

Université de Montréal

Approximation du calcul de la taille
échantillonnale pour les tests à hypothèses
multiples lorsque r parmi m hypothèses doivent
être significatives

par

Philippe Delorme

Département de mathématiques et de statistique
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en statistique

Décembre 2012

Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

**Approximation du calcul de la taille
échantillonnale pour les tests à hypothèses
multiples lorsque r parmi m hypothèses doivent
être significatives**

présenté par

Philippe Delorme

a été évalué par un jury composé des personnes suivantes :

Martin Bilodeau

(président-rapporteur)

Pierre Lafaye de Michaux

(directeur de recherche)

François Perron

(membre du jury)

Mémoire accepté le:

20 décembre 2012

SOMMAIRE

Généralement, dans les situations d'hypothèses multiples on cherche à rejeter toutes les hypothèses ou bien une seule d'entre d'elles. Depuis quelques temps on voit apparaître le besoin de répondre à la question : « Peut-on rejeter au moins r hypothèses ? ». Toutefois, les outils statistiques pour répondre à cette question sont rares dans la littérature. Nous avons donc entrepris de développer les formules générales de puissance pour les procédures les plus utilisées, soit celles de *Bonferroni*, de *Hochberg* et de *Holm*. Nous avons développé un *package R* pour le calcul de la taille échantillonnale pour les tests à hypothèses multiples (*multiple endpoints*), où l'on désire qu'au moins r des m hypothèses soient significatives. Nous nous limitons au cas où toutes les variables sont continues et nous présentons quatre situations différentes qui dépendent de la structure de la matrice de variance-covariance des données.

Mots clés : Taille d'échantillon, Puissance, Hypothèses multiples.

SUMMARY

Generally, in multiple endpoints situations we want to reject all hypotheses or at least only one of them. For some time now, we see emerge the need to answer the question : "Can we reject at least r hypotheses ?" However, the statistical tools to answer this new problem are rare in the litterature. We decide to develop general power formulas for the principals procedures : *Bonferroni's*, *Hochberg's* and *Holm's* procedures. We also develop an R package for the sample size calculation for multiple endpoints, when we want to reject at least r hypotheses. We limit ourselves in the case where all the variables are continuous and we present four different situations depending on the structure of the data's variance-covariance matrix.

Keywords : Sample size, Power, Multiple endpoints.

TABLE DES MATIÈRES

Sommaire	v
Summary	vii
Liste des figures	xiii
Liste des tableaux	xv
Remerciements	1
Introduction	3
Chapitre 1. Définitions	7
1.1. Risques d'erreur à contrôler.....	8
1.1.1. Taux d'erreurs par comparaisons (<i>PCER</i>).....	9
1.1.2. Taux d'erreurs par-famille (<i>PFER</i>).....	10
1.1.3. <i>Familywise error rate</i> (<i>FWER</i>).....	10
1.1.4. <i>Generalized familywise error rate</i> (gFWER).....	12
1.1.5. Taux de fausses découvertes (<i>FDR</i>).....	12
1.1.6. Taux positif de fausses découvertes (<i>pFDR</i>).....	12
1.2. Puissance.....	13
1.3. Seuil de significativité et grandeur d'effet.....	14
Chapitre 2. Les procédures de tests	17
2.1. Procédures <i>Single-Step</i>	17
2.1.1. La procédure de <i>Simes</i>	17
2.1.2. La procédure de <i>Bonferroni</i>	18

2.1.3.	La procédure de <i>Dunn-Šidák</i>	19
2.1.4.	La procédure de <i>Dunnnett</i>	19
2.2.	Procédures <i>Step-Up</i>	20
2.2.1.	Procédure de <i>Hochberg</i>	20
2.2.2.	Procédure de <i>Rom</i>	21
2.2.3.	Procédure de <i>Hommel</i>	22
2.2.4.	Procédure <i>Dunnnett-Tamhane</i>	23
2.3.	Procédures <i>Step-Down</i>	24
2.3.1.	Procédure de <i>Holm</i>	24
2.3.2.	Procédure <i>Step-Down Dunnnett</i>	25
2.4.	Autre procédure <i>Step-Wise</i>	26
2.4.1.	Procédure <i>Step-Up-Down</i> de Tamhane-Liu-Dunnnett	26
Chapitre 3.	Rappels pour le cas d'une hypothèse unique sur la	
	différence entre deux moyennes	29
3.1.	Test d'égalité	29
3.2.	Test de non-infériorité	32
3.3.	Test de supériorité	33
3.4.	Test d'équivalence	34
Chapitre 4.	Présentation de la problématique	37
4.1.	Cas où $\text{Var}(X_{ijk}) = \sigma^2$	39
4.2.	Cas où $\text{Var}(X_{ijk}) = \sigma_i^2$	41
4.3.	Cas où $\text{Var}(X_{ijk}) = \sigma_k^2$	42
4.4.	Cas où $\text{Var}(X_{ijk}) = \sigma_{ik}^2$	44
Chapitre 5.	Équations générales pour la puissance	45

5.1.	Équation de la puissance pour la procédure <i>Single Step Bonferroni</i>	47
5.1.1.	Cas Non-échangeable	47
5.1.2.	Cas échangeable	47
5.2.	Équation de la puissance pour la procédure <i>Step-Up Hochberg</i>	48
5.3.	Équation de la puissance pour la procédure <i>Step-Down Holm</i>	49
Chapitre 6.	Algorithme pour le calcul de la taille d'échantillon	53
Chapitre 7.	Simulations et résultats	55
7.1.	Reprise des résultats de Chen & al.	55
7.2.	Impact de la variation de r	59
7.3.	Impact lorsque les variances entre deux variables sont très différentes l'une de l'autre	63
Chapitre 8.	Exemple pratique	67
Chapitre 9.	Conclusion	69
Bibliographie		71
Annexe A.	Formules de combinatoire et de probabilité	A-i
Annexe B.	Formules sur la probabilité conjointe de K statistiques d'ordre	B-i
B.1.		B-i
B.2.		B-ii
B.3.		B-iii
Annexe C.	Loi de <i>Student</i> multivariée	C-i
C.1.	Loi de <i>Student</i> multivariée de type I	C-i

C.2. Loi de *Student* multivariée de type II C-ii

Annexe D. Calcul des estimateurs pour μ et σ^2 , lois des statistiques

D-i

D.1. Lorsque $\text{Var}(X_{ijk}) = \sigma^2$ D-i

D.1.0.1. Déterminant et inverse d'une matrice de variances-covariances
ayant une structure compound-symmetry D-iv

D.1.1. Calcul de la statistique T_k lorsque $\text{Var}(X_{ijk}) = \sigma^2$ D-v

D.2. Lorsque $\text{Var}(X_{ijk}) = \sigma_i^2$ D-v

D.2.1. Calcul de la statistique T_k lorsque $\text{Var}(X_{ijk}) = \sigma_i^2$ D-vii

Annexe E. Calcul des équations de puissance E-i

E.1. Calcul de la puissance pour la méthode *Single Step Bonferroni*
échangeable E-i

E.2. Calcul de la puissance pour la méthode *Step-Up Hochberg* E-iv

E.3. Calcul de la puissance pour la méthode *Step-Down Holm* E-v

LISTE DES FIGURES

7.1	Graphiques de la taille n en fonction de la corrélation ρ pour $m=7$ et $r=1,2,3$.	61
7.2	Graphiques de la taille n en fonction de la corrélation ρ pour $m=7$ et $r=4$.	61
7.3	Graphiques de la taille n en fonction de la corrélation ρ pour $m=7$ et $r=5$.	62
7.4	Graphiques de la taille n en fonction de la corrélation ρ pour $m=7$ et $r=6$.	62
7.5	Graphiques de la taille n en fonction de la corrélation ρ pour $m=7$ et $r=7$.	62
7.6	Graphiques de la puissance calculée et de la puissance simulée pour $m=4$ et $r=2$.	63
7.7	Graphiques de la puissance calculée et de la puissance simulée pour $m=4$ et $r=3$.	64
E.1	Graphique des indices j par rapport aux indices t .	E-iii

LISTE DES TABLEAUX

1.1	Probabilité de commettre au moins une erreur de type I en fonction du nombre m d'hypothèses.	8
1.2	Scénarios possibles en présence de m hypothèses.	9
4.1	Schéma des données.	37
7.1	<i>Table 1</i> dans Chen & al. [8].....	56
7.2	7.1 recalculée avec notre <i>package</i>	56
7.3	<i>Table 2</i> dans Chen & al. [8].....	57
7.4	7.3 recalculée avec notre <i>package</i>	57
7.5	<i>Table 3</i> dans Chen & al. [8].....	58
7.6	7.5 recalculée avec notre <i>package</i>	58
7.7	Impact de r , 1 ^{ère} partie $r = 1, 2, 3, 4$	60
7.8	Impact de r , 2 ^e partie $r = 5, 6, 7$	60
8.1	Statistiques de l'exemple (moyennes et écart-types empiriques).	68
8.2	Résultats de l'exemple.....	68

REMERCIEMENTS

Merci à Benoit et Jérémie pour leurs contributions.

Merci à Pierre pour son aide et ses précieux conseils.

Merci à mes parents qui m'ont permis de m'épanouir et de devenir l'homme que je suis aujourd'hui.

Un très gros merci à Peggy pour ton amour, ton support et ta patience, sans toi je n'aurais pu passer à travers toutes ces épreuves.

Finalement, merci à Axelle qui illumine ma vie depuis maintenant un peu plus d'un an.

INTRODUCTION

Un projet de recherche se doit d'être bien planifié afin d'obtenir les meilleurs résultats possibles et d'éviter une perte de temps, d'énergie et de ressources. Selon Lenth [40], la planification d'une étude est composée de plusieurs aspects. D'abord la problématique, qui doit être bien définie et documentée. Ensuite, les unités d'expérimentation ou d'observation doivent être sélectionnées parmi une population cible appropriée. Le chercheur doit suivre et respecter attentivement les procédures établies et choisir des instruments de mesures fiables. Puis, l'échantillon doit être de taille adéquate, afin que les buts fixés par le chercheur soient atteints. La taille échantillonnale joue en effet un rôle primordial dans la planification d'un projet de recherche. Un échantillon trop petit risque de conduire à des résultats non significatifs, alors qu'un échantillon trop grand risque d'entraîner un gaspillage de ressources. De plus, dans le cas d'études portant sur des sujets vivants, si la taille de l'échantillon n'est pas adéquate, il est probable que celles-ci soient jugées non-éthiques, car elles exposeraient alors les participants à des risques inutiles [34], [40].

Puisque la taille de l'échantillon revêt une si grande importance, plusieurs chercheurs consultent des statisticiens afin de déterminer le nombre de sujets requis pour réaliser leurs études. Il suffit donc de répondre à la question : « De combien de sujets ai-je besoin pour mon étude ? ». Hanley et Shapiro [26] répondraient à la manière de Socrate : « Combien de briques faut-il pour construire un mur ? ». Bien entendu, la réponse dépend beaucoup du genre de mur que l'on désire construire. Dans le cas de projets de recherches, le choix dépend surtout des objectifs et des besoins de l'étude, donc des procédures statistiques utilisées afin de confirmer ou d'infirmer les hypothèses émises par le chercheur. Plusieurs

autres facteurs entrent aussi en ligne de compte lors du calcul du nombre de sujets requis : le niveau de significativité, la puissance, le type de test, le seuil critique ou seuil clinique, la distribution de la population cible, le coût de recrutement de chaque sujet, etc. Afin de mieux comprendre l'impact de ces facteurs sur le calcul de la taille échantillonnale, les définitions de certains termes et de leur influence seront données plus loin.

Il est d'abord important de mentionner que l'étude de la puissance et l'approche bayésienne sont les techniques les plus répandues pour le calcul de taille d'échantillons. Un débat entre statisticiens sur le choix de l'approche à privilégier existe d'ailleurs depuis plusieurs années. À ce sujet, l'édition spéciale « Sample size determination » du *Journal of the Royal Statistical Society* (1997) est des plus intéressante. Pour des raisons pratiques, l'approche retenue dans ce mémoire favorise l'étude de la puissance. Pour ce qui est des méthodes bayésiennes, la littérature est cependant très abondante à ce sujet, citons ; Adcock [1], Pezeshk [53], Lindley [42] et Lee and Zelen [37] [38]. La méthode basée sur l'efficacité proposée dans Bachetti *et al.* [4] et Bachetti [3] a fait aussi couler beaucoup d'encre ces derniers temps et mériterait que l'on s'y intéresse. Toutefois, elle n'entre pas dans le cadre de ce mémoire.

L'approche par la puissance dépend principalement des tests d'hypothèses, car le calcul de la puissance varie selon le test employé. Ce mémoire se concentre sur le calcul de la taille échantillonnale pour les tests à hypothèses multiples (*multiple endpoints*), où l'on désire qu'au moins r des m hypothèses soient significatives. La *multiplicité* est aussi un sujet qui suscite beaucoup de discussions. Certains sont d'avis qu'il n'est pas nécessaire de faire quelque ajustement que ce soit ; [16], [51] et [61]. Leur principal argument est que si les tests avaient été réalisés par plusieurs chercheurs séparément, les résultats auraient été acceptés ; alors pourquoi faire des modifications lorsqu'un seul chercheur fait tout le travail ? D'autres préfèrent suivre les directives ICH E9 [22] de l'EMA (*European Medicines Evaluation Agency*) qui recommande de n'utiliser qu'une seule hypothèse primaire (*primary endpoint*) et de considérer les autres hypothèses comme secondaires. Ces hypothèses secondaires sont alors utilisées à titre exploratoire

seulement et ne sont testées que si l'hypothèse primaire est rejetée. Toutefois, en pratique, il existe plusieurs cas où une seule variable n'est pas suffisante et des ajustements sont donc nécessaires pour tenir compte de la *multiplicité* ; [12] [54]. Le « *multiple endpoints expert team of the pharmaceutical research and manufacturers of America* » présente d'ailleurs une liste de maladies pour lesquelles les agences de réglementation requièrent au moins deux *endpoints* ou plus ; [52]. Il y est mentionné que les vaccins représentent un cas spécial, car ceux-ci nécessitent au moins 10 *endpoints* et parfois plus.

À ce jour, les agences de réglementation demandent à ce que l'efficacité soit démontrée sur au moins un *endpoint*, soit $r = 1$, ou bien sur tous les *endpoints*, donc $r = m$. Néanmoins, une nouvelle tendance semble apparaître dans la littérature. On commence en effet à s'intéresser à la situation où l'efficacité serait démontrée sur la majorité des *endpoints*, donc $r \geq m/2$; [62]. Tamhane, Liu et Dunnett [70] mentionnent que « dans certaines applications les chercheurs veulent répondre à la question 'Peut-on rejeter au moins r hypothèses ?' » Ils ont d'ailleurs développé une procédure statistique « utile lorsque l'objectif est de rejeter un nombre minimum spécifié r d'une famille de m hypothèses. » Ils ont appelé cette procédure : *Step-Up-Down* ; celle-ci sera présentée plus en détails dans la section 3. Chen *et al.* [8] s'inspirant des formules générales de puissance de Dunnett et Tamhane [20] ont développé une méthode permettant de calculer la taille d'échantillon requise lorsque l'objectif est de rejeter au moins r hypothèses parmi m , et ce pour les procédures de *Bonferroni*, de *Hochberg* et de *Holm*. Chen *et al.* [8] se sont limités au cas $r = 1$ ou $r = 2$ lorsque $m = 2$ et $m = 3$. Nous avons toutefois remarqué que ceux-ci avaient commis des erreurs dans le développement des formules pour les procédures de *Hochberg* et de *Holm*. Ce mémoire propose donc de corriger ces erreurs et de généraliser les formules pour des valeurs quelconques de r et de m .

Ce mémoire est organisé de la façon suivante. La section 2 introduit certaines définitions utiles dans le calcul de la puissance. La section 3 présente une revue des procédures de tests multiples les plus utilisés. Un rappel du cas d'une hypothèse unique pour la différence entre deux moyennes est proposé en section 4. La présentation de la problématique se trouve dans la section 5. L'algorithme pour

le calcul de la taille d'échantillon est donné en section 6. Les formules pour le calcul des puissances sont disponibles dans la section 7. La section 8 comporte la discussion, les ouvertures et la conclusion.

Chapitre 1

DÉFINITIONS

Dans cette section, nous définissons les principaux concepts qui interviennent dans le calcul de la taille échantillonnale en présence de *multiplicité*. Avant de poursuivre, il est important de bien distinguer les deux types de *multiplicité* ; soit les *comparaisons multiples* et les *hypothèses multiples*.

Le terme *comparaisons multiples* désigne la comparaison unique de $K \geq 3$ groupes pour lesquels on a mesuré une variable d'intérêt dépendant d'un paramètre θ , soit ;

$$H_0 : \theta_1 = \theta_2 = \dots = \theta_K \quad \text{vs} \quad H_1 : \theta_i \neq \theta_j \text{ pour au moins un } i \neq j,$$

alors que le terme *hypothèses multiples* désigne la comparaison de deux groupes sur la base de m variables d'intérêt dont la $k^{\text{ième}}$ variables d'intérêt dépend d'un paramètre θ_k , soit ;

$$\begin{array}{l} H_0^1 : \theta_{11} = \theta_{21} \quad \text{vs} \quad H_1^1 : \theta_{11} \neq \theta_{21} \\ \dots \\ H_0^k : \theta_{1k} = \theta_{2k} \quad \text{vs} \quad H_1^k : \theta_{1k} \neq \theta_{2k} \\ \dots \\ H_0^m : \theta_{1m} = \theta_{2m} \quad \text{vs} \quad H_1^m : \theta_{1m} \neq \theta_{2m} \end{array}$$

Comme ce mémoire est exclusivement consacré aux hypothèses multiples, les définitions qui suivent sont présentées en tenant compte de ce contexte.

1.1. RISQUES D'ERREUR À CONTRÔLER

Rappelons d'abord qu'en présence d'une seule hypothèse, une erreur de type I est commise lorsqu'on rejette l'hypothèse nulle H_0 , alors qu'en réalité celle-ci s'avère vraie. En d'autres termes, l'interprétation des résultats indique l'existence erronée d'un effet du traitement lorsqu'il n'y en a aucun.

$$\text{Erreur de type I} = \{\text{rejet de } H_0 \text{ alors que } H_0 \text{ est vraie}\}.$$

Le niveau de significativité α , aussi appelé taux d'erreur, est une borne supérieure de la probabilité de commettre une erreur de type I fixée par le chercheur afin de savoir si l'effet observé est dû ou non à la chance. Les valeurs les plus courantes pour α sont 0,01, 0,05 et 0,10. On dit d'une procédure statistique qu'elle contrôle l'erreur de type I au niveau α lorsque :

$$\mathbf{P}(\text{rejet de } H_0 \text{ alors que } H_0 \text{ est vraie}) \leq \alpha.$$

Comme le niveau α est une borne supérieure ; plus celui-ci sera grand, plus la probabilité de commettre une erreur de type I sera grande. Ainsi, la taille requise pour l'expérience diminue lorsque α augmente et vice-versa.

Dans le cas de m hypothèses multiples, la méthode naïve consiste à ignorer la multiplicité et à effectuer m tests au niveau α . Toutefois, en procédant de la sorte on augmente la probabilité de commettre au moins une erreur de type I. Ainsi, plus le nombre d'hypothèses à tester sera grand, plus le risque qu'au moins l'une d'entre elles soit rejetée sera grand. La table 1.1 suivante montre la probabilité de commettre au moins une erreur de type I en fonction du nombre m d'hypothèses indépendantes, lorsque toutes les hypothèses sont testées au niveau α [39].

TAB. 1.1. Probabilité de commettre au moins une erreur de type I en fonction du nombre m d'hypothèses.

m	1	2	5	10	50
$\mathbf{P}(\text{commettre au moins une erreur de type I})$,05	,10	,23	,40	,92

Il devient clair qu'une telle procédure ne contrôle pas la probabilité de commettre au moins une erreur de type I au niveau 0,05. C'est pourquoi il existe

plusieurs définitions de taux d'erreur à contrôler en présence d'hypothèses multiples. Les taux les plus courants dans la littérature, et les plus utilisés en pratique, sont présentés plus loin.

La table 1.2 suivante aidera à la visualisation des scénarios possibles avec m hypothèses.

TAB. 1.2. Scénarios possibles en présence de m hypothèses.

	Sur la base d'un échantillon		Total
	Accepter H_0	Rejeter H_0	
H_0 vraie	U	V	m_0
H_0 fausse	T	S	m_1
	W	R	m

Dans ce tableau, U représente le nombre de vraies hypothèses nulles qui sont acceptées, V représente le nombre de vraies hypothèses nulles qui sont rejetées, T représente le nombre de fausses hypothèses nulles qui sont acceptées et S représente le nombre de fausses hypothèses nulles qui sont rejetées. Ainsi, $W = U + T$ représente le nombre total d'hypothèses nulles acceptées et $R = V + S$ représente le nombre total d'hypothèses nulles rejetées, alors que $m_0 = U + V$ représente le nombre total de vraies hypothèses nulles et $m_1 = T + S$ représente le nombre total de fausses hypothèses nulles. En pratique, m_0 et m_1 sont inconnus.

1.1.1. Taux d'erreurs par comparaisons ($PCER$)

Le taux d'erreur par comparaisons ($PCER$: *per-comparison error rate*) est l'espérance de la proportion d'erreurs de type I parmi les m hypothèses. Par rapport à la table 1.2, le $PCER$ vaut $\mathbb{E}(V)/m$. Il peut surtout servir lorsque les hypothèses ont toutes la même importance aux yeux du chercheur, alors celui-ci décide du niveau α_{PCER} qu'il désire pour tester chacune des hypothèses séparément.

Par exemple, pour tester trois hypothèses au niveau fixé de 5%, trois tests sont effectués, chacun étant réalisé au niveau $\alpha_{PCER}=0,05$. Ce taux ne tient pas compte de la multiplicité des tests.

1.1.2. Taux d'erreurs par-famille (*PFER*)

Le taux d'erreurs par-famille (*PFER* : *per-family error rate*) correspond à l'espérance du nombre d'erreurs de type I faites dans un groupe (famille) d'hypothèses : soit $PFER = \mathbb{E}(V)$.

On remarque la relation entre le *PFER* et le *PCER* : $PFER = m * PCER$. Ainsi, pour trois hypothèses testées au niveau $\alpha_{PCER} = 0,05$, l'espérance du nombre d'erreurs de type I commises sera de $3 * 0,05 = 0,15$.

1.1.3. *Familywise error rate* (*FWER*)

Le *familywise error rate* (*FWER*) est défini comme étant la probabilité de commettre au moins une erreur de type I lorsque l'on teste les hypothèses $H_0^1, H_0^2, \dots, H_0^m$ et que l'on suppose que certaines d'entre elles sont vraies tandis que les autres sont fausses : soit $FWER = \mathbf{P}(V \geq 1)$. Lorsque les m hypothèses sont testées via des statistiques de test indépendantes et que toutes les hypothèses sont vraies, nous obtenons la relation suivante [29] :

$$\begin{aligned}
 FWER &= \mathbf{P}(V \geq 1) \\
 &= 1 - \mathbf{P}(V = 0) \\
 &= 1 - \mathbf{P}\left(\bigcap_{i=1}^m \{\text{ne pas rejeter } H_0^i | H_0^i \text{ vraie}\}\right) \\
 &= 1 - \prod_{i=1}^m \mathbf{P}\{\text{ne pas rejeter } H_0^i | H_0^i \text{ vraie}\} \\
 &= 1 - \underbrace{(1 - \alpha_{PCER})(1 - \alpha_{PCER}) \dots (1 - \alpha_{PCER})}_{m \text{ fois}} \\
 &= 1 - (1 - \alpha_{PCER})^m.
 \end{aligned}$$

De plus, l'inégalité suivante est vraie :

$$\mathbf{P}(\text{commettre au moins une erreur de type I}) \leq 1 - (1 - \alpha_{PCER})^m, \quad (1.1.1)$$

si les m hypothèses testées sont issues d'une loi multinormale centrée ou bien d'une loi de *Student* multivariée avec une matrice de covariance ayant une certaine structure spécifique [67].

De plus, dans la littérature [29] [75] on distingue deux façons de contrôler le taux d'erreurs *FWER*; le contrôle faible (*weak*) et le contrôle fort (*strong*).

On parle de contrôle faible lorsque la procédure contrôle le *FWER* uniquement sous l'hypothèse nulle complète; soit lorsque l'on suppose que toutes les m hypothèses sont vraies.

$$\mathbf{P} \left(\text{rejeter au moins une hypothèse } H_0^k | H_0^C \text{ est vraie} \right) \leq \alpha,$$

où $H_0^C = \bigcap_{k=1}^m H_0^k$. Dans cette situation, rejeter au moins une des hypothèses est équivalent à rejeter H_0^C [28].

Notons que l'on ne connaît pas en pratique la véracité ou non des hypothèses testées.

Le contrôle fort fait référence au contrôle du *FWER* pour l'ensemble H de toutes les combinaisons possibles de vraies et de fausses hypothèses :

$$\mathbf{P} (\text{rejeter au moins une vraie hypothèse} | H') \leq \alpha,$$

pour tout $H' \subseteq H$.

Par exemple, si trois hypothèses sont testées au niveau $\alpha_{PCER} = 5\%$, alors le niveau α_{FWER} du test sera de $1 - (1 - 0,05)^3 = 0,14$. Inversement, afin d'obtenir un niveau α_{FWER} de 5%, il faut que le taux d'erreurs par comparaisons soit d'environ 0,017.

En résumé, si on utilise un niveau α' pour chacune des m hypothèses, on obtient :

$$\alpha_{PCER} = \alpha'; \alpha_{FWER} \leq 1 - (1 - \alpha')^m \text{ et } \alpha_{PFER} = m\alpha'. \quad (1.1.2)$$

Ceci conduit facilement à la relation suivante [29] :

$$PCER \leq FWER \leq PFER. \quad (1.1.3)$$

En pratique, le $FWER$ est le taux le plus couramment utilisé lorsque le nombre d'hypothèses n'est pas trop élevé.

1.1.4. *Generalized familywise error rate (gFWER)*

Le *generalized familywise error rate (gFWER)* est la probabilité de commettre au moins k , pour $k = 1, \dots, m$, erreurs de type I : soit $gFWER(k) = \mathbf{P}(V \geq k)$. À noter que $gFWER(1) = \mathbf{P}(V \geq 1) = FWER$. Ce taux n'est utilisé que lorsqu'il est acceptable de commettre plus d'une erreur de type I. En pratique, on lui préfère le $FWER$, car on cherche le plus souvent à minimiser les erreurs commises.

1.1.5. Taux de fausses découvertes (FDR)

Benjamini et Hochberg [6] ont introduit le taux de fausses découvertes (FDR : *False discovery rate*) qui est l'espérance de la proportion d'erreurs de type I parmi les hypothèses rejetées.

En général, $FDR \leq FWER$, avec égalité sous l'hypothèse nulle complète et lorsque toutes les m hypothèses rejetées sont des erreurs de type I ; i.e. $V = R = m$. Ainsi le $FDR = \mathbf{P}(V > 0) = FWER$ et les procédures qui contrôlent le FDR contrôlent aussi le $FWER$ au sens faible [15].

$$\begin{aligned} FDR &= \mathbb{E} \left(\frac{V}{R} \mid R > 0 \right) \mathbf{P}(R > 0) \quad \text{avec } V/R = 0 \text{ si } R = 0, \\ &= \mathbb{E} \left(\frac{V}{R} \mid V > 0 \right) \mathbf{P}(V > 0). \end{aligned}$$

Le FDR sert lorsque le nombre d'hypothèses m est très grand. On le retrouve le plus souvent dans les études génétiques utilisant des puces à ADN (*microarray*), où l'on teste des milliers d'hypothèses simultanément.

1.1.6. Taux positif de fausses découvertes ($pFDR$)

Le taux positif de fausses découvertes ($pFDR$: *Positive false discovery rate*) est une modification du FDR. Il s'agit en fait du FDR conditionné au fait qu'au moins une découverte positive est réalisée. La formule est donc : $pFDR = \mathbb{E} \left(\frac{V}{R} \mid R > 0 \right)$.

Le $pFDR$ sert lorsque le nombre d'hypothèses m est très grand. Tout comme le FDR , on le retrouve le plus souvent dans les études *microarray*.

1.2. PUISSANCE

La puissance joue un rôle primordial dans le calcul de la taille d'échantillon. On parle d'ailleurs souvent de *calcul de puissance* pour faire référence au calcul de taille.

Pour bien comprendre la notion de puissance, il faut d'abord introduire l'erreur de type II qui est définie comme étant l'action de ne pas rejeter l'hypothèse nulle, alors que celle-ci est fautive. En d'autres termes, conclure qu'il n'existe aucun effet du traitement, alors que le traitement a un effet.

Erreur de type II = {ne pas rejeter H_0 alors que H_0 est fautive}.

On définit alors β comme étant la probabilité de commettre une erreur de type II. La puissance est quant à elle définie comme la probabilité de correctement rejeter l'hypothèse nulle lorsque celle-ci est fautive et bien fautive [9] :

$$\begin{aligned} \text{Puissance} &= \mathbf{P}(\text{rejeter } H_0 \text{ alors que } H_0 \text{ est fautive}) & (1.2.1) \\ &= 1 - \beta. \end{aligned}$$

Pour ce qui est de l'impact de la puissance sur la taille de l'échantillon, la relation est monotone. Ainsi, plus la puissance requise sera élevée, plus la taille sera grande et inversement.

Tout comme pour les taux d'erreur de type I, il existe différents types de puissances pour les tests d'hypothèses multiples ; [32], [56] et [73].

D'abord la puissance par paire (*per-pair power*), comme son nom l'indique, correspond au concept de puissance pour tester une seule hypothèse ; tel qu'indiqué par l'équation (1.2.1).

La puissance de n'importe quelle paire (*any-pair power*), ou puissance minimale, est quant à elle la probabilité de rejeter au moins une fautive hypothèse

nulle :

\mathbf{P} (rejeter au moins 1 des m_1 fausses hypothèses).

Puis, la puissance de toutes les paires (*all-pair power*), ou puissance complète, est la probabilité de rejeter toutes les fausses hypothèses nulles :

\mathbf{P} (rejeter toutes les m_1 fausses hypothèses).

La relation entre ces trois puissances s'exprime de la façon suivante :

All-pair power \leq Per-pair power \leq Any-pair power.

Chen & al. (2011) [8] réfèrent au *r-power* comme étant la probabilité de rejeter au moins r fausses hypothèses nulles. Il s'agit en fait de la définition générale de la puissance donnée par Dunnett et Tamhane [20] :

\mathbf{P} (rejeter au moins r des m_1 fausses hypothèses).

1.3. SEUIL DE SIGNIFICATIVITÉ ET GRANDEUR D'EFFET

Le *seuil de significativité*, ou seuil clinique (Δ), est une borne établie par le chercheur pour confirmer que la différence entre deux moyennes est significative du point de vue du domaine de recherche. En d'autres termes, une borne qui permet de s'assurer que la différence est suffisamment grande pour que celle-ci soit digne d'intérêt.

La *grandeur d'effet* est simplement le seuil de significativité standardisé ; c'est-à-dire le seuil divisé par la racine carrée de la variance combinée des deux groupes étudiés ($\frac{\Delta}{\sigma}$). Selon Coe [10], c'est une façon simple de quantifier la différence entre deux groupes. Cela permet d'aller au-delà de la question « Est-ce que le traitement fonctionne ? », et de répondre à l'interrogation « À quel point le traitement fonctionne-t-il ? ».

Les termes *small*, *medium* et *large* sont souvent employés pour caractériser la grandeur d'effet. Cette convention est due à Cohen [11] qui a divisé le concept de grandeur d'effet en trois catégories : *small* lorsque la grandeur d'effet est inférieure à 0,3, *medium* lorsqu'elle tourne autour de 0,5 et *large* si elle est supérieure à 0,8.

Plus l'effet est grand, plus il est facile de le détecter et moins de sujets sont alors requis. Inversement, si l'effet est petit, la taille en sera décuplée ; réduire l'effet de moitié nécessitera approximativement une taille d'échantillon quatre fois plus grande [40]. Il est donc très important que le chercheur spécifie la grandeur de l'effet représentant un intérêt scientifique avant de procéder au calcul de la taille de l'échantillon.

Cependant, comme la grandeur de l'effet dépend de la distribution de la population et de paramètres qui peuvent être inconnus, comme la variance, il peut être nécessaire d'avoir recours à une étude pilote afin d'estimer ces informations. Il est aussi possible que ces informations soient disponibles dans la littérature suite à des expériences antérieures.

Chapitre 2

LES PROCÉDURES DE TESTS

Les procédures de tests multiples se séparent principalement en deux classes : les méthodes globales (*Single-Step*) et les méthodes pas à pas (*Stepwise*), qui se distinguent par leurs régions critiques pour les tests statistiques qui sont constantes ou aléatoires [15]. Puis, les méthodes *Stepwise* se divisent en deux grandes catégories : les *Step-Up* et les *Step-Down*.

2.1. PROCÉDURES *Single-Step*

Dans une procédure *Single-Step*, la région de rejet de chaque hypothèse est indépendante du résultat des autres hypothèses [15]. En d'autres termes, la décision de rejeter ou non une des hypothèses ne dépend pas du rejet d'une autre hypothèse. Comme l'ordre dans lequel les hypothèses sont testées n'est pas important, nous pouvons alors considérer qu'elles sont évaluées simultanément, d'où l'utilisation du terme *single-step* (une seule étape) [14]. Les méthodes les plus connues sont celles de *Simes*, *Bonferroni*, *Dunn-Šidák* et *Dunnett*.

2.1.1. La procédure de *Simes*

Le test de *Simes* [68] consiste à tester l'hypothèse globale qu'il n'existe aucun effet ; c'est-à-dire qu'aucune hypothèse n'est significative :

$$H_I = \bigcap_{i=1}^m H_i.$$

Ainsi, H_I est rejetée aussitôt qu'une des m hypothèses ordonnées est rejetée au niveau $i\alpha/m$. Pour effectuer un test global de *Simes*, il faut d'abord ordonner

les valeurs- p , $p_{(1)} < \dots < p_{(m)}$, puis H_I est rejetée si $p_{(i)} \leq i\alpha/m$ pour au moins un $i = 1, \dots, m$ [14].

Il est important de noter que cette procédure ne permet pas de conclure quoi que ce soit par rapport aux hypothèses individuelles. De plus, cette méthode est rarement utilisée, car elle ne contrôle pas le *FWER* [31]. Cette procédure est tout de même présentée ici, car elle a inspiré la méthode *stepwise* de Hommel, ainsi que celle de Hochberg.

2.1.2. La procédure de *Bonferroni*

Certains affirment que cette méthode est due à Fisher [29] [55], alors que d'autres prétendent que c'est plutôt Dunn [18] qui en est le père [41] [73]. Tous s'entendent cependant pour dire qu'elle se base sur l'inégalité de *Boole*, laquelle stipule que la probabilité d'un ou de plusieurs événements est plus petite ou égale à la somme des probabilités de ces événements pris séparément :

$$\mathbf{P}\left(\bigcup_{i=1}^n E_i\right) \leq \sum_{i=1}^n \mathbf{P}(E_i).$$

L'inégalité a pour fonction de fournir une borne supérieure pour le taux d'erreur *familywise* (FWER) [35] :

$$\mathbf{P}(\text{au moins une erreur de type I}) = \alpha \leq \sum_i \alpha'.$$

La méthode consiste donc à tester chacune des hypothèses en prenant $\alpha' = \alpha/m$, où m est le nombre d'hypothèses. En termes de statistiques- t , toutes les statistiques seront comparées à la même valeur critique $t_{\alpha/m}$. De cette façon, cette technique contrôle à la fois les niveaux *FWER* et *PFER* sous toutes les configurations. En effet par (1.1.2) :

$$PFER \leq m\alpha' = \frac{m\alpha}{m} = \alpha$$

et comme nous l'avons vu en (1.1.3); $FWER \leq PFER$, il devient clair que le taux d'erreurs *familywise* est aussi protégé.

2.1.3. La procédure de *Dunn-Šidák*

La méthode de *Dunn-Šidák* est presque identique à la procédure de *Bonferroni*. Elle se base toutefois sur les travaux de Dunn [17] et sur l'inégalité multiplicative (1.1.1) de Šidák [66], alors que la méthode de *Bonferroni* est basée sur l'inégalité additive de *Bonferroni*.

Rappelons que, sous certaines hypothèses et par (1.1.1), nous avons :

$$\alpha_{FWER} \leq 1 - (1 - \alpha_{PCER})^m.$$

Il en découle alors α'' qui est défini comme :

$$\alpha'' = 1 - (1 - \alpha_{FWER})^{1/m}.$$

Ainsi, l'inégalité permet de contrôler le niveau d'erreur *FWER* et de fournir une approximation pour la borne supérieure de celui-ci. De plus, cette approximation est légèrement meilleure que celle obtenue avec l'inégalité de *Bonferroni* [35] :

$$\begin{aligned} \mathbf{P}(\text{au moins une erreur de type I}) &\leq 1 - (1 - \alpha'')^m \\ &= 1 - (1 - [1 - (1 - \alpha)^{1/m}])^m \\ &= 1 - [(1 - \alpha)^{1/m}]^m \\ &= 1 - (1 - \alpha) \\ &= \alpha. \end{aligned}$$

Il est cependant à noter que la méthode de *Dunn-Šidák*, contrairement à l'approche de *Bonferroni*, requiert l'indépendance entre les tests. C'est pourquoi elle est beaucoup plus utilisée pour des tests de comparaisons multiples que pour des hypothèses multiples.

2.1.4. La procédure de *Dunnett*

Contrairement aux méthodes vues précédemment, la procédure de *Dunnett* [19] ne se base pas sur l'ajustement des valeurs-*p* ou la modification du niveau

α pour assurer le contrôle du *FWER*. Cette méthode est plutôt basée sur la distribution conjointe de tous les tests. De cette façon, elle tient ainsi compte de la corrélation entre les statistiques- t . Pour ce faire, Dunnett a développé la distribution du maximum de plusieurs variables aléatoires t qui suivent une loi de *Student*. Il s'agit d'une distribution multivariée analogue à la distribution de *Student* multivariée (voir annexe C). La fonction $F(x|m, \nu)$ suivante est la fonction de distribution cumulative d'une distribution unilatérale de *Dunnett* [14] :

$$F(x|m, \nu) = \mathbf{P}[\max(t_1, \dots, t_m) \leq x], \quad (2.1.1)$$

avec $\nu = (m + 1)(n - 1)$. La procédure rejette les hypothèses H_i lorsque $t_i \geq u_\alpha(m, \nu)$, où $u_\alpha(m, \nu)$ est le quantile $(1 - \alpha)$ de la distribution de *Dunnett*.

Il est toutefois important de noter que cette méthode est surtout utilisée en comparaisons multiples ; pour la comparaison de plusieurs groupes avec un contrôle ou lors d'études de recherche de dose maximale (*dose-finding studies*).

2.2. PROCÉDURES *Step-Up*

Dans une procédure *Step-Up*, les hypothèses sont testés une à une de façon croissante, i.e. de la moins significative à la plus significative, jusqu'à ce que l'une d'elles soit rejetée. Ainsi, toutes les hypothèses restantes seront aussi rejetées sans qu'il soit nécessaire de les tester. Le fait de commencer le test par l'hypothèse la moins significative permet de répondre à la question : « Est-ce que toutes les hypothèses peuvent être rejetées ? ». Lorsque la réponse est négative, le reste de la procédure permet d'identifier toutes les hypothèses pouvant être « acceptées » [69]. Les procédures *Step-Up* les plus connues sont celles de *Hochberg*, de *Rom*, de *Hommel* et *Dunnett-Tamhane*.

2.2.1. Procédure de *Hochberg*

Hochberg s'est inspiré du test global de *Simes* [68] pour développer sa procédure *Step-Up*. Il est important de noter que cette méthode utilise des valeurs- p issues de tests univariés. La procédure commence donc par tester l'hypothèse associée à la plus grande valeur- p et ne rejette pas celle-ci si elle est supérieure à

α . Si la première hypothèse n'est pas rejetée, la procédure teste au niveau $\alpha/2$ l'hypothèse suivante associée à la deuxième plus grande valeur- p et ainsi de suite jusqu'à ce que toutes les hypothèses soient « acceptées » ou jusqu'à ce qu'une hypothèse soit rejetée. Lorsqu'une hypothèse est rejetée, toutes les hypothèses suivantes le sont aussi. À chaque étape, le niveau de significativité diminue pour atteindre α/m lorsqu'il ne reste qu'une hypothèse.

En termes de valeurs- p , l'algorithme pour la procédure de *Hochberg* est le suivant [63] :

- (1) On ordonne les valeurs- p , $p_{(1)} \geq p_{(2)} \geq \dots \geq p_{(m)}$, qui correspondent aux hypothèses ordonnées $H_0^{(1)}, H_0^{(2)}, \dots, H_0^{(m)}$.
- (2) On ne rejette pas $H_0^{(1)}$ si $p_{(1)} > \alpha$, puis on passe à l'étape suivante. Sinon, la procédure s'arrête et on rejette toutes les hypothèses nulles.
- (3) On ne rejette pas $H_0^{(2)}$ si $p_{(2)} > \alpha/2$, puis on passe à l'étape suivante. Sinon, la procédure s'arrête et on rejette toutes les hypothèses nulles $H_0^{(j)}$, $j = 2, \dots, m$.
- (4) En général, on ne rejette pas $H_0^{(i)}$ si $p_{(i)} > \alpha/i$. Sinon, la procédure s'arrête et on rejette toutes les hypothèses nulles $H_0^{(j)}$, $j = i, i + 1, \dots, m$.
- (5) Si toutes les $m - 1$ premières hypothèses n'ont pas été rejetées, on ne rejette pas $H_0^{(m)}$ si $p_{(m)} > \alpha/m$. Dans le cas contraire, seule $H_0^{(m)}$ est rejetée.

2.2.2. Procédure de *Rom*

La procédure de *Rom* [59] est presque identique à celle de *Hochberg*. Seules les valeurs critiques α/i sont remplacées par d_i , où les d_i sont calculés à l'aide de la formule de récurrence suivante [21] :

$$d_1 = \alpha$$

et

$$\sum_{i=1}^{k-1} d_1^i = \sum_{i=1}^{k-1} \binom{k}{i} d_{i+1}^{k-i}, \quad \text{pour } k = 2, \dots, m.$$

En termes de valeurs- p , l'algorithme pour la procédure de *Rom* est donc le suivant :

- (1) On ordonne les valeurs- p , $p_{(1)} \geq p_{(2)} \geq \dots \geq p_{(m)}$, qui correspondent aux hypothèses ordonnées $H_0^{(1)}, H_0^{(2)}, \dots, H_0^{(m)}$.
- (2) On ne rejette pas $H_0^{(1)}$ si $p_{(1)} > d_1$, puis on passe à l'étape suivante. Sinon, la procédure s'arrête et on rejette toutes les hypothèses nulles.
- (3) On ne rejette pas $H_0^{(2)}$ si $p_{(2)} > d_2$, puis on passe à l'étape suivante. Sinon, la procédure s'arrête et on rejette toutes les hypothèses nulles $H_0^{(j)}$, $j = 2, \dots, m$.
- (4) En général, on ne rejette pas $H_0^{(i)}$ si $p_{(i)} > d_i$. Sinon, la procédure s'arrête et on rejette toutes les hypothèses nulles $H_0^{(j)}$, $j = i, i + 1, \dots, m$.
- (5) Si toutes les $m - 1$ premières hypothèses n'ont pas été rejetées, on ne rejette pas $H_0^{(m)}$ si $p_{(m)} > d_m$. Dans le cas contraire, seule $H_0^{(m)}$ est rejetée.

Selon Dunnett et Tamhane [21], cette procédure permet un certain gain en puissance, environ 2%, par rapport à la méthode de *Hochberg*. Toutefois, selon eux, il est préférable d'utiliser cette dernière, car son calcul est plus simple et le gain en puissance n'est pas suffisant pour justifier l'utilisation de la procédure de *Rom*.

2.2.3. Procédure de *Hommel*

La procédure de *Hommel* [31] est aussi inspirée du test global de *Simes*, mais sa mise en oeuvre est beaucoup plus complexe que la méthode de *Hochberg* ou celle de *Rom*.

En termes de valeurs- p , l'algorithme pour la procédure de *Hommel* est le suivant [14] :

- (1) On ordonne les valeurs- p , $p_{(1)} \leq p_{(2)} \leq \dots \leq p_{(m)}$, qui correspondent aux hypothèses ordonnées $H_0^{(1)}, H_0^{(2)}, \dots, H_0^{(m)}$.
- (2) On ne rejette pas $H_0^{(m)}$ si $p_{(m)} > \alpha$, puis on passe à l'étape suivante. Sinon, la procédure s'arrête et on rejette toutes les hypothèses nulles.

- (3) En général, pour $i = 2, \dots, m - 1$ on ne rejette pas $H_0^{(m-i+1)}$ si $p_{(m-j+1)} > (i - j + 1)\alpha/i$ pour $j = 1, \dots, i$. Sinon, la procédure s'arrête et on rejette toutes les hypothèses nulles restantes.
- (4) Si toutes les $m - 1$ premières hypothèses n'ont pas été rejetées, on ne rejette pas $H_0^{(1)}$ si $p_{(m-j+1)} > (m - j + 1)\alpha/m$ pour $j = 1, \dots, m$. Dans le cas contraire, seule $H_0^{(1)}$ est rejetée.

Bien que cette procédure soit uniformément plus puissante que celle de *Hochberg*, *Dunnnett* et *Tamhane* [21] conseillent l'utilisation de la procédure de *Hochberg*, car le gain en puissance d'environ 2% n'est pas suffisant pour justifier l'utilisation de la procédure plus complexe de *Hommel*. Ils préfèrent aussi celle de *Rom* à celle de *Hommel* pour la simplicité de son algorithme et parce que toutes les valeurs critiques d_i ne doivent être calculées qu'une seule fois pour un α donné.

2.2.4. Procédure *Dunnnett-Tamhane*

La procédure *Dunnnett-Tamhane* [20] s'apparente à celles de *Hochberg* et de *Rom*. Alors que ces dernières utilisent les valeurs- p univariées ordonnées, la méthode *Dunnnett-Tamhane* utilise les statistiques de test ordonnées et comparent celles-ci à des valeurs critiques issues d'une distribution de *Student* multivariée de type I (voir annexe C.1).

Par la suite, nous noterons $t_{k,\nu,\rho}^{(\alpha)}$ la valeur (point critique supérieur) telle que la probabilité qu'une variable aléatoire de loi Student k -multivariée (centrée, de corrélation commune ρ et de degré de liberté ν) soit supérieure ou égale à $(t_{k,\nu,\rho}^{(\alpha)}, \dots, t_{k,\nu,\rho}^{(\alpha)})$ soit égale à α . En termes de statistiques- t , l'algorithme pour la procédure de *Dunnnett-Tamhane* est le suivant [69] :

- (1) On ordonne les statistiques- t , $T_{(1)} \leq T_{(2)} \leq \dots \leq T_{(m)}$, qui correspondent aux hypothèses ordonnées $H_0^{(1)}, H_0^{(2)}, \dots, H_0^{(m)}$.
- (2) On ne rejette pas $H_0^{(1)}$ si $T_{(1)} \leq t_{\nu}^{(\alpha)}$, puis on passe à l'étape suivante. Sinon, la procédure s'arrête et on rejette toutes les hypothèses nulles.
- (3) On ne rejette pas $H_0^{(2)}$ si $T_{(2)} \leq t_{2,\nu,\rho}^{(\alpha)}$, puis on passe à l'étape suivante. Sinon, la procédure s'arrête et on rejette toutes les hypothèses nulles $H_{(0j)}$, $j = 2, \dots, m$.

- (4) En général, on ne rejette pas $H_0^{(i)}$ si $T_{(i)} \leq t_{i,\nu,\rho}^{(\alpha)}$. Sinon, la procédure s'arrête et on rejette toutes les hypothèses nulles $H_0^{(j)}$, $j = i, i + 1, \dots, m$.
- (5) Si toutes les $m - 1$ premières hypothèses n'ont pas été rejetées, on ne rejette pas $H_0^{(m)}$ si $T_{(m)} \leq t_{m,\nu,\rho}^{(\alpha)}$. Dans le cas contraire, seule $H_0^{(m)}$ est rejetée.

Dunnett et Tamhane [21] montrent que cette méthode est plus puissante que les précédentes en grande partie parce qu'elle tient compte de la corrélation entre les statistiques. Ils affirment en effet que les trois méthodes *Step-Up* présentées plus haut sont plutôt conservatrices parce qu'elles négligent la corrélation.

2.3. PROCÉDURES *Step-Down*

Dans une procédure *Step-Down*, les hypothèses sont testées une-à-une de façon décroissante, i.e. de la plus significative à la moins significative, jusqu'à ce que l'une d'elles soit « acceptée ». Ainsi, toutes les hypothèses restantes seront aussi « acceptées » sans qu'il soit nécessaire de les tester. Le fait de commencer le test par l'hypothèse la plus significative permet de répondre à la question : « Est-ce qu'au moins une hypothèse peut être rejetée ? ». Lorsque la réponse est positive, le reste de la procédure permet d'identifier toutes les hypothèses pouvant être rejetées [69]. Les procédures *Step-Down* les plus connues sont celles de *Holm*, de *Shaffer* et *Step-Down Dunnett*.

2.3.1. Procédure de *Holm*

La méthode de Holm [30] est dite séquentiellement *rejective*, car les hypothèses sont rejetées une à la fois, selon certaines règles, en utilisant des valeurs- p univariées. La procédure commence par tester l'hypothèse associée à la plus petite valeur- p et la rejette si celle-ci est inférieure à α/m . Si la première hypothèse est ainsi rejetée, la procédure teste, au niveau $\alpha/(m - 1)$, l'hypothèse suivante associée à la deuxième plus petite valeur- p et ainsi de suite jusqu'à ce que toutes les hypothèses soient rejetées ou jusqu'à ce qu'une hypothèse soit acceptée. Lorsqu'une hypothèse est acceptée, toutes les hypothèses suivantes le sont aussi. À

chaque étape, le niveau de significativité augmente pour atteindre α lorsqu'il ne reste qu'une hypothèse.

En termes de valeurs- p , l'algorithme pour la procédure de *Holm* est le suivant [63] :

- (1) On ordonne les valeurs- p , $p_{(1)} \geq p_{(2)} \geq \dots \geq p_{(m)}$, qui correspondent aux hypothèses ordonnées $H_0^{(1)}, H_0^{(2)}, \dots, H_0^{(m)}$.
- (2) On rejette $H_0^{(m)}$ si $p_{(m)} < \alpha/m$, puis on passe à l'étape suivante. Sinon, la procédure s'arrête et on ne rejette aucune des hypothèses nulles.
- (3) On rejette $H_0^{(m-1)}$ si $p_{(m-1)} < \alpha/(m-1)$, puis on passe à l'étape suivante. Sinon, la procédure s'arrête et on ne rejette aucune des hypothèses nulles $H_0^{(j)}$, $j = 1, 2, \dots, m-1$.
- (4) En général, on rejette $H_0^{(i)}$ si $p_{(i)} < \alpha/i$. Sinon, la procédure s'arrête et on ne rejette aucune des hypothèses nulles $H_0^{(j)}$, $j = 1, 2, \dots, i$.
- (5) Si toutes les $m-1$ premières hypothèses ont été rejetées, on rejette $H_0^{(1)}$ si $p_{(1)} < \alpha$. Dans le cas contraire, seule $H_0^{(1)}$ n'est pas rejetée.

Shaffer [65] a modifié la procédure de *Holm* afin de tenir compte de la dépendance logique entre les hypothèses. On parle de dépendance logique lorsque la véracité d'un sous-groupe d'hypothèses implique nécessairement la véracité d'une ou plusieurs autres hypothèses. Ceci se produit plus souvent en comparaisons multiples ; pour la comparaison de plusieurs groupes avec un contrôle ou lors d'étude de recherche de dose maximale (*dose-finding studies*).

2.3.2. Procédure *Step-Down Dunnett*

D'abord proposée par Miller [46], cette procédure a été développée par Naik [50] et Marcus, Peritz, et Gabriel [43]. Cette procédure s'apparente à celle de *Holm*. Alors que la méthode de *Holm* utilise les valeurs- p univariées ordonnées, le *Step-Down Dunnett* utilise les statistiques de test ordonnées et comparent celles-ci à des valeurs critiques issues d'une distribution de *Student* multivariée de type I (voir annexe C.1).

L'algorithme pour la procédure *Step-Down Dunnett* est le suivant [69] :

- (1) On ordonne les statistiques- t , $T_{(1)} \leq T_{(2)} \leq \dots \leq T_{(m)}$, qui correspondent aux hypothèses ordonnées $H_0^{(1)}, H_0^{(2)}, \dots, H_0^{(m)}$.
- (2) On rejette $H_0^{(m)}$ si $T_{(m)} > t_{m,\nu,\rho}^{(\alpha)}$, puis on passe à l'étape suivante. Sinon, la procédure s'arrête et on ne rejette aucune des hypothèses nulles.
- (3) On rejette $H_0^{(m-1)}$ si $T_{(m-1)} > t_{(m-1),\nu,\rho}^{(\alpha)}$, puis on passe à l'étape suivante. Sinon, la procédure s'arrête et on ne rejette aucune des hypothèses nulles $H_0^{(j)}$, $j = 1, 2, \dots, m - 1$.
- (4) En général, on rejette $H_0^{(i)}$ si $T_{(i)} > t_{i,\nu,\rho}^{(\alpha)}$. Sinon, la procédure s'arrête et on ne rejette aucune des hypothèses nulles $H_0^{(j)}$, $j = 1, 2, \dots, i$.
- (5) Si toutes les $m - 1$ premières hypothèses ont été rejetées, on rejette $H_0^{(1)}$ si $T_{(1)} > t_{\nu}^{(\alpha)}$. Dans le cas contraire, seule $H_0^{(1)}$ n'est pas rejetée.

Cette méthode tient donc compte de la corrélation entre les tests, contrairement à celle de Holm. Pour cette raison, cette procédure est légèrement plus puissante que celle de Holm.

2.4. AUTRE PROCÉDURE *Step-Wise*

2.4.1. Procédure *Step-Up-Down* de Tamhane-Liu-Dunnnett

Une généralisation des procédures *Step-Up* et *Step-Down* a été proposée par Tamhane, Liu & Dunnnett [70]. Elle permet de répondre à la question : « Est-ce qu'au moins r hypothèses peuvent être rejetées ? ». La procédure *Step-Up-Down* permet de répondre à cette question en utilisant la statistique $t_{(q)}$, où $q = m + 1 - r$ et $1 \leq q \leq m$. La procédure teste d'abord l'hypothèse $H_0^{(q)}$ qui correspond à la statistique $t_{(q)}$; si celle-ci est rejetée, alors $H_0^{(q+1)}, \dots, H_0^{(m)}$ sont aussi rejetés et la procédure continue en procédure *Step-Down* jusqu'à ce qu'une hypothèse soit « acceptée ». Si par contre $H_0^{(q)}$ est « acceptée », alors $H_0^{(q-1)}, \dots, H_0^{(1)}$ sont aussi « acceptées » et la procédure continue en procédure *Step-Up* jusqu'à ce qu'une hypothèse soit rejetée.

L'algorithme pour cette procédure est donc le suivant [70] :

- (1) On ordonne les statistiques- t , $T_{(1)} \leq T_{(2)} \leq \dots \leq T_{(m)}$, qui correspondent aux hypothèses ordonnées $H_0^{(1)}, H_0^{(2)}, \dots, H_0^{(m)}$. Puis on choisit les constantes critiques $c_1 \leq c_2 \leq \dots \leq c_m$ (Voir Théorème 2.4.1).
- (2) (A) Si $t_{(q)} \leq c_q$, on ne rejette pas $H_0^{(1)}, H_0^{(2)}, \dots, H_0^{(q)}$, puis on poursuit en *Step-Up* avec l'étape 3 (A).
 (B) Si $t_{(q)} > c_q$, on rejette $H_0^{(q)}, H_0^{(q+1)}, \dots, H_0^{(m)}$, puis on poursuit en *Step-Down* avec l'étape 3 (B).
- (3) (A) Soit $H_0^{(i)}$ la dernière hypothèse acceptée (en 2 (A) $i = q$). Si $i = m$ la procédure est terminée, dans le cas contraire on teste $H_0^{(i+1)}$. Si $t_{(i+1)} > c_{i+1}$, on rejette $H_0^{(i+1)}, H_0^{(i+2)}, \dots, H_0^{(m)}$ et la procédure est terminée. Si, par contre, $t_{(i+1)} \leq c_{i+1}$, on ne rejette pas $H_0^{(i+1)}$, puis on recommence l'étape 3(A) en posant $i = i + 1$.
 (B) Soit $H_0^{(i)}$ la dernière hypothèse rejetée (en 2 (B) $i = q$). Si $i = 1$ la procédure est terminée, dans le cas contraire on teste $H_0^{(i-1)}$. Si $t_{(i-1)} \leq c_{i-1}$, on accepte $H_0^{(i-1)}, H_0^{(i-2)}, \dots, H_0^{(1)}$ et la procédure est terminée. Si, par contre, $t_{(i-1)} > c_{i-1}$, on rejette $H_0^{(i-1)}$, puis on recommence l'étape 3(B) en posant $i = i - 1$.

Il est alors facile de remarquer que lorsque $r = 1$ la procédure se résume à la procédure *Step-Down Dunnett* et pour $r = m$ nous retrouvons la procédure *Step-Up* de Dunnett-Tamhane.

Théorème 2.4.1 (Tamhane & al. (1998) [70], thm 4.1, p.356). *Les constantes critiques $c_1 \leq c_2 \leq \dots \leq c_k$ qui doivent satisfaire l'équation :*

$$\mathbf{P}(\text{accepter } H_0^1, \dots, H_0^k) \geq 1 - \alpha, \quad \text{pour } k = 1, \dots, m$$

s'obtiennent par la résolution de l'équation :

$$\mathbf{P} \left\{ (T_1, T_2, \dots, T_k) \leq \underbrace{(c_q, \dots, c_q)}_q, c_{q+1}, \dots, c_k \right\} = 1 - \alpha, \quad \text{pour } k = q + 1, \dots, m, \quad (2.4.1)$$

où $c_q = t_{q, \nu, \rho}^{(\alpha)}$ (et $c_1 = t_{\nu}^{(\alpha)}$) et où la notation $(T_1, T_2, \dots, T_k) \leq (c_1, c_2, \dots, c_k)$ signifie que le plus petit T_i est $\leq c_1$ et le suivant est $\leq c_2$ et ainsi de suite.

Tamhane & al. (1998) mentionnent que l'équation (2.4.1) peut être résolue de manière récursive en commençant par $k = q + 1$ et donnent une expression de (2.4.1) pour l'évaluer de façon plus efficace à l'aide d'un ordinateur.

Chapitre 3

RAPPELS POUR LE CAS D'UNE HYPOTHÈSE UNIQUE SUR LA DIFFÉRENCE ENTRE DEUX MOYENNES

Il est utile de faire un petit rappel sur la comparaison de deux moyennes pour une seule hypothèse. Supposons une expérience où l'on compare deux traitements. Notons X_{ij} les valeurs observées du $j^{\text{ème}}$ sujet du $i^{\text{ème}}$ traitement, où $j = 1, \dots, n_i$ et $i = 1, 2$. Supposons aussi que toutes les valeurs X_{ij} sont indépendantes et suivent une loi normale de moyenne μ_i et de variance σ^2 . Nous définissons $d = \mu_1 - \mu_2$ comme étant la vraie différence entre les moyennes et δ_h comme l'hypothétique vraie différence. En général, la valeur de δ_h sera issue d'études antérieures. Toutefois, dans la littérature, on pose souvent $\delta_h = 0$ pour alléger la notation et simplifier les calculs. Nous définissons aussi Δ_c comme étant la borne clinique, ou seuil clinique établi par le chercheur pour confirmer que la différence est significative. Il est donc logique et nécessaire que Δ_c soit supérieur ou égal à δ_h .

Lors de la comparaison de deux moyennes, il existe plusieurs possibilités de tests pouvant être effectués; tests classiques, tests de non-infériorité, tests de supériorité et tests d'équivalence.

3.1. TEST D'ÉGALITÉ

L'objectif du test d'égalité peut être de vérifier s'il n'y a aucune différence entre les deux moyennes, donc $\Delta_c = 0$, ou bien de vérifier que si la valeur supposée

de l'hypothétique vraie différence se confirme, donc $\Delta_c = \delta_h$. Les hypothèses prennent la forme suivante :

$$H_0 : \mu_1 - \mu_2 = \Delta_c \text{ vs } \begin{cases} H_1 : \mu_1 - \mu_2 \neq \Delta_c \text{ (test bilatéral)} \\ H_1 : \mu_1 - \mu_2 > (<) \Delta_c \text{ (test unilatéral)} \end{cases} \quad (3.1.1)$$

Lorsque σ^2 est connue, l'hypothèse nulle pour un test bilatéral est rejetée au niveau α si :

$$\left| \frac{\bar{X}_1 - \bar{X}_2}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| > z_{1-\alpha/2}, \quad (3.1.2)$$

où $z_{1-\alpha/2}$ est la valeur critique de niveau $\alpha/2$ d'une loi Normale(0,1). Pour effectuer un test unilatéral, il suffit d'utiliser la valeur critique $z_{1-\alpha}$ dans la formule suivante :

$$\frac{\bar{X}_1 - \bar{X}_2}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > z_{1-\alpha}. \quad (3.1.3)$$

Soit le paramètre de décentralité $\delta = \frac{\mu_1 - \mu_2}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$. La puissance du test sous l'hypothèse alternative bilatérale prend la forme :

$$\Phi(\delta - z_{1-\alpha/2}) + \Phi(-\delta - z_{1-\alpha/2}) \quad (3.1.4)$$

qui peut être approchée par

$$\Phi(|\delta| - z_{1-\alpha/2}),$$

en négligeant le terme de (3.1.4) dont la valeur est $\leq \alpha/2$. Il en découle que la taille échantillonale peut être obtenue par la résolution de la formule :

$$|\delta| - z_{1-\alpha/2} = z_{1-\beta}.$$

Il est à noter qu'en pratique, il est parfois préférable, pour des raisons éthiques, d'allouer plus de sujets au groupe contrôle qu'au groupe recevant le traitement ; nous fixons ainsi le ratio $r = n_1/n_2$. La formule pour calculer la taille n_i prend donc la forme générale suivante :

$$n_1 = rn_2 \text{ et } n_2 = \frac{(1 + 1/r)\sigma^2(z_{1-\alpha/2} + z_{1-\beta})^2}{(\mu_1 - \mu_2)^2}. \quad (3.1.5)$$

En général, comme la différence $d = \mu_1 - \mu_2$ est inconnue, nous devons en utiliser une approximation $\delta_h \neq 0$, issue de la littérature ou d'études antérieures.

Lorsque σ^2 est inconnue, on peut la remplacer par l'estimateur de la variance s^2 . L'hypothèse nulle est alors rejetée si :

$$\left| \frac{\bar{X}_1 - \bar{X}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| > t_{1-\alpha/2, n_1+n_2-2}, \quad (3.1.6)$$

où $t_{1-\alpha/2, n_1+n_2-2}$ est la valeur critique de niveau $\alpha/2$ d'une loi t de *Student* avec $n_1 + n_2 - 2$ degrés de liberté.

Sous l'hypothèse nulle la puissance de ce test peut donc s'écrire [9] :

$$1 - T_{n_1+n_2-2} \left(t_{1-\alpha/2, n_1+n_2-2} | \delta \right) - T_{n_1+n_2-2} \left(t_{1-\alpha/2, n_1+n_2-2} | -\delta \right).$$

en négligeant le terme dont la valeur est $\leq \alpha/2$, la puissance est approximativement :

$$1 - T_{n_1+n_2-2} \left(t_{1-\alpha/2, n_1+n_2-2} \left| \frac{|\mu_1 - \mu_2|}{\sigma\sqrt{1/n_1 + 1/n_2}} \right. \right),$$

En utilisant le ratio $r = n_1/n_2$, la taille nécessaire peut être obtenue en trouvant la solution de :

$$T_{(1+r)n_2-2} \left(t_{1-\alpha/2, (1+r)n_2-2} \left| \frac{\sqrt{n_2}|\mu_1 - \mu_2|}{\sigma\sqrt{1 + 1/r}} \right. \right) = \beta. \quad (3.1.7)$$

Le calcul de la taille requiert alors l'utilisation de tables disponibles dans plusieurs ouvrages, notamment ceux de Cohen [11] et Chow, Shao et Wang [9]. L'emploi de logiciels de calcul tels que PASS, SAS, Splus ou GPower 3.1 est très répandu et facilite grandement le calcul de la taille échantillonnale. L'équation (3.1.5) peut aussi servir d'approximation pour obtenir la taille nécessaire en (3.1.6) lorsque n_1 et n_2 sont grands.

3.2. TEST DE NON-INFÉRIORITÉ

Grouin et Coste [25] définissent la non-infériorité ainsi : « Le traitement 1 est dit *non inférieur* au traitement 2 si l'efficacité de 1 est *au moins équivalente* à celle de 2, c'est-à-dire si l'efficacité de 1 est soit équivalente à celle de 2, soit non équivalente et supérieur à celle de 2 ».

L'objectif est donc de vérifier si la moyenne du groupe 1 n'est pas plus *mauvaise* que celle du groupe 2, et les hypothèses prennent la forme suivante :

$$H_0 : \mu_1 - \mu_2 \leq -\Delta_c \text{ vs } H_1 : \mu_1 - \mu_2 > -\Delta_c, \quad (3.2.1)$$

où $\Delta_c > 0$. Lorsque σ^2 est connue, l'hypothèse nulle pour un test bilatéral est rejetée au niveau α si :

$$\frac{(\bar{X}_1 - \bar{X}_2) - (-\Delta_c)}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > z_{1-\alpha}, \quad (3.2.2)$$

où $z_{1-\alpha}$ est la valeur critique de niveau α d'une loi Normale(0,1). La puissance de ce test sous l'hypothèse alternative prend la forme :

$$\Phi \left(\frac{(\mu_1 - \mu_2) - (-\Delta_c)}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} - z_{1-\alpha} \right).$$

Il en découle que la taille échantillonnale peut être obtenue par la résolution de la formule :

$$\frac{(\mu_1 - \mu_2) + \Delta_c}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} - z_{1-\alpha} = z_{1-\beta},$$

donc

$$n_1 = rn_2 \text{ et } n_2 = \frac{(1 + 1/r)\sigma^2(z_{1-\alpha} + z_{1-\beta})^2}{((\mu_1 - \mu_2) + \Delta_c)^2}. \quad (3.2.3)$$

Et comme la différence $d = \mu_1 - \mu_2$ est en général inconnue, nous devons en utiliser une approximation $\delta_h \neq 0$ issue de la littérature ou d'études antérieures.

Lorsque σ^2 est inconnue, on peut la remplacer par s^2 . L'hypothèse nulle est alors rejetée si :

$$\frac{(\bar{X}_1 - \bar{X}_2) - (-\Delta_c)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > t_{1-\alpha, n_1+n_2-2}, \quad (3.2.4)$$

où $t_{1-\alpha, n_1+n_2-2}$ est la valeur critique de niveau α d'une loi *t* de *Student* avec $n_1 + n_2 - 2$ degrés de liberté. La taille nécessaire peut alors être obtenue en trouvant la solution de :

$$T_{n_1+n_2-2} \left(t_{1-\alpha, n_1+n_2-2} \left| \frac{(\mu_1 - \mu_2) + \Delta_c}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right. \right) = \beta. \quad (3.2.5)$$

Le calcul de la taille requiert alors l'utilisation de tables disponibles dans plusieurs ouvrages, notamment ceux de Cohen [11] et Chow, Shao et Wang [9], ou d'un logiciel comme mentionné précédemment.

3.3. TEST DE SUPÉRIORITÉ

L'objectif du test de supériorité est de vérifier si la moyenne du groupe 1 est *meilleure* que celle du groupe 2. Les hypothèses ressemblent beaucoup à celles du test de non-infériorité et prennent la forme suivante :

$$H_0 : \mu_1 - \mu_2 \leq \Delta_c \text{ vs } H_1 : \mu_1 - \mu_2 > \Delta_c, \quad (3.3.1)$$

avec $\Delta_c > 0$.

Pour le calcul de la taille, la théorie de la section précédente est toujours valide. Il suffit de remplacer $-\Delta_c$ par Δ_c . Ainsi, l'équation (3.2.2) devient :

$$\frac{(\bar{X}_1 - \bar{X}_2) - (\Delta_c)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} > z_{1-\alpha}$$

et la formule (3.2.3) devient :

$$n_1 = rn_2 \text{ et } n_2 = \frac{(1 + 1/r)\sigma^2(z_{1-\alpha} + z_{1-\beta})^2}{((\mu_1 - \mu_2) - \Delta_c)^2}.$$

Encore une fois, comme la différence $d = \mu_1 - \mu_2$ est inconnue, nous devons en utiliser une approximation $\delta_h \neq 0$ issue de la littérature ou d'études antérieures.

3.4. TEST D'ÉQUIVALENCE

Selon Grouin et Coste [25], « deux traitements 1 et 2 sont dits *équivalents*, si la différence d entre ces deux traitements est inférieure à une borne Δ_c en deçà de laquelle elle peut être considérée comme négligeable du point de vue clinique. Plus précisément, la différence d doit être plus petite en valeur absolue que la borne $\Delta_c > 0$: $|d| < \Delta_c$ ou $-\Delta_c < d < \Delta_c$ ». Ils ajoutent aussi qu'il est parfois pertinent d'introduire une dissymétrie dans les bornes, soit $-\Delta_1 < d < \Delta_2$.

Les hypothèses sont donc de la forme suivante :

$$H_0 : |\mu_1 - \mu_2| \geq \Delta_c \text{ vs } H_1 : |\mu_1 - \mu_2| < \Delta_c. \quad (3.4.1)$$

Pour tester ces hypothèses, la méthode la plus courante est celle du double test unilatéral. Nous pouvons en effet réécrire les hypothèses ainsi :

$$\begin{aligned} H_0^1 : \mu_1 - \mu_2 \leq -\Delta_c \text{ vs } H_1^1 : \mu_1 - \mu_2 > -\Delta_c \\ \text{et} \\ H_0^2 : \mu_1 - \mu_2 \geq \Delta_c \text{ vs } H_1^2 : \mu_1 - \mu_2 < \Delta_c \end{aligned} \quad (3.4.2)$$

L'hypothèse (3.4.1) sera rejetée au niveau α si les deux hypothèses en (3.4.2) sont chacune rejetées au niveau $\alpha/2$. On remarque que l'hypothèse H_0^1 est en fait celle de non-infériorité telle que vue précédemment, alors que H_0^2 est l'hypothèse de supériorité.

Étant donné que le test se fait en deux étapes, il n'existe aucune formule explicite pour le calcul de la taille exacte [76].

Toutefois, [9] proposent une approximation qui ressemble aux formules des sections précédentes. En effet, lorsque σ^2 est connue la puissance du test d'équivalence sous l'hypothèse nulle est :

$$\begin{aligned} & \Phi\left(\frac{\Delta_c - (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} - z_{1-\alpha}\right) + \Phi\left(\frac{\Delta_c + (\mu_1 - \mu_2)}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} - z_{1-\alpha}\right) - 1 \\ & \approx 2\Phi\left(\frac{\Delta_c - |\mu_1 - \mu_2|}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} - z_{1-\alpha}\right) - 1. \end{aligned}$$

La formule pour calculer la taille n_i en utilisant le ratio $r = n_1/n_2$ prend donc la forme suivante :

$$n_1 = rn_2 \text{ et } n_2 = \frac{(1 + 1/r)\sigma^2(z_{1-\alpha} + z_{1-\beta/2})^2}{(\Delta_c - |\mu_1 - \mu_2|)^2}. \quad (3.4.3)$$

En général, comme la différence $d = \mu_1 - \mu_2$ est inconnue, nous devons en utiliser une approximation δ_h issue de la littérature ou d'études antérieures.

Lorsque σ^2 est inconnue, on peut la remplacer par l'estimateur de la variance s^2 et la puissance s'écrit plutôt ainsi [9] :

$$\begin{aligned} 1 - T_{n_1+n_2-2}\left(t_{1-\alpha, n_1+n_2-2} \left| \frac{\Delta_c - (\mu_1 - \mu_2)}{\sigma\sqrt{1/n_1 + 1/n_2}} \right| \right) - T_{n_1+n_2-2}\left(t_{1-\alpha, n_1+n_2-2} \left| \frac{\Delta_c + (\mu_1 - \mu_2)}{\sigma\sqrt{1/n_1 + 1/n_2}} \right| \right) \\ \geq 1 - 2T_{n_1+n_2-2}\left(t_{1-\alpha, n_1+n_2-2} \left| \frac{\Delta_c - |\mu_1 - \mu_2|}{\sigma\sqrt{1/n_1 + 1/n_2}} \right| \right). \end{aligned}$$

Ainsi, en utilisant le ratio $r = n_1/n_2$, la solution de la formule suivante donne une approximation plutôt conservatrice de la taille nécessaire :

$$T_{(1+r)n_2-2}\left(t_{1-\alpha, (1+r)n_2-2} \left| \frac{\sqrt{n_2}(\Delta_c - |\mu_1 - \mu_2|)}{\sigma\sqrt{1 + 1/r}} \right| \right) = \frac{\beta}{2}. \quad (3.4.4)$$

Chapitre 4

PRÉSENTATION DE LA PROBLÉMATIQUE

Considérons une expérience avec un groupe traitement (1) et un groupe contrôle (2), avec n_1 et n_2 sujets respectivement. Supposons qu'il y a $m \geq 2$ variables d'intérêt mesurées pour chaque sujet. On note le vecteur de données aléatoires provenant du groupe i par $\underline{X}_{ij} = (X_{ij1}, X_{ij2}, \dots, X_{ijm})^\top$ avec $i = 1, 2$ et $j = 1, 2, \dots, n_i$. On suppose que les \underline{X}_{ij} sont indépendants, issus d'une distribution normale m -multivariée avec un vecteur de moyenne $\underline{\mu}_i = (\mu_{i1}, \mu_{i2}, \dots, \mu_{im})^\top$, une matrice de variances-covariances $\underline{\Sigma}_i$ et une corrélation $\rho = \frac{Cov(X_{ijk}, X_{ijl})}{\sigma_{ik}\sigma_{il}}$ ($k \neq l$) que l'on suppose constante et connue. Cette dernière supposition est très présente dans la littérature [8], [69], [70].

TAB. 4.1. Schéma des données.

	Traitement ($i=1$)	Contrôle ($i=2$)
$k = 1 :$	$\{X_{111}, \dots, X_{1n_11}\}$ $\mathbb{E}(X_{1j1}) = \mu_{11}$	$\{X_{211}, \dots, X_{2n_21}\}$ $\mathbb{E}(X_{2j1}) = \mu_{21}$
$k = 2 :$	$\{X_{112}, \dots, X_{1n_12}\}$ $\mathbb{E}(X_{1j2}) = \mu_{12}$	$\{X_{212}, \dots, X_{2n_22}\}$ $\mathbb{E}(X_{2j2}) = \mu_{22}$
\vdots	\vdots	\vdots
$k = m :$	$\{X_{11m}, \dots, X_{1n_1m}\}$ $\mathbb{E}(X_{1jm}) = \mu_{1m}$	$\{X_{21m}, \dots, X_{2n_2m}\}$ $\mathbb{E}(X_{2jm}) = \mu_{2m}$

Soit $\Delta_k \geq 0$ le seuil clinique d'intérêt préalablement spécifié, et δ_k l'hypothétique vraie différence entre les moyennes pour la $k^{\text{ième}}$ variable. Ainsi, l'hypothèse pour tester la supériorité du traitement pour la $k^{\text{ième}}$ variable ($1 \leq k \leq m$) est la suivante :

$$H_0^k : \mu_{1k} - \mu_{2k} \leq \Delta_k \quad \text{vs} \quad H_{1k} : \mu_{1k} - \mu_{2k} > \Delta_k. \quad (4.0.5)$$

Posons $\bar{X}_{1k} = 1/n_1 \sum_{j=1}^{n_1} X_{1jk}$ et $\bar{X}_{2k} = 1/n_2 \sum_{j=1}^{n_2} X_{2jk}$ les moyennes échantillonnelles pour la $k^{\text{ième}}$ variable des groupes 1 et 2 respectivement.

Définissons la statistique- t pour tester la supériorité du traitement pour la $k^{\text{ième}}$ variable :

$$T_k = \frac{\bar{X}_{1k} - \bar{X}_{2k} - \Delta_k}{\sqrt{\widehat{\text{Var}}(\bar{X}_{1k} - \bar{X}_{2k})}}, \quad (4.0.6)$$

où l'indépendance entre les groupes permet de réécrire $\text{Var}(\bar{X}_{1k} - \bar{X}_{2k})$ ainsi :

$$\begin{aligned} \text{Var}(\bar{X}_{1k} - \bar{X}_{2k}) &= \text{Var}(\bar{X}_{1k}) + \text{Var}(\bar{X}_{2k}) \\ &= \frac{1}{n_1^2} \sum_{j=1}^{n_1} \text{Var}(X_{1jk}) + \frac{1}{n_2^2} \sum_{j=1}^{n_2} \text{Var}(X_{2jk}). \end{aligned} \quad (4.0.7)$$

Lorsque les variances $\text{Var}(X_{ijk})$ sont connues, T_k suit une distribution normale. Toutefois, lorsque la variance est inconnue, nous devons l'estimer. Survient alors quatre situations possibles.

- a) Premièrement, nous pouvons supposer que la variance est la même pour les deux groupes et pour toutes les variables, soit $\text{Var}(X_{ijk}) = \sigma^2$.
- b) Ensuite, nous pouvons supposer que les variances sont les mêmes pour toutes les variables, mais qu'elles diffèrent d'un groupe à l'autre ; $\text{Var}(X_{ijk}) = \sigma_i^2$.
- c) Nous pouvons aussi supposer que les variances sont les mêmes pour les deux groupes, mais qu'elles diffèrent d'une variable à l'autre, ainsi $\text{Var}(X_{ijk}) = \sigma_k^2$.
- d) Finalement, nous pouvons supposer que les variances sont différentes entre les deux groupes et différentes pour toutes les variables, soit $\text{Var}(X_{ijk}) = \sigma_{ik}^2$.

Nous traitons ces quatre cas dans les sections qui suivent, c'est-à-dire que nous donnons l'estimateur de $\text{Var}(\bar{X}_{1k} - \bar{X}_{2k})$, ainsi que la loi de $\underline{T} = (T_1, \dots, T_m)^\top$, lorsque cela est possible.

4.1. CAS OÙ $\text{VAR}(X_{ijk}) = \sigma^2$

Lorsque nous considérons que toutes les variances sont égales, mais inconnues, alors (4.0.7) devient :

$$\text{Var}(\bar{X}_{1k} - \bar{X}_{2k}) = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}.$$

Lorsque σ^2 est inconnue, la statistique (4.0.6) peut donc être réécrite comme suit :

$$T_k = \frac{\bar{X}_{1k} - \bar{X}_{2k} - \Delta_k}{\hat{\sigma} \sqrt{1/n_1 + 1/n_2}}, \quad (4.1.1)$$

où $\hat{\sigma}$ est défini ci-après. Posons $\bar{X}_i = 1/n_i \sum_{j=1}^{n_i} X_{ij}$. L'estimateur $\hat{\sigma}$ de la variance dépend de la structure présumée de $\underline{\Gamma} = \text{Var}(\bar{X}_1 - \bar{X}_2)$, qui peut prendre différentes formes. Nous ne nous pencherons que sur deux structures particulières, car celles-ci sont les plus souvent utilisées [48], [49] et elles facilitent grandement le calcul d'un estimateur pour σ^2 .

La plus simple expression de cette matrice, lorsque $\underline{\Sigma}_i = \underline{\Sigma} = \sigma^2 \mathcal{I}_m$, est $\underline{\Gamma} = \sigma^2(1/n_1 + 1/n_2) \mathcal{I}_m$, où \mathcal{I}_m est la matrice identité d'ordre m . Dans cette situation, toutes les variances sont égales et la corrélation ρ entre les variables est nulle. Nous sommes alors en présence de *multisample sphericity* et l'estimateur sans biais $\hat{\sigma}^2$ prend la forme suivante (preuve en annexe D.1) :

$$\hat{\sigma}^2 = \frac{1}{m(n_1 + n_2 - 2)} \text{tr} \left\{ (\underline{Y} - \underline{\hat{\mu}})(\underline{Y} - \underline{\hat{\mu}})^\top \right\}$$

où $\text{tr}(\cdot)$ est l'opérateur de trace matricielle et où

$$\underline{\underline{Y}} = \begin{pmatrix} \underline{\underline{X}}_{11}^\top \\ \vdots \\ \underline{\underline{X}}_{1n_1}^\top \\ \underline{\underline{X}}_{21}^\top \\ \vdots \\ \underline{\underline{X}}_{2n_2}^\top \end{pmatrix} \quad \text{et} \quad \underline{\underline{\hat{\mu}}} = \begin{pmatrix} \underline{\underline{\bar{X}}}_1^\top \\ \vdots \\ \underline{\underline{\bar{X}}}_1^\top \\ \underline{\underline{\bar{X}}}_2^\top \\ \vdots \\ \underline{\underline{\bar{X}}}_2^\top \end{pmatrix}.$$

Les matrices $\underline{\underline{\Sigma}}_i = \underline{\underline{\Sigma}}$ peuvent aussi prendre la forme :

$$\underline{\underline{\Sigma}} \equiv \sigma^2 K_\rho = \sigma^2 [(1 - \rho)\mathcal{I}_m + \rho J] = \begin{pmatrix} \sigma^2 & \dots & \sigma^2 \rho \\ . & \sigma^2 & . \\ \sigma^2 \rho & \dots & \sigma^2 \end{pmatrix}, \quad \text{avec} \quad J = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & 1 & \vdots \\ 1 & \dots & 1 \end{pmatrix}.$$

Dans cette situation, toutes les variances sont égales et toutes les corrélations ρ entre les variables sont les mêmes. Nous sommes alors en présence de *multisample compound symmetry*.

La matrice $\underline{\underline{\Gamma}}$ s'écrit alors $\sigma^2(1/n_1 + 1/n_2)[(1 - \rho)\mathcal{I}_m + \rho J]$ et l'estimateur sans biais $\hat{\sigma}^2$ prend la forme suivante (preuve en annexe D.1) :

$$\hat{\sigma}^2 = \frac{1}{(n_1 + n_2 - 2)m(1 - \rho)} \text{tr} \left\{ (\underline{\underline{Y}} - \underline{\underline{\hat{\mu}}}) \left[\mathcal{I}_m - \frac{\rho}{1 - \rho(1 - m)} J \right] (\underline{\underline{Y}} - \underline{\underline{\hat{\mu}}})^\top \right\}.$$

En annexe D.1, nous avons démontré que lorsque $\underline{\underline{\Sigma}}$ prend l'une des deux formes présentées ci-haut, les statistiques T_k de (4.1.1) suivent une loi de *Student* décentrée de paramètre $\delta_k = \frac{\mu_{1k} - \mu_{2k} - \Delta_k}{\sigma \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ avec $(n_1 + n_2 - 2)m$ degrés de liberté.

Nous pouvons donc écrire le vecteur $\underline{\underline{T}} = (T_1, \dots, T_m)^\top$, qui contient les statistiques des m hypothèses, comme suit :

$$\underline{\underline{T}} = \left(\frac{Y_1}{\sqrt{S/\nu}}, \dots, \frac{Y_m}{\sqrt{S/\nu}} \right)^\top,$$

où le vecteur $(Y_1, \dots, Y_m)^\top$ suit une loi multi-normale $\mathcal{N}_m(\underline{\delta}, K_\rho)$, pour laquelle K_ρ est la matrice de covariance des Y_j , et où S est une variable aléatoire de loi χ^2 à $\nu = (n_1 + n_2 - 2)m$ degrés de liberté qui est indépendante des Y_j ($1 \leq j \leq m$). La distribution de \underline{T} est donc celle d'une loi de *Student* m -multivariée de type I (de type Kshirsagar) avec ν degrés de liberté, ce que l'on écrit :

$$\underline{T} \sim t_m^I(\nu, \underline{\delta}, K_\rho).$$

4.2. CAS OÙ $\text{VAR}(X_{ijk}) = \sigma_i^2$

Nous traitons ici du cas où l'on suppose l'homoscédasticité à l'intérieur des groupes et l'hétérosécédasticité entre les groupes de traitements. À partir de l'équation (4.0.7), nous obtenons l'expression :

$$\text{Var}(\bar{X}_{1k} - \bar{X}_{2k}) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

La statistique (4.0.6) peut donc être réécrite comme suit, lorsque les variances sont inconnues :

$$T_k = \frac{\bar{X}_{1k} - \bar{X}_{2k} - \Delta_k}{\sqrt{\hat{\sigma}_1^2/n_1 + \hat{\sigma}_2^2/n_2}}, \quad (4.2.1)$$

où les estimateurs $\hat{\sigma}_1^2$ et $\hat{\sigma}_2^2$ sont définis ci-dessous.

Comme dans la section précédente, l'estimateur $\hat{\sigma}_i$ dépendra de la structure présumée de $\underline{\Sigma}_i$, la matrice des variances-covariances dans le groupe i .

Si nous supposons la présence de *multisample sphericity* dans chaque groupe, soit $\underline{\Sigma}_i = \sigma_i^2 \mathcal{I}_m$, alors les estimateurs $\hat{\sigma}_i^2$ seront de la forme suivante (preuve en annexe D.2) :

$$\hat{\sigma}_i^2 = \frac{1}{(n_i - 1)m} \text{tr} \left\{ (\underline{Y}_i - \underline{\hat{\mu}}_i)(\underline{Y}_i - \underline{\hat{\mu}}_i)^\top \right\},$$

où

$$\underline{Y}_i = \begin{pmatrix} \underline{X}_{i1}^\top \\ \vdots \\ \underline{X}_{in_i}^\top \end{pmatrix} \text{ et } \underline{\hat{\mu}}_i = \begin{pmatrix} \bar{X}_i^\top \\ \vdots \\ \bar{X}_i^\top \end{pmatrix} \text{ avec } n_i \bar{X}_i = \underline{Y}_i \mathbf{1}.$$

Si nous supposons être en présence de *multisample compound symmetry* dans chaque groupe, la matrice $\underline{\underline{\Sigma}}_i$ prendra la forme $\underline{\underline{\Sigma}}_i = \sigma_i^2[(1 - \rho)\mathcal{I}_m + \rho J]$. L'expression pour l'estimateur $\hat{\sigma}_i^2$ sera donc la suivante (preuve en annexe D.2) :

$$\hat{\sigma}_i^2 = \frac{1}{(n_i - 1)m(1 - \rho)} \text{tr} \left\{ (\underline{Y}_i - \underline{\hat{\mu}}_i) \left[\mathcal{I}_m - \frac{\rho}{1 - \rho(1 - m)} J \right] (\underline{Y}_i - \underline{\hat{\mu}}_i)^\top \right\}.$$

D'après l'approximation de Welch, citée par Feiveson et Delaney [23], la statistique (4.2.1) suit approximativement une loi de *Student* avec f degrés de liberté :

$$T_k = \frac{\bar{X}_{1k} - \bar{X}_{2k} - \Delta_k}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} \approx t_f(\delta_k),$$

où f est défini par la relation suivante :

$$f = \frac{m \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right)^2}{\frac{\sigma_1^4}{n_1^2(n_1-1)} + \frac{\sigma_2^4}{n_2^2(n_2-1)}},$$

et où

$$\delta_k = \frac{\mu_{1k} - \mu_{2k} - \Delta_k}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

Ainsi, nous pouvons approximer la distribution de $\underline{T} = (T_1, \dots, T_m)^\top$ par une loi de *Student* m -multivariée de type I avec f degrés de liberté :

$$\underline{T} \approx t_m^I(f, \underline{\delta}, K_\rho).$$

4.3. CAS OÙ $\text{VAR}(X_{ijk}) = \sigma_k^2$

Ici, nous retrouvons le cas d'une seule hypothèse univariée homoscédastique et l'équation (4.0.7) s'écrit :

$$\text{Var}(\bar{X}_{1k} - \bar{X}_{2k}) = \frac{\sigma_k^2}{n_1} + \frac{\sigma_k^2}{n_2}.$$

Nous pouvons estimer σ_k^2 à l'aide de la variance appariée (dite «poolée») de l'échantillon, notée S_k^2 [58] :

$$\hat{\sigma}_k^2 = S_k^2 = \frac{(n_1 - 1)S_{1k}^2 + (n_2 - 1)S_{2k}^2}{n_1 + n_2 - 2} \sim \chi_{n_1 + n_2 - 2}^2,$$

avec $S_{ik}^2 = \frac{1}{n_i-1} \sum_{j=1}^{n_i} (X_{ijk} - \bar{X}_{ik})^2$, $i = 1, 2$. Il est donc facile de voir que T_k de (4.0.6) devient :

$$T_k = \frac{\bar{X}_{1k} - \bar{X}_{2k} - \Delta_k}{\hat{\sigma}_k \sqrt{1/n_1 + 1/n_2}} \sim t_{n_1+n_2-2}(\delta_k).$$

La statistique T_k suit une loi de *Student* décentrée de paramètre δ_k avec $n_1 + n_2 - 2$ degrés de liberté. Nous pouvons donc écrire le vecteur $\underline{T} = (T_1, \dots, T_m)^\top$ qui contient les statistiques des m hypothèses comme suit :

$$\underline{T} = \left(\frac{Y_1}{\sqrt{S_1/\nu}}, \dots, \frac{Y_m}{\sqrt{S_m/\nu}} \right), \quad (4.3.1)$$

où le vecteur $(Y_1, \dots, Y_m)^\top$ suit une loi multi-normale $\mathcal{N}_m(\underline{\delta}, \underline{\Sigma})$ et les S_1, \dots, S_m sont des variables aléatoires de loi χ^2 à $\nu = n_1 + n_2 - 2$ degrés de liberté. La distribution de \underline{T} est celle d'une loi de *Student* m -multivariée de type II avec ν degrés de liberté (voir annexe C.2).

Selon Hasler et Horton [27], la distribution de \underline{T} peut être approximée par une loi de *Student* m -multivariée de type I :

$$\underline{T} \approx t_m^I(\nu, \underline{\delta}, \underline{R}),$$

où la matrice de corrélation \underline{R} vaut $D_{\underline{\Sigma}}^{-1/2} \underline{\Sigma} D_{\underline{\Sigma}}^{-1/2}$ et où $D_{\underline{\Sigma}}$ est la matrice diagonale $m \times m$ des éléments diagonaux de $\underline{\Sigma} = \underline{\Sigma}_1/n_1 + \underline{\Sigma}_2/n_2$. [36]

Toutefois, il est important de noter que les auteurs ne donnent aucune explication ni aucune source afin de justifier ce résultat. En outre, ceux-ci présentent un cas de comparaisons multiples, où l'on compare plusieurs groupes avec un contrôle ; nous pouvons donc supposer que les valeurs S_k au dénominateur sont plutôt semblables, soit $S_i \approx S_j$ pour $\forall i \neq j$, ce qui permet de penser que leur approximation est valable.

Il faut par contre se poser la question de savoir si l'approximation convient toujours en situation d'hypothèses multiples ? Et que vaut cette approximation lorsque les variances entre deux variables sont très différentes l'une de l'autre ; par exemple si $S_i \ll S_j$?

Ceci sera étudié au chapitre 7 au moyen de simulations de Monte-Carlo.

4.4. CAS OÙ $\text{VAR}(X_{ijk}) = \sigma_{ik}^2$

Lorsque nous supposons des variances non nécessairement égales entre les variables et entre les deux groupes, nous retrouvons alors pour chaque k , le cas d'une seule hypothèse univariée hétéroscédastique. L'équation (4.0.7) s'écrit alors ainsi :

$$\text{Var}(\bar{X}_{1k} - \bar{X}_{2k}) = \frac{\sigma_{1k}^2}{n_1} + \frac{\sigma_{2k}^2}{n_2}.$$

Les estimateurs $\hat{\sigma}_{ik}^2$ prennent la forme [58] :

$$\hat{\sigma}_{ik}^2 = S_{ik}^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (X_{ijk} - \bar{X}_{ik})^2.$$

Alors nous avons :

$$T_k = \frac{\bar{X}_{1k} - \bar{X}_{2k} - \Delta_k}{\left[\frac{S_{1k}^2}{n_1} + \frac{S_{2k}^2}{n_2} \right]^{1/2}} \approx t_{f_k}(\delta_k),$$

où $f_k = \frac{\left[\frac{S_{1k}^2}{n_1} + \frac{S_{2k}^2}{n_2} \right]^2}{\frac{(S_{1k}^2/n_1)^2}{n_1-1} + \frac{(S_{2k}^2/n_2)^2}{n_2-1}}$, qui est en général arrondi à l'entier le plus proche [64], [74].

De cette façon, la statistique T_k suit approximativement une loi de *Student* décentrée de paramètre δ_k avec f degrés de liberté [58]. Dans ce cas, il devient beaucoup plus difficile de trouver la distribution du vecteur $\underline{T} = (T_1, \dots, T_m)^\top$, car le nombre de degrés de liberté f_k n'est pas nécessairement le même pour tout k . Il serait donc risqué de supposer que \underline{T} suit approximativement une loi de *Student* m -multivariée de type II, qui par la suite serait elle-même approchée par une loi de *Student* m -multivariée de type I. Cette situation reste donc sans solution à ce jour et devra être étudiée plus en détails dans l'avenir.

Chapitre 5

ÉQUATIONS GÉNÉRALES POUR LA PUISSANCE

Dans cette section, nous présentons les formules pour le calcul de la puissance des trois méthodes présentes dans le *package* R que nous avons développé, et ce pour le cas *échangeable* et le cas *non-échangeable*.

Nous considérons que deux statistiques de test, T_i et T_j , sont *échangeables*, si elles répondent à la condition suivante : $P(T_i \leq x, T_j \leq y) = P(T_i \leq y, T_j \leq x), \forall x, y$. Dans le cas contraire, T_i et T_j sont dites *non-échangeables*. De cette façon, le cas dit *échangeable* sera celui où toutes les paires de statistiques T_i et T_j seront *échangeables*, alors que dans le cas dit *non-échangeable*, il y aura présence d'au moins une paire de statistiques *non-échangeables*. En pratique, le cas *non-échangeable* est beaucoup plus fréquent, car la *non-échangeabilité* survient lorsqu'il y a une différence de variances entre au moins deux variables d'intérêts et/ou une différence entre au moins deux seuils cliniques ; $\Delta_i \neq \Delta_j$.

Nous développons d'abord, une expression générale pour la puissance d'une procédure de tests multiples basée sur les statistiques- t_i ; inspirée des articles de Dunnett et Tamhane [20] et de Chen *et al.* [8]. La formule générale proposée par Dunnett et Tamhane [20] permet de calculer la puissance d'une procédure qui rejette au moins r des m_1 hypothèses nulles. En pratique, m_1 est inconnu ; seul le nombre total d'hypothèses m est connu. Chen *et al.*[8] mentionnent que lors d'essais cliniques on suppose généralement que le traitement est supérieur ou non-inférieur au contrôle pour tous les paramètres étudiés, donc $m_1 = m$.

Comme Chen *et al.*[8], nous supposons donc que toutes les m hypothèses sont fausses. Ainsi, la puissance $\Psi_{r,m}$ pour rejeter au moins r des m hypothèses peut s'écrire :

$$\begin{aligned}
\Psi_{r,m} &= \mathbf{P}(\text{rejeter au moins } r \text{ fausses hypothèses parmi } m) \\
&= \mathbf{P}\left(\bigcup_{j=0}^{m-r} \text{rejeter exactement } m-j \text{ fausses hypothèses parmi } m\right) \\
&= \sum_{j=0}^{m-r} \mathbf{P}(\text{rejeter exactement } m-j \text{ fausses hypothèses parmi } m) \\
&= \sum_{j=0}^{m-r} \mathbf{P}\left(\bigcup_{(l_1 < \dots < l_{m-j}) \in F_{m-j,m}} \text{rejeter uniquement } H_0^{l_1}, \dots, H_0^{l_{m-j}}\right) \\
&= \sum_{j=0}^{m-r} \sum_{(l_1 < \dots < l_{m-j}) \in F_{m-j,m}} \mathbf{P}(\text{rejeter uniquement } H_0^{l_1}, \dots, H_0^{l_{m-j}}) \quad (5.0.1)
\end{aligned}$$

où $F_{m-j,m} = \{(l_1 < \dots < l_{m-j}); l_k \in \{1, \dots, m\}, \forall k = 1, \dots, m-j\}$ et où il est facile de voir que $\text{card}(F_{m-j,m}) \equiv |F_{m-j,m}| = \binom{m}{m-j} = \binom{m}{j}$.

Dans le cas échangeable, quelque soit j ($1 \leq j \leq m$), la probabilité $\mathbf{P}(\text{rejeter uniquement } H_0^{l_1}, \dots, H_0^{l_{m-j}})$ est constante pour tous les éléments dans $F_{m-j,m}$. Alors, l'équation (5.0.1) se simplifie de la façon suivante :

$$\Psi_{r,m}^e = \sum_{j=0}^{m-r} \binom{m}{m-j} \mathbf{P}(\text{ne pas rejeter } H_0^1, \dots, H_0^j \text{ et rejeter } H_0^{j+1}, \dots, H_0^m), \quad (5.0.2)$$

où l'exposant e signifie échangeable. Dans ce mémoire, les hypothèses H_0^1, \dots, H_0^m sont de la forme (4.0.5) vue précédemment. Les formules (5.0.1) et (5.0.2) sont développées pour les procédures de *Bonferroni*, de *Hochberg* et de *Holm* dans les sous-sections qui suivent, afin de faciliter leur programmation.

5.1. ÉQUATION DE LA PUISSANCE POUR LA PROCÉDURE *Single Step Bonferroni*

5.1.1. Cas Non-échangeable

Rappelons que lors de la procédure *Single-Step Bonferroni*, toutes les hypothèses sont testées indépendemment l'une de l'autre au niveau α/m . De cette façon, une hypothèse k est rejetée lorsque $(T_k < c_{km})$ et n'est pas rejetée lorsque $(T_k \geq c_{km})$, où c_{km} est la valeur critique de niveau α/m associée à l'hypothèse k .

Donc, à partir de la formule (5.0.1), nous obtenons :

$$\Psi_{r,m}^s = \sum_{j=0}^{m-r} \sum_{(l_1 < \dots < l_{m-j}) \in F_{m-j,m}} \left\{ \mathbf{P} \left[\left(\bigcap_{k \in \mathcal{I}_m \setminus \{l_1, \dots, l_{m-j}\}} (T_k < c_{km}) \right) \cap \left(\bigcap_{l \in \{l_1, \dots, l_{m-j}\}} (T_l \geq c_{lm}) \right) \right] \right\} \quad (5.1.1)$$

où $\mathcal{I}_m = \{1, \dots, m\}$.

5.1.2. Cas échangeable

Dans cette section, nous développons la formule de puissance (5.0.2) pour le cas échangeable, pour la méthode *Single Step de Bonferroni*.

Il est à noter que dans le cas échangeable pour la méthode *Single Step Bonferroni*, on a : $\{\text{ne pas rejeter } H_0^1, \dots, H_0^j\} = \bigcap_{k=1}^j (T_k < c_m)$ et $\{\text{rejeter } H_0^{j+1}, \dots, H_0^m\} = \bigcap_{l=j+1}^m (T_l \geq c_m)$. Alors, l'équation (5.0.2) devient :

$$\Psi_{r,m}^{s,e} = \sum_{j=0}^{m-r} \binom{m}{j} \mathbf{P} \left[\left(\bigcap_{k=1}^j (T_k < c_m) \right) \cap \left(\bigcap_{l=j+1}^m (T_l \geq c_m) \right) \right], \quad (5.1.2)$$

où c_m est la valeur critique au niveau α/m qui est la même pour toutes les hypothèses puisque nous sommes dans le cas échangeable.

Pour des fins de programmation, nous développons le résultat en (5.1.2) afin d'obtenir une équation plus rapide à calculer. La formule pour la puissance devient donc (voir Annexe E.1) :

$$\begin{aligned} \Psi_{r,m}^{s,e} &= J_0 + \sum_{j=1}^{m-r} \binom{m}{j} \left\{ \mathbf{P} \left(\bigcap_{k=1}^j (T_k < c_m) \right) \right. \\ &\quad \left. - \sum_{t=1}^{m-j} (-1)^{t+1} \binom{m-j}{t} \mathbf{P} \left(\bigcap_{k=1}^{j+t} (T_k < c_m) \right) \right\}, \end{aligned} \quad (5.1.3)$$

où le terme J_0 , correspondant à $j = 0$ dans l'équation (5.1.2), s'exprime ainsi :

$$J_0 = 1 + \sum_{r=1}^m (-1)^r \binom{m}{r} \mathbf{P} \left(\bigcap_{k=1}^r (T_k < c_m) \right), \quad (5.1.4)$$

Il est possible de simplifier encore d'avantage l'équation (5.1.3) afin de lui donner la forme suivante (voir Annexe E.1) :

$$\Psi_{r,m}^{s,e} = 1 + \sum_{k=1}^m (a_k + b_k + c_k) P_k \quad \text{où} \quad \begin{cases} a_k = \binom{m}{k}, k = 1, \dots, m-r \\ a_k = 0, k = m-r+1, \dots, m \\ b_1 = 0 \\ b_k = \sum_{j=0}^{\min(k-1, m-r)} \binom{m}{j} \binom{-(m-k+1)}{k-j}, k = 2, \dots, m \\ c_k = (-1)^k \binom{m}{k}, k = 1, \dots, m \end{cases} \quad (5.1.5)$$

et où nous avons posé $P_k = \mathbf{P} \left(\bigcap_{j=1}^k T_j < c_m \right)$.

5.2. ÉQUATION DE LA PUISSANCE POUR LA PROCÉDURE *Step-Up Hochberg*

Dans cette section, nous développons les formules de puissance (5.0.1) et (5.0.2), pour la méthode *Step-Up* proposée par Hochberg [28].

Notons que les statistiques ordonnées $T_{1:m} \leq T_{2:m} \leq \dots \leq T_{m:m}$ correspondent aux hypothèses ordonnées $H_{0(1)}, H_{0(2)}, \dots, H_{0(m)}$.

Il est à noter que pour les méthodes *Step-Up*, on a : $\{\text{ne pas rejeter } H_{0(1)}, \dots, H_{0(j)}\} = \bigcap_{k=1}^j (T_{k:m} < u_k)$ et $\{\text{rejeter } H_{0(j+1)}, \dots, H_{0(m)}\} = \{\min(T_{j+1:m}, \dots, T_{m:m}) \geq u_{j+1}\} = \{T_{j+1:m} \geq u_{j+1}\}$.

L'équation pour la puissance prend donc la forme :

$$\Psi_{r,m}^u = \sum_{j=0}^{m-r} \mathbf{P} \left[\left(\bigcap_{k=1}^j (T_{k:m} < u_k) \right) \cap (T_{(j+1):m} \geq u_{j+1}) \right]. \quad (5.2.1)$$

Il est important de noter que lorsque $j = 0$, alors $\bigcap_{k=1}^j (T_{k:m} < u_k) = \Omega$, l'événement certain.

Pour des fins de programmation, nous développons le résultat (5.2.1), afin d'obtenir une équation qui ne comporte que des inégalités strictes ($<$) et afin de transformer les statistiques d'ordre $T_{k:m}$ en statistiques T_k . Nous obtenons (voir Annexe E.2) :

$$\Psi_{r,m}^u = 1 - (-1)^{\sum_{k=1}^{m-r+1} k} \sum_{\underline{a}=\underline{w}}^{a^*} (-1)^{a_+} \prod_{h=1}^{m-r+1} \binom{(\Delta a_h) - 1}{a_h - w_h} P_{\underline{a}}, \quad (5.2.2)$$

avec $\underline{w} = (1, 2, \dots, m - r + 1)$. Dans le cas non-échangeable, nous avons :

$$P_{\underline{a}} = \sum_{\underline{l}}' \mathbf{P} \left[\bigcap_{h=0}^{m-r} \left(\bigcap_{k=a_h+1}^{a_{h+1}} T_{l_k} < u_{h+1} \right) \right]$$

alors que dans le cas échangeable, nous obtenons plutôt :

$$P_{\underline{a}} = \binom{m}{\Delta^* \underline{a}} \mathbf{P} \left[\bigcap_{h=0}^{m-r} \left(\bigcap_{k=a_h+1}^{a_{h+1}} T_k < u_{h+1} \right) \right].$$

La puissance sera alors notée $\Psi_{r,m}^{u,e}$.

5.3. ÉQUATION DE LA PUISSANCE POUR LA PROCÉDURE *Step-Down Holm*

Dans cette section, nous développons les formules de puissance (5.0.1) et (5.0.2), pour la méthode *Step-Down* de *Holm* [30].

Notons que les statistiques ordonnées $T_{1:m} \leq T_{2:m} \leq \dots \leq T_{m:m}$ correspondent aux hypothèses ordonnées $H_{0(1)}, H_{0(2)}, \dots, H_{0(m)}$.

Il est à noter que pour les méthodes *Step-Down*, on a : $\{\text{ne pas rejeter } H_{0(1)}, \dots, H_{0(j)}\} = \{\max(T_{1:m}, \dots, T_{j:m}) < d_j\} = \{T_{j:m} < d_j\}$ et $\{\text{rejeter } H_{0(j+1)}, \dots, H_{0(m)}\} = \bigcap_{k=j+1}^m (T_{k:m} \geq d_k)$.

Nous pouvons donc exprimer la puissance de cette façon :

$$\Psi_{r,m}^d = \sum_{j=0}^{m-r} \mathbf{P} \left[(T_{j:m} < d_j) \cap \left(\bigcap_{k=j+1}^m (T_{k:m} \geq d_k) \right) \right]. \quad (5.3.1)$$

Il est important de noter que lorsque $j = 0$, alors $T_{j:m} < d_j = \Omega$, l'événement certain.

Pour des fins de programmation, nous développons le résultat (5.3.1), afin d'obtenir une équation qui ne comporte que des inégalités strictes ($<$) et afin de transformer les statistiques d'ordre $T_{k:m}$ en statistiques T_k . Nous obtenons (voir Annexe E.3) :

$$\begin{aligned} \Psi_{r,m}^d = & J_0 + \sum_{j=1}^{m-r} \left\{ \left[(-1)^j \sum_{a=j}^m (-1)^a \binom{a-1}{a-j} P_a \right] \right. \\ & - \sum_{t=1}^{m-j} (-1)^{t+1} \sum_{(i_1 < i_2 < \dots < i_t) \in F_{t,m-j}} \left[(-1)^{((t+1)j + \sum_{h=1}^t i_h)} \sum_{\underline{b}=\underline{w}_j}^{\underline{b}^*} (-1)^{b_+} \right. \\ & \left. \left. \prod_{h=1}^{t+1} \binom{(\Delta b_h) - 1}{b_h - (j + i_{h-1})} P_{\underline{b}} \right] \right\}, \quad (5.3.2) \end{aligned}$$

où $\underline{w}_j = (j + i_0, j + i_1, j + i_2, \dots, j + i_t)$ et $i_0 = 0$.

Le terme J_0 , correspondant à $j = 0$ dans l'équation (5.3.1), s'exprime ainsi :

$$J_0 = 1 - \left\{ \sum_{t=1}^m (-1)^{t+1} \sum_{(i_1 < i_2 < \dots < i_t) \in F_{t,m}} \left[(-1)^{\sum_{h=1}^t i_h} \sum_{\underline{c}=\underline{s}_j}^{\underline{c}^*} (-1)^{c_+} \prod_{h=1}^t \binom{(\Delta c_h) - 1}{c_h - i_h} P_{\underline{c}} \right] \right\}, \quad (5.3.3)$$

où $\underline{s}_j = (j + i_1, j + i_2, \dots, j + i_t) = (i_1, i_2, \dots, i_t)$.

Dans le cas non-échangeable, P_a , $P_{\underline{b}}$ et $P_{\underline{c}}$ s'expriment ainsi :

$$\begin{aligned} P_a &= \sum_{\underline{l}}' \mathbf{P} \left(\bigcap_{k=1}^a T_{l_k} < d_j \right), \\ P_{\underline{b}} &= \sum_{\underline{l}}' \mathbf{P} \left[\bigcap_{h=0}^t \left(\bigcap_{k=b_h+1}^{b_{h+1}} T_{l_k} < d_{j+i_h} \right) \right], \\ P_{\underline{c}} &= \sum_{\underline{l}}' \mathbf{P} \left[\bigcap_{h=0}^{t-1} \left(\bigcap_{k=c_h+1}^{c_{h+1}} T_{l_k} < d_{i_{h+1}} \right) \right]. \end{aligned}$$

Dans le cas échangeable, P_a , $P_{\underline{b}}$ et $P_{\underline{c}}$ s'expriment ainsi :

$$\begin{aligned} P_a &= \binom{m}{a} \mathbf{P} \left(\bigcap_{k=1}^a T_k < d_j \right), \\ P_{\underline{b}} &= \binom{m}{\Delta^* \underline{b}} \mathbf{P} \left[\bigcap_{h=0}^t \left(\bigcap_{k=b_h+1}^{b_{h+1}} T_k < d_{j+i_h} \right) \right], \\ P_{\underline{c}} &= \binom{m}{\Delta^* \underline{c}} \mathbf{P} \left[\bigcap_{h=0}^{t-1} \left(\bigcap_{k=c_h+1}^{c_{h+1}} T_k < d_{i_{h+1}} \right) \right], \end{aligned}$$

et la puissance sera notée $\Psi_{r,m}^{d,e}$.

Chapitre 6

ALGORITHME POUR LE CALCUL DE LA TAILLE D'ÉCHANTILLON

L'algorithme pour le calcul de la taille échantillonnale proposé dans Chen & al. [8] est plutôt simple. Nous avons d'abord programmé les formules de puissance du chapitre 5 qui s'applique à la problématique présentée au chapitre 4, plus particulièrement au cas où $Var(X_{ijk}) = \sigma_k^2$ vu en section 4.3. Puis, à l'aide de la fonction `uniroot()` de R, nous trouvons la taille n par groupe dont la puissance calculée, p_{cn} , est plus grande ou égale à la puissance désirée, $1 - \beta$; soit $\min_n \{p_{cn} \geq 1 - \beta\}$.

Nous avons repris les travaux de Chen & al. [8], afin de créer un *package* R permettant de calculer la taille échantillonnale en entrant les paramètres de l'expérience : le nombre de variables (m), le nombre de variables qui doivent être significatives (r), le niveau de significativité (α), la puissance ($1 - \beta$), les vecteurs des moyennes des 2 groupes ($\underline{\mu}_1$ et $\underline{\mu}_2$), le vecteur des variances *poolées* ($\underline{\sigma}^2$), la corrélation ρ entre les statistiques de test, le vecteur de seuils cliniques des variables ($\underline{\Delta}$) et la procédure d'analyse choisie. Les procédures disponibles dans le *package* sont les mêmes que dans Chen & al. [8], soit la procédure *Single Step* de *Bonferroni*, la procédure *Step-Down* de *Holm* et la procédure *Step-Up* de *Hochberg*. Il est à noter qu'il suffit de connaître la formule pour le calcul de la puissance d'une procédure pour ajouter celle-ci au *package*.

Chapitre 7

SIMULATIONS ET RÉSULTATS

7.1. REPRISE DES RÉSULTATS DE CHEN & AL.

Dans leur article, Chen & al. [8] se sont concentrés sur les cas particuliers $m = 2$ et $m = 3$. Étant donné que nos résultats généralisent leurs formules de puissance pour n'importe quelle valeur de m , nous avons pu nous rendre compte que ces auteurs avaient commis des erreurs en développant les formules de puissance pour la procédure *Step-Down* de *Holm* et la procédure *Step-Up* de *Hochberg*. Nous avons donc recalculé les tailles nécessaires pour les cas *échangeables* particuliers $m = 2$ et $m = 3$, en utilisant les mêmes paramètres que Chen & al. [8], à savoir $\Delta_k = 0, 2$ et $\widehat{\sigma}_k^2 = 1$. Nous sommes donc dans le cas de figure de la sous-section 4.1.

Les résultats que nous avons obtenus avec notre *package* sont présentés dans les tables 7.2, 7.4 et 7.6. Nous avons également reproduit les résultats obtenus par Chen & al. [8] dans les tables 7.1, 7.3 et 7.5.

TAB. 7.1. *Table 1* dans Chen & al. [8]

Tailles nécessaires dans chaque groupe pour atteindre la puissance indiquée pour les procédures de Bonferroni (Bonf), de Hochberg (HC) et de Holm (HM) sous différentes valeurs de corrélation ρ pour détecter un effet

$\Delta/\hat{\sigma} = 0,2$ au niveau $\alpha_{FWER} = 0,05$, cas *échangeable* où $m = 2$ et $r = 1$.

ρ	Puissance											
	0,80			0,85			0,90			0,95		
	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM
0	221	90	93	254	100	103	299	111	114	371	123	127
0,1	230	98	101	265	109	112	313	121	125	390	135	140
0,2	240	107	110	277	119	123	327	133	138	409	150	156
0,3	251	118	122	289	132	137	342	149	155	428	168	176
0,4	262	131	136	303	148	154	358	168	176	448	192	202
0,5	274	148	155	317	168	176	375	192	202	470	224	237
0,6	287	170	179	332	194	205	393	225	238	492	266	284
0,7	302	198	210	349	228	242	413	267	285	517	324	348
0,8	320	235	252	369	273	292	436	323	346	545	401	432
0,9	342	284	308	394	330	356	464	392	422	579	493	531

TAB. 7.2. 7.1 recalculée avec notre *package*

Tailles nécessaires dans chaque groupe pour atteindre la puissance indiquée pour les procédures de Bonferroni (Bonf), de Hochberg (HC) et de Holm (HM) sous différentes valeurs de corrélation ρ pour détecter un effet

$\Delta/\hat{\sigma} = 0,2$ au niveau $\alpha_{FWER} = 0,05$, cas *échangeable* où $m = 2$ et $r = 1$.

ρ	Puissance											
	0,80			0,85			0,90			0,95		
	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM
0	221	212	221	254	244	254	299	287	299	371	357	371
0,1	230	221	230	265	255	265	313	301	313	390	375	390
0,2	240	231	240	277	266	277	327	315	327	409	394	409
0,3	251	241	251	289	278	289	342	329	342	428	412	428
0,4	262	251	262	303	291	303	358	344	358	448	432	448
0,5	274	262	274	317	304	317	375	360	375	470	452	470
0,6	287	275	287	332	318	332	393	377	393	492	474	492
0,7	302	288	302	349	333	349	413	395	413	517	496	517
0,8	320	303	320	369	350	369	436	415	436	545	521	545
0,9	342	319	342	394	369	394	464	437	464	579	548	579

TAB. 7.3. *Table 2* dans Chen & al. [8]

Tailles nécessaires dans chaque groupe pour atteindre la puissance indiquée pour les procédures de Bonferroni (Bonf), de Hochberg (HC) et de Holm (HM) sous différentes valeurs de corrélation ρ pour détecter un effet

$\Delta/\widehat{\sigma} = 0,2$ au niveau $\alpha_{FWER} = 0,05$, cas *échangeable* où $m = 3$ et $r = 1$.

ρ	Puissance											
	0,80			0,85			0,90			0,95		
	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM
0	185	38	39	211	41	42	247	45	46	305	48	49
0,1	198	40	42	227	44	45	267	48	49	330	52	54
0,2	212	44	46	244	48	50	287	53	55	357	57	59
0,3	226	49	51	261	54	56	308	59	62	384	65	67
0,4	242	56	59	280	62	65	330	68	72	412	75	78
0,5	260	67	71	300	74	78	354	81	86	443	89	95
0,6	279	83	89	322	91	99	380	101	110	475	112	122
0,7	302	108	120	347	121	134	409	135	151	511	151	170
0,8	327	156	180	376	176	205	443	201	236	552	231	277
0,9	361	251	299	414	289	346	486	340	409	603	421	512

TAB. 7.4. 7.3 recalculée avec notre *package*

Tailles nécessaires dans chaque groupe pour atteindre la puissance indiquée pour les procédures de Bonferroni (Bonf), de Hochberg (HC) et de Holm (HM) sous différentes valeurs de corrélation ρ pour détecter un effet

$\Delta/\widehat{\sigma} = 0,2$ au niveau $\alpha_{FWER} = 0,05$, cas *échangeable* où $m = 3$ et $r = 1$.

ρ	Puissance											
	0,80			0,85			0,90			0,95		
	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM
0	185	179	185	212	205	212	248	239	248	305	294	305
0,1	198	192	198	228	220	228	267	258	267	331	320	331
0,2	212	205	212	244	236	244	287	278	287	357	345	357
0,3	227	220	227	261	253	262	308	298	308	384	372	384
0,4	243	235	243	280	271	280	331	320	331	413	399	413
0,5	260	251	260	300	290	300	354	342	355	443	428	443
0,6	280	269	280	322	310	322	381	367	381	475	459	475
0,7	302	289	302	347	333	347	410	393	410	511	492	511
0,8	328	311	328	377	358	377	443	422	443	552	527	552
0,9	362	334	362	415	385	415	486	453	486	603	566	603

TAB. 7.5. *Table 3* dans Chen & al. [8]

Tailles nécessaires dans chaque groupe pour atteindre la puissance indiquée pour les procédures de Bonferroni (Bonf), de Hochberg (HC) et de Holm (HM) sous différentes valeurs de corrélation ρ pour détecter un effet

$\Delta/\hat{\sigma} = 0,2$ au niveau $\alpha_{FWER} = 0,05$, cas *échangeable* où $m = 3$ et $r = 2$.

ρ	Puissance											
	0,80			0,85			0,90			0,95		
	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM
0	363	143	149	399	150	157	447	159	166	523	167	175
0,1	372	146	154	411	155	163	462	165	173	544	175	184
0,2	381	151	160	422	162	171	477	173	183	565	185	196
0,3	390	159	170	433	172	183	492	185	197	585	199	213
0,4	398	170	182	444	184	198	506	201	216	605	219	236
0,5	406	184	200	454	202	220	519	222	242	624	246	269
0,6	413	204	225	464	226	250	532	253	280	642	286	318
0,7	421	232	259	474	260	292	545	296	333	660	345	392
0,8	428	269	306	484	307	348	558	356	404	678	430	493
0,9	435	316	364	493	363	417	570	427	487	693	531	602

TAB. 7.6. 7.5 recalculée avec notre *package*

Tailles nécessaires dans chaque groupe pour atteindre la puissance indiquée pour les procédures de Bonferroni (Bonf), de Hochberg (HC) et de Holm (HM) sous différentes valeurs de corrélation ρ pour détecter un effet

$\Delta/\hat{\sigma} = 0,2$ au niveau $\alpha_{FWER} = 0,05$, cas *échangeable* où $m = 3$ et $r = 2$.

ρ	Puissance											
	0,80			0,85			0,90			0,95		
	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM
0	363	307	321	399	339	355	447	383	400	523	452	472
0,1	373	315	330	411	350	367	463	397	415	545	472	492
0,2	382	323	339	423	360	378	478	410	429	565	491	512
0,3	390	330	347	434	370	388	492	423	443	585	509	532
0,4	398	337	355	444	379	399	506	435	457	605	527	551
0,5	406	343	363	455	387	409	520	447	471	624	543	570
0,6	414	349	371	465	395	419	533	458	484	643	559	588
0,7	421	353	379	474	402	429	546	467	497	661	574	607
0,8	428	357	387	484	407	440	558	475	511	679	586	626
0,9	436	357	397	493	408	452	571	479	526	696	594	647

Nous remarquons d’abord que pour la méthode de *Bonferroni* nous obtenons passablement les mêmes résultats que Chen & al. [8]. Les quelques écarts sont dus au fait que nous avons ajouté la fonction `set.seed(1)` dans notre programme, car la fonction `pmvt()` comporte une partie aléatoire.

Toutefois, pour les procédures de *Hochberg* et de *Holm*, les différences sont beaucoup plus importantes. Les tailles que nous avons calculées pour ces procédures sont plus grandes et s’approchent des tailles pour la méthode de *Bonferroni*. Ceci tombe sous le sens, puisque ces deux procédures ne sont que légèrement plus puissantes que celle de *Bonferroni*. De plus, lorsque $r = 1$ la procédure *Step-Down* de *Holm* requiert la même taille que celle de *Bonferroni*, ce qui est logique, car la première étape de cette méthode rejette la plus petite valeur-p au niveau α/m . Donc, $\mathbf{P}(\text{rejetter au moins une hypothèse}) = 1 - \mathbf{P}(\text{ne rejeter aucune hypothèse})$, or avec la procédure de *Holm* on ne rejette aucune hypothèse que si la plus petite valeur-p est supérieure à α/m tout comme la procédure de *Bonferroni*. Les quelques petites différences qui apparaissent parfois sont attribuables à la portion aléatoire de la fonction `pmvt()`.

7.2. IMPACT DE LA VARIATION DE r

Dans cette section, nous vérifions l’impact de la variation du nombre r d’hypothèses qui doivent être significatives sur le calcul de la taille. Nous sommes toujours dans le cas de figure de la sous-section 4.1. Nous utilisons le cas $m = 7$ avec r qui varie de 1 à 7 et les résultats sont présentés dans les tables 7.7 et 7.8.

TAB. 7.7. Impact de r , 1^{ère} partie $r = 1, 2, 3, 4$

Tailles nécessaires dans chaque groupe pour atteindre une puissance de 0,80 pour les procédures de Bonferroni (Bonf), de Hochberg (HC) et de Holm (HM) sous différentes valeurs de corrélation ρ et de r pour détecter un effet $\Delta/\hat{\sigma} = 0,2$ au niveau $\alpha_{FWER} = 0,05$, cas *échangeable* où $m = 7$.

ρ	r											
	1			2			3			4		
	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM
0	136	135	136	227	213	216	313	276	283	404	336	349
0,1	154	153	154	248	233	237	334	296	304	423	353	367
0,2	173	172	175	269	253	258	353	314	324	441	368	383
0,3	194	193	196	290	273	280	372	331	343	457	381	399
0,4	218	216	218	312	294	302	391	348	362	472	392	413
0,5	244	242	241	336	315	322	410	364	379	486	402	426
0,6	273	270	268	361	338	344	429	379	395	498	410	438
0,7	307	303	300	388	361	368	449	393	412	510	416	449
0,8	349	340	341	419	383	398	472	403	432	520	417	461
0,9	405	378	400	457	396	439	499	404	462	530	408	477

TAB. 7.8. Impact de r , 2^e partie $r = 5, 6, 7$

Tailles nécessaires dans chaque groupe pour atteindre une puissance de 0,80 pour les procédures de Bonferroni (Bonf), de Hochberg (HC) et de Holm (HM) sous différentes valeurs de corrélation ρ et de r pour détecter un effet $\Delta/\hat{\sigma} = 0,2$ au niveau $\alpha_{FWER} = 0,05$, cas *échangeable* où $m = 7$.

ρ	r								
	5			6			7		
	Bonf	HC	HM	Bonf	HC	HM	Bonf	HC	HM
0	512	397	418	660	473	500	931	616	629
0,1	528	410	432	667	479	508	920	607	623
0,2	540	419	443	672	482	513	906	596	614
0,3	550	426	453	673	483	516	889	582	603
0,4	558	432	460	672	481	517	868	565	590
0,5	564	436	466	669	476	515	844	546	575
0,6	569	438	473	662	467	512	816	523	557
0,7	572	436	478	652	455	506	782	496	536
0,8	574	428	482	637	438	498	740	463	514
0,9	570	404	483	614	410	491	684	419	493

Pour $r = 1, 2, 3$, nous remarquons que la corrélation et la taille suivent une relation croissante pour les trois méthodes. En effet, plus la corrélation augmente, plus la taille nécessaire augmente (voir table 7.7 et figure 7.1).

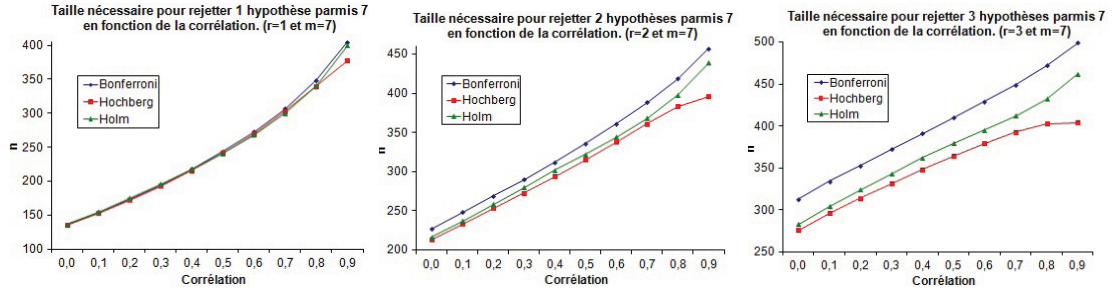


FIG. 7.1. Graphiques de la taille n en fonction de la corrélation ρ pour $m=7$ et $r=1,2,3$.

Pour $r = 4$, il y a toujours une relation linéaire croissante entre la corrélation et la taille pour les méthodes de *Bonferroni* et de *Holm*. Par contre, ce n'est plus le cas pour la méthode de *Hochberg*, où lorsque $\rho = 0,9$ la taille nécessaire $n = 408$ est inférieure à celle pour $\rho = 0,8$ (voir table 7.7 et figure 7.2).

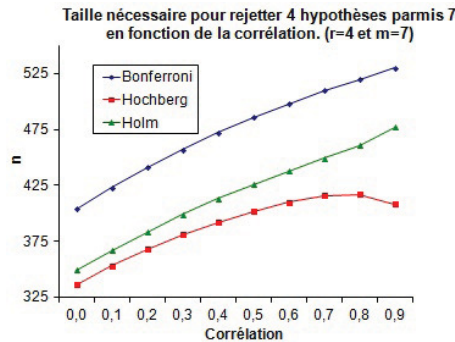


FIG. 7.2. Graphiques de la taille n en fonction de la corrélation ρ pour $m=7$ et $r=4$.

Pour $r = 5$, seule la méthode de *Holm* suit toujours une relation croissante. Les deux autres procédures semblent «plafonner» avant de redescendre (voir table 7.8 et figure 7.3).

Pour $r = 6$, les trois procédures présentent des courbes semblables ; la relation entre la corrélation et la taille est d'abord croissante, jusqu'à une certaine valeur de ρ , qui diffère d'une méthode à l'autre, où la relation se met à décroître (voir table 7.8 et figure 7.4).

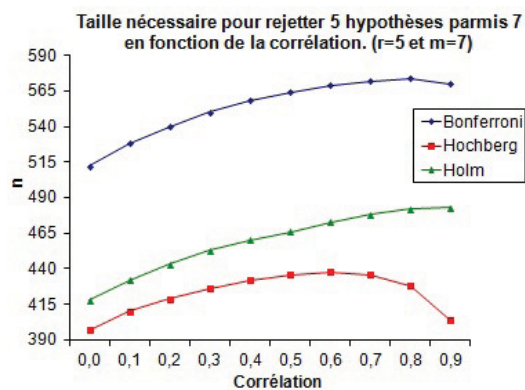


FIG. 7.3. Graphiques de la taille n en fonction de la corrélation ρ pour $m=7$ et $r=5$.

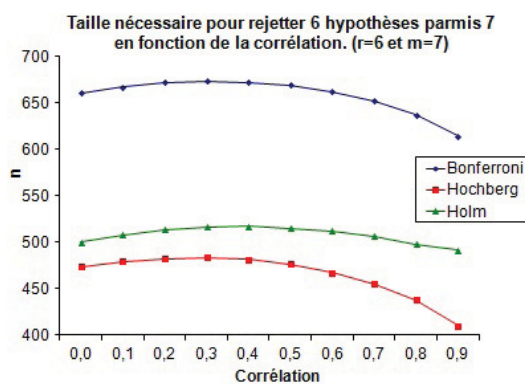


FIG. 7.4. Graphiques de la taille n en fonction de la corrélation ρ pour $m=7$ et $r=6$.

Pour $r = 7$, nous remarquons que cette fois-ci la relation entre la corrélation et la taille est décroissante pour les trois méthodes (voir table 7.8 et figure 7.5).

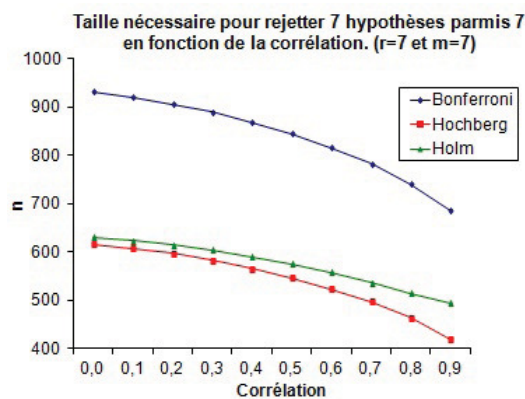


FIG. 7.5. Graphiques de la taille n en fonction de la corrélation ρ pour $m=7$ et $r=7$.

7.3. IMPACT LORSQUE LES VARIANCES ENTRE DEUX VARIABLES SONT TRÈS DIFFÉRENTES L'UNE DE L'AUTRE

Nous nous attardons ici au problème soulevé en section 4.3, afin de valider que l'estimation par une loi de *Student* multivariée de type I est appropriée. Pour ce faire, nous avons calculé à l'aide de notre *package* les tailles n_ρ nécessaires pour atteindre une puissance de 0,80%, selon une valeur de ρ variant de 0 à 0,9. Pour ces tailles n_ρ , nous avons ensuite recalculé les puissances, au moyen de simulations de Monte-Carlo, pour tous les différents scénarios.

Nous présentons ici plus en détails le cas avec $m = 4$ hypothèses pour détecter un effet $\Delta_k = \Delta = 0,2$ au niveau $\alpha_{FWER} = 0,05$. Nous avons d'abord calculé les puissances lorsque toutes les variances sont égales à 1 ($\sigma_k^2 = \sigma^2 = 1$). Puis, nous avons repris les calculs en changeant l'une des quatre variances par 10, puis par 100. Voyons d'abord les résultats lorsque $r = 2$:

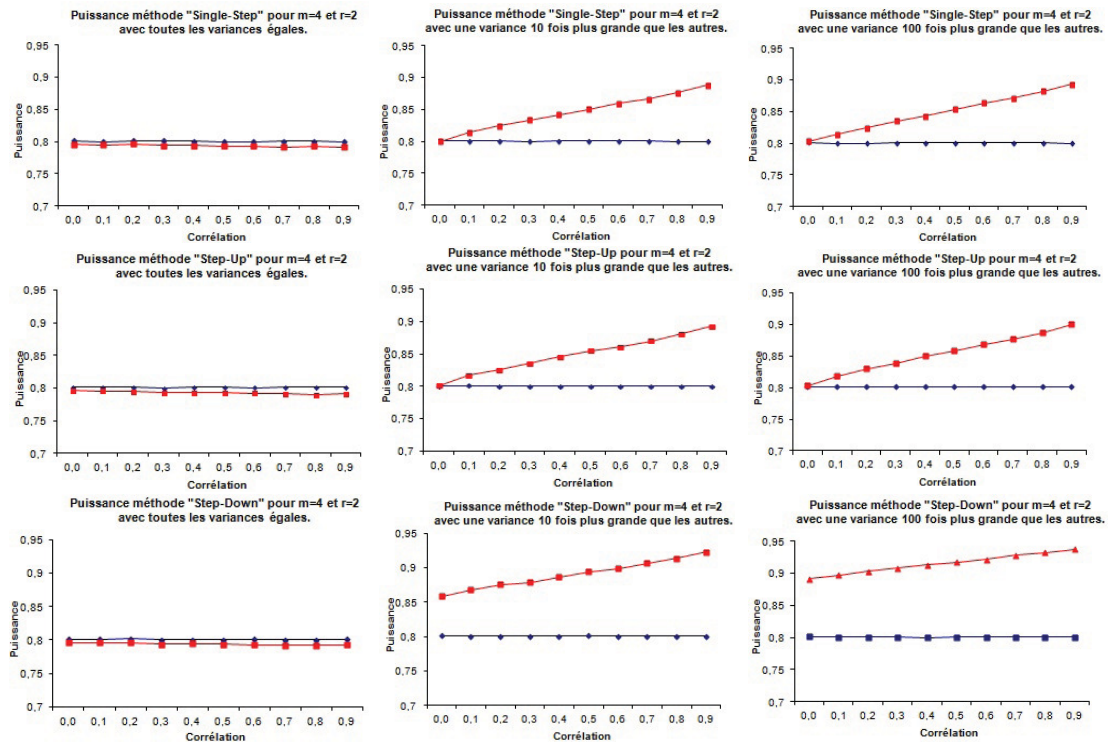


FIG. 7.6. Graphiques de la puissance calculée et de la puissance simulée pour $m=4$ et $r=2$.

La puissance calculée est en **bleu** et la puissance simulée est en **rouge**.

Nous remarquons que pour les trois méthodes, lorsque toutes les variances sont égales l'estimation est très bonne ; la différence entre la variance calculée et celle simulée n'excédant jamais 1%. Par contre, lorsqu'une des variances est 10 fois ou 100 fois plus grande que les autres, la puissance simulée est souvent beaucoup plus grande que celle calculée. Ceci laisse croire que l'estimation par une loi de *Student* multivariée de type I est plutôt conservatrice. Toutefois, nous remarquons que pour les méthodes *Bonferroni* et *Hochberg* lorsque les m variables d'intérêts sont indépendantes, i.e. $\rho = 0$, les deux puissances concordent. Donc dans cette situation, l'estimation est appropriée.

Voyons maintenant les résultats lorsque $r = 3$:

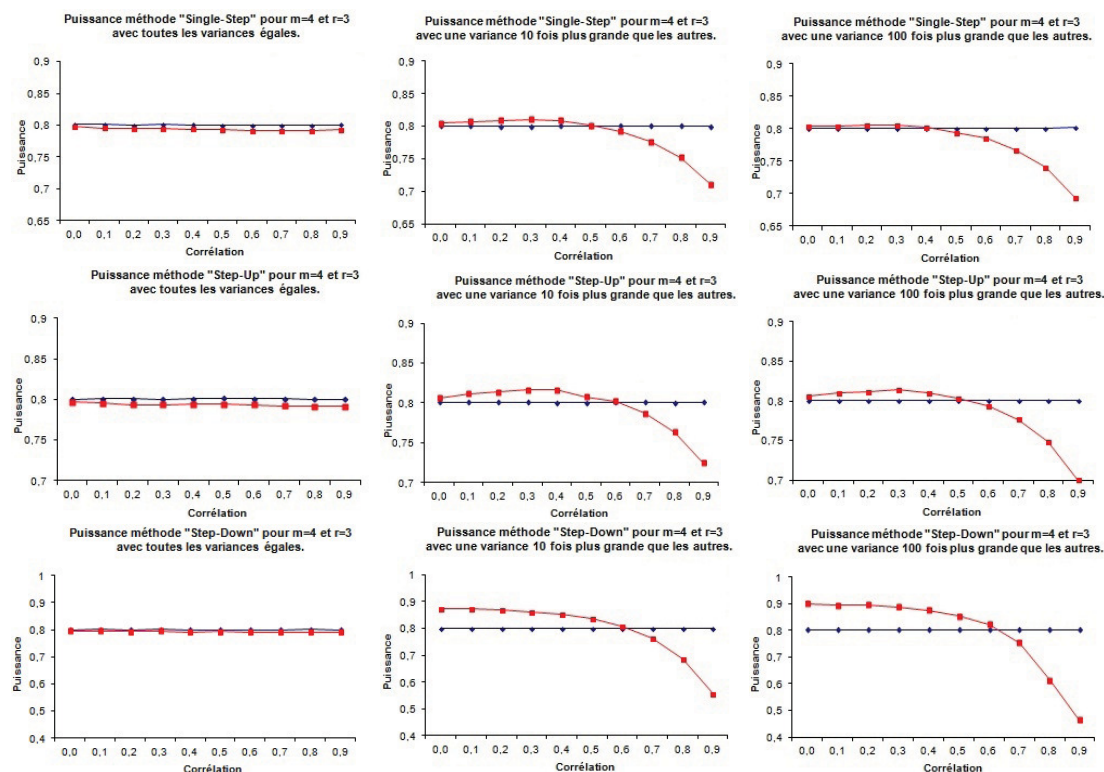


FIG. 7.7. Graphiques de la puissance calculée et de la puissance simulée pour $m=4$ et $r=3$.

La puissance calculée est en **bleu** et la puissance simulée est en **rouge**.

Ici aussi, nous remarquons que pour les trois méthodes, lorsque toutes les variances sont égales l'estimation est très bonne ; la différence entre la variance calculée et celle simulée n'excédant jamais 1%. Nous remarquons aussi que pour les

méthodes *Bonferroni* et *Hochberg* lorsque $\rho = 0$, les deux puissances concordent peu importe que les variances soient égales ou non. Donc dans cette situation, l'estimation est toujours appropriée. Par contre, lorsqu'une des variances est 10 fois ou 100 fois plus grande que les autres, la puissance simulée est parfois plus élevée et parfois beaucoup plus petite. Il devient donc clair que dans certaines situations, l'estimation par une loi de *Student* multivariée de type I ne fonctionne tout simplement pas. Selon nos observations, cela semble se produire lorsque le ratio r/m est supérieur ou égale à $1/2$, lorsque tous les autres paramètres sont constants.

En résumé, nous concluons que lorsque toutes les variances sont égales l'estimation par une loi de *Student* multivariée de type I est appropriée. Par contre, lorsque les variances sont inégales, la fiabilité de la puissance calculée est beaucoup trop dépendante des paramètres, mis à part lorsque $\rho = 0$ pour les méthodes *Bonferroni* et *Hochberg*. Il est donc risqué d'utiliser l'approximation par une loi de *Student* multivariée de type I lorsque les variances sont inégales. Il faudra attendre le développement d'un programme capable de calculer les probabilité d'une loi de *Student* multivariée de type II pour obtenir des résultats plus fiables.

Chapitre 8

EXEMPLE PRATIQUE

Dans leur article, Davison & al. [13] mentionnent que plusieurs études *proof-of-concept* en médecine cardiovasculaire sont planifiées afin de démontrer l'effet de nouvelles interventions sur un seul endpoint et nombreuses sont les études qui utilisent des endpoints de substitution ou composées pour pallier à un manque de puissance lorsqu'il y a plusieurs variables d'intérêts. Toutefois, de tels endpoints peuvent mener à de fausses conclusions ; un résultat faux-négatif mènera à l'abandon de l'intervention sans qu'il y ait d'étude confirmatoire, alors qu'un faux-positif mènera à une étude confirmatoire inutile et une perte de temps et de ressources. Les auteurs proposent donc comme solution de concevoir des études *proof-of-concept* afin d'explorer si la thérapie démontre des indications d'un possible bénéfice clinique sur plusieurs endpoints pertinents à la maladie. Ainsi, si un seul endpoint est significatif, cela peut être dû à la chance, alors que si une majorité d'endpoints sont significatifs, cela laisse croire que le traitement est efficace. Il devient alors intéressant de faire une étude confirmatoire qui n'incluerait que ces endpoints.

Davison & al. [13] se sont servi de l'étude : «Relaxin for the treatment of patients with acute heart failure (Pre-RELAX-AHF) : a multicentre, randomised, placebo-controlled, parallel-group, dose-finding phase IIb study» [71]. Dans cette étude, on teste trois doses d'un nouveau traitement pour venir en aide aux personnes souffrant d'AHF. On y étudie entre autres les variables suivantes : les symptômes, la réapparition de la maladie, la réadmission à l'hôpital, la mort et la qualité de vie. Davison & al. [13] ont démontré dans leur article qu'il y a une

tendance favorable (unilatéral $p < 0, 10$) pour 6 des 9 endpoints cliniques dans le groupe ayant reçu une dose de Relaxin de $30 \mu\text{g}/\text{kg}/\text{jour}$. Malheureusement, nous n'avons trouvé les informations que pour 7 des 9 endpoints utilisés par Davison & al. [13]. Nous proposons donc ici, de calculer la taille requise pour démontrer l'efficacité d'une dose de $30 \mu\text{g}/\text{kg}/\text{jour}$ de Relaxin par rapport à un placebo sur 5 endpoints parmi 7 au niveau $\alpha = 0, 10$, nous poserons donc $\Delta_k = 0$ pour tout k . Certains de ces endpoints sont des proportions, pour les besoin de notre exemple nous les considérerons comme des moyennes dont la variance est de 1.

Voici donc les informations des 7 variables disponibles : [71]

TAB. 8.1. Statistiques de l'exemple (moyennes et écart-types empiriques).

	Placebo	$30 \mu\text{g}/\text{kg}/\text{jour}$
Proportion with moderately or markedly better dyspnoea at 6h, 12h and 24h	23 (1)	40 (1)
Dyspnoea AUC change from baseline to day 5 (VAS[mm×h])	1679 (2556)	2567 (2898)
Worsening heart failure through day 5 (%)	21 (1)	12 (1)
Length of stay (days)	12,0 (7.3)	10,2 (6,1)
Days alive out of hospital	44,2 (14,2)	47,9 (10,1)
KM cardiovascular death or readmission after 60 days (%)	17,2 (1,0)	2,6 (1,0)
KM cardiovascular death after 180 days (%)	14,3 (1,0)	0,0 (1,0)

Le tableau ci-dessous présente les tailles que nous avons obtenues pour les trois méthodes. Nous avons considéré une corrélation nulle ($\rho = 0$) et avons estimé les m variances supposées identiques entre les deux groupes au moyen de la formule classique de la variance empirique *poolée* à partir des écart-types empiriques donnés dans le tableau précédent.

TAB. 8.2. Résultats de l'exemple

Bonf	HC	HM
114	81	84

Chapitre 9

CONCLUSION

Dans ce mémoire, nous nous sommes attardé au calcul de la taille échantillonnale pour les tests à hypothèses multiples, où au moins r des m hypothèses doivent être significatives. Après avoir fait un rappel des termes importants lors d'un calcul de puissance au chapitre 1, nous avons présenté les principales procédures de tests pour des hypothèses multiples dans le chapitre 2. Comme nous nous sommes concentré sur les tests de différence de moyennes, nous avons fait un petit rappel au chapitre 3 du calcul de la taille échantillonnale dans le cas d'une seule hypothèse. Nous avons constaté au chapitre 4 que la problématique peut se décliner en quatre cas distincts : soit toutes les variances sont égales, soit les variances sont les mêmes pour toutes les variables, mais qu'elles diffèrent d'un groupe à l'autre, soit les variances sont les mêmes pour les deux groupes, mais différentes d'une variable à l'autre, soit les variances sont différentes pour les deux groupes et pour toutes les variables. Nous avons vu qu'en théorie dans les deux premiers cas, les statistiques de tests pouvaient être approchées par une loi de *Student* m -multivariée de type I. Alors que pour le cas où les variances sont les mêmes pour les deux groupes, mais différentes d'une variable à l'autre, les statistiques de tests pouvaient être approchées par une loi de *Student* m -multivariée de type II. Pour le dernier cas, où toutes les variances diffèrent les une des autres, la distribution des statistiques de tests reste inconnue à ce jour, ce qui rend le calcul de la puissance impossible. Dans cette situation, le calcul de la taille n'est pas possible puisque l'algorithme de calcul de taille d'échantillon que nous avons développé au chapitre 6 se base sur les équations générales de la

puissance calculées dans le chapitre 5 pour les procédures de *Bonferroni*, *Hochberg* et *Holm*. La programmation de ces équations dans R a nécessité l'utilisation de la formule `pmvt()` qui calcule la fonction de distribution d'une loi de *Student* m -multivariée de type I. Les résultats que nous avons obtenus nous permettent d'affirmer que pour les deux premiers cas du chapitre 4 notre *package* fonctionne très bien. Pour le troisième cas, nous avons soulevé une interrogation à savoir si l'approximation d'une loi de *Student* m -multivariée de type II par une loi de type I était appropriée lorsqu'il existe une grande différence entre les variances. Les résultats de nos simulations obtenus au chapitre 7 et ceux de l'exemple pratique du chapitre 8 nous démontrent bien que l'approximation manque de fiabilité. En effet, la puissance calculée dans cette situation est beaucoup trop dépendante des paramètres. Il serait donc intéressant pour l'avenir de développer un programme pouvant calculer la fonction de distribution d'une loi de *Student* m -multivariée de type II.

BIBLIOGRAPHIE

- [1] C.J. ADCOCK : Sample size determination : A review. *The Statistician*, 46:261–283, 1997.
- [2] P. ARMITAGE et T. COLTON : *Multivariate t Distribution in Encyclopedia of Biostatistics*, volume 2. Wiley, second édition, 2005.
- [3] P. BACCHETTI : Current sample size conventions : Flaws, harms, and alternatives. *BMC Medicine*, 8(17), 2010.
- [4] P. BACCHETTI, C.E. MCCULLOCH et M.R. SEGAL : Simple, defensible sample sizes based on cost efficiency. *Biometrics*, 64(2):577–585, 669, 2008.
- [5] P. BAI : Sphericity test in a GMANOVA-MANOVA model with normal error. *J. Multivariate Anal.*, 100(10):2305–2312, 2009.
- [6] Y. BENJAMINI et Y. HOCHBERG : Controlling the false discovery rate : a practical and powerful approach to multiple testing. *J. Roy. Statist. Soc. Ser. B*, 57(1):289–300, 1995.
- [7] M. BILODEAU et D. BRENNER : *Theory of multivariate statistics*. Springer Texts in Statistics. Springer-Verlag, New York, 1999.
- [8] J. CHEN, J. LUO, K. LIU et D.V. MEHROTRA : On power and sample size computation for multiple testing procedures. *Computational Statistics and Data Analysis*, 55:110–122, 2011.
- [9] S.-C. CHOW, J. SHAO et H. WANG : *Sample size calculations in clinical research*, volume 20 de *Chapman & Hall/CRC Biostatistics Series*. Chapman & Hall/CRC, Boca Raton, FL, second édition, 2008.
- [10] R. COE : It's the effect size, stupid what effect size is and why it is important. *Paper presented at the Annual Conference of the British Educational Research Association, University of Exeter, England, 12-14 September 2002*, 2002.

- [11] J. COHEN : *Statistical power analysis for the behavioral sciences*. Lawrence Erlbaum Associates, Hillsdale, New Jersey, USA., second édition, 1988.
- [12] CPMP/EWP/908/99 : *Points to consider on multiplicity issues in clinical trials (adopted September 2002)*. Available at : www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500003640.pdf.
- [13] B.A. DAVISON, G. COTTER, H. SUN, L. CHEN, J.R. TEERLINK, Metra M., G.M. FELKER, A.A. VOORS, P. PONIKOWSKI, G. FILIPPATOS, B. GREENBERG, S.L. TEICHMAN, E. UNEMORI et G.G. KOCH : Permutation criteria to evaluate multiple clinical endpoints in a proof-of-concept study : lessons from pre-relax-ahf. *Clin Res Cardiol*, 100:745–753, 2011.
- [14] A. DMITRIENKO, A.C. TAMHANE et F. BRETZ : *Multiple Testing Problems in Pharmaceutical Statistics*. Chapman & Hall/CRC Biostatistics Series. Chapman & Hall/CRC, New York, 2010.
- [15] S. DUDOIT et M.J. van der LAAN : *Multiple testing procedures with applications to genomics*. Springer Series in Statistics. Springer, New York, 2008.
- [16] D.B. DUNCAN : Multiple range and multiple F tests. *Biometrics*, 11:1–42, 1955.
- [17] O.J. DUNN : Estimation of the means of dependent variables. *Ann. Math. Statist.*, 29:1095–1111, 1958.
- [18] O.J. DUNN : Multiple comparisons among means. *J. Amer. Statist. Assoc.*, 56:52–64, 1961.
- [19] C.W. DUNNETT : A multiple comparison procedure for comparing several treatments with a control. *American Statistical Association*, 50(272):1096–1121, 1955.
- [20] C.W. DUNNETT et A.C. TAMHANE : A step-up multiple test procedure. *J. Amer. Statist. Assoc.*, 87(417):162–170, 1992.
- [21] C.W. DUNNETT et A.C. TAMHANE : Power comparisons of some step-up multiple test procedures. *Statistics & Probability Letters*, 16:55–58, 1993.
- [22] ICH E9 : *Statistical Principles for Clinical Trials. Step 5 Note for guidance on statistical principles for clinical trials*, 2000. Available at : www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2009/09/WC500002928.pdf.
- [23] A.H. FEIVESON et F.C. DELANEY : The distribution and properties of a weighted sum of chi squares. *NASA Technical Note*, TN D-4575, 1968.

- [24] F.A. GRAYBILL : *Matrices with applications in statistics*. Wadsworth Statistics/Probability Series. Wadsworth Advanced Books and Software, Belmont, Calif., second édition, 1983.
- [25] J.-M. GROUIN et M. COSTE : *Essais de non-infériorité et d'équivalence : méthodologie et analyse* dans Bouvenot, Gilles et Vray, Muriel eds. *Essais Cliniques; Théorie, Pratique et Critique*. Flammarion Médecine-Sciences, Paris, 2006. 4e Édition.
- [26] J.A. HANLEY et S.H. SHAPIRO : Discussions [simple, defensible sample sizes based on cost efficiency]. *Biometrics*, 64(2):586–587, 2008.
- [27] M. HASLER et L.A. HORTON : A dunnett-type procedure for multiple endpoints. *The International Journal of Biostatistics*, 7 : Iss.1, Article 3, 2011. Available at : <http://www.bepress.com/ijb/vol7/iss1/3>.
- [28] Y. HOCHBERG : A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, 75(4):800–802, 1988.
- [29] Y. HOCHBERG et A.C. TAMHANE : *Multiple comparison procedures*. Wiley Series in Probability and Mathematical Statistics : Applied Probability and Statistics. John Wiley & Sons Inc., New York, 1987.
- [30] S. HOLM : A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.
- [31] G. HOMMEL : A stagewise rejective multiple test procedure based on a modified bonferroni test. *Biometrika*, 75:383–386, 1988.
- [32] M. HORN et R. VOLLANDT : Sample sizes for comparisons of k treatments with a control based on different definitions of the power. *Biom. J.*, 40(5):589–612, 1998.
- [33] N.L. JOHNSON et S. KOTZ : *Distributions in statistics : continuous multivariate distributions*. John Wiley & Sons Inc., New York, 1972. Wiley Series in Probability and Mathematical Statistics.
- [34] A. KIRBY, V. GEBSKI et A.C. KEECH : Determining the sample size in a clinical trial. *MJA*, 5(177):256–257, 2002.
- [35] R.E. KIRK : *Experimental design : Procedures for the behavioral sciences*. Brooks/Cole Publishing Co. Wadsworth Publishing Co., Inc., Belmont, Calif., 1982.

- [36] L. KONG, R.C. KOHBERGER et G.G. KOCH : Type i error and power in noninferiority / equivalence trials with correlated multiple endpoints : An example from vaccine development trials. *Journal of Biopharmaceutical statistics*, 14(4):893–907, 2004.
- [37] S.J. LEE et M. ZELEN : Clinical trials and sample size considerations : Another perspective. *Statistical Science*, 15(2):95–103, 2000.
- [38] S.J. LEE et M. ZELEN : Clinical trials and sample size considerations : Another perspective [rejoinder]. *Statistical Science*, 15(2):108–110, 2000.
- [39] E.L. LEHMANN et J.P. ROMANO : *Testing Statistical Hypotheses*. Springer, third édition, 2005.
- [40] R.V. LENTH : Some practical guidelines for effective sample size determination. *Amer. Statist.*, 55(3):187–193, 2001.
- [41] Y.-F. LI : The development of power estimates and sample size requirements for seven multiple procedures. *In Dissertation submitted to the Faculty or the Graduate School of the University of Maryland*. Unpublished, 1997.
- [42] D.V. LINDLEY : The choice of sample size. *J. Roy. Statist. Soc. Ser. B*, 46(2):129–138, 1997.
- [43] R. MARCUS, E. PERITZ et K.R. GABRIEL : On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*, 63:655–660, 1976.
- [44] W. MAURER et B.H. MARGOLIN : The multivariate inclusion-exclusion formula and order statistics from dependent variates. *Ann. Statist.*, 4(6):1190–1199, 1976.
- [45] J.L. MENDOZA : A significance test for multisample sphericity. *Psychometrika*, 45(4):495–498, 1980.
- [46] R.G. JR. MILLER : *Simultaneous Statistical Inference*. McGraw Hill, New York, 1966.
- [47] R.J. MUIRHEAD : *Aspects of multivariate statistical theory*. John Wiley & Sons Inc., New York, 1982. Wiley Series in Probability and Mathematical Statistics.
- [48] D.K. NAGAR et M.E. CASTAÑEDA : Distribution and percentage points of LRC for testing multisample compound symmetry in the bivariate and the trivariate cases. *Metron*, 64(2):217–238, 2006.
- [49] D.K. NAGAR et C.C. SANCHEZ : Exact percentage points for testing multisample sphericity in the bivariate case. *Communications in Statistics*, 33(2):447–457, 2004.

- [50] U.D. NAIK : Some selection rules for comparing p processes with a standard. *Communications in Statistics*, 4(6):519–535, 1975.
- [51] P.C. O'BRIEN : The appropriateness of analysis of variance and multiple comparison procedures. *Biometrics*, 39:787–794, 1983.
- [52] W. OFFEN, C. CHUANG-STEIN, A. DMITRIENKO, G. LITTMAN, J. MACA, L. MEYERSON, R. MUIRHEAD, P. STRYSZAK, A. BODDY, K. CHEN, K. COPLEY-MERRIMAN, W. DERE, S. GIVENS, D. HALL, D. HENRY, J.D. JACKSON, A. KRISHEN, T. LIU, S. RYDER, A.J. SANKOH, J. WANG et C.-H. YEH : Multiple co-primary endpoints : Medical and statistical solutions. a report from the multiple endpoints expert team of the pharmaceutical research and manufacturers of america. *Drug Informal Journal*, 41:31–46, 2007.
- [53] H. PEZESHK : Bayesian techniques for sample size determination in clinical trials : a short review. *Stat. Methods Med. Res.*, 12(6):489–504, 2003.
- [54] M.A. PROSCHAN et M.A. WACLAWIW : Practical guidelines for multiplicity adjustment in clinical trials. *Controlled Clinical Trials*, 21:527–539, 2000.
- [55] J.A. RAFTER, M.L. ABELL et J.P. BRASELTON : Multiple comparison methods for means. *SIAM Rev.*, 44(2):259–278 (electronic), 2002.
- [56] P.H. RAMSEY : Power differences between pairwise multiple comparisons. *J. Amer. Statist. Assoc.*, 73:479–485, 1978.
- [57] C.R. RAO : *Linear statistical inference and its applications*. John Wiley & Sons, New York-London-Sydney, second édition, 1973. Wiley Series in Probability and Mathematical Statistics.
- [58] J.A. RICE : *Mathematical Statistics and Data Analysis*. Duxbury Press, Belmont, California, 1995.
- [59] D.M. ROM : A sequentially rejective test procedure based on a modified Bonferroni inequality. *Biometrika*, 77(3):663–665, 1990.
- [60] S.M. ROSS : *Initiation aux probabilités*. Presses Polytechniques et Universitaires Romandes, Lausanne, 1996. With a preface by Peter Nüesch, Translated from the 4th (1994) English edition by Christian Hofer and Frédéric Dorsaz.
- [61] K. ROTHMAN : No adjustments are needed for multiple comparisons. *Epidemiology*, 1:43–46, 1990.

- [62] A.J. SANKOH, R.B. D'AGOSTINO et M.F. HUQUE : Efficacy endpoint selection and multiplicity adjustment methods in clinical trials with inherent multiple endpoint issues. *Statistics in medicine*, 22:3133–3150, 2003.
- [63] A.J. SANKOH, M.F. HUQUE et S.F. DUBEY : Some comments on frequently used multiple endpoint adjustment methods in clinical trials. *Statistics in medicine*, 16:2529–2542, 1997.
- [64] F.E. SATTERTHWAITTE : An approximate distribution of estimates of variance components. *Biometrics Bulletin*, 2:110–114, 1946.
- [65] J.P. SHAFFER : Modified sequentially rejective multiple test procedures. *Journal of the American Statistical Association*, 81(395):826–831, 1986.
- [66] Z. ŠIDÁK : Rectangular confidence regions for the means of multivariate normal distributions. *J. Amer. Statist. Assoc.*, 62:626–633, 1967.
- [67] Z. ŠIDÁK : On probabilities of rectangles in multivariate Student distributions : their dependence on correlations. *Ann. Math. Statist.*, 42:169–175, 1971.
- [68] R.J. SIMES : An improved Bonferroni procedure for multiple tests of significance. *Biometrika*, 73(3):751–754, 1986.
- [69] A.C. TAMHANE et C.W. DUNNETT : Stepwise multiple test procedures with biometric applications. *Journal of Statistical Planning and Inference*, 82:55–68, 1997.
- [70] A.C. TAMHANE, W. LIU et C.W. DUNNETT : A generalized step-up-down multiple test procedure. *Canad. J. Statist.*, 26(2):353–363, 1998.
- [71] J.R. TEERLINK, M. METRA, G.M. FELKER, P. PONIKOWSKI, A.A. VOORS, B.D. WEATHERLEY, A. MARMOR, A. KATZ, J. GRZYBOWSKI, E. UNEMORI, S.L. TEICHMAN et G. COTTER : Relaxin for the treatment of patients with acute heart failure (pre-relax-ahf) : a multicentre, randomised, placebo-controlled, parallel-group, dose-finding phase iib study. *Lancet*, 373:1429–39, 2009. Published Online.
- [72] Y.L. TONG : *The multivariate normal distribution*. Springer Series in Statistics. Springer-Verlag, New York, 1990.
- [73] L.E. TOOTHAKER : *Multiple Comparisons for Researchers*. SAGE Publications, sage, newbury park, ca, édition, 1991.
- [74] B.L. WELCH : The generalization of "student's" problem when several different population variances are involved. *Biometrika*, 34:28–35, 1947.

- [75] P.H. WESTFALL et S.S. YOUNG : *Resampling-Based Multiple Testing : Examples and Methods for p-Value Adjustment*. Wiley, New York, 1993.
- [76] P. ZHANG : A simple formula for sample size calculation in equivalence studies. *Journal of Biopharmaceutical Statistics*, 13(3):529–538, 2003.

Annexe A

FORMULES DE COMBINATOIRE ET DE PROBABILITÉ

Nous présentons ici quelques formules utilisées dans nos démonstrations.

$$\binom{n}{p} = \frac{n!}{(n-p)!p!}, \quad \text{si } 0 \leq p \leq n. \quad (\text{A.0.1})$$

$$\binom{n}{p} = 0, \quad \text{si } p > n. \quad (\text{A.0.2})$$

$$\binom{m+1}{j+1} = \binom{m}{j} + \binom{m}{j+1}. \quad (\text{A.0.3})$$

$$\binom{m}{j} = \binom{m}{m-j}. \quad (\text{A.0.4})$$

$$\binom{-n}{p} = \binom{n+p-1}{p}(-1)^p = \binom{n+p-1}{n-1}(-1)^p \quad \text{Binôme négatif.} \quad (\text{A.0.5})$$

$$\mathbf{P}(A) = 1 - \mathbf{P}(A^C). \quad (\text{A.0.6})$$

$$\mathbf{P}(A \cap B) = \mathbf{P}(A) - \mathbf{P}(A \cap B^C). \quad (\text{A.0.7})$$

$$\begin{aligned}
\mathbf{P}[A \cap (B \cap C)^c] &= \mathbf{P}(A) - \mathbf{P}[A \cap (B \cup C)] \\
&= \mathbf{P}(A) - \mathbf{P}[(A \cap B) \cup (A \cap C)] \\
&= \mathbf{P}(A) - [\mathbf{P}(A \cap B) + \mathbf{P}(A \cap C) \\
&\quad - \mathbf{P}(A \cap B \cap C)]. \tag{A.0.8}
\end{aligned}$$

Théorème A.0.1 (Ross[60], théorème 2.7, p.34). *La formule de Poincaré s'écrit :*

$$\begin{aligned}
\mathbf{P}(E_1 \cup E_2 \cup \dots \cup E_n) &= \sum_{i=1}^n \mathbf{P}(E_i) - \sum_{i_1 < i_2} \mathbf{P}(E_{i_1} E_{i_2}) + \dots \\
&\quad + (-1)^{r+1} \sum_{i_1 < i_2 < \dots < i_r} \mathbf{P}(E_{i_1} E_{i_2} \dots E_{i_r}) \\
&\quad + \dots + (-1)^{n+1} \mathbf{P}(E_1 E_2 \dots E_n) \\
&= \sum_{r=1}^n (-1)^{r+1} \sum_{(i_1 < i_2 < \dots < i_r) \in F_{r,n}} \mathbf{P}(\cap_{j=1}^r E_{i_j}) \tag{A.0.9}
\end{aligned}$$

où $F_{r,n} = \left\{ (i_1 < i_2 < \dots < i_r); i_k \in \{1, \dots, n\}, \forall k = 1, \dots, r \right\}$ est l'ensemble de toutes les combinaisons de r entiers pris parmi les n premiers entiers.

Annexe B

FORMULES SUR LA PROBABILITÉ CONJOINTE DE K STATISTIQUES D'ORDRE

B.1

Nous présentons ici certains résultats tirés de Maurer et Margolin [44].

Considérons les ensembles d'événements $\mathcal{A}_i = \{A_{i,j}\}_{j=1}^n$, $i = 1, 2, \dots, K$, tels que pour tout j , $A_{1,j} \subseteq A_{2,j} \subseteq \dots \subseteq A_{K,j}$. Soit $\underline{T} = (t_1, t_2, \dots, t_K)^\top$ un vecteur de K entiers positifs et notons $\mathbf{P}(\underline{T} : n)$ la probabilité qu'*au moins* t_i événements de \mathcal{A}_i se produisent, $i = 1, 2, \dots, K$.

Pour tout vecteur $\underline{x} = (x_1, x_2, \dots, x_K)^\top$, on pose $x_0 = 0$ et $x_{K+1} = n$, à moins d'avis contraire. Ensuite, on définit $\underline{x}^* = (x_2, x_3, \dots, x_{K+1})^\top$; $\Delta x_i = x_i - x_{i-1}$, $i=1, 2, \dots, K+1$; $\Delta \underline{x} = (\Delta x_1, \Delta x_2, \dots, \Delta x_K)^\top$; $\Delta^* \underline{x} = (\Delta x_1, \Delta x_2, \dots, \Delta x_{K+1})^\top$ et $x_+ = \sum_{i=1}^K x_i$. Un vecteur de longueur K dont tous les éléments sont des uns sera noté $\underline{1}^\top$. Pour deux vecteurs de même longueur \underline{x} et \underline{y} , l'interprétation de $\underline{x} \leq \underline{y}$ sera une inégalité *composante par composante* faible. Finalement, $\sum_{\underline{x}}$ dénote la sommation sur tous les vecteurs $\underline{x} = (x_1, x_2, \dots, x_K)^\top$ qui sont ordonnés, i.e. $0 \leq x_1 \leq x_2 \leq \dots \leq x_K \leq n$ et $\sum_{\underline{x}=\underline{s}}$ dénote la sommation sur tous les vecteurs \underline{x} qui sont ordonnés comme ci-dessus, mais qui en plus satisfont $\underline{s} \leq \underline{x} \leq \underline{r}$ pour \underline{r} et \underline{s} de longueur K .

B-ii

Théorème B.1.1 (Maurer et Margolin [44], théorème 2.1, p.1191).

Pour un vecteur d'entiers $\underline{a} = (a_1, a_2, \dots, a_K)^\top$ tel que $0 = a_0 \leq a_1 \leq a_2 \dots \leq a_K \leq a_{K+1} = n$ on définit :

$$P_{\underline{a}} = \sum' \mathbf{P} \left[\bigcap_{i=0}^{K-1} \left(\bigcap_{m=a_i+1}^{a_{i+1}} A_{i+1, j_m} \right) \right] \quad (\text{B.1.1})$$

où \sum' dénote la sommation sur tous les

$$\binom{n}{\Delta^* \underline{a}}$$

ensembles d'indices $(j_1, j_2, \dots, j_{a_K})$ tels que $j_r < j_{r+1}$ pour $r \in \{a_i + 1, a_i + 2, \dots, a_{i+1} - 1\}$, $i = 0, 1, 2, \dots, K - 1$ et $j_r \neq j_s$ quand $1 \leq r < s \leq a_K$ et où $\binom{n}{\Delta^* \underline{a}}$ est un coefficient multinomial.

Alors, on a :

$$\mathbf{P}(\underline{T} : n) = (-1)^{t_+} \sum_{\underline{a}=\underline{t}}^{\underline{a}^*} (-1)^{a_+} \prod_{i=1}^K \binom{(\Delta a_i) - 1}{a_i - t_i} P_{\underline{a}} \quad (\text{B.1.2})$$

où $t_+ = \sum_i^K t_i$, $\Delta a_i = a_i - a_{i-1}$ et $\underline{a}^* = (a_2, a_3, \dots, a_{K+1})^\top$.

Dans le cas échangeable (voir (2.8) p.1193), on a :

$$P_{\underline{a}} = \binom{n}{\Delta^* \underline{a}} \mathbf{P} \left[\bigcap_{i=0}^{K-1} \left(\bigcap_{m=a_i+1}^{a_{i+1}} A_{i+1, m} \right) \right]. \quad (\text{B.1.3})$$

B.2

Dans le cadre de ce mémoire, nous utiliserons les ensembles d'événements $\mathcal{A}_h = \{A_{h,k}\}_{k=1}^m$, $h = 1, 2, \dots, q$ avec $A_{h,k} = \{T_k < c_{i_h}\}$.

Voyons ceci sur un exemple avec $m = 3$, $q = 2$ et $i_1 = 1$, $i_2 = 3$. On obtient donc les ensembles $\mathcal{A}_1 = \left\{ \{T_1 < c_1\}, \{T_2 < c_1\}, \{T_3 < c_1\} \right\}$ et $\mathcal{A}_2 = \left\{ \{T_1 < c_3\}, \{T_2 < c_3\}, \{T_3 < c_3\} \right\}$.

Pour un vecteur d'entiers $\underline{a} = (a_1, a_2, \dots, a_q)^\top$ tel que $0 = a_0 \leq a_1 \leq a_2 \dots \leq a_q \leq a_{q+1} = m$ la formule (B.1.1) devient :

$$P_{\underline{a}} = \sum_l' \mathbf{P} \left[\bigcap_{h=0}^{q-1} \left(\bigcap_{k=a_h+1}^{a_{h+1}} T_{l_k} < c_{t_{h+1}} \right) \right]. \quad (\text{B.2.1})$$

Dans le cas échangeable, la formule (B.1.3) devient :

$$P_{\underline{a}} = \binom{m}{\Delta^* \underline{a}} \mathbf{P} \left[\bigcap_{h=0}^{q-1} \left(\bigcap_{k=a_h+1}^{a_{h+1}} T_k < c_{t_{h+1}} \right) \right]. \quad (\text{B.2.2})$$

La formule (B.1.2) s'écrit alors

$$\mathbf{P}(T : n) = \mathbf{P} \left[\bigcap_{u=1}^q (T_{t_u:m} < c_{t_u}) \right] = (-1)^{t_+} \sum_{\underline{a}=\underline{t}}^{\underline{a}^*} (-1)^{a_+} \prod_{h=1}^q \binom{(\Delta a_h) - 1}{a_h - t_h} P_{\underline{a}} \quad (\text{B.2.3})$$

avec $\underline{t} = (t_1, t_2, \dots, t_q)^\top$ et $t_+ = \sum_{h=1}^q t_h$

Il est utile de noter que

$$\sum_{\underline{a}=\underline{t}}^{\underline{a}^*} f(\underline{a}) = \begin{cases} \sum_{a_1=t_1}^m f(a_1) & \text{pour } \underline{t} = (t_1) \\ \sum_{a_2=t_2}^m \sum_{a_1=t_1}^{a_2} f(a_1, a_2) & \text{pour } \underline{t} = (t_1, t_2)^\top \\ \sum_{a_3=t_3}^m \sum_{a_2=t_2}^{a_3} \sum_{a_1=t_1}^{a_2} f(a_1, a_2, a_3) & \text{pour } \underline{t} = (t_1, t_2, t_3)^\top \\ \text{etc ...} \end{cases}$$

B.3

Explicitons la définition de \sum . Il s'agit de la sommation sur tous les

$$\binom{m}{\Delta^* \underline{a}}$$

ensembles d'indices $(l_1, l_2, \dots, l_{a_q})$ tels que $l_r < l_{r+1}$ pour $r \in \{a_h + 1, a_h + 2, \dots, a_{h+1} - 1\}$, $h = 0, 1, 2, \dots, q - 1$ et $l_r \neq l_s$ quand $1 \leq r < s \leq a_q$.

En d'autres termes, on a :

$$\sum'_{\underline{l}} f(\underline{l}) = \sum_{\underline{l}=l_1}^{l_p} f(\underline{l}), \quad \text{avec } p = \binom{m}{\Delta^* \underline{a}} \in \mathbb{N}$$

où l'ensemble $\{l_1, \dots, l_p\}$ de cardinal $\binom{m}{\Delta^* \underline{a}}$ est constitué des vecteurs d'indices de la forme $(l_1, \dots, l_{a_q})^\top$.

Voyons ceci sur un exemple. Posons $m = 7$ et $\underline{a} = (1, 4, 6)^\top$. Ici $q = 3$, $a_q = 6$ et $a_0 = 0$. Alors $\Delta^* \underline{a} = (1 - 0, 4 - 1, 6 - 4)^\top = (1, 3, 2)^\top$. Ainsi, il y aura $\binom{m}{\Delta^* \underline{a}} = \binom{7}{1,3,2} = 420$ ensembles d'indices (k_1, k_2, \dots, k_6) .

B-iv

Voyons comment construire ces ensembles. Il faut d'abord trouver les $r \in (a_h + 1, a_h + 2, \dots, a_{h+1} - 1)$, $h = 0, 1, \dots, 3 - 1$ pour trouver les conditions $k_r < k_{r+1}$.

$$r_{h=0} = (a_0 + 1, \dots, a_1 - 1) = (0 + 1, \dots, 1 - 1) = (1, \dots, 0)$$

$$\Rightarrow r_{h=0} \in \{\} \Rightarrow \text{donc aucune condition sur } k_1,$$

$$r_{h=1} = (a_1 + 1, \dots, a_2 - 1) = (1 + 1, \dots, 4 - 1) = (2, \dots, 3)$$

$$\Rightarrow r_{h=1} \in \{2, 3\} \Rightarrow k_2 < k_3 < k_4,$$

$$r_{h=2} = (a_2 + 1, \dots, a_3 - 1) = (4 + 1, \dots, 6 - 1) = (5, \dots, 5)$$

$$\Rightarrow r_{h=2} \in \{5\} \Rightarrow k_5 < k_6.$$

Construisons maintenant les $q = 3$ blocs pour chaque r_h incluant toutes les valeurs possibles de k_r qui respectent les conditions ci-haut.

Bloc 1	k_1	1	2	3	4	5	6	7	7 possibilités
	k_2	1	1	...	2	...	4	5	
Bloc 2	k_3	2	2	...	3	...	6	6	35 possibilités
	k_4	3	4	...	4	...	7	7	
Bloc 3	k_5	1	1	...	2	...	5	5	
	k_6	2	3	...	3	...	6	7	21 possibilités

Il suffit maintenant d'agencer une colonne de chaque bloc en respectant la condition $k_r \neq k_s$ pour $\forall r, s \in (1, \dots, a_q)$. Dans l'exemple ci-haut, les colonnes grises forment l'un des 420 ensembles d'indices possibles ; soit $(1, 4, 6, 7, 2, 3)$.

Annexe C

LOI DE *STUDENT* MULTIVARIÉE

Il existe dans la littérature deux types de loi de *Student* multivariée que l'on nomme type I et type II [2], [33] et [72].

C.1. LOI DE *Student* MULTIVARIÉE DE TYPE I

Soit un vecteur $\underline{X} = (X_1, \dots, X_m)^\top$ qui suit une loi normale multivariée, $\mathcal{N}_m(\underline{\mu}, \sigma^2 \underline{Z})$ et soit $\nu S^2 / \sigma^2 \sim \chi_\nu^2$ indépendante de \underline{X} . En pratique, S^2 est l'erreur quadratique moyenne provenant d'une analyse de la variance avec ν degrés de liberté. Alors, le vecteur \underline{T} dont les éléments sont $T_i = X_i/S$, ($1 \leq i \leq m$) suit une loi de *Student* multivariée de type I, que l'on notera $t_m^I(\nu, \underline{\mu}, \underline{Z})$.

La fonction de densité prend la forme (notez qu'il y a une erreur dans [2] à ce sujet) :

$$f(\underline{t}) = C(m, \nu) \left[1 + \frac{(\underline{t} - \underline{\mu})^\top \underline{Z}^{-1} (\underline{t} - \underline{\mu})}{\nu} \right]^{-(\nu+m)/2} \quad (\text{C.1.1})$$

où $C(m, \nu) = \frac{\Gamma[(\nu+m)/2]}{(\pi\nu)^{m/2} \Gamma[\nu/2] |\underline{Z}|^{1/2}}$. Ici, $\underline{\mu}$ est le centre de symétrie, \underline{Z} est une matrice de paramètres d'échelle et ν est le nombre de degrés de liberté. Cette distribution est dite *centrée* lorsque $\underline{\mu} = \underline{0}$ et *décentrée* autrement. Les deux premiers moments de cette distribution sont $\mathbb{E}(\underline{T}) = \underline{\mu}$ pour $\nu > 1$ et $\mathbb{V}\text{ar}(\underline{T}) = \frac{\nu}{\nu-2} \underline{Z}$, pour $\nu > 2$. Pour \underline{Z} arbitraire, les marginales de la distribution sont des distributions t d'échelle. Toutefois, si l'on remplace \underline{Z} par une matrice de corrélation définie positive $\underline{R} = [\rho_{ij}]$ dans (C.1.1), alors toutes les marginales suivent une distribution de *Student* univariée.

C-ii

C.2. LOI DE *Student* MULTIVARIÉE DE TYPE II

Soit un vecteur $\underline{X} = (X_1, \dots, X_m)^\top$ qui suit une loi normale multivariée, $\mathcal{N}_m(\underline{\mu}, \underline{\Sigma})$ et soit $\nu \underline{S} = \nu[S_{ij}]$ une matrice aléatoire indépendante de \underline{X} et d'ordre $m \times m$ qui suit une distribution de Wishart $W_m(\nu, \underline{\Xi})$, où $\underline{\Xi}$ est une matrice de paramètres d'échelle dont les éléments de la diagonale sont les mêmes que ceux de $\underline{\Sigma}$. Alors, le vecteur \underline{T} dont les éléments sont $T_i = X_i/S_i$, ($1 \leq i \leq m$), où $S_i^2 = S_{ii}$, ($1 \leq i \leq m$) sont les éléments sur la diagonale de S , suit une loi de *Student* multivariée de type II que l'on notera $t_m^{II}(\nu, \underline{\mu}, \underline{\Sigma}, \underline{\Xi})$.

Nous pouvons noter que $\frac{\nu S_{ii}}{\sigma_i^2} \sim \chi_\nu^2$. En effet, par le théorème 3.2.5 de Muirhead [47], nous avons :

$$\frac{\nu S_{ii}}{\sigma_i^2} = \frac{\underline{e}_i \underline{S} \underline{e}_i^\top}{\sigma_i^2} \sim \frac{1}{\sigma_i^2} W_1(\nu, \underline{e}_i \underline{\Xi} \underline{e}_i^\top),$$

où \underline{e}_i est le vecteur de longueur m avec 1 en $i^{\text{ème}}$ position et 0 partout ailleurs. Or, $\underline{e}_i \underline{\Xi} \underline{e}_i^\top = \underline{\Xi}_{ii} = \underline{\Sigma}_{ii} = \sigma_i^2$. Donc, $\frac{\nu S_{ii}}{\sigma_i^2} \sim \frac{1}{\sigma_i^2} W_1(\nu, \sigma_i^2) = \chi_\nu^2$, en utilisant le théorème 8.3.4 de Muirhead [47].

Aucune expression de la fonction de densité de \underline{T} n'existe à ce jour. Toutefois, il existe des propriétés d'inégalité des probabilités pour certaines structures de dépendance entre \underline{X} et \underline{S} (voir Sidák [67]).

Annexe D

CALCUL DES ESTIMATEURS POUR μ ET σ^2 , LOIS DES STATISTIQUES

D.1. LORSQUE $\text{VAR}(X_{ijk}) = \sigma^2$

Nous faisons l'hypothèse (*multisample compound symmetry*) que $\underline{\underline{\Sigma}}_1 = \underline{\underline{\Sigma}}_2 \equiv \underline{\underline{\Sigma}} = \sigma^2[(1 - \rho)\mathcal{I}_m + \rho J] \equiv \sigma^2 K_\rho$,

$$\text{avec } J = \begin{pmatrix} 1 & \dots & 1 \\ \vdots & 1 & \vdots \\ 1 & \dots & 1 \end{pmatrix}.$$

Il faut considérer le problème comme étant celui d'une MANOVA à mesures répétées, [7] et [5] :

$$\underline{\underline{Y}} = \begin{pmatrix} \underline{X}_{11}^\top \\ \vdots \\ \underline{X}_{1n_1}^\top \\ \underline{X}_{21}^\top \\ \vdots \\ \underline{X}_{2n_2}^\top \end{pmatrix} \sim \mathcal{N}_{(n_1+n_2) \times m} \left(\begin{pmatrix} \underline{\mu}_1^\top \\ \vdots \\ \underline{\mu}_1^\top \\ \underline{\mu}_2^\top \\ \vdots \\ \underline{\mu}_2^\top \end{pmatrix}, \underbrace{\begin{pmatrix} \underline{I}_{n_1} \otimes \underline{\Sigma}_1 & \underline{0} \\ \underline{0} & \underline{I}_{n_2} \otimes \underline{\Sigma}_2 \end{pmatrix}}_{\underline{\underline{\Omega}}} \right),$$

avec $\underline{X}_{ij} \sim \mathcal{N}_m(\underline{\mu}_i, \underline{\Sigma}_i)$, où $j = 1, \dots, n_i$ et $i = 1, 2$.

Ainsi, $\underline{\underline{Y}} = \underbrace{\underline{\underline{\mu}}}_{\underline{\underline{\mu}}} + \underbrace{\mathcal{N}_{(n_1+n_2) \times m}(\underline{0}, \underline{I}_{n_1+n_2} \otimes \underline{\underline{\Sigma}})}_{\underline{\underline{\Sigma}}}$, si $\underline{\underline{\Sigma}}_1 = \underline{\underline{\Sigma}}_2 = \underline{\underline{\Sigma}}$.

D-ii

Nous pouvons donc écrire la vraisemblance afin de calculer $\hat{\underline{\mu}}$ et $\hat{\sigma}^2$:

$$\begin{aligned} L(\underline{\mu}, \underline{\Sigma}; \underline{Y}) &= (2\pi)^{-m(n_1+n_2)/2} |\underline{\Sigma}|^{-(n_1+n_2)/2} \text{etr} \left\{ -(1/2)(\underline{Y} - \underline{\mu})\underline{\Sigma}^{-1}(\underline{Y} - \underline{\mu})^\top \right\} \\ &= (2\pi)^{-m(n_1+n_2)/2} |\underline{\Sigma}|^{-(n_1+n_2)/2} \exp \left\{ -(1/2) \sum_{i=1}^2 \sum_{j=1}^{n_i} (\underline{X}_{ij} - \underline{\mu}_i)^\top \underline{\Sigma}^{-1} (\underline{X}_{ij} - \underline{\mu}_i) \right\}. \end{aligned}$$

Notons que Mendoza et Georgia [45] donnent l'équation de la vraisemblance lorsque $\Sigma_1 \neq \Sigma_2$.

Commençons par calculer $\hat{\underline{\mu}}$. Nous avons

$$\begin{aligned} l(\underline{\mu}_i, \underline{\Sigma}; \underline{X}_{ij}) &\equiv \log L(\underline{\mu}_i, \underline{\Sigma}; \underline{X}_{ij}) \\ &= \frac{-m(n_1+n_2)}{2} \log(2\pi) - \frac{(n_1+n_2)}{2} \log \underline{\Sigma} - (1/2) \sum_{i=1}^2 \sum_{j=1}^{n_i} (\underline{X}_{ij} - \underline{\mu}_i)^\top \underline{\Sigma}^{-1} (\underline{X}_{ij} - \underline{\mu}_i). \end{aligned}$$

Par conséquent, on peut écrire :

$$\begin{aligned} \frac{\partial l(\underline{\mu}_i, \underline{\Sigma}; \underline{X}_{ij})}{\partial \underline{\mu}_i} &= \frac{\partial \left\{ -(1/2) \sum_{i=1}^2 \sum_{j=1}^{n_i} (\underline{X}_{ij} - \underline{\mu}_i)^\top \underline{\Sigma}^{-1} (\underline{X}_{ij} - \underline{\mu}_i) \right\}}{\partial \underline{\mu}_i} = 0 \\ &\Rightarrow -\frac{1}{2} \frac{\partial \left\{ \sum_{j=1}^{n_i} (\underline{X}_{ij} - \underline{\mu}_i)^\top \underline{\Sigma}^{-1} (\underline{X}_{ij} - \underline{\mu}_i) \right\}}{\partial \underline{\mu}_i} = 0 \\ &\Rightarrow -\frac{1}{2} \sum_{j=1}^{n_i} 2 \underline{\Sigma}^{-1} (\underline{X}_{ij} - \underline{\mu}_i) (-1) = 0 \\ &\Rightarrow \sum_{j=1}^{n_i} (\underline{X}_{ij} - \underline{\mu}_i) = 0 \\ &\Rightarrow \hat{\underline{\mu}}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \underline{X}_{ij} = \bar{\underline{X}}_i \end{aligned} \tag{D.1.1}$$

et on obtient $\hat{\underline{\mu}} = \begin{pmatrix} \hat{\underline{\mu}}_1^\top \\ \hat{\underline{\mu}}_2^\top \end{pmatrix}$. Notez que nous avons utilisé la formule suivante :

$$\begin{aligned} \frac{\partial (\underline{\nu}^\top \underline{\mathcal{M}} \underline{\nu})}{\partial \underline{\nu}} &= (\underline{\mathcal{M}} + \underline{\mathcal{M}}^\top) \underline{\nu} \\ &= (2\underline{\mathcal{M}}) \underline{\nu} \quad \text{si } \underline{\mathcal{M}} \text{ est symétrique.} \end{aligned}$$

Nous pouvons aussi écrire la log-vraisemblance concentrée :

$$\begin{aligned}
l(\underline{\hat{\mu}}, \sigma^2; \underline{Y}) &= \log L(\underline{\hat{\mu}}, \sigma^2; \underline{Y}) \\
&= \frac{-m(n_1 + n_2)}{2} \log(2\pi) - \frac{(n_1 + n_2)}{2} \log |\underline{\Sigma}| + \text{tr} \left\{ -(1/2)(\underline{Y} - \underline{\hat{\mu}})\underline{\Sigma}^{-1}(\underline{Y} - \underline{\hat{\mu}})^\top \right\} \\
&= \underbrace{\frac{-m(n_1 + n_2)}{2} \log(2\pi)}_{cste} - \frac{(n_1 + n_2)}{2} \log \{ \sigma^{2m} |K_\rho| \} - \frac{1}{2\sigma^2} \text{tr} \left\{ (\underline{Y} - \underline{\hat{\mu}})K_\rho^{-1}(\underline{Y} - \underline{\hat{\mu}})^\top \right\} \\
&= cste - \frac{(n_1 + n_2)}{2} \left[\log \{ (\sigma^2)^m |K_\rho| \} - \frac{2}{(n_1 + n_2)} \frac{1}{2\sigma^2} \text{tr} \left\{ (\underline{Y} - \underline{\hat{\mu}})K_\rho^{-1}(\underline{Y} - \underline{\hat{\mu}})^\top \right\} \right] \\
&= cste - \frac{(n_1 + n_2)}{2} \log \left[[\sigma^2(1 - \rho)]^m \left(1 + \frac{\rho m}{(1 - \rho)} \right) \exp \left\{ \frac{1}{(n_1 + n_2)\sigma^2(1 - \rho)} \text{tr}(\dots) \right\} \right] \\
&= cste - \frac{(n_1 + n_2)}{2} \log \left[(1 - \rho)^m \left(1 + \frac{\rho m}{(1 - \rho)} \right) [\sigma^2]^m \exp \left\{ \frac{1}{(n_1 + n_2)(1 - \rho)} \frac{\text{tr}(\dots)}{\sigma^2} \right\} \right].
\end{aligned}$$

Maintenant, puisque $\arg \min_{\lambda > 0} \lambda^m e^{\frac{\alpha}{\lambda}} = \frac{\alpha}{m}$, nous avons :

$$\begin{aligned}
\tilde{\sigma}^2 &= \frac{1}{(n_1 + n_2)m(1 - \rho)} \text{tr} \left\{ (\underline{Y} - \underline{\hat{\mu}}) \left[\underline{\mathcal{I}}_m - \frac{\rho}{1 - \rho(1 - m)} J \right] (\underline{Y} - \underline{\hat{\mu}})^\top \right\} \\
&= \frac{\sigma^2}{(n_1 + n_2)m} \text{tr} \left\{ (\underline{Y} - \underline{\hat{\mu}}) \frac{1}{\sigma^2(1 - \rho)} \left[\underline{\mathcal{I}}_m - \frac{\rho}{1 - \rho(1 - m)} J \right] (\underline{Y} - \underline{\hat{\mu}})^\top \right\} \\
&= \frac{\sigma^2}{(n_1 + n_2)m} \text{tr} \left\{ (\underline{Y} - \underline{\hat{\mu}})\underline{\Sigma}^{-1}(\underline{Y} - \underline{\hat{\mu}})^\top \right\} \\
&= \frac{\sigma^2}{(n_1 + n_2)m} \text{tr} \left\{ (\underline{Y} - \underline{\hat{\mu}})\underline{\Sigma}^{-1/2}\underline{\Sigma}^{-1/2^\top}(\underline{Y} - \underline{\hat{\mu}})^\top \right\} \\
&= \frac{\sigma^2}{(n_1 + n_2)m} \text{tr} \left\{ \underline{\Sigma}^{-1/2^\top}(\underline{Y} - \underline{\hat{\mu}})^\top(\underline{Y} - \underline{\hat{\mu}})\underline{\Sigma}^{-1/2} \right\}.
\end{aligned}$$

Par les théorèmes 3.1.2 et 3.2.4 de Muirhead [47], on peut montrer que

$$(\underline{Y} - \underline{\hat{\mu}})^\top(\underline{Y} - \underline{\hat{\mu}}) \sim W_m(n_1 + n_2 - 2, \underline{\Sigma})$$

puis, par le théorème 3.2.20 de Muirhead [47], nous avons :

$$\begin{aligned}
\tilde{\sigma}^2 &= \frac{\sigma^2}{(n_1 + n_2)m} \text{tr} \left\{ W_m(n_1 + n_2 - 2, \underline{\Sigma}^{-1/2^\top} \underline{\Sigma} \underline{\Sigma}^{-1/2}) \right\} \\
&= \frac{\sigma^2}{(n_1 + n_2)m} \text{tr} \left\{ W_m(n_1 + n_2 - 2, \underline{\mathcal{I}}_m) \right\} \sim \frac{\sigma^2}{(n_1 + n_2)m} \chi_{(n_1 + n_2 - 2)m}^2.
\end{aligned}$$

Cet estimateur étant biaisé, nous prendrons plutôt :

$$\hat{\sigma}^2 = \frac{\sigma^2}{(n_1 + n_2 - 2)m} \text{tr} \left\{ (\underline{Y} - \underline{\hat{\mu}})\underline{\Sigma}^{-1}(\underline{Y} - \underline{\hat{\mu}})^\top \right\} \sim \frac{\sigma^2}{(n_1 + n_2 - 2)m} \chi_{(n_1 + n_2 - 2)m}^2.$$

D-iv

D.1.0.1. *Déterminant et inverse d'une matrice de variances-covariances ayant une structure compound-symetry*

Déterminant de $\underline{\underline{\Sigma}}$ [57] :

$$\begin{aligned} |\underline{\underline{\Sigma}}| &= |\sigma^2[(1-\rho)\underline{\underline{\mathcal{I}}}_m + \rho\underline{\underline{J}}]| \\ &= [\sigma^2(1-\rho)]^m \left| \underline{\underline{\mathcal{I}}}_m + \frac{\rho}{(1-\rho)}\underline{\underline{J}} \right|. \end{aligned}$$

$$\begin{aligned} \text{Or, } |\underline{\underline{\mathcal{I}}}_m + c\underline{\underline{J}}| &= |\underline{\underline{\mathcal{I}}}_m| + c\underline{\underline{V}}^\top \text{adj}(\underline{\underline{\mathcal{I}}}_m)\underline{\underline{V}} \\ &= 1 + cm. \end{aligned}$$

$$\text{Donc, } |\underline{\underline{\Sigma}}_*| = [\sigma^2(1-\rho)]^m \left(1 + \frac{\rho m}{(1-\rho)} \right).$$

Inverse de $\underline{\underline{\Sigma}}$ (selon le théorème 8.3.3 de Graybill [24]) :

$$\underline{\underline{\Sigma}} = \sigma^2(1/n_1 + 1/n_2)(1-\rho)\underline{\underline{\mathcal{I}}}_m + \sigma^2\rho\underline{\underline{J}},$$

qui est de la forme : $\underline{\underline{D}} + \alpha\underline{\underline{a}}\underline{\underline{b}}^\top$.

$$\underline{\underline{\Sigma}}^{-1} = \underline{\underline{D}}^{-1} + \gamma\underline{\underline{a}}^*\underline{\underline{b}}^{*\top}.$$

$$\underline{\underline{D}}^{-1} = \frac{1}{\sigma^2(1-\rho)}\underline{\underline{\mathcal{I}}}_m.$$

$$\gamma = -\sigma^2\rho \left[1 + \sigma^2\rho \frac{m}{\sigma^2(1-\rho)} \right]^{-1}.$$

$$\underline{\underline{a}}^* = \frac{1}{\sigma^2(1-\rho)}\underline{\underline{1}} \text{ et } \underline{\underline{b}}^* = \frac{1}{\sigma^2(1-\rho)}\underline{\underline{1}}.$$

$$\underline{\underline{a}}^*\underline{\underline{b}}^{*\top} = \frac{1}{\sigma^4(1-\rho)^2}\underline{\underline{J}}.$$

$$\gamma\underline{\underline{a}}^*\underline{\underline{b}}^{*\top} = -\frac{\sigma^2\rho}{\sigma^4(1-\rho)^2} \left[1 + \rho \frac{m}{(1-\rho)} \right]^{-1} \underline{\underline{J}}$$

$$= -\frac{\rho}{\sigma^2(1-\rho)} \frac{1}{1-\rho(1-m)} \underline{\underline{J}},$$

$$\text{donc } \underline{\underline{\Sigma}}^{-1} = \frac{1}{\sigma^2(1-\rho)} \left[\underline{\underline{\mathcal{I}}}_m - \frac{\rho}{1-\rho(1-m)}\underline{\underline{J}} \right].$$

Voir aussi à ce sujet l'énoncé de l'exercice 3.9 page 114 de Muirhead [47].

D.1.1. Calcul de la statistique T_k lorsque $\text{Var}(X_{ijk}) = \sigma^2$

Nous remarquons que peu importe l'hypothèse de *multi-sample sphericity* ou l'hypothèse de *multi-sample compound symmetry*, $\hat{\sigma}^2$ suit la même distribution.

Nous pouvons réécrire T_k (4.1.1) comme suit [23] :

$$\begin{aligned}
 T_k &= \frac{(\bar{X}_{1k} - \bar{X}_{2k} - \Delta_k)}{\hat{\sigma} \sqrt{1/n_1 + 1/n_2}} \\
 &= \frac{\frac{(\bar{X}_{1k} - \bar{X}_{2k} - \Delta_k)}{\sqrt{\sigma^2(\frac{1}{n_1} + \frac{1}{n_2})}}}{\sqrt{(n_1 + n_2 - 2)m \frac{\hat{\sigma}^2}{\sigma^2(n_1 + n_2 - 2)m}}} \\
 &\sim \frac{\mathcal{N}(\delta_k, 1)}{\sqrt{\frac{1}{(n_1 + n_2 - 2)m} (\chi_{m(n_1 + n_2 - 2)}^2)}} = t_{(n_1 + n_2 - 2)m}(\delta_k).
 \end{aligned}$$

La statistique T_k suit une loi de *Student* décentrée de paramètre δ_k avec $(n_1 + n_2 - 2)m$ degrés de liberté.

D.2. LORSQUE $\text{VAR}(X_{ijk}) = \sigma_i^2$

En posant le problème de façon similaire à la section D.1, nous posons :

$$\underline{\underline{Y}}_i = \begin{pmatrix} \underline{\underline{X}}_{i1}^\top \\ \vdots \\ \underline{\underline{X}}_{in_i}^\top \end{pmatrix} \sim \mathcal{N}_{n_i \times m} \left(\begin{pmatrix} \underline{\underline{\mu}}_i^\top \\ \vdots \\ \underline{\underline{\mu}}_i^\top \end{pmatrix}, \left(\underline{\underline{I}}_{n_i} \otimes \underline{\underline{\Sigma}}_i \right) \right),$$

avec $\underline{\underline{X}}_{ij} \sim \mathcal{N}_m(\underline{\underline{\mu}}_i, \underline{\underline{\Sigma}}_i)$, où $j = 1, \dots, n_i$ et $i = 1, 2$.

Ainsi, $\underline{\underline{Y}}_i = \underline{\underline{\mu}}_i + \underbrace{\mathcal{N}_{n_i \times m}(\underline{\underline{0}}, \underline{\underline{I}}_{n_i} \otimes \underline{\underline{\Sigma}}_i)}_{\underline{\underline{\varepsilon}}}$.

De la même façon que (D.1.1), nous obtenons $\hat{\underline{\underline{\mu}}}_i = \bar{\underline{\underline{X}}}_i$ et nous pouvons donc écrire la fonction de vraisemblance concentrée afin de calculer $\tilde{\sigma}_i^2$:

$$\begin{aligned}
L(\hat{\underline{\mu}}_i, \underline{\Sigma}_i; \underline{X}_{ij}) &= (2\pi)^{-mn_i/2} |\underline{\Sigma}_i|^{-n_i/2} \text{etr} \left\{ -(1/2) (\underline{Y}_i - \hat{\underline{\mu}}_i) \underline{\Sigma}_i^{-1} (\underline{Y}_i - \hat{\underline{\mu}}_i)^\top \right\} \\
&= (2\pi)^{-mn_i/2} |\underline{\Sigma}_i|^{-n_i/2} \exp \left\{ -(1/2) \sum_{j=1}^{n_i} (\underline{X}_{ij} - \bar{\underline{X}}_i)^\top \underline{\Sigma}_i^{-1} (\underline{X}_{ij} - \bar{\underline{X}}_i) \right\}.
\end{aligned}$$

Sous l'hypothèse de *multi-sample compound symmetry*, $\underline{\Sigma}_i = \sigma_i^2 [(1-\rho_i) \underline{\mathcal{I}}_m + \rho_i J]$, nous avons :

$$\begin{aligned}
l(\hat{\underline{\mu}}_i, \sigma_i^2; \underline{X}_{ij}) &\equiv \log L(\hat{\underline{\mu}}_i, \sigma_i^2; \underline{X}_{ij}) \\
&= \frac{-mn_i}{2} \log(2\pi) - \frac{n_i}{2} \log |\underline{\Sigma}_i| - (1/2) \sum_{j=1}^{n_i} (\underline{X}_{ij} - \bar{\underline{X}}_i)^\top \underline{\Sigma}_i^{-1} (\underline{X}_{ij} - \bar{\underline{X}}_i) \\
&= \frac{-mn_i}{2} \log(2\pi) - \frac{n_i}{2} \log \left\{ [\sigma_i^2(1-\rho)]^m \left(1 + \frac{\rho m}{(1-\rho)} \right) \right\} \\
&\quad - (1/2) \sum_{j=1}^{n_i} \left\{ (\underline{X}_{ij} - \bar{\underline{X}}_i)^\top \frac{1}{\sigma_i^2(1-\rho)} \left[\underline{\mathcal{I}}_m - \frac{\rho}{1-\rho(1-m)} J \right] (\underline{X}_{ij} - \bar{\underline{X}}_i) \right\} \\
&= \frac{-mn_i}{2} \log(2\pi) - \frac{n_i}{2} \log \left\{ [\sigma_i^2(1-\rho)]^m \left(1 + \frac{\rho m}{(1-\rho)} \right) \right\} \\
&\quad - \frac{1}{2\sigma_i^2(1-\rho)} \sum_{j=1}^{n_i} \left\{ (\underline{X}_{ij} - \bar{\underline{X}}_i)^\top \left[\underline{\mathcal{I}}_m - \frac{\rho}{1-\rho(1-m)} J \right] (\underline{X}_{ij} - \bar{\underline{X}}_i) \right\} \\
&= \underbrace{\frac{-mn_i}{2} \log(2\pi)}_{cste} - \frac{n_i}{2} \left[\log \left\{ [\sigma_i^2(1-\rho)]^m \left(1 + \frac{\rho m}{(1-\rho)} \right) \right\} \right. \\
&\quad \left. + \frac{2}{n_i} \frac{1}{2\sigma_i^2(1-\rho)} \sum_{j=1}^{n_i} \left\{ (\underline{X}_{ij} - \bar{\underline{X}}_i)^\top \left[\underline{\mathcal{I}}_m - \frac{\rho}{1-\rho(1-m)} J \right] (\underline{X}_{ij} - \bar{\underline{X}}_i) \right\} \right] \\
&= cste - \frac{n_i}{2} \log \left[[\sigma_i^2(1-\rho)]^m \left(1 + \frac{\rho m}{(1-\rho)} \right) \exp \left\{ \frac{1}{n_i \sigma_i^2(1-\rho)} \sum_{j=1}^{n_i} (\dots) \right\} \right] \\
&= cste - \frac{n_i}{2} \log \left[(1-\rho)^m \left(1 + \frac{\rho m}{(1-\rho)} \right) [\sigma_i^2]^m \exp \left\{ \frac{1}{n_i(1-\rho)} \sum_{j=1}^{n_i} (\dots) \right\} \right].
\end{aligned}$$

Maintenant, puisque $\arg \min_{\lambda > 0} \lambda^m e^{\frac{\alpha}{\lambda}} = \frac{\alpha}{m}$, nous avons :

$$\begin{aligned}
\tilde{\sigma}_i^2 &= \frac{1}{mn_i(1-\rho)} \sum_{j=1}^{n_i} \left\{ (\underline{X}_{ij} - \bar{\underline{X}}_i)^\top \left[\underline{\mathcal{I}}_m - \frac{\rho}{1-\rho(1-m)} J \right] (\underline{X}_{ij} - \bar{\underline{X}}_i) \right\} \\
&= \frac{1}{n_i m(1-\rho)} \text{tr} \left\{ (\underline{Y}_i - \hat{\underline{\mu}}_i) \left[\underline{\mathcal{I}}_m - \frac{\rho}{1-\rho(1-m)} J \right] (\underline{Y}_i - \hat{\underline{\mu}}_i)^\top \right\}. \\
&= \frac{\sigma^2}{mn_i} \text{tr} \left\{ \underline{\Sigma}_i^{-1/2 \top} (\underline{Y}_i - \hat{\underline{\mu}}_i)^\top (\underline{Y}_i - \hat{\underline{\mu}}_i) \underline{\Sigma}_i^{-1/2} \right\}.
\end{aligned}$$

De plus, nous avons vu plus tôt que $(\underline{Y}_i - \hat{\underline{\mu}}_i)^\top (\underline{Y}_i - \hat{\underline{\mu}}_i) \sim W_m(n_i, \underline{\Sigma}_i)$.

Ainsi, $\hat{\sigma}_i^2 = \frac{\sigma^2}{m(n_i-1)} \text{tr} \left\{ \underline{\Sigma}_i^{-1/2} (\underline{Y}_i - \hat{\underline{\mu}}_i)^\top (\underline{Y}_i - \hat{\underline{\mu}}_i) \underline{\Sigma}_i^{-1/2} \right\} \sim \frac{\sigma_i^2}{m(n_i-1)} \chi_{m(n_i-1)}^2$.

D.2.1. Calcul de la statistique T_k lorsque $\text{Var}(X_{ijk}) = \sigma_i^2$

Nous remarquons que peu importe l'hypothèse de *multi-sample sphericity* ou l'hypothèse de *multi-sample compound symmetry*, les σ_i^2 suivent la même distribution.

Nous pouvons donc réécrire T_k (4.1.1) comme suit [23] :

$$\begin{aligned} T_k &= \frac{\bar{X}_{1k} - \bar{X}_{2k} - \Delta_k}{\sqrt{\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}}} \\ &= \frac{(\bar{X}_{1k} - \bar{X}_{2k} - \Delta_k)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \\ &= \frac{(\bar{X}_{1k} - \bar{X}_{2k} - \Delta_k)}{\sqrt{\left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}\right)^{-1} \left[\frac{\hat{\sigma}_1^2}{n_1} + \frac{\hat{\sigma}_2^2}{n_2}\right]}} \\ &\approx \frac{\mathcal{N}(\delta_k, 1)}{\sqrt{\frac{\chi_f^2}{f}}} \approx t_f(\delta_k), \end{aligned}$$

où $\delta_k = \frac{\mu_{1k} - \mu_{2k} - \Delta_k}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ et où f est défini par la relation suivante au moyen de l'approximation de Cochran-Satterthwaite :

$$f = \frac{m \left(\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \right)^2}{\frac{\sigma_1^4}{n_1^2(n_1-1)} + \frac{\sigma_2^4}{n_2^2(n_2-1)}}.$$

La statistique T_k suit, à une constante près, une loi de *Student* décentrée de paramètre δ_k avec f degrés de liberté.

Annexe E

CALCUL DES ÉQUATIONS DE PUISSANCE

E.1. CALCUL DE LA PUISSANCE POUR LA MÉTHODE *Single Step* *Bonferroni* ÉCHANGEABLE

Calcul de l'équation (5.1.3) :

$$\begin{aligned}
 \Psi_{r,m}^{s,e} &= J_0 + \sum_{j=1}^{m-r} \binom{m}{j} \left\{ \mathbf{P} \left(\bigcap_{k=1}^j (T_k < c_m) \right) - \mathbf{P} \left[\left(\bigcap_{k=1}^j (T_k < c_m) \right) \cap \left(\bigcup_{l=j+1}^m (T_l < c_m) \right) \right] \right\} \quad \text{de (A.0.7)} \\
 &= J_0 + \sum_{j=1}^{m-r} \binom{m}{j} \left\{ \mathbf{P} \left(\bigcap_{k=1}^j (T_k < c_m) \right) - \mathbf{P} \left[\bigcup_{l=j+1}^m \left[\left(\bigcap_{k=1}^j (T_k < c_m) \right) \cap (T_l < c_m) \right] \right] \right\} \\
 &= J_0 + \sum_{j=1}^{m-r} \binom{m}{j} \left\{ \mathbf{P} \left(\bigcap_{k=1}^j (T_k < c_m) \right) - \mathbf{P} \left[\bigcup_{l=1}^{m-j} \left[\left(\bigcap_{k=1}^j (T_k < c_m) \right) \cap (T_{j+l} < c_m) \right] \right] \right\} \\
 &= J_0 + \sum_{j=1}^{m-r} \binom{m}{j} \left\{ \mathbf{P} \left(\bigcap_{k=1}^j (T_k < c_m) \right) - \sum_{t=1}^{m-j} (-1)^{t+1} \sum_{(i_1 < i_2 < \dots < i_t) \in F_{t,m-j}} \mathbf{P}(\cap_{u=1}^t E_{i_u}) \right\}, \quad \text{de (A.0.9)}
 \end{aligned}$$

$$\text{où } E_{i_u} = \left(\bigcap_{k=1}^j (T_k < c_m) \right) \cap (T_{j+i_u} < c_m)$$

$$\begin{aligned}
 &= J_0 + \sum_{j=1}^{m-r} \binom{m}{j} \left\{ \mathbf{P} \left(\bigcap_{k=1}^j (T_k < c_m) \right) \right. \\
 &\quad \left. - \sum_{t=1}^{m-j} (-1)^{t+1} \sum_{(i_1 < i_2 < \dots < i_t) \in F_{t,m-j}} \mathbf{P} \left[\left(\bigcap_{k=1}^j (T_k < c_m) \right) \cap \left(\bigcap_{u=1}^t (T_{j+i_u} < c_m) \right) \right] \right\} \\
 &= J_0 + \sum_{j=1}^{m-r} \binom{m}{j} \left\{ \mathbf{P} \left(\bigcap_{k=1}^j (T_k < c_m) \right) - \sum_{t=1}^{m-j} (-1)^{t+1} \binom{m-j}{t} \mathbf{P} \left(\bigcap_{k=1}^{j+t} (T_k < c_m) \right) \right\}, \quad \text{(E.1.1)}
 \end{aligned}$$

car nous sommes dans le cas échangeable.

E-ii

Expressions du terme $J_0 = \mathbf{P}[\text{rejeter } H_0^1, \dots, H_0^m]$ (5.1.4) :

$$\begin{aligned}
J_0 &= \mathbf{P} \left[\bigcap_{i=1}^m (T_i \geq c_m) \right] \\
&= 1 - \mathbf{P} \left[\bigcup_{i=1}^m (T_i < c_m) \right] \\
&= 1 - \sum_{r=1}^m (-1)^{r+1} \sum_{(i_1 < \dots < i_r) \in F_{r,m}} \mathbf{P} \left(\bigcap_{j=1}^r E_{i_j} \right) \\
&= 1 - \sum_{r=1}^m (-1)^{r+1} \binom{m}{r} \mathbf{P} \left(\bigcap_{k=1}^r (T_k < c_m) \right) \\
&= 1 + \sum_{r=1}^m \underbrace{(-1)^r}_{c_r} \binom{m}{r} \mathbf{P} \left(\bigcap_{k=1}^r (T_k < c_m) \right). \tag{E.1.2}
\end{aligned}$$

Calcul de (5.1.5). Pour simplifier la programmation, nous posons $P_j = \mathbf{P} \left(\bigcap_{k=1}^j T_k < c_m \right)$, ce qui nous permet de réécrire (5.1.3) ainsi :

$$\begin{aligned}
\Psi_{r,m}^{s,e} &= J_0 + \sum_{j=1}^{m-r} \binom{m}{j} \left\{ P_j - \sum_{t=1}^{m-j} (-1)^{t+1} \binom{m-j}{t} P_{j+t} \right\} \\
&= J_0 + \sum_{j=1}^{m-r} \binom{m}{j} P_j - \sum_{j=1}^{m-r} \left\{ \binom{m}{j} \sum_{t=1}^{m-j} (-1)^{t+1} \binom{m-j}{t} P_{j+t} \right\}.
\end{aligned}$$

Nous pouvons ensuite réécrire le terme de droite :

$$\begin{aligned}
\sum_{j=1}^{m-r} \binom{m}{j} \sum_{t=1}^{m-j} (-1)^{t+1} \binom{m-j}{t} P_{j+t} &= \sum_{j=1}^{m-r} \sum_{t=1}^{m-j} \binom{m}{j} (-1)^{t+1} \binom{m-j}{t} P_{j+t} \\
&= \sum_{k=2}^m \sum_{j=1}^{\min(k-1, m-r)} \binom{m}{j} (-1)^{k-j+1} \binom{m-j}{k-j} P_k \\
&\quad (\text{en posant } k = j + t) \\
&= \sum_{k=2}^m (-1) \underbrace{\sum_{j=1}^{\min(k-1, m-r)} \binom{m}{j} \binom{-(m-k+1)}{k-j}}_{b_k} P_k \quad \text{de (A.0.5)}. \tag{E.1.3}
\end{aligned}$$

Nous obtenons donc l'équation (5.1.3) sous la forme suivante :

$$\Psi_{r,m}^{s,e} = 1 + \sum_{k=1}^m (c_k + a_k + b_k) P_k \quad \text{où} \quad \begin{cases} a_k = \binom{m}{k}, k = 1, \dots, m-r \\ a_k = 0, k = m-r+1, \dots, m \\ b_1 = 0 \\ b_k = \sum_{j=0}^{\min(k-1, m-r)} \binom{m}{j} \binom{-(m-k+1)}{k-j}, k = 2, \dots, m \\ c_k = (-1)^k \binom{m}{k}, k = 1, \dots, m \end{cases} \quad (\text{E.1.4})$$

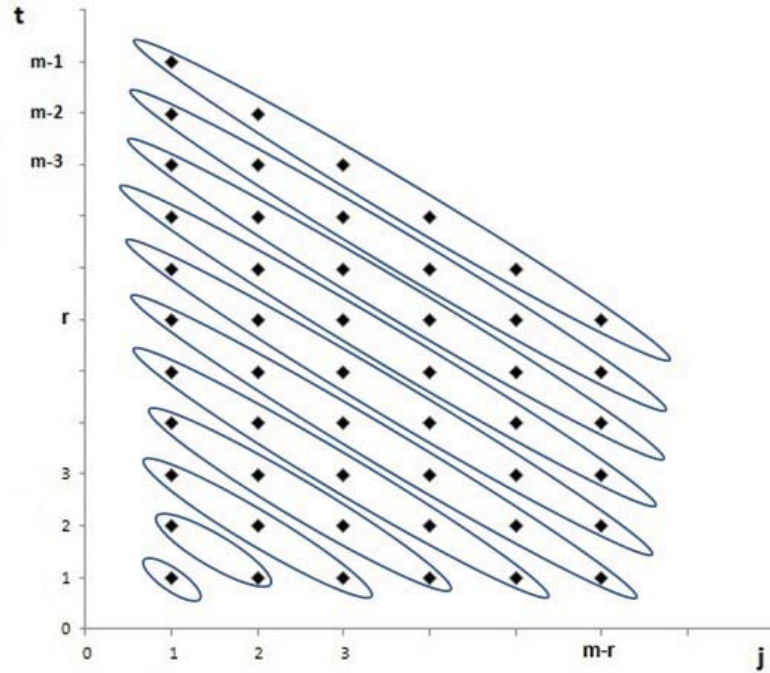


FIG. E.1. Graphique des indices j par rapport aux indices t

E-iv

E.2. CALCUL DE LA PUISSANCE POUR LA MÉTHODE *Step-Up Hochberg*

Calcul de l'équation (5.2.1)

$$\begin{aligned}
\Psi_{r,m}^{u,e} &= \sum_{j=0}^{m-r} \left\{ \mathbf{P} \left(\bigcap_{k=1}^j (T_{k:m} < u_k) \right) - \mathbf{P} \left[\left(\bigcap_{k=1}^j (T_{k:m} < u_k) \right) \cap (T_{j+1:m} < u_{j+1}) \right] \right\} \quad \text{de (A.0.7)} \\
&= \sum_{j=0}^{m-r} \left\{ \mathbf{P} \left(\bigcap_{k=1}^j (T_{k:m} < u_k) \right) - \mathbf{P} \left(\bigcap_{k=1}^{j+1} (T_{k:m} < u_k) \right) \right\} \\
&= 1 - \mathbf{P} \left(\bigcap_{k=1}^{m-r+1} (T_{k:m} < u_k) \right) \quad (\text{somme télescopique}) \\
&= 1 - (-1)^{\sum_{k=1}^{m-r+1} k} \sum_{\substack{\underline{a} \\ \underline{a}=\underline{w}}}^{a^*} (-1)^{a_+} \prod_{h=1}^{m-r+1} \binom{(\Delta a_h) - 1}{a_h - w_h} P_{\underline{a}},
\end{aligned}$$

avec $\underline{w} = (1, 2, \dots, m - r + 1)$. Dans le cas non-échangeable, nous avons :

$$P_{\underline{a}} = \sum_{\underline{l}}^l \mathbf{P} \left[\bigcap_{h=0}^{m-r} \left(\bigcap_{k=a_h+1}^{a_{h+1}} T_{l_k} < u_{h+1} \right) \right]$$

alors que dans le cas échangeable, nous obtenons plutôt :

$$P_{\underline{a}} = \binom{m}{\Delta^* \underline{a}} \mathbf{P} \left[\bigcap_{h=0}^{m-r} \left(\bigcap_{k=a_h+1}^{a_{h+1}} T_k < u_{h+1} \right) \right].$$

E.3. CALCUL DE LA PUISSANCE POUR LA MÉTHODE *Step-Down* *Holm*

Calcul de (E.3.1)

$$\begin{aligned}
\Psi_{r,m}^{d,e} &= J_0 + \sum_{j=1}^{m-r} \left\{ \mathbf{P}(T_{j:m} < d_j) - \mathbf{P} \left[(T_{j:m} < d_j) \cap \left(\bigcup_{k=j+1}^m (T_{k:m} < d_k) \right) \right] \right\} \quad \text{de (A.0.7)} \\
&= J_0 + \sum_{j=1}^{m-r} \left\{ \mathbf{P}(T_{j:m} < d_j) - \mathbf{P} \left[\bigcup_{k=j+1}^m [(T_{j:m} < d_j) \cap (T_{k:m} < d_k)] \right] \right\} \\
&= J_0 + \sum_{j=1}^{m-r} \left\{ \mathbf{P}(T_{j:m} < d_j) - \mathbf{P} \left[\bigcup_{t=1}^{m-j} E_{j+t} \right] \right\} \\
\text{où } E_k &= [(T_{j:m} < d_j) \cap (T_{k:m} < d_k)] \\
&= J_0 + \sum_{j=1}^{m-r} \left\{ \mathbf{P}(T_{j:m} < d_j) - \sum_{t=1}^{m-j} (-1)^{t+1} \sum_{(i_1 < i_2 < \dots < i_t) \in F_{t,m-j}} \mathbf{P}(\cap_{u=1}^t E_{j+i_u}) \right\} \quad \text{de (A.0.9)} \\
\text{où } F_{t,m-j} &= \{(i_1 < i_2 < \dots < i_t); i_k \in \{1, \dots, m-j\}, \forall k = 1, \dots, t\} \\
&= J_0 + \sum_{j=1}^{m-r} \left\{ \mathbf{P}(T_{j:m} < d_j) \right. \\
&\quad \left. - \sum_{t=1}^{m-j} (-1)^{t+1} \sum_{(i_1 < i_2 < \dots < i_t) \in F_{t,m-j}} \mathbf{P} \left[(T_{j:m} < d_j) \cap \left(\bigcap_{u=1}^t (T_{(j+i_u):m} < d_{j+i_u}) \right) \right] \right\} \\
&= J_0 + \sum_{j=1}^{m-r} \left\{ \mathbf{P}(T_{j:m} < d_j) \right. \\
&\quad \left. - \sum_{t=1}^{m-j} (-1)^{t+1} \sum_{(i_1 < i_2 < \dots < i_t) \in F_{t,m-j}} \mathbf{P} \left(\bigcap_{u=0}^t (T_{(j+i_u):m} < d_{j+i_u}) \right) \right\} \\
\text{en posant } i_0 &= 0 \\
&= J_0 + \sum_{j=1}^{m-r} \left\{ \mathbf{P}(T_{j:m} < d_j) \right. \\
&\quad \left. - \sum_{t=1}^{m-j} (-1)^{t+1} \sum_{(i_1 < i_2 < \dots < i_t) \in F_{t,m-j}} \mathbf{P} \left(\bigcap_{u=1}^{t+1} (T_{(j+i_{u-1}):m} < d_{j+i_{u-1}}) \right) \right\}. \quad \text{(E.3.1)}
\end{aligned}$$

Puis, en utilisant (B.2.3) dans (E.3.1)

$$\begin{aligned} \Psi_{r,m}^{d,e} &= J_0 + \sum_{j=1}^{m-r} \left\{ \left[(-1)^j \sum_{a=j}^m (-1)^a \binom{a-1}{a-j} P_a \right] \right. \\ &\quad \left. - \sum_{t=1}^{m-j} (-1)^{t+1} \sum_{(i_1 < i_2 < \dots < i_t) \in F_{t,m-j}} \left[(-1)^{\sum_{h=0}^t (j+i_h)} \sum_{\underline{b}=\underline{w}}^{\underline{b}^*} (-1)^{b_+} \prod_{h=1}^{t+1} \binom{(\Delta b_h) - 1}{b_h - t_h} P_{\underline{b}} \right] \right\} \end{aligned}$$

où $\underline{w} = (j + i_0, j + i_1, j + i_2, \dots, j + i_t)$ et $i_0 = 0$.

$$\begin{aligned} \Psi_{r,m}^{d,e} &= J_0 + \sum_{j=1}^{m-r} \left\{ \left[(-1)^j \sum_{a=j}^m (-1)^a \binom{a-1}{a-j} P_a \right] \right. \\ &\quad \left. - \sum_{t=1}^{m-j} (-1)^{t+1} \sum_{(i_1 < i_2 < \dots < i_t) \in F_{t,m-j}} \left[(-1)^{((t+1)j + \sum_{h=1}^t i_h)} \sum_{\underline{b}=\underline{w}}^{\underline{b}^*} (-1)^{b_+} \prod_{h=1}^{t+1} \binom{(\Delta b_h) - 1}{b_h - (j + i_{h-1})} P_{\underline{b}} \right] \right\}. \end{aligned} \tag{E.3.2}$$

Expressions du terme $J_0 = \mathbf{P}[\text{rejeter } H_0^1, \dots, H_0^m]$ (5.3.3) :

$$\begin{aligned} J_0 &= \mathbf{P} \left[\bigcap_{k=1}^m (T_{k:m} \geq d_k) \right] \\ &= 1 - \mathbf{P} \left[\bigcup_{k=1}^m (T_{k:m} < d_k) \right] \quad \text{de (A.0.9) on obtient :} \\ &= 1 - \left\{ \sum_{t=1}^m (-1)^{t+1} \sum_{(i_1 < i_2 < \dots < i_t) \in F_{t,m}} \left[(-1)^{\sum_{h=1}^t i_h} \sum_{\underline{c}=\underline{s}}^{\underline{c}^*} (-1)^{c_+} \prod_{h=1}^t \binom{(\Delta c_h) - 1}{c_h - i_h} P_{\underline{c}} \right] \right\}, \end{aligned} \tag{E.3.3}$$

où $\underline{s} = (j + i_1, j + i_2, \dots, j + i_t) = (i_1, i_2, \dots, i_t)$.

Dans le cas non-échangeable :

$$\begin{aligned} P_a &= \sum_{\underline{l}}' \mathbf{P} \left(\bigcap_{k=1}^a T_{l_k} < d_j \right) \\ \text{et } P_{\underline{b}} &= \sum_{\underline{l}}' \mathbf{P} \left[\bigcap_{h=0}^t \left(\bigcap_{k=b_h+1}^{b_{h+1}} T_{l_k} < d_{j+i_h} \right) \right], \\ \text{et } P_{\underline{c}} &= \sum_{\underline{l}}' \mathbf{P} \left[\bigcap_{h=0}^{t-1} \left(\bigcap_{k=c_h+1}^{c_{h+1}} T_{l_k} < d_{i_{h+1}} \right) \right]. \end{aligned}$$

Dans le cas échangeable :

$$P_a = \binom{m}{a} \mathbf{P} \left(\bigcap_{k=1}^a T_k < d_j \right)$$

$$\text{et } P_{\underline{b}} = \binom{m}{\Delta^* \underline{b}} \mathbf{P} \left[\bigcap_{h=0}^t \left(\bigcap_{k=b_h+1}^{b_{h+1}} T_k < d_{j+i_h} \right) \right],$$

$$\text{et } P_{\underline{c}} = \binom{m}{\Delta^* \underline{c}} \mathbf{P} \left[\bigcap_{h=0}^{t-1} \left(\bigcap_{k=c_h+1}^{c_{h+1}} T_k < d_{i_{h+1}} \right) \right].$$