigma_0^2\leq\ell_1 \text{ ou } n\widetilde{s}^2/\sigma_0^2\geq\ell_2\right\},$$

où $\widetilde{s}^2=\frac{1}{n}\sum_{i=1}^n\left(x_i-\overline{x}\right)^2$, et ℓ_i, $i=1,2$ sont deux constantes telles que

$$\mathbf{P}\left(\chi_{n-1}^2\leq\ell_1\right)=\mathbf{P}\left(\chi_{n-1}^2\geq\ell_2\right)=\alpha/2.$$

- **Le cas unilatéral à droite** : Soit à tester

$$\begin{cases} H_0: & \sigma^2\leq\sigma_0^2 \\ H_1: & \sigma^2>\sigma_0^2. \end{cases}$$

La région critique est

$$W = \left\{ (x_1, ..., x_n) \in \mathbb{R}^n \mid \begin{array}{ll} v^2 \geq k\sigma_0^2/n & \text{si } \mu \text{ est connue} \\ \widetilde{s}^2 \geq \ell\sigma_0^2/n & \text{si } \mu \text{ est inconnue} \end{array} \right\}.$$

où k et ℓ sont deux constantes telles que

$$\mathbf{P}\left(\chi_n^2 \geq k\right) = \mathbf{P}\left(\chi_{n-1}^2 \geq \ell\right) = \alpha.$$

- **Le cas unilatéral à gauche** : Soit à tester

$$\begin{cases} H_0 : & \sigma^2 \geq \sigma_0^2 \\ H_1 : & \sigma^2 < \sigma_0^2. \end{cases}$$

La région critique est

$$W = \left\{ (x_1, ..., x_n) \in \mathbb{R}^n \mid \begin{array}{ll} v^2 \leq k\sigma_0^2/n & \text{si } \mu \text{ est connue} \\ \widetilde{s}^2 \leq \ell\sigma_0^2/n & \text{si } \mu \text{ est inconnue} \end{array} \right\},$$

où k et ℓ sont deux constantes telles que

$$\mathbf{P}\left(\chi_n^2 \leq k\right) = \mathbf{P}\left(\chi_{n-1}^2 \leq \ell\right) = \alpha.$$

Remarque : Les deux tests

$$\begin{cases} H_0 : & \sigma^2 = \sigma_0^2 \\ H_1 : & \sigma^2 > \sigma_0^2. \end{cases} \quad \text{et} \quad \begin{cases} H_0 : & \sigma^2 \leq \sigma_0^2 \\ H_1 : & \sigma^2 > \sigma_0^2. \end{cases}$$

se traitent de la même manière. Idem, les deux tests

$$\begin{cases} H_0: & \sigma^2 = \sigma_0^2 \\ H_1: & \sigma^2 < \sigma_0^2. \end{cases} \quad \text{et} \quad \begin{cases} H_0: & \sigma^2 \geq \sigma_0^2 \\ H_1: & \sigma^2 < \sigma_0^2. \end{cases}$$

se traitent aussi de la même façon.

Chapitre 5

Tests de comparaison (de homogénéité)

Ces tests se transforment en test de conformité par changement de variables.

Deux échantillons de tailles n_1 et n_2 sont extraits, indépendamment l'une de l'autre, de deux populations $X_1 \rightsquigarrow \mathcal{N}(\mu_1, \sigma_1^2)$ et $X_2 \rightsquigarrow \mathcal{N}(\mu_2, \sigma_2^2)$. Les deux échantillons sont notés $\left(X_1^{(1)}, ..., X_1^{(n_1)}\right)$ et $\left(X_2^{(1)}, ..., X_2^{(n_2)}\right)$.

4.1 **Tests d'égalité de deux moyennes de populations normales :**

a) **Hypothèse alternative bilatérale (**Cas où σ_1^2 et σ_2^2 sont connues)

$$\begin{cases} H_0 : & \mu_1 = \mu_2 \\ H_1 : & \mu_1 \neq \mu_2. \end{cases}$$

En utilisant aussi, le principe du rapport de vraisemblance maximale, on montre aussi que la statistique de test utilisée est

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \rightsquigarrow \mathcal{N}(0,1), \text{ (sous } H_0), \tag{5.1}$$

où $\overline{X}_j = n_j^{-1} \sum_{i=1}^{n_j} X_j^{(i)}$, $j = 1, 2$. La région critique associée est

$$W = \left\{ \left(x_1^{(1)}, ..., x_1^{(n_1)}; x_2^{(1)}, ..., x_2^{(n_2)} \right) \in \mathbb{R}^{n_1+n_2} \mid \frac{|\overline{x}_1 - \overline{x}_2|}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \geq z_{1-\alpha/2} \right\},$$

où $z_{1-\alpha/2}$ (la valeur critique) est le quantile d'ordre $1-\alpha/2$ de $\mathcal{N}(0,1)$, c'est à dire $z_{1-\alpha/2} = \Phi^{-1}(1-\alpha/2)$.

b) **Hypothèse alternative unilatérale à gauche (**Cas où σ_1^2 et σ_2^2 sont connues)

$$\begin{cases} H_0: & \mu_1 = \mu_2 \\ H_1: & \mu_1 < \mu_2. \end{cases} \quad \text{ou} \quad \begin{cases} H_0: & \mu_1 \geq \mu_2 \\ H_1: & \mu_1 < \mu_2. \end{cases}$$

La région critique associée est

$$W = \left\{ \left(x_1^{(1)}, ..., x_1^{(n_1)}; x_2^{(1)}, ..., x_2^{(n_2)} \right) \in \mathbb{R}^{n_1+n_2} \mid \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \leq z_\alpha \right\},$$

où z_α (la valeur critique) est le quantile d'ordre α de $\mathcal{N}(0,1)$, c'est à dire $z_\alpha = \Phi^{-1}(\alpha)$.

c) **Hypothèse alternative unilatérale à droite (**Cas où σ_1^2 et σ_2^2 sont connues)

$$\begin{cases} H_0: & \mu_1 = \mu_2 \\ H_1: & \mu_1 > \mu_2. \end{cases} \quad \text{ou} \quad \begin{cases} H_0: & \mu_1 \leq \mu_2 \\ H_1: & \mu_1 > \mu_2. \end{cases}$$

La région critique associée est

$$W = \left\{ \left(x_1^{(1)}, ..., x_1^{(n_1)}; x_2^{(1)}, ..., x_2^{(n_2)} \right) \in \mathbb{R}^{n_1+n_2} \mid \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \geq z_{1-\alpha} \right\},$$

où $z_{1-\alpha}$ (la valeur critique) est le quantile d'ordre $1-\alpha$ de $\mathcal{N}(0,1)$, c'est à dire $z_{1-\alpha} = \Phi^{-1}(1-\alpha)$.

a) **Hypothèse alternative bilatérale** (Cas où $\sigma_1^2 = \sigma_2^2$ sont inconnues)

$$\begin{cases} H_0: & \mu_1 = \mu_2 \\ H_1: & \mu_1 \neq \mu_2. \end{cases}$$

L'évident est de remplacer, dans la statistique (5.1), les variances σ_j^2 par leurs estimateurs $\widetilde{S}_j^2$, pour avoir

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\widetilde{S}_1^2}{n_1} + \frac{\widetilde{S}_2^2}{n_2}}}.$$

Comme la loi de cette statistique n'as pas une forme analytique bien déterminée, car $\widetilde{S}_1^2/n_1 + \widetilde{S}_2^2/n_2$ ne peux être une v.a de qui-deux que si $\sigma_1^2 = \sigma_2^2$. En effet supposons que $\sigma_1^2 = \sigma_2^2 = \sigma^2$, alors

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{\overline{X}_1 - \overline{X}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \rightsquigarrow \mathcal{N}(0,1).$$

D'être autre coté, on peu vérifier facilement que

$$\widetilde{S}^2 = \frac{n_1\widetilde{S}_1^2 + n_2\widetilde{S}_2^2}{n_1 + n_2 - 2},$$

est un estimateur sans biais de σ^2, où

$$\widetilde{S}_j^2 = \frac{1}{n_j}\sum_{i=1}^{n_j}\left(X_j^{(i)} - \overline{X}_j\right)^2, \ j = 1, 2.$$

En effet,

$$\begin{aligned} \mathbf{E}\left[\widetilde{S}^2\right] &= \mathbf{E}\left[\frac{n_1\widetilde{S}_1^2 + n_2\widetilde{S}_2^2}{n_1 + n_2 - 2}\right] = \frac{n_1\mathbf{E}\left[\widetilde{S}_1^2\right] + n_2\mathbf{E}\left[\widetilde{S}_2^2\right]}{n_1 + n_2 - 2} \\ &= \frac{n_1\left(\frac{n_1-1}{n_1}\right)\sigma_1^2 + n_2\left(\frac{n_2-1}{n_2}\right)\sigma_2^2}{n_1 + n_2 - 2} = \sigma^2. \end{aligned}$$

En outre nous avons

$$
\begin{aligned}
& (n_1+n_2-2)\,\widetilde{S}^2/\sigma^2 \\
& = n_1\widetilde{S}_1^2/\sigma_1^2 + n_2\widetilde{S}_2^2/\sigma_2^2 \\
& = \sum_{i=1}^{n_1}\left(X_1^{(i)}-\overline{X}_1\right)^2/\sigma_1^2 + \sum_{i=1}^{n_2}\left(X_2^{(i)}-\overline{X}_2\right)^2/\sigma_2^2 \\
& = \chi^2_{n_1-1} + \chi^2_{n_2-1} = \chi^2_{n_1+n_2-2},
\end{aligned}
$$

où $\chi^2_{n_1-1}$ et $\chi^2_{n_2-1}$ sont, respectivement, deux v.a de qui-deux indépendantes à n_1-1 et n_2-1 degrés de liberté. Rappelons que, étant donné $Z \rightsquigarrow \mathcal{N}(0,1)$ et $Q \rightsquigarrow \chi^2_m$ indépendantes, alors

$$\frac{Z}{\sqrt{Q/m}} \rightsquigarrow t\,(m)\,,$$

où $t\,(m)$ est une loi de Student à m degré de liberté. En prenant

$$Z := \frac{\overline{X}_1-\overline{X}_2}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}},\ Q := (n_1+n_2-2)\,\frac{\widetilde{S}^2}{\sigma^2}$$

et $m := n_1+n_2-2$, on obtient (après simplifications)

$$\frac{\overline{X}_1-\overline{X}_2}{\widetilde{S}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \rightsquigarrow t\,(n_1+n_2-2)\,,$$

où $t\,(n_1+n_2-2)$ est la loi de Student à n_1+n_2-2 degré de liberté. La région critique associée est

$$W = \left\{\left(x_1^{(1)},...,x_1^{(n_1)};x_2^{(1)},...,x_2^{(n_2)}\right)\in\mathbb{R}^{n_1+n_2} \mid \frac{|\overline{x}_1-\overline{x}_2|}{\widetilde{s}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \geq t_{1-\alpha/2}\right\},$$

où

$$\widetilde{s} := \sqrt{\frac{n_1\widetilde{s}_1^2+n_2\widetilde{s}_2^2}{n_1+n_2-2}},$$

et $t_{1-\alpha/2}$ (la valeur critique) est le quantile d'ordre $1-\alpha/2$ de $t\left(n_1+n_2-2\right)$.

b) **Hypothèse alternative unilatérale à droite (**cas où σ_1^2 et σ_2^2 inconnues mais sont égales)

$$\left\{\begin{array}{ll} H_0: & \mu_1=\mu_2 \\ & \\ H_1: & \mu_1>\mu_2. \end{array}\right. \text{ ou } \left\{\begin{array}{ll} H_0: & \mu_1<\mu_2 \\ & \\ H_1: & \mu_1>\mu_2. \end{array}\right.$$

La région critique associée est

$$W=\left\{\left(x_1^{(1)},...,x_1^{(n_1)};x_2^{(1)},...,x_2^{(n_2)}\right)\in\mathbb{R}^{n_1+n_2}\mid\frac{\overline{x}_1-\overline{x}_2}{\widetilde{s}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\geq t_{1-\alpha}\right\},$$

où $t_{1-\alpha}$ (la valeur critique) est le quantile d'ordre $1-\alpha$ de $t\left(n_1+n_2-2\right)$.

c) **Hypothèse alternative unilatérale à gauche (**cas où σ_1^2 et σ_2^2 inconnues mais sont égales)

$$\left\{\begin{array}{ll} H_0: & \mu_1=\mu_2 \\ & \\ H_1: & \mu_1<\mu_2. \end{array}\right. \text{ ou } \left\{\begin{array}{ll} H_0: & \mu_1>\mu_2 \\ & \\ H_1: & \mu_1<\mu_2. \end{array}\right.$$

La région critique associée est

$$W=\left\{\left(x_1^{(1)},...,x_1^{(n_1)};x_2^{(1)},...,x_2^{(n_2)}\right)\in\mathbb{R}^{n_1+n_2}\mid\frac{\overline{x}_1-\overline{x}_2}{\widetilde{s}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\leq t_{\alpha}\right\}, \tag{5.2}$$

où t_α (la valeur critique) est le quantile d'ordre α de $t\left(n_1+n_2-2\right)$.

Test de Welch (cas $\sigma_1^2\neq\sigma_2^2$). Si n_1 et n_2 sont supérieurs à 20, la région critique (5.2) reste encore approximativement valable. Pour les petites échantillons, ceci devient un problème délicat. En effet, comme les variances ne sont pas égales, on ne peut pas effectuer une modification pour avoir une variable de Student. Pour résoudre ce problème, Welch (1947) a obtenu une approximation de la loi de cette statistique par la loi de Student à ν degré de liberté où

$$\nu := \frac{\left(\widetilde{s}_1^2/n_1 + \widetilde{s}_2^2/n_2\right)^2}{\widetilde{s}_1^4/\left(n_1^2\left(n_1-1\right)\right) + \widetilde{s}_2^4/\left(n_2^2\left(n_2-1\right)\right)}. \tag{5.3}$$

Plus précisément, la statistique de test utilisée est

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\widetilde{S}_1^2/n_1 + \widetilde{S}_2^2/n_2}} \simeq t\left(\nu\right),$$

où $t\left(\nu\right)$ est la loi de Student à ν degré de liberté. On note qu'en vertu de l'équation (5.3), on montre que

$$\min\left(n_1, n_2\right) - 1 \leq \nu \leq n_1 + n_2 - 2. \tag{5.4}$$

La région critique associée est

$$W = \left\{ \left(x_1^{(1)}, ..., x_1^{(n_1)}; x_2^{(1)}, ..., x_2^{(n_2)}\right) \in \mathbb{R}^{n_1+n_2} \mid \frac{\left|\overline{x}_1 - \overline{x}_2\right|}{\sqrt{\widetilde{s}_1^2/n_1 + \widetilde{s}_2^2/n_2}} \geq t_{1-\alpha/2}\left(\nu\right) \right\},$$

où $t_{1-\alpha/2}\left(\nu\right)$ (la valeur critique) est le quantile d'ordre $1-\alpha/2$ de $t\left(\nu\right)$. Les régions critiques pour les tests unilatéraux se font de la même manière. En conclusion le plus proche de la réalité est le test de Welch.

Il est évident que la valeur de ν calculée, dans (5.3), n'est pas nécessairement un entier naturel. Dans ce cas, on effectue une interpolation linéaire afin de calculer la valeur critique associée $t_{1-\alpha/2}\left(\nu\right)$. Rappelons que l'interpolation linéaire d'une fonction f au point $x \in [a, b]$ est donnée par :

$$f\left(x\right) \simeq f\left(a\right) + \left(x - a\right)\frac{f\left(b\right) - f\left(a\right)}{b - a}.$$

En prenant $f = t_{1-\alpha/2}$, $x = \nu$, $a = \lfloor \nu \rfloor$, et $b = \lceil \nu \rceil$, on écrit :

$$t_{1-\alpha/2}(\nu) \simeq t_{1-\alpha/2}(\lfloor \nu \rfloor) + (\nu - \lfloor \nu \rfloor) \frac{t_{1-\alpha/2}(\lceil \nu \rceil) - t_{1-\alpha/2}(\lfloor \nu \rfloor)}{\lceil \nu \rceil - \lfloor \nu \rfloor},$$

où $\lfloor \nu \rfloor$ et $\lceil \nu \rceil$ désignent la partie inférieure et la partie supérieure de ν respectivement.

Application : Prenons le test

$$\begin{cases} H_0 : & \mu_1 = \mu_2 \\ H_1 : & \mu_1 \neq \mu_2, \end{cases}$$

avec $\alpha = 0.05$, $(n_1 = 34, \overline{x}_1 = 8.103, \widetilde{s}_1 = 0.507)$ et $(n_2 = 25, \overline{x}_2 = 8.503, \widetilde{s}_2 = 0.816)$.
En remplaçant dans la formule (5.3) on obtient

$$\nu = \frac{\left((0.507)^2 /34 + (0.816)^2 /25\right)^2}{(0.507)^4 / (34^2 (34-1)) + (0.816)^4 / (25^2 (25-1))} = 37.369.$$

Il est clair que les deux inégalités (5.4) sont vérifiées. On effet

$$24 = (25-1) \leq 37.369 \leq 34 + 25 - 2 = 57.$$

Nous avons $\lfloor \nu \rfloor = 37$, $\lceil \nu \rceil = 38$, et en utilisant le langage R on obtient

$$t_{0.975}(37) = 2.026 \text{ et } t_{0.975}(38) = 2.024.$$

Ainsi

$$\begin{aligned} t_{0.975}(37.369) &= \frac{37.369 - 37}{38 - 37}(2.024 - 2.026) + 2.026 \\ &= 2.025. \end{aligned}$$

La région critique est donc

$$W = \left\{ \left(x_1^{(1)}, ..., x_1^{(34)}; x_2^{(1)}, ..., x_2^{(25)} \right) \in \mathbb{R}_+^{59} \mid \frac{|\overline{x}_1 - \overline{x}_2|}{\sqrt{\widetilde{s}_1^2/34 + \widetilde{s}_2^2/25}} \geq 2.025 \right\}.$$

La valeur observée de la statistique de test est

$$\frac{|\overline{x}_{1,obs} - \overline{x}_{2,obs}|}{\sqrt{\widetilde{s}_1^2/34 + \widetilde{s}_2^2/25}} = \frac{|8.103 - 8.503|}{\sqrt{0.507^2/34 + 0.816^2/25}} = 2.163 > 2.025.$$

Donc on rejette H_0 pour accepter $H_1 : \mu_1 \neq \mu_2$.

Application en langage R :

```
x = rnorm(34,8,0.5)

y = rnorm(25,8.5,0.8)

t.test(x,y)
```

—————————————————————————

Résultat :

```
    Welch Two Sample t-test

data : x and y

t = -2.7032, df = 30.104, p-value = 0.01118

alternative hypothesis : true difference in means is not equal to 0

95 percent confidence interval :

-0.8850528 -0.1233278

sample estimates :

mean of x mean of y

7.957330 8.461521
```

On remarque que la $p-value = 0.01118 < 0.05$ (seuil de signification), ce qui conduit à rejeter H_0 pour accepter $H_1 : \mu_1 \neq \mu_2$.

Remarque. Si les populations ne sont pas normales (populations quelconques) avec de grandes échantillons ($n_1, n_2 > 30$) on utilise les approximations gaussiennes (théorème central limite) suivantes :

– Si σ_1^2 et σ_2^2 connues (non nécessairement égales) :

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \simeq \mathcal{N}(0,1) \ \text{(sous } \mu_1 = \mu_2\text{)}.$$

– Si σ_1^2 et σ_2^2 inconnues (non nécessairement égales) :

$$\frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\widetilde{S}_1^2/n_1 + \widetilde{S}_2^2/n_2}} \simeq \mathcal{N}(0,1) \ \text{(sous } \mu_1 = \mu_2\text{)}.$$

4.1 **Tests d'égalité de deux variances de populations normales : (moyennes inconnues)**

$$\begin{cases} H_0 : & \sigma_1^2 = \sigma_2^2 \\ H_1 : & \sigma_1^2 \neq \sigma_2^2. \end{cases}$$

La statistique de test (variable de décision) à utiliser est :

$$\frac{\frac{n_1 \widetilde{S}_1^2/\sigma_1^2}{n_1 - 1}}{\frac{n_2 \widetilde{S}_2^2/\sigma_2^2}{n_2 - 1}} \rightsquigarrow F(n_1 - 1, n_2 - 1), \quad \left(\text{sous } H_0 : \sigma_1^2 = \sigma_2^2\right), \tag{5.5}$$

où $F(n_1 - 1, n_2 - 1)$ désigne la loi de Fisher à $(n_1 - 1, n_2 - 1)$ degrés de liberté. A un seuil de signification α, la région critique associée est

$$W = \left\{ \left(x_1^{(1)}, ..., x_1^{(n_1)}; x_2^{(1)}, ..., x_2^{(n_2)}\right) \in \mathbb{R}^{n_1+n_2} \mid \frac{\widetilde{s}_1^2}{\widetilde{s}_2^2} \geq \frac{n_2(n_1-1)}{n_1(n_2-1)} c_1 \text{ ou } \frac{\widetilde{s}_1^2}{\widetilde{s}_2^2} \leq \frac{n_2(n_1-1)}{n_1(n_2-1)} c_2 \right\},$$

où c_1 et c_2 sont deux constantes telles que

$$\mathbf{P}(F(n_1 - 1, n_2 - 1) \geq c_1) = \mathbf{P}(F(n_1 - 1, n_2 - 1) \leq c_2) = \alpha/2.$$

Exemple. On veut comparer les dispersions des dépenses hebdomadaires des étudiants de deux Universités A et B a un niveau de signification de 5%. Pour cela il sélectionne deux échantillons aléatoires de 20 et 30 étudiants respectivement et obtient les réponses suivantes :

A	$120, 150, 180, 200, 130, 150, 170, 160, 190, 100,$ $125, 145, 175, 200, 120, 130, 135, 165, 150, 180$
B	$115, 118, 135, 185, 195, 170, 155, 180, 191, 200$ $100, 98, 105, 135, 145, 155, 118, 120, 112, 130$ $118, 125, 135, 155, 165, 156, 187, 198, 127, 130$

Note test est de la forme

$$\begin{cases} H_0 : \quad \sigma_A^2 = \sigma_B^2 \\ H_1 : \quad \sigma_A^2 \neq \sigma_B^2. \end{cases}$$

Tout d'abord on doit vérifier la normalité de deux échantillons. Pour cela on va utiliser **le test de normalité de Shapiro&Wilk** (Shapiro and Wilk, 1964) à l'aide du logiciel R :

```
A<-c(120,150,180,200,130,150,170,160,190,100,125,145,
175,200,120,130,135,165,150,180)

B<-c(115,118,135,185,195,170,155,180,191,200,100,98,105,135,145,155,118,
120,112,130,118,125,135,155,165,156,187,198,127,130)

shapiro.test(A)

shapiro.test(B)
```

Après l'excusions, voici les résultats obtenus

```
Shapiro-Wilk normality test

data : A
```

W = 0.96807, **p-value = 0.7138**

et

Shapiro-Wilk normality test

data : B

W = 0.93314, p-value=0.05952

Pour les deux cas la p-valeur est supérieure à 0.05, donc on accepte la normalité des deux échantillons.

La région critique de ce test est

$$W = \left\{ \left(x_1^{(1)}, ..., x_1^{(20)}; x_2^{(1)}, ..., x_2^{(30)} \right) \in \mathbb{R}_+^{50} \mid \frac{\widetilde{s}_1^2}{\widetilde{s}_2^2} \geq \frac{30\,(20-1)}{20\,(30-1)} c_1 \text{ ou } \frac{\widetilde{s}_1^2}{\widetilde{s}_2^2} \leq \frac{30\,(20-1)}{20\,(30-1)} c_2 \right\}.$$

En d'autres termes

$$W = \left\{ \left(x_1^{(1)}, ..., x_1^{(20)}; x_2^{(1)}, ..., x_2^{(30)} \right) \in \mathbb{R}_+^{50} \mid \frac{\widetilde{s}_1^2}{\widetilde{s}_2^2} \geq 0.982\, c_1 \text{ ou } \frac{\widetilde{s}_1^2}{\widetilde{s}_2^2} \leq 0.982\, c_2 \right\},$$

où $\mathbf{P}\left(F\left(19,29\right) \geq c_1\right) = \mathbf{P}\left(F\left(19,29\right) \leq c_2\right) = 0.05/2 = 0.025$. En utilisant le logiciel R on écrit :

```
##### le quantile d'ordre 0.975 de Fisher à (19,29) degŕe de liberté

c1<-qf(0.975,19,29)

##### le quantile d'ordre 0.025 de Fisher à (19,29) degré deliberté

c2<-qf(0.025,19,29)
```

Après l'excusion, on obtient : $c_1 = 2.231$ et $c_2 = 0.416$, ainsi la région critique est

$$W = \left\{ \left(x_1^{(1)}, ..., x_1^{(20)}; x_2^{(1)}, ..., x_2^{(30)} \right) \in \mathbb{R}_+^{50} \mid \right.$$
$$\left. \frac{\widetilde{s}_1^2}{\widetilde{s}_2^2} \geq 2.190 \text{ ou } \frac{\widetilde{s}_1^2}{\widetilde{s}_2^2} \leq 0.408 \right\}.$$

Les valeurs observées des variances des deux échantillons A et B sont, respectivement, $\widetilde{s}_{1,obs}^2 = 772.187$ et $\widetilde{s}_{2,obs}^2 = 936.630$. On remarque que le rapport

$$\frac{\widetilde{s}_{1,obs}^2}{\widetilde{s}_{2,obs}^2} = \frac{772.187}{953.0622} = 0.810\,22,$$

ni ≥ 2.190 ni ≤ 0.408, donc les deux échantillons n'appartiennent pas à la région de rejet W. Ce qui signifie que l'hypothèse qui prétende que les dispersions des dépenses hebdomadaires des étudiants de deux Universités A et B sont égales est en effet vraie.

p-value : dans le cas des distributions asymétriques (comme Qui-deux ou Fisher) la règle de décision (de rejet), basée sur la p-value, est

$$\mathbf{P}(\mathbb{U} \leq u_{obs}) < \alpha/2 \text{ ou } \mathbf{P}(\mathbb{U} \geq u_{obs}) < \alpha/2, \tag{5.6}$$

où $\mathbb{U}$ est la statistique de test utilisée et u_{obs} sa valeur observée correspondante. Plus précisement, si $\mathbf{P}(\mathbb{U} \leq u_{obs}) < \alpha/2$ ou $\mathbf{P}(\mathbb{U} \geq u_{obs}) < \alpha/2$, ont rejette l'hypothèse nulle H_0; autrement dit on accepte l'hypothèse alternative H_1. Les deux inégalités précédentes (5.6) peuvent être résumées en une seule, comme suit :

$$\min\{\mathbf{P}(\mathbb{U} \leq u_{obs}), \mathbf{P}(\mathbb{U} \geq u_{obs})\} < \alpha/2.$$

Ainsi la règle de décision est

$$p - value := 2\min\{\mathbf{P}(\mathbb{U} \leq u_{obs}), \mathbf{P}(\mathbb{U} \geq u_{obs})\} < \alpha. \tag{5.7}$$

Comme application à notre cas, on écrit

$$p - value = 2\min\left\{\mathbf{P}\left(\frac{\widetilde{S}_1^2}{\widetilde{S}_2^2} \leq 0.81022\right), \mathbf{P}\left(\frac{\widetilde{S}_1^2}{\widetilde{S}_2^2} \geq 0.81022\right)\right\}.$$

En multipliant les deux membres des deux inégalités par $\frac{20(30-1)}{30(20-1)}$ on obtient

$$\begin{aligned} p - value &= 2\min\left\{\mathbf{P}\left(F\left(19,29\right) \leq 0.824\,43\right), \mathbf{P}\left(F\left(19,29\right) \geq 0.824\,43\right)\right\} \\ &= 2\min\left\{0.6646258, 0.3353742\right\} = 0.670\,75. \end{aligned}$$

La p-value ici est $0.670\,75 > 0.05$ ce qui conduit à la même conclusion. Rappelons que la p-valeur, c'est la plus petite des valeurs de α pour lesquelles la valeur observée de la statistique de test conduit au rejet de H_0. C'est donc la probabilité d'obtenir, sous H_0, la valeur observée de la statistique de test ou une valeur plus extrême.

Nous allons vérifier notre résultat en utilisant le test de Fisher implimenté au langage R.

```
var.test(A,B)

    F test to compare two variances

data : A and B

F = 0.82443, num df = 19, denom df = 29, p-value = 0.6708

alternative hypothesis : true ratio of variances is not equal to 1

95 percent confidence interval :

 0.3694892 1.9802375

sample estimates :

ratio of variances

 0.8244316
```

La $p - value = 0.6708 > 0.05$, ceci nous conduit en effet à garder l'hypothèse nulle H_0.

Remarque : comme vous remarquez, pour le calcul des quantiles, j'ai utilisé le langage R, car les tableaux statistiques sont limités. En effet, dans le tableau statistique des quantiles de la loi de Fisher, vous pouvez vérifier que le quantile d'ordre 0.025 pour le couple de degrés de liberté $(19, 29)$ n'est pas tabulé. Dans ce cas on ait obligé de passer par *l'interpolation bilinéaire*, et ceci demande un calcul manuel relativement pénible.

4.3 **Tests d'égalité (homogénéité) de deux proportions :**

On extrait de deux échantillons de grandes tailles n_1 et n_2 (indépendants) de 2 populations différentes de proportions p_1 and p_2 par rapport à certain caractère. Sous l'hypothèse "$p_1 = p_2$", la statistique de test est

$$\frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}\,(1-\widehat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \rightsquigarrow \mathcal{N}(0,1),$$

où

$$\widehat{p} := \frac{n_1\widehat{p}_1 + n_2\widehat{p}_2}{n_1+n_2},$$

et l'estimateur commun de $p_1$et p_2, à condition que

$$n_1 + n_2 > 30,\ \ n_i p_i \geq 5 \text{ et } n_i\,(1-p_i) \geq 5,\ \text{ pour } i = 1, 2.$$

Les régions critiques sont :

– Pour le test bilatéral

$$\begin{cases} H_0 : & p_1 = p_2 \\ H_1 : & p_1 \neq p_2 \end{cases}$$

est

$$W = \left\{ \left(x_1^{(1)}, ..., x_{n_1}^{(1)}; x_2^{(2)}, ..., x_{n_2}^{(2)} \right) \in \{0,1\}^{n_1+n_2} \mid \frac{|\widehat{p}_1 - \widehat{p}_2|}{\sqrt{\widehat{p}\,(1-\widehat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \geq z_{1-\alpha/2} \right\}.$$

– Pour le test unilatéral à droite

$$\begin{cases} H_0 : & p_1 \leq p_2 \\ H_1 : & p_1 > p_2 \end{cases}$$

est

$$W = \left\{ \left(x_1^{(1)}, ..., x_{n_1}^{(1)}; x_2^{(2)}, ..., x_{n_2}^{(2)} \right) \in \{0,1\}^{n_1+n_2} \mid \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}\,(1-\widehat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \geq z_{1-\alpha} \right\}.$$

– Pour le test unilatéral à gauche :

$$\begin{cases} H_0 : & p_1 \geq p_2 \\ H_1 : & p_1 < p_2 \end{cases}$$

est

$$W = \left\{ \left(x_1^{(1)}, ..., x_{n_1}^{(1)}; x_2^{(2)}, ..., x_{n_2}^{(2)} \right) \in \{0,1\}^{n_1+n_2} \mid \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\widehat{p}\,(1-\widehat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} \leq z_{\alpha} \right\}.$$

Bibliographie

[1] Borovkov, A. A (1998). *Mathematical Statistics.* The Gordon and Breach Science Publisher.

[2] Le Jeune, M. (2010). *Statistique : la théorie et ses applications.* $2^{\text{ème}}$édition, Edition Springer.

[3] Rohatgi, V. K. & Ehsanes Saleh, A. K. Md. (2015). *An Introduction to Probability and Statistics.* Wiley, Third Edition.

[4] Saporta, G. (2006). *Probabilités Analyse des données et Statistique*, $2^{\text{ème}}$édition, Edition Technip.