

# Tests multiples d'un continuum d'hypothèses

---

GRELA Fabrice

## **Rapport de stage**

*Stage effectué du 1er avril au 22 juin 2018*

IRMAR  
Université Rennes 2

Encadrants: Magalie FROMONT et Ronan LE GUÉVEL



# Table des matières

0.1	Introduction . . . . .	4
<b>1</b>	<b><math>p</math>-valeurs et tests d'une seule hypothèse.</b>	<b>5</b>
1.1	Fonction de répartition et inverse généralisée. . . . .	5
1.2	Notion de $p$ -valeur pour un test simple. . . . .	9
1.2.1	Erreur de première espèce, puissance d'un test et valeur critique. . . . .	9
1.2.2	Notion de $p$ -valeur. . . . .	10
1.2.3	Valeur critique versus $p$ -valeur. . . . .	13
<b>2</b>	<b>Tests multiples : procédure de tests et erreurs de première espèce.</b>	<b>17</b>
2.1	Procédure de tests multiples. . . . .	17
2.1.1	Idées générales. . . . .	17
2.1.2	Procédures de tests multiples. . . . .	18
2.1.3	Les erreurs de première espèce. . . . .	20
2.1.4	Performance du test. . . . .	22
2.2	Contrôle adaptatif de la $k$ -FWER. . . . .	22
2.2.1	Hypothèses sur les procédures. . . . .	23
2.2.2	Procédure <i>single-step</i> . . . . .	23
2.2.3	Procédure <i>step-down</i> pour la FWER. . . . .	24
2.2.4	Procédure <i>step-down</i> pour la $k$ -FWER. . . . .	25
2.2.5	Opérateur Next. . . . .	27
2.3	Contrôle de la FDR. . . . .	31
2.3.1	Procédure de Benjamini Hochberg. . . . .	31
2.3.2	Conditions de dépendance et contrôle de l'erreur de première espèce. . . . .	34
<b>3</b>	<b>Erreurs de seconde espèce et théorie minimax pour les tests multiples.</b>	<b>39</b>
3.1	Premières définitions des erreurs de seconde espèce pour les tests multiples. . . . .	39
3.1.1	Rappels sur les tests multiples et procédure $\min p$ . . . . .	39
3.1.2	Définitions de la puissance d'un test multiple. . . . .	42
3.1.3	Critères maximin. . . . .	44
3.2	Correspondance entre les tests agrégés et les tests multiples. . . . .	46
3.2.1	Tests agrégés, erreur de première espèce et théorie minimax. . . . .	46
3.2.2	Correspondance entre la théorie des tests multiples et celle des tests agrégés. . . . .	48
3.3	Théorie minimax pour les tests multiples. . . . .	49
3.3.1	Erreur de première espèce et premières identifications. . . . .	49
3.3.2	Family Wise Separation Rates pour les tests multiples. . . . .	50
3.3.3	Exemples : le cas gaussien. . . . .	54
<b>4</b>	<b>Tests multiples sur un continuum d'hypothèses et contrôle de la FDR.</b>	<b>61</b>
4.1	Tests multiples d'un continuum d'hypothèses. . . . .	61
4.1.1	Procédure de tests multiples sur un continuum d'hypothèses. . . . .	61
4.1.2	False Discovery Rate. . . . .	66
4.1.3	Procédures Step-Up. . . . .	66
4.1.4	Condition de Positively Regressively Dependent on each one of a Subset. . . . .	68
4.2	Contrôle de la FDR. . . . .	69
4.2.1	Enoncé du théorème et exemples. . . . .	69
4.2.2	Démonstration du théorème 4.2.1. . . . .	70
4.3	Conclusion. . . . .	75

## 0.1 Introduction

Les problèmes de tests multiples interviennent lors de l'analyse d'un grand nombre de données. On utilise les tests multiples lorsque l'on souhaite savoir quels paramètres agissant sur nos données ont un rôle significatif dans nos observations. Par exemple, on peut mesurer l'expression d'un ensemble de gènes chez un individu sain et un individu malade et chercher quels gènes interviennent dans la maladie.

Les tests multiples interviennent donc dans des problèmes de grande dimension : on observe beaucoup de données et l'on veut tester un grand nombre d'hypothèses simultanément. Plutôt que d'avoir un grand nombre d'observations, on peut aussi observer la trajectoire d'un processus stochastique sur un intervalle de temps  $[0, T]$  et vouloir tester une propriété en chaque instant  $t \in [0, T]$ . Par exemple, si on observe une trajectoire continue d'un processus, on peut être amené à s'intéresser au signe de la valeur moyenne  $(\mu_t)_{t \in [0, T]}$  du processus à chaque instant. Ici, on a donc besoin de tester simultanément un nombre infini indénombrable d'hypothèses... De plus, lorsque l'on effectue un test statistique, on cherche à construire une procédure de décision pour répondre à un problème (par exemple, ce gène est-il responsable de l'apparition d'une maladie?). Il faut alors prendre garde à ce que la probabilité de commettre des erreurs soit "la plus petite possible". Le but de ce stage est de définir la notion de tests multiples sur une infinité non dénombrable d'hypothèses et de chercher à contrôler la probabilité de se tromper.

Dans le premier chapitre, nous ferons essentiellement des rappels sur les tests simples, i.e. les tests d'une seule hypothèse. Après avoir défini de manière générale et abstraite la notion de  $p$ -valeur, nous rappellerons comment construire un test simple à partir d'une statistique de test ou à partir d'une  $p$ -valeur. Nous chercherons ensuite à comparer les tests construits avec ces deux approches, voir les liens et les différences qui existent entre ces tests ainsi que leurs avantages et inconvénients en pratique. Nous définirons aussi les erreurs qui peuvent être commises lors d'un test, à savoir l'erreur de première espèce (i.e. rejeter l'hypothèse alors qu'elle était vraie) et l'erreur de seconde espèce (i.e. accepter l'hypothèse alors qu'elle était fausse). Les discussions sur les  $p$ -valeurs seront notamment inspirées des annexes des articles de M. Fromont, M. Lerasle et P. Reynaud-Bouret [6] et de E. Roquain [14].

Dans un second chapitre, nous définirons ce qu'est une procédure de tests multiples d'un nombre fini d'hypothèses. Nous définirons aussi des erreurs de première espèce pour les tests multiples et établirons plusieurs résultats qui assurent le contrôle de ces erreurs, notamment la FDR et la FWER. Dans ce chapitre, nous illustrerons ces notions à travers différentes procédures de tests multiples (procédure de Bonferroni, procédure de Benjamini Hochberg...) et des problèmes de tests classiques (notamment les tests gaussiens) sous différents cadres théoriques (conditions de dépendance entre les  $p$ -valeurs, différentes définitions de procédures - procédures step-down ou step-up). Ce panorama est essentiellement dressé à partir de l'article d'Etienne Roquain [14]. Nous avons également généralisé le contrôle de la FWER à celui de la  $k$ -FWER des procédures définies de façon itérative par un opérateur Next (c.f. l'article de J. Goeman et A. Solari [7]).

Dans le troisième chapitre, nous chercherons à définir un contrôle de l'erreur de seconde espèce pour les tests multiples. Plusieurs approches issues de la littérature seront présentées mais nous nous concentrerons sur la notion de vitesse de séparation. Ce critère lié à l'erreur de seconde espèce est basé sur une correspondance entre les tests multiples et les tests agrégés et il permet de fournir une théorie d'optimalité pour les tests multiples. Ce chapitre est basé sur l'article de M. Fromont, M. Lerasle et P. Reynaud-Bouret [6].

Enfin dans un quatrième chapitre, nous verrons comment définir des tests multiples sur un continuum d'hypothèses, autrement dit, d'une infinité non dénombrable d'hypothèses. Nous généraliserons aussi au cas continu, la définition de FDR, et établirons un résultat assurant son contrôle. Nous verrons que cette généralisation impose des contraintes de mesurabilité sur lesquelles nous insisterons à travers des exemples et des contre-exemples. Ce chapitre est basé sur la lecture de l'article de G. Blanchard, S. Delattre, E. Roquain [3].

# Chapitre 1

## $p$ -valeurs et tests d'une seule hypothèse.

### 1.1 Fonction de répartition et inverse généralisée.

On rappelle les propriétés de la fonction de répartition.

#### Proposition 1.1.1

Soit  $F$  la fonction de répartition d'une variable aléatoire réelle  $X$  :

$$\forall t \in \mathbb{R}, F(t) = \mathbb{P}(X \leq t).$$

1. La fonction  $F$  est croissante (donc mesurable), continue à droite et vérifie :

$$\lim_{t \rightarrow -\infty} F(t) = 0 \quad \text{et} \quad \lim_{t \rightarrow +\infty} F(t) = 1.$$

2. Pour tout  $t \in \mathbb{R}$ , la limite à gauche de  $F$  en  $t$ ,  $F(t^-) = \lim_{x \rightarrow t, x < t} F(x)$  existe et vaut  $\mathbb{P}(X < t)$ . Ainsi, on a  $F(t) - F(t^-) = \mathbb{P}(X = t)$ . La fonction de répartition possède au plus un nombre dénombrable de points de discontinuité.
3. La fonction de répartition caractérise la loi.

Soit  $F$  une fonction définie sur  $\mathbb{R}$  à valeurs dans  $[0, 1]$ , croissante et continue à droite. Par convention, on pose  $F(-\infty) := \lim_{t \rightarrow -\infty} F(t) = 0$  et  $F(+\infty) = 1$ . Son **inverse généralisée**, noté  $F^{-1}$ , est définie par

$$F^{-1}(u) = \inf\{t \in \mathbb{R}, F(t) \geq u\}, \quad u \in [0, 1].$$

On remarque que  $F^{-1}$  est croissante par construction et que si  $\{t \in \mathbb{R}, F(t) \geq u\} = \emptyset$  alors  $u = 1$  et  $F^{-1}(u) = +\infty$ .

Si la fonction  $F$  est bijective de  $\mathbb{R}$  dans  $]0, 1[$  (par exemple dans le cas où  $X$  a une densité  $f$ ), alors elle est continue sur  $\mathbb{R}$  et l'inverse généralisée et l'inverse de  $F$  coïncident. L'inverse généralisée reste définie même lorsque  $F$  n'est pas bijective.

L'inverse généralisée de la fonction de répartition permet d'introduire la notion de quantile.

#### Définition 1.1.2 : Quantile

Soit  $X$  une variable aléatoire réelle de fonction de répartition  $F$  et  $F^{-1}$  l'inverse généralisée de  $F$ . Pour  $u \in ]0, 1]$ , la quantité  $F^{-1}(u)$  s'appelle le quantile d'ordre  $u$  ou le  $u$ -quantile de la loi de  $X$ .

**Remarque 1.1.3 :** Autre définition pour la fonction de répartition

Soit une variable aléatoire réelle  $X$ , on peut définir la fonction de répartition  $\bar{F}$  de  $X$  et  $\bar{F}^{-1}$  son inverse généralisée par<sup>1</sup> :

$$\begin{aligned}\forall t \in \mathbb{R}, \quad \bar{F}(t) &= \mathbb{P}(X < t), \\ \forall u \in [0, 1], \quad \bar{F}^{-1}(u) &= \sup\{t \in \mathbb{R}, \bar{F}(t) \leq u\}.\end{aligned}$$

Dans ce cas,  $\bar{F}$  est croissante, continue à gauche, admet une limite à droite (une telle fonction sera notée càglad) et  $\mathbb{P}(X = t) = \bar{F}(t^+) - \bar{F}(t)$ .

On remarque que  $\bar{F}^{-1}$  est croissante par construction, que si  $\{t \in \mathbb{R}, \bar{F}(t) \leq u\} = \emptyset$  alors  $u = 0$  et  $\bar{F}^{-1}(u) = -\infty$  et par passage à la limite, on a aussi  $\bar{F}(+\infty) = 1$  et  $\bar{F}(-\infty) = 0$ .

Nous allons maintenant établir des relations entre ces deux définitions de fonction de répartition.

**Proposition 1.1.4 :** Fromont, Lerasle, Reynaud-Bouret (2015) [6]

Pour tout  $t \in \mathbb{R}$  et pour tout  $u \in [0, 1]$ ,

1.  $F^{-1}(F(t)) \leq t \leq \bar{F}^{-1}(\bar{F}(t))$ ,
2.  $\bar{F}(\bar{F}^{-1}(u)) \leq u \leq F(F^{-1}(u))$ .
3.  $F^{-1}(u) \leq t \Leftrightarrow u \leq F(t)$  et  $\bar{F}^{-1}(u) \geq t \Leftrightarrow u \geq \bar{F}(t)$ .  
De plus,  $F^{-1}$  est continue à gauche et  $\bar{F}^{-1}$  est continue à droite,
4.  $\bar{F}(F^{-1}(u)) \leq u \leq F(\bar{F}^{-1}(u))$ ,
5.  $F^{-1}(u) \leq \bar{F}^{-1}(u)$ ,
6.  $F^{-1}(u) < \bar{F}^{-1}(u) \Rightarrow \bar{F}(\bar{F}^{-1}(u)) = F(F^{-1}(u)) = u$ .

*Démonstration.*

Pour tout  $u \in [0, 1]$ , on pose  $\mathcal{F}_u = \{t, F(t) \geq u\}$  et  $\mathcal{F}_u^- = \{t, \bar{F}(t) \leq u\}$  de sorte que  $F^{-1}(u) = \inf \mathcal{F}_u$  et  $\bar{F}^{-1}(u) = \sup \mathcal{F}_u^-$ .

Soit  $t \in \mathbb{R}$  et  $u \in [0, 1]$ .

1. Comme  $t \in \mathcal{F}_{F(t)}$  et  $t \in \mathcal{F}_{\bar{F}(t)}^-$ , on a

$$F^{-1}(F(t)) = \inf \mathcal{F}_{F(t)} \leq t \leq \sup \mathcal{F}_{\bar{F}(t)}^- = \bar{F}^{-1}(\bar{F}(t)).$$

2. Par définition, si  $\mathcal{F}_u \neq \emptyset$ , il existe une suite décroissante  $(t_n)_{n \in \mathbb{N}}$  de  $\mathcal{F}_u$  telle que  $t_n \rightarrow_{n \rightarrow +\infty} F^{-1}(u)$ . Comme  $F(t_n) \geq u$  pour tout  $n \in \mathbb{N}$  et comme  $F$  est càdlàg, on obtient, en faisant tendre  $n$  vers  $+\infty$  que  $F(F^{-1}(u)) \geq u$ . Si  $\mathcal{F}_u = \emptyset$  alors  $u = 1$ ,  $F^{-1}(u) = +\infty$  et  $F(F^{-1}(u)) = F(+\infty) = 1 = u$ .

De même, si  $\mathcal{F}_u^- \neq \emptyset$ , il existe une suite croissante  $(t_n^-)_{n \in \mathbb{N}}$  de  $\mathcal{F}_u^-$  telle que  $t_n^- \rightarrow_{n \rightarrow +\infty} \bar{F}^{-1}(u)$ . Comme  $\bar{F}(t_n^-) \leq u$  pour tout  $n \in \mathbb{N}$  et comme  $\bar{F}$  est càglad, on obtient  $\bar{F}(\bar{F}^{-1}(u)) \leq u$ . Si  $\mathcal{F}_u^- = \emptyset$  alors  $u = 0$ ,  $F^{-1}(u) = -\infty$  et  $\bar{F}(\bar{F}^{-1}(u)) = \bar{F}(-\infty) = 0 = u$ .

3. Supposons que  $u \leq F(t)$ . Alors  $t \in \mathcal{F}_u$  (et  $\mathcal{F}_u \neq \emptyset$ ) ce qui implique que  $F^{-1}(u) = \inf \mathcal{F}_u \leq t$ . Supposons maintenant que  $F^{-1}(u) \leq t$ . Comme  $F$  est croissante,  $F(F^{-1}(u)) \leq F(t)$  et le point 2 assure que  $F(t) \geq u$ . Supposons à présent que  $u \geq \bar{F}(t)$ . Alors  $t \in \mathcal{F}_u^-$  (et  $\mathcal{F}_u^- \neq \emptyset$ ) ce qui implique que  $F^{-1}(u) \geq t$ . Enfin, supposons que  $\bar{F}^{-1}(u) \geq t$ . Comme  $\bar{F}$  est croissante,  $\bar{F}(\bar{F}^{-1}(u)) \geq \bar{F}(t)$  et le point 2 assure que  $u \geq \bar{F}(t)$ . Ainsi pour tout  $u \in ]0, 1]$ ,

$$F^{-1}(u - \varepsilon) \leq t, \quad \forall \varepsilon \in ]0, u[ \Leftrightarrow u - \varepsilon \leq F(t), \quad \forall \varepsilon \in ]0, u[ \Leftrightarrow u \leq F(x) \Leftrightarrow F^{-1}(u) \leq t,$$

1. Cette définition pour la fonction de répartition est notamment utilisée dans l'ouvrage fondateur de Kolmogorov, *Foundations of the theory of probability*, Springer, 1933.

ce qui assure la continuité à gauche de  $F^{-1}$  sur  $[0, 1]$ . De même,  $\overline{F}^{-1}$  est continue à droite sur  $[0, 1]$  car pour tout  $u \in [0, 1[$ ,

$$\overline{F}^{-1}(u + \varepsilon) \geq t, \quad \forall \varepsilon \in ]0, 1 - u[ \iff u + \varepsilon \geq \overline{F}(t), \quad \forall \varepsilon \in ]0, 1 - u[ \iff u \geq \overline{F}(t) \iff \overline{F}^{-1}(u) \geq t.$$

4. On suppose tout d'abord que  $F^{-1}(u) \neq \pm\infty$ . Alors, pour tout  $n \geq 1$ ,  $F^{-1}(u) - 1/n < \inf \mathcal{F}_u$  donc  $F^{-1}(u) - 1/n \notin \mathcal{F}_u$  et  $\overline{F}(F^{-1}(u) - 1/n) \leq F(F^{-1}(u) - 1/n) < u$ . De plus,  $\overline{F}$  est càglàd, donc en passant à la limite sur  $n$ , on obtient  $\overline{F}(F^{-1}(u)) \leq u$ . Maintenant, si  $F^{-1}(u) = +\infty$  alors  $u = 1$  et  $\overline{F}(F^{-1}(u)) = \overline{F}(+\infty) = 1 = u$  et si  $F^{-1}(u) = -\infty$ ,  $u = 0$  et  $\overline{F}(F^{-1}(u)) = \overline{F}(-\infty) = 0 = u$ .

De la même manière, si on suppose que  $\overline{F}^{-1}(u) \neq \pm\infty$ . Alors pour tout  $n \geq 1$ ,  $\overline{F}^{-1}(u) + 1/n > \sup \mathcal{F}_u^-$  et on obtient comme précédemment que  $F(\overline{F}^{-1}(u) + 1/n) \geq \overline{F}(\overline{F}^{-1}(u) + 1/n) > u$  et par passage à la limite,  $F(\overline{F}^{-1}(u)) \geq u$ . Enfin, si  $\overline{F}^{-1}(u) = +\infty$  alors  $u = 1$  et  $F(\overline{F}^{-1}(u)) = F(+\infty) = -1 = u$  et si  $\overline{F}^{-1}(u) = -\infty$ ,  $u = 0$  et  $F(\overline{F}^{-1}(u)) = F(-\infty) = 0 = u$  ce qui termine la preuve.

5. D'après le point 4, on a d'une part,  $u \leq F(\overline{F}^{-1}(u))$  et donc d'après le point 2,  $\overline{F}^{-1}(u) \geq F^{-1}(u)$ . D'autre part,  $\overline{F}(F^{-1}(u)) \leq u$  équivaut, d'après 2, à  $F^{-1}(u) \leq \overline{F}^{-1}(u)$ .

6. Supposons que  $F^{-1}(u) < \overline{F}^{-1}(u)$ . Alors :

$$F(F^{-1}(u)) = \mathbb{P}(X \leq F^{-1}(u)) \leq \mathbb{P}(X < \overline{F}^{-1}(u)) = \overline{F}(\overline{F}^{-1}(u)).$$

Le point 2 permet de conclure. □

La proposition précédente est à la base de la méthode d'inversion de la fonction de répartition, qui peut permettre de simuler des variables aléatoires réelles comme l'indique le prochain corollaire.

### Corollaire 1.1.5

Soit  $F$  et  $F^{-1}$  (respectivement  $\overline{F}$  et  $\overline{F}^{-1}$ ) la fonction de répartition et l'inverse généralisée d'une variable aléatoire réelle  $X$ .

1. Soit  $U$  une variable aléatoire de loi uniforme sur  $[0, 1]$ . Alors la fonction  $F$  (resp.  $\overline{F}$ ) est la fonction de répartition de la variable aléatoire  $F^{-1}(U)$  (resp.  $\overline{F}^{-1}(U)$ ).
2. Soit  $X$  une variable aléatoire à valeurs dans  $\mathbb{R}$  de fonction de répartition  $F$  (resp.  $\overline{F}$ ). Si  $F$  (resp.  $\overline{F}$ ) est continue alors  $F(X)$  (resp.  $\overline{F}(X)$ ) est de loi uniforme sur  $[0, 1]$ .

*Démonstration.*

On va rédiger cette preuve pour la fonction de répartition  $F$ , la démonstration étant identique pour  $\overline{F}$ .

1. Soit  $U$  une variable uniforme sur  $[0, 1]$ . D'après la propriété 3 du lemme 1.1.4, on a pour tout  $t \in \mathbb{R}$ ,

$$\mathbb{P}(F^{-1}(U) \leq t) = \mathbb{P}(U \leq F(t)) = F(t).$$

Donc  $F$  est la fonction de répartition de la variable aléatoire  $F^{-1}(U)$ .

2. D'après le point précédent, les variables aléatoires  $X$  et  $F^{-1}(U)$  ont même fonction de répartition. Elles ont donc la même loi. Ainsi, comme  $F$  est supposée continue,  $F(X)$  à la même loi que  $F(F^{-1}(U))$ .

D'après 1.1.4.2, on a aussi que  $F(F^{-1}(U)) \geq U$ . De plus, comme  $F$  est continue,  $X$  n'a pas d'atome et on a  $\overline{F} = F$ . Ainsi, d'après 1.1.4.4,  $F(F^{-1}(U)) = \overline{F}(F^{-1}(U)) \leq U$ .

On en déduit que  $F(F^{-1}(U)) = U$  et donc que  $F(X)$  et  $U$  ont même loi. □

On termine cette section par un corollaire qui nous sera utile pour comparer des tests entre-eux.

**Corollaire 1.1.6**

Soit  $\mathcal{D}_F$  l'ensemble des points de discontinuité de  $F$ . Pour tout  $u \in [0, 1]$ , on a :

$$\begin{aligned}\mathbb{1}_{\overline{F}(X)=u} &= \mathbb{1}_{X=\overline{F}^{-1}(u)} \mathbb{1}_{\overline{F}(\overline{F}^{-1}(u))=u} \mathbb{1}_{\overline{F}^{-1}(u) \in \mathcal{D}_F} \\ &= \mathbb{1}_{X=\overline{F}^{-1}(u)} \mathbb{1}_{\overline{F}(\overline{F}^{-1}(u))=u} \text{ p.s.}\end{aligned}$$

*Démonstration.*

Soit  $u \in [0, 1]$ .

On rappelle que l'on note  $\mathcal{F}_u = \{t, F(t) \geq u\}$  et  $\mathcal{F}_u^- = \{t, \overline{F}(t) \leq u\}$ .

Si  $\overline{F}(X) = u$  alors  $X \in \mathcal{F}_u^- \cap \mathcal{F}_u$  donc  $F^{-1}(u) \leq X \leq \overline{F}^{-1}(u)$ . Or,

$$\begin{aligned}\mathbb{P}(F^{-1}(u) < X < \overline{F}^{-1}(u)) &= \overline{F}(\overline{F}^{-1}(u)) - F(F^{-1}(u)) \\ &= 0 \text{ d'après 1.1.4.6.}\end{aligned}$$

Donc  $\mathbb{1}_{\overline{F}(X)=u} = \mathbb{1}_{\overline{F}(X)=u} \mathbb{1}_{X \in \{F^{-1}(u), \overline{F}^{-1}(u)\}} \text{ p.s.}$

On note  $\mathcal{C}_F := \mathbb{R} \setminus \mathcal{D}_F$  l'ensemble des points de continuité de  $F$ .

On a :

$$\mathbb{1}_{\overline{F}(X)=u} \mathbb{1}_{X \in \{F^{-1}(u), \overline{F}^{-1}(u)\}} = \mathbb{1}_{\overline{F}(X)=u} \mathbb{1}_{X \in \{F^{-1}(u), \overline{F}^{-1}(u)\} \cap \mathcal{C}_F} + \mathbb{1}_{\overline{F}(X)=u} \mathbb{1}_{X \in \{F^{-1}(u), \overline{F}^{-1}(u)\} \cap \mathcal{D}_F} \text{ p.s.}$$

On remarque alors que :

$$\mathbb{P}(X = F^{-1}(u)) = \mathbb{P}(F^{-1}(u) \leq X \leq F^{-1}(u)) = F(F^{-1}(u)) - \overline{F}(F^{-1}(u));$$

$$\mathbb{P}(X = \overline{F}^{-1}(u)) = F(\overline{F}^{-1}(u)) - \overline{F}(\overline{F}^{-1}(u)). \quad (\#)$$

Donc

$$\begin{aligned}\mathbb{P}(X \in \{F^{-1}(u), \overline{F}^{-1}(u)\} \cap \mathcal{C}_F) &\leq \mathbb{P}(X = F^{-1}(u)) \mathbb{1}_{F^{-1}(u) \in \mathcal{C}_F} + \mathbb{P}(X = \overline{F}^{-1}(u)) \mathbb{1}_{\overline{F}^{-1}(u) \in \mathcal{C}_F} \\ &= (F(F^{-1}(u)) - \overline{F}(F^{-1}(u))) \mathbb{1}_{F^{-1}(u) \in \mathcal{C}_F} + (F(\overline{F}^{-1}(u)) - \overline{F}(\overline{F}^{-1}(u))) \mathbb{1}_{\overline{F}^{-1}(u) \in \mathcal{C}_F} \\ &= 0.\end{aligned}$$

Donc  $\mathbb{1}_{\overline{F}(X)=u} \mathbb{1}_{X \in \{F^{-1}(u), \overline{F}^{-1}(u)\} \cap \mathcal{C}_F} = 0 \text{ p.s.}$

De plus, si  $\overline{F}(x) = u$  et  $x \in \mathcal{D}_F$  alors  $x \geq \overline{F}^{-1}(u)$ .

En effet, supposons que  $x < \overline{F}^{-1}(u)$ . Comme  $\overline{F}^{-1}(u)$  est le plus petit des majorants de  $\mathcal{F}_u^-$ ,  $x$  n'est pas un majorant de  $\mathcal{F}_u^-$  alors il existe  $t \in \mathcal{F}_u^-$  tel que  $t > x$ . Or  $t \in \mathcal{F}_u^-$ , donc  $\overline{F}(t) \leq u$ . Comme  $x \in \mathcal{D}_F$ , on a  $\overline{F}(x) < F(x) = \mathbb{P}(X \leq x) \leq \mathbb{P}(X < t) = \overline{F}(t) \leq u$ .

Absurde car on a supposé que  $\overline{F}(x) = u$ .

Donc d'après 1.1.4.5,

$$\begin{aligned}\mathbb{1}_{\overline{F}(X)=u} \mathbb{1}_{X \in \{F^{-1}(u), \overline{F}^{-1}(u)\} \cap \mathcal{D}_F} &= \mathbb{1}_{\overline{F}(X)=u} \mathbb{1}_{X=\overline{F}^{-1}(u)} \mathbb{1}_{X \in \mathcal{D}_F} \\ &= \mathbb{1}_{X=\overline{F}^{-1}(u)} \mathbb{1}_{\overline{F}(\overline{F}^{-1}(u))=u} \mathbb{1}_{\overline{F}^{-1}(u) \in \mathcal{D}_F}.\end{aligned}$$

Enfin, d'après (#),

$$\mathbb{1}_{\overline{F}(X)=u} \mathbb{1}_{X \in \{F^{-1}(u), \overline{F}^{-1}(u)\} \cap \mathcal{D}_F} = \mathbb{1}_{X=\overline{F}^{-1}(u)} \mathbb{1}_{\overline{F}(\overline{F}^{-1}(u))=u} \text{ p.s.}$$

□

## 1.2 Notion de $p$ -valeur pour un test simple.

On considère dans cette section, une procédure de test simple.

### 1.2.1 Erreur de première espèce, puissance d'un test et valeur critique.

Un test statistique permet de construire une stratégie de décision : rejeter ou non une hypothèse.

#### Définition 1.2.1 : Test statistique

Soit  $(\mathbb{X}, \mathcal{X}, \mathcal{P})$  un modèle statistique et soit  $\theta_0$  un sous ensemble de  $\mathcal{P}$ . On considère le problème de test  $H_0 : "P \in \theta_0"$  contre  $H_1 : "P \in \theta_0^c := \theta_1"$ . Une zone de rejet (ou région critique)  $R$  est un sous-ensemble de l'espace des observations, la stratégie de décision est alors :

- si  $x \notin R$ ,  $H_0$  n'est pas rejetée ;
- si  $x \in R$ ,  $H_0$  est rejetée.

Le test statistique de  $H_0$  contre  $H_1$  de zone de rejet  $R$  est la stratégie de décision ci-dessus.

Plus formellement, un test statistique est caractérisé par la fonction indicatrice de sa zone de rejet.

Il convient à présent de relier la zone de rejet d'un test à l'hypothèse nulle.

Rejeter ou non l'hypothèse nulle au profit d'une hypothèse alternative peut conduire à deux erreurs : l'erreur de première espèce consiste à rejeter  $H_0$  à tort et l'erreur de seconde espèce consiste à accepter  $H_0$  à tort.

<i>Décision \ Vérité</i>	<i>Hypothèse nulle vraie</i>	<i>Hypothèse nulle fausse</i>
<i>Hypothèse nulle non rejetée</i>	Décision juste	Erreur de deuxième espèce
<i>Hypothèse nulle rejetée</i>	Erreur de première espèce	Décision juste

#### Définition 1.2.2 : Risque de première espèce et taille d'un test

Soit  $(\mathbb{X}, \mathcal{X}, \mathcal{P})$  un modèle statistique et  $R$  une zone de rejet pour  $H_0$  vs  $H_1$  définie comme ci-dessus.

1. Le risque de 1ère espèce est la fonction définie sur  $\Theta_0$  par  $P \mapsto P(R)$  ;
2. La taille du test est la quantité

$$\sup_{P \in \Theta_0} P(R).$$

3. Le test est de niveau  $\alpha$  si le risque de 1ère espèce maximal est inférieure à  $\alpha$ , ie :

$$\sup_{P \in \Theta_0} P(R) \leq \alpha.$$

Il est toujours possible de construire une zone de rejet de manière à ce que  $H_0$  soit toujours acceptée (en prenant une zone de rejet vide). Dans ce cas, le test est de niveau nul. Cependant, on accepte toujours  $H_0$ , même si cette hypothèse n'est pas vraie. Il convient donc de définir un autre type d'erreur.

#### Définition 1.2.3 : Risque de deuxième espèce et puissance d'un test

Soit  $(\mathbb{X}, \mathcal{X}, \mathcal{P})$  un modèle statistique et  $R$  une zone de rejet pour  $H_0$  vs  $H_1$  définie comme ci-dessus.

1. Le risque de 2ième espèce est la fonction définie sur  $\Theta_1$  par  $P \mapsto P(R^c)$  ;
2. La puissance du test est la fonction définie sur  $\Theta_1$  par  $P \mapsto P(R)$ .

Puisque  $H_0$  est l'hypothèse privilégiée, rejeter  $H_0$  à tort est plus préjudiciable qu'accepter  $H_0$  à tort. On adopte alors le **principe de Neyman-Pearson** :

- se fixer un niveau  $\alpha$  ;

► rechercher ensuite, parmi les tests de niveau  $\alpha$ , ceux dont la puissance est maximale.

On cherche alors un compromis lors de la construction du test en choisissant convenablement le niveau (plus il est faible, plus la zone de rejet à tendance à être petite ce qui diminue la puissance du test) et les hypothèses (plus  $\Theta_1$  est grand, plus la puissance diminue).

### 1.2.2 Notion de $p$ -valeur.

On introduit dans cette sous-section une procédure de décision basée sur la notion de  $p$ -valeur.

A partir d'une observation  $X$  de loi inconnue  $P$ , on souhaite tester l'hypothèse nulle  $H_0 : "P \in \Theta_0"$ .

De manière heuristique, une  $p$ -valeur traduit la probabilité, sous  $H_0$ , d'observer pire que les observations. Si la  $p$ -valeur est faible, ce que l'on observe est rare sous  $H_0$ , ce qui remet en cause l'hypothèse  $H_0$ . En particulier, pour un seuil  $\alpha$  fixé,  $H_0$  est rejetée au niveau inférieur à  $\alpha$  dès que la  $p$ -valeur est plus petite que  $\alpha$ . En revanche, si la  $p$ -valeur est assez grande, cela signifie qu'il est raisonnable d'observer cette valeur sous  $H_0$ . De manière plus formelle, on a :

#### Définition 1.2.4 : $p$ -valeur

Une statistique  $p(X)$  est appelée  $p$ -valeur associée à  $X$  si cette variable aléatoire est minorée en probabilité par une variable aléatoire uniforme sur  $[0, 1]$  sous l'hypothèse nulle ie :

$$\forall P \in \Theta_0, \forall u \in [0, 1], \mathbb{P}_P(p(X) \leq u) \leq u.$$

On peut définir un test statistique de niveau  $\alpha$  à partir d'une  $p$ -valeur  $p(X)$  de la façon suivante :

$$\phi(x) = \mathbb{1}_{(p(x) \leq \alpha)}.$$

Cette définition de  $p$ -valeur est abstraite et n'est pas explicite pour réaliser un test en pratique. Cependant, on va voir que l'on peut définir une  $p$ -valeur à partir d'une statistique de test  $S(X)$ . On suppose que  $H_0$  est rejetée pour "des grandes valeurs" de  $S(X)$ . On se concentrera donc sur des zones de rejet de la forme :

$$R = \{X \in \mathbb{X} : S(X) > v_c\};$$

où  $v_c$  est la valeur critique associée au test. La valeur critique est une constante à déterminer en fonction de la loi de  $S(X)$  sous ( $H_0$ ) (son quantile par exemple) et du niveau  $\alpha \in (0, 1)$  fixé.

La procédure de décision repose sur la construction d'une zone de rejet.

On note :

- $T_P(s) = \mathbb{P}_P(S(X) \geq s)$  ;
- $F_P(s) = \mathbb{P}_P(S(X) \leq s)$  ;
- $\bar{F}_P(s) = \mathbb{P}_P(S(X) < s)$  ;
- $F_P^{-1}(u) = \inf\{s \in \mathbb{R} : F_P(s) \geq u\}$ .

#### Lemme 1.2.5 : d'après [14]

Pour tout  $P \in \mathcal{P}$ , pour tout  $\alpha \in [0, 1]$ , on a

$$\{T_P(S(X)) < \alpha\} = \{S(X) > \bar{F}_P^{-1}(1 - \alpha)\};$$

$$\{T_P(S(X)) \leq \alpha\} = \begin{cases} \{S(X) \geq F_P^{-1}(1 - \alpha)\} & \text{si } \bar{F}_P(F_P^{-1}(1 - \alpha)) = 1 - \alpha \\ \{S(X) > F_P^{-1}(1 - \alpha)\} & \text{si } \bar{F}_P(F_P^{-1}(1 - \alpha)) < 1 - \alpha. \end{cases}$$

*Démonstration.*

Soit  $P \in \mathcal{P}$ , soit  $\alpha \in [0, 1]$ .

La première égalité du lemme est immédiate car d'après le point 3 de 1.1.4,

$$\{S(X) > \bar{F}_P^{-1}(1 - \alpha)\} = \{\bar{F}_P(S(X)) > 1 - \alpha\} = \{T_P(S(X)) < \alpha\}.$$

Pour prouver la deuxième partie du lemme, on remarque d'abord que

$$\{T_P(S(X)) \leq \alpha\} = \{\bar{F}_P(S(X)) \geq 1 - \alpha\} = \{S(X) > \bar{F}_P^{-1}(1 - \alpha)\} \cup \{\bar{F}_P(S(X)) = 1 - \alpha\} \quad \text{d'après 1.1.4.3}$$

Or, on a que  $\{\bar{F}_P(S(X)) = 1 - \alpha\} \subset \{F_P(S(X)) \geq 1 - \alpha\} = \{F_P^{-1}(1 - \alpha) \leq S(X)\}$  d'après 1.1.4.3. De plus, le point 1.1.4.5 donne directement que  $\{S(X) > \bar{F}_P^{-1}(1 - \alpha)\} \subset \{S(X) \geq F_P^{-1}(1 - \alpha)\}$ . Ainsi,

$$\{T_P(S(X)) \leq \alpha\} \subset \{S(X) \geq F_P^{-1}(1 - \alpha)\}.$$

Supposons que  $\bar{F}_P(F_P^{-1}(1 - \alpha)) = 1 - \alpha$ .

On a immédiatement que

$$\{S(X) \geq F_P^{-1}(1 - \alpha)\} \subset \{\bar{F}_P(S(X)) \geq \bar{F}_P(F_P^{-1}(1 - \alpha))\} = \{\bar{F}_P(S(X)) \geq 1 - \alpha\} = \{T_P(S(X)) \leq \alpha\}.$$

Supposons maintenant que  $\bar{F}_P(F_P^{-1}(1 - \alpha)) < 1 - \alpha$ .

Alors on a d'une part que

$$\{\bar{F}_P(S(X)) \geq 1 - \alpha\} \subset \{\bar{F}_P(S(X)) > \bar{F}_P(F_P^{-1}(1 - \alpha))\} = \{S(X) > F_P^{-1}(1 - \alpha)\};$$

(par hypothèse et par croissance de  $\bar{F}_P$ ) et d'autre part que

$$\{S(X) > F_P^{-1}(1 - \alpha)\} \subset \{\bar{F}_P(S(X)) \geq F_P(F_P^{-1}(1 - \alpha))\} \subset \{\bar{F}_P(S(X)) \geq 1 - \alpha\} \quad \text{d'après 1.1.4.2}$$

La première inclusion vient du fait que, pour tout  $u, t \in \mathbb{R}$ ,  $\{u > t\} \subset \{\bar{F}_P(u) \geq F_P(t)\}$  (en revenant aux définitions de  $F_P$  et  $\bar{F}_P$ ).

Cela achève la preuve du lemme. □

**Proposition 1.2.6 :** *Exemple important de  $p$ -valeur*

En supposant que le supremum conserve la mesurabilité, la  $p$ -valeur définie par une statistique de test  $S(X)$  est donnée par  $p_S(X) := \sup_{P \in \Theta_0} T_P(S(X))$ .

*Démonstration.*

Soit  $P \in \Theta_0$ ,  $u \in [0, 1]$ . On a  $\mathbb{P}_P(\sup_{P \in \Theta_0} T_P(S(X)) \leq u) \leq \mathbb{P}_P(T_P(S(X)) \leq u)$ . On applique le lemme 1.2.5 : si  $\bar{F}_P(F_P^{-1}(1 - u)) = 1 - u$ ,

$$\mathbb{P}_P(p_S(X) \leq u) \leq \mathbb{P}_P(S(X) \geq F_P^{-1}(1 - u)) = 1 - \bar{F}_P(F_P^{-1}(1 - u)) = u;$$

et si  $\bar{F}_P(F_P^{-1}(1 - u)) < 1 - u$ , alors

$$\mathbb{P}_P(p_S(X) \leq u) \leq \mathbb{P}_P(S(X) > F_P^{-1}(1 - u)) = 1 - F_P(F_P^{-1}(1 - u)) \leq u \quad \text{d'après 1.1.4.2.}$$

□

On dispose alors d'un corollaire qui permet de donner une interprétation d'une  $p$ -valeur.

**Corollaire 1.2.7 :** *Roquain (2011) [14]*

La  $p$ -valeur  $p_S(X)$  vérifie les propriétés suivantes :

(i) Si pour tout  $P \in \Theta_0$ ,  $F_P$  est continue alors on a, pour toute réalisation  $x$  de  $X$ ,

$$p_S(x) = \inf \left( \alpha \in [0, 1] : S(x) \geq \sup_{P \in \Theta_0} F_P^{-1}(1 - \alpha) \right).$$

Si on suppose de plus que  $\Theta_0$  est un singleton, on a, pour  $P \in \Theta_0$ ,  $p_S(X) \sim \mathcal{U}(0, 1)$ .

(ii) Si pour tout  $P \in \Theta_0$ , la variable  $S(X)$  est discrète alors on a, pour toute réalisation  $x$  de  $X$ ,

$$p_S(x) = \inf \left( \alpha \in [0, 1] : S(x) > \sup_{P \in \Theta_0} F_P^{-1}(1 - \alpha) \right).$$

En particulier, si  $S(X)$  est une variable aléatoire discrète à valeurs dans  $\mathbb{N}$ , on a,

$$p_S(x) = \inf \left( \alpha \in [0, 1] : S(x) \geq \sup_{P \in \Theta_0} F_P^{-1}(1 - \alpha) + 1 \right).$$

*Démonstration.*

Montrons (i).

Supposons que pour tout  $P \in \Theta_0$ ,  $F_P$  est continue. Dans ce cas,  $F_P = \bar{F}_P$  et d'après 1.1.4.2 et 1.1.4.4, pour tout  $\alpha \in [0, 1]$ ,  $\bar{F}_P(F_P^{-1}(1 - \alpha)) = F_P(F_P^{-1}(1 - \alpha)) = 1 - \alpha$ . Le résultat 1.2.5 fournit  $\{T_P(S(X)) \leq \alpha\} = \{S(X) \geq F_P^{-1}(1 - \alpha)\}$ . On obtient donc, pour tout  $x \in X$ ,

$$\begin{aligned} p_S(x) &= \inf \{ \alpha \in [0, 1] : \forall P \in \Theta_0, T_P(S(x)) \leq \alpha \} \\ &= \inf \{ \alpha \in [0, 1] : \forall P \in \Theta_0, S(x) \geq F_P^{-1}(1 - \alpha) \}. \end{aligned}$$

Montrons (ii).

Supposons que la loi de  $S(X)$  soit discrète. On a alors, pour tout  $\alpha \in [0, 1]$ ,  $\bar{F}_P(F_P^{-1}(1 - \alpha)) < 1 - \alpha$ . La preuve est alors identique au point précédent.  $\square$

A partir d'une statistique de test, une  $p$ -valeur calcule donc l'erreur de première espèce réelle du test.

On en déduit une caractérisation d'une  $p$ -valeur : c'est la plus petite des valeurs de risque de première espèce pour lesquelles la décision serait de rejeter  $H_0$ .

### Remarque 1.2.8 : Notion de $p$ -valeur au sens de Lehmann

Cette proposition nous permet de remarquer qu'une  $p$ -valeur peut être vue comme le niveau minimum à partir duquel on change de décision dans un test statistique.

C'est de cette manière que Lehmann définit une  $p$ -valeur dans son ouvrage *Testing Statistical Hypotheses*, Springer.

### Exemple 1.2.9 : Cas d'une loi continue.

On observe une variable aléatoire de loi gaussienne d'espérance inconnue  $\mu$  et de variance égale à 1.

On souhaite tester l'hypothèse  $\mathcal{H}_0 : \mu = 0$  contre l'alternative  $\mathcal{H}_1 : \mu \neq 0$ .

On estime  $\mu$  par la variable  $X$  : l'hypothèse nulle étant rejetée si  $|X|$  est trop grande.

Sous  $\mathcal{H}_0$ ,  $X \sim \mathcal{N}(0, 1)$  et on note  $F$  la fonction de répartition de  $X$  sous  $\mathcal{H}_0$ . On remarque que

$$\forall x \in \mathbb{R}, \quad \mathbb{P}_{\mathcal{H}_0}(|X| \leq x) = F(x) - F(-x) = 2F(x) - 1.$$

En prenant  $x = F^{-1}(1 - \alpha/2)$ , le quantile d'ordre  $1 - \alpha/2$  d'une loi normale centrée et réduite, on obtient

$$\mathbb{P}_{\mathcal{H}_0}(|X| \geq F^{-1}(1 - \alpha/2)) = \alpha.$$

Alors :

$$\begin{aligned} p_S(x) &= \inf\{\alpha \in [0, 1] : |x| \geq F^{-1}(1 - \frac{\alpha}{2})\} \\ &= \inf\{\alpha \in [0, 1] : 1 - F(|x|) \leq \frac{\alpha}{2}\} \\ &= 2(1 - F(|x|)). \end{aligned}$$

On en déduit alors que

$$\begin{aligned} p_S(X) &= 2(1 - F(|X|)) \\ &= 2F(-|X|). \end{aligned}$$

On sait aussi que la loi de  $p_S(X)$  est uniforme sur  $[0, 1]$ .

La zone de rejet au niveau  $\alpha \in (0, 1)$  sera alors

$$R = \{p_S(X) \leq \alpha\}.$$

### Exemple 1.2.10 : Cas d'une loi discrète.

On réalise 400 expériences indépendantes de Bernoulli de paramètre  $p$ . On souhaite tester  $\mathcal{H}_0 : "p = 0,005"$  contre l'alternative  $\mathcal{H}_1 : "p > 0,005"$ . On observe  $X$  (le nombre d'événements de probabilité  $p$ ) de sorte que l'hypothèse nulle soit rejetée si  $X$  est trop grand.

Sous  $\mathcal{H}_0$ ,  $X \sim \mathcal{B}(400, 0.005)$  et on peut approcher la loi binomiale par une loi de Poisson :  $X \sim \mathcal{P}(2)$ . Soit  $F$  la fonction de répartition d'une loi de Poisson de paramètre 2. Alors

$$p_S(x) = \inf\{\alpha \in [0, 1] : x \geq F^{-1}(1 - \alpha)\}.$$

On en déduit alors que  $x = F^{-1}(1 - p_S(x))$  ie

$$1 - p_S(x) = F(x^-).$$

Comme les lois considérées sont à valeurs dans  $\mathbb{N}$ , on trouve

$$p_S(X) = 1 - F(X - 1).$$

La zone de rejet au niveau  $\alpha \in (0, 1)$  sera alors

$$R = \{p_S(X) \leq \alpha\}.$$

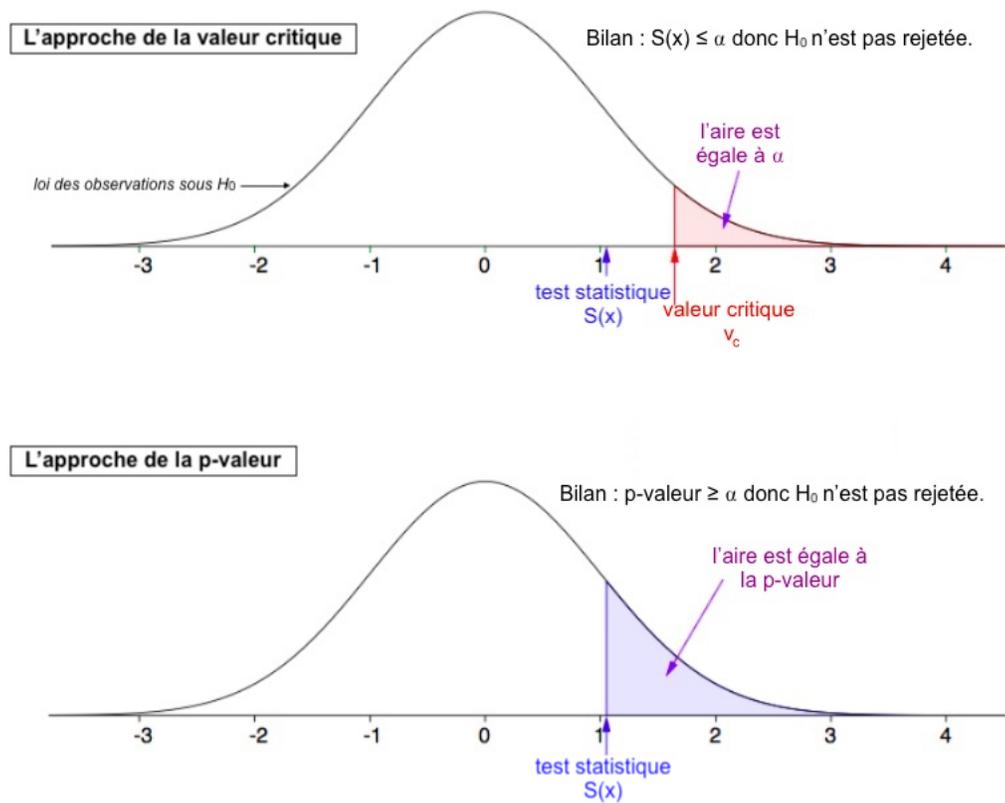
### 1.2.3 Valeur critique versus $p$ -valeur.

La  $p$ -valeur et la valeur critique associées à un test sont très liées : elles donnent le niveau de décision du test. Cependant, la valeur critique donne uniquement une valeur de comparaison pour la valeur pratique issue des observations tandis que la  $p$ -valeur donne la valeur exacte de la significativité du test. L'interprétation de ces deux valeurs est donc différente :

- la comparaison avec la valeur critique donne uniquement le fait que le test soit valide ou non ;
- la connaissance de la  $p$ -valeur donne aussi le niveau de fiabilité que l'on peut accorder à la décision.

La démarche pour le calcul de la  $p$ -valeur et de la valeur critique diffère également :

- Pour la  $p$ -valeur : à partir de la valeur pratique de notre statistique de test et de la loi que suit la statistique de test, on cherche la valeur du risque en fonction de notre valeur pratique ;
- Pour la valeur critique : à partir de la loi de la statistique de test et du niveau de risque choisi, on cherche (dans les tables) la valeur critique en fonction du niveau fixé.

FIGURE 1.1 –  $p$ -valeur ou probabilité critique.

Il reste maintenant à voir quelles sont les liens entre ces deux valeurs et quelles conséquences a le choix de la  $p$ -valeur ou de la valeur critique sur la performance du test.

Afin de définir un test, on a vu qu'il existe deux manières de procéder à partir d'une statistique de test : en déterminant une valeur critique ou en considérant une  $p$ -valeur. Le résultat suivant permet de faire le lien entre ces 2 approches, les tests multiples utilisant les  $p$ -valeurs et les tests agrégés utilisant les valeurs critiques de la statistique de test.

**Proposition 1.2.11 :** *d'après Fromont, Lerasle, Reynaud-Bouret (2015) [6]*

On suppose que la loi de  $S(X)$  ne dépend pas de  $P$  sous  $H_0$  (par exemple si  $\Theta_0$  est un singleton). On note  $F_0$  la fonction de répartition de  $S(X)$ . On a donc que  $p_S(X) = 1 - \bar{F}_0(S(X))$ . Pour tout niveau  $\alpha \in ]0, 1[$  fixé, on définit les tests

$$\phi_S = \mathbb{1}_{\{S(X) > F_0^{-1}(1-\alpha)\}}; \quad \bar{\phi}_S = \mathbb{1}_{\{S(X) > \bar{F}_0^{-1}(1-\alpha)\}}; \quad \phi_p = \mathbb{1}_{\{p_S(X) \leq \alpha\}} \quad \text{et} \quad \bar{\phi}_p = \mathbb{1}_{\{p_S(X) < \alpha\}}.$$

Alors

1. Ces tests sont de niveau  $\alpha$ ,
2. On a, presque sûrement,

$$\bar{\phi}_p = \bar{\phi}_S \leq \phi_p;$$

$$\bar{\phi}_p = \bar{\phi}_S \leq \phi_S;$$

$$\bar{\phi}_p = \bar{\phi}_S = \phi_p = \phi_S \quad \text{si} \quad \bar{F}_0(\bar{F}_0^{-1}(1-\alpha)) < 1-\alpha.$$

Plus précisément, en notant  $\mathcal{D}_{F_0}$  l'ensemble des points de discontinuité de  $F_0$ , on a,

$$\begin{aligned}\phi_p &= \bar{\phi}_S + \mathbb{1}_{S(X)=\bar{F}_0^{-1}(1-\alpha)} \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha))=1-\alpha} \mathbb{1}_{\bar{F}_0^{-1}(1-\alpha) \in \mathcal{D}_{F_0}} \\ &= \phi_S \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha)) < 1-\alpha} + \mathbb{1}_{S(X) \geq \bar{F}_0^{-1}(1-\alpha)} \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha))=1-\alpha}.\end{aligned}$$

3. La  $p$ -valeur au sens de Lehmann de ces quatre tests est égale à  $p_S(X)$ , c'est-à-dire,

$$\begin{aligned}p_S(X) &= \inf\{\alpha \in [0, 1], S(X) > F_0^{-1}(1-\alpha)\} \\ &= \inf\{\alpha \in [0, 1], S(X) > \bar{F}_0^{-1}(1-\alpha)\} \\ &= \inf\{\alpha \in [0, 1], p_S(X) \leq \alpha\} \\ &= \inf\{\alpha \in [0, 1], p_S(X) < \alpha\}.\end{aligned}$$

*Démonstration.*

1. Soit  $P \in \mathcal{P}$ . On a d'après les points 2 et 4 de 1.1.4, sous  $H_0$ ,

- $\mathbb{P}_p(\phi_S = 1) = \mathbb{P}_p(S(X) > F_0^{-1}(1-\alpha)) = 1 - F_0(F_0^{-1}(1-\alpha)) \leq \alpha$ ;
- $\mathbb{P}_p(\bar{\phi}_S = 1) = \mathbb{P}_p(S(X) > \bar{F}_0^{-1}(1-\alpha)) = 1 - F_0(\bar{F}_0^{-1}(1-\alpha)) \leq \alpha$ .
- Par définition d'une  $p$ -valeur, on a sous  $H_0$ ,  $\mathbb{P}_p(\phi_p = 1) = \mathbb{P}_p(p_S(X) \leq \alpha) \leq \alpha$ ;
- $\mathbb{P}_p(\bar{\phi}_p = 1) = \mathbb{P}_p(p_S(X) < \alpha) \leq \mathbb{P}_p(p_S(X) \leq \alpha) \leq \alpha$ .

Donc les tests  $\phi_S, \bar{\phi}_S, \phi_p$  et  $\bar{\phi}_p$  sont de niveau  $\alpha$ .

2. D'après 1.2.5,

$$\{S(X) > \bar{F}_0^{-1}(1-\alpha)\} = \{p_S(X) < \alpha\}.$$

On obtient donc que  $\bar{\phi}_p = \bar{\phi}_S$ .

De plus, il est clair que  $\bar{\phi}_p \leq \phi_p$  et que  $\bar{\phi}_S \leq \phi_S$ .

Montrons à présent que

$$\begin{aligned}\phi_p &= \bar{\phi}_S + \mathbb{1}_{S(X)=\bar{F}_0^{-1}(1-\alpha)} \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha))=1-\alpha} \mathbb{1}_{\bar{F}_0^{-1}(1-\alpha) \in \mathcal{D}_{F_0}} \\ &= \phi_S \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha)) < 1-\alpha} + \mathbb{1}_{S(X) \geq \bar{F}_0^{-1}(1-\alpha)} \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha))=1-\alpha}.\end{aligned}$$

L'égalité  $\bar{\phi}_p = \bar{\phi}_S = \phi_p = \phi_S$  si  $\bar{F}_0(\bar{F}_0^{-1}(1-\alpha)) < 1-\alpha$  découlera alors immédiatement.

On a,

$$\begin{aligned}\phi_p &= \mathbb{1}_{\bar{F}_0(S(X)) \geq 1-\alpha} \text{ par définition de } p_S(X) \\ &= \mathbb{1}_{\bar{F}_0(S(X)) > 1-\alpha} + \mathbb{1}_{\bar{F}_0(S(X)) = 1-\alpha} \\ &= \mathbb{1}_{S(X) > \bar{F}_0^{-1}(1-\alpha)} + \mathbb{1}_{\bar{F}_0(S(X)) = 1-\alpha} \text{ d'après 1.1.4.3.}\end{aligned}$$

D'après le corollaire 1.1.6,

$$\phi_p = \mathbb{1}_{S(X) > \bar{F}_0^{-1}(1-\alpha)} + \mathbb{1}_{S(X)=\bar{F}_0^{-1}(1-\alpha)} \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha))=1-\alpha} \mathbb{1}_{\bar{F}_0^{-1}(1-\alpha) \in \mathcal{D}_{F_0}}.$$

On a aussi,

$$\begin{aligned}\phi_p &= \mathbb{1}_{S(X) > \bar{F}_0^{-1}(1-\alpha)} + \mathbb{1}_{S(X)=\bar{F}_0^{-1}(1-\alpha)} \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha))=1-\alpha} \text{ p.s.} \\ &= \mathbb{1}_{S(X) > \bar{F}_0^{-1}(1-\alpha)} \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha)) < 1-\alpha} + \mathbb{1}_{S(X) > \bar{F}_0^{-1}(1-\alpha)} \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha))=1-\alpha} \\ &+ \mathbb{1}_{S(X)=\bar{F}_0^{-1}(1-\alpha)} \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha))=1-\alpha} \text{ p.s.} \\ &= \mathbb{1}_{S(X) > \bar{F}_0^{-1}(1-\alpha)} \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha)) < 1-\alpha} + \mathbb{1}_{S(X) \geq \bar{F}_0^{-1}(1-\alpha)} \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha))=1-\alpha} \text{ p.s.} \\ &= \mathbb{1}_{S(X) > \bar{F}_0^{-1}(1-\alpha)} \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha)) < 1-\alpha} + \mathbb{1}_{S(X) \geq \bar{F}_0^{-1}(1-\alpha)} \mathbb{1}_{\bar{F}_0(\bar{F}_0^{-1}(1-\alpha))=1-\alpha} \text{ p.s. d'après 5. et 6. de 1.1.4.}\end{aligned}$$

3. Tout d'abord, il est clair que

$$\begin{aligned}p_S(X) &= \inf\{\alpha \in [0, 1], p_S(X) \leq \alpha\} \\ &= \inf\{\alpha \in [0, 1], p_S(X) < \alpha\}.\end{aligned}$$

De plus, on a vu au point 2 que  $\{S(X) > \overline{F}_0^{-1}(1 - \alpha)\} = \{p_S(X) < \alpha\}$  donc

$$p_S(X) = \inf\{\alpha \in [0, 1], S(X) > \overline{F}_0^{-1}(1 - \alpha)\}.$$

Il reste à montrer que  $p_S(X) = \inf\{\alpha \in [0, 1], S(X) > F_0^{-1}(1 - \alpha)\}$ .

Pour cela, montrons que  $\inf\{\alpha \in [0, 1], S(X) > F_0^{-1}(1 - \alpha)\} = \inf\{\alpha \in [0, 1], S(X) > \overline{F}_0^{-1}(1 - \alpha)\}$ .

On va noter

$$A = \{\alpha \in [0, 1], S(X) > F_0^{-1}(1 - \alpha)\},$$

$$B = \{\alpha \in [0, 1], S(X) > \overline{F}_0^{-1}(1 - \alpha)\}.$$

D'après 1.1.4.5,  $B \subseteq A$  donc  $\inf_\alpha A \leq \inf_\alpha B$ .

Par l'absurde, supposons que  $\inf_\alpha A < \inf_\alpha B$ .

Il existe alors  $\eta \in (\inf_\alpha A, \inf_\alpha B)$ .

Alors d'une part,  $\eta \in A$  donc  $S(X) > F_0^{-1}(1 - \eta)$  (\*) et d'autre part,  $\eta \notin B$  donc  $S(X) \leq \overline{F}_0^{-1}(1 - \eta)$  (\*\*).

On a donc  $F_0^{-1}(1 - \eta) < S(X) \leq \overline{F}_0^{-1}(1 - \eta)$ .

Comme  $F_0^{-1}(1 - \eta) < \overline{F}_0^{-1}(1 - \eta)$ , le lemme 1.1.4.6 assure que  $\overline{F}_0(\overline{F}_0^{-1}(1 - \eta)) = F_0(F_0^{-1}(1 - \eta)) = 1 - \eta$ .

Alors d'après (\*),  $F_0(S(X)) \geq 1 - \eta$  et donc  $\eta \geq 1 - F_0(S(X))$ .

► Supposons que  $S(X)$  ne soit pas un atome de  $P_0$ .

Alors  $F_0(S(X)) = \overline{F}_0(S(X))$  et  $\eta \geq 1 - \overline{F}_0(S(X)) = p_S(X) = \inf_\alpha B$  ce qui est absurde.

► Supposons que  $S(X)$  soit un atome de la loi  $P_0$ .

Alors  $F_0^{-1}$  est une fonction constante à  $S(X)$  sur  $(\overline{F}_0(S(X)), F_0(S(X))]$ . De plus, d'après (\*\*), on a  $\overline{F}_0(S(X)) \leq 1 - \eta$  et donc  $\overline{F}_0(S(X)) \leq 1 - \eta \leq F_0(S(X))$ .

Si  $\overline{F}_0(S(X)) = 1 - \eta$  alors  $\eta = p_S(X)$  ce qui est absurde.

Sinon,  $1 - \eta \in (\overline{F}_0(S(X)), F_0(S(X))]$  donc  $F_0^{-1}(1 - \eta) = S(X)$ . Cela implique que  $\eta \notin A$  ce qui est absurde.

Finalement,  $\inf_\alpha A = \inf_\alpha B$  ce qui achève la preuve.

□

### Remarque 1.2.12

Lorsque la fonction de répartition de la statistique de test est continue, les quatre tests  $\phi_S, \overline{\phi}_S, \phi_p$  et  $\overline{\phi}_p$  sont presque sûrement égaux.

## Chapitre 2

# Tests multiples : procédure de tests et erreurs de première espèce.

Ce chapitre est essentiellement basée sur l'article d'Etienne Roquain [14].

### 2.1 Procédure de tests multiples.

#### 2.1.1 Idées générales.

Soit  $X$  une variable aléatoire définie d'un ensemble mesurable  $(\Omega, \mathcal{F})$  à l'espace des observations  $(\mathbb{X}, \mathcal{X})$ . On suppose qu'il existe une famille de mesures de probabilité sur  $(\Omega, \mathcal{F})$  qui induit un sous-ensemble  $\mathcal{P}$  de lois de probabilité sur  $(\mathbb{X}, \mathcal{X})$ . Le triplet  $(\mathbb{X}, \mathcal{X}, \mathcal{P})$  est appelé modèle statistique. La loi d'une observation  $X$  sur  $(\mathbb{X}, \mathcal{X})$  est notée  $P$  et pour tout  $P \in \mathcal{P}$ , il existe une loi sur  $(\Omega, \mathcal{F})$  pour laquelle  $X \sim P$ ; on notera  $\mathbb{P}_P$  la mesure image associée et  $\mathbb{E}_P$  l'espérance correspondante.

On considère une famille  $(\Theta_{0,i})_{1 \leq i \leq m}$  de  $m \geq 2$  sous-ensembles de  $\mathcal{P}$ . A partir d'une observation  $X = (X_1, \dots, X_n)$  où  $X_i \in \mathbb{R}^d$  pour tout  $i \in \{1, \dots, n\}$ , on veut tester les hypothèses nulles  $H_{0,i} : "P \in \Theta_{0,i}"$  contre l'alternative  $H_{1,i} : "P \in \Theta_{0,i}^c"$ , simultanément pour tout  $i \in \{1, \dots, m\}$ .

#### Remarque 2.1.1

En pratique, on dispose du cadre d'étude suivant (en reprenant les notations ci-dessus) :

- observation :  $n$  expériences et  $d$  candidats ;
- problème de grande dimension :  $d \gg n$  ;
- dépendances potentiellement complexes entre les candidats.

Par exemple, les "candidats" peuvent être des relations entre les gènes, des régions du cerveau (chaque pixel d'une imagerie peut être examiné), des motifs dans une séquence d'ADN...

Pour tout  $P \in \mathcal{P}$ , on pose  $\mathcal{H}_0(P) := \{1 \leq i \leq m : P \in \Theta_{0,i}\}$ , l'ensemble des indices  $i$  pour lesquels  $P$  satisfait  $H_{0,i}$ . Son cardinal  $|\mathcal{H}_0(P)|$  est noté  $m_0(P)$  et l'ensemble  $\{1, \dots, m\}$  est noté  $\mathcal{H}$ . On pose également  $\mathcal{H}_1(P) := \mathcal{H} \setminus \mathcal{H}_0(P)$ . Selon la loi de l'observation  $X$ , on a donc deux types d'hypothèses nulles :

- les hypothèses vraies  $\mathcal{H}_0$  ;
- les hypothèses fausses  $\mathcal{H}_1$ .

L'objectif est de trouver quelles hypothèses nulles sont vraies-faussees à partir de  $X$ .

Intuitivement, une *procédure de tests multiples* va consister à rejeter ou à ne pas rejeter  $H_{0,i}$  à partir des données  $X$  pour chaque  $i \in \{1, \dots, m\}$ . Plus précisément, à partir de l'observation  $x = X(\omega)$ , on va construire l'ensemble des hypothèses jugées fausses (ie l'ensemble des *hypothèses rejetées*)  $R(x) \subset \{1, \dots, m\}$  de sorte que

$$i \in R(x) \iff H_{0,i} \text{ est rejetée par la procédure.}$$

**Vocabulaire** : on utilise les dénominations suivantes pour évoquer les erreurs commises dans les tests multiples.

Vérité Décision	Hypothèse nulle vraie	Hypothèse nulle fausse	Total
Hypothèse nulle non rejetée	vrai négatif	faux négatif	$m -  R(X) $
Hypothèse nulle rejetée	faux positif	vrai positif	$ R(X) $
Total	$m_0$	$m - m_0$	$m$

On suppose enfin que chaque hypothèse  $H_{0,i}$  est testée par une statistique de test  $S_i(X)$  qui rejette l'hypothèse nulle pour de "grandes" valeurs de  $S_i(X)$ .

### 2.1.2 Procédures de tests multiples.

Nous allons commencer par donner une définition formelle d'une procédure de tests multiples.

**Définition 2.1.2 :** Procédure de tests multiples.

Une procédure de tests multiples  $R$  sur  $\mathcal{H}$  est une fonction  $R : x \in \mathbb{X} \mapsto R(x) \subset \mathcal{H}$  telle que pour tout  $i \in \mathcal{H}$ , la fonction indicatrice  $\mathbb{1}\{i \in R(x)\}$  est mesurable. Les hypothèses  $i \in R$  sont les *hypothèses nulles rejetées* par la procédure  $R$ .

Formalisons l'idée précédente à l'aide des  $p$ -valeurs.

D'après 1.2.6, chaque statistique de test  $S_i(X)$  peut être vue comme la donnée d'une  $p$ -valeur  $p_i(X)$ . Ainsi, pour chaque  $i \in \mathcal{H}$ , on effectue un test simple de l'hypothèse  $H_{0,i}$  avec la  $p$ -valeur  $p_i(X)$ .

On peut donc à présent supposer que la connaissance de  $X$  est équivalente à la connaissance de la famille de  $p$ -valeurs  $\mathbf{p}(X) = \{p_i(X), 1 \leq i \leq m\}$  que l'on notera plus simplement  $\mathbf{p} = \{p_i, 1 \leq i \leq m\} \in [0, 1]^m$ . On rappelle cependant qu'une loi  $P \in \mathcal{P}$  est celle de l'observation  $X$  et pas celle de  $\mathbf{p}$ .

Autrement dit, on suppose que pour chaque hypothèse nulle  $i \in \mathcal{H}$ , il existe une  $p$ -valeur  $p_i$ , définie comme étant une fonction mesurable  $p_i : \mathbb{X} \rightarrow [0, 1]$ , telle que, si  $i \in \mathcal{H}_0(P)$ ,

$$\forall P \in \mathcal{P}, \forall u \in [0, 1], \mathbb{P}_P(p_i(X) \leq u) \leq u.$$

**Exemple 2.1.3 :** Cas gaussien

On suppose que  $X \sim P = \mathcal{N}(\mu, \Gamma)$  où  $\mu \in \mathbb{R}^m$  est inconnu et  $\Gamma \in M_m(\mathbb{R})$  est une matrice de covariance (connue ou inconnue) qui vérifie  $\Gamma_{jj} = 1 \quad \forall j \in \{1, \dots, m\}$ .

► **Test unilatère**

On suppose ici que  $\Theta_{0,j} = \{P = \mathcal{N}(\mu, \Gamma) : \mu_j \leq 0\}$  c'est-à-dire que  $H_{0,j} : "$  $\mu_j \leq 0$ ".

On considère alors  $S_j(X) = X_j$  et  $p_j(X) = \phi(X_j)$  où  $\phi(x) = \mathbb{P}(Z \geq x)$  pour  $Z \sim \mathcal{N}(0, 1)$ .

► **Test bilatère**

On suppose ici que  $\Theta_{0,j} = \{P = \mathcal{N}(\mu, \Gamma) : \mu_j = 0\}$  c'est-à-dire que  $H_{0,j} : "$  $\mu_j = 0$ ".

On considère alors  $S_j(X) = |X_j|$  et  $p_j(X) = 2\phi(|X_j|)$ .

**Exemple 2.1.4 :** Cas des deux groupes

On observe  $X = (X^{(1)}, \dots, X^{(n)}) = (Y^{(1)}, \dots, Y^{(n_0)}, Z^{(1)}, \dots, Z^{(n_1)}) \in \mathbb{R}^{m \times n}$  où  $X_i \in \mathbb{R}^m$ .

On suppose que

$$Y^{(i)} = \mu_0 + \eta_i, \quad 1 \leq i \leq n_0 \quad \text{et} \quad Z^{(i)} = \mu_1 + \xi_i, \quad 1 \leq i \leq n_1$$

avec  $(\xi_i)_i, (\eta_i)_i$  des suites de variables iid de loi  $Q$ , une loi centrée sur  $\mathbb{R}^m$ .

On veut tester l'hypothèse  $H_{0,j} : \mu_{0,j} = \mu_{1,j}$  contre  $H_{1,j} : \mu_{0,j} \neq \mu_{1,j}$ .

On considère alors la statistique de Student :

$$S_j(X) = \frac{1}{\sqrt{1/n_0 + 1/n_1}} \frac{|\widehat{\mu}_{0,j} - \widehat{\mu}_{1,j}|}{\widehat{\sigma}_j} \quad \text{où} \quad \widehat{\mu}_{0,j} = \frac{1}{n_0} \sum_{i=1}^{n_0} Y_j^{(i)} \quad \text{et} \quad \widehat{\mu}_{1,j} = \frac{1}{n_1} \sum_{i=1}^{n_1} Z_j^{(i)}.$$

► Si  $Q = \mathcal{N}(0, \Gamma)$

Alors  $p_j(X) = 2\phi(S_j(X))$  où  $\phi(x) = \mathbb{P}(Z \geq x)$  avec  $Z \sim \tau(n-2)$ .

► Si  $Q$  est une loi inconnue et arbitraire

On peut obtenir des  $p$ -valeurs en utilisant des permutations :

$$p_j(X) = \frac{1}{B+1} \left( 1 + \sum_{b=1}^B \mathbb{1}_{S_j(X^{\sigma_b}) \geq S_j(X)} \right)$$

où  $\sigma_1, \dots, \sigma_B$  sont iid de loi uniforme sur  $\mathfrak{S}_n$  et  $X^\sigma = (X^{(\sigma(1))}, \dots, X^{(\sigma(n))})$ .

### Remarque 2.1.5

Le calcul des  $p$ -valeurs n'est pas toujours possible car il nécessite de connaître la loi de la statistique de test sous  $\mathcal{H}_0$ . Dans certains cas, on peut utiliser une *technique de randomisation ou de permutation* qui fournit une famille de  $p$ -valeurs<sup>1</sup>.

Une procédure de tests multiples est donc une fonction

$$R : q = (q_i)_{1 \leq i \leq m} \in [0, 1]^m \mapsto R(q) \subset \{1, \dots, m\},$$

telle que pour tout  $i \in \{1, \dots, m\}$ , l'application  $x \in \mathbb{X} \mapsto \mathbb{1}_{i \in R(\mathbf{p}(x))}$  soit mesurable. Les coordonnées sélectionnées par  $R(\mathbf{p})$  correspondent aux hypothèses nulles rejetées.

*Bilan* : A partir de l'observation  $X$ , on calcule une famille de  $p$ -valeurs  $\mathbf{p}$  afin de déterminer la procédure de tests multiples  $R(\mathbf{p}) \subset \{1, \dots, m\}$ .

Voyons maintenant comment choisir une procédure de tests multiples.

D'après 1.2.4, pour tout  $i \in \{1, \dots, m\}$ , rejeter  $H_{0,i}$  lorsque  $p_i(X) \leq \alpha$  définit un test de niveau  $\alpha$ . On a aussi vu que si la  $p$ -valeur  $p_i(X)$  est distribuée selon une loi uniforme sur  $(0, 1)$  sous  $H_{0,i}$  alors on obtient un test de niveau exactement  $\alpha$ . De plus, rejeter une hypothèse nulle pour de "grandes" valeurs de la statistique de test revient à rejeter l'hypothèse nulle pour de "petites"  $p$ -valeurs. Ainsi, une  $p$ -valeur proche de zéro est interprétée comme une preuve issue des observations contre la validité de l'hypothèse nulle.

En conclusion, une région de rejet pour chaque  $H_{0,i}$  est de la forme  $\{1 \leq i \leq m : p_i \leq t_i(x, \mathbf{p})\}$  où la famille  $(t_i)_{1 \leq i \leq m} \in [0, 1]$  définit des *seuils*. On s'intéressera ici au cas où le seuil, noté  $t$ , est le même pour chaque  $p$ -valeur. Ce type de procédure s'appelle procédure par **méthode de seuillage**.

*Bilan* : La région de rejet est de la forme  $R(\mathbf{p}) = \{1 \leq i \leq m : p_i \leq t(\mathbf{p})\}$ , où le seuil  $t(\cdot) \in [0, 1]$  peut dépendre des observations.

1. une référence pour ces techniques de randomisation est : Romano et Wolf, Exact and approximate stepdown methods for multiple hypothesis testing, *J. Amer. Statist. Assoc.*, 2005.

**Exemple 2.1.6 : Procédure de Bonferroni**

La procédure de Bonferroni de niveau  $\alpha \in (0, 1)$  rejette l'hypothèse nulle pour une  $p$ -valeur plus petite que  $\alpha/m$ . La procédure correspondante est alors :

$$R^{\text{Bonf}}(\mathbf{p}) = \{1 \leq i \leq m : p_i \leq \alpha/m\}.$$

**Remarque 2.1.7 : Procédures step-down et step-up**

Dans la suite, nous allons considérer ou construire des procédures itératives particulières :

- les procédures *step-down* sont des procédures de tests multiples où l'ensemble des hypothèses rejetées croît à chaque étape de l'algorithme (on accepte de moins en moins d'hypothèses) ;
- les procédures *step-up* sont des procédures de tests multiples où l'ensemble des hypothèses rejetées décroît à chaque étape de l'algorithme (on accepte de plus en plus d'hypothèses).

*Objectifs* : Choisir le seuil  $t$  afin d'assurer la performance du test. Pour cela, il convient d'abord de définir des critères de qualité pour  $R$ .

**2.1.3 Les erreurs de première espèce.**

Avec l'approche de Neyman-Pearson, on s'intéresse d'abord à l'erreur de première espèce ou erreur de type I, ie la probabilité de rejeter  $\mathcal{H}_0$  à tort. On s'intéresse donc aux éléments de l'ensemble  $R(\mathbf{p}) \cap \mathcal{H}_0(P)$ .

Pour une procédure de test  $R$ , on définit plusieurs erreurs de type I :

- la *k-Family-Wise Error Rate* ( $k$ -FWER) de  $R$  : c'est la probabilité que la procédure  $R$  retourne au moins  $k$  rejets à tort ie au moins  $k$  faux positifs ;

$$\forall P \in \mathcal{P}, \quad k\text{-FWER}(R, P) := \mathbb{P}_P(|R(\mathbf{p}) \cap \mathcal{H}_0(P)| \geq k),$$

où  $k \in \{1, \dots, m\}$  est un paramètre fixé.

Dans le cas particulier où  $k = 1$ , l'erreur est appelée la *family-wise error rate* de  $R$  et est notée  $\text{FWER}(R, P)$  :

$$\forall P \in \mathcal{P}, \quad \text{FWER}(R, P) := \mathbb{P}_P(\exists k \in \mathcal{H} : k \in R(\mathbf{p}) \cap \mathcal{H}_0(P)).$$

- la *Per-Comparison Error Rate* (PCER) de  $R$  (introduit par Blanchard, Delattre, Roquain [3]) : c'est l'espérance du nombre d'erreurs commises ;

$$\forall P \in \mathcal{P}, \quad \text{PCER}(R, P) = \mathbb{E}_P[|R(\mathbf{p}) \cap \mathcal{H}_0(P)|].$$

- la *False Discovery Proportion* (FDP) de  $R$  : c'est la proportion d'erreurs commises dans l'ensemble des hypothèses rejetées ie le taux moyen de faux positifs ;

$$\forall P \in \mathcal{P}, \quad \text{FDP}(R(\mathbf{p}), P) = \frac{|R(\mathbf{p}) \cap \mathcal{H}_0(P)|}{|R(\mathbf{p})| \vee 1}.$$

Ainsi, la FDP est égale à zéro lorsque la procédure ne rejette aucune hypothèse à tort. Comme cette quantité est aléatoire, elle ne définit pas un taux d'erreur ; deux erreurs peuvent découler de la FDP :

- ➔ la queue de la distribution de la FDP de  $R$  : pour un paramètre  $\gamma \in (0, 1)$  fixé,

$$\forall P \in \mathcal{P}, \quad \mathbb{P}_P(\text{FDP}(R(\mathbf{p}), P) > \gamma).$$

➔ la *False Discovery Rate* (FDR) : c'est l'espérance de la FDP de  $R$  ;

$$\forall P \in \mathcal{P}, \text{FDR}(R, P) := \mathbb{E}_P[\text{FDP}(R(\mathbf{p}), P)] = \mathbb{E}_P \left[ \frac{|R(\mathbf{p}) \cap \mathcal{H}_0(P)|}{|R(\mathbf{p})| \vee 1} \right].$$

*Remarques* : Le contrôle de la FDR est beaucoup moins fort que celui de la FWER.

Dans la suite, on cherchera à contrôler l'erreur de première espèce ER au niveau  $\alpha$  c'est-à-dire :

$$\text{pour } \alpha \in (0, 1), \text{ trouver } R = R_\alpha \text{ tel que } \forall P \in \mathcal{P}, \text{ER}(R, P) \leq \alpha.$$

**Exemple 2.1.8** : Procédure de Bonferroni et contrôle de la FWER (d'après [15]).

La procédure de Bonferroni  $R^{\text{Bonf}}(\mathbf{p}) = \{1 \leq i \leq m : p_i \leq \alpha/m\}$  satisfait :

$$\begin{aligned} \mathbb{E}_P[|R^{\text{Bonf}}(\mathbf{p}) \cap \mathcal{H}_0(P)|] &= \mathbb{E}_P[|\{i \in \mathcal{H}_0(P) : p_i \leq \alpha/m\}|] = \mathbb{E}_P \left[ \sum_{i \in \mathcal{H}_0(P)} \mathbf{1}_{p_i \leq \alpha/m} \right] \\ &= \sum_{i \in \mathcal{H}_0(P)} \mathbb{P}_P(p_i \leq \alpha/m) \leq \alpha m_0(P)/m \leq \alpha. \end{aligned}$$

D'après l'inégalité de Markov,

$$\mathbb{P}_P(|R^{\text{Bonf}}(\mathbf{p}) \cap \mathcal{H}_0(P)| \geq 1) \leq \alpha,$$

c'est-à-dire que pour tout  $P \in \mathcal{P}$ ,  $\text{FWER}(R^{\text{Bonf}}, P) \leq \alpha$ .

De plus, si toute loi définie sur  $[0, 1]^m$  dont les marginales sont uniformes correspond à la loi d'une famille de  $p$ -valeur pour un certain  $P_0 \in \mathcal{P}$  sous  $\mathcal{H}_0$  (la vraie hypothèse nulle) alors

$$\sup_{P \in \mathcal{P}} \{\text{FWER}(R^{\text{Bonf}}, P)\} = \alpha.$$

En effet, soit  $U_1, \dots, U_m$  iid  $\sim \mathcal{U}(0, 1)$ . On pose, pour tout  $1 \leq j \leq m$ ,

$$U'_j = \frac{j-1 + U_j}{m} \in \left[ \frac{j-1}{m}, \frac{j}{m} \right].$$

Soit une permutation aléatoire  $\sigma \in \mathfrak{S}_n$  uniformément distribuée sur  $\mathfrak{S}_n$  et indépendante des autres variables aléatoires considérées. On considère alors la loi de  $(U'_{\sigma(j)}, 1 \leq j \leq m)$ .

Par hypothèse, il existe une loi  $P_0$  pour laquelle  $P_0 \in \Theta_{0,j}$  pour tout  $j$  et  $(p_j(X))_j \sim (U'_{\sigma(j)})_j$ . Enfin,

$$\begin{aligned} \text{FWER}(R^{\text{Bonf}}, P_0) &= \mathbb{P}(\exists j \in \{1, \dots, m\} : U'_{\sigma(j)} \leq \frac{\alpha}{m}) \\ &= \mathbb{P} \left( \bigcup_{j=1}^m \bigcup_{k=1}^m \{\sigma(j) = k, U'_k \leq \frac{\alpha}{m}\} \right) \\ &= \mathbb{P} \left( \bigcup_{k=1}^m \{U'_k \leq \frac{\alpha}{m}\} \right) = \mathbb{P}(U'_1 \leq \frac{\alpha}{m}) = \mathbb{P}(U_1 \leq \alpha) = \alpha. \end{aligned}$$

*Remarque* : Si pour une loi  $P_0$  sous  $\mathcal{H}_0$ , les  $(p_j(X))_j$  sont indépendantes alors  $\text{FWER}(R^{\text{Bonf}}, P_0) = 1 - (1 - \alpha/m)^m \simeq \alpha$  pour  $\alpha$  petit.

**Remarque 2.1.9** : Défaut d'adaptativité de la procédure de test

On considère la procédure de Bonferroni  $R^{\text{Bonf}}$  précédente. Pour des  $P_0$  particuliers,  $\text{FWER}(R^{\text{Bonf}}, P_0)$  peut être très inférieure à  $\alpha$ . Par exemple :

► *Forte dépendance entre les tests*

Le test gaussien bilatère avec  $\Gamma_{ij} = 1$  pour tout  $i, j$  donne  $p_i(X) = p_j(X)$  pour tout  $i, j$  d'où

$$\text{FWER}(R^{\text{Bonf}}, P_0) = \frac{\alpha}{m} \ll \alpha.$$

► *Beaucoup d'hypothèses nulles rejetées*

Si  $m_0/m$  n'est pas proche de 1 alors  $\text{FWER}(R^{\text{Bonf}}, P) \leq \alpha m_0/m$  n'est pas proche de  $\alpha$ .

*Objectif* : Construire une procédure qui fournit un contrôle adaptatif de la  $k$ -FWER.

### 2.1.4 Performance du test.

Soit  $\alpha \in (0, 1)$ .

On veut contrôler chaque erreur de première espèce au niveau  $\alpha$  pour une collection de lois  $\mathcal{P}' \subset \mathcal{P}$  aussi large que possible, ie construire une procédure  $R$  telle que

$$\forall P \in \mathcal{P}', \quad \text{ER}(R, P) \leq \alpha, \quad (*)$$

pour  $\mathcal{P}' \subset \mathcal{P}$  aussi large que possible. Comme dans le cas des tests simples, on peut contrôler l'erreur de type I avec  $\mathcal{P} = \mathcal{P}'$  pour  $R = \emptyset$  (test trivial) sans pour autant avoir "un bon test" :  $\mathcal{H}_0$  est systématiquement accepté à tort.

**Définition 2.1.10** : *Procédure conservative*

On dit qu'une procédure  $R$  est moins conservative que  $R'$  si  $R' \subset R$  avec  $\text{ER}(R, P) \leq \alpha$  et  $\text{ER}(R', P) \leq \alpha$ .

**Remarque 2.1.11**

On remarque que si  $R$  est moins conservative que  $R'$  alors  $R$  est plus puissante que  $R'$ .

On va donc chercher à construire une procédure qui soit peu conservative.

On précise aussi que l'on souhaite contrôler l'erreur de première espèce pour tout  $m \geq 2$  et pas seulement pour  $m$  qui tend vers l'infini.

En résumé, afin d'obtenir une procédure vérifiant (\*) et étant peu conservative, on propose une méthode générique :

1. on construit une famille de procédure  $(R_h)_h$  qui dépendent d'un paramètre  $h$  ;
2. on trouve un ensemble de paramètres  $h$  pour lesquels  $R_h$  satisfait (\*) ;
3. on détermine parmi ces valeurs, le paramètre  $h$  qui rend la région de rejet  $R_h$  "la plus grande possible" (ie la procédure la moins conservative).

## 2.2 Contrôle adaptatif de la $k$ -FWER.

Cette section est essentiellement basée sur l'article d'Etienne Roquain [14] et celui de Jelle J. Goeman et Aldo Solari [7].

### 2.2.1 Hypothèses sur les procédures.

On suppose qu'il existe une famille  $\{R_{\mathcal{C}}\}_{\mathcal{C} \subset \mathcal{H}}$  de procédures de test multiple qui satisfait les hypothèses suivantes :

► (P1) :  $\mathcal{C} \mapsto R_{\mathcal{C}}$  est décroissante :

$$\forall \mathcal{C}, \mathcal{C}' \subset \mathcal{H}, \quad \mathcal{C} \subset \mathcal{C}' \Rightarrow R_{\mathcal{C}'} \subset R_{\mathcal{C}};$$

► (P2) :  $R_{\mathcal{C}}$  contrôle la  $k$ -FWER si  $\mathcal{C} = \mathcal{H}_0(P)$  :

$$\forall P \in \mathcal{P}, \quad k\text{-FWER}(R_{\mathcal{H}_0(P)}, P) \leq \alpha.$$

On rappelle que l'on peut décrire la famille  $\{R_{\mathcal{C}}\}_{\mathcal{C} \subset \mathcal{H}}$  par une famille de seuils de la forme

$$R_{\mathcal{C}} = \{1 \leq i \leq m : p_i \leq t_{\mathcal{C}}\},$$

où  $t_{\mathcal{C}} \in [0, 1]$  est un seuil qui peut dépendre des observations et donc de  $\mathbf{p} = (p_i)_{1 \leq i \leq m}$ .

Ainsi, l'hypothèse (P1) est vérifiée dès que l'application  $\mathcal{C} \mapsto t_{\mathcal{C}}$  est décroissante. Il reste alors à choisir  $t_{\mathcal{C}}$  de telle sorte que l'hypothèse (P2) soit vérifiée.

#### Exemple 2.2.1 : Procédure de Bonferroni.

Pour tout  $\mathcal{C} \subset \mathcal{H}$ , on va vérifier que le seuil  $t_{\mathcal{C}} := \min(\alpha k / |\mathcal{C}|, 1)$  satisfait bien les hypothèses (P1) (c'est clair) et (P2).

En effet, d'après l'inégalité de Markov,

$$\mathbb{P}_P(|\mathcal{H}_0(P) \cap R_{\mathcal{H}_0(P)}^{\text{Bonf}}| \geq k) \leq \frac{1}{k} \sum_{i \in \mathcal{H}_0(P)} \mathbb{P}_P(p_i \leq t_{\mathcal{H}_0(P)}) \leq \frac{|\mathcal{H}_0(P)| t_{\mathcal{H}_0(P)}}{k} \leq \alpha.$$

### 2.2.2 Procédure *single-step*.

D'après l'hypothèse (P2), la procédure  $R_{\mathcal{H}_0}$  contrôle la  $k$ -FWER. Cette procédure est appelée procédure *single step* associée à la famille  $\{R_{\mathcal{C}}\}_{\mathcal{C} \subset \mathcal{H}}$ .

Cependant, cette procédure ne peut pas être utilisée en pratique car  $\mathcal{H}_0(P)$  dépend de la loi inconnue  $P$ . Par contre, on peut utiliser la procédure  $R_{\mathcal{H}}$ , car d'après les hypothèses (P1) et (P2), on a  $k\text{-FWER}(R_{\mathcal{H}}, P) \leq k\text{-FWER}(R_{\mathcal{H}_0}, P) \leq \alpha$ . Cependant, cette procédure n'est pas performante en générale car  $R_{\mathcal{H}}$  est plus conservative que  $R_{\mathcal{H}_0}$  dès que  $\mathcal{H}_0(P)$  est plus petit que  $\mathcal{H}$ .

#### Exemple 2.2.2 : Cas gaussien avec dépendance connue (d'après [15]).

On considère le cas gaussien de matrice de covariance  $\Gamma$  connue avec  $\Gamma_{ii} = 1$  pour tout  $i$ .

On considère la statistique de test, pour tout  $1 \leq j \leq m$ ,

$$S_j(X) = \begin{cases} X_j & \text{test unilatère} \\ |X_j| & \text{test bilatère} \end{cases}$$

On considère le seuil

$$t_{\alpha}(\Gamma) = \min\{x \in \mathbb{R} : \mathbb{P}_{\mathcal{H}_0} \left( \max_{1 \leq j \leq m} S_j(X) \leq x \right) \geq 1 - \alpha\}.$$

**Proposition 2.2.3**

Dans le cadre gaussien (unilatère ou bilatère),

$$\forall P \in \mathcal{P}, \text{FWER}(t_\alpha(\Gamma), P) \leq \alpha$$

et il existe  $P_0$  tel que  $\text{FWER}(t_\alpha(\Gamma), P_0) = \alpha$ .

*Démonstration.*

Dans le cas unilatère et bilatère on a,  $S_j(X) \leq S_j(X - \mu)$  pour tout  $j \in \mathcal{H}_0$  alors

$$\begin{aligned} \text{FWER}(t_\alpha(\Gamma), P) &= \mathbb{P}_P(\exists j \in \mathcal{H}_0(P) : S_j(X) > t_\alpha(\Gamma)) = \mathbb{P}_P\left(\sup_{j \in \mathcal{H}_0(P)} \{S_j(X)\} > t_\alpha(\Gamma)\right) \\ &\leq \mathbb{P}_P\left(\sup_{j \in \mathcal{H}_0(P)} \{S_j(X - \mu)\} > t_\alpha(\Gamma)\right) \leq \mathbb{P}_P\left(\sup_{1 \leq j \leq m} \{S_j(\underbrace{X - \mu}_{\sim \mathcal{N}(0, \Gamma)})\} > t_\alpha(\Gamma)\right) \leq \alpha. \end{aligned}$$

De plus, pour  $P_0$  tel que  $\mu = 0$ ,

$$\text{FWER}(t_\alpha(\Gamma), P_0) = \mathbb{P}_{P_0}\left(\sup_{1 \leq j \leq m} \{S_j(X - \mu)\} > t_\alpha(\Gamma)\right) = \alpha$$

car  $\sup_{1 \leq j \leq m} \{S_j(X - \mu)\}$  a une loi continue. □

Nous allons présenter des procédures itératives *step-down*, c'est-à-dire des procédures où toutes les hypothèses sont acceptées initialement et où des hypothèses sont rejetées à chaque étape de l'algorithme.

**2.2.3 Procédure *step-down* pour la FWER.**

On va construire une procédure "proche" de  $R_{\mathcal{H}_0(P)}$  qui contrôle la FWER à partir de la famille  $\{R_C\}_{C \subset \mathcal{H}}$ , via un algorithme itératif.

On note  $A_C$  l'ensemble  $(R_C)^c = \{1, \dots, m\} \setminus R_C$  des hypothèses non rejetées par la procédure  $R_C$ .

**Théorème 2.2.4 : Roquain (2011) [14]**

On suppose que la famille  $\{R_C\}_{C \subset \mathcal{H}}$  de procédures de tests multiples satisfait les conditions (P1) et (P2) et on considère la famille associée  $\{A_C\}_{C \subset \mathcal{H}}$ .

On définit  $\hat{\mathcal{C}}$  par la procédure "step-down" définie comme suit :

- Initialisation :  $\mathcal{C}_0 = \mathcal{H}$  ;
- Étape  $j \geq 1$  : on pose  $\mathcal{C}_j = A_{\mathcal{C}_{j-1}}$ . Si  $\mathcal{C}_j = \mathcal{C}_{j-1}$ , on pose  $\hat{\mathcal{C}} = \mathcal{C}_j$  et l'algorithme est terminé. Sinon, on passe à l'étape  $j + 1$ .

Alors

$$\forall P \in \mathcal{P}, \text{FWER}(R_{\hat{\mathcal{C}}}) \leq \alpha.$$

*Démonstration.*

On considère l'évènement

$$\Omega_0 = \{R_{\mathcal{H}_0(P)} \cap \mathcal{H}_0(P) = \emptyset\} = \{\mathcal{H}_0(P) \subset A_{\mathcal{H}_0(P)}\}.$$

D'après l'hypothèse (P2),

$$\mathbb{P}_P(\Omega_0) = 1 - \mathbb{P}_P(|\mathcal{H}_0(P) \cap R_{\mathcal{H}_0(P)}| \geq 1) \geq 1 - \alpha.$$

D'après l'hypothèse (P1),  $\mathcal{C} \mapsto A_C$  est une application croissante.

Sur  $\Omega_0$ , on a, pour tout  $\mathcal{C} \subset \mathcal{H}$ ,

$$\mathcal{H}_0 \subset \mathcal{C} \Rightarrow A_{\mathcal{H}_0(P)} \subset A_C \Rightarrow \mathcal{H}_0(P) \subset A_C \text{ (car } \mathcal{H}_0(P) \subset A_{\mathcal{H}_0(P)} \subset A_C).$$

Sur  $\Omega_0$  :

► en prenant  $\mathcal{C} = \mathcal{C}_0 = \mathcal{H}$ , on obtient  $\mathcal{H}_0(P) \subset A_{\mathcal{C}_0}$ ;

► en prenant  $\mathcal{C} = \mathcal{C}_1 = A_{\mathcal{C}_0}$ , on obtient  $\mathcal{H}_0(P) \subset A_{\mathcal{C}_1}$  avec  $\mathcal{C}_1 \subset \mathcal{C}_0$ .

En répétant cet algorithme, on construit récursivement l'ensemble  $\widehat{\mathcal{C}}$  et une procédure dont la *FWER* est contrôlée au niveau  $\alpha$ . En effet,  $\widehat{\mathcal{C}} \subset \mathcal{C}_0$  d'où  $\mathcal{C}_0^c \subset \widehat{\mathcal{C}}^c$  et donc  $R_{\widehat{\mathcal{C}}} \subset R_{\mathcal{C}_0^c}$ .

Comme l'ensemble des hypothèses rejetées ne peut que croître au cours de l'algorithme, la procédure obtenue est toujours moins conservatrice que la procédure obtenue par *single step* pour le même contrôle de la *FWER*.  $\square$

**Exemple 2.2.5 :** Procédure step-down de Bonferroni pour la *FWER*.

On peut utiliser le théorème précédent avec

$$R_{\mathcal{C}}^{Bonf} = \{1 \leq i \leq m : p_i \leq \alpha/|\mathcal{C}|\}.$$

**Exemple 2.2.6 :** Modèle gaussien avec covariance  $\Gamma$  connue et contrôle de la *FWER*.

On définit une procédure par

$$R_{\mathcal{C}} = \{1 \leq j \leq m : S_j(X) > t_{\mathcal{C}}(\Gamma)\}$$

où

$$t_{\mathcal{C}}(\Gamma) = \min\{x \in \mathbb{R} : \mathbb{P}_{Z \sim \mathcal{N}(0, \Gamma)}(\max_{j \in \mathcal{C}} \{S_j(Z)\} \leq x) \geq 1 - \alpha\}.$$

Cette "step down" procédure utilise la connaissance de  $\Gamma$  et améliore ainsi la procédure "single step" de la section précédente.

**Remarque 2.2.7 :** Procédure de Holm

On propose ici une version "simplifiée" du théorème précédent. On définit une procédure *step-down* de valeurs critiques  $T_l$ ,  $1 \leq l \leq m$  en rejetant les hypothèses nulles correspondantes aux  $p$ -valeurs réordonnées  $p_{(1)} \leq \dots \leq p_{(m)}$  avec

$$\widehat{l} := \max\{l \in \{0, \dots, m\} : \forall l' \leq l, p_{(l')} \leq T_{l'}\}.$$

Par exemple, la procédure de Holm correspond à la procédure *step down* de Bonferroni pour les valeurs critiques

$$T_l = \frac{\alpha}{m - l + 1}, \quad 1 \leq l \leq m.$$

## 2.2.4 Procédure *step-down* pour la $k$ -FWER.

Le but de cette section est de généraliser le théorème 2.2.4 (ie l'algorithme récursif) au cas de la  $k$ -FWER. Pour tout sous-ensemble  $\mathcal{C} \subset \mathcal{H}$ , on pose :

$$\phi(\mathcal{C}) := \bigcup_{I \subset \mathcal{H}, |I|=k-1} A_{\mathcal{C} \cup I} = \bigcup_{I \subset \mathcal{C}^c, |I| \leq k-1} A_{\mathcal{C} \cup I}.$$

**Théorème 2.2.8 :** Roquain (2011) [14]

On suppose que la famille  $\{R_{\mathcal{C}}\}_{\mathcal{C} \subset \mathcal{H}}$  de procédures de tests multiples satisfait les conditions (P1) et (P2) et on considère la famille associée  $\{A_{\mathcal{C}}\}_{\mathcal{C} \subset \mathcal{H}}$ .

On définit  $\widehat{\mathcal{C}}$  par la procédure "step-down" définie comme suit :

► Initialisation :  $\mathcal{C}_0 = \mathcal{H}$  et calculer  $\phi(\mathcal{C}_0) (= A_{\mathcal{C}_0})$ ;

► Etape  $j \geq 1$  : on pose  $\mathcal{C}_j = \phi(\mathcal{C}_{j-1})$  et calculer  $\phi(\mathcal{C}_j)$ . Si  $\phi(\mathcal{C}_j) = \mathcal{C}_j$ , alors  $\widehat{\mathcal{C}} = \mathcal{C}_j$  et l'algorithme est terminé. Sinon, on passe à l'étape  $j + 1$ .

Alors

$$\forall P \in \mathcal{P}, k\text{-FWER}(R_{\widehat{\mathcal{C}}}) \leq \alpha.$$

*Démonstration.*

On considère l'évènement

$$\Omega_0 := \{|R_{\mathcal{H}_0(P)} \cap \mathcal{H}_0(P)| \leq k - 1\}$$

qui satisfait, par hypothèse,  $\mathbb{P}_P(\Omega_0) \geq 1 - \alpha$ . On remarque tout d'abord que

$$\Omega_0 = \{\exists I_0 \subset \mathcal{H}, |I_0| = k - 1 : \mathcal{H}_0(P) \subset A_{\mathcal{H}_0(P)} \cup I_0\}.$$

On a alors :

$$\exists I \subset \mathcal{H}, |I| = k - 1 : \mathcal{H}_0(P) \subset \mathcal{C} \cup I \implies \exists I \subset \mathcal{H}, |I| = k - 1 : A_{\mathcal{H}_0(P)} \subset A_{\mathcal{C} \cup I} \subset \phi(\mathcal{C}) \quad (*)$$

$$\implies \exists I' \subset \mathcal{H}, |I'| = k - 1 : \mathcal{H}_0(P) \subset \phi(\mathcal{C}) \cup I'. \quad (**)$$

En effet, (\*) découle de la croissance de  $A_{\mathcal{C}}$  sur  $\mathcal{C}$  et il suffit de considérer  $I' = I_0$  pour prouver (\*\*).

Ainsi, sur l'évènement  $\Omega_0$ , pour tout  $\mathcal{C} \subset \mathcal{H}$ ,

$$|\mathcal{C}^c \cap \mathcal{H}_0(P)| \leq k - 1 \implies |(\phi(\mathcal{C}))^c \cap \mathcal{H}_0(P)| \leq k - 1.$$

Par définition,  $\phi$  est une application croissante :

$$\forall \mathcal{C} \subset \mathcal{C}', \phi(\mathcal{C}) \leq \phi(\mathcal{C}').$$

Ainsi, à partir de l'algorithme, comme  $\mathcal{C}_1 \subset \mathcal{C}_0$ , on obtient que  $\mathcal{C}_j \subset \mathcal{C}_{j-1}$  pour tout  $j \geq 1$ . On a donc que la zone de rejet ne peut que croître à chaque étape de l'algorithme. En particulier, la procédure issue de l'algorithme,

$$(\widehat{\mathcal{C}})^c = \bigcap_{|I|=k-1} R_{\widehat{\mathcal{C}} \cup I}$$

est toujours moins conservative que la *single step method*  $R_{\mathcal{C}}$  pour le même contrôle de la  $k$ -FWER ce qui achève la preuve.  $\square$

Appliquons cet algorithme à la procédure  $R_{\mathcal{C}} = \{1 \leq i \leq m : p_i \leq t_{\mathcal{C}}\}$  pour laquelle (P2) est vérifiée et tel que  $\mathcal{C} \mapsto t_{\mathcal{C}}$  est décroissante (ie que l'hypothèse (P1) est valide). La relation de récursivité,  $\mathcal{C}' = \phi(\mathcal{C})$  devient

$$\begin{aligned} (\mathcal{C}')^c &= \bigcap_{I \subset \mathcal{C}^c, |I| \leq k-1} R_{\mathcal{C} \cup I} = \bigcap_{I \subset \mathcal{C}^c, |I| \leq k-1} \{1 \leq i \leq m : p_i \leq t_{\mathcal{C} \cup I}\} \\ &= \{1 \leq i \leq m : p_i \leq \min_{I \subset \mathcal{C}^c, |I| \leq k-1} \{t_{\mathcal{C} \cup I}\}\}. \end{aligned}$$

**Exemple 2.2.9 :** Procédure step-down de Bonferroni pour le contrôle de la  $k$ -FWER.

En considérant les seuils  $t_{\mathcal{C}} := \min(\alpha k / |\mathcal{C}|, 1)$  pour tout  $\mathcal{C} \subset \mathcal{H}$ , on obtient

$$\min_{I \subset \mathcal{C}^c, |I| \leq k-1} \{t_{\mathcal{C} \cup I}\} = \frac{\alpha k}{m \wedge (|\mathcal{C}| + k - 1)}.$$

Il existe une procédure de Holm généralisée reprenant les mêmes idées que la procédure de Holm décrite précédemment.

**Remarque 2.2.10 :** Complexité de l'algorithme

Un problème majeur de cet algorithme concerne le calcul de  $\phi(\cdot)$  qui peut être long si  $k$  est grand car on doit considérer tous les ensembles  $I$  de  $\mathcal{C}^c$  de cardinal au plus  $k - 1$ .

### 2.2.5 Opérateur Next.

#### Procédure de rejet itérative *step-down* et contrôle de la FWER.

Afin de contrôler la FWER, une approche équivalente consiste à introduire un opérateur  $\mathcal{N} : \mathcal{H} \rightarrow \mathcal{H}$ , appelé fonction *Next*. On interprêtera  $\mathcal{N}(R)$  comme les hypothèses rejetées dans la prochaine étape de la procédure, après avoir rejeté  $R$  dans l'étape précédente. La procédure itérative est alors définie par

1.  $R_0 = \emptyset$ ,
2.  $R_{i+1} = R_i \cup \mathcal{N}(R_i)$ ,
3.  $R := \lim_{i \rightarrow +\infty} R_i$ .

Comme  $\mathcal{H}$  est un ensemble fini, la procédure s'arrête lorsque  $R_{n+1} = R_n$ . Ainsi, la limite  $R$  est atteinte en un nombre fini d'étapes.

De plus, on dispose d'un contrôle de la FWER à chaque étape de l'algorithme :

**Proposition 2.2.11 :** *Goeman, Solarì (2010) [7]*

Soit  $\alpha \in (0, 1)$ .

On suppose que l'application  $\mathcal{N} : \mathcal{H} \rightarrow \mathcal{H}$  vérifie les deux hypothèses suivantes :

(H1) - Monotonie :

$$\forall R \subseteq S \subset \mathcal{H}, \quad \mathcal{N}(R) \subseteq \mathcal{N}(S) \cup S,$$

(H2) - Contrôle des faux :

$$\forall P \in \mathcal{P}, \quad \mathbb{P}_P(\mathcal{N}(\mathcal{H}_1(P)) \subseteq \mathcal{H}_1(P)) \geq 1 - \alpha.$$

Alors, pour tout  $P \in \mathcal{P}$ ,

$$\mathbb{P}_P(R \subseteq \mathcal{H}_1(P)) \geq 1 - \alpha.$$

En particulier, pour tout  $P \in \mathcal{P}$ ,  $\text{FWER}(R, P) \leq \alpha$ .

*Démonstration.*

Soit  $P \in \mathcal{P}$ . On considère l'évènement  $E = \{\mathcal{N}(\mathcal{H}_1(P)) \subset \mathcal{H}_1(P)\}$ .

D'après (H2),  $\mathbb{P}_P(E) \geq 1 - \alpha$ .

On se place sur l'évènement  $E$ . Par récurrence, montrons que  $R_n \subseteq \mathcal{H}_1(P)$  pour tout  $n \in \mathbb{N}$ .

►  $R_0 = \emptyset \subseteq \mathcal{H}_1(P)$ .

► Soit  $n \in \mathbb{N}$ . Supposons  $R_n \subseteq \mathcal{H}_1(P)$ .

$$R_{n+1} \cap \mathcal{H}_0(P) = \mathcal{N}(R_n) \cap \mathcal{H}_0(P) \subseteq \mathcal{N}(\mathcal{H}_1(P)) \cap \mathcal{H}_0(P) = \emptyset \quad \text{d'après (H1)}.$$

Cela prouve l'hérédité.

On en déduit que

$$\mathbb{P}_P(R_n \subseteq \mathcal{H}_1(P)) \geq \mathbb{P}_P(E) \geq 1 - \alpha \quad \text{pour tout } n \in \mathbb{N}.$$

Donc,

$$\mathbb{P}_P(R \subseteq \mathcal{H}_1(P)) \geq 1 - \alpha.$$

En particulier,

$$\{\mathbb{P}_P(R \subseteq \mathcal{H}_1(P)) \geq 1 - \alpha\} = \{1 - \mathbb{P}_P(R \cap \mathcal{H}_0(P) \neq \emptyset) \geq 1 - \alpha\} = \{\mathbb{P}_P(R \cap \mathcal{H}_0(P) \neq \emptyset) \leq \alpha\},$$

d'où  $\text{FWER}(R, P) \leq \alpha$ . □

**Exemple 2.2.12 : Procédure de Bonferroni**

On définit l'opérateur, pour tout  $\mathcal{G} \subset \mathcal{H}$ ,

$$\mathcal{N}(\mathcal{G}) := \{i \in \mathcal{H} : p_i \leq \frac{\alpha}{m - |\mathcal{G}|\}\}.$$

► Vérification de (H1) :

Si  $\mathcal{G} \subseteq \mathcal{G}'$  alors  $\frac{\alpha}{m - |\mathcal{G}|} \leq \frac{\alpha}{m - |\mathcal{G}'|}$  donc  $\mathcal{N}(\mathcal{G}) \subseteq \mathcal{N}(\mathcal{G}')$ .

► Vérification de (H2) :

$$\mathbb{P}_P(\mathcal{N}(\mathcal{H}_1(P)) \cap \mathcal{H}_0(P) \neq \emptyset) = \mathbb{P}_P\left(\exists i \in \mathcal{H}_0(P), p_i \leq \frac{\alpha}{m - |\mathcal{H}_1(P)|} = \frac{\alpha}{m_0(P)}\right) \leq \sum_{i \in \mathcal{H}_0(P)} \frac{\alpha}{|\mathcal{H}_0(P)|} = \alpha.$$

**Exemple de procédure de rejet itérative *step-up* et contrôle de la FWER.**

Les procédures itératives *step-up* consistent à rejeter toutes les hypothèses initialement et d'en accepter à chaque itération. On travaille ici avec une famille de  $p$ -valeurs  $(p_i)_{1 \leq i \leq m}$ .

Supposons que pour tout  $R \subset \mathcal{H}$ , on choisisse une suite ordonnée de seuils

$$t_1(R) \geq \dots \geq t_{|\mathcal{H} \setminus R|}(R).$$

On définit alors une procédure itérative en posant

$$\mathcal{N}(R) = \bigcup \{I \subseteq \mathcal{H} \setminus R : p_i \leq t_{|\mathcal{H} \setminus (R \cup I)|+1}(R) \quad \forall i \in I\}.$$

Cette procédure permet aussi de contrôler la FWER :

**Corollaire 2.2.13 : Goeman, Solari (2010)**

Supposons que :

► Si  $R \subseteq S$  alors

$$t_i(S) \geq t_i(R) \quad \text{pour tout } i = 1, \dots, |\mathcal{H} \setminus S|,$$

► Pour tout  $P \in \mathcal{P}$ , si  $R = \mathcal{H}_1(P)$  alors

$$\mathbb{P}_P\left(\bigcup_{I \subseteq \mathcal{H}_0(P)} \{p_i \leq t_{|\mathcal{H}_0(P) \setminus I|+1}(\mathcal{H}_1(P)) \quad \forall i \in I\}\right) \leq \alpha.$$

Alors la procédure de rejet itérative définie ci-dessus satisfait

$$\mathbb{P}_P(R \subseteq \mathcal{H}_1(P)) \geq 1 - \alpha.$$

*Démonstration.*

Pour démontrer ce corollaire, nous allons montrer que l'application  $\mathcal{N}$  vérifie les hypothèses (H1) et (H2) afin d'appliquer la proposition 2.2.11.

► Hypothèse (H1) :

Soit  $R \subseteq S \subset \mathcal{H}$ .

Si  $\mathcal{N}(R) = \emptyset$  alors l'hypothèse de monotonie est claire. Supposons donc que  $\mathcal{N}(R) \neq \emptyset$ .

Soit  $H \in \mathcal{N}(R)$ . Il existe  $I \subseteq \mathcal{H} \setminus R$  tel que  $H \in I$  et  $p_i \leq t_{|\mathcal{H} \setminus (R \cup I)|+1}(R)$  pour tout  $i \in I$ . Si  $H \in S$  alors l'hypothèse de monotonie est vérifiée.

Supposons que  $H \notin S$ . On sait déjà que  $H \in I \setminus S := \tilde{I}$  et que  $\tilde{I} \subseteq \mathcal{H} \setminus S$ . Or, comme les seuils sont rangés par ordre décroissant, on a, par hypothèse du corollaire,

$$t_{|\mathcal{H} \setminus (R \cup I)|+1}(R) \leq t_{|\mathcal{H} \setminus (R \cup I)|+1}(S) \leq t_{|\mathcal{H} \setminus (S \cup I)|+1}(S).$$

Donc, pour tout  $i \in \tilde{I}$ ,

$$p_i \leq t_{|\mathcal{H} \setminus (S \cup I)|+1}(S) = t_{|\mathcal{H} \setminus (S \cup \tilde{I})|+1}(S).$$

Donc  $H \in \mathcal{N}(S)$ .

► Hypothèse (H2) :

$$\mathbb{P}_P(\mathcal{N}(\mathcal{H}_1(P)) \subseteq \mathcal{H}_1(P)) = 1 - \mathbb{P}_P(\mathcal{N}(\mathcal{H}_1(P)) \cap \mathcal{H}_0(P) \neq \emptyset).$$

Or, on a

$$\begin{aligned} \mathbb{P}_P(\mathcal{N}(\mathcal{H}_1(P)) \cap \mathcal{H}_0(P) \neq \emptyset) &= \mathbb{P}_P \left( \bigcup_{I \subseteq \mathcal{H} \setminus \mathcal{H}_1(P)} \{p_i \leq t_{|\mathcal{H} \setminus (\mathcal{H}_1(P) \cup I)|+1}(R) \ \forall i \in I\} \cap \mathcal{H}_0(P) \right) \\ &= \mathbb{P}_P \left( \bigcup_{I \subseteq \mathcal{H}_0(P)} \{p_i \leq t_{|\mathcal{H}_0(P) \setminus I|+1}(R) \ \forall i \in I\} \right) \\ &\leq \alpha \quad \text{par hypothèse du corollaire.} \end{aligned}$$

□

### Procédure de rejet itérative et contrôle de la $k$ -FWER.

A partir de la même procédure itérative que précédemment,

1.  $R_0 = \emptyset$ ,
2.  $R_{i+1} = R_i \cup \mathcal{N}(R_i)$ ,
3.  $R := \lim_{i \rightarrow +\infty} R_i$ ,

on dispose d'un contrôle de la  $k$ -FWER :

#### Proposition 2.2.14

Soit  $\alpha \in (0, 1)$ .

On suppose que l'application  $\mathcal{N} : \mathcal{H} \rightarrow \mathcal{H}$  vérifie les deux hypothèses suivantes :

(A1) - Monotonie :

$$\forall R \subseteq S \subseteq \mathcal{H}, \quad \mathcal{N}(R) \subseteq \mathcal{N}(S),$$

(A2) - Contrôle des faux :

$$\forall P \in \mathcal{P}, \quad \mathbb{P}_P \left( \bigcap_{\substack{I \subseteq \mathcal{H}_0(P), \\ |I| \leq k-1}} \{|\mathcal{H}_0(P) \cap \mathcal{N}(\mathcal{H}_1(P)) \cup I| \leq k-1\} \right) \geq 1 - \alpha.$$

Alors, pour tout  $P \in \mathcal{P}$ ,

$$\mathbb{P}_P(|R \cap \mathcal{H}_0(P)| \geq k) \leq \alpha.$$

En particulier, pour tout  $P \in \mathcal{P}$ ,  $k$ -FWER( $R, P$ )  $\leq \alpha$ .

*Démonstration.*

Soit  $P \in \mathcal{P}$ .

Sous l'hypothèse (A1), on montre par récurrence que  $\mathcal{N}^n(\emptyset) \subset \mathcal{N}^{n+1}(\emptyset)$  pour tout  $n \geq 0$ . On en déduit donc que, pour tout  $n \geq 0$ ,

$$R_{n+1} = \mathcal{N}^{n+1}(\emptyset) = \mathcal{N}(R_n).$$

On considère l'évènement  $E = \{\bigcap_{I \subset \mathcal{H}_0(P), |I| \leq k-1} \{|\mathcal{H}_0(P) \cap \mathcal{N}(\mathcal{H}_1(P) \cup I)| \leq k-1\}\}$ .

D'après (A2),  $\mathbb{P}_P(E) \geq 1 - \alpha$ .

On se place sur l'évènement  $E$ . Par récurrence, montrons que  $|R_n \cap \mathcal{H}_0(P)| \leq k-1$  pour tout  $n \in \mathbb{N}$ .

►  $|R_0 \cap \mathcal{H}_0(P)| = |\emptyset| \leq k-1$ .

► Soit  $n \in \mathbb{N}$ . Supposons que  $|R_n \cap \mathcal{H}_0(P)| \leq k-1$ .

Il existe donc  $I \subset \mathcal{H}_0(P)$ ,  $|I| \leq k-1$  tel que  $R_n \subset \mathcal{H}_1(P) \cup I$ .

Alors

$$\begin{aligned} R_{n+1} &= \mathcal{N}(R_n) \subset \mathcal{N}(\mathcal{H}_1(P) \cup I) \text{ d'après (A1),} \\ &= \mathcal{N}(\mathcal{H}_1(P) \cup I) \cap [\mathcal{H}_0(P) \cup \mathcal{H}_1(P)] \\ &= [\mathcal{N}(\mathcal{H}_1(P) \cup I) \cap \mathcal{H}_0(P)] \cup [\mathcal{N}(\mathcal{H}_1(P) \cup I) \cap \mathcal{H}_1(P)] \\ &\subset [\mathcal{N}(\mathcal{H}_1(P) \cup I) \cap \mathcal{H}_0(P)] \cup \mathcal{H}_1(P). \end{aligned}$$

Posons  $J := \mathcal{N}(\mathcal{H}_1(P) \cup I) \cap \mathcal{H}_0(P)$ . On a  $J \subset \mathcal{H}_0(P)$  et par hypothèse de récurrence,  $|J| \leq k-1$ .

Cela prouve l'hérédité.

On en déduit que

$$\mathbb{P}_P(|R_n \cap \mathcal{H}_0(P)| \leq k-1) \geq \mathbb{P}_P(E) \geq 1 - \alpha \text{ pour tout } n \in \mathbb{N}.$$

Ainsi,

$$\mathbb{P}_P(|R \cap \mathcal{H}_0(P)| \leq k-1) \geq 1 - \alpha,$$

d'où  $k$ -FWER( $R, P$ )  $\leq \alpha$ . □

### Exemple 2.2.15 : Procédure de Bonferroni

Soit  $\mathcal{N} : \mathcal{H} \rightarrow \mathcal{H}$  l'opérateur défini par

$$\forall \mathcal{G} \subset \mathcal{H}, \mathcal{N}(\mathcal{G}) = \{i \in \mathcal{H} : p_i \leq \frac{\alpha k}{m \wedge (m+k-1-|\mathcal{G}|)}\}.$$

► L'hypothèse (A1) est clairement vérifiée.

► Vérification de l'hypothèse (A2) :

Supposons que l'on se place sur l'évènement

$$E = \left\{ \bigcup_{\substack{I \subset \mathcal{H}_0(P), \\ |I| \leq k-1}} |\mathcal{H}_0(P) \cap \mathcal{N}(\mathcal{H}_1(P) \cup I)| \geq k \right\}.$$

Alors il existe  $I \subset \mathcal{H}_0(P)$ ,  $|I| \leq k-1$  tel que  $|\mathcal{H}_0(P) \cap \mathcal{N}(\mathcal{H}_1(P) \cup I)| \geq k$ .

Or

$$\mathcal{H}_0(P) \cap \mathcal{N}(\mathcal{H}_1(P) \cup I) = \{i \in \mathcal{H}_0(P) : p_i \leq \frac{\alpha k}{m \wedge (m+k-1-|\mathcal{H}_1(P)|-|I|)}\} \subset \{i \in \mathcal{H}_0(P) : p_i \leq \frac{\alpha k}{m_0}\}.$$

On a alors :

$$|\mathcal{H}_0(P) \cap \mathcal{N}(\mathcal{H}_1(P) \cup I)| \leq |\{i \in \mathcal{H}_0(P) : p_i \leq \frac{\alpha k}{m_0}\}|.$$

Ainsi,

$$\mathbb{P}_P(E) \leq \mathbb{P}_P\left(|\{i \in \mathcal{H}_0(P) : p_i \leq \frac{\alpha k}{m_0}\}| \geq k\right)$$

$$\begin{aligned} &\leq \frac{1}{k} \sum_{i \in \mathcal{H}_0(P)} \mathbb{P}_P \left( p_i \leq \frac{\alpha k}{m_0} \right) \\ &\leq \frac{1}{k} m_0 \frac{\alpha k}{m_0} = \alpha. \end{aligned}$$

D'après le théorème 2.2.14, la procédure de Bonferroni définie à partir de cet opérateur contrôle la  $k$ -FWER au niveau  $\alpha$ .

## 2.3 Contrôle de la FDR.

Cette section est essentiellement basée sur le cours d'Etienne Roquain [15].

### 2.3.1 Procédure de Benjamini Hochberg.

On rappelle qu'à partir d'une famille de  $p$ -valeurs  $\mathbf{p} = \{p_i, 1 \leq i \leq m\} \in [0, 1]^m$ , la FDP est la proportion d'erreurs commises dans l'ensemble des hypothèses rejetées ie le taux moyen de faux positifs ;

$$\forall P \in \mathcal{P}, \text{FDP}(R(\mathbf{p}), P) = \frac{|R(\mathbf{p}) \cap \mathcal{H}_0(P)|}{|R(\mathbf{p})| \vee 1}.$$

On définit alors la FDR :

$$\forall P \in \mathcal{P}, \text{FDR}(R, P) := \mathbb{E}_P[\text{FDP}(R(\mathbf{p}), P)] = \mathbb{E}_P \left[ \frac{|R(\mathbf{p}) \cap \mathcal{H}_0(P)|}{|R(\mathbf{p})| \vee 1} \right].$$

La procédure de Benjamini Hochberg (BH) introduit en 1995 dans [2] est la suivante :

- On considère  $\{p_{(i)}, 1 \leq i \leq m\}$ , la famille des  $p$ -valeurs rangées dans l'ordre croissant.
- On se donne un critère d'arrêt donné par :

$$\hat{l} = \max\{l \in \{1, \dots, m\} \mid p_{(l)} \leq \alpha \frac{l}{m}\},$$

$$\text{et } \hat{l} = 0 \text{ si } \{l \in \{1, \dots, m\} \mid p_{(l)} \leq \alpha \frac{l}{m}\} = \emptyset.$$

- On rejette alors les hypothèses correspondants à  $p_{(1)}, \dots, p_{(\hat{l})}$  :

$$R^{BH} = \{1 \leq j \leq m \mid p_j \leq \underbrace{\frac{\alpha(\hat{l} \vee 1)}{m}}_{:=t(\mathbf{p}) \text{ seuil}}\}.$$

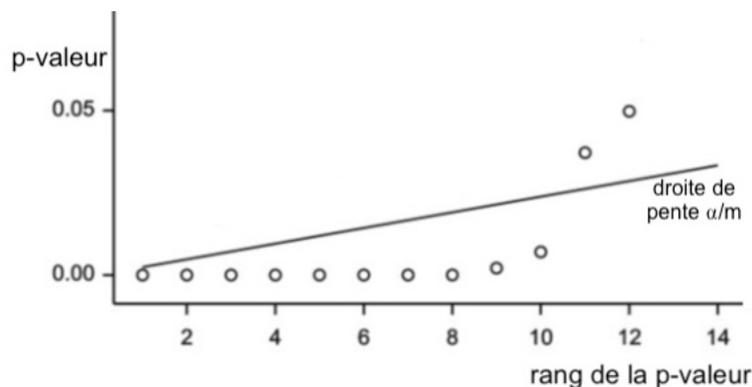


FIGURE 2.1 – Illustration de la procédure de Benjamini Hochberg avec  $\hat{l} = 10$ .

Cette procédure permet de contrôler la FDR au niveau  $\alpha$  (on notera parfois cette procédure  $BH(\alpha)$ ).

**Proposition 2.3.1 :** *Benjamini, Hochberg (1995) et Benjamini, Yekutieli (2001)*

Supposons qu'il existe  $P \in \mathcal{P}$  tel que  $(p_j, 1 \leq j \leq m)$  soit une famille de  $p$ -valeurs *mutuellement indépendantes*. Soit  $\alpha \in (0, 1)$ . Alors :

1.  $\text{FDR}(R^{BH}, P) \leq \alpha \frac{m_0(P)}{m} \leq \alpha$ .
2. De plus, si chaque  $p$ -valeur suit une loi uniforme sous  $P$  alors

$$\text{FDR}(R^{BH}, P) = \alpha \frac{m_0(P)}{m} \leq \alpha.$$

*Démonstration.* (d'après Ferreira et Zwinderman (2006))

On définit la famille  $(\theta_j)_{1 \leq j \leq m}$  par

$$\theta_j = \begin{cases} 0 & \text{si } P \in \Theta_{0,j}, \\ 1 & \text{sinon.} \end{cases}$$

On a donc

$$|R^{BH} \cap \mathcal{H}_0(P)| = \sum_{j=1}^m (1 - \theta_j) \mathbf{1}_{p_j \leq \alpha(\widehat{l} \vee 1)/m}.$$

Ainsi,

$$\text{FDR}(R^{BH}, P) = \sum_{j=1}^m (1 - \theta_j) \mathbb{E}_P \left[ \frac{\mathbf{1}_{\{p_j \leq \alpha(\widehat{l} \vee 1)/m\}}}{\widehat{l} \vee 1} \right].$$

On définit à présent  $\mathbf{p}^{-j} := \{p_1, \dots, p_m\} \setminus \{p_j\}$ . On notera  $\tilde{\mathbf{p}}^{-j} := \{\tilde{p}_1, \dots, \tilde{p}_m\}$ , c'est-à-dire que  $p_1 = \tilde{p}_1, \dots, p_{j-1} = \tilde{p}_{j-1}, p_{j+1} = \tilde{p}_j, \dots, p_m = \tilde{p}_{m-1}$ .

On note alors  $\widehat{l}^{(-j)} := \max\{l \in \{1, \dots, m-1\}, \tilde{p}_{(l)} \leq \alpha \frac{l+1}{m}\}$ .

Montrons que

$$p_j \leq \frac{\alpha}{m}(\widehat{l} \vee 1) \iff \widehat{l} = \widehat{l}^{(-j)} + 1.$$

Supposons que  $p_j \leq \frac{\alpha}{m}(\widehat{l} \vee 1)$  ie que  $p_j$  soit rejetée. Si  $\widehat{l} \geq 2$ , comme

$$\widehat{l} = \max\{l \in \{1, \dots, m\} \mid p_{(l)} \leq \alpha \frac{l}{m}\},$$

alors

$$\begin{aligned} p_{(\widehat{l})} \leq \frac{\alpha}{m}\widehat{l} \text{ et } p_{(s)} > \frac{\alpha}{m}s \quad \forall s > \widehat{l} &\implies \tilde{p}_{(\widehat{l}-1)} \leq \frac{\alpha}{m}\widehat{l} \text{ et } \tilde{p}_{(s-1)} > \frac{\alpha}{m}s \quad \forall s > \widehat{l} \text{ car } j \leq \widehat{l} \\ &\implies \tilde{p}_{(\widehat{l}-1)} \leq \frac{\alpha}{m}\widehat{l} \text{ et } \tilde{p}_{(s)} > \frac{\alpha}{m}(s+1) \quad \forall s > \widehat{l}-1 \\ &\implies \widehat{l}^{(-j)} = \widehat{l}-1 \\ &\implies \widehat{l} = \widehat{l}^{(-j)} + 1. \end{aligned}$$

Enfin, si  $\widehat{l} = 1$  ou si  $\widehat{l} = 0$  alors  $p_j \leq \alpha/m$  et  $\tilde{p}_{(s)} > \frac{\alpha}{m}(s+1) \quad \forall s \geq 1$  d'où  $\widehat{l}^{(-j)} = 0$ .

Supposons à présent que  $p_j > \frac{\alpha}{m}(\widehat{l} \vee 1)$  ie que  $p_j$  ne soit pas rejetée. Comme  $\widehat{l} = \max\{l \in \{1, \dots, m\} \mid p_{(l)} \leq \alpha \frac{l}{m}\}$ , alors,

$$\begin{aligned} p_{(\widehat{l})} \leq \frac{\alpha}{m}\widehat{l} &\implies \tilde{p}_{(\widehat{l})} \leq \frac{\alpha}{m}\widehat{l} \text{ car } j > \widehat{l} \\ &\implies \tilde{p}_{(\widehat{l})} \leq \frac{\alpha}{m}\widehat{l} + \frac{\alpha}{m} \\ &\implies \widehat{l} \leq \widehat{l}^{(-j)} \\ &\implies \widehat{l} < \widehat{l}^{(-j)} + 1. \end{aligned}$$

On a alors

$$\begin{aligned}
\text{FDR}(R^{BH}, P) &= \sum_{j=1}^m (1 - \theta_j) \mathbb{E}_P \left[ \frac{\mathbf{1}\{p_j \leq \alpha(\widehat{l}^{(-j)} + 1)/m\}}{\widehat{l}^{(-j)} + 1} \right] \quad \text{avec } p_j \text{ et } \widehat{l}^{(-j)} \text{ indépendantes} \\
&= \sum_{j=1}^m (1 - \theta_j) \mathbb{E}_P \left[ \mathbb{E} \left[ \frac{\mathbf{1}\{p_j \leq \alpha(\widehat{l}^{(-j)} + 1)/m\}}{\widehat{l}^{(-j)} + 1} \mid \widehat{l}^{(-j)} \right] \right] \\
&= \sum_{j=1}^m (1 - \theta_j) \mathbb{E}_P \left[ \frac{1}{\widehat{l}^{(-j)} + 1} \underbrace{\mathbb{P}_P \left[ p_j \leq \alpha(\widehat{l}^{(-j)} + 1)/m \mid \widehat{l}^{(-j)} \right]}_{\leq \alpha(\widehat{l}^{(-j)} + 1)/m} \right] \\
&\leq \alpha m_0 / m.
\end{aligned}$$

Si on suppose de plus que les  $(p_j)_j$  sont de loi uniforme sous  $P$ , alors l'inégalité ci-dessus est une égalité.  $\square$

On ne peut pas s'affranchir de l'hypothèse d'indépendance des  $p$ -valeurs.

**Proposition 2.3.2 :** *Benjamini, Yekutieli (2001)*

Soit  $(p_j, 1 \leq j \leq m)$  une famille de  $p$ -valeurs associée au test multiple.  
Posons  $\gamma_m = 1 + 1/2 + \dots + 1/m$ . Alors :

$$\forall P \in \mathcal{P}, \text{FDR}(R^{BH}, P) \leq \alpha \frac{m_0(P)}{m} \gamma_m \leq \alpha \gamma_m.$$

*Démonstration.*

On a vu que

$$\text{FDR}(R^{BH}, P) = \sum_{j=1}^m (1 - \theta_j) \mathbb{E}_P \left[ \frac{\mathbf{1}\{p_j \leq \alpha(\widehat{l} \vee 1)/m\}}{\widehat{l} \vee 1} \right].$$

De plus, on a, pour tout  $l \in \{1, \dots, m\}$ ,

$$\frac{1}{l} = \sum_{k \geq l} \left( \frac{1}{k} - \frac{1}{k+1} \right) = \sum_{k \geq l} \frac{1}{k(k+1)}.$$

Ainsi,

$$\begin{aligned}
\mathbb{E}_P \left[ \frac{\mathbf{1}\{p_j \leq \alpha(\widehat{l} \vee 1)/m\}}{\widehat{l} \vee 1} \right] &= \mathbb{E}_P \left[ \sum_{k=1}^{\infty} \frac{1}{k(k+1)} \mathbf{1}\{p_j \leq \alpha(\widehat{l} \vee 1)/m\} \mathbf{1}\{k \geq \widehat{l} \vee 1\} \right] \\
&\leq \sum_{k \geq 1} \frac{1}{k(k+1)} \mathbb{P}_P \left( p_j \leq \frac{\alpha(k \wedge m)}{m}, k \geq \widehat{l} \vee 1 \right) \quad \text{par Fubini-Tonelli} \\
&\leq \frac{\alpha}{m} \sum_{k \geq 1} \frac{k \wedge m}{k(k+1)}.
\end{aligned}$$

Enfin, il reste à remarquer que

$$\sum_{k \geq 1} \frac{k \wedge m}{k(k+1)} = \sum_{k=1}^{m-1} \frac{1}{k+1} + \sum_{k \geq m} m \frac{1}{k(k+1)} = 1 + \frac{1}{2} + \dots + \frac{1}{m} = \gamma_m.$$

$\square$

*Conséquence :* La procédure de Benjamini Yekutieli définie par  $R^{BH(\alpha/\gamma_m)}$  contrôle toujours la FDR au niveau  $\alpha$ . Mais cette procédure a le défaut d'être trop conservatrice en pratique.

**Remarque 2.3.3 : Défaut d'adaptativité**

La procédure BH appliquée à un niveau  $\alpha$  garantit un contrôle de la FDR au niveau  $m_0(P)/m$  qui peut être strictement inférieur au niveau prescrit dès que  $m_0(P)/m < 1$ .

Afin de corriger ce défaut d'adaptativité, une idée est de considérer un estimateur de la proportion  $m_0(P)/m$  d'hypothèses nulles vraies.

**2.3.2 Conditions de dépendance et contrôle de l'erreur de première espèce.**

Dans cette section, on va d'abord chercher à affaiblir la condition d'indépendance des  $p$ -valeurs de la proposition 2.3.1 pour contrôler l'erreur de première espèce. Dans un second temps, on va vouloir contrôler les fluctuations de la FDP dans un cadre général de dépendance entre les  $p$ -valeurs.

**Théorème 2.3.4 : Roquain (2011) [14]**

Soit  $(p_j, 1 \leq j \leq m)$  une famille de  $p$ -valeurs associée au test multiple. On considère la procédure  $BH(\alpha)$  associée aux  $\hat{l}$  rejets.

Supposons qu'il existe  $P \in \mathcal{P}$  vérifiant la condition :

$$\forall j \in \{1, \dots, m\} \ P \in H_{0,j}, \ \forall l \in \{2, \dots, m\} \ \mathbb{P}_P \left( \hat{l} \leq l - 1 \mid p_j \leq \frac{\alpha(l-1)}{m} \right) \leq \mathbb{P}_P \left( \hat{l} \leq l - 1 \mid p_j \leq \frac{\alpha l}{m} \right). \quad (\dagger)$$

Alors

$$\text{FDR}(R^{BH}, P) \leq \alpha \frac{m_0(P)}{m}.$$

*Démonstration.*

$$\begin{aligned} \text{FDR}(R^{BH}, P) &= \sum_{j=1}^m (1 - \theta_j) \sum_{l=1}^m \frac{1}{l} \mathbb{P}_P(\hat{l} = l, p_j \leq \frac{\alpha l}{m}) \\ &= \sum_{j=1}^m (1 - \theta_j) \sum_{l=1}^m \frac{1}{l} \left( \mathbb{P}_P(\hat{l} \leq l, p_j \leq \frac{\alpha l}{m}) - \mathbb{P}_P(\hat{l} \leq l - 1, p_j \leq \frac{\alpha l}{m}) \right) \\ &= \sum_{j=1}^m (1 - \theta_j) \sum_{l=1}^m \frac{1}{l} \left( \mathbb{P}_P(\hat{l} \leq l \mid p_j \leq \frac{\alpha l}{m}) - \underbrace{\mathbb{P}_P(\hat{l} \leq l - 1 \mid p_j \leq \frac{\alpha l}{m})}_{\geq \mathbb{P}_P(\hat{l} \leq l - 1 \mid p_j \leq \alpha(l-1)/m) \text{ par } (\dagger)} \right) \times \underbrace{\mathbb{P}_P(p_j \leq \frac{\alpha l}{m})}_{\leq \alpha l/m} \\ &\leq \frac{\alpha}{m} \sum_{j=1}^m (1 - \theta_j) \sum_{l=1}^m \left( \mathbb{P}_P(\hat{l} \leq l \mid p_j \leq \frac{\alpha l}{m}) - \mathbb{P}_P(\hat{l} \leq l - 1 \mid p_j \leq \alpha(l-1)/m) \right) \\ &= \frac{\alpha}{m} \sum_{j=1}^m (1 - \theta_j) \mathbb{P}_P(\hat{l} \leq m \mid p_j \leq \alpha) = \frac{\alpha m_0}{m}. \end{aligned}$$

□

Nous allons maintenant présenter quelques cas particuliers pour lesquels on dispose de la condition  $(\dagger)$ .

**Cas 1 : Dépendance positive (Benjamini, Yekutieli (2001)).**

On dit qu'un ensemble  $D \subset [0, 1]^m$  est *croissant* si

$$\forall p, p' \in [0, 1]^m, (p \in D, p_j \leq p'_j \ \forall j) \implies p' \in D.$$

(†) est vraie si pour tout ensemble  $D \subset [0, 1]^m$  mesurable et croissant, pour tout  $j \in \mathcal{H}_0(P)$ , l'application  $t \in [0, 1] \mapsto \mathbb{P}_P(p \in D \mid p_j \leq t)$  est croissante. Cette condition est appelée *wPRDS* (weak Positively Regressively Dependant on each one of a Subset).

**Définition 2.3.5 :** *Conditions de PRDS pour une famille (finie) de  $p$ -valeurs - Blanchard, Delattre, Roquain (2014) [3]*

Soit  $\mathcal{H}$  un ensemble fini. Pour tout sous-ensemble  $\mathcal{H}'$  de  $\mathcal{H}$ , on dit que la famille de  $p$ -valeurs  $\mathbf{p}(X) = (p_j(X))_{j \in \mathcal{H}}$  satisfait la condition de *PRDS faible* (*wPRDS*) sur  $\mathcal{H}'$  pour la loi  $P$ , si pour tout  $j \in \mathcal{H}'$ , pour tout ensemble mesurable croissant  $D$  de  $[0, 1]^{|\mathcal{H}'|}$ , la fonction

$$t \in [0, 1] \mapsto \mathbb{P}_P(\mathbf{p}(X) \in D \mid p_j(X) \leq t) \text{ est croissante sur } \{t \in [0, 1] : \mathbb{P}_P(p_j(X) \leq t) > 0\}.$$

On dit que la famille satisfait la condition de *PRDS (forte)* si la fonction

$$t \in [0, 1] \mapsto \mathbb{P}_P(\mathbf{p}(X) \in D \mid p_j(X) = t) \text{ est croissante.}$$

On remarque que la probabilité conditionnelle par rapport à l'évènement  $\{p_j(X) \leq t\}$  est définie (comme rapport de probabilités) pour tout  $t \in [0, 1]$  dès lors que cet évènement est de probabilité strictement positive; tandis que la probabilité conditionnelle par rapport à  $p_j(X) = t$  peut seulement être définie comme une espérance conditionnelle. Ainsi, la définition de condition de PRDS forte est définie à ensemble  $p_j(X)$ -négligeable près.

On remarque que les propriétés de *wPRDS* et de *PRDS (forte)* sont invariantes par transformations monotones : si  $\mathbf{p}(X)$  est PRDS (faible ou forte) alors  $f(\mathbf{p}(X))$  est PRDS (respectivement faible ou forte) si les  $f_j$  sont toutes croissantes ou décroissantes.

**Lemme 2.3.6 :** *Benjamini, Yekutieli (2001)*

La condition de PRDS (forte) implique la condition de PRDS faible.

*Démonstration.*

Posons  $f(u) = \mathbb{P}_P(p \in D \mid p_j = u)$ . Par hypothèse, l'application  $f$  est croissante.

Soit  $u \geq 0$  tel que  $\mathbb{P}_P(p_j \leq u) > 0$ .

Pour tout  $u' \geq u$ , posons  $\gamma = \mathbb{P}_P(p_j \leq u \mid p_j \leq u')$  (cette quantité est bien définie car  $\mathbb{P}_P(p_j \leq u) > 0$ ).

Montrons que  $\mathbb{P}_P(p \in D \mid p_j \leq u') \geq \mathbb{P}_P(p \in D \mid p_j \leq u)$ .

$$\begin{aligned} \mathbb{P}_P(p \in D \mid p_j \leq u') &= \mathbb{E}_P(\mathbf{1}_{p \in D} \mid p_j \leq u') \\ &= \mathbb{E}_P \left[ \underbrace{\mathbb{E}_P(\mathbf{1}_{p \in D} \mid p_j)}_{=f(p_j)} \mid p_j \leq u' \right] \text{ car } \{p_j \leq u'\} \subset \sigma(p_j) \\ &= \mathbb{E}_P(f(p_j) \mid p_j \leq u') \\ &= \mathbb{E}_P(f(p_j)\mathbf{1}_{p_j \leq u} \mid p_j \leq u') + \mathbb{E}_P(f(p_j)\mathbf{1}_{p_j > u} \mid p_j \leq u') \\ &= \frac{\mathbb{E}_P(f(p_j)\mathbf{1}_{p_j \leq u} \mathbf{1}_{p_j \leq u'}) \mathbb{P}_P(p_j \leq u)}{\mathbb{P}_P(p_j \leq u') \mathbb{P}_P(p_j \leq u)} + \frac{\mathbb{E}_P(f(p_j)\mathbf{1}_{p_j > u} \mathbf{1}_{p_j \leq u'}) \mathbb{P}_P(u < p_j \leq u')}{\mathbb{P}_P(p_j \leq u') \mathbb{P}_P(u < p_j \leq u')} \\ &= \mathbb{E}_P(f(p_j) \mid p_j \leq u) \frac{\mathbb{P}_P(\{p_j \leq u\} \cap \{p_j \leq u'\})}{\mathbb{P}_P(p_j \leq u')} \\ &\quad + \mathbb{E}_P(f(p_j) \mid u < p_j \leq u') \frac{\mathbb{P}_P(\{u < p_j\} \cap \{p_j \leq u'\})}{\mathbb{P}_P(p_j \leq u')} \\ &= \gamma \mathbb{E}_P(f(p_j) \mid p_j \leq u) + (1 - \gamma) \mathbb{E}_P(f(p_j) \mid u < p_j \leq u') \\ &\geq \mathbb{E}_P(f(p_j) \mid p_j \leq u) \text{ car } f \text{ est croissante} \\ &= \mathbb{P}_P(p \in D \mid p_j \leq u) \end{aligned}$$

□

**Exemple 2.3.7 : Vecteurs gaussiens**

Soit  $X = (X_1, \dots, X_n)$  un vecteur gaussien de moyenne  $\nu \in \mathbb{R}^n$  et de matrice de covariance  $\Sigma \in M_n(\mathbb{R})$ . Soit  $\mathcal{H}' \subset \{1, \dots, n\}$ . Si pour tout  $i \in \mathcal{H}'$  et pour tout  $j \in \{1, \dots, m\}$ ,  $\Sigma_{i,j} \geq 0$  alors  $X$  est PRDS sous  $\mathcal{H}'$ .

*Démonstration.*

Soit  $i \in \mathcal{H}'$ . Pour fixer les idées, on suppose que  $i = 1$ .

On pose

$$X = \begin{pmatrix} X_1 \\ X_{(-1)} \end{pmatrix}, \quad \nu = \begin{pmatrix} \nu_1 \\ \nu_{(-1)} \end{pmatrix} \quad \text{et} \quad \Sigma = \begin{pmatrix} \Sigma_{1,1} & \Sigma_{1,(-1)} \\ \Sigma_{(-1),1} & \Sigma_{(-1),(-1)} \end{pmatrix}.$$

D'après le résultat sur le conditionnement gaussien, la loi de  $X_{(-1)}$  sachant  $X_1 = x$  est donnée par

$$\mathcal{L}(X_{(-1)} \mid X_1 = x) = \mathcal{N}(\nu_{(-1)} + \Sigma_{(-1),1} \Sigma_{1,1}^{-1} (x - \nu_1), \Sigma_{(-1),(-1)} - \Sigma_{(-1),1} \Sigma_{1,1}^{-1} \Sigma_{1,(-1)}).$$

Par hypothèse,  $\Sigma_{(-1),1} \geq 0$  et donc la moyenne conditionnelle est croissante en  $x$ . Comme la matrice de covariance de la loi conditionnelle ne dépend pas de  $x$ , la loi conditionnelle est croissante en probabilité par rapport à  $x$ . Alors, pour toute fonction croissante  $f$ , d'après le lemme de transfert,  $x \leq x'$  implique

$$\mathbb{E}_P(f(X_{(-1)}) \mid X_1 = x) \leq \mathbb{E}_P(f(X_{(-1)}) \mid X_1 = x').$$

Soit  $D$  un ensemble croissant. Par définition de  $D$ , la fonction indicatrice de  $D$  est une fonction croissante. Pour une telle fonction  $f$ , on obtient la propriété de PRDS. □

*Remarques :*

1. Autrement dit, pour des données gaussiennes, la condition de PRDS est équivalente à des corrélations positives.
2. De plus, la fonction définie pour tout  $j \in \{1, \dots, n\}$  par  $G_j : x \in \mathbb{R} \mapsto \mathbb{P}_P(X_j \geq x)$  est décroissante donc la famille de  $p$ -valeurs  $(p_j(X_j))_j$  satisfait aussi la condition de PRDS.

On remarque par ailleurs que  $D_l = \{p \in [0, 1]^m, \widehat{l}(p) \leq l - 1\}$  est un ensemble croissant de  $[0, 1]^m$ .

**Cas 2 : Dépendance martingale (Storey et al (2004) - Heesen, Janssen (2015)).**

On considère la filtration  $(\mathcal{F}_l)_{1 \leq l \leq m}$  définie par

$$\mathcal{F}_l := \sigma \left( \mathbf{1}\{p_j \leq \frac{\alpha l'}{m}\}, l \leq l' \leq m, 1 \leq j \leq m \right), \quad 1 \leq l \leq m.$$

On remarque que  $D_l = \{\widehat{l} \leq l - 1\}$  est mesurable par rapport à la filtration  $\mathcal{F}_l$ .

En effet,

$$\widehat{l} \leq l - 1 \Leftrightarrow \forall l' \geq l, p_{(l')} > \frac{\alpha l'}{m} \Leftrightarrow \forall l' \geq l, \sum_{j=1}^m \mathbf{1}\{p_j \leq \frac{\alpha l'}{m}\} < l'.$$

On a donc

$$\begin{aligned} \mathbb{P}_P(\widehat{l} \leq l - 1 \mid p_j \leq \frac{\alpha(l-1)}{m}) &= \mathbb{E}_P \left[ \mathbf{1}\{\widehat{l} \leq l - 1\} \frac{\mathbf{1}\{p_j \leq \alpha(l-1)/m\}}{\mathbb{P}_P(p_j \leq \alpha(l-1)/m)} \right] \\ &= \mathbb{E}_P \left[ \mathbf{1}\{\widehat{l} \leq l - 1\} \mathbb{E}_P \left[ \underbrace{\frac{\mathbf{1}\{p_j \leq \alpha(l-1)/m\}}{\mathbb{P}_P(p_j \leq \alpha(l-1)/m)}}_{=: M_{l-1}^{(j)}} \mid \mathcal{F}_l \right] \right]. \end{aligned}$$

On suppose que les variables aléatoires  $(M_{l-1}^{(j)})_{2 \leq l \leq m}$  satisfont une condition de type martingale ie

$$\forall l \in \{2, \dots, m\}, \forall j \in \mathcal{H}_0(P), \mathbb{E}_P[M_{l-1}^{(j)} \mid \mathcal{F}_l] \leq M_l^{(j)}.$$

On a donc

$$\mathbb{P}_P(\widehat{l} \leq l-1 \mid p_j \leq \frac{\alpha(l-1)}{m}) \leq \mathbb{E}_P[\mathbb{1}\{\widehat{l} \leq l-1\} M_l^{(j)}] = \mathbb{P}_P(\widehat{l} \leq l-1 \mid p_j \leq \frac{\alpha l}{m}) \text{ et (†) est vérifiée.}$$

Un cadre général de dépendance entre les  $p$ -valeurs ne permet pas de contrôler facilement la FDR. On va alors chercher à contrôler les fluctuations de la FDP. Formellement, on veut contrôler la FDP, c'est-à-dire trouver un seuil  $t = t(\mathbf{p})$  tel que  $\forall P \in \mathcal{P}, \mathbb{P}_P(\text{FDP}(R(t), P) \leq \alpha) \geq 1 - \xi$ .

Pour cela on va utiliser une procédure step-up corrigée qui consiste à trouver de nouvelles valeurs critiques  $\tau_l, 1 \leq l \leq m$  telles que pour  $\widehat{l} = \max\{l \in \{0, 1, \dots, m\}, p_{(l)} \leq \tau_l\}$ , le seuil  $t(\mathbf{p}) = \tau_{\widehat{l}}$  contrôle la FDP.

**Proposition 2.3.8 :** *Guo et al (2014)*

On considère une famille de  $p$ -valeurs et une probabilité  $P \in \mathcal{P}$  satisfaisant l'hypothèse de  $w$ PRDS. On considère la procédure de step-up corrigée avec les valeurs critiques de Lehmann Romano :

$$\tau_l := \xi \frac{(\lfloor \alpha l \rfloor + 1)}{m}, \quad 1 \leq l \leq m.$$

Alors

$$\mathbb{P}_P(\text{FDP}(\tau_{\widehat{l}}, P) \leq \alpha) \geq 1 - \xi.$$

*Démonstration.*

On a

$$\{\text{FDP}(\tau_{\widehat{l}}) > \alpha\} = \left\{ \sum_{j=1}^m (1 - \theta_j) \mathbb{1}\{p_j \leq \tau_{\widehat{l}}\} \geq \lfloor \alpha \widehat{l} \rfloor + 1 \right\}.$$

On a donc  $m_0(P) \geq \lfloor \alpha \widehat{l} \rfloor + 1$ . Ainsi,

$$\{\text{FDP}(\tau_{\widehat{l}}) > \alpha\} \subset \left\{ \exists k \in \{1, \dots, m_0(P)\}, p_{(k:\mathcal{H}_0)} \leq \frac{\xi k}{m} \right\},$$

où  $p_{(k:\mathcal{H}_0)}$  est la  $k$ -ième plus petite  $p$ -valeur.

En effet, cette condition est en particulier vérifiée pour  $k = \lfloor \alpha \widehat{l} \rfloor + 1$ .

On a donc,

$$\mathbb{P}_P(\text{FDP}(\tau_{\widehat{l}}, P) > \alpha) \leq \mathbb{P}_P(\exists k \in \{1, \dots, m_0(P)\}, p_{(k:\mathcal{H}_0)} \leq \frac{\xi k}{m}) \leq \mathbb{P}_P(p_{(k:\mathcal{H}_0)} \leq \frac{\xi k}{m}).$$

D'après l'hypothèse de  $w$ PRDS,

$$\mathbb{P}_P(\text{FDP}(\tau_{\widehat{l}}, P) > \alpha) \leq \frac{\xi m_0(P)}{m} \leq \xi.$$

□



# Chapitre 3

## Erreurs de seconde espèce et théorie minimax pour les tests multiples.

En théorie des tests classiques d'une hypothèse, la construction de tests par agrégation permet de contrôler l'erreur de première espèce tout en satisfaisant des critères d'optimalité liés à l'erreur de seconde espèce sous certaines classes d'alternatives. Ces propriétés de type minimax sont basées sur la notion de vitesse de séparation introduite par Ingster sous un angle asymptotique et par Baraud sous un angle non asymptotique. Le but de ce chapitre est de donner quelques définitions des erreurs de seconde espèce pour les tests multiples présentes dans la littérature puis de montrer qu'il existe une correspondance entre les tests agrégés et les tests multiples permettant d'étendre la théorie des vitesses de séparation pour les tests multiples. Cela permettra de définir un nouveau critère d'optimalité lié à l'erreur de seconde espèce et une théorie minimax pour les tests multiples.

### 3.1 Premières définitions des erreurs de seconde espèce pour les tests multiples.

#### 3.1.1 Rappels sur les tests multiples et procédure $\min p$ .

Dans cette section, basée sur l'article [6], nous allons présenter la procédure de tests multiples  $\min p$ . Cela nous permettra de fixer les notations pour la suite du chapitre.

Soit  $(\mathbb{X}, \mathcal{X}, \mathcal{P})$  un modèle statistique et  $X$  une observation de loi  $P \in \mathcal{P}$ .

Une hypothèse  $H$  est un sous-ensemble de  $\mathcal{P}$ . On dit que  $H$  est vraie sous une probabilité  $P \in \mathcal{P}$  si  $P \in H$  et fausse sous  $P$  si  $P \notin H$ .

On se donne  $\mathcal{H}$  une collection finie d'hypothèses  $H$ . On veut tester simultanément  $H$  contre  $\mathcal{P} \setminus H$  pour tout  $H \in \mathcal{H}$  ie " $H$  est vraie sous  $P$ " contre " $H$  est fausse sous  $P$ " pour tout  $H \in \mathcal{H}$ . On définit alors l'ensemble des hypothèses vraies sous  $P$  par

$$\mathcal{T}(P) := \{H \in \mathcal{H} \mid P \in H\},$$

et l'ensemble des hypothèses fausses sous  $P$  par

$$\mathcal{F}(P) := \mathcal{H} \setminus \mathcal{T}(P) = \{H \in \mathcal{H} \mid P \notin H\}.$$

Une procédure de tests multiples est la donnée d'un ensemble d'hypothèses rejetées  $\mathcal{R} \subset \mathcal{H}$ .  $\mathcal{R}$  ne dépend que de l'observation  $X$  sous  $P$ .

On va chercher à contrôler une erreur de première espèce.

On rappelle que la FWER est définie par

$$\text{FWER}(\mathcal{R}) := \sup_{P \in \mathcal{P}} P(\mathcal{R} \cap \mathcal{T}(P) \neq \emptyset).$$

On définit aussi la *weak Family Wise Error Rate* par

$$\begin{aligned} w\text{FWER} &:= \sup_{P \in \mathcal{P}, \mathcal{T}(P) = \mathcal{H}} P(\mathcal{R} \cap \mathcal{T}(P) \neq \emptyset) \\ &= \sup_{P \in \bigcap_{H \in \mathcal{H}} H} P(\mathcal{R} \cap \tau(P) \neq \emptyset). \end{aligned}$$

On notera  $\bigcap_{H \in \mathcal{H}} H$ .

On suppose connue une famille des  $p$ -valeurs  $p_H$  pour tout  $H \in \mathcal{H}$ .

### Définition 3.1.1 : Procédure min $p$

Pour tout sous-ensemble  $\mathcal{G}$  de  $\mathcal{H}$  et tout  $\alpha \in (0, 1)$ , on définit  $q_{mp, \mathcal{G}, \alpha}$  une fonction décroissante de  $\mathcal{G}$  telle que

$$\forall P \in \bigcap_{H \in \mathcal{G}} H, P(\min_{H \in \mathcal{G}} p_H < q_{mp, \mathcal{G}, \alpha}) \leq \alpha.$$

La procédure min  $p$  est alors définie par la zone de rejet

$$\forall P \in \bigcap_{H \in \mathcal{G}} H, \mathcal{R}_{mp} := \{H \in \mathcal{G} \mid p_H < q_{mp, \mathcal{G}, \alpha}\}.$$

### Exemple 3.1.2

On note  $F_0$  la fonction de répartition càdlàg et  $\bar{F}_0$  la fonction de répartition càglàd de  $\min_{H \in \mathcal{H}} p_H$  pour  $P \in \bigcap_{H \in \mathcal{H}} H$ . On a donc

$$\mathcal{R}_{mp} = \{H \mid p_H < \bar{F}_0^{-1}(\alpha)\}.$$

► *Contrôle du wFWER.*  
Soit  $P \in \bigcap_{H \in \mathcal{H}} H$ .

$$\begin{aligned} P(\mathcal{R}_{mp} \neq \emptyset) &= P(\exists H \in \mathcal{H}, p_H < \bar{F}_0^{-1}(\alpha)) \\ &= P(\min_{H \in \mathcal{H}} p_H < \bar{F}_0^{-1}(\alpha)) \\ &= \bar{F}_0(\bar{F}_0^{-1}(\alpha)) \\ &\leq \alpha. \end{aligned}$$

► *Contrôle du FWER.*

On fait une hypothèse supplémentaire :

$$(†) \forall \mathcal{G} \subset \mathcal{H} \text{ la loi de } \min_{H \in \mathcal{G}} p_H \text{ est la même } \forall P \in \bigcap_{H \in \mathcal{G}} H.$$

Cette hypothèse est par exemple vérifiée si les  $p_H$  sont indépendantes pour tout  $P \in \mathcal{P}$ . Alors, pour tout  $P \in \mathcal{P}$ ,

$$\begin{aligned} P(\mathcal{R}_{mp} \cap \mathcal{T}(P) \neq \emptyset) &= P(\exists H \in \mathcal{T}(P), p_H < \bar{F}_0^{-1}(\alpha)) \\ &= P(\min_{H \in \mathcal{T}(P)} p_H < \bar{F}_0^{-1}(\alpha)). \end{aligned}$$

On prend  $\mathcal{G} = \mathcal{T}(P)$ .

D'après (†), on peut remplacer  $P$  par  $Q \in \bigcap_{H \in \mathcal{T}(P)} H$ . Donc :

$$\begin{aligned} P(\mathcal{R}_{mp} \cap \mathcal{T}(P) \neq \emptyset) &= Q(\min_{H \in \mathcal{T}(P)} p_H < \bar{F}_0^{-1}(\alpha)) \\ &\leq Q(\min_{H \in \mathcal{H}} p_H < \bar{F}_0^{-1}(\alpha)) \end{aligned}$$

$$= \bar{F}_0(\bar{F}_0^{-1}(\alpha)) \quad \text{car sous } Q, \text{ la fonction de répartition càglàd de } \min_{H \in \mathcal{H}} p_H \text{ est } \bar{F}_0.$$

$$\leq \alpha.$$

On peut "améliorer" les procédures (ie les rendre moins conservatives) par *step-down* dont on rappelle l'algorithme. Soit  $\mathcal{N} : \mathcal{H} \rightarrow \mathcal{H}$  une application vérifiant les hypothèses suivantes :

(H1) - Monotonie :

$$\forall R \subseteq S \subset \mathcal{H}, \quad \mathcal{N}(R) \subseteq \mathcal{N}(S) \cup S,$$

(H2) - Contrôle des faux :

$$\forall P \in \mathcal{P}, \quad P(\mathcal{N}(\mathcal{F}(P)) \subseteq \mathcal{F}(P)) \geq 1 - \alpha.$$

Alors la procédure  $\mathcal{R}$  définie par

1.  $\mathcal{R}_0 = \emptyset$ ,
2.  $\mathcal{R}_{i+1} = \mathcal{R}_i \cup \mathcal{N}(\mathcal{R}_i)$ ,
3.  $\mathcal{R} := \lim_{i \rightarrow +\infty} \mathcal{R}_i$ ;

contrôle la FWER.

### Remarque 3.1.3

On pourrait définir la procédure *min p* comme la procédure itérative de rejet associée à l'application  $\mathcal{N}$  définie par

$$\mathcal{N}_{mp} : \mathcal{S} \mapsto \{H \in \mathcal{H} \setminus \mathcal{S} : p_H < q_{mp, \mathcal{H} \setminus \mathcal{S}, \alpha}\}.$$

Cette application vérifie (H1) et (H2) donc la procédure a une FWER contrôlée par  $\alpha$ .

En effet, (H1) est vérifiée par monotonie de  $\mathcal{G} \mapsto q_{mp, \mathcal{G}, \alpha}$ . Montrons (H2).

On a, pour tout  $P \in \mathcal{P}$ ,

$$\mathcal{N}(\mathcal{F}(P)) = \{H \in \mathcal{T}(P) : p_H < q_{mp, \mathcal{T}(P), \alpha}\}.$$

Alors, pour tout  $P \in \mathcal{P}$ ,

$$\begin{aligned} P(\mathcal{N}(\mathcal{F}(P)) \cap \mathcal{T}(P) \neq \emptyset) &= P(\exists H \in \mathcal{T}(P) : p_H < q_{mp, \mathcal{T}(P), \alpha}) \\ &= P\left(\min_{H \in \mathcal{T}(P)} p_H < q_{mp, \mathcal{T}(P), \alpha}\right) \\ &< \alpha \quad \text{d'après la définition 3.1.1.} \end{aligned}$$

La dernière inégalité est vraie car si  $P \in \mathcal{P}$  alors  $P \in \cap \mathcal{T}(P)$  (si on suppose que  $P \notin \mathcal{P}$  alors il existe  $H \in \mathcal{T}(P)$  telle que  $P \notin H$  ce qui est absurde).

### Exemple 3.1.4 : Amélioration de la procédure par *step-down*

Supposons (†). Si l'on connaît  $F_{\mathcal{G}}$ , la fonction de répartition de  $\min_{H \in \mathcal{G}} p_H$  pour  $P \in \cap_{H \in \mathcal{G}} H$ , on peut poser,

$$\mathcal{N}(\mathcal{G}) := \{H : p_H < \bar{F}_{\mathcal{G}^c}^{-1}(\alpha)\} \quad \text{où } \mathcal{G}^c = \mathcal{H} \setminus \mathcal{G}.$$

► Monotonie :

Si  $\mathcal{G} \subseteq \mathcal{G}'$  alors  $\min_{H \in \mathcal{G}^c} p_H \leq \min_{H \in \mathcal{G}'^c} p_H$ .

On prend  $Q \in \cap \mathcal{H}$ . Donc

$$Q\left(\min_{H \in \mathcal{G}^c} p_H < t\right) \geq Q\left(\min_{H \in \mathcal{G}'^c} p_H < t\right)$$

et

$$\bar{F}_{\mathcal{G}^c}(t) \geq \bar{F}_{\mathcal{G}'^c}(t).$$

D'où

$$\bar{F}_{\mathcal{G}'^c}^{-1}(\alpha) \geq \bar{F}_{\mathcal{G}^c}^{-1}(\alpha).$$

Donc  $\mathcal{N}(\mathcal{G}) \subseteq \mathcal{N}(\mathcal{G}')$ .

► Contrôle pour les faux :

Soit  $P \in \mathcal{P}$ .

On a  $\mathcal{N}(\mathcal{F}(P)) = \{H : p_H < \overline{F}_{\mathcal{T}(P)}^{-1}(\alpha)\}$ .

Donc

$$\begin{aligned} P(\mathcal{N}(\mathcal{F}(P)) \cap \mathcal{T}(P) \neq \emptyset) &= P(\exists H \in \mathcal{T}(P), p_H < \overline{F}_{\mathcal{T}(P)}^{-1}(\alpha)) \\ &= P(\min_{H \in \mathcal{T}(P)} p_H < \overline{F}_{\mathcal{T}(P)}^{-1}(\alpha)) \\ &= \overline{F}_{\mathcal{T}(P)}(\overline{F}_{\mathcal{T}(P)}^{-1}(\alpha)) \leq \alpha. \end{aligned}$$

On remarque que  $\mathcal{R}_1 = \mathcal{R}_{mp}$ .

### Exemple 3.1.5 : Procédures particulières

Pour des valeurs particulières de  $q_{mp, \mathcal{G}, \alpha}$ , on obtient des procédures usuelles :

1. Procédure de Holm  $\mathcal{R}_{Holm}$  associée à l'application  $\mathcal{N}_{mp}$  avec  $q_{mp, \mathcal{G}, \alpha} = \alpha/|\mathcal{G}|$ . L'application Next pour Holm sera désormais notée  $\mathcal{N}_{Holm}$ .
2. Procédure de Bonferroni avec  $\mathcal{R}_{Bonf} := \mathcal{N}_{Holm}(\emptyset)$ .
3. Si la loi de  $\min_{H \in \mathcal{G}} p_H$  ne dépend pas de  $P$  sur  $\cap \mathcal{G}$  et est connue, on définit  $\mathcal{R}_{mp}$  comme dans les exemples précédents en prenant

$$\mathcal{N}_{mp} : \mathcal{S} \mapsto \{H \in \mathcal{H} \setminus \mathcal{S} : p_H < \overline{F}_{\mathcal{H} \setminus \mathcal{S}}^{-1}(\alpha)\}.$$

On peut remarquer que  $\mathcal{R}_{mp}$  est moins conservative que  $\mathcal{R}_{Holm}$  ie

$$\mathcal{R}_{Holm} \subset \mathcal{R}_{mp}.$$

On peut étendre les procédures  $\min p$  en des procédures  $\min p$  pondérées en définissant

$$\mathcal{N}_{wmp} : \mathcal{S} \mapsto \{H \in \mathcal{H} \setminus \mathcal{S} : p_H < w_H q_{wmp, \mathcal{H} \setminus \mathcal{S}, \alpha}\},$$

où  $(w_H)_{H \in \mathcal{H}}$  est une famille de poids positifs satisfaisant  $\sum_{H \in \mathcal{H}} w_H = 1$  et où  $q_{wmp, \mathcal{G}, \alpha}$  vérifie, pour tout  $\alpha \in (0, 1)$ ,

$$\forall P \in \cap \mathcal{G}, P(\min_{H \in \mathcal{G}} w_H^{-1} p_H < q_{wmp, \mathcal{G}, \alpha}) \leq \alpha.$$

Quand la loi de  $\min_{H \in \mathcal{G}} w_H^{-1} p_H$  ne dépend pas de  $P$  sur  $\cap \mathcal{G}$  et est connue, on peut prendre  $q_{wmp, \mathcal{G}, \alpha} = \overline{F}_{w, \mathcal{G}}^{-1}(\alpha)$ . On définit alors une zone de rejet  $\mathcal{R}_{wmp}$ .

## 3.1.2 Définitions de la puissance d'un test multiple.

Une première approche afin de définir une erreur de seconde espèce pour les tests multiples est de chercher à définir une notion de puissance par analogie avec les définitions des erreurs de première espèce vues en 2.1.3.

Intuitivement, définir une erreur de seconde espèce, c'est s'intéresser aux hypothèses de l'ensemble  $\mathcal{R}^c \cap \mathcal{F}(P)$ . Aussi, afin de définir une puissance, on peut s'intéresser aux hypothèses rejetées par la procédure qui sont dans  $\mathcal{F}(P)$ , ie aux éléments de l'ensemble  $\mathcal{R} \cap \mathcal{F}(P)$ .

Comme pour les erreurs de première espèce, il existe plusieurs définitions de puissance :

Ces deux premières définition de puissance sont définies et étudiées (théoriquement et numériquement) dans [17] et [5].

► la *puissance complète* ou *all pairs power* : c'est la probabilité de rejeter tous les vrais positifs ;

$$\forall P \in \mathcal{P}, P(\mathcal{R} \subseteq \mathcal{F}(P)).$$

► la *puissance minimale* ou *any pair power* : c'est la probabilité de rejeter au moins un vrai positif ;

$$\forall P \in \mathcal{P}, P(\mathcal{F}(P) \cap \mathcal{R} \neq \emptyset);$$

et la *r-puissance* (introduit dans [5]) : c'est la probabilité de rejeter au moins  $r$  fausses hypothèses ;

$$\forall P \in \mathcal{P}, P(|\mathcal{F}(P) \cap \mathcal{R}| \geq r).$$

► le nombre moyen de vrais positifs (introduit dans [16]) :

$$\forall P \in \mathcal{P}, \mathbb{E}[|\mathcal{F}(P) \cap \mathcal{R}|].$$

Les définitions suivantes sont définies et étudiées dans [2] et [9].

► la *True Positive Fraction* : c'est la proportion d'hypothèses fausses rejetées dans l'ensemble des hypothèses fausses ie le taux moyen de vrais positifs ;

$$\forall P \in \mathcal{P}, \text{TPF}(\mathcal{R}, P) := \frac{|\mathcal{R} \cap \mathcal{F}(P)|}{|\mathcal{F}(P)| \vee 1}.$$

Ainsi, la TPF est égale à zéro lorsque la procédure ne rejette aucune hypothèse fausse. Afin d'obtenir un test performant, on va donc chercher à maximiser cette quantité. Comme elle est aléatoire, elle ne définit pas convenablement une puissance ; deux notions de puissance peuvent découler de la TPF :

► la  *$\lambda$ -puissance* pour un paramètre  $\lambda \in (0, 1)$  fixé, c'est la queue de la distribution de la TPF,

$$\forall P \in \mathcal{P}, \mathbb{P}(\text{TPF}(\mathcal{R}, P) \geq \lambda).$$

► la *puissance moyenne*, c'est l'espérance de la TPF,

$$\forall P \in \mathcal{P}, \mathbb{E}[\text{TPF}(\mathcal{R}, P)] = \mathbb{E} \left[ \frac{|\mathcal{F}(P) \cap \mathcal{R}|}{|\mathcal{F}(P)| \vee 1} \right].$$

On remarque que dans le cas d'un test simple, ces définitions de puissance coïncident avec la notion de puissance usuelle pour les tests d'une seule hypothèse.

De plus, comme on cherche à avoir "la plus grande puissance possible", on cherchera souvent à maximiser la quantité

$$\inf_{P \in \mathcal{P}} \pi_1(\mathcal{R}, P),$$

où  $\pi_1$  désigne l'une des définitions de puissance données ci-dessus.

### Exemple 3.1.6

On veut tester les hypothèses  $H_1, H_2, H_3$  et  $H_4$  en utilisant des statistiques de test indépendantes dont les puissances de chaque test individuel respectifs sont 0.5, 0.5, 0.8 et 0.8. Alors

- *Puissance complète* =  $0.5 \times 0.5 \times 0.8 \times 0.8 = 0.16$
- *Puissance minimale* =  $1 - (1 - 0.5)(1 - 0.5)(1 - 0.8)(1 - 0.8) = 0.99$
- *Puissance moyenne* =  $(0.5 + 0.5 + 0.8 + 0.8)/4 = 0.65$

Pour un même problème, on obtient des puissances très variées selon la définition que l'on prend. En pratique, il convient donc de choisir la notion de puissance la plus adaptée au problème que l'on traite.

Enfin, dans la pratique, pour comparer plusieurs procédures et choisir une procédure "performante" ou "optimale" dans le sens où elle contrôle l'erreur de première espèce au niveau  $\alpha$  et maximise une puissance, on effectue des simulations numériques. Dans la plupart des cas, on trace l'évolution de la TPF (ou de la puissance moyenne) en fonction du nombre d'hypothèses dans plusieurs configurations : on fait varier le nombre d'hypothèses nulles vraies et on choisit différentes alternatives (d'après [2]).

### 3.1.3 Critères maximin.

Dans cette section, nous allons définir des erreurs de seconde espèce qui permettent d'étudier l'optimalité des procédures de tests multiples. Des critères d'optimalité, appelés optimalité maximin, ont été établis pour certaines procédures step-down. Cette section s'inspire des références [10] et [13].

On se donne une distance  $d$  définie sur  $\mathcal{P}$ . Pour tout  $P \in \mathcal{P}$  et tout sous-ensemble  $\mathcal{Q}$  de  $\mathcal{P}$ , on définit :

$$d(P, \mathcal{Q}) := \inf_{Q \in \mathcal{Q}} d(P, Q).$$

Pour  $P \in \mathcal{P}$ , on suppose qu'il existe  $H \in \mathcal{H}$  telle que

$$d(P, H) \geq r_H, \quad \text{où } r_H > 0.$$

Cela signifie que l'hypothèse  $H$  est "éloignée" de  $P$  et donc qu'il est peu probable que  $P$  appartienne à  $H$ . Une première idée est donc de maximiser (sur l'ensemble des procédures  $\mathcal{R}$ ) la plus petite probabilité de rejeter au moins une hypothèse, compte tenu du fait que  $d(P, H) \geq r_H$ . Autrement dit, on veut maximiser

$$\inf_{P \in \mathcal{P}: \exists H \in \mathcal{H}, d(P, H) \geq r_H} P(\mathcal{R} \neq \emptyset).$$

Plus précisément, on peut souhaiter que l'hypothèse rejetée soit une hypothèse fautive ; ie que l'on souhaite maximiser

$$\inf_{P \in \mathcal{P}: \exists H \in \mathcal{H}, d(P, H) \geq r_H} P(\mathcal{R} \cap \mathcal{F}(P) \neq \emptyset).$$

On peut aussi supposer que pour tout  $H \in \mathcal{H}$ ,

$$d(P, H) \geq r_H.$$

On demande alors de maximiser (sur l'ensemble des procédures  $\mathcal{R}$ ) la probabilité de rejeter au moins une hypothèse :

$$\inf_{P \in \mathcal{P}: \forall H \in \mathcal{H}, d(P, H) \geq r_H} P(\mathcal{R} \neq \emptyset).$$

#### Définition 3.1.7 : Critères de seconde espèce - Lehmann, Romano, Shaffer (2005)

On définit une erreur de seconde espèce pour les tests multiples par l'une des quantités suivantes

$$\pi_1^1(\mathcal{R}) := \inf_{P \in \mathcal{P}: \forall H \in \mathcal{H}, d(P, H) \geq r_H} P(\mathcal{R} \neq \emptyset);$$

$$\pi_1^2(\mathcal{R}) := \inf_{P \in \mathcal{P}: \exists H \in \mathcal{H}, d(P, H) \geq r_H} P(\mathcal{R} \cap \mathcal{F}(P) \neq \emptyset).$$

Une procédure *optimale maximin* est une procédure  $\mathcal{R}_{r_H}$  qui maximise l'un des deux critères précédents parmi toutes les procédures de tests multiples qui contrôlent la FWER au niveau  $\alpha$ .

On peut montrer qu'il n'est pas possible, a priori, de déterminer une procédure  $\mathcal{R}$  qui maximise à la fois  $\pi_1^1(\mathcal{R})$  et  $\pi_1^2(\mathcal{R})$ . On peut néanmoins prouver que si les procédures de tests multiples satisfont une hypothèse de monotonie, alors on dispose de résultats d'optimalité (au sens d'optimalité maximin) pour certaines procédures step-down ou step-up usuelles et pour une certaine alternative.

On peut aussi démontrer l'existence de procédures de tests multiples optimales maximin avec un critère de seconde espèce moins générale :

#### Définition 3.1.8 : Critère de seconde espèce - Romano, Shaikh, Wolf (2011)

On définit une erreur de seconde espèce pour les tests multiples par

$$\pi_1^3 := \inf_{P \in \mathcal{P}: \exists H \in \mathcal{H}, d(P, H) \geq r} P(\mathcal{R} \neq \emptyset);$$

$$\pi_1^4 := \inf_{P \in \mathcal{P}: \exists H \in \mathcal{H}, d(P, H) \geq r} P(\mathcal{R} \cap \mathcal{F}(P) \neq \emptyset).$$

Une procédure *optimale maximin* est une procédure  $\mathcal{R}_r$  qui maximise un des critères précédents parmi toutes les procédures de tests multiples qui contrôlent la FWER au niveau  $\alpha$ .

Ces critères d'optimalité pour les tests multiples reposent sur le lien entre tests multiples et *closure method*.

**Remarque 3.1.9 :** *Closure method - Hochberg, Tamhane (1987) [8]*

Le but de cette méthode est de faire le lien entre un problème de tests multiples contrôlant la FWER et un problème de test simple qui contrôle l'erreur de première espèce. On présente ici les idées générales de cette méthode.

On se donne une procédure de tests multiples.

Pour  $\mathcal{G} \subset \mathcal{H}$ , on considère le test simple d'hypothèse nulle  $H_{0, \mathcal{G}} := P \in \bigcap_{H \in \mathcal{G}} H$  contre l'alternative  $P \notin \bigcap_{H \in \mathcal{G}} H$ . On note  $\cap \mathcal{G} := \bigcap_{H \in \mathcal{G}} H$ .

Ce problème de test est associé au test statistique  $\phi_{\mathcal{G}} := \mathbb{1}_{\mathcal{R}_{\mathcal{G}}}$  de niveau  $\alpha$  :

$$\sup_{P \in \cap \mathcal{G}} E_P[\phi_{\mathcal{G}}] = \sup_{P \in \cap \mathcal{G}} P(\mathcal{R}_{\mathcal{G}}) \leq \alpha.$$

La closure method consiste à rejeter l'hypothèse  $H$  ssi  $H_{0, \mathcal{G}}$  est rejetée par le test  $\phi_{\mathcal{G}}$  pour tout  $\mathcal{G} \subseteq \mathcal{H}$  tel que  $H \in \mathcal{G}$ .

*Remarques :*

► L'évènement A : "au moins une des hypothèses  $H \in \mathcal{T}(P)$  est rejetée par la closure method" implique l'évènement B : " $H_{0, \mathcal{T}(P)}$  est rejetée par le test de niveau  $\alpha$   $\phi_{\mathcal{T}(P)}$ "; on obtient donc un contrôle de l'erreur de première espèce pour le test multiple à partir du contrôle de l'erreur de première espèce du test simple car

$$\text{FWER}(\mathcal{R}) = P(A) \leq P(B) \leq \alpha.$$

► On peut montrer que toute procédure de tests multiples est équivalente à la donnée d'un test simple par closure method.

On dispose alors du théorème suivant qui fournit la procédure optimale maximin pour un problème de tests multiples à l'aide de la closure method pour l'erreur  $\pi_1^3$ .

**Théorème 3.1.10 :** *Romano, Shaikh, Wolf (2011)*

Soit  $X \sim P \in \mathcal{P}$  et le problème de tests multiples d'hypothèses nulles  $P \in H$  contre  $P \notin H$  pour tout  $H \in \mathcal{H}$ .

On s'intéresse au problème de test simple d'hypothèse nulle  $H_{0, \mathcal{H}}$ . On considère alors le test de niveau  $\alpha$   $\phi_{\mathcal{H}} := \mathbb{1}_{\mathcal{R}_{\mathcal{H}}}$  qui maximise la puissance minimale sous l'alternative  $\mathcal{Q} \subset (\cap \mathcal{H})^c$ , ie qui maximise

$$\min_{P \in \mathcal{Q}} P(\mathcal{R}_{\mathcal{H}} \neq \emptyset).$$

On suppose de plus que la procédure de tests multiples  $\mathcal{R}$  associée à cette procédure de test simple par closure method est *consonant*, c'est-à-dire que si l'hypothèse  $H$  n'est pas rejetée alors toute hypothèse  $H'$  vérifiant  $H \subset H'$  n'est pas rejetée par la procédure  $\mathcal{R}$ .

Alors cette procédure de tests multiples maximise

$$\inf_{P \in \mathcal{Q}} P(\mathcal{R} \neq \emptyset)$$

parmi toutes les procédures de tests multiples qui contrôlent la FWER au niveau  $\alpha$ .

On remarque que  $P \in (\cap \mathcal{H})^c$  équivaut à dire qu'il existe  $H \in \mathcal{H}$  telle que  $P \in H^c$ . Avec nos notations, cela revient à dire que  $d(P, H) \geq r$  pour  $r$  suffisamment grand. Le choix de  $r$  intervient pour assurer la *consonance* de la procédure  $\mathcal{R}$ .

## 3.2 Correspondance entre les tests agrégés et les tests multiples.

Dans cette section, nous allons reprendre l'idée développée dans la section précédente, c'est-à-dire que nous allons établir un parallèle entre une procédure de tests multiples qui contrôle la FWER et une procédure de test simple qui contrôle l'erreur de première espèce ainsi qu'une erreur de seconde espèce. L'objectif est alors de définir une erreur de seconde espèce et démontrer des critères d'optimalité pour la procédure de tests multiples à partir du test simple.

Cette section est basée sur l'article de [6].

### 3.2.1 Tests agrégés, erreur de première espèce et théorie minimax.

Soit  $(\mathbb{X}, \mathcal{X}, \mathcal{P})$  un modèle statistique.

On considère le problème de test simple d'hypothèse nulle  $H_0$  contre l'alternative  $H_1 := \mathcal{P} \setminus H_0$ .

On se donne, en général, une collection finie d'hypothèses  $\mathcal{H}$  telle que  $H_0 \subset \cap \mathcal{H}$ . Pour chaque hypothèse  $H$  de la collection  $\mathcal{H}$ , on construit un test  $\phi_H$  d'hypothèse nulle  $H$  contre l'alternative  $\mathcal{P} \setminus H$ . On obtient alors une collection de tests notée  $\Phi_{\mathcal{H}} = \{\phi_H, H \in \mathcal{H}\}$ . L'idée est de rejeter l'hypothèse nulle  $H_0$  si au moins un des tests de la collection  $\Phi_{\mathcal{H}}$  vaut 1 ie au moins une des hypothèses  $H \in \mathcal{H}$  est rejetée par  $\phi_H$ . On définit alors le test agrégé par :

$$\bar{\Phi}_{\mathcal{H}} := \sup_{H \in \mathcal{H}} \phi_H.$$

#### Exemple 3.2.1 : Cas Gaussien

On observe  $X = (X_1, \dots, X_n)'$  un vecteur gaussien.

On suppose que  $X \sim P_f$  avec  $P_f = N(f, \sigma^2 I_n)$  où l'espérance  $f \in \mathbb{R}^n$  est inconnue et la variance  $\sigma^2 > 0$  est connue.

On note  $f = (f_1, \dots, f_n)$  et  $\mathcal{P} := \{P_f, f \in \mathbb{R}^n\}$ .

On veut tester  $H_0 = \{P_0\}$  contre  $H_1 = \mathcal{P} \setminus \{P_0\}$ , ie  $H_0 : "f = 0"$  contre  $H_1 = "f \neq 0"$ .

► On se donne une collection finie  $\mathcal{S}$  de sev  $S \subset \mathbb{R}^n$ . Pour tout  $S \in \mathcal{S}$ , on note  $\Pi_S$  la projection orthogonale sur  $S$  et on construit un test  $\phi_S$  de  $H_S$  contre  $\mathcal{P} \setminus H_S$  par

$$\phi_S = \begin{cases} 1 & \text{si } \|\Pi_S(X)\| \text{ est grande,} \\ 0 & \text{sinon ;} \end{cases}$$

où

$$H_S := \{P_f, \Pi_S f = 0\} = \{P_f, f \in S^\perp\}.$$

► A partir d'une collection finie d'hypothèses  $\mathcal{H} := \{H_S, S \in \mathcal{S}\}$  vérifiant  $H_0 \subseteq \cap \mathcal{H}$ , le test agrégé d'hypothèse nulle  $H_0$  contre  $H_1$ , basé sur la collection de tests  $\Phi_{\mathcal{H}} = \{\phi_S, S \in \mathcal{S}\}$  est définie par

$$\bar{\Phi}_{\mathcal{H}} := \sup_{H_S \in \mathcal{H}} \phi_S.$$

#### Erreur de première espèce.

L'erreur de première espèce d'un test agrégé  $\bar{\Phi}_{\mathcal{H}}$  d'hypothèse nulle  $H_0$  est définie par

$$\text{ER}(\bar{\Phi}_{\mathcal{H}}, H_0) := \sup_{P \in H_0} P(\bar{\Phi}_{\mathcal{H}} = 1) = \sup_{P \in H_0} P(\sup_{H \in \mathcal{H}} \phi_H = 1).$$

On fixe un niveau  $\alpha \in (0, 1)$ .

Pour tout  $H \in \mathcal{H}$ , le test individuel  $\phi_H$  est défini à partir d'une statistique  $S_H$  dont la loi ne dépend pas de  $P$

lorsque  $P \in H_0$ . On rappelle que l'on note  $F_H$  et  $\bar{F}_H$  les fonctions de répartition càdlàg et càglàd de la loi de  $S_H$  sous  $H_0$ . Plus précisément, on a,

$$\phi_H = \mathbb{1}_{\{S_H > \bar{F}_H^{-1}(1-u_{H,\alpha})\}}$$

où  $u_{H,\alpha}$  est choisi de telle sorte que le test agrégé soit de niveau  $\alpha$  ie,

$$\text{ER}(\bar{\Phi}_{\mathcal{H}}, H_0) \leq \alpha.$$

Plusieurs choix de  $u_{H,\alpha}$  sont alors possibles :

► *Test de Bonferroni.*

Avec  $u_{H,\alpha} = \alpha/N$  où  $N = |\mathcal{H}|$  est le nombre d'hypothèses dans  $\mathcal{H}$ , on obtient le test agrégé de Bonferroni  $\bar{\Phi}_{\mathcal{H}}^{Bonf}$  basé sur la collection

$$\Phi_{\mathcal{H}}^{Bonf} = \{\phi_H^{Bonf} = \mathbb{1}_{\{S_H > \bar{F}_H^{-1}(1-\alpha/N)\}}, H \in \mathcal{H}\}.$$

► *Test de Bonferroni pondéré.*

Avec  $u_{H,\alpha} = w_H \alpha$  où  $(w_H)_{H \in \mathcal{H}}$  est une famille de poids positifs telle que  $\sum_{H \in \mathcal{H}} w_H \leq 1$ , on obtient le test agrégé de Bonferroni pondéré  $\bar{\Phi}_{\mathcal{H}}^{wBonf}$  basé sur la collection

$$\Phi_{\mathcal{H}}^{wBonf} = \{\phi_H^{wBonf} = \mathbb{1}_{\{S_H > \bar{F}_H^{-1}(1-w_H \alpha)\}}, H \in \mathcal{H}\}.$$

► *Test proposé par Baraud, Huet et Laurent.*

Un choix moins conservatif en pratique garantissant un contrôle au niveau  $\alpha$  est  $u_{H,\alpha} = w_H u_{\alpha}$  où

$$u_{\alpha} = \sup\{u \mid \sup_{P \in H_0} P\left(\exists H \in \mathcal{H} \setminus S_H > \bar{F}_H^{-1}(1-w_H u)\right) \leq \alpha\}.$$

Avec  $w_H = 1/N$ , on obtient le test agrégé  $\bar{\Phi}_{\mathcal{H}}^{wBHL}$  basé sur la collection

$$\Phi_{\mathcal{H}}^{wBHL} = \{\phi_H^{wBHL} = \mathbb{1}_{\{S_H > \bar{F}_H^{-1}(1-w_H u_{\alpha})\}}, H \in \mathcal{H}\}.$$

*Démonstration.*

On va montrer que ce choix de  $u_{H,\alpha}$  permet de contrôler l'erreur de première espèce au niveau  $\alpha$ .

Pour tout  $P \in H_0$ , on note  $\bar{G}_{H,P}$  la fonction de répartition de  $\min_{H \in \mathcal{H}} \frac{1 - \bar{F}_H(S_H)}{w_H}$  et  $\bar{G}_H := \sup_{P \in H_0} \bar{G}_{H,P}$ . On a

$$\begin{aligned} u_{\alpha} &= \sup\{u \mid \sup_{P \in H_0} P\left(\exists H \in \mathcal{H} \setminus S_H > \bar{F}_H^{-1}(1-w_H u)\right) \leq \alpha\} \\ &= \sup\{u \mid \sup_{P \in H_0} P\left(\exists H \in \mathcal{H} \setminus \bar{F}_H(S_H) > 1-w_H u\right) \leq \alpha\} \\ &= \sup\{u \mid \sup_{P \in H_0} P\left(\min_{H \in \mathcal{H}} \frac{1 - \bar{F}_H(S_H)}{w_H} < u\right) \leq \alpha\} \\ &= \sup\{u \mid \sup_{P \in H_0} \bar{G}_{H,P}(u) \leq \alpha\} \\ &= \sup\{u \mid \bar{G}_H(u) \leq \alpha\} \\ &= \bar{G}_H^{-1}(\alpha). \end{aligned}$$

Ainsi,

$$\begin{aligned} \text{ER}(\bar{\Phi}_{\mathcal{H}}, H_0) &= \sup_{P \in H_0} P(\exists H \in \mathcal{H}, S_H > \bar{F}_H^{-1}(1-u_{H,\alpha})) \\ &= \sup_{P \in H_0} P(\exists H \in \mathcal{H}, \bar{F}_H(S_H) > 1-w_H u_{\alpha}) \\ &= \sup_{P \in H_0} P\left(\min_{H \in \mathcal{H}} \frac{1 - \bar{F}_H(S_H)}{w_H} < u_{\alpha}\right) \\ &= \sup_{P \in H_0} \bar{G}_{H,P}(u_{\alpha}) \\ &= \bar{G}_H(u_{\alpha}) \\ &= \bar{G}_H(\bar{G}_H^{-1}(\alpha)) \leq \alpha. \end{aligned}$$

□

**Erreur de seconde espèce.**

Définissons à présent l'erreur de seconde espèce pour les tests agrégés.

Soit  $d$  une distance définie sur  $\mathcal{P}$ . Pour tout  $P \in \mathcal{P}$  et tout sous-ensemble  $\mathcal{Q}$  de  $\mathcal{P}$ , on définit :

$$d(P, \mathcal{Q}) := \inf_{Q \in \mathcal{Q}} d(P, Q).$$

Comme  $\mathcal{P}$  peut être "trop grand" pour définir convenablement la notion de vitesse de séparation, on va considérer un sous-ensemble  $\mathcal{Q}$  de  $\mathcal{P}$ .

**Définition 3.2.2 :** *Vitesse de séparation uniforme - Fromont, Lerasle, Reynaud-Bouret (2015) [6]*

Soit  $\alpha, \beta \in (0, 1)$ , un sous-ensemble de lois de probabilité  $\mathcal{Q}$  de  $\mathcal{P}$  et un test  $\bar{\Phi}$  d'hypothèse nulle  $H_0$  de niveau  $\alpha$ .

La *vitesse de séparation uniforme* de  $\bar{\Phi}$  sous  $\mathcal{Q}$  avec une erreur de seconde espèce fixée  $\beta$  est définie par

$$\text{SR}_d^\beta(\bar{\Phi}, \mathcal{Q}, H_0) = \inf\{r > 0 \mid \sup_{P \in \mathcal{Q}, d(P, H_0) \geq r} P(\bar{\Phi} = 0) \leq \beta\}.$$

La vitesse de séparation uniforme  $\text{SR}_d^\beta(\bar{\Phi}, \mathcal{Q}, H_0)$  est donc définie comme étant le plus petit réel positif  $r$  tel que l'erreur de seconde espèce du test soit au plus égale à  $\beta$  pour toutes les alternatives  $P \in \mathcal{Q}$  à distance  $r$  de  $H_0$ . Cette vitesse de séparation s'interprète comme la plus petite distance entre les alternatives  $P \in \mathcal{Q} \subseteq \mathcal{P}$  et  $H_0$  pour que les hypothèses  $(H_0)$  et  $(H_1)$  soit suffisamment "séparées", distinguables par le test.

On remarque que lorsque  $H_0 \subset \cap \mathcal{H}$  alors

$$\text{SR}_d^\beta(\bar{\Phi}, \mathcal{Q}, H_0) \geq \text{SR}_d^\beta(\bar{\Phi}, \mathcal{Q}, \cap \mathcal{H}).$$

On peut donc développer une théorie minimax pour les tests agrégés.

**Définition 3.2.3 :** *Vitesse de séparation minimax, test minimax et test adaptatif - Fromont, Lerasle, Reynaud-Bouret (2015) [6]*

1. La *vitesse de séparation minimax* sous  $\mathcal{Q}$  de niveau  $\alpha$  et d'erreur de seconde espèce  $\beta$  fixés est défini par :

$$m\text{SR}_d^{\alpha, \beta}(\mathcal{Q}, H_0) := \inf_{\bar{\Phi}, \text{ER}(\bar{\Phi}, H_0) \leq \alpha} \text{SR}_d^\beta(\bar{\Phi}, \mathcal{Q}, H_0),$$

où l'infimum est pris sur tous les tests possibles de niveau  $\alpha$ .

2. Un test  $\bar{\Phi}$  de niveau  $\alpha$  est dit *minimax* sous  $\mathcal{Q}$  si  $\text{SR}_d^\beta(\bar{\Phi}, \mathcal{Q}, H_0)$  est égale à  $m\text{SR}_d^{\alpha, \beta}(\mathcal{Q}, H_0)$  à une constante (dépendant de  $\alpha$  et  $\beta$ ) multiplicative près.
3. Un test  $\bar{\Phi}$  de niveau  $\alpha$  est dit *adaptatif au sens minimax* sous une collection de classes  $\mathcal{Q}$ , si  $\text{SR}_d^\beta(\bar{\Phi}, \mathcal{Q}, H_0)$  atteint (à une constante mutiplicative près)  $m\text{SR}_d^{\alpha, \beta}(\mathcal{Q}, H_0)$  simultanément sur toutes les classes  $\mathcal{Q}$  de la collection, sans savoir a priori dans quelle classe de la collection appartient la loi  $P$ .

La vitesse de séparation minimax correspond à la plus petite distance  $r > 0$  pour laquelle il existe un test  $\bar{\Phi}$  de niveau  $\alpha$  dont l'erreur de seconde espèce est plus petite que  $\beta$  sur une classe d'alternative  $\mathcal{Q}$ . Le test minimax est celui qui atteint la borne minimax sur  $\mathcal{Q}$ . On dit qu'il est adaptatif s'il atteint la borne minimax sur plusieurs classes d'alternatives simultanément.

**3.2.2 Correspondance entre la théorie des tests multiples et celle des tests agrégés.**

On suppose que l'on dispose d'une collection finie  $\mathcal{H}$  d'hypothèses  $H$  et une hypothèse nulle simple  $H_0$  telle que  $H_0 \subseteq \cap \mathcal{H}$ .

On a la correspondance suivante :

$$\begin{aligned} \text{Test agrégé de } H_0 &\longleftrightarrow \text{Test multiple de } \mathcal{H} \\ \Phi_{\mathcal{H}} = \{\phi_H, H \in \mathcal{H}\} &\longrightarrow \mathcal{R}(\Phi_{\mathcal{H}}) = \{H \in \mathcal{H} \mid \phi_H = 1\} \\ \bar{\Phi}(\mathcal{R}) = \mathbb{1}_{\{\mathcal{R} \neq \emptyset\}} &\longleftarrow \mathcal{R} \subset \mathcal{H} \end{aligned}$$

**Exemple 3.2.4 : Cas gaussien**

La procédure de tests multiples associée au test agrégé définie précédemment dans 3.2.1 est

$$\mathcal{R}(\Phi_{\mathcal{H}}) = \{H_S \in \mathcal{H}, \phi_S = 1\}.$$

### 3.3 Théorie minimax pour les tests multiples.

#### 3.3.1 Erreur de première espèce et premières identifications.

**Remarque 3.3.1 : weak Family Wise Error Rate et erreur de première espèce**

On définit la *weak Family Wise Error Rate* (*wFWER*) de  $\mathcal{R}(\Phi_{\mathcal{H}})$  par

$$w\text{FWER}(\mathcal{R}(\Phi_{\mathcal{H}})) = \sup_{P, \mathcal{T}(P)=\mathcal{H}} P(\mathcal{R}(\Phi_{\mathcal{H}}) \cap \mathcal{T}(P) \neq \emptyset).$$

On a

$$w\text{FWER}(\mathcal{R}(\Phi_{\mathcal{H}})) = \sup_{P \in \cap \mathcal{H}} P(\mathcal{R}(\Phi_{\mathcal{H}}) \neq \emptyset) = \sup_{P \in \cap \mathcal{H}} P(\bar{\Phi}_{\mathcal{H}} = 1) = \text{ER}(\bar{\Phi}_{\mathcal{H}}, \cap \mathcal{H}).$$

Comme  $H_0 \subset \cap \mathcal{H}$ ,

$$w\text{FWER}(\mathcal{R}(\Phi_{\mathcal{H}})) \geq \text{ER}(\bar{\Phi}_{\mathcal{H}}, H_0).$$

Il est donc plus difficile de contrôler  $w\text{FWER}(\mathcal{R}(\Phi_{\mathcal{H}}))$  que  $\text{ER}(\bar{\Phi}_{\mathcal{H}}, H_0)$  (sauf dans le cas où  $H_0 = \cap \mathcal{H}$ ).

On suppose désormais que pour tout  $H \in \mathcal{H}$ , on dispose d'une statistique de test  $S_H$ , dont la loi ne dépend pas de  $P$  lorsque  $P \in H$ . On note  $p_H$  la  $p$ -valeur correspondante.

**Proposition 3.3.2 : Fromont, Lerasle, Reynaud-Bouret (2015) [6]**

On a :

$$\mathcal{R}(\Phi_{\mathcal{H}}^{\text{Bonf}}) = \mathcal{R}_{\text{Bonf}},$$

et

$$\bar{\Phi}_{\mathcal{H}}^{\text{Bonf}} = \bar{\Phi}(\mathcal{R}_{\text{Bonf}}) = \bar{\Phi}(\mathcal{R}_{\text{Holm}}).$$

Si de plus, on suppose que la loi de  $\min_{H \in \mathcal{H}} w_H^{-1} p_H$  ne dépend pas de  $P$  lorsque  $P$  appartient à  $\cap \mathcal{H}$ , alors

$$\mathcal{N}_{wmp}(\emptyset) = \mathcal{R}(\Phi_{\mathcal{H}}^{w\text{BHL}}) \text{ et } \bar{\Phi}_{\mathcal{H}}^{w\text{BHL}} = \bar{\Phi}(\mathcal{R}_{wmp}).$$

*Démonstration.*

On rappelle que l'on a :

$$S_H > \bar{F}_H^{-1}(1 - \alpha) \iff p(S_H) < \alpha.$$

On a tout d'abord, en notant  $N = |\mathcal{H}|$ ,

$$R(\Phi_{\mathcal{H}}^{Bonf}) = \{H \in \mathcal{H} : \phi_H^{Bonf} = 1\} = \{H : S_H > \bar{F}_H^{-1}(1 - \alpha/N)\} = \{H : p(S_H) < \alpha/N\} = \mathcal{N}_{Holm}(\emptyset) = \mathcal{R}_{Bonf}.$$

De plus,

$$\bar{\Phi}_{\mathcal{H}}^{Bonf} = \sup_{H \in \mathcal{H}} \{\mathbb{1}\{S_H > \bar{F}_H^{-1}(1 - \alpha/N)\}\} = \mathbb{1}\{\exists H \in \mathcal{H} : p(S_H) < \alpha/N\} = \bar{\Phi}(\mathcal{R}_{Bonf}) = \bar{\Phi}(\mathcal{R}_{Holm}).$$

Supposons que la loi de  $\min_{H \in \mathcal{H}} w_H^{-1} p_H$  ne dépend pas de  $P$  lorsque  $P$  appartient à  $\cap \mathcal{H}$ . On note  $F$  la fonction de répartition de  $\min_{H \in \mathcal{H}} w_H^{-1} p_H$ .

On a :

$$\begin{aligned} P(\exists H \in \mathcal{H} : S_H > \bar{F}_H^{-1}(1 - w_H u)) &= P(\exists H \in \mathcal{H} : w_H^{-1} p_H < u) \\ &= P(\min_{H \in \mathcal{H}} w_H^{-1} p_H < u) \\ &= \bar{F}(u) \end{aligned}$$

De plus, par définition de  $u_\alpha$ , on a  $u_\alpha = \bar{F}^{-1}(\alpha)$ . Enfin,

$$\mathcal{R}(\Phi_{\mathcal{H}}^{wBHL}) = \{H : S_H > \bar{F}_H^{-1}(1 - w_H u_\alpha)\} = \{H : p_H < w_H u_\alpha\} = \{H : p_H < w_H \bar{F}^{-1}(\alpha)\} = \mathcal{N}_{wmp}(\emptyset);$$

et

$$\bar{\Phi}_{\mathcal{H}}^{wBHL} = \mathbb{1}\{\exists H : p_H < w_H u_\alpha\} = \mathbb{1}\{\exists H : w_H^{-1} p_H < \bar{F}^{-1}(\alpha)\} = \mathbb{1}\{\min_H w_H^{-1} p_H < \bar{F}^{-1}(\alpha)\} = \bar{\Phi}(\mathcal{R}_{wmp}).$$

□

### 3.3.2 Family Wise Separation Rates pour les tests multiples.

On considère à présent une procédure de tests multiples  $\mathcal{R}$  et le test agrégé correspondant  $\bar{\Phi}(\mathcal{R})$  pour  $H_0 \subset \cap \mathcal{H}$ . Soit  $\beta \in (0, 1)$ . On se donne une classe  $\mathcal{Q} \subset \mathcal{P}$ . D'après 3.2.2, la vitesse de séparation uniforme de  $\bar{\Phi}(\mathcal{R})$  sous  $\mathcal{Q}$  avec l'erreur de seconde espèce  $\beta$  et la distance  $d$  fixées est définie par

$$SR_d^\beta(\bar{\Phi}(\mathcal{R}), \mathcal{Q}, \cap \mathcal{H}) = \inf\{r > 0 : \sup_{P \in \mathcal{Q}, d(P, \cap \mathcal{H}) \geq r} P(\mathcal{R} = \emptyset) \leq \beta\}.$$

En s'inspirant de la définition de la weak Family Wise Error Rate de  $\mathcal{R}$  (qui est égale à  $ER(\bar{\Phi}_{\mathcal{H}}, H_0)$  d'après la remarque 3.3.1), on aimerait définir la notion de weak Family Wise Separation Rate de  $\mathcal{R}$  comme étant  $SR_d^\beta(\bar{\Phi}(\mathcal{R}), \mathcal{Q}, \cap \mathcal{H})$ . Cependant, cette définition ne serait pas satisfaisante car elle ne pourrait pas s'étendre en une notion forte de FWSR et donc ne permettrait pas d'introduire une théorie minimax. En effet, dans la théorie des tests agrégés, on veut rejeter  $\cap \mathcal{H}$ , s'il existe au moins un  $H \in \mathcal{H}$  telle que  $d(P, H) \geq r$ .

On définit alors l'ensemble des hypothèses fausses à distance  $r$  de  $P$  par

$$\mathcal{F}_r(P) := \{H \in \mathcal{H} : d(P, H) \geq r\}.$$

On introduit alors la notion suivante :

**Définition 3.3.3 :** *weak Family Wise Separation Rate - Fromont, Lerasle, Reynaud-Bouret (2015) [6]*

Soit  $\beta \in (0, 1)$  et une classe de lois de probabilité  $\mathcal{Q} \subset \mathcal{P}$ . La *weak Family Wise Separation Rate* d'une procédure de tests multiples  $\mathcal{R}$  sous  $\mathcal{Q}$  d'erreur de seconde espèce  $\beta$  est définie par

$$wFWSR_d^\beta(\mathcal{R}, \mathcal{Q}) := \inf\{r > 0 : \sup_{P \in \mathcal{Q} : \mathcal{F}_r(P) \neq \emptyset} P(\mathcal{R} = \emptyset) \leq \beta\}.$$

**Remarque 3.3.4**

Cette définition semble rappeler le critère maximin  $\pi_1^3(r) = \inf_{P \in \mathcal{P}: \exists H \in \mathcal{H}, d(P, H) \geq r} P(\mathcal{R} \neq \emptyset)$  défini par Romano, Shaikh et Wolf (2011) (3.1.8). En effet, Romano, Shaikh et Wolf cherchent à maximiser "la puissance"  $\pi_1^3$  tandis que Fromont, Lerasle et Reynaud-Bouret veulent minimiser la  $w\text{FWSR}_d^\beta(\mathcal{R}, \mathcal{Q})$  (quantité liée à l'erreur de seconde espèce) sur l'ensemble des procédures qui contrôlent l'erreur de première espèce. Cependant, l'approche entre ces deux critères est différente : pour Romano, Shaikh et Wolf, le paramètre  $r$  est fixé au préalable et on cherche la "meilleure" erreur de seconde espèce tandis que pour Fromont, Lerasle et Reynaud-Bouret, l'erreur de seconde espèce  $\beta$  est fixée et on cherche le "meilleur"  $r > 0$  assurant le contrôle  $\beta(r) := \sup_{P \in \mathcal{Q}: \exists H \in \mathcal{H}, d(P, H) \geq r} P(\mathcal{R} = \emptyset) \leq \beta$ .

On en déduit alors le résultat suivant :

**Proposition 3.3.5** : Fromont, Lerasle, Reynaud-Bouret (2015) [6]

Pour tout sous-ensemble  $\mathcal{Q}$  de  $\mathcal{P}$  et  $\beta \in (0, 1)$ ,

$$w\text{FWSR}_d^\beta(R, \mathcal{Q}) \leq \text{SR}_d^\beta(\bar{\Phi}(R), \mathcal{Q}, \cap \mathcal{H}).$$

De plus, si la collection d'hypothèses  $\mathcal{H}$  et la distance  $d$  vérifient

$$(*) \quad \forall r > 0, \quad d(P, \cap \mathcal{H}) \geq r \iff \mathcal{F}_r(P) \neq \emptyset,$$

alors, pour tout  $\beta \in (0, 1)$ , et toute classe  $\mathcal{Q} \subset \mathcal{P}$ ,

$$w\text{FWSR}_d^\beta(R, \mathcal{Q}) = \text{SR}_d^\beta(\bar{\Phi}(R), \mathcal{Q}, \cap \mathcal{H}).$$

*Démonstration.*

Soit  $\mathcal{Q} \subset \mathcal{P}$ . Soit  $P \in \mathcal{Q}$ .

Si  $\mathcal{F}_r(P) \neq \emptyset$  alors il existe  $H \in \mathcal{H}$  tel que  $d(P, H) \geq r$ , autrement dit, pour tout  $Q \in H$ ,  $d(P, Q) \geq r$ .

En particulier, en prenant  $Q \in \cap \mathcal{H} \subset H$ , on obtient  $d(P, \cap \mathcal{H}) \geq r$ .

Donc

$$\sup_{P \in \mathcal{Q}, \mathcal{F}_r(P) \neq \emptyset} P(\mathcal{R} = \emptyset) \leq \sup_{P \in \mathcal{Q}, d(P, \cap \mathcal{H}) \geq r} P(\mathcal{R} = \emptyset).$$

D'où

$$\{r > 0, \sup_{P \in \mathcal{Q}, d(P, \cap \mathcal{H}) \geq r} P(\mathcal{R} = \emptyset) \leq \beta\} \subset \{r > 0, \sup_{P \in \mathcal{Q}, \mathcal{F}_r(P) \neq \emptyset} P(\mathcal{R} = \emptyset) \leq \beta\}.$$

Si la condition (\*) est vérifiée, les deux ensembles sont égaux et on obtient l'égalité recherchée.  $\square$

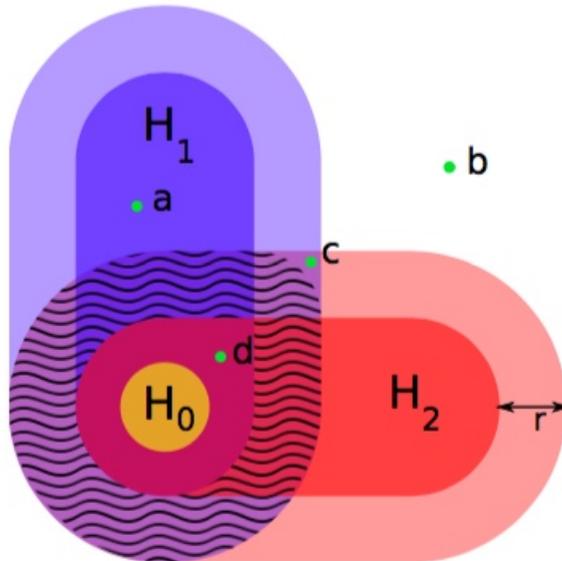


FIGURE 3.1 – Visualisation d'un problème de tests multiples avec deux hypothèses  $H_1$  et  $H_2$  (représentées en bleu et rouge foncées). Les hypothèses situées à une distance au plus  $r$  de  $H_1$  et  $H_2$  sont de couleur plus claires et celles situées à une distance au plus  $r$  de  $H_1 \cap H_2$  sont hachurées. L'hypothèse  $H_0$  est strictement incluse dans  $H_1 \cap H_2$ . Les points  $a, b, c$  et  $d$  correspondent à des lois de probabilités.

### Remarque 3.3.6

1. Description des probabilités  $P$  correspondant aux points  $a, b, c$  et  $d$  de la figure :

- ▶ Point  $a$  :  $\mathcal{T}(P) = \emptyset$  et  $\mathcal{F}(P) = \mathcal{F}_r(P) = \{H_2\}$ .
- ▶ Point  $b$  :  $\mathcal{T}(P) = \emptyset$  et  $\mathcal{F}(P) = \mathcal{F}_r(P) = \{H_1, H_2\}$ .
- ▶ Point  $c$  :  $\mathcal{T}(P) = \emptyset$ ,  $\mathcal{F}(P) = \{H_1, H_2\}$  et  $\mathcal{F}_r(P) = \emptyset$  mais  $d(P, H_1 \cap H_2) \geq r$ .
- ▶ Point  $d$  :  $\mathcal{T}(P) = \{H_1, H_2\}$  et  $\mathcal{F}(P) = \mathcal{F}_r(P) = \emptyset$  mais  $P \in H_0$ .

2. Il est plus difficile de contrôler  $\text{SR}_d^\beta(\overline{\Phi}(R), \mathcal{Q}, \cap \mathcal{H})$  que  $w\text{FWSR}_d^\beta(R, \mathcal{Q})$ .

3. Si la collection d'hypothèses  $\mathcal{H}$  est fermée, ie

$$H \in \mathcal{H} \text{ et } H' \in \mathcal{H} \implies H \cap H' \in \mathcal{H},$$

alors la condition (\*) est vérifiée.

En effet, supposons que  $H, H' \in \mathcal{H} \implies H \cap H' \in \mathcal{H}$ . Soit  $r > 0$ .

Si  $d(P, \cap \mathcal{H}) \geq r$  alors  $\mathcal{F}_r(P) = \{H \in \mathcal{H} : d(H, P) \geq r\} \neq \emptyset$  car  $\cap \mathcal{H} \in \mathcal{H}$ .

S'il existe  $H \in \mathcal{H}$  telle que  $d(H, P) \geq r$  alors  $d(P, \cap \mathcal{H}) = \inf_{H' \in \cap \mathcal{H}} d(P, H') \geq d(P, H) \geq r$  car  $\cap \mathcal{H} \subset H$  (cette implication est toujours vraie).

### Exemple 3.3.7 : Cas Gaussien

On suppose ici que  $\mathcal{H} = \{H_i, 1 \leq i \leq n\}$  où  $H_i = \{P_f, f_i = 0\}$  pour tout  $i \in \{1, \dots, n\}$ .

▶ On considère la distance  $d = d_\infty$  définie par

$$d(P_f, P_g) = \|f - g\|_\infty = \max_{i \in \{1, \dots, n\}} |f_i - g_i|.$$

Alors la condition (\*) de 3.3.5 est vérifiée.

En effet, soit  $r > 0$  et supposons que  $d(P_f, \cap \mathcal{H}) \geq r$ . On remarque que  $P_g \in \cap \mathcal{H}$  est équivalent à  $g = 0$ . Ainsi, si on note  $f_j := \max_{i \in \{1, \dots, n\}} |f_i|$  alors  $|f_j| \geq r$ . On a donc que  $d(P_f, H_j) \geq r$  et donc  $\mathcal{F}_r(P) \neq \emptyset$ .

*Conclusion* : Pour tout  $\beta \in (0, 1)$ ,  $w\text{FWSR}_{d_\infty}^\beta(R(\Phi_{\mathcal{H}}), \mathcal{P}) = \text{SR}_d^\beta(\overline{\Phi}_{\mathcal{H}}, \mathcal{P}, \cap \mathcal{H})$ .

▶ On considère la distance  $d = d_2$  définie par

$$d_2(P_f, P_g) = \left( \sum_{i=1}^n |f_i - g_i|^2 \right)^{1/2}.$$

Alors la condition (\*) n'est pas vérifiée (cf le point  $c$  de la figure 1 ci-dessus représentée avec  $d_2$ ).

Nous pouvons à présent définir une notion plus forte que la  $w\text{FWSR}$  ; la *Family Wise Separation Rate* qui permettra d'établir une erreur de seconde espèce pour les procédures de tests multiples.

### Définition 3.3.8 : Family Wise Separation Rate - Fromont, Lerasle, Reynaud-Bouret (2015) [6]

Soit  $\beta \in (0, 1)$  et une classe de lois de probabilité  $\mathcal{Q} \subset \mathcal{P}$ .

La *Family Wise Separation Rate* d'un test multiple  $\mathcal{R}$  sous  $\mathcal{Q}$  d'erreur de seconde espèce  $\beta$  est définie par

$$\begin{aligned} \text{FWSR}_d^\beta(R, \mathcal{Q}) &:= \inf\{r > 0 : \sup_{P \in \mathcal{Q}} P(\mathcal{F}_r(P) \cap (\mathcal{H} \setminus R) \neq \emptyset) \leq \beta\} \\ &= \inf\{r > 0 : \inf_{P \in \mathcal{Q}} P(\mathcal{F}_r(P) \subset R) \geq 1 - \beta\}. \end{aligned}$$

**Remarque 3.3.9**

Par définition, la  $\text{FWSR}_d^\beta(R, \mathcal{Q})$  est monotone en  $R$  ie

$$R \subset R' \implies \text{FWSR}_d^\beta(R', \mathcal{Q}) \leq \text{FWSR}_d^\beta(R, \mathcal{Q}).$$

La FWSR offre un contrôle plus fort que la  $w\text{FWSR}$  comme l'affirme la proposition suivante :

**Proposition 3.3.10** : *Fromont, Lerasle, Reynaud-Bouret (2015) [6]*

Pour tout sous-ensemble  $\mathcal{Q}$  de  $\mathcal{P}$  et  $\beta \in (0, 1)$ ,

$$w\text{FWSR}_d^\beta(R, \mathcal{Q}) \leq \text{FWSR}_d^\beta(R, \mathcal{Q}).$$

*Démonstration.*

Soit  $\mathcal{Q} \subset \mathcal{P}$ . Soit  $P \in \mathcal{Q}$ .

On remarque que  $P(R = \emptyset) \leq P(\mathcal{F}_r(P) \cap (\mathcal{H} \setminus R) \neq \emptyset)$  pour tout  $P \in \mathcal{Q}$  telle que  $\mathcal{F}_r(P) \neq \emptyset$ .

On en déduit donc que

$$\begin{aligned} \sup_{P \in \mathcal{Q}, \mathcal{F}_r(P) \neq \emptyset} P(R = \emptyset) &\leq \sup_{P \in \mathcal{Q}, \mathcal{F}_r(P) \neq \emptyset} P(\mathcal{F}_r(P) \cap (\mathcal{H} \setminus R) \neq \emptyset) \\ &\leq \sup_{P \in \mathcal{Q}} P(\mathcal{F}_r(P) \cap (\mathcal{H} \setminus R) \neq \emptyset). \end{aligned}$$

Ainsi,

$$\{r > 0, \sup_{P \in \mathcal{Q}} P(\mathcal{F}_r(P) \cap (\mathcal{H} \setminus R) \neq \emptyset) \leq \beta\} \subset \{r > 0, \sup_{P \in \mathcal{Q}, \mathcal{F}_r(P) \neq \emptyset} P(R = \emptyset) \leq \beta\}.$$

□

On peut à présent introduire une approche minimax pour les tests multiples.

**Définition 3.3.11** : *minimax Family Wise Separation Rate, test multiple minimax et adaptatif - Fromont, Lerasle, Reynaud-Bouret (2015) [6]*

1. Soit  $\alpha, \beta \in (0, 1)$  et une classe de lois de probabilité  $\mathcal{Q} \subset \mathcal{P}$ . La *FWSR minimax* sous  $\mathcal{Q}$  dont la FWER est de niveau  $\alpha$  et d'erreur de seconde espèce  $\beta$  est définie par

$$m\text{FWSR}_d^{\alpha, \beta}(\mathcal{Q}) := \inf_{R : \text{FWER}(R) \leq \alpha} \text{FWSR}_d^\beta(R, \mathcal{Q})$$

où l'infimum est pris sur tous les tests multiples dont la FWER est contrôlée par  $\alpha$ .

2. Une *procédure de tests multiples*  $\mathcal{R}$ , dont la FWER est contrôlée par  $\alpha$ , est dit *minimax* sous  $\mathcal{Q}$  si la  $\text{FWSR}_d^\beta(\mathcal{R}, \mathcal{Q})$  est égale à la  $m\text{FWSR}_d^{\alpha, \beta}(\mathcal{Q})$  à une constante multiplicative près dépendant de  $\alpha$  et  $\beta$ .
3. Une *procédure de tests multiples*  $\mathcal{R}$ , dont la FWER est contrôlée par  $\alpha$ , est dite *adaptative* (au sens minimax sous une collection de classes  $\mathcal{Q}$ ) si la  $\text{FWSR}_d^{\alpha, \beta}(\mathcal{R}, \mathcal{Q})$  atteint (éventuellement à une constante près) la  $m\text{FWSR}_d^{\alpha, \beta}(\mathcal{Q})$  simultanément pour toutes les classes  $\mathcal{Q}$  de la collection, sans connaître a priori dans quelle classe appartient la loi  $P$ .

Cette approche minimax peut être reliée, dans certains cas, à la théorie minimax pour les tests d'une seule hypothèse.

**Théorème 3.3.12 :** *Fromont, Lerasle, Reynaud-Bouret (2015) [6]*

Si la distance  $d$  et la collection d'hypothèses  $\mathcal{H}$  satisfont (\*), alors pour tout sous-ensemble  $\mathcal{Q}$  de  $\mathcal{P}$  et pour tout  $\beta \in (0, 1)$ ,

$$mFWSR_d^{\alpha, \beta}(\mathcal{Q}) \geq mSR_d^{\alpha, \beta}(\mathcal{Q}, \cap \mathcal{H}).$$

*Démonstration.*

On suppose que la condition (\*) est vérifiée.

Pour toute procédure de tests multiples  $\mathcal{R}$ , on a :

$$FWER(\mathcal{R}) \geq wFWER(\mathcal{R}) = ER(\bar{\Phi}(\mathcal{R}), \cap \mathcal{H}).$$

Ainsi,

$$\begin{aligned} mFWER_d^{\alpha, \beta}(\mathcal{Q}) &= \inf_{\mathcal{R}, FWER(\mathcal{R}) \leq \alpha} FWER_d^\beta(\mathcal{R}, \mathcal{Q}) \\ &\geq \inf_{\mathcal{R}, ER(\bar{\Phi}(\mathcal{R}), \cap \mathcal{H}) \leq \alpha} FWER_d^\beta(\mathcal{R}, \mathcal{Q}) \\ &\geq \inf_{\mathcal{R}, ER(\bar{\Phi}(\mathcal{R}), \cap \mathcal{H}) \leq \alpha} wFWSR_d^\beta(\mathcal{R}, \mathcal{Q}) \quad \text{d'après 3.3.10} \\ &= \inf_{\mathcal{R}, ER(\bar{\Phi}(\mathcal{R}), \cap \mathcal{H}) \leq \alpha} SR_d^\beta(\bar{\Phi}(\mathcal{R}), \mathcal{Q}, \cap \mathcal{H}) \quad \text{d'après 3.3.5} \\ &= \inf_{\bar{\Phi}, ER(\bar{\Phi}, \cap \mathcal{H}) \leq \alpha} SR_d^\beta(\bar{\Phi}, \mathcal{Q}, \cap \mathcal{H}) \quad \text{par correspondance entre le test multiple } \mathcal{R} \text{ et le test simple } \bar{\Phi}(\mathcal{R}) \text{ de } \cap \mathcal{H}, \\ &= mSR_d^{\alpha, \beta}(\mathcal{Q}, \cap \mathcal{H}). \end{aligned}$$

□

**Remarque 3.3.13 :** *Bornes inférieures pour la minimax FWSR*

Ce théorème permet de déterminer des bornes inférieures pour la minimax FWSR sous certaines classes  $\mathcal{Q}$  en utilisant les résultats minimax des tests simples.

### 3.3.3 Exemples : le cas gaussien.

**Exemple 1 :** condition (\*) non vérifiée.

Soit  $(e_1, \dots, e_n)$  une base de  $\mathbb{R}^n$ .

On suppose ici que  $\mathcal{H} = \{H_{S_i}, i = 1, \dots, n\}$  où  $S_i = \text{Vect}(e_i)$  et  $H_{S_i} = \{P_f, f_i = 0\} = \{P_f, f \in S_i^\perp\}$ .

On teste l'hypothèse nulle  $H_0 = \cap \mathcal{H} = \{P_0\}$  avec  $d = d_2$  (on a vu que pour cette distance, (\*) n'est pas vérifiée), sous la classe d'alternatives  $\mathcal{Q} = \mathcal{P}_k$  définie pour tout entier  $k \leq n$  par

$$\mathcal{P}_k = \{P_f, |f|_0 \leq k\},$$

avec  $|f|_0$  le nombre de coefficients non nuls de  $f$ .

Dans ce cadre, on a :

**Théorème 3.3.14 :** *Test simple - Baraud (2002) [1]*

Pour tous  $\alpha, \beta \in (0, 1)$  tels que  $\alpha + \beta \leq 0.5$  et  $k \geq 1$ ,

$$mSR_{d_2}^{\alpha,\beta}(\mathcal{P}_k, H_0) \geq \sigma \left( k \ln \left( 1 + \frac{n}{k^2} \vee \sqrt{\frac{n}{k^2}} \right) \right)^{1/2}.$$

On peut alors montrer que

**Théorème 3.3.15 :** *Tests multiples - Fromont, Lerasle, Reynaud-Bouret (2015) [6]*

Pour tous  $\alpha, \beta \in (0, 1)$  tels que  $\alpha + \beta \leq 0.5$ , pour tout  $s \in [1, \infty]$ , pour tout  $k \in \{1, \dots, n\}$ ,

$$mFWSR_{d_s}^{\alpha,\beta}(\mathcal{P}_k) \geq \sigma \sqrt{\ln(1+n)}.$$

*Démonstration.*

On commence par remarquer que pour tout  $\mathcal{Q} \subset \mathcal{A} \subset \mathcal{P}$ ,

$$\{r > 0, \sup_{P \in \mathcal{A}} P(\mathcal{F}_r(P) \cap (\mathcal{H} \setminus \mathcal{R}) \neq \emptyset) \leq \beta\} \subset \{r > 0, \sup_{P \in \mathcal{Q}} P(\mathcal{F}_r(P) \cap (\mathcal{H} \setminus \mathcal{R}) \neq \emptyset) \leq \beta\}.$$

On en déduit alors que  $FWSR_d^\beta(\mathcal{R}, \mathcal{A}) \geq FWSR_d^\beta(\mathcal{R}, \mathcal{Q})$  et donc que

$$mFWSR_d^{\alpha,\beta}(\mathcal{A}) \geq mFWSR_d^{\alpha,\beta}(\mathcal{Q}).$$

Comme  $\mathcal{P}_1 \subset \mathcal{P}_k$  pour tout  $k \in \{1, \dots, n\}$ , alors

$$mFWSR_{d_s}^{\alpha,\beta}(\mathcal{P}_k) \geq mFWSR_{d_s}^{\alpha,\beta}(\mathcal{P}_1).$$

Or, pour tout  $P_f \in \mathcal{P}_1$  et tout  $i = 1, \dots, n$ ,

$$d_s(P_f, H_{S_i}) = d_\infty(P_f, H_{S_i}) = d_2(P_f, H_{S_i}).$$

Alors, on a d'une part,

$$mFWSR_{d_s}^{\alpha,\beta}(\mathcal{P}_1) = mFWSR_{d_\infty}^{\alpha,\beta}(\mathcal{P}_1).$$

D'autre part, comme (\*) est vérifiée par  $d_\infty$ , on a

$$mFWSR_{d_\infty}^{\alpha,\beta}(\mathcal{P}_1) \geq mSR_{d_\infty}^{\alpha,\beta}(\mathcal{P}_1, \cap \mathcal{H}) = mSR_{d_2}^{\alpha,\beta}(\mathcal{P}_1, \cap \mathcal{H}).$$

En conclusion, d'après 3.3.14, on obtient la borne inférieure souhaitée.  $\square$

On va à présent montrer que cette borne inférieure est atteinte pour un test que nous allons préciser.

Pour tout  $i = 1, \dots, n$ , on considère les  $p$ -valeurs  $p_i$  correspondant aux statistiques de test  $S_i(X) = |X_i| \sigma^{-1}$  et aux zones de rejet  $H_{S_i}$  définies par  $S_i(X) > F^{-1}(1 - \alpha/2)$  avec  $F$  la fonction de répartition d'une loi normale standard. Comme  $F$  est continue, on sait que les trois tests suivants sont égaux :

$$\mathbb{1}_{\{S_i(X) > F^{-1}(1-\alpha/2)\}} = \mathbb{1}_{\{S_i(X) > \bar{F}^{-1}(1-\alpha/2)\}} = \mathbb{1}_{p_i \leq \alpha}.$$

On a alors

**Théorème 3.3.16 :** *Fromont, Lerasle, Reynaud-Bouret (2015) [6]*

Soit  $\alpha \in (0, 1)$  et soit  $\mathcal{R}$ , une des quatre procédures de tests multiples suivantes :  $\mathcal{R}_{Bonf}$ ,  $\mathcal{R}_{Holm}$ ,  $\mathcal{R}_{mp}$  ou  $\mathcal{R}(\Phi_{\mathcal{H}}^{BHL})$ , basée sur les  $p$ -valeurs  $p_i$  définies précédemment et telle que  $\text{FWER}(\mathcal{R}) \leq \alpha$ .

Alors, pour tout  $s \in [1, \infty]$ ,  $k \in \{1, \dots, n\}$  et  $\beta \in (0, 1)$ ,

$$FWSR_{d_s}^\beta(\mathcal{R}, \mathcal{P}_k) \leq \sigma \left( \sqrt{2 \ln \left( \frac{k}{2\beta} \right)} + \sqrt{2 \ln \left( \frac{2n}{\alpha} \right)} \right).$$

*Démonstration.*

Il suffit de montrer qu'il existe  $r_0 > 0$  tel que pour tout  $r \geq r_0$  et pour tout  $P_f \in \mathcal{P}_k$ ,

$$P_f(\mathcal{F}_r(P_f) \subset \mathcal{R}_{Bonf}) \geq 1 - \beta.$$

En effet, on a, par construction :  $\mathcal{R}_{Bonf} \subset \mathcal{R}_{Holm} \subset \mathcal{R}_{mp}$ .

Par 3.3.2, on a aussi que  $\mathcal{R}_{Bonf} = \mathcal{N}_{Holm}(\emptyset) \subset \mathcal{N}_{mp}(\emptyset) = \mathcal{R}(\Phi_{\mathcal{H}}^{BHL})$ .

Ainsi, d'après 3.3.9 (monotonie de la FWSR), l'obtention de la borne supérieure pour  $\mathcal{R}_{Bonf}$  permettra de prouver le théorème pour les trois autres procédures de tests multiples.

Comme  $d_s(P_f, H_{S_i}) = \inf_{P_g \in H_{S_i}} d_s(P_f, P_g) = |f_i|$ , on a

$$\begin{aligned} P_f(\mathcal{F}_r(P_f) \subset \mathcal{R}_{Bonf}) &= P_f(\forall i \in \{1, \dots, n\} \text{ tel que } d_s(P_f, H_{S_i}) \geq r, H_{S_i} \in \mathcal{R}_{Bonf}) \\ &= P_f(\forall i \in \{1, \dots, n\} \text{ tel que } |f_i| \geq r, p_i \leq \alpha/n) \\ &= \prod_{i, |f_i| \geq r} P_f(p_i \leq \alpha/n) \text{ par indépendance des } p_i. \end{aligned}$$

On note  $F$  la fonction de répartition d'une variable aléatoire gaussienne standard  $N$ . On notera  $N_1, \dots, N_n$  des variables aléatoires gaussiennes standard iid et indépendante de  $N$ . On a vu que, pour tout  $i \in \{1, \dots, n\}$ ,  $p_i = 2F(-|N_i|)$  et comme  $X_i = \sigma N_i$ , alors

$$p_i = 2F(-\sigma^{-1}|X_i|) = 2F\left(-\frac{f_i}{\sigma} + N_i\right).$$

De plus, on remarque que pour tout  $a, b \in \mathbb{R}$ ,

$$P(|a + N| > b) \geq F(|a| - b). \quad (\#)$$

En effet, on montre que

$$\mathbb{P}(|a + N| > b) = 1 - \mathbb{P}(N \leq b - a) + \mathbb{P}(N \leq -b - a)$$

$$\text{et que } \mathbb{P}(N \leq |a| - b) = 1 - \mathbb{P}(N \leq b - a) - \mathbb{P}(N \leq -b - a),$$

ce qui permet de conclure.

Alors, comme  $F$  est continue,

$$\begin{aligned} P_f(\mathcal{F}_r(P_f) \subset \mathcal{R}_{Bonf}) &= \prod_{i, |f_i| \geq r} P_f\left(\frac{f_i}{\sigma} + N_i \geq -F^{-1}\left(\frac{\alpha}{2n}\right)\right) \\ &\geq \prod_{i, |f_i| \geq r} \left(F\left(\frac{|f_i|}{\sigma} + F^{-1}\left(\frac{\alpha}{2n}\right)\right)\right) \quad \text{d'après } (\#), \\ &\geq \left(F\left(\frac{r}{\sigma} + F^{-1}\left(\frac{\alpha}{2n}\right)\right)\right)^{|\mathcal{F}_r(P_f)|}. \end{aligned}$$

Alors  $P_f(\mathcal{F}_r(P_f) \subset \mathcal{R}_{Bonf}) \geq 1 - \beta$  si

$$r \geq \sigma \left( F^{-1}\left((1 - \beta)^{1/|\mathcal{F}_r(P_f)|}\right) - F^{-1}\left(\frac{\alpha}{2n}\right) \right),$$

ou de manière équivalente, si

$$r \geq \sigma \left( F^{-1}\left((1 - \beta)^{1/|\mathcal{F}_r(P_f)|}\right) + F^{-1}\left(1 - \frac{\alpha}{2n}\right) \right).$$

En effet, on a  $F^{-1}(1 - \alpha) = -F^{-1}(\alpha)$  car  $\mathbb{P}(N \leq -F^{-1}(\alpha)) = \mathbb{P}(N \geq F^{-1}(\alpha)) = 1 - F(F^{-1}(\alpha)) = 1 - \alpha$ .

On rappelle à présent que pour une loi normale, on a :

$$\forall u > 0, \quad 1 - F(u) \leq \frac{1}{2} e^{-u^2/2}.$$

Cette inégalité implique, par définition de  $F^{-1}$ , que

$$\forall u \in (0, 1), \quad F^{-1}(u) \leq \sqrt{-2 \ln(2(1 - u))}. \quad (\#\#)$$

Ainsi  $P_f(\mathcal{F}_r(P_f) \subset \mathcal{R}_{Bonf}) \geq 1 - \beta$  si

$$r \geq \sigma \left( \sqrt{-2 \ln(2(1 - (1 - \beta)^{1/|\mathcal{F}_r(P_f)|}))} + \sqrt{-2 \ln \left( \frac{\alpha}{2n} \right)} \right).$$

Enfin, on a, pour tout  $u \in (0, 1)$  et  $x \in [0, 1]$ ,

$$x(1 - u) \leq 1 - u^x.$$

En effet, en introduisant la fonction  $g(u) = 1 - u^x - x(1 - u)$ , on montre que  $g'(u) = x(1 - e^{x \ln(u)})/u \leq 0$  et on obtient le résultat en remarquant que  $g(1) = 0$ .

On obtient donc que  $P_f(\mathcal{F}_r(P_f) \subset \mathcal{R}_{Bonf}) \geq 1 - \beta$  si

$$r \geq \sigma \left( \sqrt{2 \ln \left( \frac{|\mathcal{F}_r(P_f)|}{2\beta} \right)} + \sqrt{2 \ln \left( \frac{2n}{\alpha} \right)} \right).$$

Enfin, comme  $P_f \in \mathcal{P}_k$ , on a  $|\mathcal{F}_r(P_f)| \leq k$  et on obtient le résultat voulu.  $\square$

### Remarque 3.3.17

1. On a montré que les quatre procédures de tests multiples précédentes sont minimax sous les classes  $\mathcal{P}_k$  avec une FWSR de l'ordre de  $\sigma(\ln(n))^{1/2}$  à une constante multiplicative près. Comme les procédures ne dépendent pas de  $k$ , ces tests sont adaptatifs en un sens minimax sur toutes les classes  $\mathcal{P}_k$  pour  $k = 1, \dots, n$  simultanément.
2. On peut interpréter l'hypothèse (\*) comme une condition pour que l'étude de la performance d'une procédure de tests multiples soit plus difficile que l'étude d'un test d'une seule hypothèse.

### Exemple 2 : condition (\*) vérifiée.

Soit  $(e_1, \dots, e_n)$  une base de  $\mathbb{R}^n$ .

On suppose ici que  $\mathcal{H} = \{H_{\bar{S}_i}, i = 1, \dots, n\}$  où  $\bar{S}_i = \text{Vect}(e_1, \dots, e_i)$  et  $H_{\bar{S}_i} = \{P_f, f_1 = \dots = f_i = 0\} = \{P_f, \Pi_{\bar{S}_i}(f) = 0\}$ .

On teste l'hypothèse nulle  $H_0 = \cap \mathcal{H} = \{P_0\}$  sous la classe d'alternatives  $\mathcal{Q} = \mathcal{P}_k$  définie précédemment.

Comme la collection d'hypothèses  $\mathcal{H}$  est fermée, la condition (\*) est vérifiée pour toute distance d'après 3.3.12 et on a, pour toute distance  $d$ ,

$$mFWSR_d^{\alpha, \beta}(\mathcal{P}_k) \geq mSR_d^{\alpha, \beta}(\mathcal{P}_k, H_0).$$

En particulier, pour  $d = d_2$ , pour tous  $\alpha, \beta \in (0, 1)$  tels que  $\alpha + \beta \leq 0.5$ , pour tout  $k \in \{1, \dots, n\}$ , on a, d'après 3.3.14,

$$mFWSR_d^{\alpha, \beta}(\mathcal{P}_k) \geq \sigma \left( k \ln \left( 1 + \frac{n}{k^2} \vee \sqrt{\frac{n}{k^2}} \right) \right)^{1/2}.$$

On va à présent définir une procédure de tests multiples, qui ne dépend pas de  $k$  et dont la FWSR sous  $\mathcal{P}_k$  atteint cette borne inférieure (à une constante multiplicative près).

Comme pour l'exemple précédent, on définit, pour tout  $i \in \{1, \dots, n\}$ , les  $p$ -valeurs  $p_i$  associées aux tests simples qui rejettent les hypothèses nulles  $H_i = \{P_f, f_i = 0\}$  lorsque  $S_i(X) = |X_i| \sigma^{-1}$  prend de "grandes" valeurs.

On introduit alors une procédure de tests multiples :

$$\bar{\mathcal{R}} = \{H_{\bar{S}_i}, \min_{j \leq i} p_j \leq \alpha/n\}.$$

Comme la fonction de répartition d'une loi normale standard  $F$  est continue, on remarque que l'on a

$$\begin{aligned} \bar{\mathcal{R}} &= \{H_{\bar{S}_i}, \min_{j \leq i} S_j(X) > F^{-1} \left( 1 - \frac{\alpha}{2n} \right)\} \\ &= \{H_{\bar{S}_i}, \min_{j \leq i} S_j(X) > \bar{F}^{-1} \left( 1 - \frac{\alpha}{2n} \right)\}. \end{aligned}$$

On a alors :

**Théorème 3.3.18 :** *Fromont, Lerasle, Reynaud-Bouret (2015) [6]*

Soit  $\alpha \in (0, 1)$ . On considère le test multiple  $\overline{\mathcal{R}}$  définie précédemment.

Alors

$$\text{FWER}(\overline{\mathcal{R}}) \leq \alpha,$$

et pour tout  $k \in \{1, \dots, n\}$ ,  $\beta \in (0, 0.5)$ ,

$$\text{FWSR}_{d_2}^\beta(\overline{\mathcal{R}}, \mathcal{P}_k) \leq \sigma\sqrt{k} \left( \sqrt{-2 \ln(2\beta)} + \sqrt{2 \ln\left(\frac{2n}{\alpha}\right)} \right).$$

*Démonstration.*

Nous allons commencer par montrer que  $\text{FWER}(\overline{\mathcal{R}}) \leq \alpha$ .

Pour tout  $f \in \mathbb{R}^n$ , on note  $i_0$  le plus grand indice  $i$  pour lequel  $f_1 = \dots = f_i = 0$ . Alors

$$\begin{aligned} P_f(\overline{\mathcal{R}} \cap \tau(P_f) \neq \emptyset) &= P_f(\exists i \leq i_0, \exists j \leq i, p_j \leq \alpha/n) \\ &\leq P_f(\exists j \leq i_0, p_j \leq \alpha/n) \\ &\leq \alpha. \end{aligned}$$

On considère à présent  $d = d_2$ . On va trouver  $r_0$  tel que pour tout  $r \geq r_0$  et pour tout  $P_f \in \mathcal{P}_k$ ,

$$P_f(\mathcal{F}_r(P_f) \subset \overline{\mathcal{R}}) \geq 1 - \beta.$$

Supposons que  $P_f \in \mathcal{P}_k$ . Soit  $r > 0$ .

$$\mathcal{F}_r(P_f) = \{H_{\overline{S}_i}, d_2(P_f, H_{\overline{S}_i}) \geq r\} = \{H_{\overline{S}_i}, \sum_{j \in \{1, \dots, i\}} f_j^2 \geq r^2\}.$$

Si  $\sum_{j=1}^n f_j^2 < r^2$  alors  $P_f(\mathcal{F}_r(P_f) \subset \overline{\mathcal{R}}) = 1$  et on peut conclure directement (en effet, sous cette condition, si  $H_{\overline{S}_i} \in \mathcal{F}_r(P_f)$  alors  $H_{\overline{S}_i} = \emptyset \subset \overline{\mathcal{R}}$ ).

Sinon, notons  $i_0$ , le plus petit entier de  $\{1, \dots, n\}$  tel que  $\sum_{j=1}^{i_0} f_j^2 \geq r^2$ .

Comme cette somme a au plus  $i_0 \wedge k$  termes non nuls, il existe  $j_0 \in \{1, \dots, i_0\}$  tel que  $f_{j_0}^2 \geq r^2 / (i_0 \wedge k) \geq r^2 / k$ . De plus,

$$P_f(\mathcal{F}_r(P_f) \subset \overline{\mathcal{R}}) = P_f \left( \forall i \text{ tel que } \sum_{j \in \{1, \dots, i\}} f_j^2 \geq r^2, \min_{j \in \{1, \dots, i\}} p_j \leq \frac{\alpha}{n} \right).$$

Supposons que  $p_{j_0} \leq \alpha/n$ .

On a alors  $\min_{j=1, \dots, i_0} p_j \leq \alpha/n$ . Aussi, si pour tout  $i \in \{1, \dots, n\}$ ,  $\sum_{j=1}^i f_j^2 \geq r^2$  alors on a  $i \geq i_0 \geq j_0$  (sinon cela contredirait la minimalité de  $i_0$ ).

Donc  $\min_{j=1, \dots, i} p_j \leq \alpha/n$ .

On a donc l'inclusion des évènements

$$\{p_{j_0} \leq \alpha/n\} \subset \{\forall i : \sum_{j=1}^i f_j^2 \geq r^2, \min_{j=1, \dots, i} p_j \leq \alpha/n\}.$$

Alors,

$$\begin{aligned} P_f(\mathcal{F}_r(P_f) \subset \overline{\mathcal{R}}) &\geq P_f(p_{j_0} \leq \alpha/n) \\ &\geq P_f(2F(-\sigma^{-1}|X_{j_0}|) \leq \alpha/n) \\ &\geq P_f \left( 2F\left(-\left|\frac{f_{j_0}}{\sigma} + N_i\right|\right) \leq \frac{\alpha}{n} \right) \\ &= P_f \left( \left|\frac{f_{j_0}}{\sigma} + N_i\right| \geq -F^{-1}\left(\frac{\alpha}{2n}\right) \right). \end{aligned}$$

Ainsi, par (#),

$$\begin{aligned} P_f(\mathcal{F}_r(P_f) \subset \overline{\mathcal{R}}) &\geq F\left(\frac{|f_{j_0}|}{\sigma} + F^{-1}\left(\frac{\alpha}{2n}\right)\right) \\ &\geq F\left(\frac{r}{\sqrt{k}\sigma} + F^{-1}\left(\frac{\alpha}{2n}\right)\right). \end{aligned}$$

On a donc  $P_f(\mathcal{F}_r(P_f) \subset \overline{\mathcal{R}}) \geq 1 - \beta$  si

$$F\left(\frac{r}{\sqrt{k}\sigma} + F^{-1}\left(\frac{\alpha}{2n}\right)\right) \geq 1 - \beta.$$

Enfin, d'après (##),  $P_f(\mathcal{F}_r(P_f) \subset \overline{\mathcal{R}}) \geq 1 - \beta$  si

$$\frac{r}{\sigma\sqrt{k}} \geq \sqrt{-2\ln(2\beta)} + \sqrt{2\ln(2n/\alpha)},$$

ce qui conclut la preuve. □

### Remarque 3.3.19

Pour  $k$  proportionnel à  $n^\gamma$  où  $\gamma \in [0, 1/2)$ , on remarque que cette borne supérieure coïncide avec la borne inférieure obtenue en 3.3.14. On conclut alors que  $\overline{\mathcal{R}}$  est adaptatif au sens minimax sous les classes  $\mathcal{P}_k$  pour tout  $k \in \{1, \dots, n\}$  simultanément.



# Chapitre 4

## Tests multiples sur un continuum d'hypothèses et contrôle de la FDR.

L'objectif de ce chapitre est de définir les tests multiples d'un ensemble infini non dénombrable d'hypothèses. Nous insisterons sur les problèmes de mesurabilité liés à cette théorie et généraliserons aussi certains résultats sur les tests multiples d'un ensemble fini d'hypothèses à ce cadre plus général. Nous nous concentrerons particulièrement sur la définition et le contrôle de la FDR dans le cas d'un continuum d'hypothèses.

### 4.1 Tests multiples d'un continuum d'hypothèses.

#### 4.1.1 Procédure de tests multiples sur un continuum d'hypothèses.

Une motivation importante pour définir des tests multiples sur un ensemble infini non dénombrable d'hypothèses est la présence de données continues dans nos observations ou nos hypothèses. Ainsi, on veut généraliser des problèmes de tests lorsque l'observation du modèle statistique est celle d'un processus stochastique sur un intervalle de temps. Aussi, on peut être amené à tester des hypothèses qui sont indexées par le temps (des hypothèses sur la densité de probabilité, sur la fonction de répartition par exemple) : on a alors un nombre indénombrable d'hypothèses à tester simultanément. Par abus de langage, un *ensemble non dénombrable d'hypothèses* sera appelé *ensemble continu d'hypothèses*. Ce chapitre est basé sur l'article de Blanchard, Delattre, Roquain [3].

Soit  $(\mathbb{X}, \mathcal{X}, \mathcal{P})$  un modèle statistique et  $X$  une observation de loi  $P \in \mathcal{P}$ . On considère le problème de tests multiples pour  $P$  défini précédemment.

Soit  $\mathcal{H}$  l'espace des indices des hypothèses. Pour tout  $h \in \mathcal{H}$ , on définit une hypothèse  $H_h \subset \mathcal{P}$ . On notera  $\mathcal{H}_0(P) := \{h \in \mathcal{H} : P \in H_h\}$  et  $\mathcal{H}_1(P) := \mathcal{H} \setminus \mathcal{H}_0(P)$ . Ici,  $\mathcal{H}$  est a priori un ensemble quelconque (infini indénombrable en toute généralité).

On va à présent définir formellement la notion de procédure de tests multiples.

**Condition 4.1.1 :** *Mesurabilité de l'ensemble  $\mathcal{H}$ .*

L'espace des indices  $\mathcal{H}$  est muni d'une tribu  $\mathfrak{H}$  et pour tout  $P \in \mathcal{P}$ , l'ensemble  $\mathcal{H}_0(P)$  est supposé  $\mathfrak{H}$ -mesurable, c'est-à-dire que  $\mathcal{H}_0(P) \in \mathfrak{H}$ .

**Définition 4.1.2 :** *Procédure de tests multiples.*

Soit  $X : (\Omega, \mathcal{F}) \rightarrow (\mathbb{X}, \mathcal{X})$  une variable aléatoire,  $(\mathbb{X}, \mathcal{X}, \mathcal{P})$  un modèle statistique et  $\mathcal{H}$  un ensemble d'indices pour les hypothèses, satisfaisant la condition 4.1.1.

Une procédure de tests multiples sur  $\mathcal{H}$  est une application  $R : X(\Omega) \subset \mathbb{X} \rightarrow \mathfrak{h}$  telle que l'ensemble

$$\{(\omega, h) \in \Omega \times \mathcal{H} : h \in R(X(\omega))\}$$

est un ensemble  $\mathcal{F} \otimes \mathfrak{H}$ -mesurable.

Autrement dit, cela signifie que le processus  $(\mathbb{1}\{h \in R(X)\})_{h \in \mathcal{H}}$  est un processus mesurable pour la tribu produit, c'est-à-dire que l'application

$$\begin{aligned} (\Omega \times \mathcal{H}, \mathcal{F} \otimes \mathfrak{H}) &\longrightarrow (\{0, 1\}, \mathcal{P}(\{0, 1\})) \\ (\omega, h) &\longmapsto \mathbb{1}\{h \in R(X(\omega))\} \end{aligned}$$

est une application mesurable pour la tribu produit.

Comme dans les chapitres précédents, on veut construire une procédure de tests multiples à partir d'une famille de  $p$ -valeurs. Rappelons et précisons la définition d'une  $p$ -valeur.

**Définition 4.1.3 :**  *$p$ -valeur.*

La fonctionnelle  $p$ -valeur est définie comme une application  $\mathbf{p} : \mathbb{X} \rightarrow [0, 1]^{\mathcal{H}}$ , ou de manière équivalente, comme une collection de fonctions  $\mathbf{p} = (p_h(\cdot))_{h \in \mathcal{H}}$  avec pour tout  $h \in \mathcal{H}$ ,  $p_h : \mathbb{X} \rightarrow [0, 1]$  une statistique calculée à partir de l'observation  $x \in \mathbb{X}$ .

Pour tout  $P \in \mathcal{P}$ , les lois marginales des  $p$ -valeurs sous  $\mathcal{H}_0(P)$  sont minorées en probabilité par une loi uniforme sur  $[0, 1]$  :

$$\forall h \in \mathcal{H}_0(P) \quad \forall u \in [0, 1], \quad \mathbb{P}_P(p_h(X) \leq u) \leq u.$$

On sera amené à considérer "la variable aléatoire  $p$ -valeur",  $\omega \in \Omega \mapsto p_h(X(\omega))$ , et "le processus  $p$ -valeur",  $\omega \in \Omega \mapsto \mathbf{p}(X(\omega))$ .

On fera l'hypothèse suivante sur le processus  $p$ -valeur, spécifique au cas d'un continuum d'hypothèses :

**Condition 4.1.4 :** *Mesurabilité du processus  $p$ -valeur.*

On suppose que le processus stochastique  $(p_h(X))_{h \in \mathcal{H}}$  est un processus mesurable (pour la tribu produit), c'est-à-dire que l'application

$$(\omega, h) \in (\Omega \times \mathcal{H}, \mathcal{F} \otimes \mathfrak{H}) \mapsto p_h(X(\omega)) \in [0, 1]$$

est mesurable (pour la tribu produit).

On rappelle également qu'à partir d'une statistique de test  $S_h(x)$  à valeurs dans  $\mathbb{R}$ , on définit (d'après 1.2.6) une  $p$ -valeur par :

$$p_h(x) = \sup_{P \in \mathcal{H}_h} T_{P,h}(S_h(x))$$

où  $T_{P,h}(t) = \mathbb{P}_P(S_h(X) \geq t)$  et où le supremum est supposé conserver la mesurabilité en  $x$  pour tout  $h$  fixé.

**Remarque 4.1.5 :** *Modification de processus*

Afin de vérifier la condition de mesurabilité 4.1.4, on pourra être amené à considérer une *modification*<sup>1</sup> du processus stochastique observé. Ainsi, s'il existe une modification de  $X$  telle que la condition 4.1.4 soit vérifiée alors on supposera que l'on observe cette modification de  $X$  afin que les hypothèses de mesurabilité soient satisfaites (voir l'exemple 4.1.11). On peut aussi montrer que le contrôle de la FDR étudié dans la section 4.2 ne dépend pas du choix de la modification du processus  $X$ .

1. On rappelle qu'un processus  $(Y_h)_{h \in \mathcal{H}}$  est une *modification* ou *version* d'un processus  $(X_h)_{h \in \mathcal{H}}$  si  $\mathbb{P}(X_h = Y_h)$  pour tout  $h \in \mathcal{H}$ . Cela implique notamment que les processus  $X$  et  $Y$  ont même lois infindimensionnelles.

**Remarque 4.1.6 :** *Indépendance des  $p$ -valeurs et mesurabilité.*

On dispose du résultat suivant qui montre que, pour un espace  $\mathcal{H}$  non dénombrable, le fait que les  $p$ -valeurs  $\{p_h, h \in \mathcal{H}\}$  soient mutuellement indépendantes empêche la condition 4.1.4 d'être vérifiée. Cela est une différence majeure avec le cas classique où  $\mathcal{H}$  est finie !

**Lemme 4.1.7 :** *Revuz, Yor (1991) [12]*

Soit  $(Z_t)_{t \in [0,1]}$  un processus stochastique réel sur  $(\Omega, \mathcal{F}, \mathbb{P})$  où  $[0,1]$  est muni de la tribu borélienne et de la mesure de Lebesgue  $\lambda$ . Supposons que, pour tout  $t \in [0,1]$ ,  $\mathbb{E}[Z_t] = 0$ ,  $Z_t$  est de carré intégrable et  $\text{Var}(Z_t) = 1$ . Alors, si les variables  $(Z_t)_t$  sont mutuellement indépendantes, l'application  $(\omega, t) \mapsto Z_t(\omega)$  n'est pas mesurable conjointement en ces deux variables.

La preuve de ce lemme est un raisonnement par l'absurde ; si on suppose la mesurabilité de l'application  $(\omega, t) \mapsto Z_t(\omega)$  alors le théorème de Fubini serait licite et on pourrait montrer que  $\text{Var}(Z_t) = 0$ .

Nous allons chercher à affaiblir la condition 4.1.4.

► *Cas où  $\mathcal{H}$  est au plus dénombrable :* si  $\mathcal{H}$  est muni de la tribu de l'ensemble des parties de  $\mathcal{H}$  alors la condition 4.1.4 est vérifiée (par composition d'applications mesurables) dès lors que l'on suppose que, pour tout  $h \in \mathcal{H}$ ,  $p_h : \mathbb{X} \rightarrow [0,1]$  est mesurable (ce qui est une hypothèse usuelle en tests multiples, voir 2.1.2).

► *Cas où  $\mathcal{H}$  n'est pas dénombrable :* une condition suffisante pour vérifier 4.1.4 est la mesurabilité conjointe de la fonction  $p$ -valeur (par composition d'applications mesurables) au sens suivant :

**Condition 4.1.8 :** *Mesurabilité conjointe de la fonctionnelle  $p$ -valeur.*

On suppose que

$$(x, h) \in (\mathbb{X} \times \mathcal{H}, \mathcal{X} \otimes \mathfrak{H}) \mapsto p_h(x) \in [0,1]$$

est mesurable pour la tribu produit.

Si les hypothèses de mesurabilité sont facilement vérifiées dans le cas fini ou dénombrable, ce n'est pas toujours le cas dans le cadre infini non dénombrable :

**Contre-exemple 4.1.9**

On suppose que l'on observe un processus stochastique indexé par l'espace des hypothèses  $\mathcal{H}$ ,  $X = \{X_h, h \in \mathcal{H}\}$ . Dans ce cas, l'espace des observations  $\mathbb{X}$  est inclu dans  $\mathbb{R}^{\mathcal{H}}$ . De plus, supposons que la fonctionnelle  $p$ -valeur est de la forme  $p_h(x) := \psi(x_h)$  pour tout  $h \in \mathcal{H}$  où  $\psi$  est une application mesurable qui associe à  $x \in \mathbb{X}$  sa valeur au point  $h : x_h \in \mathbb{R}$ .

Pour assurer la condition de mesurabilité conjointe 4.1.8, il suffit de prouver la mesurabilité (par rapport aux deux variables) de l'application  $(x, h) \in \mathbb{X} \times \mathcal{H} \mapsto x_h$ .

Il est important de remarquer ici que cette mesurabilité dépend totalement de l'espace  $\mathbb{X}$  (l'exemple 4.1.11 illustrera ce constat).

Prenons par exemple,  $\mathbb{X} := \mathbb{R}^{\mathcal{H}}$  que l'on munit de la tribu produit canonique (explicitée dans l'exemple 4.1.11). Dans ce cas, la condition 4.1.8 n'est pas vérifiée, sinon cela impliquerait que toute fonction de  $\mathcal{H}$  dans  $\mathbb{R}$  est mesurable, ce qui est absurde.

Ces questions de mesurabilité sont importantes comme le souligne la remarque suivante :

**Remarque 4.1.10 :** *Importance des hypothèses de mesurabilité*

Définir une procédure de tests multiples pour une infinité non dénombrable d'hypothèses impose une condition

de mesurabilité sur la procédure de tests multiples 4.1.2 et sur le processus  $p$ -valeur 4.1.4. Ces conditions de mesurabilité seront nécessaires pour appliquer le théorème de Fubini et ainsi "permuter les espérances définies à partir de  $\omega$  et les intégrales sur  $\mathcal{H}$ ".

**Exemple 4.1.11 :** *Test sur la moyenne d'un processus*

On observe un processus stochastique  $X = (X_t)_{t \in [0,1]^d}$ .

La fonction moyenne  $\mu : t \in [0, 1]^d \mapsto \mu(t) := \mathbb{E}[X_t]$  est inconnue.

On suppose que pour tout  $t \in [0, 1]^d$ ,  $X_t \sim \mathcal{N}(\mu(t), 1) := P$ .

Le modèle statistique associé est donc  $(\mathbb{X}, \mathcal{X}, \mathcal{P}) := (\mathbb{X}, \mathcal{X}, \mathcal{N}(\mu(t), 1))_{t \in [0,1]^d}$  où l'espace des observations  $(\mathbb{X}, \mathcal{X})$  est à préciser selon les propriétés du processus.

Ici,  $\mathcal{H} = [0, 1]^d$  et on considère le problème de test suivant :

$$H_t : " \mu(t) \leq 0 " \quad \forall t \in [0, 1]^d.$$

On a donc, pour tout  $P \in \mathcal{P}$ ,

$$\mathcal{H}_0(P) = \{t \in [0, 1]^d : \mu(t) \leq 0\}.$$

On remarque que  $X_t - \mu(t) \sim \mathcal{N}(0, 1)$ ; intuitivement, sous  $\mathcal{H}_0(P)$ ,  $X_t - \mu(t)$  est "grand".

On note  $G(x) := \mathbb{P}(Z \geq x)$  pour tout  $x \in \mathbb{R}$  où  $Z \sim \mathcal{N}(0, 1)$ . On définit alors le processus de  $p$ -valeurs :

$$p_t(X) := G(X_t) \quad \forall t \in [0, 1]^d.$$

Ce processus définit bien un processus de  $p$ -valeurs (définition 4.1.3) car

$$\forall u \in [0, 1], \forall t \in [0, 1]^d, \mu(t) \leq 0, \quad \mathbb{P}_P(p_t(X) \leq u) = \mathbb{P}_P(G(X_t) \leq u) \leq \mathbb{P}_P(G(X_t - \mu(t)) \leq u) \leq u.$$

Montrons la condition de mesurabilité 4.1.4 pour différents espaces d'observations.

1.  **$X$  est à trajectoires continues.**

On a alors :

$$\begin{aligned} X : (\Omega, \mathcal{F}) &\longrightarrow \mathcal{C}([0, 1]^d, \mathbb{R}) := \mathbb{X} \\ \omega &\longmapsto (X_t(\omega))_{t \in [0,1]^d} \end{aligned}$$

Précisons maintenant la tribu sur  $\mathcal{C}([0, 1]^d, \mathbb{R})$ . On a :

$$\mathcal{C}([0, 1]^d, \mathbb{R}) \subset \mathbb{R}^{[0,1]^d}.$$

► On munit  $\mathbb{R}^{[0,1]^d}$  de la tribu cylindrique  $\sigma(Cyl)$  engendrée par la famille des cylindres

$$Cyl = \{ \{x : [0, 1]^d \rightarrow \mathbb{R} : x(t_1) \in A_1, \dots, x(t_p) \in A_p\} \mid A_1, \dots, A_p \in \mathcal{B}(\mathbb{R}), p \in \mathbb{N} \}.$$

Il s'agit de la tribu sur  $\mathbb{R}^{[0,1]^d}$  qui rend mesurable les applications coordonnées

$$\pi_t : x \in \mathbb{R}^{[0,1]^d} \mapsto x(t) \in \mathbb{R}.$$

► Comme  $X$  est à trajectoires continues,  $X$  est à valeurs dans  $(\mathcal{C}([0, 1]^d, \mathbb{R}), \|\cdot\|_\infty)$ , espace normé que l'on munit de la tribu borélienne  $\mathcal{B}(\mathcal{C}([0, 1]^d, \mathbb{R}))$ .

► On dispose enfin d'un résultat affirmant que la tribu borélienne est égale à la tribu trace de  $\mathbb{R}^{[0,1]^d}$  :

$$\mathcal{B}(\mathcal{C}([0, 1]^d, \mathbb{R})) = \sigma(Cyl) \cap \mathcal{C}([0, 1]^d, \mathbb{R}).$$

Alors la fonction

$$(x, t) \in \mathcal{C}([0, 1]^d, \mathbb{R}) \times [0, 1]^d \mapsto p_t(x) = G(x(t)) \in [0, 1]$$

vérifie la condition 4.1.8 (et donc le processus  $p$ -valeur vérifie 4.1.4). En effet,

✓  $(x, t) \in \mathcal{C}([0, 1]^d, \mathbb{R}) \times [0, 1]^d \mapsto x(t)$  est mesurable car continue;

✓  $x \in \mathbb{R} \mapsto G(x) \in [0, 1]$  est mesurable car continue;

✓ donc  $p_t(x)$  est mesurable par composition d'applications mesurables.

## 2. $d = 1$ , $X$ est continue à droite et admet une limite à gauche.

Ici,  $\mathbb{X} := D([0, 1], \mathbb{R})$  espace de Skorohod des fonctions càdlàg de  $[0, 1]$  dans  $\mathbb{R}$ .

► Il existe une métrique  $d_0$  sur  $D([0, 1], \mathbb{R})$  permettant de munir  $\mathbb{X}$  d'une tribu borélienne.

► On a toujours  $D([0, 1], \mathbb{R}) \subset \mathbb{R}^{[0, 1]}$  avec  $(\mathbb{R}^{[0, 1]}, \sigma(Cyl))$ .

► On dispose d'un résultat affirmant que la tribu borélienne est égale à la tribu trace de  $D([0, 1], \mathbb{R})$  (voir par exemple Billingsley) :

$$\mathcal{B}(D([0, 1], \mathbb{R})) = \sigma(Cyl) \cap D([0, 1], \mathbb{R}).$$

Alors la fonction

$$(x, t) \in D([0, 1], \mathbb{R}) \times [0, 1] \mapsto p_t(x) = G(x(t)) \in [0, 1]$$

vérifie la condition 4.1.8 (et donc le processus  $p$ -valeur vérifie 4.1.4). En effet,

✓  $(x, t) \in D([0, 1], \mathbb{R}) \times [0, 1] \mapsto x(t)$  est mesurable comme limite simple d'une suite de fonctions mesurables :

$$f_n : (x, t) \mapsto f_n(x, t) := \sum_{k=1}^{2^n} x(k2^{-n}) \mathbf{1}\{(k-1)2^{-n} \leq t < k2^{-n}\} + x(1) \mathbf{1}\{t = 1\};$$

✓  $x \in \mathbb{R} \mapsto G(x) \in [0, 1]$  est mesurable car décroissante;

✓ donc  $p_t(x)$  est mesurable par composition d'applications mesurables.

## 3. $X$ est un processus Gaussien

Ici,  $\mathbb{X} := \mathbb{R}^{[0, 1]^d}$  est muni de la tribu  $\sigma(Cyl)$ .

On suppose de plus que la fonction de covariance  $\Sigma(t, t')$  vérifie les hypothèses suivantes :  $\Sigma$  est continue en  $(t, t')$  et  $\Sigma(t, t) := \sigma^2$  est connue.

Alors  $X$  est continue dans  $L^2$  donc  $X$  et en particulier continue en probabilité. On dispose alors d'un résultat qui affirme que  $X$  admet une modification telle que  $(x, t) \mapsto x(t)$  soit mesurable.

On en déduit comme précédemment que le processus de  $p$ -valeurs vérifie la condition de mesurabilité 4.1.4.

### Exemple 4.1.12 : Test sur la fonction de répartition

On observe  $X = (X_1, \dots, X_m) \in \mathbb{X} := \mathbb{R}^m$  un  $m$ -uplet de variables aléatoires réelles i.i.d. de même fonction de répartition continue  $F$  inconnue.

Le modèle statistique associé est donc  $(\mathbb{X}, \mathcal{X}, \mathcal{P}) = (\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m), \mathcal{P})$ . Pour  $\mathcal{H} = I$  un intervalle de  $\mathbb{R}$  et une fonction  $F_0$  connue, le problème de test considéré est :

$$H_t : "F(t) \leq F_0(t)" \quad \forall t \in I.$$

On note  $G_t(k) = \mathbb{P}(Z_t \geq k)$  pour tout  $k \in \mathbb{N}$  où  $Z_t \sim \mathcal{B}(m, F_0(t))$ .

On note aussi  $F_m(X, t) = \frac{1}{m} \sum_{i=1}^m \mathbf{1}\{X_i \leq t\}$  la fonction de répartition empirique de  $X_1, \dots, X_m$ .

Chaque hypothèse  $H_t$  peut être testée en utilisant la  $p$ -valeur :

$$p_t(X) = G_t(mF_m(X, t)).$$

Montrons que le processus  $(p_t(X))_{t \in I}$  définit un processus de  $p$ -valeurs mesurable.

► Définition 4.1.3.

Soit  $Y^{(p)} \sim \mathcal{B}(m, p)$ . Posons, pour  $k \in [0, m]$ ,  $f : p \in [0, 1] \mapsto \mathbb{P}(Y^{(p)} \geq k)$ . On a,

$$\begin{aligned}
f'(p) &= \sum_{j=k}^m \binom{m}{j} (jp^{j-1}(1-p)^{m-j} - p^j(m-j)(1-p)^{m-j-1}) \\
&= \sum_{j=k}^m \underbrace{\binom{m}{j} j}_{=m \binom{m-1}{j-1}} p^{j-1}(1-p)^{m-j} - \sum_{j=k}^{m-1} \underbrace{\binom{m}{j} (m-j)}_{=m \binom{m-1}{j}} p^j(1-p)^{m-j-1} \\
&= m \binom{m-1}{k-1} p^{k-1}(1-p)^{m-k} > 0
\end{aligned}$$

On a montré que  $f$  est croissante.

Si on note  $T_{P,t}(s) = \mathbb{P}_P(mF_m(X, t) \geq s)$ , on en déduit que, sous  $\mathcal{H}_0(P)$  (ie que pour tout  $t \in I$ ,  $F(t) \leq F_0(t)$ ),

$$T_{P,t}(mF_m(X, t)) = G_t(mF_m(X, t)).$$

On conclut donc, d'après 1.2.6, que

$$\forall t \in I, \quad \mathbb{P}_P(p_t(X) \leq u) \leq u.$$

► Hypothèse 4.1.4.

L'application

$$(x, t) \in \mathbb{R}^m \times I \mapsto p_t(x) = G_t(mF_m(x, t)) \in [0, 1]$$

est mesurable par composition d'applications mesurables. La condition 4.1.8 est donc vérifiée d'où 4.1.4.

## 4.1.2 False Discovery Rate.

Dans cette section, nous allons généraliser à un continuum d'hypothèses la *False Discovery Rate* définie en 2.1.3. Pour cela, nous allons quantifier la proportion de faux positifs par un rapport de volumes à l'aide d'une mesure finie  $\Lambda$  sur  $(\mathcal{H}, \mathfrak{H})$ .

### Définition 4.1.13 : False Discovery Proportion, False Discovery Rate

Soit  $\Lambda$  une mesure finie définie sur  $(\mathcal{H}, \mathfrak{H})$ . Soit  $R$  une procédure de tests multiples sur  $\mathcal{H}$ . La *False Discovery Rate* de  $R$  est définie comme l'espérance de la *False Discovery Proportion* (FDP) :

$$\forall P \in \mathcal{P}, \forall x \in X(\Omega), \quad \text{FDP}(R(x), P) := \frac{\Lambda(R(x) \cap \mathcal{H}_0(P))}{\Lambda(R(x))} \mathbf{1}_{\{\Lambda(R(x)) > 0\}}$$

et

$$\forall P \in \mathcal{P}, \quad \text{FDR}(R(X), P) := \mathbb{E}_P[\text{FDP}(R(X), P)].$$

Dans le cas d'un nombre fini d'hypothèses, on retrouve les définitions de FDP et FDR présentées dans la section 2.1.3 en prenant  $\Lambda$  égale à la mesure de comptage.

On remarque également que, compte tenu de l'hypothèse de mesurabilité pour la définition de  $R$  (définition 4.1.2), les quantités précédentes sont bien définies.

## 4.1.3 Procédures Step-Up.

Dans la suite, nous allons nous intéresser à une forme particulière de procédure définie à partir de la famille de  $p$ -valeurs  $\mathbf{p}(x) = (p_h(x))_{h \in \mathcal{H}}$ .

On commence par définir une famille de zones de rejet basée sur une *fonction seuil*  $\Delta : (h, r) \in \mathcal{H} \times \mathbb{R}^+ \mapsto \Delta(h, r) \in \mathbb{R}^+$  :

$$\forall r \geq 0, \forall x \in \mathbb{X}, \quad L_\Delta(r, x) := \{h \in \mathcal{H}, p_h(x) \leq \Delta(h, r)\}.$$

On s'intéressera plus particulièrement aux seuils de la forme  $\Delta(h, r) = \alpha\pi(h)\beta(r)$  où  $\alpha \in (0, 1)$  est le *niveau* du test multiple,  $\pi : \mathcal{H} \rightarrow \mathbb{R}^+$  est mesurable et est appelée *fonction poids*, et  $\beta : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  une fonction croissante et continue à droite (*shape function*).

En choisissant convenablement le paramètre  $r$  comme une fonction de l'observation  $x = X(\omega)$ , on détermine une zone de rejet.

**Exemple 4.1.14 : Procédure Step-Up.**

Soit  $\Delta(h, r) = \alpha\pi(h)\beta(r)$  une fonction seuil avec  $\alpha \in (0, 1)$ ;  $\pi : \mathcal{H} \rightarrow \mathbb{R}^+$  une fonction mesurable et  $\beta : \mathbb{R}^+ \rightarrow \mathbb{R}^+$  une fonction croissante et continue à droite. On considère la procédure de tests multiples *step-up*  $R$  sur  $(\mathcal{H}, \Lambda)$  associée à la famille de seuils  $\Delta$ , définie par

$$\forall x \in X(\Omega), \quad R(x) = L_\Delta(x, \hat{r}(x)) \quad \text{où } \hat{r}(x) := \max\{r \geq 0 : \Lambda(L_\Delta(x, r)) \geq r\}.$$

On remarque que, compte tenu de l'hypothèse 4.1.4 sur les  $p$ -valeurs,  $\hat{r}$  est une quantité bien définie. En effet, d'après 4.1.4, comme  $x \in X(\Omega)$ , la fonction  $h \mapsto p_h(x) - \alpha\pi(h)\beta(r)$  est mesurable. Donc  $L_\Delta(x, r)$  est mesurable sur  $\mathcal{H}$  et donc  $\Lambda(L_\Delta(x, r))$  est bien définie. De plus, 0 appartient à l'ensemble  $\{r \geq 0 : \Lambda(L_\Delta(x, r)) \geq r\}$  et  $M = \Lambda(\mathcal{H})$  en est une borne supérieure. Enfin, cette borne supérieure est un maximum car la fonction  $h \mapsto \Lambda(L_\Delta(x, r))$  est croissante.

**Exemple 4.1.15 : Procédure de Benjamini Hochberg**

Dans le cas où  $\mathcal{H}$  est fini (disons que  $\mathcal{H} = \{1, \dots, m\}$ ),  $\Lambda$  est la mesure de comptage,  $\beta(r) = r$  et  $\pi(h) = 1/m$ , on retrouve la procédure de Benjamini Hochberg définie en 2.3.1.

En effet, il suffit pour cela de remarquer (en raisonnant par double inclusion et en reprenant les notations de la section 2.3.1) que

$$\{l \in \{1, \dots, m\} : p_{(l)} \leq \frac{\alpha l}{m}\} = \{r \in \{1, \dots, m\} : |\{h \in \mathcal{H} : p_h \leq \frac{\alpha r}{m}\}| \geq r\}.$$

Dans le cas où  $\mathcal{H}$  est infini non dénombrable, il faut montrer que la procédure step-up de tests multiples  $R$  définie par 4.1.14 satisfait les bonnes conditions de mesurabilité.

**Lemme 4.1.16**

La procédure step-up définie par 4.1.14 est une procédure de tests multiples au sens de 4.1.2.

*Démonstration.*

Vérifions que l'application

$$(\omega, h) \mapsto \mathbf{1}\{h \in R(X(\omega))\} = \mathbf{1}\{p_h(X(\omega)) \leq \alpha\pi(h)\beta(\hat{r}(X(\omega)))\}$$

est mesurable.

D'après l'hypothèse 4.1.4 et comme  $\beta$  et  $\pi$  sont mesurables, il suffit de montrer que  $\omega \mapsto \hat{r}(X(\omega))$  est mesurable.

Pour tout  $x \in X(\Omega)$ , on considère la fonction

$$f : r \in \mathbb{R}^+ \mapsto \Lambda(L_\Delta(x, r)) = \int_{\mathcal{H}} \mathbf{1}\{p_h(x) \leq \alpha\pi(h)\beta(r)\} d\Lambda(h).$$

On remarque d'abord que  $f$  est continue à droite et décroissante (par hypothèses sur  $\beta$ ) et bornée par  $\hat{r} = \max\{r \geq 0 : f(r) \geq r\}$ . On peut aussi montrer que, pour tout  $x \in X(\Omega)$ ,

$$\begin{aligned} \hat{r}(x) &= \inf_{\varepsilon > 0, \varepsilon \in \mathbb{Q}} \sup\{r \in \mathbb{Q}^+ : \Lambda(L_\Delta(x, r)) \geq r - \varepsilon\} \\ &= \inf_{\varepsilon > 0, \varepsilon \in \mathbb{Q}} \sup(r \mathbf{1}\{\Lambda(L_\Delta(x, r)) \geq r - \varepsilon\}). \end{aligned}$$

Or d'après 4.1.4, pour tout  $\varepsilon > 0, \varepsilon \in \mathbb{Q}$  et  $r \in \mathbb{Q}^+$ , la fonction

$$\omega \mapsto r \mathbf{1}\{\Lambda(L_\Delta(X(\omega), r)) \geq r - \varepsilon\} = r \mathbf{1}\{\Lambda(\{h \in \mathcal{H} : p_h(X(\omega)) \leq \alpha\pi(h)\beta(r)\}) \geq r - \varepsilon\}$$

est mesurable, et l'expression de  $\hat{r}$  précédente implique que  $\omega \mapsto \hat{r}(X(\omega))$  est mesurable.

En conclusion, la procédure step-up  $R$  satisfait l'hypothèse de mesurabilité de la définition 4.1.2.  $\square$

#### 4.1.4 Condition de Positively Regressively Dependent on each one of a Subset.

Comme nous l'avons vu dans le second chapitre en 2.3.2, le contrôle de la FDR est assuré par des hypothèses sur la dépendance entre les  $p$ -valeurs. On rappelle également que si le cas d'indépendance des  $p$ -valeurs est le cadre "idéal" pour  $\mathcal{H}$  fini, dans notre contexte où  $\mathcal{H}$  est infini non dénombrable, on ne peut pas supposer l'indépendance mutuelle des  $p$ -valeurs sans contredire l'hypothèse de mesurabilité (d'après 4.1.6).

Dans cette section, nous allons généraliser la condition de dépendance positive, ou PRDS (voir 2.3.5) au cas d'une infinité d'hypothèses.

On rappelle que pour tout ensemble fini  $\mathcal{I}$ , un sous-ensemble  $D \subset [0, 1]^{\mathcal{I}}$  est dit *croissant* si pour tout  $\mathbf{z}, \mathbf{z}' \in [0, 1]^{\mathcal{I}}$  tels que  $\mathbf{z} \leq \mathbf{z}'$  (ie  $\forall h \in \mathcal{I}, z_h \leq z'_h$ ), on a  $\mathbf{z} \in D \Rightarrow \mathbf{z}' \in D$ .

**Définition 4.1.17 :** *Condition de PRDS finie dimensionnelle pour un processus de  $p$ -valeurs.*

Pour  $\mathcal{H}'$  un sous-ensemble de  $\mathcal{H}$ , le processus  $p$ -valeur  $\mathbf{p}(X) = (p_h(X))_{h \in \mathcal{H}}$  vérifie la condition de  $w$ PRDS finie dimensionnelle sur  $\mathcal{H}'$  (respectivement la condition de PRDS (forte) finie dimensionnelle sur  $\mathcal{H}'$ ) pour la loi  $P$ , si pour tout sous-ensemble fini  $S$  de  $\mathcal{H}$ , la famille finie de  $p$ -valeur  $\mathbf{p}_S(X) = (p_h(X))_{h \in \mathcal{H} \cap S}$  vérifie la condition de  $w$ PRDS sur  $\mathcal{H}' \cap S$  (respectivement la condition de PRDS (forte) sur  $\mathcal{H}' \cap S$ ) pour la loi  $P$ .

La propriété de  $w$ PRDS finie dimensionnelle sera suffisante pour prouver le contrôle de la FDR mais il sera parfois plus simple de vérifier la propriété "forte" de PRDS. En effet, on a prouvé que la propriété de PRDS forte implique la propriété de PRDS faible (cf 2.3.6).

**Exemple 4.1.18 :** *Processus gaussien*

Soit  $\mathbf{p}(X) = (p_h(X))_{h \in \mathcal{H}}$  un processus de  $p$ -valeurs de la forme  $p_h(X) = G(X_h)$ ,  $h \in \mathcal{H}$ , où  $X = (X_h)_{h \in \mathcal{H}}$  est un processus gaussien et où  $G$  est une fonction continue et décroissante de  $\mathbb{R}$  dans  $[0, 1]$ . On suppose que la fonction de covariance  $\Sigma$  de  $X$  vérifie

$$\forall h, h' \in \mathcal{H}, \quad \Sigma(h, h') \geq 0.$$

Alors le processus de  $p$ -valeurs satisfait la condition de PRDS finie dimensionnelle forte.

La preuve découle de l'exemple 2.3.7.

**Exemple 4.1.19 :** *Fonction de répartition*

Dans l'exemple 4.1.12, nous avons introduit un processus de  $p$ -valeurs défini sur  $I$  un intervalle de  $\mathbb{R}$ ,

$$p_t(X) = G_t(mF_m(X, t)),$$

où  $G_t(k) = \mathbb{P}(Z_t \geq k)$  pour tout  $k \in [0, m]$  avec  $Z_t \sim \mathcal{B}(m, F_0(t))$  et  $F_m(X, t) = \frac{1}{m} \sum_{i=1}^m \mathbb{1}\{X_i \leq t\}$  la fonction de répartition empirique de l'observation  $X_1, \dots, X_m$ .

Montrons que ce processus de  $p$ -valeurs est  $w$ PRDS finie dimensionnelle.

*Démonstration.*

Soit  $(t_j)_{0 \leq j \leq N-1}$  un ensemble fini de points de  $I$  et  $D$  un ensemble mesurable croissant de  $[0, 1]^N$ . Nous allons montrer que la fonction

$u \in [0, 1] \mapsto \mathbb{P}(\mathbf{p}(X) \in D \mid p_{t_0}(X) \leq u)$  est croissante sur  $\{u \in [0, 1] : \mathbb{P}(p_{t_0}(X) \leq u) > 0\}$ .

Si  $F(t_0) \in \{0, 1\}$  alors le résultat est clair.

Supposons à présent que  $F(t_0) \in (0, 1)$ . Posons

$$\mathcal{U}_{t_0} = \{G_{t_0}(k), k = m, m-1, \dots, 0\},$$

un ensemble qui contient alors des points croissants de  $(0, 1]$ . Il suffit donc de montrer la propriété de monotonie pour  $u \in \mathcal{U}_{t_0}$ .

Comme  $G_{t_0}$  est décroissante de  $\{0, \dots, m\}$  dans  $\mathcal{U}_{t_0}$ , on a

$$p_{t_0}(X) \leq G_{t_0}(k) \iff mF_m(X, t_0) \geq k \iff X_{(k)} \leq t_0 \text{ en posant } X_{(0)} = -\infty.$$

On doit donc prouver que pour tout  $1 \leq k \leq m$ ,

$$\mathbb{P}((X_{(1)}, \dots, X_{(m)}) \in D' \mid X_{(k-1)} \leq t_0) \geq \mathbb{P}((X_{(1)}, \dots, X_{(m)}) \in D' \mid X_{(k)} \leq t_0),$$

où  $D' = \{x \in \mathbb{R}^m : (p_{t_j}(x))_{0 \leq j \leq N-1} \in D\}$  est un ensemble croissant de  $\mathbb{R}^m$  (car  $\mathbf{p}$  vérifie  $x \leq x' \Rightarrow \forall t, p_t(x) \leq p_t(x')$ ).

On peut alors montrer que la famille des statistiques d'ordre  $\{X_{(i)}\}_i$  vérifie la condition de monotonie suivante : la fonction  $f$  définie par

$$f(a, b) = \mathbb{E}[(X_{(1)}, \dots, X_{(m)}) \in D' \mid X_{(k-1)} = a, X_{(k)} = b]$$

est croissante en  $a$  et  $b$ .

Ainsi, en notant,  $\gamma = \mathbb{P}(X_{(k)} \leq t_0 \mid X_{(k-1)} \leq t_0)$ , on a (par un raisonnement similaire à la preuve de 2.3.6),

$$\begin{aligned} \mathbb{P}[(X_{(1)}, \dots, X_{(m)}) \in D' \mid X_{(k-1)} \leq t_0] &= \gamma \mathbb{E}[f(X_{(k-1)}, X_{(k)}) \mid X_{(k-1)} \leq t_0, X_{(k)} \leq t_0] \\ &\quad + (1 - \gamma) \mathbb{E}[f(X_{(k-1)}, X_{(k)}) \mid X_{(k-1)} \leq t_0 < X_{(k)}] \\ &\geq \mathbb{E}[f(X_{(k-1)}, X_{(k)}) \mid X_{(k-1)} \leq t_0, X_{(k)} \leq t_0], \end{aligned}$$

ce qui montre la propriété de monotonie et conclut la preuve.  $\square$

## 4.2 Contrôle de la FDR.

### 4.2.1 Énoncé du théorème et exemples.

Le théorème suivant établit le contrôle de la FDR à un niveau  $\alpha$  fixé pour des procédures step-up définies comme dans l'exemple 4.1.14 pour deux situations différentes :

- une situation où la dépendance entre les  $p$ -valeurs est totalement arbitraire ;
- une situation où une dépendance positive est supposée entre les  $p$ -valeurs (condition PRDS).

#### Théorème 4.2.1

On suppose que l'espace  $\mathcal{H}$  vérifie l'hypothèse 4.1.1 et qu'il est muni d'une mesure finie  $\Lambda$ .

On considère  $\mathbf{p}(X) = (p_h(X))_{h \in \mathcal{H}}$  un processus de  $p$ -valeurs satisfaisant la condition 4.1.4.

On appelle  $R$  la procédure step-up sur  $(\mathcal{H}, \Lambda)$  associée à la fonction seuil  $\Delta(h, r) = \alpha \pi(h) \beta(r)$ , avec  $\alpha \in (0, 1)$ ,  $\beta$  une fonction croissante continue à droite et  $\pi$  une densité de probabilité sur  $\mathcal{H}$  par rapport à  $\Lambda$ . On suppose que l'une des deux hypothèses suivantes est vraie. Pour tout  $P \in \mathcal{P}$ ,

1.  $\beta(x) = x$  et le processus de  $p$ -valeurs  $\mathbf{p}$  satisfait la condition de  $w$ PRDS finie dimensionnelle sur  $\mathcal{H}_0(P)$  pour la loi  $P$  ;
2. la fonction  $\beta$  est de la forme

$$\beta_\mu(x) = \int_0^x u d\mu(u), \quad (b)$$

où  $\mu$  est une loi de probabilité arbitraire sur  $(0, \infty)$ .

Alors

$$FDR(R, P) \leq \alpha \Pi(\mathcal{H}_0(P)) \leq \alpha,$$

$$\text{où } \Pi(\mathcal{H}_0(P)) := \int_{h \in \mathcal{H}_0(P)} \pi(h) d\Lambda(h).$$

Lorsque l'on cherche une procédure de tests multiples, on veut d'une part contrôler l'erreur de première espèce et d'autre part avoir une grande "puissance". Comme les notions de puissance vues en section 3.1.2 utilisent l'ensemble  $R \cap \mathcal{H}_1(P)$ , on souhaite que la procédure rejette un volume le plus large possible d'hypothèses. Dans cette optique, choisir une procédure step-up avec  $\beta(x) = x$  conduit toujours à des procédures plus performantes qu'une procédure step-up où  $\beta$  est de la forme (b). En effet, on a  $\int_0^x u d\mu(u) \leq x$ .

Ainsi, la condition de PRDS dans le théorème offre un résultat moins conservatif que la condition de dépendance arbitraire.

#### Exemple 4.2.2 : Test sur la moyenne d'un processus

On reprend l'exemple 4.1.11 dans le cas où l'observation  $X = (X_t)_{t \in [0,1]^d}$  est celle d'un processus gaussien de fonction moyenne mesurable  $\mu$ , de fonction de covariance  $\Sigma$  et de variance connue (et égale à 1). On rappelle que le problème est de tester pour tout  $t \in [0, 1]^d$  l'hypothèse  $H_t$  : " $\mu(t) \leq 0$ ".

On considère  $\Lambda$  la mesure de Lebesgue  $d$ -dimensionnelle. Soit  $\alpha \in (0, 1)$ ,  $\beta$  une fonction croissante et continue à droite. On note également  $G$  la queue de distribution d'une loi normale standard et

$$\hat{r}(X) = \max\{r \in [0, 1] : \Lambda(\{t \in [0, 1]^d : G(X_t) \leq \alpha\beta(r)\}) \geq r\}.$$

Alors on a le contrôle de la FDR au niveau  $\alpha$  de la procédure step-up associée à la fonction  $\beta$  et à la densité  $\pi(h) = 1$ ,

$$\mathbb{E} \left[ \frac{\Lambda(\{t \in [0, 1]^d : \mu(t) \leq 0, G(X_t) \leq \alpha\beta(\hat{r}(X))\})}{\Lambda(\{t \in [0, 1]^d : G(X_t) \leq \alpha\beta(\hat{r}(X))\})} \right] \leq \alpha,$$

si  $(X_t)_{t \in [0,1]^d}$  est un processus mesurable et si une des deux conditions suivantes est vérifiée :

- $\beta(x) = x$  et la fonction de covariance du processus est positive (d'après 4.1.18) ;
- $\beta$  est de la forme (b).

#### Exemple 4.2.3 : Test sur la fonction de répartition

On reprend l'exemple 4.1.12 dans lequel on veut tester " $F(t) \leq F_0(t)$ " pour tout  $t$  d'un intervalle  $I \subset \mathbb{R}$ . L'exemple 4.1.19 affirme que le processus de  $p$ -valeurs

$$p_t(X) = G_t(mF_m(X, t))$$

vérifie la condition de  $w$ PRDS finie dimensionnelle. Ainsi le théorème 4.2.1 s'applique et la procédure de step-up associée contrôle la FDR.

## 4.2.2 Démonstration du théorème 4.2.1.

Cette preuve repose sur trois lemmes et une proposition.

**Lemme 4.2.4 :** Deux conditions qui assurent le contrôle de la FDR - Blanchard, Roquain (2008) [4]

Pour tout  $P \in \mathcal{P}$ , supposons que les deux hypothèses suivantes sont vérifiées :

1. La procédure de tests multiples  $R$  satisfait la "self-consistency condition" :

$$R(x) \subset \{h \in \mathcal{H} : p_h(x) \leq \alpha\pi(h)\beta(\Lambda(R(x)))\} \text{ pour } P\text{-presque tout } x \in \mathcal{X} \quad (\text{SC}(\alpha, \pi, \beta))$$

2. Pour tout  $h \in \mathcal{H}_0(P)$ , le couple de variables aléatoires  $(U_h, V) := (p_h(X), \Lambda(R(X)))$  satisfait la "condition de contrôle de la dépendance" :

$$\forall c > 0, \quad \mathbb{E} \left[ \frac{\mathbb{1}\{U_h \leq c\beta(V)\}}{V} \mathbb{1}\{V > 0\} \right] \leq c. \quad (\text{DC}(\beta))$$

*Démonstration.*

On a

$$\begin{aligned} \text{FDR}(R, P) &= \mathbb{E} \left[ \frac{\Lambda(R \cap \mathcal{H}_0(P))}{\Lambda(R)} \mathbb{1}\{\Lambda(R) > 0\} \right] \\ &= \mathbb{E} \left[ \int_{h \in \mathcal{H}_0(P)} \frac{\mathbb{1}\{h \in R\}}{\Lambda(R)} \mathbb{1}\{\Lambda(R) > 0\} d\Lambda(h) \right] \\ &= \int_{h \in \mathcal{H}_0(P)} \mathbb{E} \left[ \frac{\mathbb{1}\{h \in R\}}{\Lambda(R)} \mathbb{1}\{\Lambda(R) > 0\} \right] d\Lambda(h) \text{ d'après le théorème de Fubini } (R \text{ étant mesurable par 4.1.2}) \\ &\leq \int_{h \in \mathcal{H}_0(P)} \mathbb{E} \left[ \frac{\mathbb{1}\{p_h \leq \alpha\pi(h)\beta(\Lambda(R))\}}{\Lambda(R)} \mathbb{1}\{\Lambda(R) > 0\} \right] d\Lambda(h) \text{ d'après SC}(\alpha, \pi, \beta) \\ &\leq \alpha \int_{h \in \mathcal{H}_0(P)} \pi(h) d\Lambda(h) \text{ d'après DC}(\beta). \end{aligned}$$

□

Nous allons à présent montrer la procédure step-up définie auparavant satisfait les deux conditions de ce lemme.

**Lemme 4.2.5 :** Condition  $\text{SC}(\alpha, \pi, \beta)$

La procédure step-up définie précédemment satisfait la condition  $\text{SC}(\alpha, \pi, \beta)$ .

*Démonstration.*

On a déjà montré que cette procédure step-up fournit une procédure de tests multiples au sens de 4.1.2.

D'après la définition de la procédure step-up, pour tout  $\varepsilon > 0$ ,

$$\begin{aligned} \Lambda(L_\Delta(\hat{r})) &\leq \Lambda(L_\Delta(\hat{r} + \varepsilon)) \text{ par décroissance de la fonction } \beta \\ &< \hat{r} + \varepsilon \text{ car } \hat{r} < \hat{r} + \varepsilon \text{ et par définition de } \hat{r}(\varepsilon). \end{aligned}$$

On en déduit alors que  $\Lambda(L_\Delta(\hat{r})) \leq \hat{r}$ . De plus, d'après la définition de  $\hat{r}$ ,  $\Lambda(L_\Delta(\hat{r})) \geq \hat{r}$ . Ainsi  $\Lambda(L_\Delta(\hat{r})) = \hat{r}$ .

Enfin, la procédure step-up  $R$  satisfait la condition  $\text{SC}(\alpha, \pi, \beta)$  avec égalité.

□

**Lemme 4.2.6 :** Conditions pour que l'hypothèse  $\text{DC}(\beta)$  soit vérifiée - Blanchard, Roquain (2008) [4]

Soit  $(U, V)$  un couple de variables aléatoires positives et telle que  $U$  soit minorée en probabilité par une variable aléatoire uniforme sur  $[0, 1]$  (ie  $\forall t \in [0, 1], \mathbb{P}(U \leq t) \leq t$ ).

Alors la condition du contrôle de dépendance  $\text{DC}(\beta)$  est satisfaite par  $(U, V)$  si une des deux hypothèses suivantes est vérifiée :

(i)  $\beta(r) = r$  et

(v)  $\rho(u) = u \wedge \varepsilon$

$$\forall r \in \mathbb{R}^+, u \mapsto \mathbb{P}(V \leq r \mid U \leq u)$$

est une application croissante sur  $\{u : P(U \leq u) > 0\}$ .

(ii) La fonction  $\beta$  est de la forme (b).

*Démonstration.*

(i) Soit  $\varepsilon > 0$ ,  $\rho \in (0, 1)$ .

On choisit  $K$  suffisamment grand pour que  $\rho^k < \varepsilon$ .

Posons  $v_0 = 0$  et  $v_i = \rho^{K+1-i}$  pour tout  $1 \leq i \leq 2K + 1$ .

$$\begin{aligned} \mathbb{E} \left[ \frac{\mathbb{1}\{U \leq cV\}}{V \wedge \varepsilon} \right] &\leq \sum_{i=1}^{2K+1} \frac{\mathbb{P}(U \leq cv_i, V \in [v_{i-1}, v_i])}{v_{i-1} \wedge \varepsilon} + \varepsilon \\ &\leq \sum_{i=1}^{2K+1} \frac{\mathbb{P}(U \leq cv_i, V \in [v_{i-1}, v_i])}{\mathbb{P}(U \leq cv_i)} \frac{\mathbb{P}(U \leq cv_i)}{v_{i-1} \wedge \varepsilon} + \varepsilon \\ &\leq \frac{c}{\rho} \sum_{i=1}^{2K+1} \mathbb{P}(V \in [v_{i-1}, v_i] \mid U \leq cv_i) + \varepsilon \\ &= \frac{c}{\rho} \sum_{i=1}^{2K+1} (\mathbb{P}(V < v_i \mid U \leq cv_i) - \mathbb{P}(V < v_{i-1} \mid U \leq cv_i)) + \varepsilon \\ &\leq \frac{c}{\rho} \underbrace{\sum_{i=1}^{2K+1} (\mathbb{P}(V < v_i \mid U \leq cv_i) - \mathbb{P}(V < v_{i-1} \mid U \leq cv_{i-1}))}_{= -\mathbb{P}(V < v_0 \mid U \leq cv_0) + \mathbb{P}(V < v_{2K+1} \mid U \leq cv_{2K+1})} + \varepsilon \\ &\leq \frac{c}{\rho} + \varepsilon \quad \text{car } v_0 = 0. \end{aligned}$$

On fait alors tendre  $\rho$  vers 1 et par le théorème de convergence dominée, on fait tendre  $\varepsilon$  vers 0. On obtient le résultat voulu.

(ii) Pour tout  $z > 0$ , on remarque que  $\frac{1}{z} = \int_0^{+\infty} \frac{1}{v^2} \mathbb{1}\{v \geq z\} dv$ .

Alors :

$$\begin{aligned} \mathbb{E} \left[ \frac{\mathbb{1}\{U \leq c\beta(V)\}}{V} \right] &= \mathbb{E} \left[ \int_0^{+\infty} \frac{1}{v^2} \mathbb{1}\{v \geq V\} \mathbb{1}\{U \leq c\beta(v)\} dv \right] \\ &= \int_0^{+\infty} \frac{1}{v^2} \mathbb{E} [\mathbb{1}\{v \geq V\} \mathbb{1}\{U \leq c\beta(v)\}] dv \\ &\leq \int_0^{+\infty} \frac{1}{v^2} \mathbb{P}(U \leq c\beta(v)) dv \\ &\leq c \int_0^{+\infty} \frac{\beta(v)}{v^2} dv \\ &= c \int_0^{+\infty} \left( \int_0^v \frac{u}{v^2} d\mu(u) \right) dv \\ &= c \int_0^{+\infty} \int_0^\infty \frac{u}{v^2} \mathbb{1}\{u \leq v\} d\mu(u) dv \\ &= c \int_{u \geq 0} u \int_0^{+\infty} \frac{\mathbb{1}\{u \leq v\}}{v^2} dv d\mu(u) \quad \text{d'après Fubini-Tonelli} \\ &= c. \end{aligned}$$

□

A ce stade de la démonstration, le point (ii) du lemme précédent démontre le point 2. du théorème 4.2.1.

Pour démontrer le point 1. du théorème, il reste à prouver que la condition (i) du lemme précédent est vérifiée dans le cas de la  $w$ PRDS finie dimensionnelle.



**Proposition 4.2.7 :** *La condition de wPRDS finie dimensionnelle implique la condition (ii) du lemme*

On suppose que le processus de  $p$ -valeurs  $\mathbf{p} = (p_h, h \in \mathcal{H})$  satisfait la wPRDS finie dimensionnelle sur  $\mathcal{H}_0(P)$  pour tout  $P \in \mathcal{P}$ . On considère  $R$ , la procédure de step-up définie par 4.1.14 avec  $\beta(x) = x$ .

Alors, pour tout  $P \in \mathcal{P}$ , pour tout  $h \in \mathcal{H}_0(P)$ , le couple de variables aléatoires  $(U_h, V) := (p_h, \Lambda(R))$  satisfait (ii). En particulier, l'hypothèse DC( $\beta$ ) est vérifiée pour  $\beta(x) = x$ .

*Démonstration. (Idée de la démonstration)*

La preuve de cette proposition s'appuie sur le fait que la procédure step-up définie dans le cas continu par 4.1.14 peut être vue comme une limite de procédures step-up finies. Cette idée est résumée par le lemme suivant qui sera admis.

**Lemme 4.2.8 :** *Approximation finie de la procédure step-up*

On considère la procédure step-up  $R = L_\Delta(\hat{r})$  sur  $\mathcal{H}$  définie par 4.1.14.

Alors il existe une suite de mesures  $\Lambda_n$  de support finie sur  $\mathcal{H}$  telle que, si l'on note

$$\hat{r}_{n,k} = \max\{r \geq 0 : \Lambda_n(L_\Delta(r)) \geq r - \frac{1}{k}\},$$

on a,

$$\hat{r} = \lim_{k \rightarrow \infty} \left( \limsup_{n \rightarrow \infty} \hat{r}_{n,k} \right) = \lim_{k \rightarrow \infty} \left( \liminf_{n \rightarrow \infty} \hat{r}_{n,k} \right) \text{ presque sûrement.}$$

□

**Remarque 4.2.9 :** *Commentaires sur les conditions de lemme 4.2.4*

La condition SC( $\alpha, \pi, \beta$ ) est purement algorithmique et dépend de la manière dont est construit la procédure  $R$  à partir des observations. Intuitivement, quand le volume  $|R|$  devient grand, on commet plus d'erreurs de type I, on doit donc augmenter le seuil  $\Delta(h, r)$  dans la définition de  $L_\Delta(r) = \{h \in \mathcal{H} : p_h \leq \Delta(h, r)\}$ . Ainsi, le seuil doit être une fonction croissante en  $r$ . Cela justifie le fait que  $R \subset L_\Delta(|R|)$ .

La condition DC( $\beta$ ) est une condition qui caractérise les relations de dépendance des  $p$ -valeurs  $(p_h(X))_{h \in \mathcal{H}}$ .

Il faut noter que ces deux conditions sont liées car elles dépendent toutes deux de la fonction  $\beta$ . De plus, elles sont suffisantes pour contrôler la FDR mais pas nécessaires.

## 4.3 Conclusion.

Ce rapport propose un résumé des principaux résultats d'optimalité sur les tests multiples d'un nombre fini d'hypothèses. Ainsi, nous nous sommes concentrés sur le contrôle de l'erreur de première espèce et sur une définition d'un critère lié à l'erreur de seconde espèce permettant d'obtenir des résultats d'optimalité de type minimax.

Nous avons aussi généralisé la définition de tests multiples au cas d'un nombre infini non dénombrable d'hypothèses en nous posant les questions liées aux définitions de chaque objet (essentiellement des contraintes de mesurabilité). Dans la même optique que dans le cas d'un nombre fini d'hypothèses, nous avons également proposé une définition de la FDR pour un continuum d'hypothèses ainsi qu'un résultat assurant son contrôle. Plusieurs perspectives d'études sont envisagées afin de généraliser une approche "optimale" (en un sens qui n'est pas défini!) des tests d'une infinité non dénombrable d'hypothèses. Pour cela, il faudrait définir un critère de type "FWER" pour un continuum d'hypothèses afin de définir un nouveau critère lié à l'erreur de seconde espèce permettant d'étendre une théorie de type minimax comme dans le troisième chapitre présenté ici. Pour cela, nous allons nous concentrer sur des tests multiples d'homogénéité d'un processus de Poisson, pour lesquels des résultats ont déjà été établis (notamment une définition et le contrôle de la FWER) par F. Picard, P. Reynaud-Bouret et E. Roquain [11]. Nous espérons pouvoir approfondir l'étude de l'erreur de seconde espèce au travers d'une approche minimax.



# Bibliographie

- [1] Y. Baraud, Non-asymptotic minimax rates of testing in signal detection, *Bernoulli*, 2002.
- [2] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate : a practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society*, 1995.
- [3] G. Blanchard, S. Delattre, E. Roquain, Testing over a continuum of null hypotheses with False Discovery Rate control, *Bernoulli*, 2014.
- [4] G. Blanchard, E. Roquain, Two simple sufficient conditions for FDR control, *Electron. J. Stat*, 2008.
- [5] J. Chen et al, On power and sample size computation for multiple testing procedures, *Computational Statistics and Data Analysis*, 2009.
- [6] M. Fromont, M. Lerasle, P. Reynaud-Bouret, Family wise separation rates for multiple testing, *Annals of Statistics, Institute of Mathematical Statistics*, 2016.
- [7] J. Goeman et A. Solari, The sequential rejection principle of familywise error control, *The Annals of Statistics*, 2010.
- [8] Y. Hochberg, A. Tamhane, Multiple Comparison Procedures, *Wiley, New-York*, 1987.
- [9] G. Izmirlian, Average Power and  $\lambda$ -power in Multiple Testing Scenarios when the Benjamini-Hochberg False Discovery Rate Procedure is Used, *arXiv : 1801.03989*, 2018.
- [10] E.L. Lehmann, J.P. Romano, J.P. Shaffer, On optimality of stepdown and stepup multiple test procedures, *The Annals of Statistics*, 2005.
- [11] F. Picard, P. Reynaud-Bouret, E. Roquain, Continuous testing for Poisson process intensities : A new perspective on scanning statistics, *arXiv :1705.08800v1*, 2017.
- [12] D. Revuz, M. Yor, Continuous martingales and Brownian motion, *Springer*, 1991.
- [13] J.P. Romano, A. Shaikh, M. Wolf, Consonance and the closure method in multiple testing, *The International Journal of Biostatistics*, 2011.
- [14] E. Roquain, Type I error rate control for testing many hypotheses : a survey with proofs, *Journal de la Société Française de la Statistique*, 2011.
- [15] E. Roquain, On controlling the amount of false positives when making multiple simultaneous tests, *Summer school in Anger*, 2016.
- [16] J.D Storey, The optimal discovery procedure : A new approach to simultaneous significance testing, *UW Biostatistics Working Paper series*, 2005.
- [17] P.H. Westfall, R.D. Tobias, R.D. Wolfinger, *Multiple Comparisons and Multiple Tests Using SAS, Second Edition*, SAS Publishing, 2011.