

Université Mohamed Khider, Biskra  
Faculté des Sc. Exactes et Sc. de la Nature et la Vie  
Département de Mathématiques  
Master 1: 2020/2021

## Cours de Tests Statistiques (à suivre)

### 1. 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,...).

### 2. Notions générales

**Definition 1.** 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 hypothèse nulle et hypothèse alternative.

**Exemple 1:** Soient  $\Theta_0$  et  $\Theta_1$  deux sous ensembles disjoints de  $\Theta$ . On veut savoir si la valeur du paramètre  $\theta$  est dans  $\Theta_0$  ou  $\Theta_1$ , 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  $\theta$  qui n'appartiennent pas à  $\Theta_0 \cup \Theta_1$  ne sont pas envisagées, on peut admettre que  $\Theta = \Theta_0 \cup \Theta_1$ .

**2.1. Catégories de tests:** 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:

2.1.1. *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.

2.1.2. *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.

2.1.3. *Tests d'ajustement:* Ils servent à vérifier si un échantillon provient d'une population de distribution donnée.

2.1.4. *Tests d'indépendance:* Ils servent à vérifier si deux ou plusieurs populations sont indépendantes.

**2.2. 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.

**2.3. 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 est 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  $\Theta$  ayant plus d'un élément.

**2.4. 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 aboutira à choisir  $H_0$  ou  $H_1$  quatre situations peuvent être envisagées:

- accepter  $H_0$  et elle est vraie.
- rejeter  $H_0$  et elle est fautive.
- accepter  $H_0$  alors qu'elle est fautive
- 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 vraie appelée *erreur de première espèce* et sa probabilité appelée *risque de première espèce*. On le note  $\alpha$ .

L'erreur commise en acceptant une hypothèse fautive est appelée *erreur de deuxième espèce* et sa probabilité appelée *risque de deuxième espèce*. On le note  $\beta$ .

On a donc:

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

et

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

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

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

Tab.1

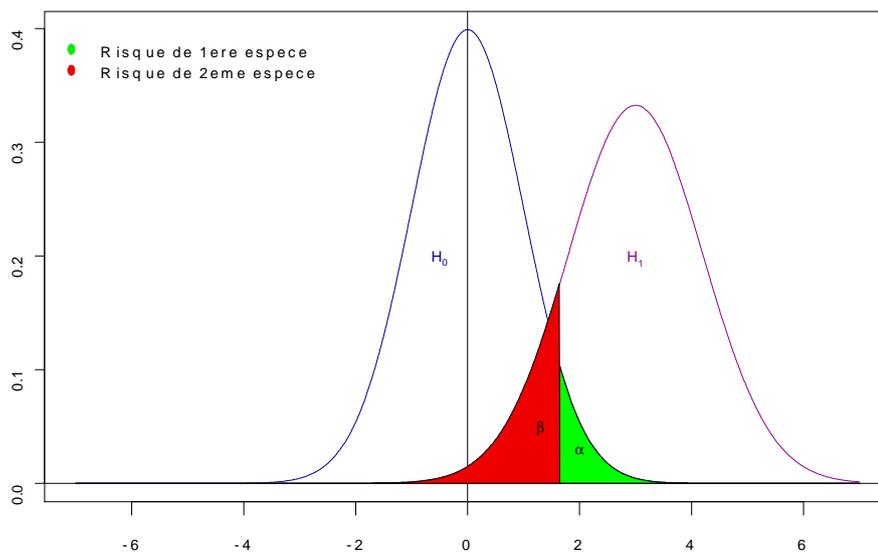


Fig.1

On remarque à la vue du graphique, que plus la différence entre les deux populations est faible, plus le risque  $\beta$  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  $\alpha$  et  $\beta$  sont toutes en sens petites. Ceci n'est pas possible car  $\alpha$  et  $\beta$  en sens contraires. Si on diminue  $\alpha$  (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  $\beta$ .

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.

2.4.1. *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

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

Lorsque  $H_1$  est composite, la puissance est variable sur  $\Theta_1$ . De même lorsque  $H_0$  est composite, le risque de première espèce est variable sur  $\Theta_0$ . On définit alors une fonction sur l'ensemble  $\Theta$  qu'on appelle fonction puissance

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

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

2.4.2. *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.

2.4.3. *Région de rejet et région critique.* La *région de rejet* d'un test est l'ensemble des points  $(X_1, \dots, X_n)$  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:

$$\mathbf{P}(W \mid H_0) = \alpha \quad (2.1)$$

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

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

**Remarque 1:**

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

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

(2) La 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 connaître la loi de probabilité de la variable de décision sous l'hypothèse  $H_0$ .

(3) Puisque la puissance est  $\pi = \mathbf{P}[W \mid H_1]$  alors, pour son calcul, il est nécessaire de connaître la loi de probabilité de la variable de décision sous l'hypothèse  $H_1$ .

(4) On a

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

(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  $\Theta_1 = ]0, 3[ \cup ]4, +\infty[$ . On suppose que la région d'acceptation du test est

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

On rappelle que  $Y_n := \max\{x_1, \dots, x_n\}$  est l'estimateur de maximum de vraisemblance de  $\theta$  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

$$\begin{aligned} \pi(\theta) &= \mathbf{P}(W) = \mathbf{P}(Y_n < 2.94) + \mathbf{P}(Y_n > 4) \\ &= F_{Y_n}(2.94) + 1 - F_{Y_n}(4). \end{aligned}$$

- 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  $\alpha \simeq 0.10$ .

2.4.4. *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  $\Theta_1$  est composite.

Pour trancher entre  $H_0$  et  $H_1$ , à niveau de signification  $\alpha$  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  $\alpha$ ).

Un test  $\delta^*$ , de niveau de signification  $\alpha$ , est upp si et seulement si:

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

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

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

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

2.4.5. *Test sans biais*: Un test  $\delta$ , niveau de signification  $\alpha$ , est dit sans biais si  $1 - \beta \geq \alpha$ ; en d'autres termes

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

2.4.6. *Tests convergents ou consistants*: Un test est dit convergent si sa puissance tend vers 1:

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

2.4.7. *Tests entre hypothèses simples*: Soit  $X$  une variable aléatoire de densité de probabilité  $f_\theta$  où  $\theta$  est un paramètre inconnu.

Il s'agit de tester

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta = \theta_1. \end{cases} \quad (2.2)$$

Soit

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

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

On pose

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

Le rapport

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

est appelé rapport de vraisemblance.

2.4.8. *Lemme de Neyman-Pearson (cas continu)*. On suppose ici que la v.a  $X$  est continue. Un test  $\delta_k$  est le test qui rejette  $H_0$  au niveau de signification  $\alpha$ , si et seulement si le rapport de vraisemblance est au moins  $k$ , où  $k = k(\alpha) \geq 0$ .

Si  $\delta$  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  $\delta_k$  est le plus puissant. En d'autres termes, la région critique optimale est

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

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

Le test  $\delta_k$  est alors défini comme suit:

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

où est la solution de l'équation

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

Autrement dit

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

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

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

**Solution:** Le rapport de vraisemblance

$$\begin{aligned} \frac{L_1(x_1, \dots, x_n)}{L_0(x_1, \dots, x_n)} &= \frac{\prod_{i=1}^9 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}(x_i - 1)^2\right\}}{\prod_{i=1}^9 \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x_i^2\right\}} \\ &= \exp\left\{9\left(\bar{x} - \frac{1}{2}\right)\right\}, \end{aligned}$$

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

$$\begin{aligned} W &= \left\{ (x_1, \dots, x_n) \in \mathbb{R}^9 : \exp\left\{9\left(\bar{x} - \frac{1}{2}\right)\right\} \geq k \right\} \\ &= \left\{ (x_1, \dots, x_n) \in \mathbb{R}^9 : \bar{x} \geq \frac{1}{2} + \frac{\log k}{9} =: k' \right\}, \end{aligned}$$

où la constante  $k'$  est telle que:

$$0.05 = P(W | \mu = 0) = P_0(W) = P_0(\bar{X} \geq k').$$

Or sous  $H_0$ ,  $\bar{X} \rightsquigarrow \mathcal{N}(0, 1/9)$ , ainsi  $Z := 3\bar{X} \rightsquigarrow \mathcal{N}(0, 1)$ , 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:

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

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

$$\delta(x_1, \dots, x_n) = \begin{cases} 1 & \text{si } \bar{x} \geq 0.54 \\ 0 & \text{si } \bar{x} < 0.54 \end{cases} \quad (2.3)$$

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

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

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

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

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

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

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

On note que  $\delta_k$  est appelé test du rapport de vraisemblance et  $\delta_k^*$  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  $H_1 : \lambda = 2$  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:

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

et sa fonction de répartition

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

Le rapport de vraisemblance qui correspond à ce test est:

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

où  $s := x_1 + \dots + x_8$ . 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:

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

où  $S := X_1 + \dots + X_8$ . 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:

$$\mathbf{P}_0(S \leq c) = 0.90 + p \mathbf{P}_0(S = c) > 0.90. \quad (2.4)$$

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  $8\lambda = 8$ . De la table statistique

de la loi de Poisson, on remarque que la petite valeur de  $c$  vérifiant  $\mathbf{P}_0(S \leq c) > 0.90$  est  $c = 12$ , qui correspond à  $\mathbf{P}_0(S \leq 12) = 0.93$ . 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 (2.4), 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  $S > 12$  on rejette  $H_0$  à 100%. Ainsi le test le plus puissant est:

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

Il est clair que

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

2.4.10. *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

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

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

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

Nous allons passer 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

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

se qui implique  $T_{obs} > k$ , et par conséquent  $(x_1, \dots, x_n) \in W$  conduisant à la rejection de l'hypothèse nulle  $H_0$ . Le cas contraire conduit évidemment 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  $\mathcal{N}(0, 1)$  :

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

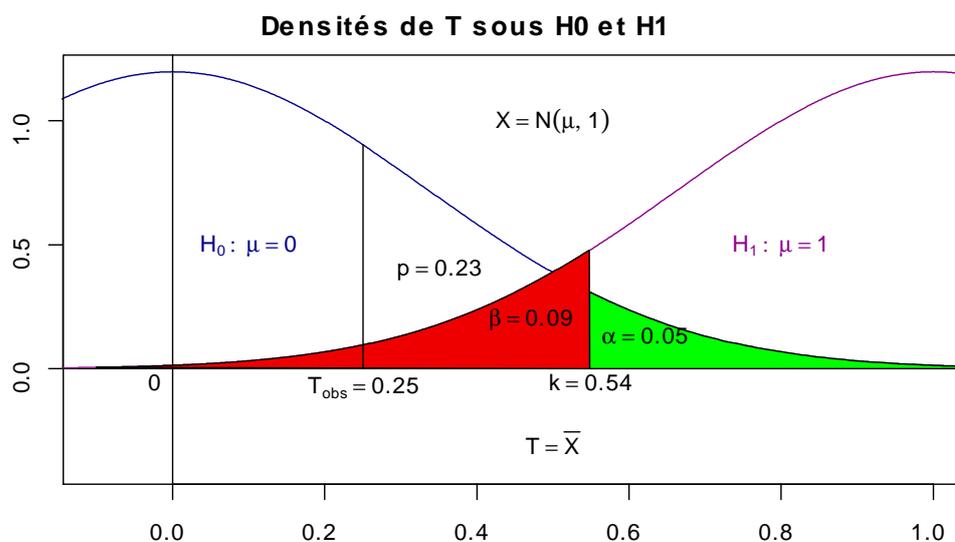
On veut tester l'hypothèse  $H_0 : \mu = 0$  contre l'hypothèse  $H_1 : \mu = 1$ . La statistique de décision dans cette exemple est  $T = \bar{X}$  et

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

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

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

Nous avons  $p = 0.23 > 0.05 = \alpha$ , ceci nous conduit à ne pas rejeter  $H_0$ . En utilisant la règle du test  $\delta$  donnée dans (2.3) en remarque que  $T_{obs} = \bar{x} = 0.25 < 0.54$  donc  $\delta = 0$ , ce qui conduit aussi à garder  $H_0$ .



### 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  $\delta$  qui correspond au hypothèses (2.2) 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  $(x_1, \dots, x_n) \in \mathbb{R}^n$ , nous examinons deux cas:

- Si  $(x_1, \dots, x_n) \in W_k$  alors  $\delta_k := \delta_k(x_1, \dots, x_n) = 1$  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  $(x_1, \dots, x_n) \notin W_k$  alors  $\delta_k(x_1, \dots, x_n) = 0$  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  $(x_1, \dots, x_n) \in \mathbb{R}^n$  :

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

Ceci implique que

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

En d'autres termes

$$\begin{aligned} & \int \delta \times L_1 dx_1 \dots dx_n - k \int \delta \times L_0 dx_1 \dots dx_n \\ & \leq \int \delta_k \times L_1 dx_1 \dots dx_n - k \int \delta_k \times L_0 dx_1 \dots dx_n, \end{aligned}$$

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  $\delta_k$  c'est le seul test le PP. En effet, soit  $\delta_k$  et  $\delta$  deux tests les PP au niveau de signification  $\alpha$ , 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

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

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  $k \geq 1$  :

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

- Si  $k < 1$  :

$$\begin{aligned} \beta &= \mathbf{P}_1(\overline{W}) = \int_{\overline{W}} L_1 dx_1 \dots dx_n < k \int_{\overline{W}} L_0 dx_1 \dots dx_n \\ &< \int_W L_0 dx_1 \dots dx_n = \mathbf{P}_0(\overline{W}) = 1 - \alpha. \end{aligned}$$

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

2.4.11. *Tests d'une hypothèse simple contre une hypothèse alternative.* 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  $\theta > \theta_0$ .

## 2.4.12. Tests entres deux hypothèses composites

. a) **Alternative unilatérale:**

$$\begin{cases} H_0 : \theta \leq \theta_0 \\ H_1 : \theta > \theta_0 \end{cases}$$

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

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

est une fonction croissante en  $t = T(x_1, \dots, x_n)$ .

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

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

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

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

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  $t \rightarrow a^n \times b^t$  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 croissant pour une statistique  $T$ , alors

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

est le test upp de  $H_0 : \theta \leq \theta_0$  contre  $H_1 : \theta > \theta_0$  au niveau de signification  $\alpha$ . Les constantes  $c$  et  $p$  se déterminent à partir de la condition

$$\mathbf{P}_{\theta=\theta_0}(T > c) + p\mathbf{P}_{\theta=\theta_0}(T = c) = \alpha.$$

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

**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

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

où  $t = \bar{x}$ ,  $a := n(\mu_1 - \mu_2) > 0$  et  $b := a - \frac{1}{2}(\mu_1 + \mu_2)$ . Il est clair que la fonction  $t \rightarrow \exp \{ at + b \}$  est une fonction croissante en  $t$ . Donc la distribution de  $X$  possède un rapport de vraisemblance croissant. Comme  $X$  est continue, le test upp le plus puissant est défini par

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

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

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

La fonction puissance du test  $\delta$  est définie par

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

En d'autres termes

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

La fonction puissance  $\pi$  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  $\mu \rightarrow \pi(\mu) = \Phi(\sqrt{65}\mu - 1.28)$ ,  $\mu \in \mathbb{R}$  est donné par la figure Fig.2

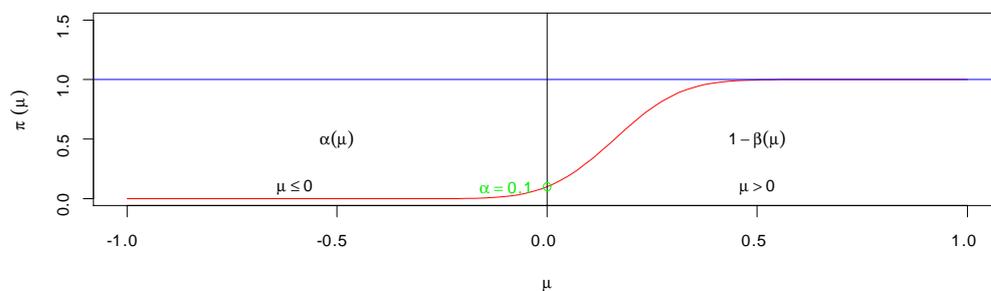


Fig.3

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))
```

---

**Remarque 3:** si nous considérons le test suivant:

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

alors le test upp sera

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

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

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

La fonction puissance du test  $\delta$  est définie par

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

En d'autres termes

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

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

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

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

Fig.4

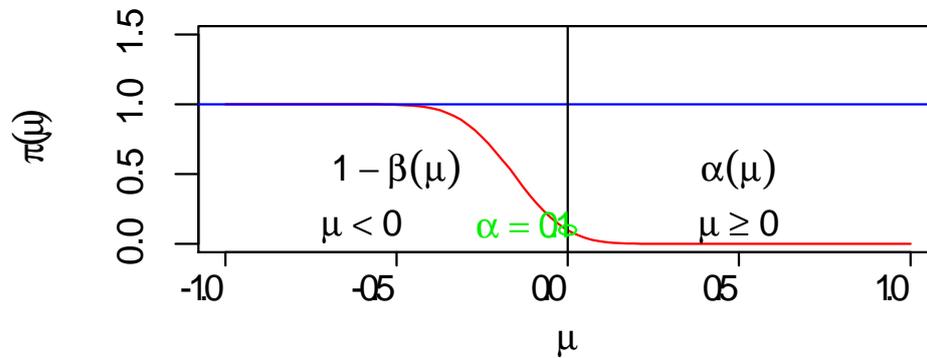


Fig.4

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  $\alpha$ . 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

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

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  $\alpha$ . 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 inférieur ou égale à  $\alpha$ . Il est clair que le test de Neyman-Pearson  $\delta_k$  est tel que  $\pi(\theta_0, \delta_k) = \alpha$  donc  $\delta_k \in \mathcal{C}$ . D'après qui précède, on peut dire que  $\delta_k$  est 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  $\delta \in \mathcal{C}' \iff \pi(\theta, \delta) \leq \alpha, \text{ pour tout } \theta \leq \theta_0 \implies \pi(\theta_0, \delta) \leq \alpha \iff \delta \in \mathcal{C}$ . D'après la première remarque ci-dessous, la fonction puissance du test  $\delta_k$  est croissante (par rapport à  $\theta$ ), d'où

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

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

#### Remarque 4:

- 1) La fonction puissance du test  $\delta_k$  [ $\pi(\theta, \delta_k, \theta \in \Xi$ )] est une fonction croissante par rapport à  $\theta$  dans  $\Xi$ , c'est à dire,  $\theta' \leq \theta'' \implies \pi(\theta', \delta_k) \leq \pi(\theta'', \delta_k)$
- 2) Pour  $\theta < \theta_0$ , le test  $\delta_k$  minimise la fonction  $\pi(\theta)$ ; dans le sens que si  $\delta$  est un autre test de niveau de signification  $\alpha$  alors pour tout  $\theta < \theta_0$ ,  $\pi(\theta, \delta_k) \leq \pi(\theta, \delta)$ .

#### Preuve de la remarque 4.

- 1) On montre que si  $\theta' \leq \theta'' \implies \pi(\theta', \delta_k) \leq \pi(\theta'', \delta_k)$ , pour 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  $\delta_k$  maximise  $\pi(\theta'', \delta)$  pour  $\delta \in \mathcal{C}$  ( $\delta_k$  est upp dans  $\mathcal{C}$ ). Soit  $\tilde{\delta}$  un test tel que  $\pi(\theta, \tilde{\delta}) = \gamma$ , pour tout  $\theta \in \Theta$ , alors  $\tilde{\delta} \in \mathcal{C}$ . Donc

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

- 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  $\delta_k^*$  est maximale pour  $\theta < \theta_0$ . La puissance du test  $\delta_k^*$  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,  $\delta_k$  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:

$$f_X(x; \theta) = \exp \{a(x) u(\theta) + b(x) + \vartheta(\theta)\}, \quad (2.5)$$

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

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

est le upp pour les hypothèses ci-dessus au niveau de signification  $\alpha$ , 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)  $\delta$  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 décroît strictement à gauche et à droite de  $\theta_0$ .
- 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 (2.5), 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

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

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

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

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:

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

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:

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

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 unilatérale:**

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

**Proposition 5.**

Soit  $X$  une population de densité:

$$f_X(x; \theta) = \exp \{a(x)u(\theta) + b(x) + \vartheta(\theta)\}, \quad (2.6)$$

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  $\alpha$ , 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

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

**Cas particulier:**  $\theta_1 = \theta_2 = \theta_0$ . Le problème de test devient:

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

Dans ce cas, la fonction puissance  $\pi(\theta)$  est minimale en  $\theta_0$ . 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:

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

**Exemple 9:**

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

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

avec  $n = 9$  et  $\alpha = 0.05$ . Le test upp est

$$\delta = \begin{cases} 1 & \text{si } \sum_{i=1}^9 x_i \leq c_1 \text{ ou } \sum_{i=1}^9 x_i \geq c_2 \\ 0 & \text{si } c_1 < \sum_{i=1}^9 x_i < c_2 \end{cases}$$

En d'autres termes

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

où  $k_i := c_i/9$ ,  $i = 1, 2$ . Nous avons

$$\begin{aligned} \pi(\mu; \delta) &= \mathbf{E}_{\mu}[\delta] = \mathbf{P}_{\mu}(\bar{X} \leq k_1 \text{ ou } \bar{X} \geq k_2) \\ &= \mathbf{P}_{\mu}(\bar{X} \leq k_1) + \mathbf{P}_{\mu}(\bar{X} \geq k_2) \\ &= \mathbf{P}_{\mu}(3(\bar{X} - \mu) \leq 3(k_1 - \mu)) + \mathbf{P}_{\mu}(3(\bar{X} - \mu) \geq (k_2 - \mu)) \\ &= \Phi(3(k_1 - \mu)) + 1 - \Phi(3(k_2 - \mu)). \end{aligned}$$

En passant à la dérivée par rapport à  $\mu$ , on trouve

$$\pi'(\mu; \delta) = -3\varphi(3k_1 - 3\mu) + 3\varphi(3k_2 - 3\mu), \text{ avec } \varphi = \Phi',$$

d'où  $\pi'(0; \delta) = 3[\varphi(3k_2) - \varphi(3k_1)]$ . En d'autres termes

$$\pi'(\mu; \delta) = \frac{3}{\sqrt{2\pi}} \left[ \exp\left\{-\frac{9k_2^2}{2}\right\} - \exp\left\{-\frac{9k_1^2}{2}\right\} \right].$$

Donc le fait que  $\pi'(\mu; \delta) = 0 \iff k_2^2 = k_1^2 \iff k_2 = -k_1$ , car  $k_2 \neq k_1$ . D'autres part

$$\begin{aligned} \mathbf{E}_{\mu=0}[\delta] = 0.05 &\iff \Phi(3k_1) + 1 - \Phi(3k_2) = 0.05 \\ &\iff 2\Phi(3k_1) = 0.05 \iff \Phi(3k_1) = 0.025. \end{aligned}$$

De la table statistique on trouve  $3k_1 = -1.96$ , d'où  $k_1 = -0.65$  et  $k_2 = 0.65$ . Le test  $\delta$  upp est donc

$$\delta = \begin{cases} 1 & \text{si } \bar{x} \leq -0.65 \text{ ou } \bar{x} \geq 0.65 \\ 0 & \text{si } -0.65 < \bar{x} < 0.65. \end{cases}$$

**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

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

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

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

Les constantes  $k_1$  et  $k_2$  solution du système

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

Ceci peut être réécrit comme suit

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

En d'autres termes

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

Ainsi nous obtenons

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

où  $\Phi$  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

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

En d'autres 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.

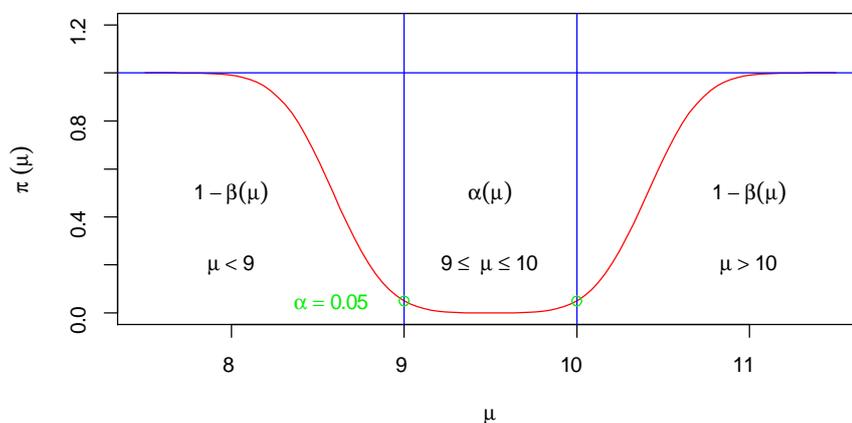


Fig.5

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)))  
-----
```