

# Le modèle linéaire à effets mixtes pour analyser des données génétiques provenant de familles

par

Catherine ALLARD

Mémoire présenté au Département de mathématiques  
en vue de l'obtention du grade de  
maître ès sciences (M. Sc.)

FACULTÉ DES SCIENCES  
UNIVERSITÉ DE SHERBROOKE

Sherbrooke, Québec, Canada, 10 avril 2015

Le 10 avril 2015,

*le jury a accepté le mémoire de madame Catherine Allard  
dans sa version finale.*

Membres du jury

Professeur Ernest Monga  
Directeur de recherche  
Département de mathématiques

Professeure Marie-France Hivert  
Codirectrice de recherche  
Département de médecine

Professeur Bernard Colin  
évaluateur interne  
Département de mathématiques

Professeur Shengrui Wang  
évaluateur externe  
Département informatique

Professeur Taoufik Bouezmarni  
Président-rapporteur  
Département de mathématiques

Je dédie ce mémoire à tous ceux  
qui m'ont soutenue dans mes choix  
qui ne faisaient pas toujours l'unanimité.

# SOMMAIRE

Nous désirons savoir quelles sont les variations génétiques qui sont associées à une tension artérielle élevée. Pour ce faire, nous avons des données provenant de plusieurs familles, c'est-à-dire qu'il y a des personnes de la même famille qui se retrouvent dans cet échantillon. Dans cette base de données, il y a de l'information sur quelques caractéristiques démographique (âge, sexe, fumeur/non fumeur), il y a aussi la pression diastolique et systolique ainsi qu'un grand nombre de variations génétiques distribuées sur tout le génome. Pour pouvoir analyser des observations qui ne sont pas indépendantes, nous devons utiliser un modèle qui diffère un peu de la régression classique. En effet, nous ne pouvons pas utiliser la régression classique, car notre échantillon ne respecte pas toutes les hypothèses du modèle. Le modèle que nous allons utiliser prend en compte la covariance entre les individus de même famille.

Nous allons donc présenter la théorie du modèle linéaire à effets mixtes simple ainsi que sa généralisation pour des données génétiques provenant de familles. Nous allons terminer par une application de ce modèle généralisé à notre base de données sur la tension artérielle pour déterminer quelles parties du génome (quelles variations génétiques) expliquent le mieux la tension artérielle de cet échantillon.

*Mots-clés : Modèle linéaire à effets mixtes, génétique, biostatistique, tension artérielle, généalogie*

# REMERCIEMENTS

Je voudrais premièrement remercier la Dre Marie-France Hivert qui, malgré le fait qu'elle était la plupart du temps à l'extérieur, a été très présente et très motivante. Je voudrais aussi remercier M. Ernest Monga pour son suivi et son soutien pour assimiler de nouvelles notions mathématiques.

Je voudrais aussi dire merci à Mme Josée Dupuis, professeure à Boston University, pour sa grande patience, ses réponses ultra rapides et sa disponibilité inégalée. Merci aussi à toute l'équipe des Drs Perron et Hivert de m'avoir accompagnée durant ma maîtrise et mes stages.

Merci au FRQ-s et aux IRSC pour leur soutien financier qui a enlevé un stress et m'a aidée à me concentrer sur mes études. Merci aussi au Centre de Recherche du CHUS et à la Faculté de médecine et des sciences de la santé pour les installations et l'accès aux ordinateurs. Cela m'a permis d'être plus proche de l'équipe de travail et de mieux les aider.

Merci aussi à l'équipe du Genetic Analysis Workshop 18 et au consortium T2D-GENES pour m'avoir permis d'utiliser leurs données dans mon mémoire, j'en suis très reconnaissante.

Et finalement, un merci tout spécial à mon copain Francis qui m'a toujours soutenue dans mes choix et qui m'a encouragée tout au long de ma maîtrise.

Mes plus sincères remerciements à toutes et à tous,

Catherine Allard

Sherbrooke, avril 2015

# TABLE DES MATIÈRES

<b>SOMMAIRE</b>	<b>iv</b>
<b>REMERCIEMENTS</b>	<b>v</b>
<b>TABLE DES MATIÈRES</b>	<b>vii</b>
<b>LISTE DES TABLEAUX</b>	<b>x</b>
<b>LISTE DES FIGURES</b>	<b>xii</b>
<b>CHAPITRE 1 — Problématique</b>	<b>1</b>
1.1 Mise en contexte . . . . .	1
1.2 La question de recherche . . . . .	2
1.3 Les variations génétiques, ce qu'il faut savoir . . . . .	2
1.4 Description des données . . . . .	4
1.5 Quel modèle choisir ? . . . . .	5
1.6 Objectifs . . . . .	5

<b>CHAPITRE 2 — Théorie du modèle linéaire à effets mixtes usuel</b>	<b>6</b>
2.1 Introduction du modèle linéaire à effets mixtes . . . . .	6
2.2 Analyse de la variance à un facteur . . . . .	9
2.3 Régression linéaire à une variable explicative . . . . .	42
<b>CHAPITRE 3 — Théorie du modèle à effets mixtes généralisé</b>	<b>101</b>
3.1 Quelle est la différence avec les modèles usuels? . . . . .	102
3.2 La matrice $K$ . . . . .	103
3.3 Théorie du nouveau modèle . . . . .	104
<b>CHAPITRE 4 — Application aux données réelles</b>	<b>123</b>
4.1 Choix du modèle . . . . .	123
4.2 Explication de la procédure . . . . .	123
4.3 Résultats . . . . .	124
<b>CHAPITRE 5 — Conclusion</b>	<b>128</b>
<b>BIBLIOGRAPHIE</b>	<b>130</b>
<b>ANNEXE A — Quelques notions mathématiques</b>	<b>132</b>
A.1 Vecteurs et matrices de uns . . . . .	132
A.2 Produit de Kronecker . . . . .	135
A.3 Calcul différentiel avec des matrices et des vecteurs . . . . .	136



A.4	Pseudo-inverse . . . . .	140
<b>ANNEXE B — Quelques démonstrations supplémentaires</b>		<b>143</b>
B.1	Loi de certaines statistiques . . . . .	143
B.2	Variance et covariance conditionnelle . . . . .	161
B.3	Autres preuves pour l'ANOVA à un facteur aléatoire . . . . .	162
B.4	Autres preuves pour la régression linéaire simple classique . . . . .	171
B.5	Autres preuves pour l'ANCOVA avec un facteur fixe . . . . .	181
B.6	Autres preuves pour la régression linéaire simple avec un effet aléatoire et des données équilibrées . . . . .	192
B.7	Autres preuves pour la régression linéaire simple avec un effet aléatoire et des données non équilibrées . . . . .	200
<b>ANNEXE C — Tableau des lois des sommes de carrés</b>		<b>222</b>

# LISTE DES TABLEAUX

2.1	Données sur la croissance des champignons . . . . .	7
2.2	Données sur la réponse à un médicament . . . . .	8
2.3	Sommes de carrés, ANOVA fixe . . . . .	10
2.4	Sommes de carrés, ANOVA aléatoire . . . . .	20
2.5	Comparaison entre les différents estimateurs - ANOVA à un facteur . . .	41
2.6	Comparaison des tests d'hypothèse - ANOVA à un facteur . . . . .	41
2.7	Sommes de carrés, régression linéaire simple . . . . .	43
2.8	Sommes de carrés, ANCOVA fixe . . . . .	48
2.9	Sommes de carrés, régression linéaire avec effet aléatoire, données équilibrées	61
2.10	Sommes de carrés, régression linéaire avec effet aléatoire, données non équilibrées . . . . .	86
2.11	Comparaison entre les différents estimateurs - régression linéaire simple .	99
2.12	Comparaison des tests d'hypothèse - régression linéaire simple . . . . .	100
3.1	Comparaison entre le modèle généralisé et le modèle usuel . . . . .	102

3.2	Sommes de carrés, modèle généralisé aux données génétiques provenant de familles . . . . .	106
C.1	Lois des sommes de carrés pour l'ANOVA . . . . .	222
C.2	Lois des sommes de carrés pour l'ANCOVA . . . . .	223

# LISTE DES FIGURES

1.1	ADN : acide désoxyribonucléique . . . . .	3
1.2	SNP : polymorphisme d'un seul nucléotide . . . . .	3
2.1	Esquisse de la vraisemblance par rapport à $\sigma^2$ . . . . .	34
4.1	Graphique Manhattan de la pression artérielle systolique . . . . .	125
4.2	Graphique Manhattan de la pression artérielle diastolique . . . . .	126

# CHAPITRE 1

## Problématique

### 1.1 Mise en contexte

Il faut savoir que plusieurs grandes études recueillent des données génétiques sur plusieurs générations à l'intérieur de grandes familles comme la Framingham Heart Study (FHS) [Nat48] et le Centre d'Étude du Polymorphisme Humain (CEPH) [Dau84]. Il existe plusieurs tests pour analyser des données génétiques provenant de familles, mais la plupart des tests ne prennent que des caractéristiques binaires comme variable dépendante. D'autres tests ont parfois des problèmes de puissance ou ils négligent certaines informations ou encore ils dépendent d'hyperparamètres qu'il faut trouver à l'aide de techniques qui nécessitent un grand temps de calcul. Par contre, il y a un modèle qui permet d'inclure des observations non indépendantes et d'introduire facilement des covariables. Ce modèle est une généralisation du modèle linéaire à effets mixtes usuel et il est considérablement utilisé en analyse génétique. Nous l'introduisons dans ce document.

L'hypertension est un fléau grandissant dans la population nord-américaine avec la montée du taux d'obésité, un des plus grands facteurs de risque de cette condition. L'hyper-

tension est définie par une pression artérielle systolique de plus de 140 mmHg ou une pression diastolique de plus de 90 mmHg et elle peut apporter d'importantes complications comme l'accident vasculaire cérébral et la crise cardiaque. C'est donc important de comprendre les mécanismes de cette maladie et d'identifier quelles sont les personnes les plus à risque. C'est pourquoi l'application du modèle servira à déterminer quelles sont les variations génétiques qui sont reliées à l'hypertension ou plus précisément à une pression sanguine élevée.

## **1.2 La question de recherche**

Nous avons accès à une base de données qui contient des données génétiques et qui comporte des participants provenant de plusieurs familles. Comme nous avons mentionné plus haut, nous aimerions savoir quelles sont les variations génétiques qui sont les plus associées à la pression artérielle de la population ciblée.

## **1.3 Les variations génétiques, ce qu'il faut savoir**

L'acide désoxyribonucléique (ADN) est une structure qui est présente dans chaque cellule et qui contient toute l'information nécessaire pour le développement et le fonctionnement du corps humain. Il est constitué d'une chaîne de nucléotides (A, C, G, ou T) en forme de double hélice (voir la figure 1.1). Les deux brins qui composent la double hélice sont complémentaires, c'est-à-dire que s'il y a un A sur un brin, alors il y a un T à la même position sur l'autre brin et vice-versa, de même avec C et G. Donc, un seul brin permet de connaître toute l'information génétique.

Puisque nous ne sommes pas tous des jumeaux identiques, il y a des différences dans

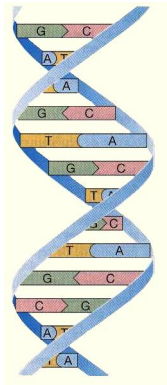


Figure 1.1 – ADN : acide désoxyribonucléique  
<http://www.pasteur-lille.fr/fr/recherche/u744/resultat/adn.html>

la séquence ADN entre les personnes dans la population. Une différence s'appelle une variation génétique ou un SNP ("single-nucleotide polymorphism" ou polymorphisme d'un seul nucléotide). La figure suivante représente un SNP. En effet, supposons que la portion d'ADN du haut représente la personne 1 et que la portion d'ADN du bas représente la personne 2, alors si nous regardons le brin inférieur de la double hélice, nous remarquons qu'à cette place précise sur le génome (la séquence de nucléotide), la personne 1 et la personne 2 n'ont pas le même nucléotide. Le nucléotide du brin supérieur s'ajuste pour être complémentaire (A avec T et C avec G).

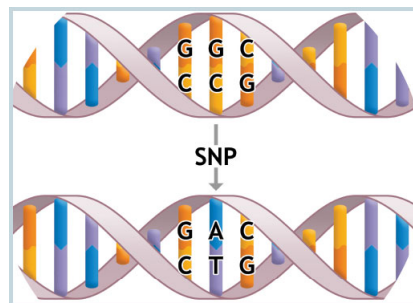


Figure 1.2 – SNP : polymorphisme d'un seul nucléotide  
<http://www.ibbl.lu/personalised-medicine/what-is-personalised-medicine/dna-genes-snps/>

Les différents nucléotides possibles pour un SNP se nomment allèles. Dans la figure 1.2, si nous regardons encore le brin inférieur de la double hélice, nous remarquons que les deux allèles ici sont T et C.

Ce qu'il faut savoir ensuite c'est que l'ADN est réparti sur 23 paires de chromosomes. La moitié de ces chromosomes viennent du père et l'autre moitié de la mère. Il peut donc arriver qu'à une place précise sur le génome, une personne ait un A sur un chromosome et un C sur l'autre chromosome de la paire puisqu'il y a un SNP dans la population. Nous dirons alors que le génotype de cette personne est AC. Ceci veut aussi dire que soit son père lui a transmis le A et sa mère le C, soit le contraire.

Finalement, lorsque nous parlons d'un phénotype, cela désigne une caractéristique physique ou biochimique observable (à l'aide d'une technique quelconque) qui est le résultat de l'interaction entre l'environnement de l'individu et son profil génétique.

## 1.4 Description des données

Les données utilisées proviennent du consortium T2D-GENES et elles ont été utilisées lors du Genetic Analysis Workshop 18 (GAW18) qui est un congrès où les chercheurs reçoivent tous la même base de données génétiques et ils doivent développer ou utiliser des méthodes pour bien analyser le type de données fournies.

La base de données contient de l'information complète sur 795 participants (génétique et tension artérielle). L'information pour chaque participant contient, entre autres, 472 049 variations génétiques en plus de leur pression systolique et diastolique (et quelques autres informations qui ne seront pas utilisées). Il y a en tout 20 grandes familles contenant chacune environ une quarantaine de personnes.



## 1.5 Quel modèle choisir ?

Nous cherchons à savoir s'il y a une association entre deux variables (SNP et pression sanguine). Nous allons traiter la variable SNP de la même façon qu'une variable quantitative discrète, car le SNP sera codé 0, 1 ou 2 selon le nombre d'allèles mineurs (allèle ayant la plus petite fréquence dans la population) qu'une personne possède (modèle additif, le plus souvent utilisé). Le modèle le plus simple serait la régression linéaire simple, par contre, une des hypothèses de la régression linéaire simple est que les observations doivent être indépendantes, ce qui n'est pas respecté lorsque nous avons des données provenant de familles. En effet, le parent et l'enfant ont la moitié de leur ADN en commun, donc si le parent est AA, l'enfant ne peut pas être BB. Pour contrer ce problème, nous pouvons utiliser le modèle linéaire avec effets mixtes que nous allons introduire au chapitre 2. Par contre, nous ne pouvons pas utiliser ce modèle selon la théorie habituelle, il faut faire une petite modification qui est en fait une généralisation pour pouvoir l'utiliser avec des participants provenant de plusieurs familles (voir le chapitre 3). Ce dernier modèle va nous permettre de prendre en compte la covariance estimée entre les différents participants de la base de données.

## 1.6 Objectifs

Les objectifs de cette recherche sont les suivants :

1. Présenter la théorie du modèle linéaire à effets mixtes usuel
2. Présenter la théorie du modèle linéaire à effets mixtes généralisé pour des données familiales
3. Répondre à la question de recherche, donc appliquer le modèle linéaire à effets mixtes généralisé à des données réelles

# CHAPITRE 2

## Théorie du modèle linéaire à effets mixtes usuel

### 2.1 Introduction du modèle linéaire à effets mixtes

Tout au long de ce document, il sera question des modèles linéaires (ML), c'est-à-dire que l'espérance de la variable étudiée, aussi appelée observation (souvent notée  $y_{ij}$ ), est une fonction linéaire de certains paramètres (inconnus, mais que nous désirons estimer). Voici quelques exemples de modèles linéaires :

**Exemple 2.1** (Modèles linéaires).

$$\mathbb{E}[y_{ij}] = \mu + \alpha_i + \beta_j$$

$$\mathbb{E}[y_{ij}] = \mu + \beta_1 x_{1ij} + \beta_2 x_{2ij}$$

$$\mathbb{E}[y_{ij}] = \mu + \beta_1 x_{1ij} + \beta_2 x_{2ij}^2$$

Remarquons que nous utilisons une notation spéciale ici en utilisant le même symbole

$(y_{ij})$  pour la variable aléatoire  $(Y_{ij})$  et l'observation, qui est la réalisation de cette variable aléatoire. Cette notation sera utilisée dans l'ensemble du document.

Une portion des modèles linéaires considère que certains paramètres ne sont pas fixes, c'est-à-dire qu'ils sont analysés comme des réalisations de variables aléatoires. Ces modèles seront nommés les modèles linéaires mixtes (MLM). À partir de maintenant, les paramètres notés par une lettre grecque seront les paramètres fixes (effets fixes) et ceux notés par une lettre romaine seront les paramètres aléatoires (effets aléatoires). Lorsqu'il y a un effet aléatoire dans le modèle, la forme de l'énoncé du modèle va un peu changer.

**Exemple 2.2** (Modèles avec effets mixtes).

$$\mathbb{E}[y_{ij}|a_i] = \mu + a_i$$

$$\mathbb{E}[y_{ij}|a_i] = \mu + a_i + \beta x_{ij}$$

Maintenant, sur quoi se base-t-on pour décider si un facteur doit être considéré comme fixe ou aléatoire? Ce n'est pas toujours simple, en règle générale, il suffit de se poser la question suivante : est-il raisonnable de penser que les niveaux du facteur proviennent d'une distribution de probabilité? Si la réponse est positive, alors ce facteur sera analysé comme aléatoire et si la réponse est négative, il sera analysé comme fixe. Voici deux exemples pour illustrer ce propos.

**Exemple 2.3.** *Nous nous intéressons à la croissance de différentes sortes de champignons pour différentes températures. Les données sont les suivantes :*

Températures	Giroles	Pleurotes	De Paris
10 ° C	1.98	4.77	5.02
20 ° C	4.48	2.70	8.13
30 ° C	3.48	6.65	2.12
40 ° C	4.66	4.43	5.34

Tableau 2.1 – Données sur la croissance des champignons

*Ici, il est raisonnable de penser que ces sortes de champignons ont été choisies aléatoirement parmi un très grand nombre d'espèces, donc ce facteur (la sorte de champignon) sera considéré comme aléatoire.*

**Exemple 2.4.** *Nous voulons tester la réponse à un médicament en utilisant deux groupes (placebo et médicamenté). Les données sont les suivantes :*

Individus	Groupe	Mesure
1	placebo	1.3
2	médicamenté	4.3
3	médicamenté	3.1
4	placebo	2.4
⋮	⋮	⋮

Tableau 2.2 – Données sur la réponse à un médicament

*Nous pouvons croire qu'il n'existe pas d'autres niveaux au facteur « Groupe » (à part placebo et traité) puisque nous n'avons pas choisi les niveaux placebo et médicamenté parmi un grand nombre de niveaux. Ce facteur sera donc considéré comme fixe.*

La théorie présentée dans ce chapitre vient en grande partie du livre *Generalized, Linear, and Mixed Models* [MSN08]. Il y a quelques ajouts, quelques modifications (surtout dans l'ordre de présentation) et plusieurs preuves ajoutées. Les prochaines sections seront divisées en général comme suit,

1. Le modèle
2. Les covariances entre les observations (s'il y a un effet aléatoire)
3. Les sommes de carrés
4. La vraisemblance
5. Les équations du maximum de vraisemblance et leurs solutions

6. Les estimateurs du maximum de vraisemblance
7. Espérance et biais des estimateurs
8. Variance des estimateurs (et leur loi)
9. Tests d'hypothèse et intervalles de confiance
10. Prédiction (s'il y a un effet aléatoire)

Les sections 7 et 8 peuvent parfois être regroupées en une section « Les lois des estimateurs » si c'est plus simple selon le modèle. Nous allons débiter avec le modèle le plus simple, c'est-à-dire l'analyse de la variance à un facteur (fixe et aléatoire), nous allons ensuite présenter la théorie de la régression linéaire simple (classique, ANCOVA et avec une ordonnée à l'origine aléatoire).

## **2.2 Analyse de la variance à un facteur**

Commençons par énoncer la théorie du modèle le plus simple, c'est-à-dire la modélisation de la situation où les observations  $(y_{ij})$  sont divisées en  $m$  groupes. La majorité du temps, les personnes qui utilisent ce modèle cherchent à savoir s'il y a une différence entre les moyennes des observations de chaque groupe. Nous supposons que les observations sont normalement distribuées.

### **2.2.1 Modèle avec un effet fixe (ANOVA)**

Nous allons débiter par le modèle usuel, c'est-à-dire avec un effet fixe, et nous regarderons ensuite le modèle avec un effet aléatoire.

**a) Le modèle**

Nous considérons le modèle suivant.

$$\begin{aligned} \mathbb{E}[y_{ij}] &= \mu + \alpha_i \\ y_{ij} &\sim \mathcal{N}(\mu + \alpha_i, \sigma^2) \text{ indépendantes} \end{aligned} \tag{2.1}$$

Sous la contrainte  $\sum_{i=1}^m n_i \alpha_i = 0$  pour ne pas qu'il y ait d'ambiguïté sur  $\mu$ . En procédant comme cela,  $\mu$  désigne la moyenne globale de la population.

Où  $i = 1, \dots, m$  représente le  $i^e$  groupe et où  $j = 1, \dots, n_i$  représente la  $j^e$  observation du  $i^e$  groupe. Ainsi,  $\mu + \alpha_i$  représente la moyenne du groupe  $i$ . Aussi, notons  $N = \sum_{i=1}^m n_i$ .

**b) Les sommes de carrés**

Voici les sommes de carrés qui interviennent dans ce modèle.

Notation	Expression	Carré moyen	Signification
$SCE$	$\sum_{i=1}^m \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y}_{..})^2$	$CME = \frac{SCE}{m-1}$	Carrés expliqués
$SCR$	$\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$	$CMR = \frac{SCR}{N-m}$	Carrés résiduels
$SCT$	$\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2$	—	Carrés totaux

Tableau 2.3 – Tableau des sommes de carrés, ANOVA avec un effet fixe

Avec  $SCT = SCE + SCR$ .

**c) La vraisemblance**

Voici une proposition concernant la vraisemblance.

**Proposition 2.1.**  $L = (2\pi\sigma^2)^{-N/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i)^2\right)$

*Démonstration.*

$$\begin{aligned} L &= \prod_{i=1}^m \prod_{j=1}^{n_i} f_{y_{ij}}(y_{ij}) \\ &= \prod_{i=1}^m \prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y_{ij} - (\mu + \alpha_i))^2}{2\sigma^2}} \\ &= (2\pi\sigma^2)^{-N/2} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i)^2\right) \quad \square \end{aligned}$$

Voici maintenant la log-vraisemblance.

$$l = \log L = -\frac{N}{2} \log(2\pi) - \frac{N}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i)^2 \quad (2.2)$$

#### d) Les équations du maximum de vraisemblance et leurs solutions

les solutions du maximum de vraisemblance seront notées avec un point et les estimateurs du maximum de vraisemblance seront notés avec un chapeau (ex.  $\hat{\sigma}^2$  la solution versus  $\hat{\sigma}^2$  l'estimateur de  $\sigma^2$ ).

**Proposition 2.2.** *Les solutions du maximum de vraisemblance sont*

$$\hat{\mu} = \bar{y}_{..} \quad \hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..} \quad \hat{\sigma}^2 = \frac{SCR}{N}$$

*Démonstration.* Voici les équations obtenues lorsque les dérivées partielles sont égalées à 0.

$$\frac{\partial l}{\partial \mu} = 0 \Rightarrow \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i) = 0$$

$$\frac{\partial l}{\partial \alpha_i} = 0 \Rightarrow \frac{1}{\sigma^2} \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i) = 0$$

$$\frac{\partial l}{\partial \sigma^2} = 0 \Rightarrow -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i)^2 = 0$$

Les solutions de ces équations sont données par

$$\hat{\mu} = \bar{y}_{..} = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij}$$

$$\hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..} = \frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij} - \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij}$$

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu} - \hat{\alpha}_i)^2 = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 = \frac{SCR}{N} \quad \square$$

Ces derniers résultats représentent les solutions du maximum de vraisemblance, mais pas nécessairement les estimateurs, car pour être un estimateur viable, la solution doit être dans le domaine possible du paramètre. C'est pourquoi les solutions sont notées différemment des estimateurs.

### e) Les estimateurs du maximum de vraisemblance

Puisque  $\hat{\mu} \in \mathbb{R}$ ,  $\hat{\alpha}_i \in \mathbb{R} \forall i$  et que  $\hat{\sigma}^2 > 0$  alors ce sont aussi les estimateurs du maximum de vraisemblance, ainsi



**Proposition 2.3.**

$$\hat{\mu} = \dot{\mu} \quad \hat{\alpha}_i = \dot{\alpha}_i \forall i \quad \hat{\sigma}^2 = \dot{\sigma}^2$$

f) **Les lois des estimateurs**

Avec un peu de calcul, il est possible de montrer que les lois de ces estimateurs sont les suivantes et qu'ils sont indépendants l'un par rapport à l'autre.

**Proposition 2.4.**

$$\hat{\mu} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{N}\right)$$

$$\hat{\alpha}_i \sim \mathcal{N}\left(\alpha_i, \sigma^2 \left(\frac{1}{n_i} - \frac{1}{N}\right)\right)$$

$$\frac{N\hat{\sigma}^2}{\sigma^2} \sim \chi_{(N-m)}^2$$

**Rappel 1.**  $Y \sim \chi_{(d)}^2$  indique que la variable aléatoire  $Y$  suit une loi du khi carré centrée avec  $d$  degrés de liberté.

$Y \sim \chi^{2l}(d, \lambda)$  indique que la variable aléatoire  $Y$  suit une loi du khi carré non centrée avec  $d$  degrés de liberté et un paramètre de décentralisation  $\lambda$ .

*Démonstration pour l'indépendance.* Il est possible de montrer l'indépendance entre ces trois estimateurs à l'aide des énoncés suivants.

$$\begin{aligned} & (y_{ij} - \bar{y}_{i.}) \text{ est indépendant de } \bar{y}_{i.} \\ \Rightarrow & (y_{ij} - \bar{y}_{i.}) \text{ est indépendant de } \bar{y}_{..} \\ \Rightarrow & (y_{ij} - \bar{y}_{i.}) \text{ est indépendant de } (\bar{y}_{i.} - \bar{y}_{..}) \end{aligned}$$

Et que

$(\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})$  est indépendant de  $\bar{y}_{\cdot\cdot}$ .

Donc  $\hat{\mu} = \bar{y}_{\cdot\cdot}$  est indépendant de  $\hat{\alpha}_i = (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})$  et  $\hat{\sigma}^2 = f(y_{ij} - \bar{y}_{i\cdot})$  est indépendant de  $\hat{\mu}$  et  $\hat{\alpha}_i$ .  $\square$

Puisque que  $\sigma^2$  est inconnu, il faut trouver d'autres lois pour  $\hat{\mu}$  et  $\hat{\alpha}_i$  qui ne dépendent pas de  $\sigma^2$ .

**Proposition 2.5.**

$$\frac{\hat{\mu} - \mu}{s/\sqrt{N}} \sim t_{(N-m)}$$

**Rappel 2.**  $Y \sim t_{(d)}$  indique que la variable aléatoire  $Y$  suit une loi de Student avec  $d$  degrés de liberté.

**Proposition 2.6.**

$$\frac{\hat{\alpha}_i - \alpha_i}{s\sqrt{\frac{1}{n_i} - \frac{1}{N}}} \sim t_{(N-m)}$$

Où  $s^2 = \frac{N}{N-m}\hat{\sigma}^2$ . Ces dernières lois découlent du fait que  $\hat{\alpha}_i$  et  $\hat{\mu}$  sont indépendants de  $\hat{\sigma}^2$  et de la loi de  $\hat{\sigma}^2$ .

**Rappel 3.**

$$\frac{Z}{\sqrt{U/k}} \sim t_{(k)}$$

Où  $Z \sim \mathcal{N}(0, 1)$  et  $U \sim \chi_{(k)}^2$  indépendantes.

Remarquons que  $\hat{\mu}$  est sans biais pour  $\mu$ , que  $\hat{\alpha}_i$  est sans biais pour  $\alpha_i$ , mais que  $\hat{\sigma}^2$  n'est pas sans biais pour  $\sigma^2$ . C'est pourquoi, plusieurs utilisent plutôt  $s^2$  qui lui est sans biais. En effet,

$$\begin{aligned}
& \mathbb{E} \left[ \frac{N\hat{\sigma}^2}{\sigma^2} \right] = N - m \\
\Rightarrow & \frac{N}{\sigma^2} \mathbb{E}[\hat{\sigma}^2] = N - m \\
\Rightarrow & \mathbb{E}[\hat{\sigma}^2] = \frac{N - m}{N} \sigma^2 \\
\Rightarrow & \mathbb{E}[s^2] = \mathbb{E} \left[ \frac{N}{N - m} \hat{\sigma}^2 \right] = \sigma^2
\end{aligned} \tag{2.3}$$

### g) Tests d'hypothèse et intervalles de confiance

Avec les lois trouvées précédemment, il est possible de dériver les tests d'hypothèse et les intervalles de confiance pour les trois paramètres. Un autre test qui peut être effectué ici est le test du rapport de vraisemblance pour l'hypothèse nulle  $H_0 : \alpha_i = 0 \forall i$ . En effet,

**Proposition 2.7.** *Il y a deux tests possibles pour tester  $H_0 : \alpha_i = 0 \forall i$ . Premièrement, il est possible de tester l'hypothèse nulle avec la statistique de Fisher suivante*

$$F = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y}_{..})^2 / (m - 1)}{\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 / (N - m)} = \frac{CME}{CMR} \sim \mathcal{F}_{N-m}^{m-1}$$

*Ou, deuxièmement, avec la statistique suivante qui représente le test du rapport de vraisemblance.*

$$-2 \log \Lambda = N \log \left( 1 + \frac{m - 1}{N - m} F \right) \approx \chi_{(m-1)}^2$$

**Rappel 4.**  $Y \sim \mathcal{F}_{d_2}^{d_1}$  indique que la variable aléatoire  $Y$  suit une loi de Fisher avec  $d_1$  (numérateur) et  $d_2$  (dénominateur) degrés de liberté.

*Démonstration.* Commençons par trouver la vraisemblance sous  $H_0$ .

$$L(\mu, \alpha_i = 0, \sigma^2) = (2\pi\sigma^2)^{-N/2} \exp \left[ -\frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu)^2 \right]$$

$$\Rightarrow l(\mu, \alpha_i = 0, \sigma^2) = -\frac{N}{2} \log(2\pi) - \frac{N}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu)^2$$

Ensuite, les équations du maximum de vraisemblance sont données par

$$\frac{\partial l}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu) = 0$$

$$\Rightarrow \sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij} = N\mu$$

$$\Rightarrow \hat{\mu}_0 = \bar{y}_{..}$$

$$\frac{\partial l}{\partial \sigma^2} = -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu)^2 = 0$$

$$\Rightarrow \frac{1}{2\sigma^4} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu)^2 = \frac{N}{2\sigma^2}$$

$$\Rightarrow \hat{\sigma}_0^2 = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu}_0)^2$$

$$\Rightarrow \hat{\sigma}_0^2 = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2$$

$$\Rightarrow \hat{\sigma}_0^2 = \frac{SCT}{N}$$

Donc calculons  $L(\hat{\theta}_0)$  et  $L(\hat{\theta})$ .

$$\begin{aligned}
L(\hat{\boldsymbol{\theta}}_0) &= (2\pi\hat{\sigma}_0^2)^{-N/2} \exp \left[ -\frac{1}{2\hat{\sigma}_0^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu}_0)^2 \right] \\
&= (2\pi\hat{\sigma}_0^2)^{-N/2} \exp \left( \frac{-N}{2} \right)
\end{aligned}$$

$$\begin{aligned}
L(\hat{\boldsymbol{\theta}}) &= (2\pi\hat{\sigma}^2)^{-N/2} \exp \left[ -\frac{1}{2\hat{\sigma}^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu} - \hat{\alpha}_i)^2 \right] \\
&= (2\pi\hat{\sigma}^2)^{-N/2} \exp \left( \frac{-N}{2} \right)
\end{aligned}$$

Donc le rapport des vraisemblances est donné par

$$\Lambda = \frac{L(\hat{\boldsymbol{\theta}}_0)}{L(\hat{\boldsymbol{\theta}})} = \left( \frac{\hat{\sigma}_0^2}{\hat{\sigma}^2} \right)^{-N/2} = \left( \frac{SCT}{SCR} \right)^{-N/2}$$

Ensuite, en prenant le logarithme naturel et en multipliant par  $-2$ , l'expression devient

$$-2 \log \Lambda = N \log \left( 1 + \frac{m-1}{N-m} F \right) \approx \chi_{(m-1)}^2$$

Où  $F = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y}_{..})^2 / (m-1)}{\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 / (N-m)} = \frac{CME}{CMR} \sim \mathcal{F}_{N-m}^{m-1}$  sous  $H_0$ . Ceci découle de la définition

de la loi de Fisher, du fait que  $\frac{SCR}{\sigma^2} \sim \chi_{(N-m)}^2$ , sous  $H_0$ ,  $\frac{SCE}{\sigma^2} \sim \chi_{(m-1)}^2$  et qu'elles sont indépendantes (voir la section B.1 des annexes et le tableau C.1 dans les annexes). De plus, les deux statistiques sont équivalentes puisqu'elles varient dans le même sens. En effet

$$\begin{aligned}
\frac{\partial}{\partial F} N \log \left( 1 + \frac{m-1}{N-m} F \right) &= \frac{N}{1 + \frac{m-1}{N-m} F} \frac{m-1}{N-m} \\
&= \frac{N \text{ SCR } m-1}{\text{SCT } N-m} \\
&= \frac{N(m-1) \text{ SCR}}{(N-m) \text{ SCT}} \\
&> 0
\end{aligned}$$

Nous avons donc deux statistiques équivalentes pour le même test ( $H_0 : \alpha_i = 0 \forall i$ ), mais une est exacte et l'autre est approximative.  $\square$

**Rappel 5.**

$$\frac{U_1/d_1}{U_2/d_2} \sim \mathcal{F}_{d_1}^{d_2}$$

Où  $U_1 \sim \chi_{(d_1)}^2$  et  $U_2 \sim \chi_{(d_2)}^2$  indépendantes.

**Rappel 6.** [Wil08] et [MGB73] ont démontré que la loi asymptotique de  $-2 \log(\Lambda)$  est la loi du khi carré avec comme degrés de liberté, la différence entre le nombre de paramètres estimés dans le modèle complet et le nombre de paramètres estimés sous  $H_0$ .

## 2.2.2 Modèle avec un effet aléatoire (ANOVA aléatoire)

Présentons maintenant la théorie du modèle de l'analyse de la variance à un facteur aléatoire.

### a) Le modèle

Le modèle est traditionnellement écrit comme suit.

$$\begin{aligned}\mathbb{E}[y_{ij}|a_i] &= \mu + a_i \\ y_{ij}|a_i &\sim \mathcal{N}(\mu + a_i, \sigma^2) \text{ indépendantes} \\ a_i &\sim \text{indépendantes et identiquement distribuées (i.i.d) } \mathcal{N}(0, \sigma_a^2)\end{aligned}\tag{2.4}$$

Où, encore,  $i = 1, \dots, m$  représente le  $i^e$  groupe et où  $j = 1, \dots, n_i$  représente la  $j^e$  observation du  $i^e$  groupe. Ainsi,  $\mu$  représente la moyenne globale et  $\mu + a_i$  représente la moyenne du groupe  $i$ . Aussi, nous noterons encore  $N = \sum_{i=1}^m n_i$ .

Il en découle que  $\mathbb{E}[y_{ij}] = \mathbb{E}[\mathbb{E}[y_{ij}|a_i]] = \mathbb{E}[\mu + a_i] = \mu$ .

## b) Les covariances entre les observations

Rappelons tout d'abord deux propriétés des variances et covariances conditionnelles.

### Proposition 2.8.

$$\text{Var}(y) = \text{Var}_u(\mathbb{E}[y|u]) + \mathbb{E}_u[\text{Var}(y|u)]$$

### Proposition 2.9.

$$\text{cov}(y, w) = \text{cov}_u(\mathbb{E}[y|u], \mathbb{E}[w|u]) + \mathbb{E}_u[\text{cov}(y, w|u)]$$

Les preuves se retrouvent à la section B.2 des annexes.

Les variances et les covariances sont donc données par

### Proposition 2.10. $\text{Var}(y_{ij}) = \sigma_a^2 + \sigma^2$

*Démonstration.*

$$\begin{aligned}\text{Var}(y_{ij}) &= \text{Var}(\mathbb{E}[y_{ij}|a_i]) + \mathbb{E}[\text{Var}(y_{ij}|a_i)] \\ &= \text{Var}(\mu + a_i) + \mathbb{E}[\sigma^2] \\ &= \sigma_a^2 + \sigma^2\end{aligned}\quad \square$$

Et,

**Proposition 2.11.**  $\text{cov}(y_{ij}, y_{il}) = \sigma_a^2$  avec  $j \neq l$

*Démonstration.*

$$\begin{aligned} \text{cov}(y_{ij}, y_{il}) &= \text{cov}(\mathbb{E}[y_{ij}|a_i], \mathbb{E}[y_{il}|a_i]) + \mathbb{E}[\text{cov}(y_{ij}, y_{il}|a_i)] \\ &= \text{cov}(\mu + a_i, \mu + a_i) + 0 \\ &= \text{cov}(a_i, a_i) \\ &= \sigma_a^2 \end{aligned}$$

□

### c) Les sommes de carrés

Voici les sommes de carrés qui interviennent dans ce modèle.

Notation	Expression	Carré moyen	Signification
$SCE$	$\sum_{i=1}^m \sum_{j=1}^{n_i} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2$	$CME = \frac{SCE}{m-1}$	Carrés expliqués
$SCR$	$\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})^2$	$CMR = \frac{SCR}{N-m}$	Carrés résiduels
$SCT$	$\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{\cdot\cdot})^2$	—	Carrés totaux

Tableau 2.4 – Tableau des sommes de carrés, ANOVA avec un effet aléatoire

Avec  $SCT = SCR + SCE$ .

### d) La vraisemblance

La vraisemblance n'est pas aussi simple que dans le cas avec un effet fixe, puisque les observations ne sont pas toutes indépendantes. Par contre,



**Proposition 2.12.**

$$\mathbf{y}_i \sim \mathcal{N}(\mu \mathbf{1}_{n_i}, \mathbf{V}_i) \quad \text{indépendants}$$

Où  $\mathbf{y}_i = [y_{i1}, y_{i2}, \dots, y_{in_i}]^t$  et  $\mathbf{V}_i = \sigma^2 \mathbf{I}_{n_i} + \sigma_a^2 \mathbf{J}_{n_i}$ , avec  $\mathbf{I}_{n_i}$ , la matrice identité d'ordre  $n_i$ ,  $\mathbf{J}_{n_i}$ , la matrice carrée d'ordre  $n_i$  contenant que des 1 et  $\mathbf{1}_{n_i}$ , le vecteur de longueur  $n_i$  contenant que des 1.

*Démonstration.* Premièrement, les  $\mathbf{y}_i$  sont indépendants par construction. En effet, pour  $i \neq k$

$$\begin{aligned} \text{cov}(y_{ij}, y_{kl}) &= \text{cov}(\mathbb{E}[y_{ij}|a_i, a_k], \mathbb{E}[y_{kl}|a_i, a_k]) + \mathbb{E}[\text{cov}(y_{ij}, y_{kl}|a_i, a_k)] \\ &= \text{cov}(\mu + a_i, \mu + a_k) + 0 \\ &= \text{cov}(a_i, a_k) \\ &= 0 \end{aligned}$$

Puisque les  $y_{ij}$  sont de loi normale, ils sont indépendants. Et les paramètres de la loi normale proviennent des calculs faits plus haut sur la variance et la covariance des  $y_{ij}$ .  $\square$

Voici deux propositions sur  $\mathbf{V}_i$ .

**Proposition 2.13.**

$$\mathbf{V}_i^{-1} = \frac{1}{\sigma^2} \mathbf{I}_{n_i} - \frac{\sigma_a^2}{\sigma^2(\sigma^2 + n_i \sigma_a^2)} \mathbf{J}_{n_i}$$

Et

**Proposition 2.14.**

$$|\mathbf{V}_i| = (\sigma^2 + n_i \sigma_a^2) (\sigma^2)^{n_i - 1}$$

Les preuves sont directes en utilisant les propriétés de la section A.1 des annexes sur les vecteurs et matrices de uns.

Voici maintenant la vraisemblance.

**Proposition 2.15.**

$$L = \prod_{i=1}^m \left[ (2\pi)^{-n_i/2} (\sigma^2 + n_i \sigma_a^2)^{-1/2} (\sigma^2)^{-(n_i-1)/2} \right] \\ \cdot \prod_{i=1}^m \left[ \exp \left( -\frac{1}{2} (\mathbf{y}_i - \mu \mathbf{1}_{n_i})^t \left( \frac{1}{\sigma^2} \mathbf{I}_{n_i} - \frac{\sigma_a^2}{\sigma^2 (\sigma^2 + n_i \sigma_a^2)} \mathbf{J}_{n_i} \right) (\mathbf{y}_i - \mu \mathbf{1}_{n_i}) \right) \right]$$

Et donc,

$$l = \log L = -\frac{N}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^m \log(\sigma^2 + n_i \sigma_a^2) - \frac{N-m}{2} \log(\sigma^2) \\ - \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu)^2 + \frac{\sigma_a^2}{2\sigma^2} \sum_{i=1}^m \frac{1}{\sigma^2 + n_i \sigma_a^2} (y_{i\cdot} - n_i \mu)^2$$

*Démonstration.*

$$L = \prod_{i=1}^m (2\pi)^{-n_i/2} |\mathbf{V}_i|^{-1/2} \exp \left( -\frac{1}{2} (\mathbf{y}_i - \mu \mathbf{1}_{n_i})^t \mathbf{V}_i^{-1} (\mathbf{y}_i - \mu \mathbf{1}_{n_i}) \right) \\ = \prod_{i=1}^m \left[ (2\pi)^{-n_i/2} (\sigma^2 + n_i \sigma_a^2)^{-1/2} (\sigma^2)^{-(n_i-1)/2} \right] \\ \cdot \prod_{i=1}^m \left[ \exp \left( -\frac{1}{2} (\mathbf{y}_i - \mu \mathbf{1}_{n_i})^t \left( \frac{1}{\sigma^2} \mathbf{I}_{n_i} - \frac{\sigma_a^2}{\sigma^2 (\sigma^2 + n_i \sigma_a^2)} \mathbf{J}_{n_i} \right) (\mathbf{y}_i - \mu \mathbf{1}_{n_i}) \right) \right]$$

Et donc,

$$l = \log L = \sum_{i=1}^m \log \left[ (2\pi)^{-n_i/2} (\sigma^2 + n_i \sigma_a^2)^{-1/2} (\sigma^2)^{-(n_i-1)/2} \exp(\star) \right] \\ = \sum_{i=1}^m \left[ -\frac{n_i}{2} \log(2\pi) - \frac{1}{2} \log(\sigma^2 + n_i \sigma_a^2) - \frac{n_i-1}{2} \log(\sigma^2) \right] \\ + \sum_{i=1}^m \left[ -\frac{1}{2\sigma^2} (\mathbf{y}_i - \mu \mathbf{1}_{n_i})^t (\mathbf{y}_i - \mu \mathbf{1}_{n_i}) + \frac{\sigma_a^2}{2\sigma^2 (\sigma^2 + n_i \sigma_a^2)} (\mathbf{y}_i - \mu \mathbf{1}_{n_i})^t \mathbf{J}_{n_i} (\mathbf{y}_i - \mu \mathbf{1}_{n_i}) \right]$$

$$\begin{aligned}
&= -\frac{N}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^m \log(\sigma^2 + n_i \sigma_a^2) - \frac{N-m}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu)^2 \\
&\quad + \frac{\sigma_a^2}{2\sigma^2} \sum_{i=1}^m \frac{1}{\sigma^2 + n_i \sigma_a^2} \left[ \sum_{j=1}^{n_i} (y_{ij} - \mu), \dots, \sum_{j=1}^{n_i} (y_{ij} - \mu) \right] (\mathbf{y}_i - \mu \mathbf{1}_{n_i}) \\
&= -\frac{N}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^m \log(\sigma^2 + n_i \sigma_a^2) - \frac{N-m}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu)^2 \\
&\quad + \frac{\sigma_a^2}{2\sigma^2} \sum_{i=1}^m \frac{1}{\sigma^2 + n_i \sigma_a^2} (y_{i\cdot} - n_i \mu) \mathbf{1}_{n_i}^t (\mathbf{y}_i - \mu \mathbf{1}_{n_i}) \\
&= -\frac{N}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^m \log(\sigma^2 + n_i \sigma_a^2) - \frac{N-m}{2} \log(\sigma^2) \\
&\quad - \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu)^2 + \frac{\sigma_a^2}{2\sigma^2} \sum_{i=1}^m \frac{1}{\sigma^2 + n_i \sigma_a^2} (y_{i\cdot} - n_i \mu)^2
\end{aligned}$$

Où  $\star = -\frac{1}{2} (\mathbf{y}_i - \mu \mathbf{1}_{n_i})^t \left( \frac{-\sigma_a^2}{\sigma^2(\sigma^2 + n_i \sigma_a^2)} \mathbf{J}_{n_i} + \frac{1}{\sigma^2} \mathbf{I}_{n_i} \right) (\mathbf{y}_i - \mu \mathbf{1}_{n_i})$  □

e) **Estimation dans le cas où les données sont équilibrées**

### La vraisemblance

La vraisemblance est beaucoup plus belle dans le cas où les données sont équilibrées (« balanced data » en anglais). Par données équilibrées, nous parlons du cas où  $n_i = n \forall i$ . Cela simplifie beaucoup le calcul des solutions du maximum de vraisemblance. En effet, la log-vraisemblance devient

**Proposition 2.16.**  $l = -\frac{N}{2} \log(2\pi) - \frac{m}{2} \log(\lambda) - \frac{m(n-1)}{2} \log(\sigma^2) - \frac{SCR}{2\sigma^2} - \frac{SCE}{2\lambda} - mn \frac{(\bar{y}_{\cdot} - \mu)^2}{2\lambda}$

Où  $\lambda = \sigma^2 + n\sigma_a^2$ .

*Démonstration.*

$$\begin{aligned}
l = \log L &= -\frac{N}{2} \log(2\pi) - \frac{m}{2} \log(\sigma^2 + n\sigma_a^2) - \frac{mn - m}{2} \log(\sigma^2) \\
&\quad - \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \mu)^2 + \frac{\sigma_a^2}{2\sigma^2(\sigma^2 + n\sigma_a^2)} \sum_{i=1}^m (n\bar{y}_i - n\mu)^2 \\
&= -\frac{N}{2} \log(2\pi) - \frac{m}{2} \log(\sigma^2 + n\sigma_a^2) - \frac{m(n-1)}{2} \log(\sigma^2) \\
&\quad - \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \mu)^2 + \frac{n^2\sigma_a^2}{2\sigma^2(\sigma^2 + n\sigma_a^2)} \sum_{i=1}^m (\bar{y}_i - \mu)^2 \\
&= -\frac{N}{2} \log(2\pi) - \frac{m}{2} \log(\lambda) - \frac{m(n-1)}{2} \log(\sigma^2) \\
&\quad - \frac{SCR}{2\sigma^2} - \frac{SCE}{2\lambda} - mn \frac{(\bar{y}_.. - \mu)^2}{2\lambda}
\end{aligned}$$

La dernière équation provient du fait que

$$\begin{aligned}
\sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \mu)^2 &= \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2 + \sum_{i=1}^m \sum_{j=1}^n (\bar{y}_i - \mu)^2 \\
&= SCR + \sum_{i=1}^m \sum_{j=1}^n (\bar{y}_i - \mu)^2
\end{aligned}$$

et

$$\begin{aligned}
\sum_{i=1}^m \sum_{j=1}^n (\bar{y}_i - \mu)^2 &= \sum_{i=1}^m \sum_{j=1}^n (\bar{y}_i - \bar{y}_..)^2 + \sum_{i=1}^m \sum_{j=1}^n (\bar{y}_.. - \mu)^2 \\
&= SCE + mn(\bar{y}_.. - \mu)^2
\end{aligned}$$

□

## Les équations du maximum de vraisemblance et leurs solutions

**Proposition 2.17.** *Les solutions du maximum de vraisemblance sont données par*

$$\hat{\mu} = \bar{y}_..$$

$$\dot{\sigma}^2 = \frac{SCR}{m(n-1)}$$

$$\dot{\sigma}_a^2 = \frac{(1-1/m)CME - CMR}{n}$$

*Démonstration.* Les dérivées partielles par rapport à chaque variable ( $\mu$ ,  $\sigma^2$  et  $\lambda$ ) sont les suivantes.

$$l_\mu = \frac{\partial l}{\partial \mu} = \frac{mn(\bar{y}_. - \mu)}{\lambda}$$

$$l_{\sigma^2} = \frac{\partial l}{\partial \sigma^2} = \frac{-m(n-1)}{2\sigma^2} + \frac{SCR}{2\sigma^4} = \frac{-m(n-1)}{2\sigma^4} \left( \sigma^2 - \frac{SCR}{m(n-1)} \right)$$

$$l_\lambda = \frac{\partial l}{\partial \lambda} = \frac{-m}{2\lambda} + \frac{SCE}{2\lambda^2} + \frac{mn(\bar{y}_. - \mu)^2}{2\lambda^2} = \frac{-m}{2\lambda^2} \left( \lambda - \frac{SCE}{m} \right) + \frac{mn(\bar{y}_. - \mu)^2}{2\lambda^2}$$

En égalant ces dernières dérivées à 0, les solutions sont les suivantes.

$$\dot{\mu} = \bar{y}_.$$

$$\dot{\sigma}^2 = CMR = \frac{SCR}{m(n-1)}$$

$$\dot{\lambda} = \frac{SCE}{m} = \left(1 - \frac{1}{m}\right) CME$$

Il en découle que

$$\dot{\sigma}_a^2 = \frac{\dot{\lambda} - \dot{\sigma}^2}{n} = \frac{SCE/m - SCR/m(n-1)}{n} = \frac{(1-1/m)CME - CMR}{n} \quad \square$$

### Les estimateurs du maximum de vraisemblance

Puisque le domaine possible pour  $\mu$  est les réels et que sa solution ( $\dot{\mu}$ ) ne dépend pas de  $\dot{\sigma}^2$  ou  $\dot{\sigma}_a^2$ ,  $\hat{\mu} = \dot{\mu} = \bar{y}_. \in \mathbb{R}$ . Pour l'estimateur de  $\sigma^2$ , il faut qu'il soit dans les réels positifs. Puisque  $\dot{\sigma}^2 = CMR > 0$ , à moins que  $y_{ij} = \bar{y}_. \forall i$  ce qui est peu probable et

qui n'est pas intéressant, il est raisonnable de dire qu'il est dans le bon domaine. Par contre, puisque  $\dot{\sigma}_a^2$  dépend de  $\dot{\sigma}^2$ , ils doivent être étudiés ensemble. C'est-à-dire que la paire d'estimateur  $(\hat{\sigma}^2, \hat{\sigma}_a^2)$  doit être dans l'espace de dimension deux possible de la paire de paramètres  $(\sigma^2, \sigma_a^2)$ . Lorsque  $\dot{\sigma}_a^2 \geq 0$ , il n'y a pas de problème, les solutions sont également les estimateurs. Lorsque  $\dot{\sigma}_a^2 < 0$ , l'estimateur de  $\sigma_a^2$  est posé à 0, car il doit être dans les réels non négatifs. Ensuite, pour trouver l'estimateur de  $\sigma^2$  dans cette situation, il y a la proposition suivante.

**Proposition 2.18.** *La solution du maximum de vraisemblance pour  $\sigma^2$  lorsque  $\sigma_a^2 = 0$  est donnée par  $\dot{\sigma}_0^2 = \frac{SCT}{mn}$*

*Démonstration.* Il suffit de remplacer  $\sigma_a^2$  dans la formule de la vraisemblance par 0 et de trouver une nouvelle solution du maximum de vraisemblance pour  $\sigma^2$ . Le logarithme de la vraisemblance devient

$$\begin{aligned} l(\mu, \sigma^2, \sigma_a^2 = 0) &= -\frac{N}{2} \log(2\pi) - \frac{m}{2} \log(\sigma^2) - \frac{mn - m}{2} \log(\sigma^2) - \frac{SCR}{2\sigma^2} - \frac{SCE}{2\sigma^2} - \frac{mn(\bar{y}_{..} - \mu)^2}{2\sigma^2} \\ &= -\frac{N}{2} \log(2\pi) - \frac{mn}{2} \log(\sigma^2) - \frac{SCT}{2\sigma^2} - \frac{mn(\bar{y}_{..} - \mu)^2}{2\sigma^2} \end{aligned}$$

Donc, il en découle que

$$\begin{aligned} \frac{\partial l}{\partial \mu} &= \frac{mn(\bar{y}_{..} - \mu)}{\sigma^2} = 0 \\ &\Rightarrow \dot{\mu}_0 = \dot{\mu} = \bar{y}_{..} \end{aligned}$$

$$\begin{aligned}
\frac{\partial l}{\partial \sigma^2} &= -\frac{mn}{2\sigma^2} + \frac{SCT}{2\sigma^4} + \frac{mn(\bar{y}_{..} - \mu)^2}{2\sigma^4} = 0 \\
&\Rightarrow \frac{SCT}{2\sigma_0^4} = \frac{mn}{2\sigma_0^2} \\
&\Rightarrow \hat{\sigma}_0^2 = \frac{SCT}{mn}
\end{aligned}$$

□

En résumé, les estimateurs sont les suivants.

$$\begin{aligned}
\hat{\mu} &= \bar{y}_{..} \\
\hat{\sigma}^2 &= \begin{cases} CMR & \text{si } \left(1 - \frac{1}{m}\right) CME \geq CMR \\ \frac{SCT}{mn} & \text{si } \left(1 - \frac{1}{m}\right) CME < CMR \end{cases} \\
\hat{\sigma}_a^2 &= \begin{cases} \frac{(1-1/m)CME - CMR}{n} & \text{si } \left(1 - \frac{1}{m}\right) CME \geq CMR \\ 0 & \text{si } \left(1 - \frac{1}{m}\right) CME < CMR \end{cases}
\end{aligned} \tag{2.5}$$

### Espérance et biais des estimateurs

Regardons maintenant si ces estimateurs sont sans biais.

**Proposition 2.19.**  $\mathbb{E}[\hat{\mu}] = \mu$

*Démonstration.*

$$\mathbb{E}[\hat{\mu}] = \mathbb{E}[\bar{y}_{..}] = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^n \mathbb{E}[y_{ij}] = \frac{N\mu}{N} = \mu$$

□

**Proposition 2.20.**  $\mathbb{E}[\hat{\sigma}^2] = \sigma^2$

*Démonstration.*

$$\mathbb{E}[\hat{\sigma}^2] = \mathbb{E}[CMR] = \sigma^2$$

Les détails pour  $\mathbb{E}[CMR]$  se retrouvent à la section B.3.1 des annexes.

□

**Proposition 2.21.**  $\mathbb{E}[\hat{\sigma}_a^2] = \left(1 - \frac{1}{m}\right) \sigma_a^2 - \frac{\sigma^2}{mn}$

*Démonstration.*

$$\begin{aligned}
\mathbb{E}[\hat{\sigma}_a^2] &= \frac{1}{n} \mathbb{E} \left[ \left(1 - \frac{1}{m}\right) CME - CMR \right] \\
&= \frac{1}{n} \left(1 - \frac{1}{m}\right) \mathbb{E}[CME] - \frac{1}{n} \mathbb{E}[CMR] \\
&= \frac{\left(1 - \frac{1}{m}\right) (\sigma^2 + n\sigma_a^2) - \sigma^2}{n} \\
&= \left(1 - \frac{1}{m}\right) \sigma_a^2 - \frac{\sigma^2}{mn}
\end{aligned}$$

Les détails pour  $\mathbb{E}[CME]$  se retrouvent à la section B.3.1 des annexes.  $\square$

Donc l'estimateur de  $\mu$  est sans biais, mais pour les deux autres, les calculs précédents représentent l'espérance des solutions et non des estimateurs. Les espérances des estimateurs s'écrivent comme suit.

$$\begin{aligned}
\mathbb{E}[\hat{\sigma}^2] &= (1-p) \mathbb{E}[CMR | \hat{\sigma}_a^2 \geq 0] + p \mathbb{E} \left[ \frac{SCT}{mn} | \hat{\sigma}_a^2 < 0 \right] \\
\mathbb{E}[\hat{\sigma}_a^2] &= (1-p) \mathbb{E}[\hat{\sigma}_a^2 | \hat{\sigma}_a^2 \geq 0]
\end{aligned} \tag{2.6}$$

avec

$$\begin{aligned}
p &= \mathbb{P} \left\{ \left(1 - \frac{1}{m}\right) CME < CMR \right\} \\
&= \mathbb{P} \left\{ \mathcal{F}_{m-1}^{m(n-1)} > \left(1 - \frac{1}{m}\right) \left(1 + \frac{n\sigma_a^2}{\sigma^2}\right) \right\}
\end{aligned} \tag{2.7}$$

Car  $\frac{SCR}{\sigma^2} \sim \chi_{(m(n-1))}^2$  et  $\frac{SCE}{n\sigma_a^2 + \sigma^2} \sim \chi_{(m-1)}^2$  et ils sont indépendants (voir la section B.1 des annexes et le tableau C.1 dans les annexes) donc

$$\frac{\frac{SCR/\sigma^2}{m(n-1)}}{\frac{SCE/(n\sigma_a^2 + \sigma^2)}{m-1}} = \frac{CMR/\sigma^2}{CME/(n\sigma_a^2 + \sigma^2)} \sim \mathcal{F}_{m-1}^{m(n-1)} \tag{2.8}$$



### Variance des estimateurs (et leur loi)

Pour  $\hat{\mu}$ , la variance (exacte) peut être calculée, en effet,

**Proposition 2.22.**  $Var(\hat{\mu}) = Var(\bar{y}_{..}) = \frac{\sigma^2 + n\sigma_a^2}{mn}$

*Démonstration.* Voir la section B.3.1 des annexes. □

Donc la loi de  $\hat{\mu}$  est donnée par

$$\hat{\mu} \sim \mathcal{N}\left(\mu, \frac{\sigma^2 + n\sigma_a^2}{mn}\right) \quad (2.9)$$

Pour les autres estimateurs ( $\sigma^2$  et  $\sigma_a^2$ ) la variance exacte n'a pas de forme explicite comme leur espérance. Par contre, il est possible de calculer leur variance asymptotique. Il y a un résultat concernant la loi asymptotique des estimateurs du maximum de vraisemblance qui dit que (présenté pour deux estimateurs, mais se généralise facilement pour plus)

**Théorème 2.1** (Loi asymptotique des estimateurs du maximum de vraisemblance).

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} \approx \mathcal{N}\left(\begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \left(-\mathbb{E}\begin{pmatrix} l_{\alpha\alpha} & l_{\alpha\beta} \\ l_{\alpha\beta} & l_{\beta\beta} \end{pmatrix}\right)^{-1}\right)$$

Où  $\hat{\alpha}$  et  $\hat{\beta}$  sont les estimateurs du maximum de vraisemblance de  $\alpha$  et  $\beta$  respectivement. Aussi,  $l_{\alpha\alpha}$ ,  $l_{\alpha\beta}$  et  $l_{\beta\beta}$  sont les dérivées secondes du logarithme de la vraisemblance par rapport aux variables en indice.

Il en découle que

**Proposition 2.23.**

$$\begin{bmatrix} \hat{\sigma}^2 \\ \hat{\sigma}_a^2 \end{bmatrix} \approx \mathcal{N}\left(\begin{bmatrix} \sigma^2 \\ \sigma_a^2 \end{bmatrix}, 2\sigma^4 \begin{pmatrix} \frac{1}{m(n-1)} & \frac{-1}{mn(n-1)} \\ \frac{-1}{mn(n-1)} & \frac{1}{n^2} \left(\frac{\lambda^2}{m\sigma^4} + \frac{1}{m(n-1)}\right) \end{pmatrix}\right)$$

*Démonstration.* Le théorème 2.1 dit que

$$\text{Var} \begin{bmatrix} \hat{\sigma}^2 \\ \hat{\lambda} \end{bmatrix} \approx \left( -\mathbb{E} \begin{pmatrix} l_{\sigma^2\sigma^2} & l_{\sigma^2\lambda} \\ l_{\sigma^2\lambda} & l_{\lambda\lambda} \end{pmatrix} \right)^{-1}$$

Où « Var » qui désigne la matrice de variance-covariance pour un vecteur de variable aléatoire ou pour une variable aléatoire multidimensionnelle. Donc, en dérivant encore  $l_\lambda$  et  $l_{\sigma^2}$ , il s'ensuit que

$$l_{\sigma^2\sigma^2} = \frac{\partial^2 l}{(\partial\sigma^2)^2} = \frac{m(n-1)}{2\sigma^4} - \frac{SCR}{\sigma^6}$$

$$l_{\sigma^2\lambda} = \frac{\partial^2 l}{\partial\lambda\partial\sigma^2} = 0$$

$$l_{\lambda\lambda} = \frac{\partial^2 l}{(\partial\lambda)^2} = \frac{m}{2\lambda^2} - \frac{SCE}{\lambda^3} - \frac{mn(\bar{y}_.. - \mu)^2}{\lambda^3}$$

Ensuite, en prenant le négatif de l'espérance de ces dérivées secondes, il s'ensuit que

$$-\mathbb{E}[l_{\sigma^2\sigma^2}] = \frac{-m(n-1)}{2\sigma^4} + \frac{2m(n-1)\sigma^2}{2\sigma^6} = \frac{m(n-1)}{2\sigma^4}$$

$$-\mathbb{E}[l_{\sigma^2\lambda}] = 0$$

$$\begin{aligned} -\mathbb{E}[l_{\lambda\lambda}] &= \frac{-m}{2\lambda^2} + \frac{2(m-1)\lambda}{2\lambda^3} + \frac{mn\mathbb{E}[\bar{y}_..^2 - 2\bar{y}_..\mu]}{\lambda^3} + \frac{mn\mu^2}{\lambda^3} \\ &= \frac{-m}{2\lambda^2} + \frac{2(m-1)}{2\lambda^2} + \frac{mn}{\lambda^3} \left( \mu^2 + \frac{\lambda}{mn} - 2\mu^2 + \mu^2 \right) \\ &= \frac{m-2}{2\lambda^2} + \frac{2}{2\lambda^2} \\ &= \frac{m}{2\lambda^2} \end{aligned}$$

Car  $\mathbb{E}[\bar{y}_{..}^2] = \mu^2 + \frac{\lambda}{mn}$ , voir la section B.3.1 des annexes.

Avec tous ces résultats, il est possible de calculer la matrice de variance-covariance asymptotique de chaque estimateur.

$$\text{Var} \begin{bmatrix} \hat{\sigma}^2 \\ \hat{\lambda} \end{bmatrix} \approx \begin{pmatrix} \frac{m(n-1)}{2\sigma^4} & 0 \\ 0 & \frac{m}{2\lambda^2} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{2\sigma^4}{m(n-1)} & 0 \\ 0 & \frac{2\lambda^2}{m} \end{pmatrix}$$

Ensuite, puisque  $\hat{\sigma}_a^2 = (\hat{\lambda} - \hat{\sigma}^2)/n$ , la matrice de variance-covariance asymptotique des estimateurs du maximum de vraisemblance des variances est la suivante.

$$\text{Var} \begin{bmatrix} \hat{\sigma}^2 \\ \hat{\sigma}_a^2 \end{bmatrix} \approx 2\sigma^4 \begin{pmatrix} \frac{1}{m(n-1)} & \frac{-1}{mn(n-1)} \\ \frac{-1}{mn(n-1)} & \frac{1}{n^2} \left( \frac{\lambda^2}{m\sigma^4} + \frac{1}{m(n-1)} \right) \end{pmatrix} \quad \square$$

## f) Estimation dans le cas où les données sont non équilibrées

### La vraisemblance

Procédons de façon semblable en posant  $\lambda_i = \sigma^2 + n_i\sigma_a^2$ . La vraisemblance est donc donnée par (après quelques manipulations)

**Proposition 2.24.**  $l = -\frac{1}{2}N \log(2\pi) - \frac{1}{2} \sum_{i=1}^m \log(\lambda_i) - \frac{1}{2}(N-m) \log(\sigma^2) - \frac{SCR}{2\sigma^2} - \sum_{i=1}^m \frac{n_i(\bar{y}_{i.} - \mu)^2}{2\lambda_i}$

### Les équations du maximum de vraisemblance et leurs solutions

**Proposition 2.25.** *Les solutions du maximum de vraisemblance ne sont pas toutes explicites, mais voici les équations les plus simplifiées*

$$\dot{\mu} = \frac{\sum_{i=1}^m \frac{n_i \bar{y}_{i.}}{\dot{\sigma}^2 + n_i \dot{\sigma}_a^2}}{\sum_{i=1}^m \frac{n_i}{\dot{\sigma}^2 + n_i \dot{\sigma}_a^2}}$$

$$\frac{SCR}{\dot{\sigma}^4} - \frac{N-m}{\dot{\sigma}_2} + \sum_{i=1}^m \frac{n_i(\bar{y}_i - \dot{\mu})^2}{(\dot{\sigma}^2 + n_i\dot{\sigma}_a^2)^2} - \sum_{i=1}^m \frac{1}{\dot{\sigma}^2 + n_i\dot{\sigma}_a^2} = 0$$

$$\sum_{i=1}^m \frac{n_i^2(\bar{y}_i - \dot{\mu})^2}{(\dot{\sigma}^2 + n_i\dot{\sigma}_a^2)^2} = \sum_{i=1}^m \frac{n_i}{\dot{\sigma}^2 + n_i\dot{\sigma}_a^2}$$

*Démonstration.* Trouvons premièrement les dérivées partielles pour chaque paramètre.

$$l_\mu = \frac{\partial l}{\partial \mu} = \sum_{i=1}^m \frac{n_i(\bar{y}_i - \mu)}{\lambda_i}$$

$$\begin{aligned} l_{\sigma^2} &= \frac{\partial l}{\partial \sigma^2} = \frac{-(N-m)}{2\sigma^2} + \frac{SCR}{2\sigma^4} - \frac{1}{2} \sum_{i=1}^m \frac{\partial}{\partial \lambda_i} (\log(\lambda_i)) \frac{\partial \lambda_i}{\partial \sigma^2} - \sum_{i=1}^m \frac{\partial}{\partial \lambda_i} \left( \frac{n_i(\bar{y}_i - \mu)^2}{2\lambda_i} \right) \frac{\partial \lambda_i}{\partial \sigma^2} \\ &= \frac{-(N-m)}{2\sigma^2} + \frac{SCR}{2\sigma^4} - \frac{1}{2} \sum_{i=1}^m \left( \frac{1}{\lambda_i} \cdot 1 \right) + \sum_{i=1}^m \left( \frac{n_i(\bar{y}_i - \mu)^2}{2\lambda_i^2} \cdot 1 \right) \\ &= \frac{-(N-m)}{2\sigma^2} + \frac{SCR}{2\sigma^4} - \frac{1}{2} \sum_{i=1}^m \frac{1}{\lambda_i} + \sum_{i=1}^m \frac{n_i(\bar{y}_i - \mu)^2}{2\lambda_i^2} \end{aligned}$$

$$\begin{aligned} l_{\sigma_a^2} &= \frac{\partial l}{\partial \sigma_a^2} = -\frac{1}{2} \sum_{i=1}^m \frac{\partial}{\partial \lambda_i} (\log(\lambda_i)) \frac{\partial \lambda_i}{\partial \sigma_a^2} - \sum_{i=1}^m \frac{\partial}{\partial \lambda_i} \left( \frac{n_i(\bar{y}_i - \mu)^2}{2\lambda_i} \right) \frac{\partial \lambda_i}{\partial \sigma_a^2} \\ &= -\frac{1}{2} \sum_{i=1}^m \left( \frac{1}{\lambda_i} \cdot n_i \right) + \sum_{i=1}^m \left( \frac{n_i(\bar{y}_i - \mu)^2}{2\lambda_i^2} \cdot n_i \right) \\ &= -\frac{1}{2} \sum_{i=1}^m \frac{n_i}{\lambda_i} + \sum_{i=1}^m \frac{n_i^2(\bar{y}_i - \mu)^2}{2\lambda_i^2} \end{aligned}$$

En égalant ces équations à 0, il s'ensuit que

$$\dot{\mu} = \frac{\sum_{i=1}^m \frac{n_i \bar{y}_i}{\dot{\lambda}_i}}{\sum_{i=1}^m \frac{n_i}{\dot{\lambda}_i}} = \frac{\sum_{i=1}^m \frac{n_i \bar{y}_i}{\dot{\sigma}^2 + n_i \dot{\sigma}_a^2}}{\sum_{i=1}^m \frac{n_i}{\dot{\sigma}^2 + n_i \dot{\sigma}_a^2}} = \frac{\sum_{i=1}^m [\bar{y}_i / \text{Var}(\bar{y}_i)]}{\sum_{i=1}^m 1 / \text{Var}(\bar{y}_i)}$$

$$\frac{SCR}{\dot{\sigma}^4} - \frac{N - m}{\dot{\sigma}_2} + \sum_{i=1}^m \frac{n_i(\bar{y}_{i.} - \dot{\mu})^2}{\dot{\lambda}_i^2} - \sum_{i=1}^m \frac{1}{\dot{\lambda}_i} = 0$$

$$\sum_{i=1}^m \frac{n_i^2(\bar{y}_{i.} - \dot{\mu})^2}{\dot{\lambda}_i^2} = \sum_{i=1}^m \frac{n_i}{\dot{\lambda}_i}$$

avec  $\dot{\lambda}_i = \dot{\sigma}^2 + n_i \dot{\sigma}_a^2$  et  $\text{Var}(\bar{y}_{i.}) = \dot{\sigma}_a^2 + \frac{1}{n_i} \dot{\sigma}^2$ . □

La solution pour  $\mu$  a une forme explicite (mais qui contient des inconnues) et elle est une moyenne pondérée de  $\bar{y}_{i.}$  par l'inverse d'un estimateur de sa variance (car  $\text{Var}(\bar{y}_{i.}) = \sigma_a^2 + \frac{1}{n_i} \sigma^2$ ). Par contre, les solutions pour les deux paramètres de variance n'ont pas de forme explicite, il faut donc les calculer numériquement.

### Les estimateurs du maximum de vraisemblance

Après avoir trouvé numériquement les solutions des équations précédentes pour  $\dot{\sigma}^2$  et  $\dot{\sigma}_a^2$ , les estimateurs sont les suivants.

**Proposition 2.26.** *Si  $\dot{\sigma}_a^2 \geq 0$  :*

$$\hat{\sigma}^2 = \dot{\sigma}^2, \quad \hat{\sigma}_a^2 = \dot{\sigma}_a^2 \quad \text{et} \quad \hat{\mu} = \dot{\mu}$$

*Si  $\dot{\sigma}_a^2 < 0$  :*

$$\hat{\sigma}^2 = \frac{SCT}{N}, \quad \hat{\sigma}_a^2 = 0 \quad \text{et} \quad \hat{\mu} = \bar{y}_{..}$$

*Démonstration.* Lorsque  $\sigma_a^2$  est remplacé par 0 dans la formule de la vraisemblance, les  $\lambda_i$  ne dépendent plus de  $i$  donc ils deviennent seulement  $\lambda$  et la formule devient la même que dans le cas où les données sont équilibrées. Aussi, nous remarquons qu'il n'y a pas

de condition sur  $\sigma^2$ , car il est possible de montrer que  $L$  atteint son maximum en une valeur positive de  $\sigma^2$ . En effet,

$$\lim_{\sigma^2 \rightarrow 0} \log L = -\infty$$

$$\lim_{\sigma^2 \rightarrow \infty} \log L = -\infty$$

Donc  $L \rightarrow 0$  lorsque  $\sigma^2$  tend vers 0 ou vers  $\infty$ . La figure 2.1 représente une esquisse de la vraisemblance ( $L$ ) en fonction de  $\sigma^2$ . □

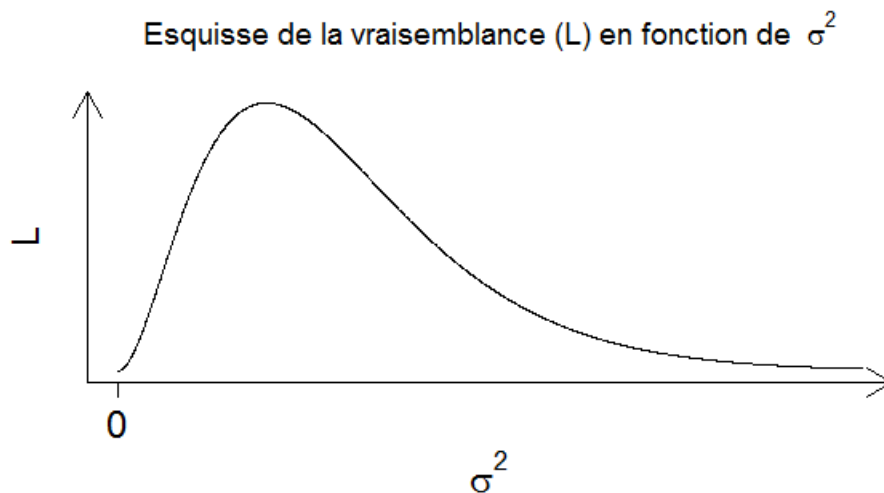


Figure 2.1 – Esquisse de la vraisemblance par rapport à  $\sigma^2$

### Espérance et biais des estimateurs

Dans le cas où les données sont non équilibrées, les espérances des estimateurs de  $\mu$ ,  $\sigma^2$  et de  $\sigma_a^2$  ne peuvent pas être dérivées analytiquement, car la probabilité que  $\hat{\sigma}_a^2$  soit négatif n'a pas une forme explicite. De plus, puisque les espérances n'ont pas de forme explicite, il est ardu de savoir si les estimateurs du maximum de vraisemblance sont sans biais ou non.

### Variance des estimateurs (et leur loi)

Par contre, il est possible de trouver la matrice de variance-covariance asymptotique des estimateurs du maximum de vraisemblance en procédant comme dans le cas où les données sont équilibrées.

#### **Proposition 2.27.**

$$\begin{bmatrix} \hat{\mu} \\ \hat{\sigma}^2 \\ \hat{\sigma}_a^2 \end{bmatrix} \approx \mathcal{N} \left( \begin{bmatrix} \mu \\ \sigma^2 \\ \sigma_a^2 \end{bmatrix}, \begin{pmatrix} \left( \sum_{i=1}^m \frac{n_i}{\sigma^2 + n_i \sigma_a^2} \right)^{-1} & 0 & 0 \\ 0 & \frac{2}{D} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} & -\frac{2}{D} \sum_{i=1}^m \frac{n_i}{\lambda_i^2} \\ 0 & -\frac{2}{D} \sum_{i=1}^m \frac{n_i}{\lambda_i^2} & \frac{2}{D} \left( \frac{N-m}{\sigma^4} + \sum_{i=1}^m \frac{1}{\lambda_i^2} \right) \end{pmatrix} \right)$$

$$\text{Où } D = \frac{N-m}{\sigma^4} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} + \sum_{i=1}^m \frac{1}{\lambda_i^2} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} - \left( \sum_{i=1}^m \frac{n_i}{\lambda_i^2} \right)^2.$$

*Démonstration.* Trouvons tout d'abord les dérivées secondes de  $l$ .

$$l_{\mu\mu} = \frac{\partial^2 l}{\partial \mu \partial \mu} = - \sum_{i=1}^m \frac{n_i}{\lambda_i}$$

$$l_{\mu\sigma^2} = \frac{\partial^2 l}{\partial \mu \partial \sigma^2} = - \sum_{i=1}^m \frac{n_i (\bar{y}_i - \mu)}{\lambda_i^2}$$

$$l_{\mu\sigma_a^2} = \frac{\partial^2 l}{\partial \mu \partial \sigma_a^2} = - \sum_{i=1}^m \frac{n_i^2 (\bar{y}_i - \mu)}{\lambda_i^2}$$

$$l_{\sigma^2\sigma_a^2} = \frac{\partial^2 l}{\partial \sigma^2 \partial \sigma_a^2} = \frac{1}{2} \sum_{i=1}^m \frac{n_i}{\lambda_i^2} - \sum_{i=1}^m \frac{n_i^2 (\bar{y}_i - \mu)^2}{\lambda_i^3}$$

$$l_{\sigma^2\sigma^2} = \frac{\partial^2 l}{\partial \sigma^2 \partial \sigma^2} = \frac{N-m}{2\sigma^4} - \frac{SCR}{\sigma^6} + \frac{1}{2} \sum_{i=1}^m \frac{1}{\lambda_i^2} - \sum_{i=1}^m \frac{n_i (\bar{y}_i - \mu)^2}{\lambda_i^3}$$

$$l_{\sigma_a^2 \sigma_a^2} = \frac{\partial^2 l}{\partial \sigma_a^2 \partial \sigma_a^2} = \frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} - \sum_{i=1}^m \frac{n_i^3 (\bar{y}_i - \mu)^2}{\lambda_i^3}$$

En prenant le négatif des espérances des dérivées partielles calculées plus haut, il s'ensuit que

$$-\mathbb{E}[l_{\mu\mu}] = \sum_{i=1}^m \frac{n_i}{\lambda_i}$$

$$-\mathbb{E}[l_{\mu\sigma^2}] = 0$$

$$-\mathbb{E}[l_{\mu\sigma_a^2}] = 0$$

$$-\mathbb{E}[l_{\sigma^2 \sigma_a^2}] = \frac{1}{2} \sum_{i=1}^m \frac{n_i}{\lambda_i^2}$$

$$-\mathbb{E}[l_{\sigma^2 \sigma^2}] = \frac{N-m}{2\sigma^4} + \frac{1}{2} \sum_{i=1}^m \frac{1}{\lambda_i^2}$$

$$-\mathbb{E}[l_{\sigma_a^2 \sigma_a^2}] = \frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2}$$

Donc,

$$\text{Var} \begin{bmatrix} \hat{\mu} \\ \hat{\sigma}^2 \\ \hat{\sigma}_a^2 \end{bmatrix} \approx \begin{pmatrix} \sum_{i=1}^m \frac{n_i}{\lambda_i} & 0 & 0 \\ 0 & \frac{N-m}{2\sigma^4} + \frac{1}{2} \sum_{i=1}^m \frac{1}{\lambda_i^2} & \frac{1}{2} \sum_{i=1}^m \frac{n_i}{\lambda_i^2} \\ 0 & \frac{1}{2} \sum_{i=1}^m \frac{n_i}{\lambda_i^2} & \frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} \end{pmatrix}^{-1}$$



Et donc, il en découle que

$$\text{Var}(\hat{\mu}) \approx \left( \sum_{i=1}^m \frac{n_i}{\sigma^2 + n_i \sigma_a^2} \right)^{-1}$$

et

$$\text{Var} \begin{bmatrix} \hat{\sigma}^2 \\ \hat{\sigma}_a^2 \end{bmatrix} \approx \frac{2}{D} \begin{pmatrix} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} & - \sum_{i=1}^m \frac{n_i}{\lambda_i^2} \\ - \sum_{i=1}^m \frac{n_i}{\lambda_i^2} & \frac{N-m}{\sigma^4} + \sum_{i=1}^m \frac{1}{\lambda_i^2} \end{pmatrix}$$

Où  $D = \frac{N-m}{\sigma^4} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} + \sum_{i=1}^m \frac{1}{\lambda_i^2} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} - \left( \sum_{i=1}^m \frac{n_i}{\lambda_i^2} \right)^2$ . Plus de détails sont fournis à la section B.3.2 des annexes. □

### g) Tests d'hypothèse et intervalles de confiance

#### Pour la moyenne globale $\mu$

Voici une proposition pour le cas où les données sont équilibrées.

#### **Proposition 2.28.**

$$\frac{\bar{y}_{..} - \mu}{\sqrt{CME/mn}} \sim t_{(m-1)}$$

*Démonstration.*

$$\frac{SCE}{n\sigma_a^2 + \sigma^2} \sim \chi_{(m-1)}^2$$

et

$$\frac{\bar{y}_{..} - \mu}{\sqrt{\frac{1}{m}(\sigma_a^2 + \frac{1}{n}\sigma^2)}} \sim \mathcal{N}(0, 1)$$

Puisque  $\hat{\mu} = \bar{y}_{..} \sim \mathcal{N}\left(\mu, \frac{\sigma^2 + n\sigma_a^2}{mn}\right)$  (voir l'équation 2.9)

Puisque  $SCE$  est une fonction de  $(\bar{y}_i - \bar{y}_{..})$  et  $\hat{\mu}$  est une fonction de  $\bar{y}_{..}$ , ces deux statistiques sont indépendantes. Ainsi, en divisant la loi normale par la racine de la loi du khi carré sur son degré de liberté, il en découle le résultat souhaité.  $\square$

De plus, avec cette statistique de Student, il est possible de construire facilement le test d'hypothèse et les intervalles de confiance sur  $\mu$ .

### Pour $\sigma^2$

**Proposition 2.29.** *Un intervalle de confiance pour  $\sigma^2$  est donné par*

$$IC(\sigma^2) = \left[ \frac{SCR}{\chi_{(N-m), 1-\alpha/2}^2}, \frac{SCR}{\chi_{(N-m), \alpha/2}^2} \right]$$

*Démonstration.* Remarquons que  $\frac{(N-m)s^2}{\sigma^2} = \frac{SCR}{\sigma^2} \sim \chi_{(N-m)}^2$  dans le cas où les données sont équilibrées ou non.  $\square$

### Pour $\sigma_a^2$

**Proposition 2.30.** *Dans le cas où les données sont équilibrées ou non, il est possible d'utiliser une statistique de Fisher pour tester l'hypothèse nulle  $H_0 : \sigma_a^2 = 0$ . Il suffit de rejeter  $H_0$  si  $F = \frac{CME}{CMR} > F_{N-m, 1-\alpha}^{m-1}$ .*

*Démonstration.* Les détails de la preuve sont présentés à la section B.3.3 des annexes.  $\square$

## h) Prédiction

Voici une définition pour la meilleure prédiction au sens des moindres carrés.

**Définition 2.1.**  $BP(Y) = \mathbb{E}[Y|X]$  dans le cas d'un modèle  $Y = X + e$ , avec  $BP$  représentant la meilleure prédiction (« Best Predictor »).

Dans ce contexte, nous voulons les meilleures prédictions des  $a_i$  en ayant accès aux données  $(y_{ij})$ . La portion des données qui est pertinente pour les  $a_i$  sont les  $\bar{y}_{i.}$ . Il s'ensuit que

**Proposition 2.31.**  $BP(a_i) = \mathbb{E}[a_i|y] = \frac{\sigma_a^2}{\sigma_a^2 + \sigma^2/n_i}(\bar{y}_{i.} - \mu)$

*Démonstration.*

$$\begin{aligned}\mathbb{E}[a_i|y] &= \mathbb{E}[a_i|\bar{y}_{i.}] = \mathbb{E}[a_i] + \text{cov}(a_i, \bar{y}_{i.})[\text{Var}(\bar{y}_{i.})]^{-1}(\bar{y}_{i.} - \mathbb{E}[\bar{y}_{i.}]) \\ &= 0 + \frac{\sigma_a^2}{\sigma_a^2 + \sigma^2/n_i}(\bar{y}_{i.} - \mu) \\ &= \frac{\sigma_a^2}{\sigma_a^2 + \sigma^2/n_i}(\bar{y}_{i.} - \mu)\end{aligned}$$

Car, premièrement,  $a_i$  et  $\bar{y}_{i.}$  sont de lois normales donc il est possible de dériver leur espérance conditionnelle à l'aide des propriétés de la loi normale multidimensionnelle.

### Rappel 7.

$$\mathbb{E}[X_2|X_1 = x_1] = \mu_2 + \Sigma_{12}\Sigma_{11}^{-1}(x_1 - \mu_1)$$

$$\text{Où } \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12} & \Sigma_{22} \end{pmatrix} \right)$$

Deuxièmement, car,

$$\begin{aligned}
\text{cov}(a_i, \bar{y}_{i.}) &= \mathbb{E}[a_i \bar{y}_{i.}] - \mathbb{E}[a_i] \mathbb{E}[\bar{y}_{i.}] \\
&= \mathbb{E}[a_i \bar{y}_{i.}] \\
&= \mathbb{E}[\mathbb{E}[a_i \bar{y}_{i.} | a_i]] \\
&= \mathbb{E}[a_i \mathbb{E}[\bar{y}_{i.} | a_i]] \\
&= \mathbb{E}[a_i (a_i + \mu)] \\
&= \sigma_a^2 \quad \square
\end{aligned}$$

Il y a un petit problème ici par contre,  $\sigma^2$  et  $\sigma_a^2$  sont très souvent inconnus donc il serait mieux d'utiliser leur estimateur. Il est donc possible d'estimer la meilleure prédiction par

$$\tilde{a}_i = \widehat{\text{BP}}(a_i) = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}^2/n_i} (\bar{y}_{i.} - \hat{\mu}) \quad (2.10)$$

Si les données sont équilibrées, l'expression précédente devient

$$\tilde{a}_i = \widehat{\text{BP}}(a_i) = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}^2/n} (\bar{y}_{i.} - \bar{y}_{..}) \quad (2.11)$$

### 2.2.3 Comparaison des deux modèles d'analyse de la variance à un facteur

Voici deux tableaux résumés qui comparent les deux méthodes d'analyse de la variance à un facteur, une en le traitant comme un effet fixe et l'autre comme un effet aléatoire. Les tableaux représentent les résultats dans le cas où les données sont équilibrées ( $n_i = n \forall i$ ) puisqu'il n'y a pas de forme explicite pour les estimateurs du maximum de vraisemblance des variances dans le modèle avec un effet aléatoire.

Estimateurs	Fixe	Aléatoire
$\hat{\mu}$	$\bar{y}_{..}$	$\bar{y}_{..}$
$\hat{\alpha}_i$	$\bar{y}_{i.} - \bar{y}_{..}$	-
$\hat{\sigma}^2$	$\frac{SCR}{N}$	$\frac{SCR}{N-m}$ ou $\frac{SCT}{N}$
$\hat{\sigma}_a^2$	-	$\frac{(1-1/m)CME-CMR}{n}$ ou 0

Tableau 2.5 – Comparaison entre les différents estimateurs - ANOVA à un facteur

Tests	Fixe	Aléatoire
$H_0 : \alpha_i = 0 \forall i$	$F = \frac{CME}{CMR} \sim F_{N-m}^{m-1}$	-
$H_0 : \sigma_a^2 = 0$	-	$F = \frac{CME}{CMR} \sim F_{N-m}^{m-1}$

Tableau 2.6 – Comparaison des tests d’hypothèse - ANOVA à un facteur

Nous remarquons que c’est exactement le même test et que les estimateurs se ressemblent. Puisque les tests sont les mêmes, nous pourrions nous demander pourquoi utiliser le modèle avec un effet aléatoire au lieu d’utiliser le modèle avec un effet fixe. La différence sera au niveau de la prédiction d’une nouvelle donnée sachant son groupe. Par exemple, étudions le cas où les données sont équilibrées et que la nouvelle donnée appartient au groupe  $i$ . Dans le cas de l’analyse de la variance avec un facteur fixe, la prédiction sera  $\hat{\mu} + \hat{\alpha}_i = \bar{y}_{i.}$  Par contre, dans le cas de l’analyse de la variance à un facteur aléatoire, la prédiction sera  $\hat{\mu} + \tilde{\alpha}_i = \bar{y}_{..} + \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}^2/n}(\bar{y}_{i.} - \bar{y}_{..}) = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}^2/n}\bar{y}_{i.} + \frac{\hat{\sigma}^2/n}{\hat{\sigma}_a^2 + \hat{\sigma}^2/n}\bar{y}_{..}$ . Donc, si nous voulons faire des prédictions, nous devons choisir le modèle minutieusement puisque les résultats vont différer d’un modèle à l’autre. Par contre, si nous désirons uniquement tester si le groupe a une influence sur l’observation, alors c’est moins crucial de choisir le bon modèle.

## 2.3 Régression linéaire à une variable explicative

Pour cette prochaine section, nous nous intéressons à la relation entre deux variables continues avec ou sans une autre variable catégorique. Nous ferons donc une régression linéaire entre la variable dépendante ( $y_{ij}$ ) et la variable explicative ( $x_{ij}$ ) avec ou sans ajustement pour une tierce variable ( $a_i$  ou  $\alpha_i$ ).

Il faut tout d'abord introduire une nouvelle notation pour écrire les vecteurs et les matrices en fonction de leurs éléments.

Soit  $u_1, \dots, u_n$  des scalaires, des vecteurs ou des matrices, alors

$$\{c u_i\}_{i=1}^n = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$$

$$\{l u_i\}_{i=1}^n = [u_1, \dots, u_n]$$

$$\{d u_i\}_{i=1}^n = \begin{pmatrix} u_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & u_n \end{pmatrix}$$

De plus, il se peut que les indices soient omis si c'est évident. Voici donc la régression linéaire sous sa forme la plus simple.

### 2.3.1 Régression linéaire simple classique

#### a) Le modèle

$$\begin{aligned} \mathbb{E}[y_i] &= \mu + \beta x_i \\ y_i &\sim \mathcal{N}(\mu + \beta x_i, \sigma^2) \text{ indépendantes} \quad i = 1, 2, \dots, N \end{aligned} \tag{2.12}$$

#### b) Les sommes de carrés

Voici les sommes de carrés qui interviennent dans ce modèle.

Notation	Expression	Carré moyen	Signification
$SCC$	$\frac{S_{xy}^2}{S_{xx}}$	$CMC = \frac{SCC}{1}$	Carrés dûs à la covariance
$SCRR$	$\sum_{i=1}^N \left[ (y_i - \bar{y}) - \frac{S_{xy}}{S_{xx}}(x_i - \bar{x}) \right]^2$	$CMRR = \frac{SCRR}{N-2}$	Carrés résiduels après la régression
$SCT$	$\sum_{i=1}^N (y_{ij} - \bar{y})^2$	—	Carrés totaux

Tableau 2.7 – Tableau des sommes de carrés, régression linéaire simple

Où  $S_{xx} = \sum_{i=1}^N (x_i - \bar{x})^2$  et où  $S_{xy} = \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})$ . Il est aussi possible de montrer que  $SCT = SCRR + SCC$ .

#### c) La vraisemblance

La vraisemblance est facile à obtenir puisque toutes les observations sont indépendantes.

**Proposition 2.32.**  $L = (2\pi\sigma)^{-N/2} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^N (y_i - \mu - \beta x_i)^2}$

*Démonstration.* Remarquons que  $L = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(y_i - \mu - \beta x_i)^2}$  et le résultat est direct. □

Ainsi,

$$l = \log(L) = -\frac{N}{2} \log(2\pi) - \frac{N}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^N (y_i - \mu - \beta x_i)^2 \quad (2.13)$$

#### d) Les équations du maximum de vraisemblance et leurs solutions

**Proposition 2.33.** *Les solutions du maximum de vraisemblance sont données par*

$$\begin{aligned} \dot{\beta} &= \frac{S_{xy}}{S_{xx}} = \sqrt{\frac{SCC}{S_{xx}}} \\ \dot{\mu} &= \bar{y} - \dot{\beta} \bar{x} \\ \dot{\sigma}^2 &= \frac{SCRR}{N} \end{aligned}$$

*Démonstration.* Il suffit de dériver  $l$  par rapport à toutes les variables, d'égaliser les expressions obtenues à 0 et de résoudre pour trouver les solutions du maximum de vraisemblance.

$$\begin{aligned} \frac{\partial l}{\partial \mu} &= \frac{1}{\sigma^2} \sum_{i=1}^N (y_i - \mu - \beta x_i) \\ \frac{\partial l}{\partial \beta} &= \frac{1}{\sigma^2} \sum_{i=1}^N x_i (y_i - \mu - \beta x_i) \\ \frac{\partial l}{\partial \sigma^2} &= -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^N (y_i - \mu - \beta x_i)^2 \end{aligned}$$

Donc en égalant ces expressions à 0, il s'ensuit les équations suivantes.



$$\begin{aligned}\frac{\partial l}{\partial \mu} = 0 &\Rightarrow \sum_{i=1}^N y_i - N\mu - \beta \sum_{i=1}^N x_i = 0 \Rightarrow \dot{\mu} = \bar{y} - \dot{\beta}\bar{x} \\ \frac{\partial l}{\partial \beta} = 0 &\Rightarrow \sum_{i=1}^N x_i y_i - \mu \sum_{i=1}^N x_i - \beta \sum_{i=1}^N x_i^2 = 0 \Rightarrow \sum_{i=1}^N x_i y_i = \dot{\mu} \sum_{i=1}^N x_i + \dot{\beta} \sum_{i=1}^N x_i^2 \\ \frac{\partial l}{\partial \sigma^2} = 0 &\Rightarrow \frac{N}{2\sigma^2} = \frac{1}{2\sigma^4} \sum_{i=1}^N (y_i - \mu - \beta x_i)^2 \Rightarrow \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N (y_i - \dot{\mu} - \dot{\beta} x_i)^2\end{aligned}$$

Le résultat est obtenu après quelques simplifications. □

### e) Les estimateurs du maximum de vraisemblance

Puisque les solutions du maximum de vraisemblance se retrouvent toutes dans le bon domaine, les estimateurs du maximum de vraisemblance sont donnés par les solutions.

**Proposition 2.34.**

$$\hat{\beta} = \dot{\beta} \quad ; \quad \hat{\mu} = \dot{\mu} \quad ; \quad \hat{\sigma}^2 = \dot{\sigma}^2$$

### f) Les lois des estimateurs

Les distributions des estimateurs du maximum de vraisemblance sont les suivantes.

**Proposition 2.35.**

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \frac{\sigma^2}{S_{xx}} \begin{pmatrix} \sum_{i=1}^N \frac{x_i^2}{N} & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix} \right)$$

*Démonstration.* La preuve est à la section B.4.1 des annexes. □

De plus,

**Proposition 2.36.**  $\hat{\mu}$  et  $\hat{\beta}$  sont indépendants de  $\hat{\sigma}^2$  qui suit une loi du khi carré. En effet,  $\frac{N\hat{\sigma}^2}{\sigma^2} \sim \chi_{(N-2)}^2$ .

*Démonstration.* La preuve de la loi de  $\hat{\sigma}^2$  et la preuve de l'indépendance sont à la section B.4.2 des annexes. □

### g) Tests d'hypothèse et intervalles de confiance

En utilisant les distributions énoncées ci-haut, il est possible de dériver des tests sur les paramètres du modèle.

Pour  $H_0 : \beta = 0$

**Proposition 2.37.** *Rejetons  $H_0$  si*

$$\left| \frac{\frac{\hat{\beta}}{N}}{\sqrt{\frac{N-2}{S_{xx}} \hat{\sigma}^2}} \right| > t_{(N-2), \alpha/2}$$

*Démonstration.*  $\frac{\hat{\beta}}{\sqrt{\sigma^2/S_{xx}}} \sim \mathcal{N}(0, 1)$ ,  $\frac{N\hat{\sigma}^2}{\sigma^2} \sim \chi_{(N-2)}^2$  et ils sont indépendants. □

Ou de façon équivalente, puisque la statistique de Fisher suivante est le carré de la statistique de Student mentionnée ci-haut, voici un autre test sur  $\hat{\beta}$ .

**Proposition 2.38.** *Un autre test (équivalent) pour tester l'hypothèse nulle sur  $\beta$  est d'utiliser la statistique de Fisher suivante*

$$F = \frac{SCC/1}{SCRR/(N-2)} \sim \mathcal{F}_{N-2}^1$$

*Démonstration.*  $\frac{\hat{\beta}^2}{\sigma^2/S_{xx}} = \frac{SCC}{\sigma^2} \sim \chi_{(1)}^2$ ,  $\frac{N\hat{\sigma}^2}{\sigma^2} = \frac{SCRR}{\sigma^2} \sim \chi_{(N-2)}^2$  et ils sont indépendants. □

Pour  $\sigma^2$

**Proposition 2.39.** *Un intervalle de confiance de niveau  $1 - \alpha$  sur  $\sigma^2$  est le suivant*

$$\left[ \frac{N\hat{\sigma}^2}{\chi_{(N-2),1-\alpha/2}^2}, \frac{N\hat{\sigma}^2}{\chi_{(N-2),\alpha/2}^2} \right]$$

*Démonstration.* La preuve est directe en se servant de la loi de  $\hat{\sigma}^2$  (propriété 2.36).  $\square$

### 2.3.2 Régression linéaire simple avec un effet fixe (groupe) (ANCOVA)

Voici la théorie de l'analyse de la covariance (ANCOVA), un modèle qui est une combinaison de la régression linéaire simple et de l'analyse de la variance. Nous allons présenter la théorie dans le cas où les données sont non nécessairement équilibrées.

#### a) Le modèle

$$\begin{aligned} \mathbb{E}[y_{ij}] &= \mu + \alpha_i + \beta x_{ij} & i = 1, \dots, m ; j = 1, \dots, n_i \\ y_{ij} &\sim \mathcal{N}(\mu + \alpha_i + \beta x_{ij}, \sigma^2) \text{ indépendantes} \end{aligned} \tag{2.14}$$

Sous la contrainte que  $\sum_{i=1}^m n_i \alpha_i = 0$ . Posons aussi  $N = \sum_{i=1}^m n_i$ .

Ici,  $\mu$  représente l'ordonnée à l'origine de la population globale et  $\alpha_i$  représente la différence entre l'ordonnée à l'origine globale et l'ordonnée à l'origine du groupe  $i$ . Il s'ensuit que l'ordonnée à l'origine du groupe  $i$  est  $\mu + \alpha_i$ .

## b) Les sommes de carrés

Voici les sommes de carrés qui interviennent dans ce modèle.

Notation	Expression	Carré moyen	Signification
$SCE$	$\sum_{i=1}^m \sum_{j=1}^{n_i} (\bar{y}_{i.} - \bar{y}_{..})^2$	$CME = \frac{SCE}{m-1}$	Carrés expliqués
$SCR$	$\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2$	$CMR = \frac{SCR}{N-m}$	Carrés résiduels
$SCT$	$\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2$	—	Carrés totaux
$SCC$	$\frac{S_{xy}^2}{S_{xx}}$	$CMC = \frac{SCC}{1}$	Carrés dûs à la covariance
$SCR$	$\sum_{i=1}^m \sum_{j=1}^{n_i} \left[ y_{ij} - \bar{y}_{i.} - \frac{S_{xy}}{S_{xx}}(x_{ij} - \bar{x}_{i.}) \right]^2$	$CMRR = \frac{SCR}{N-m-1}$	Carrés résiduels après la régression

Tableau 2.8 – Tableau des sommes de carrés, ANCOVA avec un effet fixe

Où  $S_{xy} = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})(y_{ij} - \bar{y}_{i.})$ ,  $S_{xx} = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})^2$ . Il est aussi possible de montrer que  $SCT = SCE + SCR$  et  $SCR + SCC = SCR$ .

## c) La vraisemblance

Premièrement, voici la loi de  $\mathbf{y}$  qui est le vecteur de toutes les observations.

**Proposition 2.40.**

$$\mathbf{y} \sim \mathcal{N}(\mu \mathbf{1}_N + \{c \mathbf{1}_{n_i} \alpha_i\} + \beta \mathbf{x}, \mathbf{I}_N \sigma^2)$$

Où  $\mathbf{x}$  est le vecteur  $[x_{11}, \dots, x_{mn_m}]^t$ .

Ainsi, la vraisemblance se déduit facilement puisque toutes les observations sont indépendantes. Il en découle que

**Proposition 2.41.**  $L = (2\pi\sigma^2)^{-\frac{N}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i - \beta x_{ij})^2\right)$

*Démonstration.* Remarquons que  $L = \prod_{i=1}^m \prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y_{ij} - \mu - \alpha_i - \beta x_{ij})^2\right)$  et le résultat est direct.  $\square$

Ainsi,

$$l = \log(L) = -\frac{N}{2} \log(2\pi) - \frac{N}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i - \beta x_{ij})^2 \quad (2.15)$$

#### d) Les équations du maximum de vraisemblance et leurs solutions

**Proposition 2.42.** *Les solutions du maximum de vraisemblance sont données par*

$$\dot{\mu} = \bar{y}_{..} - \dot{\beta} \bar{x}_{..}$$

$$\dot{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..} - \dot{\beta}(\bar{x}_{i.} - \bar{x}_{..})$$

$$\dot{\beta} = \frac{S_{xy}}{S_{xx}} = \sqrt{\frac{SCC}{S_{xx}}}$$

$$\dot{\sigma}^2 = \frac{SCRR}{N}$$

*Démonstration.* Pour trouver les solutions du maximum de vraisemblance, la log-vraisemblance doit être dérivée et les dérivées partielles doivent être égalées à 0. Les équations obtenues sont les suivantes.

$$\begin{aligned}
\frac{\partial l}{\partial \mu} = 0 &\Rightarrow \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i - \beta x_{ij}) = 0 \\
&\Rightarrow N\bar{y}_{..} - N\mu - \sum_{i=1}^m n_i \alpha_i - \beta N\bar{x}_{..} = 0 \\
&\Rightarrow \dot{\mu} = \bar{y}_{..} - \dot{\beta} \bar{x}_{..}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial l}{\partial \alpha_i} = 0 &\Rightarrow \frac{1}{\sigma^2} \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i - \beta x_{ij}) = 0 \\
&\Rightarrow n_i \bar{y}_{i.} - n_i \mu - n_i \alpha_i - \beta n_i \bar{x}_{i.} = 0 \\
&\Rightarrow \dot{\alpha}_i = \bar{y}_{i.} - \dot{\mu} - \dot{\beta} \bar{x}_{i.} \\
&\Rightarrow \dot{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..} + \dot{\beta} \bar{x}_{..} - \dot{\beta} \bar{x}_{i.} \\
&\Rightarrow \dot{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..} - \dot{\beta} (\bar{x}_{i.} - \bar{x}_{..})
\end{aligned}$$

$$\begin{aligned}
\frac{\partial l}{\partial \beta} = 0 &\Rightarrow \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} (y_{ij} - \mu - \alpha_i - \beta x_{ij}) = 0 \\
&\Rightarrow \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} y_{ij} - \mu N \bar{x}_{..} - \sum_{i=1}^m \alpha_i n_i \bar{x}_{i.} - \beta \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}^2 = 0 \\
&\Rightarrow \dot{\beta} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}^2 = \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} y_{ij} - \dot{\mu} N \bar{x}_{..} - \sum_{i=1}^m \dot{\alpha}_i n_i \bar{x}_{i.} \\
&\Rightarrow \dot{\beta} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}^2 - N \dot{\beta} \bar{x}_{..}^2 - \dot{\beta} \sum_{i=1}^m (\bar{x}_{i.} - \bar{x}_{..}) n_i \bar{x}_{i.} = \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} y_{ij} - N \bar{x}_{..} \bar{y}_{..} - \sum_{i=1}^m (\bar{y}_{i.} - \bar{y}_{..}) n_i \bar{x}_{i.} \\
&\Rightarrow \dot{\beta} \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij}^2 - \bar{x}_{i.}^2) = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} y_{ij} - \bar{y}_{i.} \bar{x}_{i.}) \\
&\Rightarrow \dot{\beta} = \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})(y_{ij} - \bar{y}_{i.})}{\sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})^2} = \frac{S_{xy}}{S_{xx}}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial l}{\partial \sigma^2} = 0 &\Rightarrow -\frac{N}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i - \beta x_{ij})^2 = 0 \\
&\Rightarrow N\sigma^2 = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu - \alpha_i - \beta x_{ij})^2 \\
&\Rightarrow \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta} x_{ij})^2 \\
&\Rightarrow \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} [y_{ij} - \bar{y}_{..} + \hat{\beta} \bar{x}_{..} - \bar{y}_{i.} + \bar{y}_{..} + \hat{\beta} (\bar{x}_{i.} - \bar{x}_{..}) - \hat{\beta} x_{ij}]^2 \\
&\Rightarrow \hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} [(y_{ij} - \bar{y}_{i.}) - \hat{\beta} (x_{ij} - \bar{x}_{i.})]^2 = \frac{SCRR}{N} \\
&\Rightarrow \hat{\sigma}^2 = \frac{1}{N} \left( S_{yy} - \frac{S_{xy}^2}{S_{xx}} \right)
\end{aligned}$$

Où  $S_{yy} = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 = SCR$ . □

### e) Les estimateurs du maximum de vraisemblance

Il s'ensuit que  $\hat{\mu} \in \mathbb{R}$ ,  $\hat{\alpha}_i \in \mathbb{R} \forall i$ ,  $\hat{\beta} \in \mathbb{R}$  et que  $\hat{\sigma}^2 > 0$  puisque c'est une somme de carrés (voir l'avant dernière ligne dans les calculs pour  $\hat{\sigma}^2$  de la démonstration de la propriété 2.42). Les estimateurs du maximum de vraisemblance sont donc donnés par les solutions.

#### Proposition 2.43.

$$\begin{aligned}
\hat{\beta} &= \hat{\beta} = \frac{S_{xy}}{S_{xx}} \\
\hat{\mu} &= \hat{\mu} = \bar{y}_{..} - \hat{\beta} \bar{x}_{..} \\
\hat{\alpha}_i &= \hat{\alpha}_i = (\bar{y}_{i.} - \bar{y}_{..}) - \hat{\beta} (\bar{x}_{i.} - \bar{x}_{..}) \\
\hat{\sigma}^2 &= \hat{\sigma}^2 = \frac{SCRR}{N}
\end{aligned}$$

f) Les lois des estimateurs

La distribution multidimensionnelle des estimateurs  $\hat{\mu}$ ,  $\hat{\beta}$  et de  $\hat{\alpha}_i$  est la suivante pour  $i$  fixé.

**Proposition 2.44.**

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \\ \hat{\alpha}_i \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \mu \\ \beta \\ \alpha_i \end{bmatrix}, \frac{\sigma^2}{S_{xx}} \begin{pmatrix} \bar{x}_{..}^2 + \frac{S_{xx}}{N} & -\bar{x}_{..} & \bar{x}_{..}(\bar{x}_i - \bar{x}_{..}) \\ -\bar{x}_{..} & 1 & -(\bar{x}_i - \bar{x}_{..}) \\ \bar{x}_{..}(\bar{x}_i - \bar{x}_{..}) & -(\bar{x}_i - \bar{x}_{..}) & (\bar{x}_i - \bar{x}_{..})^2 + \frac{S_{xx}(N-n_i)}{N n_i} \end{pmatrix} \right)$$

Il y a aussi que

**Proposition 2.45.**

$$\begin{bmatrix} \hat{\alpha}_i \\ \hat{\alpha}_j \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \alpha_i \\ \alpha_j \end{bmatrix}, \begin{pmatrix} \frac{(\bar{x}_i - \bar{x}_{..})^2 \sigma^2}{S_{xx}} + \frac{\sigma^2(N-n_i)}{N n_i} & \frac{(\bar{x}_i - \bar{x}_{..})(\bar{x}_j - \bar{x}_{..}) \sigma^2}{S_{xx}} - \frac{\sigma^2}{N} \\ \frac{(\bar{x}_i - \bar{x}_{..})(\bar{x}_j - \bar{x}_{..}) \sigma^2}{S_{xx}} - \frac{\sigma^2}{N} & \frac{(\bar{x}_j - \bar{x}_{..})^2 \sigma^2}{S_{xx}} + \frac{\sigma^2(N-n_j)}{N n_j} \end{pmatrix} \right)$$

*Démonstration.* Les preuves détaillées de ces lois sont à la section B.5.1 des annexes.  $\square$

Pour  $\hat{\sigma}^2$ , sa loi est donnée par

**Proposition 2.46.**  $\frac{N\hat{\sigma}^2}{\sigma^2} \sim \chi_{(N-m-1)}^2$

*Démonstration.* Nous avons montré à la section B.5.2 a) des annexes (voir aussi le tableau C.2 dans les annexes) que

$$\frac{SCRR}{\sigma^2} \sim \chi_{(N-m-1)}^2$$

Ensuite,

$$\hat{\sigma}^2 = \frac{SCRR}{N} \Rightarrow \frac{N\hat{\sigma}^2}{\sigma^2} = \frac{SCRR}{\sigma^2} \Rightarrow \frac{N\hat{\sigma}^2}{\sigma^2} \sim \chi_{(N-m-1)}^2 \quad \square$$



## g) Tests d'hypothèse et intervalles de confiance

Des tests d'hypothèse et des intervalles de confiance peuvent être déduits des lois des estimateurs. Nous présenterons aussi les tests du rapport de vraisemblance et des tests associés à une statistique de Fisher. Commençons par trouver  $L(\hat{\theta})$  qui est la vraisemblance avec comme paramètres, les estimateurs.

**Proposition 2.47.**  $L(\hat{\theta}) = \left(2\pi \frac{SCRR}{N}\right)^{-\frac{N}{2}} \exp\left(-\frac{N}{2}\right)$

*Démonstration.*

$$\begin{aligned}
 L(\hat{\theta}) &= (2\pi\hat{\sigma}^2)^{-\frac{N}{2}} \exp\left(-\frac{1}{2\hat{\sigma}^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}x_{ij})^2\right) \\
 &= (2\pi\hat{\sigma}^2)^{-\frac{N}{2}} \exp\left(-\frac{1}{2\hat{\sigma}^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..} + \hat{\beta}\bar{x}_{..} - \bar{y}_{i.} + \bar{y}_{..} + \hat{\beta}(\bar{x}_{i.} - \bar{x}_{..}) - \hat{\beta}x_{ij})^2\right) \\
 &= (2\pi\hat{\sigma}^2)^{-\frac{N}{2}} \exp\left(-\frac{1}{2\hat{\sigma}^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.} - \hat{\beta}(x_{ij} - \bar{x}_{i.}))^2\right) \\
 &= \left(2\pi \frac{SCRR}{N}\right)^{-\frac{N}{2}} \exp\left(-\frac{N}{2}\right) \quad \square
 \end{aligned}$$

### L'hypothèse $H_0 : \beta = 0$

Sous  $H_0$ , les estimateurs des paramètres deviennent

$$\hat{\mu}_0 = \bar{y}_{..} \quad \hat{\alpha}_{i0} = \bar{y}_{i.} - \bar{y}_{..} \quad \hat{\sigma}_0^2 = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 = \frac{SCR}{N} \quad (2.16)$$

Remarquons que ce sont les mêmes estimateurs que dans le cas de l'ANOVA à un facteur fixe. Ceci est cohérent, car si  $\beta = 0$  dans le modèle de l'ANCOVA, nous retombons sur l'ANOVA.

**Proposition 2.48.** *Il y a deux tests qui peuvent être utilisés pour tester l'hypothèse nulle. Il suffit de rejeter  $H_0$  si*

$$N \log \left( \frac{SCR}{SCR R} \right) > \chi_{(1), 1-\alpha}^2$$

Ou si

$$F = \frac{SCC/1}{SCR R / (N - m - 1)} > \mathcal{F}_{N-m-1, 1-\alpha}^1$$

La dernière statistique de Fisher est le carré de la statistique de Student suivante.

$$\frac{\hat{\beta}}{\sqrt{\frac{N\hat{\sigma}^2}{S_{xx}(N-m-1)}}} \sim t_{(N-m-1)}$$

Puisque sous  $H_0$ ,  $\hat{\beta} = \sqrt{\frac{SCC}{S_{xx}}} \sim \mathcal{N}\left(0, \frac{\sigma^2}{S_{xx}}\right)$  et  $\frac{N\hat{\sigma}^2}{\sigma^2} = \frac{SCR R}{\sigma^2} \sim \chi_{(N-m-1)}^2$  et qu'ils sont indépendants ( $SCC$  et  $SCR R$  sont indépendants sous  $H_0$ , voir la section B.5.2 des annexes et le tableau C.2 dans les annexes).

Démonstration.  $L(\hat{\theta}_0)$  est donnée par

$$\begin{aligned} L(\hat{\theta}_0) &= \left(2\pi \frac{SCR}{N}\right)^{-\frac{N}{2}} \exp \left[ -\frac{1}{2\hat{\sigma}_0^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..} - \bar{y}_{i.} + \bar{y}_{.j})^2 \right] \\ &= \left(2\pi \frac{SCR}{N}\right)^{-\frac{N}{2}} \exp \left[ -\frac{1}{2\hat{\sigma}_0^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 \right] \\ &= \left(2\pi \frac{SCR}{N}\right)^{-\frac{N}{2}} \exp \left[ -\frac{N}{2} \right] \\ &\Rightarrow \frac{L(\hat{\theta}_0)}{L(\hat{\theta})} = \frac{\left(2\pi \frac{SCR}{N}\right)^{-N/2}}{\left(2\pi \frac{SCR R}{N}\right)^{-N/2}} = \left(\frac{SCR}{SCR R}\right)^{-N/2} \end{aligned}$$

$$\Rightarrow -2 \log(\Lambda) = -2 \frac{-N}{2} \log \left( \frac{SCR}{SCRR} \right) = N \log \left( \frac{SCR}{SCRR} \right) \approx \chi_{(1)}^2$$

Il est aussi possible de trouver une statistique de Fisher pour ce test. Pour cela, calculons la somme des résidus totaux (modèle complet) et la somme des résidus sous  $H_0$ .

$$\begin{aligned} \text{Somme des résidus totaux} &= \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}x_{ij})^2 \\ &= \sum_{i=1}^m \sum_{j=1}^{n_i} \left[ y_{ij} - \bar{y}_{i\cdot} - \frac{S_{xy}}{S_{xx}}(x_{ij} - \bar{x}_{i\cdot}) \right]^2 \\ &= SCRR \end{aligned}$$

$$\begin{aligned} \text{Somme des résidus sous } H_0 &= \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu}_0 - \hat{\alpha}_{i0})^2 \\ &= \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})^2 \\ &= SCR \end{aligned}$$

Puisque  $\frac{SCR-SCRR}{SCRR} = \frac{SCC}{SCRR}$ , la statistique de Fisher est donnée par

$$F = \frac{SCC/1}{SCRR/(N-m-1)} \sim \mathcal{F}_{N-m-1}^1$$

Voir la section B.5.2 des annexes et le tableau C.2 dans les annexes pour les détails sur les lois de  $SCC$  et  $SCRR$ . □

**L'hypothèse  $H_0 : \alpha_1 = \dots = \alpha_m = 0$**

Sous  $H_0$ , les estimateurs des paramètres deviennent

$$\begin{aligned}
\hat{\beta}_0 &= \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{..})(y_{ij} - \bar{y}_{..})}{\sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{..})^2} \\
\hat{\mu}_0 &= \bar{y}_{..} - \hat{\beta}_0 \bar{x}_{..} \\
\hat{\sigma}_0^2 &= \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} [y_{ij} - \bar{y}_{..} - \hat{\beta}_0(x_{ij} - \bar{x}_{..})]^2 = \frac{SCRR_0}{N}
\end{aligned} \tag{2.17}$$

Remarquons que ce sont exactement les estimateurs de la régression linéaire simple puisque nous ne considérons plus l'effet du groupe. Ceci est cohérent avec le fait de poser les  $\alpha_i$  à 0. (Posons  $SCRR_0 = \sum_{i=1}^m \sum_{j=1}^{n_i} [y_{ij} - \bar{y}_{..} - \hat{\beta}_0(x_{ij} - \bar{x}_{..})]^2$ )

**Proposition 2.49.** *Il y a encore deux tests qui peuvent être utilisés pour tester l'hypothèse nulle. Il suffit de rejeter  $H_0$  si*

$$N \log \left( \frac{SCRR_0}{SCRR} \right) > \chi_{(m-1), 1-\alpha}^2$$

Où  $SCRR_0 = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..} - \hat{\beta}_0(x_{ij} - \bar{x}_{..}))^2$

Ou, si les données sont équilibrées,

$$F = \frac{SCE/(m-1)}{SCRR/(N-m-1)} > \mathcal{F}_{N-m-1, 1-\alpha}^{m-1}$$

**Rappel 8.**

$$\frac{U_1/d_1}{U_2/d_2} \sim \mathcal{F}_{d_2}^{d_1(\lambda)}$$

Où  $U_1 \sim \chi^2(d_1, \lambda)$  et  $U_2 \sim \chi^2(d_2)$  indépendantes.

*Démonstration.*  $L(\hat{\theta}_0)$  est donnée par

$$\begin{aligned}
L(\hat{\theta}_0) &= (2\pi\hat{\sigma}_0^2)^{-N/2} \exp \left[ -\frac{1}{2\hat{\sigma}_0^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..} + \hat{\beta}_0 \bar{x}_{..} - \hat{\beta}_0 x_{ij})^2 \right] \\
&= (2\pi\hat{\sigma}_0^2)^{-N/2} \exp \left[ -\frac{N}{2} \right]
\end{aligned}$$

$$\Rightarrow \Lambda = \frac{L(\hat{\theta}_0)}{L(\hat{\theta})} = \left( \frac{\hat{\sigma}_0^2}{\hat{\sigma}^2} \right)^{-N/2} = \left( \frac{SCRR_0}{SCRR} \right)$$

$$\Rightarrow -2 \log(\Lambda) = -2 \frac{-N}{2} \log \left( \frac{SCRR_0}{SCRR} \right) = N \log \left( \frac{SCRR_0}{SCRR} \right) \approx \chi_{(m-1)}^2$$

Ici, le degré de liberté est de  $m - 1$ , même s'il y a une différence de  $m$  paramètres entre les deux modèles, ceci est dû à la contrainte  $\sum_{i=1}^m n_i \alpha_i = 0$ . En effet, la contrainte fait qu'en fin de compte, dans le modèle total, nous n'estimons que  $(m - 1)$  paramètres  $\alpha_i$  puisque le dernier  $\alpha_i$  est déterminé par les autres.

Il est aussi possible de déduire une statistique de Fisher pour ce test. Nous le montrerons uniquement dans le cas où les données sont équilibrées, c'est-à-dire que  $n_i = n \ \forall i$  et  $x_{ij} = x_j \ \forall i$ . Dans ce cas, réécrivons les estimateurs.

$$\begin{aligned}
\hat{\beta} &= \frac{\sum_{i=1}^m \sum_{j=1}^n (x_j - \bar{x})(y_{ij} - \bar{y}_{i.})}{\sum_{i=1}^m \sum_{j=1}^n (x_j - \bar{x})^2} \\
\hat{\mu} &= \bar{y}_{..} - \hat{\beta} \bar{x} \\
\hat{\alpha}_i &= \bar{y}_{i.} - \bar{y}_{..} \\
\hat{\sigma}^2 &= \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \bar{y}_{i.} + \hat{\beta} \bar{x} - \hat{\beta} x_j)^2
\end{aligned}$$

$$\hat{\beta}_0 = \frac{\sum_{i=1}^m \sum_{j=1}^n (x_j - \bar{x})(y_{ij} - \bar{y}_{..})}{\sum_{i=1}^m \sum_{j=1}^n (x_j - \bar{x})^2}$$

$$\hat{\mu}_0 = \bar{y}_{..} - \hat{\beta}_0 \bar{x}$$

$$\hat{\sigma}_0^2 = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \bar{y}_{..} + \hat{\beta}_0 \bar{x} - \hat{\beta}_0 x_j)^2$$

Ensuite, calculons la somme des résidus totaux (modèle complet) et la somme des résidus sous  $H_0$ .

$$\begin{aligned} \text{Somme des résidus totaux} &= \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta} x_j)^2 \\ &= \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \bar{y}_{i.} - \hat{\beta}(x_j - \bar{x}))^2 \\ &= SCRR \end{aligned}$$

$$\begin{aligned} \text{Somme des résidus sous } H_0 &= \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \hat{\mu}_0 - \hat{\beta}_0 x_j)^2 \\ &= \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \bar{y}_{..} + \hat{\beta}_0 \bar{x} - \hat{\beta}_0 x_j)^2 \\ &= \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \bar{y}_{..} - \hat{\beta}_0(x_j - \bar{x}))^2 \\ &= SCRR_0 \end{aligned}$$

Il est possible de montrer que  $SCRR_0 - SCRR = SCE$ , avec  $SCE = \sum_{i=1}^m \sum_{j=1}^n (\bar{y}_{i.} - \bar{y}_{..})^2$ .

Donc,  $\frac{SCRR_0 - SCRR}{SCRR} = \frac{SCE}{SCRR}$ . Ainsi, la statistique de Fisher est donnée par

$$F = \frac{SCE/(m-1)}{SCRR/(N-m-1)} \sim \mathcal{F}_{N-m-1}^{m-1}$$

Voir la section B.5.2 des annexes et le tableau C.2 dans les annexes pour les détails sur les lois de  $SCE$  et  $SCRR$ . □

Et si les données sont non équilibrées,

$$F = \frac{SCE/(m-1)}{SCRR/(N-m-1)} \sim \mathcal{F}_{N-m-1}^{m-1} \left( \beta^2 \sum_{i=1}^m \sum_{j=1}^{n_i} (\bar{x}_{i.} - \bar{x}_{..})^2 \right)$$

### 2.3.3 Régression linéaire simple avec un effet aléatoire (données équilibrées)

a) Le modèle

$$\begin{aligned} \mathbb{E}[y_{ij}|a_i] &= \mu + a_i + \beta x_j & i = 1, \dots, m ; j = 1, \dots, n \\ y_{ij}|a_i &\sim \mathcal{N}(\mu + a_i + \beta x_j, \sigma^2) \text{ indépendantes} & (2.18) \\ a_i &\sim \text{i.i.d } \mathcal{N}(0, \sigma_a^2) \end{aligned}$$

Remarquons qu'il s'agit de données équilibrées puisque  $n_i = n \ \forall i$ . Remarquons aussi que la variable indépendante ( $x_j$ ) a les mêmes modalités dans chaque groupe, c'est-à-dire que l'observation ( $y$ ) a été mesurée sous les mêmes conditions ( $x_1, \dots, x_n$ ) dans chaque groupe. Un exemple de données pour lesquelles ce modèle pourrait facilement s'appliquer est l'exemple 2.3. Dans cet exemple, nous aurions  $m = 3$  pour les trois groupes et  $n = 4$  pour les 4 températures avec  $x_1 = 10^\circ \text{C}$ ,  $x_2 = 20^\circ \text{C}$ ,  $x_3 = 30^\circ \text{C}$  et  $x_4 = 40^\circ \text{C}$  et aucune donnée manquante.

Il s'ensuit que  $\mathbb{E}[y_{ij}] = \mathbb{E}[\mathbb{E}[y_{ij}|a_i]] = \mathbb{E}[\mu + a_i + \beta x_j] = \mu + \beta x_j$ .

b) Les covariances entre les observations

Voici les variances et les covariances entre les observations.

**Proposition 2.50.**  $\text{Var}(y_{ij}) = \sigma_a^2 + \sigma^2$

*Démonstration.*

$$\begin{aligned}\text{Var}(y_{ij}) &= \text{Var}(\mathbb{E}[y_{ij}|a_i]) + \mathbb{E}[\text{Var}(y_{ij}|a_i)] \\ &= \text{Var}(\mu + a_i + \beta x_j) + \mathbb{E}[\sigma^2] \\ &= \text{Var}(a_i) + \sigma^2 \\ &= \sigma_a^2 + \sigma^2\end{aligned}$$

□

**Proposition 2.51.**  $\text{cov}(y_{ij}, y_{ik}) = \sigma_a^2$  pour  $j \neq k$

*Démonstration.*

$$\begin{aligned}\text{cov}(y_{ij}, y_{ik}) &= \text{cov}(\mathbb{E}[y_{ij}|a_i], \mathbb{E}[y_{ik}|a_i]) + \mathbb{E}[\text{cov}(y_{ij}, y_{ik}|a_i)] \\ &= \text{cov}(\mu + a_i + \beta x_j, \mu + a_i + \beta x_k) + \mathbb{E}[0] \\ &= \text{Var}(a_i) \\ &= \sigma_a^2\end{aligned}$$

□

c) Les sommes de carrés

Voici les sommes de carrés qui interviennent dans ce modèle.



Notation	Expression	Carré moyen	Signification
$SCE$	$\sum_{i=1}^m \sum_{j=1}^n (\bar{y}_{i.} - \bar{y}_{..})^2$	$CME = \frac{SCE}{m-1}$	Carrés expliqués
$SCR$	$\sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2$	$CMR = \frac{SCR}{m(n-1)}$	Carrés résiduels
$SCT$	$\sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2$	—	Carrés totaux
$SCC$	$\frac{S_{xy}^2}{S_{xx}}$	$CMC = \frac{SCC}{1}$	Carrés dûs à la covariance
$SCRR$	$\sum_{i=1}^m \sum_{j=1}^n \left[ y_{ij} - \bar{y}_{i.} - \frac{S_{xy}}{S_{xx}}(x_j - \bar{x}) \right]^2$	$CMRR = \frac{SCRR}{m(n-1)-1}$	Carrés résiduels après la régression

Tableau 2.9 – Tableau des sommes de carrés, régression linéaire avec un effet aléatoire (données équilibrées)

Où  $S_{xx} = \sum_{i=1}^m \sum_{j=1}^n (x_j - \bar{x})^2$  et  $S_{xy} = \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})(x_j - \bar{x})$ . Il est aussi possible de montrer que  $SCT = SCE + SCR$  et  $SCRR + SCC = SCR$

#### d) La vraisemblance

Pour pouvoir écrire la vraisemblance, trouvons tout d'abord la loi de  $\mathbf{y} = [y_{11}, \dots, y_{mn}]^t$ , le vecteur de toutes les observations.

**Proposition 2.52.** *La loi de  $\mathbf{y}$  est donnée par*

$$\mathbf{y} \sim \mathcal{N} \left( \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \mathbf{V} \right)$$

Où  $\mathbf{X} = \mathbf{1}_m \otimes \mathbf{X}_0 = \mathbf{1}_m \otimes [\mathbf{1}_n, \mathbf{x}_0]$  et où  $\mathbf{x}_0 = [x_1, \dots, x_n]^t$ , où  $\mathbf{V} = \mathbf{I}_m \otimes \mathbf{V}_0 = \mathbf{I}_m \otimes (\sigma^2 \mathbf{I}_n + \sigma_a^2 \mathbf{J}_n)$  et où  $\otimes$  représente le produit de Kronecker (ou produit direct). Pour la définition et quelques propriétés de cet opérateur, voir la section A.2 des annexes.

*Démonstration.* Définissons

$$\mathbf{y}_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{in} \end{bmatrix} \quad \mathbf{1}_n = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}_n \quad \text{et} \quad \mathbf{x}_0 = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$\begin{aligned} \Rightarrow \mathbb{E}[\mathbf{y}_i | a_i] &= \mu \mathbf{1}_n + a_i \mathbf{1}_n + \beta \mathbf{x}_0 & \text{et} \quad \mathbb{E}[\mathbf{y}_i] &= \mu \mathbf{1}_n + \beta \mathbf{x}_0 \\ \Rightarrow \mathbb{E}[\mathbf{y}_i | a_i] &= \mathbf{X}_0 \begin{bmatrix} \mu \\ \beta \end{bmatrix} + a_i \mathbf{1}_n & \text{et} \quad \mathbb{E}[\mathbf{y}_i] &= \mathbf{X}_0 \begin{bmatrix} \mu \\ \beta \end{bmatrix} \end{aligned}$$

Où  $\mathbf{X}_0 = [\mathbf{1}_n, \mathbf{x}_0]$ . Ainsi,

$$\begin{aligned} \mathbf{y}_i | a_i &\sim \mathcal{N} \left( \mathbf{X}_0 \begin{bmatrix} \mu \\ \beta \end{bmatrix} + a_i \mathbf{1}_n, \sigma^2 \mathbf{I}_n \right) \text{ indépendantes} \\ \Rightarrow \mathbf{y}_i &\sim \mathcal{N} \left( \mathbf{X}_0 \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \mathbf{V}_0 \right) \text{ indépendantes} \end{aligned}$$

Où  $\mathbf{V}_0 = \text{Var}(\mathbf{y}_i) = \sigma^2 \mathbf{I}_n + \sigma_a^2 \mathbf{J}_n$  en vertu des résultats trouvés sur les covariances des observations. De plus, les  $\mathbf{y}_i$  sont indépendants puisque, pour  $i \neq k$ ,

$$\begin{aligned} \text{cov}(y_{ij}, y_{kl}) &= \text{cov}(\mathbb{E}[y_{ij} | a_i, a_k], \mathbb{E}[y_{kl} | a_i, a_k]) + \mathbb{E}[\text{cov}(y_{ij}, y_{kl} | a_i, a_k)] \\ &= \text{cov}(\mu + a_i + \beta x_j, \mu + a_k + \beta x_l) + \mathbb{E}[0] \\ &= \text{cov}(a_i, a_k) \\ &= 0 \end{aligned}$$

Posons maintenant

$$\mathbf{y} = \{ {}_c \mathbf{y}_i \}_{i=1}^m \quad \text{et} \quad \mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$$

Il en découle que

$$\mathbb{E}[\mathbf{y}|\mathbf{a}] = \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix} + \mathbf{Z}\mathbf{a} \quad \text{et} \quad \mathbb{E}[\mathbf{y}] = \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix}$$

Où  $\mathbf{X} = \mathbf{1}_m \otimes \mathbf{X}_0$  et où  $\mathbf{Z} = \mathbf{I}_m \otimes \mathbf{1}_n$ . Donc finalement,

$$\begin{aligned} \mathbf{y}|\mathbf{a} &\sim \mathcal{N} \left( \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix} + \mathbf{Z}\mathbf{a}, \sigma^2 \mathbf{I}_N \right) \\ \mathbf{y} &\sim \mathcal{N} \left( \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \mathbf{V} \right) \end{aligned}$$

Où  $\mathbf{V} = \text{Var}(\mathbf{y}) = \mathbf{I}_m \otimes \mathbf{V}_0$ . □

Les propositions suivantes se démontrent en utilisant les arguments des sections A.1 et A.2 des annexes.

**Proposition 2.53.**  $\mathbf{V}^{-1} = \mathbf{I}_m \otimes \left( \frac{1}{\sigma^2} (\mathbf{I}_n - \tau \bar{\mathbf{J}}_n) \right)$

Où  $\tau = \frac{n\sigma_a^2}{\sigma^2 + n\sigma_a^2}$  et  $\bar{\mathbf{J}}_n = \frac{1}{n} \mathbf{J}_n$ .

*Démonstration.* En vertu des propriétés des vecteurs et matrices de uns,  $\mathbf{V}_0^{-1} = \frac{1}{\sigma^2} (\mathbf{I}_n - \frac{\sigma_a^2}{\sigma^2 + n\sigma_a^2} \mathbf{J}_n)$ . Par la suite, remarquons que

$$\tau = \frac{n\sigma_a^2}{\sigma^2 + n\sigma_a^2} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma^2/n} \quad \Rightarrow \quad 1 - \tau = \frac{\sigma^2}{\sigma^2 + n\sigma_a^2} = \frac{\sigma^2}{n\sigma_a^2} \tau$$

En raison des propriétés du produit de Kronecker,  $\mathbf{V}^{-1} = \mathbf{I}_m \otimes \mathbf{V}_0^{-1} = \mathbf{I}_m \otimes \left( \frac{1}{\sigma^2} (\mathbf{I}_n - \tau \bar{\mathbf{J}}_n) \right)$ . □

Aussi,

**Proposition 2.54.**  $|\mathbf{V}| = (\sigma^2)^{m(n-1)} (\sigma^2 + n\sigma_a^2)^m$

*Démonstration.* En raison des propriétés du produit de Kronecker et des vecteurs et matrices de uns,

$$\begin{aligned}
|\mathbf{V}| &= |\mathbf{I}_m|^n |\mathbf{V}_0|^m \\
&= |\sigma^2 \mathbf{I}_n + \sigma_a^2 \mathbf{J}_n|^m \\
&= [(\sigma^2)^{n-1} (\sigma^2 + n\sigma_a^2)]^m \\
&= (\sigma^2)^{m(n-1)} (\sigma^2 + n\sigma_a^2)^m
\end{aligned}
\quad \square$$

Ensuite, la vraisemblance est donnée par

**Proposition 2.55.**

$$L(\mu, \beta, \sigma^2, \sigma_a^2) = \frac{1}{(2\pi)^{N/2}} |\mathbf{V}|^{-1/2} \exp \left( -\frac{1}{2} \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix} \right)^t \mathbf{V}^{-1} \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix} \right) \right)$$

*Démonstration.* La preuve est directe avec la définition de la densité de la loi normale multidimensionnelle. □

Donc, en prenant le log et en développant le terme de l'exponentielle, il en découle que

**Proposition 2.56.**

$$\begin{aligned}
l = \log(L) &= -\frac{N}{2} \log(2\pi) - \frac{1}{2} \log(|\mathbf{V}|) + \frac{1}{2} \boldsymbol{\beta}^t \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} + \frac{1}{2} \mathbf{y}^t \mathbf{V}^{-1} \mathbf{X} \boldsymbol{\beta} \\
&\quad - \frac{1}{2} \boldsymbol{\beta}^t \mathbf{X}^t \mathbf{V}^{-1} \mathbf{X} \boldsymbol{\beta} - \frac{1}{2} \mathbf{y}^t \mathbf{V}^{-1} \mathbf{y}
\end{aligned}$$

Où  $\boldsymbol{\beta} = \begin{bmatrix} \mu \\ \beta \end{bmatrix}$

Estimons premièrement  $\mu$  et  $\beta$

**Proposition 2.57.** *Les estimateurs du maximum de vraisemblance de  $\mu$  et  $\beta$  (en supposant que  $\mathbf{V}$  soit connu) sont*

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} = (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y}$$

*Démonstration.* En vertu des définitions et des propriétés de la dérivation de vecteurs de la section A.3 des annexes, la dérivée de la log-vraisemblance par rapport à  $\beta$  est donnée par

$$\begin{aligned} \frac{\partial l}{\partial \beta} &= \frac{\partial}{\partial \beta} \left( \frac{1}{2} \beta^t \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} \right) + \frac{\partial}{\partial \beta} \left( \frac{1}{2} \mathbf{y}^t \mathbf{V}^{-1} \mathbf{X} \beta \right) - \frac{\partial}{\partial \beta} \left( \frac{1}{2} \beta^t \mathbf{X}^t \mathbf{V}^{-1} \mathbf{X} \beta \right) \\ &= \frac{1}{2} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} + \frac{\partial}{\partial \beta} \left( \frac{1}{2} \beta^t \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} \right) - \frac{1}{2} (2 \mathbf{X}^t \mathbf{V}^{-1} \mathbf{X} \beta) \\ &= \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} - \mathbf{X}^t \mathbf{V}^{-1} \mathbf{X} \beta \end{aligned}$$

Le deuxième égalité provient du fait que  $\frac{1}{2} \mathbf{y}^t \mathbf{V}^{-1} \mathbf{X} \beta$  est un scalaire donc  $\frac{1}{2} \mathbf{y}^t \mathbf{V}^{-1} \mathbf{X} \beta = \left( \frac{1}{2} \mathbf{y}^t \mathbf{V}^{-1} \mathbf{X} \beta \right)^t = \frac{1}{2} \beta^t \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y}$  et que  $\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X}$  est une matrice symétrique.

Maintenant, égalons la dernière expression à 0.

$$\begin{aligned} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} - \mathbf{X}^t \mathbf{V}^{-1} \mathbf{X} \dot{\beta} &= 0 \\ \Rightarrow \mathbf{X}^t \mathbf{V}^{-1} \mathbf{X} \dot{\beta} &= \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} \\ \Rightarrow \dot{\beta} &= (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} \end{aligned}$$

Et, puisque  $\dot{\beta} \in \mathbb{R}^2$ ,  $\hat{\beta} = \dot{\beta}$ . □

Ensuite, remplaçons  $\mathbf{V}$ ,  $\mathbf{X}$  et  $\mathbf{y}$  par leur valeur.

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} \bar{y}_{..} - \hat{\beta}\bar{x} \\ \frac{S_{xy}}{S_{xx}} \end{bmatrix} \quad (2.19)$$

Les détails se retrouvent à la section B.6.1 des annexes.

### La vraisemblance modifiée pour estimer $\sigma^2$ et $\sigma_a^2$

Jusqu'ici, nous avons supposé que les variances  $\sigma^2$  et  $\sigma_a^2$  étaient connues, mais c'est rarement le cas. Il faut donc les estimer et nous allons utiliser la méthode du maximum de vraisemblance comme pour tous les autres paramètres. Même si les estimateurs  $\hat{\mu}$  et  $\hat{\beta}$  sont déjà trouvés, cette technique (maximum de vraisemblance) peut être utilisée puisque les variances ( $\sigma^2$  et  $\sigma_a^2$ ) ne se retrouvent pas dans les formules de  $\hat{\mu}$  et  $\hat{\beta}$ . Les estimateurs des variances peuvent donc être trouvés indépendamment des estimateurs  $\hat{\mu}$  et  $\hat{\beta}$ . Ainsi, nous allons utiliser une log-vraisemblance modifiée notée  $l^* = \log(L(\mu = \hat{\mu}, \beta = \hat{\beta}, \sigma^2, \sigma_a^2))$ .

La log-vraisemblance s'écrit comme suit.

$$l = -\frac{mn}{2} \log(2\pi) - \frac{1}{2} \log(|\mathbf{V}|) - \frac{1}{2} (\mathbf{y} - \boldsymbol{\Theta})^t \mathbf{V}^{-1} (\mathbf{y} - \boldsymbol{\Theta}) \quad (2.20)$$

Où  $\boldsymbol{\Theta} = \mathbb{E}[\mathbf{y}] = \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix}$ .

**Proposition 2.58.** *Développons  $\mathbf{V}$  et  $\mathbf{y} - \boldsymbol{\Theta}$  dans la précédente expression et remplaçons  $\mu$  par  $\hat{\mu}$  et  $\beta$  par  $\hat{\beta}$ .*

$$l^* = -\frac{1}{2} \left[ mn \log(2\pi) + m(n-1) \log(\sigma^2) + m \log(\sigma^2 + n\sigma_a^2) + \frac{S_1^*}{\sigma^2} - \frac{n\sigma_a^2 S_2^*}{\sigma^2(\sigma^2 + n\sigma_a^2)} \right]$$

Où  $S_1^* = \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \hat{\mu} - \hat{\beta}x_j)^2$  et où  $S_2^* = \sum_{i=1}^m n(\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x})^2$ .

*Démonstration.* Les détails des calculs se retrouvent à la section B.6.2 a) des annexes.  $\square$

Remplaçons  $\hat{\beta}$  et  $\hat{\mu}$  par leur valeur dans les expressions de  $S_1^*$  et  $S_2^*$ .

**Proposition 2.59.**  $S_1^* = SCE + SCRR$

*Démonstration.*

$$\begin{aligned}
 S_1^* &= \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \hat{\mu} - \hat{\beta}x_j)^2 \\
 &= \sum_{i=1}^m \sum_{j=1}^n [y_{ij} - \bar{y}_{..} - \hat{\beta}(x_j - \bar{x})]^2 \\
 &= SCT - \hat{\beta}^2 S_{xx} \\
 &= SCE + SCR - SCC \\
 &= SCE + SCRR
 \end{aligned}$$

Car,  $SCT = SCE + SCR$ ,  $SCC = \hat{\beta}^2 S_{xx}$  et  $SCRR = SCR - SCC$ . La dernière égalité est un peu moins évidente, alors les détails se retrouvent à la section B.6.2 b) des annexes.  $\square$

Voici maintenant une proposition pour  $S_2^*$ .

**Proposition 2.60.**  $S_2^* = SCE$

*Démonstration.*

$$\begin{aligned}
 S_2^* &= \sum_{i=1}^m n(\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x})^2 \\
 &= \sum_{i=1}^m n [\bar{y}_i - \bar{y}_{..} - \hat{\beta}(\bar{x} - \bar{x})]^2 \\
 &= \sum_{i=1}^m n(\bar{y}_i - \bar{y}_{..})^2 \\
 &= SCE
 \end{aligned}$$

$\square$

Donc la log-vraisemblance modifiée finale est

$$l^* = -\frac{1}{2} \left[ mn \log(2\pi) + m(n-1) \log(\sigma^2) + m \log(\sigma^2 + n\sigma_a^2) + \frac{SCE + SCRR}{\sigma^2} - \frac{n\sigma_a^2 SCE}{\sigma^2(\sigma^2 + n\sigma_a^2)} \right] \quad (2.21)$$

### e) Les équations du maximum de vraisemblance et leurs solutions

Énonçons ici les équations du maximum de la vraisemblance modifiée pour les variances seulement ( $\sigma^2$  et  $\sigma_a^2$ ) puisque cela a déjà été fait pour  $\mu$  et  $\beta$ .

**Proposition 2.61.** *Les solutions du maximum de la vraisemblance modifiée sont données par*

$$\begin{aligned} \dot{\sigma}^2 &= \frac{SCRR}{m(n-1)} \\ \dot{\sigma}_a^2 &= \frac{1}{n} \left[ \left(1 - \frac{1}{m}\right) CME - CMR \right] + \frac{\hat{\beta}^2 S_{xx}}{mn(n-1)} \end{aligned}$$

*Démonstration.* Lorsque  $l^*$  est dérivée par rapport à  $\sigma^2$  et  $\sigma_a^2$  et que ces expressions sont égalées à 0, il en découle que

$$\frac{\partial l^*}{\partial \sigma_a^2} = 0 \Rightarrow -\frac{mn}{2(\sigma^2 + n\sigma_a^2)} + \frac{nSCE}{2(\sigma^2 + n\sigma_a^2)^2} = 0 \Rightarrow \dot{\sigma}^2 + n\dot{\sigma}_a^2 = \frac{SCE}{m} \quad (2.22)$$

$$\begin{aligned} \frac{\partial l^*}{\partial \sigma^2} = 0 &\Rightarrow \frac{-m(n-1)}{2\sigma^2} - \frac{m}{2(\sigma^2 + n\sigma_a^2)} + \frac{SCE + SCRR}{2\sigma^4} - \frac{n\sigma_a^2 SCE(2\sigma^2 + n\sigma_a^2)}{2\sigma^4(\sigma^2 + n\sigma_a^2)^2} = 0 \\ &\Rightarrow m(n-1)\dot{\sigma}^2(\dot{\sigma}^2 + n\dot{\sigma}_a^2)^2 + m\dot{\sigma}^4(\dot{\sigma}^2 + n\dot{\sigma}_a^2) + n\dot{\sigma}_a^2 SCE(2\dot{\sigma}^2 + n\dot{\sigma}_a^2) \\ &= (SCE + SCRR)(\dot{\sigma}^2 + n\dot{\sigma}_a^2)^2 \end{aligned} \quad (2.23)$$



Si nous utilisons l'équation 2.22 pour simplifier l'équation 2.23, il s'ensuit que

$$\begin{aligned}
&\Rightarrow \frac{n-1}{m} \dot{\sigma}^2 SCE^2 + \dot{\sigma}^4 SCE + n \dot{\sigma}_a^2 SCE \left( \dot{\sigma}^2 + \frac{SCE}{m} \right) = (SCE + SCRR) \left( \frac{SCE}{m} \right)^2 \\
&\Rightarrow \frac{n-1}{m} \dot{\sigma}^2 SCE + \dot{\sigma}^4 + \left( \frac{SCE}{m} - \dot{\sigma}^2 \right) \left( \dot{\sigma}^2 + \frac{SCE}{m} \right) = \frac{SCE^2}{m^2} + \frac{SCRR SCE}{m^2} \\
&\Rightarrow \frac{n-1}{m} \dot{\sigma}^2 SCE = \frac{SCRR SCE}{m^2} \\
&\Rightarrow \dot{\sigma}^2 = \frac{SCRR}{m(n-1)}
\end{aligned}$$

Donc,

$$\dot{\sigma}_a^2 = \frac{1}{n} \left( \frac{SCE}{m} - \dot{\sigma}^2 \right) = \frac{1}{n} \left[ \left( 1 - \frac{1}{m} \right) CME - CMR \right] + \frac{\hat{\beta}^2 S_{xx}}{mn(n-1)} \quad \square$$

#### f) Les estimateurs du maximum de vraisemblance

Les estimateurs du maximum de vraisemblance sont donnés par

**Proposition 2.62.**

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} \bar{y}_{..} - \hat{\beta} \bar{x} \\ \frac{S_{xy}}{S_{xx}} \end{bmatrix}$$

$$\hat{\sigma}^2 = \begin{cases} \frac{SCRR}{m(n-1)} & si \left( 1 - \frac{1}{m} \right) CME + \frac{\hat{\beta}^2 S_{xx}}{m(n-1)} \geq CMR \\ \frac{SCE+SCRR}{mn} & si \left( 1 - \frac{1}{m} \right) CME + \frac{\hat{\beta}^2 S_{xx}}{m(n-1)} < CMR \end{cases}$$

$$\hat{\sigma}_a^2 = \begin{cases} \frac{(1-1/m)CME-CMR}{n} + \frac{\hat{\beta}^2 S_{xx}}{mn(n-1)} & si \left( 1 - \frac{1}{m} \right) CME + \frac{\hat{\beta}^2 S_{xx}}{m(n-1)} \geq CMR \\ 0 & si \left( 1 - \frac{1}{m} \right) CME + \frac{\hat{\beta}^2 S_{xx}}{m(n-1)} < CMR \end{cases}$$

*Démonstration.* Premièrement, pour  $\hat{\mu}$  et pour  $\hat{\beta}$ , il n'y a pas de problème de domaine et ils ne changent pas en fonction des valeurs que vont prendre les estimateurs des variances donc les estimateurs sont donnés par les solutions. Deuxièmement, remarquons que  $\hat{\sigma}_a^2$

peut être négatif (mais pas  $\hat{\sigma}^2$ ). Il faut donc procéder différemment que dans le cas fixe. En effet, pour trouver les estimateurs du maximum de vraisemblance pour les variances, nous procéderons de la même façon que dans le cas de l'analyse de la variance à un facteur aléatoire. C'est-à-dire que si  $\dot{\sigma}_a^2 < 0$  alors  $\hat{\sigma}_a^2 = 0$  et  $\hat{\sigma}^2$  est égal à la valeur satisfaisant le maximum de vraisemblance avec  $\sigma_a^2 = 0$  ( $\dot{\sigma}_0^2$ ). La vraisemblance modifiée dans le cas où  $\dot{\sigma}_a^2 < 0$  devient,

$$l^* = l(\hat{\mu}, \hat{\beta}, \sigma, \sigma_a^2 = 0) = -\frac{1}{2} \left[ mn \log(2\pi) + mn \log(\sigma^2) + \frac{SCE + SCRR}{\sigma^2} \right]$$

Donc la dérivée partielle par rapport à  $\sigma^2$  est donnée par

$$\begin{aligned} \frac{\partial l^*}{\partial \sigma^2} = 0 &\Rightarrow -\frac{mn}{2\sigma^2} + \frac{SCE + SCRR}{2\sigma^4} = 0 \\ &\Rightarrow \dot{\sigma}_0^2 = \frac{SCE + SCRR}{mn} \end{aligned}$$

Ce qui donne le résultat souhaité. □

### g) Espérance et biais des estimateurs

Il y a les propositions suivantes sur les espérances des estimateurs de  $\mu$  et de  $\beta$ .

**Proposition 2.63.**  *$\hat{\mu}$  et  $\hat{\beta}$  sont sans biais.*

*Démonstration.*

$$\begin{aligned}
\mathbb{E}[\hat{\beta}] &= \mathbb{E} \left[ \frac{\left( \sum_{j=1}^n x_j \bar{y}_{\cdot j} \right) - n \bar{x} \bar{y}_{\cdot}}{\left( \sum_{j=1}^n x_j^2 \right) - n \bar{x}^2} \right] \\
&= \frac{\sum_{j=1}^n \left[ x_j \left( \frac{1}{m} \sum_{i=1}^m \mathbb{E}[y_{ij}] \right) \right] - n \bar{x} \left( \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \mathbb{E}[y_{ij}] \right)}{\left( \sum_{j=1}^n x_j^2 \right) - n \bar{x}^2} \\
&= \frac{\sum_{j=1}^n \left[ x_j \left( \frac{1}{m} \sum_{i=1}^m \mu + \beta x_j \right) \right] - n \bar{x} \left( \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \mu + \beta x_j \right)}{\left( \sum_{j=1}^n x_j^2 \right) - n \bar{x}^2} \\
&= \frac{n \bar{x} \mu + \beta \left( \sum_{j=1}^n x_j^2 \right) - n \bar{x} \mu - n \beta \bar{x}^2}{\left( \sum_{j=1}^n x_j^2 \right) - n \bar{x}^2} \\
&= \beta
\end{aligned}$$

Et

$$\begin{aligned}
\mathbb{E}[\hat{\mu}] &= \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \mathbb{E}[y_{ij}] - \mathbb{E}[\hat{\beta}] \bar{x} \\
&= \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (\mu + \beta x_j) - \beta \bar{x} \\
&= \mu + \beta \bar{x} - \beta \bar{x} \\
&= \mu
\end{aligned}$$

□

Pour  $\sigma^2$  et  $\sigma_a^2$ , puisque leur estimateur n'a pas de forme explicite (ils sont définis en deux options), il est beaucoup plus compliqué de trouver leur espérance. Par contre, leur loi asymptotique sera décrite plus loin.

## h) Variance des estimateurs (et leur loi)

### Pour $\hat{\mu}$ et $\hat{\beta}$

Premièrement, il est évident que le vecteur  $[\hat{\mu}, \hat{\beta}]^t$  suit une loi normale multidimensionnelle à cause de sa forme. En effet le vecteur  $[\hat{\mu}, \hat{\beta}]^t$  est une transformation affine de  $\mathbf{y}$ , qui lui suit une loi normale multidimensionnelle. La moyenne de ces estimateurs a déjà été trouvée, il ne reste donc qu'à trouver leur matrice de variance-covariance. Voici donc une propriété concernant la covariance d'une transformation affine linéaire d'une variable aléatoire qui suit une loi normale multidimensionnelle.

**Proposition 2.64.** Si  $\mathbf{x} \sim \mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ,  $\mathbf{A}$  une matrice  $q \times p$  et  $\mathbf{b}$  un vecteur de longueur  $q$ , alors  $\text{Var}(\mathbf{Ax} + \mathbf{b}) = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^t$

*Démonstration.* Sachant que la fonction génératrice des moments d'une variable aléatoire qui suit une loi normale multidimensionnelle est la suivante

$$M_{\mathbf{x}}(\mathbf{t}) = \exp\left(\mathbf{t}^t \boldsymbol{\mu} + \frac{1}{2} \mathbf{t}^t \boldsymbol{\Sigma} \mathbf{t}\right)$$

Posons  $\mathbf{y} = \mathbf{Ax} + \mathbf{b}$ , il en découle que

$$\begin{aligned} M_{\mathbf{y}}(\mathbf{t}) &= \mathbb{E}[e^{\mathbf{t}^t \mathbf{y}}] \\ &= \mathbb{E}[e^{\mathbf{t}^t (\mathbf{Ax} + \mathbf{b})}] \\ &= e^{\mathbf{t}^t \mathbf{b}} \mathbb{E}[e^{(\mathbf{A}^t \mathbf{t})^t \mathbf{x}}] \\ &= e^{\mathbf{t}^t \mathbf{b}} M_{\mathbf{x}}(\mathbf{A}^t \mathbf{t}) \end{aligned}$$

$$\begin{aligned}
&= \exp \left( \mathbf{t}^t \mathbf{b} + \mathbf{t}^t \mathbf{A} \boldsymbol{\mu} + \frac{1}{2} \mathbf{t}^t \mathbf{A} \boldsymbol{\Sigma} \mathbf{A}^t \mathbf{t} \right) \\
&= \exp \left( \mathbf{t}^t (\mathbf{A} \boldsymbol{\mu} + \mathbf{b}) + \frac{1}{2} \mathbf{t}^t (\mathbf{A} \boldsymbol{\Sigma} \mathbf{A}^t) \mathbf{t} \right) \\
&\Rightarrow \mathbf{y} \sim \mathcal{N}(\mathbf{A} \boldsymbol{\mu} + \mathbf{b}, \mathbf{A} \boldsymbol{\Sigma} \mathbf{A}^t) \quad \square
\end{aligned}$$

En se servant de la dernière proposition, la matrice de variance-covariance du vecteur  $[\hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\beta}}]^t$  est donnée par

**Proposition 2.65.**

$$\text{Var} \left( \begin{bmatrix} \hat{\boldsymbol{\mu}} \\ \hat{\boldsymbol{\beta}} \end{bmatrix} \right) = \begin{pmatrix} \frac{\sigma^2 + n\sigma_a^2}{mn} + \frac{\bar{x}^2 \sigma^2}{S_{xx}} & -\frac{\bar{x} \sigma^2}{S_{xx}} \\ -\frac{\bar{x} \sigma^2}{S_{xx}} & \frac{\sigma^2}{S_{xx}} \end{pmatrix}$$

*Démonstration.*

$$\begin{aligned}
\text{Var} \left( \begin{bmatrix} \hat{\boldsymbol{\mu}} \\ \hat{\boldsymbol{\beta}} \end{bmatrix} \right) &= \text{Var} \left( (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} \right) \\
&= ((\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1}) \text{Var}(\mathbf{y}) ((\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1})^t \\
&= (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{V} \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \\
&= (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \\
&= (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \\
&= \frac{1}{m} (\mathbf{X}_0^t \mathbf{V}_0^{-1} \mathbf{X}_0)^{-1} \\
&= \frac{1}{m} \frac{m \sigma^2}{S_{xx}} \begin{pmatrix} \frac{\sigma_a^2 S_{xx}}{m \sigma^2 \tau} + \bar{x}^2 & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix} \\
&= \begin{pmatrix} \frac{\sigma^2 + n\sigma_a^2}{mn} + \frac{\bar{x}^2 \sigma^2}{S_{xx}} & -\frac{\bar{x} \sigma^2}{S_{xx}} \\ -\frac{\bar{x} \sigma^2}{S_{xx}} & \frac{\sigma^2}{S_{xx}} \end{pmatrix}
\end{aligned}$$

L'avant dernière équation vient de l'expression que prend  $(\mathbf{X}_0^t \mathbf{V}_0^{-1} \mathbf{X}_0)^{-1}$  lorsque  $\mathbf{X}_0$  et  $\mathbf{V}_0^{-1}$  sont remplacés par leur valeur. Pour plus de détails sur  $(\mathbf{X}_0^t \mathbf{V}_0^{-1} \mathbf{X}_0)^{-1}$ , voir la section B.6.1 des annexes (dans les calculs pour trouver une formule plus simple pour  $\hat{\boldsymbol{\mu}}$  et  $\hat{\boldsymbol{\beta}}$ ).  $\square$

Donc la loi du vecteur  $[\hat{\mu}, \hat{\beta}]^t$  est donnée par

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \begin{pmatrix} \frac{\sigma^2 + n\sigma_a^2}{mn} + \frac{\bar{x}^2 \sigma^2}{S_{xx}} & -\frac{\bar{x} \sigma^2}{S_{xx}} \\ -\frac{\bar{x} \sigma^2}{S_{xx}} & \frac{\sigma^2}{S_{xx}} \end{pmatrix} \right) \quad (2.24)$$

### Pour $\hat{\sigma}^2$ et $\hat{\sigma}_a^2$

Pour  $\hat{\sigma}^2$  et  $\hat{\sigma}_a^2$ , leur loi ne peut pas s'écrire de façon analytique à cause de la forme en deux temps que prennent ces estimateurs. Par contre, il est possible d'utiliser le résultat sur la loi asymptotique des estimateurs du maximum de vraisemblance.

$$\begin{bmatrix} \hat{\sigma}^2 \\ \hat{\sigma}_a^2 \end{bmatrix} \approx \mathcal{N} \left( \begin{bmatrix} \sigma^2 \\ \sigma_a^2 \end{bmatrix}, \left( -\mathbb{E} \begin{pmatrix} l_{\sigma^2 \sigma^2} & l_{\sigma^2 \sigma_a^2} \\ l_{\sigma_a^2 \sigma^2} & l_{\sigma_a^2 \sigma_a^2} \end{pmatrix} \right)^{-1} \right) \quad (2.25)$$

Où  $l$  est la log-vraisemblance non modifiée qui est donnée par

$$l = -\frac{1}{2} \left[ mn \log(2\pi) + m(n-1) \log(\sigma^2) + m \log(\sigma^2 + n\sigma_a^2) + \frac{S_1}{\sigma^2} - \frac{n\sigma_a^2 S_2}{\sigma^2(\sigma^2 + n\sigma_a^2)} \right] \quad (2.26)$$

Où  $S_1 = \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \mu - \beta x_j)^2$  et  $S_2 = \sum_{i=1}^m n(\bar{y}_i - \mu - \beta \bar{x})^2$ .

En effectuant les calculs, la loi des estimateurs des variances est donnée par

### **Proposition 2.66.**

$$\begin{bmatrix} \hat{\sigma}^2 \\ \hat{\sigma}_a^2 \end{bmatrix} \approx \mathcal{N} \left( \begin{bmatrix} \sigma^2 \\ \sigma_a^2 \end{bmatrix}, \frac{1}{D} \begin{pmatrix} \frac{mn^2}{2(\sigma^2 + n\sigma_a^2)^2} & -\frac{mn}{2(\sigma^2 + n\sigma_a^2)^2} \\ -\frac{mn}{2(\sigma^2 + n\sigma_a^2)^2} & \frac{m}{2(\sigma^2 + n\sigma_a^2)^2} + \frac{N-m}{2\sigma^4} \end{pmatrix} \right)$$

Où  $D = \frac{m^2 n^2 (n-1)}{4\sigma^4 (\sigma^2 + n\sigma_a^2)^2}$

Ou encore,

$$\begin{bmatrix} \hat{\sigma}^2 \\ \hat{\sigma}_a^2 \end{bmatrix} \approx \mathcal{N} \left( \begin{bmatrix} \sigma^2 \\ \sigma_a^2 \end{bmatrix}, 2\sigma^4 \begin{pmatrix} \frac{1}{m(n-1)} & -\frac{1}{mn(n-1)} \\ -\frac{1}{mn(n-1)} & \frac{1}{n^2} \left( \frac{1}{m(n-1)} + \frac{(\sigma^2 + n\sigma_a^2)^2}{m\sigma^4} \right) \end{pmatrix} \right)$$

Qui est la même loi asymptotique que dans le cas de l'ANOVA à un facteur aléatoire, voir proposition 2.23.

*Démonstration.*

$$\begin{aligned}
\frac{\partial l}{\partial \sigma_a^2} &= -\frac{mn}{2(\sigma^2 + n\sigma_a^2)} + \frac{nS_2}{2(\sigma^2 + n\sigma_a^2)^2} \\
\frac{\partial l}{\partial \sigma^2} &= -\frac{m(n-1)}{2\sigma^2} - \frac{m}{2(\sigma^2 + n\sigma_a^2)} + \frac{S_1}{2\sigma^4} - \frac{n\sigma_a^2 S_2 (2\sigma^2 + n\sigma_a^2)}{2\sigma^4 (\sigma^2 + n\sigma_a^2)^2} \\
\Rightarrow \frac{\partial^2 l}{(\partial \sigma_a^2)^2} &= \frac{mn^2}{2(\sigma^2 + n\sigma_a^2)^2} - \frac{n^2 S_2}{(\sigma^2 + n\sigma_a^2)^3} \\
\frac{\partial^2 l}{\partial \sigma_a^2 \partial \sigma^2} &= \frac{mn}{2(\sigma^2 + n\sigma_a^2)^2} - \frac{nS_2}{(\sigma^2 + n\sigma_a^2)^3} \\
\frac{\partial^2 l}{(\partial \sigma^2)^2} &= \frac{m(n-1)}{2\sigma^4} + \frac{m}{2(\sigma^2 + n\sigma_a^2)^2} - \frac{S_1}{\sigma^6} - \frac{S_2}{(\sigma^2 + n\sigma_a^2)^3} + \frac{S_2}{\sigma^6}
\end{aligned}$$

De plus,

$$\begin{aligned}
\mathbb{E}[S_1] &= \sum_{i=1}^m \sum_{j=1}^n \mathbb{E}[(y_{ij} - \mu - \beta x_j)^2] \\
&= \sum_{i=1}^m \sum_{j=1}^n \mathbb{E}[y_{ij}^2] - 2(\mu + \beta x_j) \mathbb{E}[y_{ij}] + (\mu + \beta x_j)^2 \\
&= \sum_{i=1}^m \sum_{j=1}^n \sigma_a^2 + \sigma^2 + (\mu + \beta x_j)^2 - 2(\mu + \beta x_j)(\mu + \beta x_j) + (\mu + \beta x_j)^2 \\
&= N(\sigma_a^2 + \sigma^2)
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}[S_2] &= \sum_{i=1}^m n \left( \mathbb{E}[(\bar{y}_i - \mu - \beta \bar{x})^2] \right) \\
&= \sum_{i=1}^m n \left( \mathbb{E}[\bar{y}_i^2] - 2(\mu + \beta \bar{x}) \mathbb{E}[\bar{y}_i] + (\mu + \beta \bar{x})^2 \right)
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^m n \left( \frac{\sigma^2}{n} + \sigma_a^2 + (\mu + \beta \bar{x})^2 - 2(\mu + \beta \bar{x})(\mu + \beta \bar{x}) + (\mu + \beta \bar{x})^2 \right) \\
&= m(\sigma^2 + n\sigma_a^2)
\end{aligned}$$

$$\text{Car } \bar{y}_i \sim \mathcal{N} \left( \mu + \beta \bar{x}, \frac{\sigma^2}{n} + \sigma_a^2 \right).$$

$$\begin{aligned}
\Rightarrow -\mathbb{E} \left[ \frac{\partial l}{\partial \sigma_a^2 \sigma_a^2} \right] &= -\frac{mn^2}{2(\sigma^2 + n\sigma_a^2)^2} + \frac{n^2}{(\sigma^2 + n\sigma_a^2)^3} \mathbb{E}[S_2] \\
&= -\frac{mn^2}{2(\sigma^2 + n\sigma_a^2)^2} + \frac{mn^2(\sigma^2 + n\sigma_a^2)}{(\sigma^2 + n\sigma_a^2)^3} \\
&= \frac{mn^2}{2(\sigma^2 + n\sigma_a^2)^2}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow -\mathbb{E} \left[ \frac{\partial l}{\partial \sigma_a^2 \sigma^2} \right] &= -\frac{mn}{2(\sigma^2 + n\sigma_a^2)^2} + \frac{n}{(\sigma^2 + n\sigma_a^2)^3} \mathbb{E}[S_2] \\
&= -\frac{mn}{2(\sigma^2 + n\sigma_a^2)^2} + \frac{mn(\sigma^2 + n\sigma_a^2)}{(\sigma^2 + n\sigma_a^2)^3} \\
&= \frac{mn}{2(\sigma^2 + n\sigma_a^2)^2}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow -\mathbb{E} \left[ \frac{\partial l}{\partial \sigma^2 \sigma^2} \right] &= -\frac{m(n-1)}{2\sigma^4} - \frac{m}{2(\sigma^2 + n\sigma_a^2)^2} + \frac{1}{\sigma^6} \mathbb{E}[S_1] + \frac{1}{(\sigma^2 + n\sigma_a^2)^3} \mathbb{E}[S_2] - \frac{1}{\sigma^6} \mathbb{E}[S_2] \\
&= -\frac{m(n-1)}{2\sigma^4} - \frac{m}{2(\sigma^2 + n\sigma_a^2)^2} + \frac{N(\sigma^2 + \sigma_a^2)}{\sigma^6} + \frac{m(\sigma^2 + n\sigma_a^2)}{(\sigma^2 + n\sigma_a^2)^3} - \frac{m(\sigma^2 + n\sigma_a^2)}{\sigma^6} \\
&= -\frac{m(n-1)}{2\sigma^4} + \frac{m}{2(\sigma^2 + n\sigma_a^2)^2} + \frac{N\sigma^2 + N\sigma_a^2 - m\sigma^2 - N\sigma_a^2}{\sigma^6} \\
&= -\frac{m(n-1)}{2\sigma^4} + \frac{m}{2(\sigma^2 + n\sigma_a^2)^2} + \frac{N-m}{\sigma^4} \\
&= \frac{m}{2(\sigma^2 + n\sigma_a^2)^2} + \frac{N-m}{2\sigma^4}
\end{aligned}$$



$$\begin{aligned} \Rightarrow -\mathbb{E} \begin{pmatrix} l_{\sigma^2\sigma^2} & l_{\sigma^2\sigma_a^2} \\ l_{\sigma^2\sigma_a^2} & l_{\sigma_a^2\sigma_a^2} \end{pmatrix} &= \begin{pmatrix} \frac{m}{2(\sigma^2+n\sigma_a^2)^2} + \frac{N-m}{2\sigma^4} & \frac{mn}{2(\sigma^2+n\sigma_a^2)^2} \\ \frac{mn}{2(\sigma^2+n\sigma_a^2)^2} & \frac{mn^2}{2(\sigma^2+n\sigma_a^2)^2} \end{pmatrix} \\ \Rightarrow \left( -\mathbb{E} \begin{pmatrix} l_{\sigma^2\sigma^2} & l_{\sigma^2\sigma_a^2} \\ l_{\sigma^2\sigma_a^2} & l_{\sigma_a^2\sigma_a^2} \end{pmatrix} \right)^{-1} &= \frac{1}{D} \begin{pmatrix} \frac{mn^2}{2(\sigma^2+n\sigma_a^2)^2} & -\frac{mn}{2(\sigma^2+n\sigma_a^2)^2} \\ -\frac{mn}{2(\sigma^2+n\sigma_a^2)^2} & \frac{m}{2(\sigma^2+n\sigma_a^2)^2} + \frac{N-m}{2\sigma^4} \end{pmatrix} \end{aligned}$$

Où  $D = \frac{m^2n^2(n-1)}{4\sigma^4(\sigma^2+n\sigma_a^2)^2}$

Car

$$\begin{aligned} D &= \left( \frac{m}{2(\sigma^2+n\sigma_a^2)^2} + \frac{N-m}{2\sigma^4} \right) \left( \frac{mn^2}{2(\sigma^2+n\sigma_a^2)^2} \right) - \left( \frac{mn}{2(\sigma^2+n\sigma_a^2)^2} \right) \left( \frac{mn}{2(\sigma^2+n\sigma_a^2)^2} \right) \\ &= \frac{m^2n^2}{4(\sigma^2+n\sigma_a^2)^4} + \frac{m^2n^2(n-1)}{4\sigma^4(\sigma^2+n\sigma_a^2)^2} - \frac{m^2n^2}{4(\sigma^2+n\sigma_a^2)^4} \\ &= \frac{m^2n^2(n-1)}{4\sigma^4(\sigma^2+n\sigma_a^2)^2} \quad \square \end{aligned}$$

### i) Tests d'hypothèse et intervalles de confiance

Nous allons tester deux hypothèses d'intérêt. Premièrement, nous voulons savoir s'il y a une relation linéaire significative entre les observations et les valeurs de la variable explicative indépendamment des groupes ( $H_0 : \beta = 0$ ). Deuxièmement, nous voulons tester s'il y a une différence significative entre les groupes indépendamment de la variable explicative ( $H_0 : \sigma_a^2 = 0$ ). Le test du rapport de vraisemblance sera utilisé.

$$-2\log(\Lambda) = -2l(\hat{\theta}_0) + 2l(\hat{\theta})$$

### La vraisemblance modifiée maximisée

Trouvons premièrement  $l(\hat{\theta})$  en supposant que  $\sigma_a^2 \geq 0$  (ce qui est vrai si  $N \rightarrow \infty$ ). Il

faut donc remplacer  $\sigma^2$  par  $\hat{\sigma}^2 \left( \frac{SCRR}{m(n-1)} \right)$  et  $\sigma_a^2$  par  $\hat{\sigma}_a^2 \left( \frac{(1-\frac{1}{m})CME-CMR}{n} + \frac{\hat{\beta}^2 S_{xx}}{mn(n-1)} \right)$  dans  $l^*(\hat{\mu}, \hat{\beta}, \sigma^2, \sigma_a^2)$ , il en découle que

**Proposition 2.67.**  $-2l(\hat{\theta}) = f_1(m, n) + m(n-1) \log(SCRR) + m \log(SCE)$

Où  $f_1(m, n)$  est une fonction ne dépendant que de  $m$  et  $n$ .

*Démonstration.*

$$\begin{aligned}
-2l(\hat{\theta}) &= mn \log(2\pi) + m(n-1) \log \left( \frac{SCRR}{m(n-1)} \right) + m \log \left( \frac{SCE}{m} \right) \\
&\quad + \frac{SCE + SCRR}{SCRR/(m(n-1))} - \frac{SCE/m - SCRR/(m(n-1))}{SCRR/(m(n-1))} \frac{SCE}{SCE/m} \\
&= mn \log(2\pi) + m(n-1) \log \left( \frac{SCRR}{m(n-1)} \right) + m \log \left( \frac{SCE}{m} \right) \\
&\quad + m(n-1) \frac{SCE + SCRR}{SCRR} - m \frac{(n-1)SCE - SCRR}{SCRR} \\
&= mn \log(2\pi) + m(n-1) \log \left( \frac{SCRR}{m(n-1)} \right) + m \log \left( \frac{SCE}{m} \right) \\
&\quad + \frac{m(n-1)SCE + m(n-1)SCRR - m(n-1)SCE + mSCRR}{SCRR} \\
&= mn \log(2\pi) + m(n-1) \log \left( \frac{SCRR}{m(n-1)} \right) + m \log \left( \frac{SCE}{m} \right) + mn \\
&= mn \log(2\pi) + m(n-1) \log(SCRR) - m(n-1) \log(m(n-1)) \\
&\quad + m \log(SCE) - m \log(m) + mn \\
&= f_1(m, n) + m(n-1) \log(SCRR) + m \log(SCE)
\end{aligned}$$

Où  $f_1(m, n) = mn \log(2\pi) - m(n-1) \log(m(n-1)) - m \log(m) + mn$ . □

### Tester l'hypothèse nulle $H_0 : \sigma_a^2 = 0$

Trouvons maintenant  $l(\hat{\theta}_0)$  qui est la vraisemblance évaluée en  $\mu = \hat{\mu}_0 = \hat{\mu}$ ,  $\beta = \hat{\beta}_0 = \hat{\beta}$ ,  $\sigma^2 = \hat{\sigma}_0^2 = \frac{SCE+SCRR}{mn}$  et  $\sigma_a^2 = 0$ , ce qui revient à remplacer  $\sigma^2$  par  $\frac{SCE+SCRR}{mn}$  dans la vraisemblance de la preuve de la proposition 2.62.

**Proposition 2.68.**  $-2l(\hat{\theta}_0) = f_2(m, n) + mn \log(SCE + SCRR)$

Où  $f_2(m, n)$  est une fonction ne dépendant que de  $m$  et  $n$ .

*Démonstration.*

$$\begin{aligned} -2l(\hat{\theta}_0) &= mn \log(2\pi) + mn \log\left(\frac{SCE + SCRR}{mn}\right) + \frac{SCE + SCRR}{(SCE + SCRR)/mn} \\ &= mn \log(2\pi) + mn \log(SCE + SCRR) - mn \log(mn) + mn \\ &= f_2(m, n) + mn \log(SCE + SCRR) \end{aligned}$$

Où  $f_2(m, n) = mn \log(2\pi) - mn \log(mn) + mn$ . □

Et donc,

**Proposition 2.69.**  $-2 \log(\Lambda) = mn \log\left(\frac{SCE}{SCRR} + 1\right) - m \log\left(\frac{SCE}{SCRR}\right)$  en ne tenant pas compte de  $f_1$  et  $f_2$ .

*Démonstration.*

$$\begin{aligned} -2 \log(\Lambda) &= mn \log(SCE + SCRR) - m(n-1) \log(SCRR) - m \log(SCE) \\ &= mn \log\left(\frac{SCE + SCRR}{SCRR}\right) - m \log\left(\frac{SCE}{SCRR}\right) \\ &= mn \log\left(\frac{SCE}{SCRR} + 1\right) - m \log\left(\frac{SCE}{SCRR}\right) \end{aligned}$$

En ne tenant pas compte de  $f_1$  et  $f_2$ . □

Il y a donc la proposition suivante.

**Proposition 2.70.** *Le test du rapport de vraisemblance est équivalent au test relié à la statistique de Fisher suivante*

$$F = \frac{CME}{CMRR} \sim \mathcal{F}_{N-m-1}^{m-1}$$

C'est aussi la statistique de Fisher utilisée pour l'ANCOVA pour l'hypothèse nulle  $H_0 : \alpha_i = 0 \forall i$ .

*Démonstration.* Posons  $q = \frac{SCE}{SCRR} = \frac{m-1}{m(n-1)-1}F$  où  $F = \frac{CME}{CMRR}$ . Il s'ensuit que (en ne tenant pas compte de  $f_1$  et  $f_2$ )

$$-2 \log(\Lambda) = m(n \log(q+1) - \log(q))$$

Pour étudier la monotonie de  $-2 \log(\Lambda)$  par rapport à  $q$ , considérons

$$\frac{\partial}{\partial q}(-2 \log(\Lambda)) = m \left( \frac{n}{q+1} - \frac{1}{q} \right) = \frac{m}{q(q+1)} [(n-1)q - 1]$$

Et,

$$\begin{aligned} \frac{m}{q(q+1)} [(n-1)q - 1] > 0 &\Leftrightarrow q > \frac{1}{n-1} \\ &\Leftrightarrow SCE > \frac{SCRR}{n-1} \\ &\Leftrightarrow \hat{\sigma}_a^2 > 0 \\ &\Leftrightarrow \hat{\sigma}_a^2 > 0 \end{aligned}$$

Car  $\hat{\sigma}_a^2 = \frac{1}{n} \left( \frac{SCE}{m} - \frac{SCRR}{m(n-1)} \right)$ . Donc les deux statistiques varient dans le même sens pour tout le domaine de  $\hat{\sigma}_a^2$  sauf pour un point (0), mais cela est suffisant pour l'équivalence des statistiques. Ce dernier résultat montre que le test de Fisher de l'analyse de la covariance ( $F = \frac{CME}{CMRR}$ ) est le test du rapport de vraisemblance dans le contexte de la régression linéaire simple avec un effet aléatoire (données équilibrées) pour  $H_0 : \sigma_a^2 = 0$ .

De plus, par les mêmes calculs que pour l'ANCOVA avec un facteur fixe, il est possible de montrer que sous  $H_0 : \sigma_a^2 = 0$ ,  $\frac{SCE}{\sigma^2} \sim \chi_{(m-1)}^2$  et est indépendant de  $\frac{SCRR}{\sigma^2} \sim \chi_{(N-m-1)}^2$ .

Ceci implique que sous  $H_0 : \sigma_a^2 = 0$ ,  $F = \frac{CME}{CMRR} \sim \mathcal{F}_{N-m-1}^{m-1}$ . □

### Tester l'hypothèse nulle $H_0 : \beta = 0$

Procédons de façon similaire pour  $H_0 : \beta = 0$ .  $l(\hat{\theta})$  reste la même et pour  $l(\hat{\theta}_0)$ , c'est maintenant  $\hat{\beta} = 0$  ce qui donne que  $\hat{\mu}_0 = \bar{y}..$ ,  $S_{10}^* = SCT$  et  $S_{20}^* = S_2^* = SCE$ . Il en découle que  $\dot{\sigma}_0^2 + n\dot{\sigma}_{a0}^2 = SCE/m$  et  $\dot{\sigma}_0^2 = SCR/m(n-1)$

Ainsi,

**Proposition 2.71.**  $-2l(\hat{\theta}_0) = f_3(m, n) + m(n-1) \log(SCR) + m \log(SCE)$

Où  $f_3(m, n)$  est une fonction ne dépendant que de  $m$  et  $n$ .

*Démonstration.*

$$\begin{aligned}
-2l(\hat{\theta}_0) &= mn \log(2\pi) + m(n-1) \log\left(\frac{SCR}{m(n-1)}\right) + m \log\left(\frac{SCE}{m}\right) \\
&\quad + \frac{SCT}{SCR/(m(n-1))} - \frac{SCE/m - SCR/(m(n-1))}{SCR/(m(n-1))} \frac{SCE}{SCE/m} \\
&= mn \log(2\pi) + m(n-1) \log\left(\frac{SCR}{m(n-1)}\right) + m \log\left(\frac{SCE}{m}\right) \\
&\quad + m(n-1) \frac{SCE}{SCR} + m(n-1) - m \left(\frac{(n-1)SCE}{SCR} - 1\right) \\
&= mn \log(2\pi) + m(n-1) \log\left(\frac{SCR}{m(n-1)}\right) + m \log\left(\frac{SCE}{m}\right) + mn \\
&= f_3(m, n) + m(n-1) \log(SCR) + m \log(SCE)
\end{aligned}$$

Où  $f_3(m, n) = mn \log(2\pi) - m(n-1) \log(m(n-1)) - m \log(m) + mn$ . □

Ainsi,

**Proposition 2.72.**  $-2\log(\Lambda) = -m(n-1) \log\left(1 - \frac{SCC}{SCR}\right)$  en ne tenant pas compte de  $f_1$  et  $f_3$ .

*Démonstration.*

$$\begin{aligned}
-2\log(\Lambda) &= m(n-1)\log(SCR) + m\log(SCE) - m(n-1)\log(SCRR) - m\log(SCE) \\
&= -m(n-1)\log\left(\frac{SCRR}{SCR}\right) \\
&= -m(n-1)\log\left(1 - \frac{SCC}{SCR}\right)
\end{aligned}$$

En ne tenant pas compte de  $f_1$  et  $f_3$ . □

Voici donc une proposition sur le test du rapport de vraisemblance.

**Proposition 2.73.** *Le test du rapport de vraisemblance est équivalent au test relié à la statistique de Fisher suivante*

$$F = \frac{CMC}{CMRR} \sim \mathcal{F}_{N-m-1}^1$$

*Qui est aussi la statistique de Fisher utilisée pour l'ANCOVA pour l'hypothèse nulle  $H_0 : \beta = 0$ .*

*Démonstration.* Posons  $q_b = \frac{SCC}{SCR} = \frac{F}{m(n-1)-1+F}$  où  $F = \frac{CMC}{CMRR}$ . Il en découle que

$$-2\log(\Lambda) = -m(n-1)\log(1 - q_b)$$

Pour étudier la monotonie de  $-2\log(\Lambda)$  par rapport à  $q_b$ , considérons,

$$\frac{\partial}{\partial q_b}(-2\log(\Lambda)) = \frac{m(n-1)}{1 - q_b}$$

Et,

$$\begin{aligned}
\frac{m(n-1)}{1-q_b} > 0 &\Leftrightarrow 1-q_b > 0 \\
&\Leftrightarrow q_b < 1 \\
&\Leftrightarrow \frac{SCC}{SCR} < 1 \\
&\Leftrightarrow SCR - SCC > 0 \\
&\Leftrightarrow SCRR > 0 \text{ ce qui est toujours vrai}
\end{aligned}$$

Aussi, puisqu'ici  $q_b$  n'est pas seulement un multiple de  $F$ , il faut vérifier qu'ils varient dans le même sens ( $F = \frac{(m(n-1)-1)q_b}{1-q_b}$ ).

$$\frac{\partial F}{\partial q_b} = \frac{(m(n-1)-1)(1-q_b) + (m(n-1)-1)q_b}{(1-q_b)^2} = \frac{m(n-1)-1}{(1-q_b)^2} > 0$$

Puisque  $-2 \log(\Lambda)$  varie dans le même sens que  $q_b$  qui varie dans le même sens que  $F$ , alors les deux tests sont équivalents. Ce qui montre que le test de Fisher de l'analyse de la covariance ( $F = \frac{CMC}{CMRR}$ ) est le test du rapport de vraisemblance dans le contexte de la régression linéaire simple avec un effet aléatoire (données équilibrées) pour  $H_0 : \beta = 0$ .

De plus, comme montré à la section B.7.5 des annexes dans le cas où les données sont non équilibrées (voir aussi le tableau C.2 dans les annexes), il est possible de montrer que sous  $H_0 : \beta = 0$ ,  $\frac{SCC}{\sigma^2} \sim \chi_{(1)}^2$  et est indépendant de  $\frac{SCRR}{\sigma^2} \sim \chi_{(N-m-1)}^2$ . Ceci implique que sous  $H_0 : \beta = 0$ ,  $F = \frac{CMC}{CMRR} \sim \mathcal{F}_{N-m-1}^1$ .  $\square$

## j) Prédiction

Comme auparavant, utilisons la meilleure prédiction au sens des moindres carrés  $BP(a) = \mathbb{E}[a|\mathbf{y}] \Rightarrow BP(a_i) = \mathbb{E}[a_i|\bar{y}_i]$ . Ainsi,

**Proposition 2.74.**  $BP(a_i) = \mathbb{E}[a_i|\bar{y}_i] = \frac{n\sigma_a^2}{\sigma^2+n\sigma_a^2}(\bar{y}_i - \mu - \beta\bar{x})$

*Démonstration.*

$$\begin{aligned}
BP(a_i) &= \mathbb{E}[a_i | \bar{y}_i] \\
&= \mathbb{E}[a_i] + \text{cov}(a_i, \bar{y}_i) [\text{var}(\bar{y}_i)]^{-1} (\bar{y}_i - \mathbb{E}[\bar{y}_i]) \\
&= 0 + (\mathbb{E}[a_i \bar{y}_i] - \mathbb{E}[a_i] \mathbb{E}[\bar{y}_i]) \left[ \frac{1}{n^2} \text{var} \left( \sum_{j=1}^n y_{ij} \right) \right]^{-1} \left( \bar{y}_i - \frac{1}{n} \sum_{j=1}^n \mathbb{E}[y_{ij}] \right) \\
&= (\mathbb{E}[\mathbb{E}[a_i \bar{y}_i | a_i]]) \left[ \frac{1}{n^2} \mathbf{1}_n^t (\sigma^2 \mathbf{I}_n + \sigma_a^2 \mathbf{J}_n) \mathbf{1}_n \right]^{-1} \left( \bar{y}_i - \frac{1}{n} \sum_{j=1}^n (\mu + \beta x_j) \right) \\
&= (\mathbb{E}[a_i (\mu + a_i + \beta \bar{x})]) \left[ \frac{1}{n^2} [\sigma^2 + n\sigma_a^2, \dots, \sigma^2 + n\sigma_a^2] \mathbf{1}_n \right]^{-1} (\bar{y}_i - \mu - \beta \bar{x}) \\
&= \sigma_a^2 \left[ \frac{1}{n^2} (n\sigma^2 + n^2 \sigma_a^2) \right]^{-1} (\bar{y}_i - \mu - \beta \bar{x}) \\
&= \frac{n\sigma_a^2}{\sigma^2 + n\sigma_a^2} (\bar{y}_i - \mu - \beta \bar{x}) \quad \square
\end{aligned}$$

Si les variances  $\sigma^2$  et  $\sigma_a^2$  sont connues, utilisons

$$BP^0(a_i) = \frac{n\sigma_a^2}{\sigma^2 + n\sigma_a^2} (\bar{y}_i - \hat{\mu} - \hat{\beta} \bar{x}) = \frac{n\sigma_a^2}{\sigma^2 + n\sigma_a^2} (\bar{y}_i - \bar{y}_{..}) \quad (2.27)$$

Si les variances ne sont pas connues, il est possible d'utiliser

$$\widehat{BP}(a_i) = \frac{n\hat{\sigma}_a^2}{\hat{\sigma}^2 + n\hat{\sigma}_a^2} (\bar{y}_i - \bar{y}_{..}) \quad (2.28)$$

Mais les propriétés de cet estimateur de la prédiction ne sont pas connues.



### 2.3.4 Régression linéaire simple avec un effet aléatoire (données non équilibrées)

#### a) Le modèle

$$\begin{aligned}\mathbb{E}[y_{ij}|a_i] &= \mu + a_i + \beta x_{ij} & i = 1, \dots, m; j = 1, \dots, n_i \\ y_{ij}|a_i &\sim \mathcal{N}(\mu + a_i + \beta x_{ij}, \sigma^2) \text{ indépendantes} & (2.29) \\ a_i &\sim \text{i.i.d } \mathcal{N}(0, \sigma_a^2)\end{aligned}$$

Remarquons que les deux seules différences entre ce modèle et le modèle précédent est que  $x_i$  est remplacé par  $x_{ij}$  et  $n$  est remplacé par  $n_i$ . Le fait d'autoriser les données à être non équilibrées permet d'avoir différents  $x_{ij}$  pour chaque groupe, donc c'est beaucoup moins limité comme modèle. Notons  $N = \sum_{i=1}^m n_i$ .

Il en découle que  $\mathbb{E}[y_{ij}] = \mathbb{E}[\mathbb{E}[y_{ij}|a_i]] = \mathbb{E}[\mu + a_i + \beta x_{ij}] = \mu + \beta x_{ij}$ .

#### b) Les covariances entre les observations

Ce sont les mêmes variances et covariances que dans le modèle où les données sont équilibrées. En effet,

**Proposition 2.75.**  $\text{Var}(y_{ij}) = \sigma_a^2 + \sigma^2$

*Démonstration.*

$$\begin{aligned}\text{Var}(y_{ij}) &= \text{Var}(\mathbb{E}[y_{ij}|a_i]) + \mathbb{E}[\text{Var}(y_{ij}|a_i)] \\ &= \text{Var}(\mu + a_i + \beta x_{ij}) + \mathbb{E}[\sigma^2] \\ &= \text{Var}(a_i) + \sigma^2 \\ &= \sigma_a^2 + \sigma^2\end{aligned} \quad \square$$

**Proposition 2.76.**  $cov(y_{ij}, y_{ik}) = \sigma_a^2$  avec  $j \neq k$ .

*Démonstration.*

$$\begin{aligned} cov(y_{ij}, y_{ik}) &= cov(\mathbb{E}[y_{ij}|a_i], \mathbb{E}[y_{ik}|a_i]) + \mathbb{E}[cov(y_{ij}, y_{ik}|a_i)] \\ &= cov(\mu + a_i + \beta x_{ij}, \mu + a_i + \beta x_{ik}) + \mathbb{E}[0] \\ &= Var(a_i) \\ &= \sigma_a^2 \end{aligned} \quad \square$$

### c) Les sommes de carrés

Voici les sommes de carrés qui interviennent dans ce modèle.

Notation	Expression	Carré moyen	Signification
$SCE$	$\sum_{i=1}^m \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y}_{..})^2$	$CME = \frac{SCE}{m-1}$	Carrés expliqués
$SCR$	$\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2$	$CMR = \frac{SCR}{N-m}$	Carrés résiduels
$SCT$	$\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2$	—	Carrés totaux
$SCC$	$\frac{S_{xy}^2}{S_{xx}}$	$CMC = \frac{SCC}{1}$	Carrés dûs à la covariance
$SCR R$	$\sum_{i=1}^m \sum_{j=1}^{n_i} \left[ y_{ij} - \bar{y}_i - \frac{S_{xy}}{S_{xx}}(x_{ij} - \bar{x}_{i.}) \right]^2$	$CMRR = \frac{SCR R}{N-m-1}$	Carrés résiduels après la régression

Tableau 2.10 – Tableau des sommes de carrés, régression linéaire avec un effet aléatoire (données non équilibrées)

Où  $S_{xy} = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)(x_{ij} - \bar{x}_{i.})$  et où  $S_{xx} = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})^2$ . Il est aussi possible de montrer que  $SCT = SCE + SCR$  et  $SCR R + SCC = SCR$

d) **La vraisemblance**

Pour pouvoir écrire la vraisemblance, trouvons tout d'abord la loi de  $\mathbf{y} = [y_{11}, \dots, y_{mnm}]^t$ , le vecteur de toutes les observations.

**Proposition 2.77.** *La loi de  $\mathbf{y}$  est donnée par*

$$\mathbf{y} \sim \mathcal{N}\left(\mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \mathbf{V}\right)$$

Où  $\mathbf{X} = [\mathbf{1}_N, \mathbf{x}]$  et où  $\mathbf{x} = [x_{11}, \dots, x_{mnm}]^t$ . Aussi,  $\mathbf{V} = \{d \mathbf{V}_i\}$  où  $\mathbf{V}_i = \sigma^2 \mathbf{I}_{n_i} + \sigma_a^2 \mathbf{J}_{n_i}$ .

*Démonstration.* Posons

$$\mathbf{y}_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{in_i} \end{bmatrix} \quad \mathbf{x}_i = \begin{bmatrix} x_{i1} \\ \vdots \\ x_{in_i} \end{bmatrix}$$

$$\begin{aligned} \Rightarrow \mathbb{E}[\mathbf{y}_i | a_i] &= \mu \mathbf{1}_{n_i} + a_i \mathbf{1}_{n_i} + \beta \mathbf{x}_i & \text{et} & \quad \mathbb{E}[\mathbf{y}_i] = \mu \mathbf{1}_{n_i} + \beta \mathbf{x}_i \\ \Rightarrow \mathbb{E}[\mathbf{y}_i | a_i] &= \mathbf{X}_i \begin{bmatrix} \mu \\ \beta \end{bmatrix} + a_i \mathbf{1}_{n_i} & \text{et} & \quad \mathbb{E}[\mathbf{y}_i] = \mathbf{X}_i \begin{bmatrix} \mu \\ \beta \end{bmatrix} \end{aligned}$$

Où  $\mathbf{X}_i = [\mathbf{1}_{n_i}, \mathbf{x}_i]$ . Ainsi,

$$\begin{aligned} \mathbf{y}_i | a_i &\sim \mathcal{N}\left(\mathbf{X}_i \begin{bmatrix} \mu \\ \beta \end{bmatrix} + a_i \mathbf{1}_{n_i}, \sigma^2 \mathbf{I}_{n_i}\right) \text{ indépendantes} \\ \Rightarrow \mathbf{y}_i &\sim \mathcal{N}\left(\mathbf{X}_i \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \mathbf{V}_i\right) \text{ indépendantes} \end{aligned}$$

Où  $\mathbf{V}_i = \text{Var}(\mathbf{y}_i) = \sigma^2 \mathbf{I}_{n_i} + \sigma_a^2 \mathbf{J}_{n_i}$  en vertu des résultats trouvés sur les covariances des observations. De plus, les  $\mathbf{y}_i$  sont indépendants puisque, pour  $i \neq k$ ,

$$\begin{aligned}
\text{cov}(y_{ij}, y_{kl}) &= \text{cov}(\mathbb{E}[y_{ij}|a_i, a_k], \mathbb{E}[y_{kl}|a_i, a_k]) + \mathbb{E}[\text{cov}(y_{ij}, y_{kl}|a_i, a_k)] \\
&= \text{cov}(\mu + a_i + \beta x_{ij}, \mu + a_k + \beta x_{kl}) + \mathbb{E}[0] \\
&= \text{cov}(a_i, a_k) \\
&= 0
\end{aligned}$$

Posons maintenant

$$\mathbf{y} = \{c \mathbf{y}_i\}_{i=1}^m = \begin{bmatrix} y_{11} \\ \vdots \\ y_{mn_m} \end{bmatrix} \quad \text{et} \quad \mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix}$$

Il en découle que

$$\mathbb{E}[\mathbf{y}|\mathbf{a}] = \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix} + \mathbf{Z}\mathbf{a} \quad \text{et} \quad \mathbb{E}[\mathbf{y}] = \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix}$$

Où  $\mathbf{X} = \{c \mathbf{X}_i\} = \{c [\mathbf{1}_{n_i}, \mathbf{x}_i]\} = [\mathbf{1}_N, \mathbf{x}]$  et  $\mathbf{Z} = \{d \mathbf{1}_{n_i}\}$ . Donc finalement,

$$\begin{aligned}
\mathbf{y}|\mathbf{a} &\sim \mathcal{N}\left(\mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix} + \mathbf{Z}\mathbf{a}, \sigma^2 \mathbf{I}_N\right) \\
\mathbf{y} &\sim \mathcal{N}\left(\mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \mathbf{V}\right)
\end{aligned}$$

Où  $\mathbf{V} = \text{Var}(\mathbf{y}) = \{d \mathbf{V}_i\}$ . □

Pour simplifier un peu la notation, posons,

$$\mathbf{I}_i \equiv \mathbf{I}_{n_i} \quad \mathbf{J}_i \equiv \mathbf{J}_{n_i} \quad \tau_i = \frac{n_i \sigma_a^2}{\sigma^2 + n_i \sigma_a^2} \Rightarrow 1 - \tau_i = \frac{\sigma^2}{n_i \sigma_a^2} \tau_i$$

Voici maintenant deux propositions qui se démontrent à l'aide des propriétés des matrices et vecteurs de uns.

**Proposition 2.78.**  $\mathbf{V}^{-1} = \frac{1}{\sigma^2} \{d \mathbf{I}_i - \tau_i \bar{\mathbf{J}}_i\}$  où  $\bar{\mathbf{J}}_i = \frac{1}{n_i} \mathbf{J}_i$

*Démonstration.* Commençons par trouver  $\mathbf{V}_i^{-1}$

$$\begin{aligned} \mathbf{V}_i^{-1} &= \frac{1}{\sigma^2} \left( \mathbf{I}_i - \frac{\sigma_a^2}{\sigma^2 + n_i \sigma_a^2} \mathbf{J}_i \right) \\ &= \frac{1}{\sigma^2} \left( \mathbf{I}_i - \tau_i \bar{\mathbf{J}}_i \right) \end{aligned}$$

Ceci vient des propriétés de la section A.1 des annexes sur les matrices contenant des uns et des zéros. Et donc,  $\mathbf{V}^{-1} = \{d \mathbf{V}_i^{-1}\} = \frac{1}{\sigma^2} \{d \mathbf{I}_i - \tau_i \bar{\mathbf{J}}_i\}$ . □

De plus,

**Proposition 2.79.**  $|\mathbf{V}| = (\sigma^2)^{N-m} \prod_{i=1}^m (\sigma^2 + n_i \sigma_a^2)$

*Démonstration.*

$$\begin{aligned} |\mathbf{V}| &= \prod_{i=1}^m |\mathbf{V}_i| \\ &= \prod_{i=1}^m |\sigma^2 \mathbf{I}_i + \sigma_a^2 \mathbf{J}_i| \\ &= \prod_{i=1}^m [(\sigma^2)^{n_i-1} (\sigma^2 + n_i \sigma_a^2)] \\ &= (\sigma^2)^{N-m} \prod_{i=1}^m (\sigma^2 + n_i \sigma_a^2) \end{aligned} \quad \square$$

La vraisemblance est donnée par

**Proposition 2.80.**

$$L(\mu, \beta, \sigma^2, \sigma_a^2) = \frac{1}{(2\pi)^{N/2}} |\mathbf{V}|^{-1/2} \exp \left( -\frac{1}{2} \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix} \right)^t \mathbf{V}^{-1} \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix} \right) \right)$$

*Démonstration.* La preuve est directe avec la définition de la densité de la loi normale multidimensionnelle. □

Prenons le logarithme et développons le terme de l'exponentielle.

**Proposition 2.81.**

$$\begin{aligned} l = \log(L) &= -\frac{N}{2} \log(2\pi) - \frac{1}{2} \log(|\mathbf{V}|) + \frac{1}{2} \boldsymbol{\beta}^t \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} + \frac{1}{2} \mathbf{y}^t \mathbf{V}^{-1} \mathbf{X} \boldsymbol{\beta} \\ &\quad - \frac{1}{2} \boldsymbol{\beta}^t \mathbf{X}^t \mathbf{V}^{-1} \mathbf{X} \boldsymbol{\beta} - \frac{1}{2} \mathbf{y}^t \mathbf{V}^{-1} \mathbf{y} \end{aligned}$$

Où  $\boldsymbol{\beta} = \begin{bmatrix} \mu \\ \beta \end{bmatrix}$

Estimons premièrement  $\mu$  et  $\beta$

**Proposition 2.82.** *Les estimateurs du maximum de vraisemblance de  $\mu$  et  $\beta$  (en supposant que  $V$  soit connu) sont*

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} = (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y}$$

*Démonstration.* Voir la preuve de la proposition 2.57. □

Remplaçons  $\mathbf{V}$ ,  $\mathbf{X}$  et  $\mathbf{y}$  par leur valeur.

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} \tilde{y}_{..} - \hat{\beta} \tilde{x}_{..} \\ \frac{(\sigma_a^2/\sigma^2)S_{xy} + WSM_{xy}}{(\sigma_a^2/\sigma^2)S_{xx} + WSM_{xx}} \end{bmatrix} \tag{2.30}$$

Où  $\tilde{y}_{..} = \frac{\sum_{i=1}^m \tau_i \bar{y}_i}{\sum_{i=1}^m \tau_i}$ ,  $\tilde{x}_{..} = \frac{\sum_{i=1}^m \tau_i \bar{x}_i}{\sum_{i=1}^m \tau_i}$ ,  $WSM_{xy} = \sum_{i=1}^m \tau_i (\bar{y}_i - \tilde{y}_{..})(\bar{x}_i - \tilde{x}_{..})$  et  $WSM_{xx} = \sum_{i=1}^m \tau_i (\bar{x}_i - \tilde{x}_{..})^2$ .

Les détails se retrouvent à la section B.7.1 des annexes.

### La vraisemblance modifiée pour estimer $\sigma^2$ et $\sigma_a^2$

Jusqu'ici, nous avons supposé que les variances  $\sigma^2$  et  $\sigma_a^2$  étaient connues, mais ce n'est pas toujours le cas. Il faut donc les estimer et, comme dans le cas où les données sont équilibrées, le maximum de la log-vraisemblance modifiée  $l^* = \log(L(\mu = \hat{\mu}, \beta = \hat{\beta}, \sigma^2, \sigma_a^2))$  sera utilisé.

La log-vraisemblance s'écrit comme suit.

$$l = -\frac{N}{2} \log(2\pi) - \frac{1}{2} \log(|\mathbf{V}|) - \frac{1}{2} (\mathbf{y} - \boldsymbol{\Theta})^t \mathbf{V}^{-1} (\mathbf{y} - \boldsymbol{\Theta}) \quad (2.31)$$

Où  $\boldsymbol{\Theta} = \mathbb{E}[\mathbf{y}] = \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix}$ .

**Proposition 2.83.** *Développons  $\mathbf{V}$  et  $\mathbf{y} - \boldsymbol{\Theta}$  dans la précédente expression et remplaçons  $\mu$  par  $\hat{\mu}$  et  $\beta$  par  $\hat{\beta}$ . Ainsi,*

$$l^* = -\frac{1}{2} \left[ N \log(2\pi) + (N - m) \log(\sigma^2) + \sum_{i=1}^m \log(\sigma^2 + n_i \sigma_a^2) + \frac{S_1^*}{\sigma^2} - \frac{WS_2^*}{\sigma^2} \right]$$

Où  $S_1^* = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu} - \hat{\beta} x_{ij})^2$  et  $WS_2^* = \sum_{i=1}^m \tau_i n_i (\bar{y}_i - \hat{\mu} - \hat{\beta} \bar{x}_i)^2$ .

*Démonstration.* Les détails des calculs se retrouvent à la section B.7.2 des annexes.  $\square$

e) Les équations du maximum de vraisemblance et leurs solutions

Puisque  $\hat{\mu}$  et  $\hat{\beta}$  ont déjà été trouvés, énonçons seulement les équations du maximum de vraisemblance pour les estimateurs des variances.

**Proposition 2.84.** *Les solutions du maximum de la vraisemblance modifiée sont données par les équations suivantes*

$$\frac{\partial WS_2^*}{\partial \sigma_a^2} = \frac{\partial S_1^*}{\partial \sigma_a^2} + \sum_{i=1}^m \frac{n_i \sigma^2}{\sigma^2 + n_i \sigma_a^2}$$

$$S_1^* + \sigma^2 \frac{\partial WS_2^*}{\partial \sigma^2} = \sigma^2(N - m) + \sum_{i=1}^m \frac{\sigma^4}{\sigma^2 + n_i \sigma_a^2} + \sigma^2 \frac{\partial S_1^*}{\partial \sigma^2} + WS_2^*$$

*Démonstration.*

$$\begin{aligned} \frac{\partial l^*}{\partial \sigma^2} = 0 &\Rightarrow -\frac{N - m}{2\sigma^2} - \frac{1}{2} \sum_{i=1}^m \frac{1}{\sigma^2 + n_i \sigma_a^2} - \frac{1}{2\sigma^4} \left( \sigma^2 \frac{\partial S_1^*}{\partial \sigma^2} - S_1^* \right) + \frac{1}{2\sigma^4} \left( \sigma^2 \frac{\partial WS_2^*}{\partial \sigma^2} - WS_2^* \right) = 0 \\ &\Rightarrow -\sigma^2(N - m) - \sum_{i=1}^m \frac{\sigma^4}{\sigma^2 + n_i \sigma_a^2} - \sigma^2 \frac{\partial S_1^*}{\partial \sigma^2} + S_1^* + \sigma^2 \frac{\partial WS_2^*}{\partial \sigma^2} - WS_2^* = 0 \\ &\Rightarrow S_1^* + \sigma^2 \frac{\partial WS_2^*}{\partial \sigma^2} = \sigma^2(N - m) + \sum_{i=1}^m \frac{\sigma^4}{\sigma^2 + n_i \sigma_a^2} + \sigma^2 \frac{\partial S_1^*}{\partial \sigma^2} + WS_2^* \end{aligned}$$

$$\begin{aligned} \frac{\partial l^*}{\partial \sigma_a^2} = 0 &\Rightarrow -\frac{1}{2} \sum_{i=1}^m \frac{n_i}{\sigma^2 + n_i \sigma_a^2} - \frac{1}{2\sigma^2} \frac{\partial S_1^*}{\partial \sigma_a^2} + \frac{1}{2\sigma^2} \frac{\partial WS_2^*}{\partial \sigma_a^2} = 0 \\ &\Rightarrow \frac{\partial WS_2^*}{\partial \sigma_a^2} = \frac{\partial S_1^*}{\partial \sigma_a^2} + \sum_{i=1}^m \frac{n_i \sigma^2}{\sigma^2 + n_i \sigma_a^2} \quad \square \end{aligned}$$

f) Les estimateurs du maximum de vraisemblance

Les estimateurs du maximum de vraisemblance sont donnés par



**Proposition 2.85.**

$$\begin{aligned}\hat{\mu} &= \begin{cases} \tilde{y}_{..} - \frac{(\hat{\sigma}_a^2/\hat{\sigma}^2)S_{xy}+WSM_{xy}}{(\hat{\sigma}_a^2/\hat{\sigma}^2)S_{xx}+WSM_{xx}} \tilde{x}_{..} & si \ \dot{\sigma}_a^2 \geq 0 \\ \bar{y}_{..} - \frac{S_{xy}}{S_{xx}} \bar{x}_{..} & si \ \dot{\sigma}_a^2 < 0 \end{cases} \\ \hat{\beta} &= \begin{cases} \frac{(\hat{\sigma}_a^2/\hat{\sigma}^2)S_{xy}+WSM_{xy}}{(\hat{\sigma}_a^2/\hat{\sigma}^2)S_{xx}+WSM_{xx}} & si \ \dot{\sigma}_a^2 \geq 0 \\ \frac{S_{xy}}{S_{xx}} & si \ \dot{\sigma}_a^2 < 0 \end{cases} \\ \hat{\sigma}^2 &= \begin{cases} \hat{\sigma}^2 & si \ \dot{\sigma}_a^2 \geq 0 \\ \frac{SCRR}{N} & si \ \dot{\sigma}_a^2 < 0 \end{cases} \\ \hat{\sigma}_a^2 &= \begin{cases} \dot{\sigma}_a^2 & si \ \dot{\sigma}_a^2 \geq 0 \\ 0 & si \ \dot{\sigma}_a^2 < 0 \end{cases}\end{aligned}$$

*Démonstration.* Premièrement, pour  $\hat{\mu}$  et pour  $\hat{\beta}$ , il n'y a pas de problème de domaine, mais contrairement au cas où les données sont équilibrées, ils changent en fonction des valeurs que vont prendre les estimateurs des variances donc il faut en tenir compte. Deuxièmement, remarquons que  $\dot{\sigma}_a^2$  peut être négatif (ce n'est pas clair ici si  $\dot{\sigma}^2$  aussi peut être négatif, mais nous pourrions sans doute penser à utiliser un argument semblable à celui de l'analyse de la variance avec un facteur aléatoire. Par contre c'est beaucoup plus complexe pour ce modèle, ainsi, ceci pourrait être le sujet d'une étude subséquente). Pour trouver les estimateurs du maximum de vraisemblance, procédons de façon semblable au cas de l'analyse de la variance à un facteur aléatoire, mais en tenant compte de  $\mu$  et  $\beta$ . C'est-à-dire que si  $\dot{\sigma}_a^2 < 0$  alors  $\hat{\sigma}_a^2 = 0$  et  $\hat{\sigma}^2$ ,  $\hat{\mu}$  et  $\hat{\beta}$  sont égaux aux valeurs satisfaisant le maximum de vraisemblance avec  $\sigma_a^2 = 0$ . Dans cette situation ( $\sigma_a^2 = 0$ ) la variance devient

$$\begin{aligned}\mathbf{V}_0 = \sigma^2 \mathbf{I}_N &\Rightarrow \mathbf{V}_0^{-1} = \frac{1}{\sigma^2} \mathbf{I}_N \\ &\Rightarrow |\mathbf{V}_0| = (\sigma^2)^N\end{aligned}$$

Donc,

$$\begin{aligned}
l(\mu, \beta, \sigma, \sigma_a^2 = 0) &= -\frac{N}{2} \log(2\pi) - \frac{1}{2} \log((\sigma^2)^N) - \frac{1}{2\sigma^2} \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix} \right)^t \mathbf{I}_N \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \mu \\ \beta \end{bmatrix} \right) \\
&= -\frac{N}{2} \log(2\pi) - \frac{N}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu - \beta x_{ij})^2
\end{aligned}$$

Remarquons que cette vraisemblance est exactement la même que pour la régression linéaire simple, donc le résultat s'ensuit.  $\square$

### g) Espérance et biais des estimateurs

Les espérances des estimateurs de  $\mu$  et de  $\beta$  (en supposant que les variances  $\sigma^2$  et  $\sigma_a^2$  sont connues) sont données par

**Proposition 2.86.**  $\mathbb{E}[\hat{\beta}] = \beta$  et  $\mathbb{E}[\hat{\mu}] = \mu$

*Démonstration.* Premièrement, il est possible de montrer que (voir la section B.7.3)

$$\mathbb{E}[S_{xy}] = \beta S_{xx} \quad \text{et} \quad \mathbb{E}[WSM_{xy}] = \beta WSM_{xx}$$

Donc,

$$\begin{aligned}
\mathbb{E}[\hat{\beta}] &= \mathbb{E} \left[ \frac{(\sigma_a^2/\sigma^2)S_{xy} + WSM_{xy}}{(\sigma_a^2/\sigma^2)S_{xx} + WSM_{xx}} \right] \\
&= \frac{(\sigma_a^2/\sigma^2)\mathbb{E}[S_{xy}] + \mathbb{E}[WSM_{xy}]}{(\sigma_a^2/\sigma^2)S_{xx} + WSM_{xx}} \\
&= \frac{(\sigma_a^2/\sigma^2)\beta S_{xx} + \beta WSM_{xx}}{(\sigma_a^2/\sigma^2)S_{xx} + WSM_{xx}} \\
&= \beta
\end{aligned}$$

Et,

$$\begin{aligned}
\mathbb{E}[\hat{\beta}_0] &= \mathbb{E}\left[\frac{S_{xy}}{S_{xx}}\right] \\
&= \frac{\mathbb{E}[S_{xy}]}{S_{xx}} \\
&= \frac{\beta S_{xx}}{S_{xx}} \\
&= \beta
\end{aligned}$$

Donc, pour les deux situations qui définissent  $\hat{\beta}$  l'espérance est la même ( $\beta$ ), donc cet estimateur est sans biais.

Pour  $\hat{\mu}$ ,

$$\mathbb{E}[\hat{\mu}] = \mathbb{E}[\tilde{y}_{..} - \hat{\beta}\tilde{x}_{..}] = \mu + \beta\tilde{x}_{..} - \beta\tilde{x}_{..} = \mu$$

Et,

$$\mathbb{E}[\hat{\mu}_0] = \mathbb{E}[\bar{y}_{..} - \hat{\beta}_0\bar{x}_{..}] = \mu + \beta\bar{x}_{..} - \beta\bar{x}_{..} = \mu$$

Donc, comme dans le cas de  $\hat{\beta}$ , pour les deux situations qui définissent  $\hat{\mu}$ , l'espérance est la même ( $\mu$ ). □

Pour  $\hat{\sigma}^2$  et  $\hat{\sigma}_a^2$ , il est impossible de trouver une forme explicite pour leur espérance puisqu'ils sont décrits en deux cas et eux mêmes n'ont pas de forme explicite. Par contre, il est possible de trouver leur loi asymptotique et ce sera fait dans la prochaine section.

## h) Variance des estimateurs (et leur loi)

Premièrement, voici la proposition suivante si  $\sigma^2$  et  $\sigma_a^2$  sont connus.

**Proposition 2.87.**

$$\begin{aligned} \text{Var}(\hat{\beta}) &= \frac{1}{(S_{xx}/\sigma^2) + (WSM_{xx}/\sigma_a^2)} \\ \text{Var}(\hat{\mu}) &= \frac{\sigma_a^2}{\sum_{i=1}^m \tau_i} + \tilde{x}_{..}^2 \text{Var}(\hat{\beta}) \\ \text{cov}(\hat{\mu}, \hat{\beta}) &= -\tilde{x}_{..} \text{Var}(\hat{\beta}) \end{aligned}$$

*Démonstration.* Utilisons la même façon de procéder que dans le cas où les données sont équilibrées, c'est-à-dire  $\text{Var} \left( \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \right) = (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1}$ . Les détails sont à la section B.7.3 des annexes □

La loi asymptotique des estimateurs du maximum de vraisemblance est donnée par la proposition suivante.

**Proposition 2.88.**

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \approx \mathcal{N} \left( \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \frac{1}{D_2 \sigma^2} \begin{pmatrix} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}^2 - \sum_{i=1}^m n_i \tau_i \bar{x}_i^2 & \sum_{i=1}^m n_i \tau_i \bar{x}_i - N \bar{x}_{..} \\ \sum_{i=1}^m n_i \tau_i \bar{x}_i - N \bar{x}_{..} & N - \sum_{i=1}^m n_i \tau_i \end{pmatrix} \right)$$

Et ils sont indépendants de

$$\begin{bmatrix} \hat{\sigma}^2 \\ \hat{\sigma}_a^2 \end{bmatrix} \approx \mathcal{N} \left( \begin{bmatrix} \sigma^2 \\ \sigma_a^2 \end{bmatrix}, \frac{1}{2D_1} \begin{pmatrix} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} & -\sum_{i=1}^m \frac{n_i}{\lambda_i^2} \\ -\sum_{i=1}^m \frac{n_i}{\lambda_i^2} & \frac{N-m}{\sigma^4} + \sum_{i=1}^m \frac{1}{\lambda_i^2} \end{pmatrix} \right)$$

Où  $\lambda_i = \sigma^2 + n_i \sigma_a^2$ ,  $\tau_i = \frac{n_i \sigma_a^2}{\sigma^2 + n_i \sigma_a^2}$ ,  $D_1 = \frac{N-m}{4\sigma^4} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} + \frac{1}{4} \left( \sum_{i=1}^m \frac{1}{\lambda_i} \right) \left( \sum_{i=1}^m \frac{n_i^2}{\lambda_i} \right) - \frac{1}{4} \left( \sum_{i=1}^m \frac{n_i}{\lambda_i} \right) \left( \sum_{i=1}^m \frac{n_i}{\lambda_i} \right)$   
 et  $D_2 = \left( \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}^2 - \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \bar{x}_i^2 \right) \left( \frac{N}{\sigma^2} - \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \right) - \left( \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \bar{x}_i - \frac{N \bar{x}_{..}}{\sigma^2} \right)^2$

*Démonstration.* Le preuve se retrouve à la section B.7.4 des annexes. □

i) **Tests d'hypothèse et intervalles de confiance**

**Tester l'hypothèse nulle  $H_0 : \beta = 0$**

Premièrement, sous  $H_0 : \beta = 0$ , il est possible de montrer que

**Proposition 2.89.**

$$F = \frac{SCC/1}{SCRR/N - m - 1} \sim \mathcal{F}_{N-m-1}^1$$

*Démonstration.* Les détails de la preuve se retrouve à la section B.7.5 □

Par contre, nous n'avons pas trouvé de preuve indiquant que ce test est valide pour l'hypothèse nulle  $H_0 : \beta = 0$  dans le cas où les données sont non équilibrées. Voici le test approximatif utilisé dans le logiciel R avec le « package » lme4 [R C14].

$$t = \frac{\hat{\beta}}{\sqrt{\widehat{\text{Var}}(\hat{\beta})}} = \hat{\beta} \sqrt{\frac{S_{xx}}{\hat{\sigma}^2} + \frac{WSM_{xx}}{\hat{\sigma}_a^2}} \approx t_{(d)} \quad (2.32)$$

Où  $d$  est le nombre de degré de liberté de la loi de Student. Nous soupçonnons que le degré de liberté est  $N - m - 1$ , mais nous n'avons pas la précision dans le logiciel R. Il est aussi possible d'approximer cette loi de Student par la loi normale centrée réduite si  $N$  est assez grand.

**Tester l'hypothèse nulle  $H_0 : \sigma_a^2 = 0$**

Malheureusement, pour tester l'hypothèse nulle  $H_0 : \sigma_a^2 = 0$ , la statistique utilisée dans le cas où les données sont équilibrées n'est plus valide. Par contre, il est possible de montrer, avec les mêmes calculs que dans la section B.5.2 des annexes, que, sous l'hypothèse nulle  $H_0 : \sigma_a^2 = 0$ ,  $F = \frac{CME}{CMRR}$  suit une loi de Fisher non centrée. En effet,

$$F = \frac{CME}{CMRR} \sim \mathcal{F}_{N-m-1}^{m-1} \left( \beta^2 \sum_{i=1}^m \sum_{j=1}^{n_i} (\bar{x}_{i.} - \bar{x}_{..})^2 \right)$$

Par contre, nous n'avons pas trouvé de preuve indiquant que ce test est valide pour l'hypothèse nulle  $H_0 : \sigma_a^2 = 0$  dans le cas où les données sont non équilibrées.

## j) Prédiction

Comme auparavant, utilisons la meilleure prédiction au sens des moindres carrés  $BP(a) = \mathbb{E}[a|\mathbf{y}] \Rightarrow BP(a_i) = \mathbb{E}[a_i|\bar{y}_i]$ . Ainsi,

**Proposition 2.90.**  $BP(a_i) = \mathbb{E}[a_i|\bar{y}_i] = \frac{n_i\sigma_a^2}{\sigma^2+n_i\sigma_a^2}(\bar{y}_i - \mu - \beta\bar{x}_i)$

*Démonstration.*

$$\begin{aligned}
BP(a_i) &= \mathbb{E}[a_i|\bar{y}_i] \\
&= \mathbb{E}[a_i] + \text{cov}(a_i, \bar{y}_i)[\text{var}(\bar{y}_i)]^{-1}(\bar{y}_i - \mathbb{E}[\bar{y}_i]) \\
&= 0 + (\mathbb{E}[a_i\bar{y}_i] - \mathbb{E}[a_i]\mathbb{E}[\bar{y}_i]) \left[ \frac{1}{n_i^2} \text{var} \left( \sum_{j=1}^{n_i} y_{ij} \right) \right]^{-1} \left( \bar{y}_i - \frac{1}{n_i} \sum_{j=1}^{n_i} \mathbb{E}[y_{ij}] \right) \\
&= (\mathbb{E}[\mathbb{E}[a_i\bar{y}_i|a_i]]) \left[ \frac{1}{n_i^2} \mathbf{1}_{n_i}^t (\sigma^2 \mathbf{I}_{n_i} + \sigma_a^2 \mathbf{J}_{n_i}) \mathbf{1}_{n_i} \right]^{-1} \left( \bar{y}_i - \frac{1}{n_i} \sum_{j=1}^{n_i} (\mu + \beta x_{ij}) \right) \\
&= (\mathbb{E}[a_i(\mu + a_i + \beta\bar{x}_i)]) \left[ \frac{1}{n_i^2} [\sigma^2 + n_i\sigma_a^2, \dots, \sigma^2 + n_i\sigma_a^2] \mathbf{1}_{n_i} \right]^{-1} (\bar{y}_i - \mu - \beta\bar{x}_i) \\
&= \sigma_a^2 \left[ \frac{1}{n_i^2} (n_i\sigma^2 + n_i^2\sigma_a^2) \right]^{-1} (\bar{y}_i - \mu - \beta\bar{x}_i) \\
&= \frac{n_i\sigma_a^2}{\sigma^2 + n_i\sigma_a^2} (\bar{y}_i - \mu - \beta\bar{x}_i) \quad \square
\end{aligned}$$

Si  $\sigma^2$  et  $\sigma_a^2$  sont connus, utilisons

$$BP^0(a_i) = \frac{n_i\sigma_a^2}{\sigma^2 + n_i\sigma_a^2} (\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x}_i) \quad (2.33)$$

Si les variances ne sont pas connues, il est possible d'utiliser

$$\widehat{BP}(a_i) = \frac{n_i \hat{\sigma}_a^2}{\hat{\sigma}^2 + n_i \hat{\sigma}_a^2} (\bar{y}_{i.} - \hat{\mu} - \hat{\beta} \bar{x}_{i.}) \quad (2.34)$$

Mais les propriétés de cet estimateur de la prédiction ne sont pas connues.

### 2.3.5 Comparaison des deux modèles de régression avec un facteur

Voici deux tableaux résumés qui comparent les deux méthodes de régression linéaire avec un facteur, une le traite comme un effet fixe et l'autre comme un effet aléatoire. Nous comparons donc l'ANCOVA avec la régression linéaire simple avec un effet aléatoire. Les tableaux représentent les résultats dans le cas où les données sont équilibrées ( $n_i = n \forall i$ ) puisque, pour le modèle avec un effet aléatoire et où les données sont non équilibrées, il n'y a pas de forme explicite pour les estimateurs.

Estimateurs	Fixe	Aléatoire
$\hat{\mu}$	$\bar{y}_{..} - \hat{\beta} \bar{x}$	$\bar{y}_{..} - \hat{\beta} \bar{x}$
$\hat{\beta}$	$\frac{S_{xy}}{S_{xx}}$	$\frac{S_{xy}}{S_{xx}}$
$\hat{\alpha}_i$	$\bar{y}_{i.} - \bar{y}_{..}$	-
$\hat{\sigma}^2$	$\frac{SCR}{N}$	$\frac{SCR}{N-m}$ ou $\frac{SCE+SCR}{N}$
$\hat{\sigma}_a^2$	-	$\frac{(1-1/m)CME-CMR}{n} + \frac{\hat{\beta}^2 S_{xx}}{N(n-1)}$ ou 0

Tableau 2.11 – Comparaison entre les différents estimateurs - régression linéaire simple

Tests	Fixe	Aléatoire
$H_0 : \alpha_i = 0 \ \forall i$	$F = \frac{CME}{CMRR} \sim F_{N-m-1}^{m-1}$	-
$H_0 : \sigma_a^2 = 0$	-	$F = \frac{CME}{CMRR} \sim F_{N-m-1}^{m-1}$
$H_0 : \beta = 0$	$F = \frac{CMC}{CMRR} \sim \mathcal{F}_{N-m-1}^1$	$F = \frac{CMC}{CMRR} \sim \mathcal{F}_{N-m-1}^1$

Tableau 2.12 – Comparaison des tests d’hypothèse - régression linéaire simple

Nous remarquons que c’est exactement les mêmes tests et que les estimateurs se ressemblent. Comme dans le cas de l’analyse de la variance à un facteur, nous pourrions nous demander pourquoi utiliser le modèle avec un effet aléatoire au lieu d’utiliser le modèle avec un effet fixe. La différence sera encore au niveau de la prédiction d’une nouvelle donnée sachant son groupe et sa valeur pour la variable indépendante. Par exemple, étudions le cas où les données sont équilibrées et que la nouvelle donnée appartient au groupe  $i$  et a  $x_{nouw}$  comme valeur de la variable indépendante. Dans le cas de l’ANCOVA à un facteur fixe, la prédiction sera  $\hat{\mu} + \hat{\alpha}_i + \hat{\beta}x_{nouw} = \bar{y}_{..} - \hat{\beta}\bar{x} + \bar{y}_i - \bar{y}_{..} + \hat{\beta}x_{nouw} = \bar{y}_i - \frac{S_{xy}}{S_{xx}}(\bar{x} - x_{nouw})$ . Par contre, dans le cas de la régression linéaire avec une ordonnée à l’origine aléatoire, la prédiction sera  $\hat{\mu} + \tilde{\alpha}_i + \hat{\beta}x_{nouw} = \bar{y}_{..} - \hat{\beta}\bar{x} + \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}^2/n}(\bar{y}_i - \bar{y}_{..}) + \hat{\beta}x_{nouw} = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}^2/n}\bar{y}_i + \frac{\hat{\sigma}^2/n}{\hat{\sigma}_a^2 + \hat{\sigma}^2/n}\bar{y}_{..} - \frac{S_{xy}}{S_{xx}}(\bar{x} - x_{nouw})$ . Donc, si nous voulons faire des prédictions nous devons choisir le modèle minutieusement puisque les résultats vont différer d’un modèle à l’autre comme dans le cas de l’analyse de la variance à un facteur. Par contre, si nous désirons uniquement tester si le groupe a une influence sur l’observation indépendamment de la variable indépendante ou si la variable indépendante a un effet sur l’observation indépendamment du groupe, alors c’est moins crucial de choisir le bon modèle.



# CHAPITRE 3

## Théorie du modèle à effets mixtes généralisé pour des données génétiques provenant de familles

Maintenant que nous avons présenté la théorie des modèles à effets mixtes usuels, nous allons présenter le modèle utilisé lorsque les données proviennent de familles et que la variable indépendante représente une caractéristique génétique. Nous allons donc appliquer la régression linéaire simple avec un effet aléatoire à ces données avec une légère modification. Le problème avec des données provenant de famille est que l'hypothèse d'indépendance entre les observations n'est pas respectée. L'astuce ici est que l'effet aléatoire  $a_i$  va servir à mettre chaque personne comprise dans l'étude dans un groupe différent. Il va donc y avoir autant de niveaux à ce facteur qu'il y a d'observations (ou de personnes dans l'étude). Avec cette astuce, nous modifierons la loi des  $a_i$  pour que nous soyons capables de gérer la covariance entre les personnes d'une même famille.

### 3.1 Quelle est la différence avec les modèles usuels ?

Voici la notation qui sera utilisée tout au long de la section.

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \beta \end{bmatrix}$$

$$\mathbf{Z} = \mathbf{I}_n \quad \mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \quad \boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_n \end{bmatrix} \quad \mathbf{K} \in M_{n \times n}$$

Il y a donc  $n$  personnes incluses dans le modèle et donc  $n$  groupes. Les personnes peuvent avoir un lien de parenté entre elles (ex. parents, grands-parents, enfants, etc.) ou elles peuvent être indépendantes. Voici un tableau qui présente le modèle généralisé comparé au modèle étudié au chapitre précédent avec  $n$  personnes et  $n$  groupes pour que les modèles soient comparables.

Modèle Vu	Modèle proposé
$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \boldsymbol{\epsilon}$	$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a} + \boldsymbol{\epsilon}$
$\mathbf{a} \sim \mathcal{N}(0, \sigma_a^2 \mathbf{I}_n)$	$\mathbf{a} \sim \mathcal{N}(0, \sigma_a^2 \mathbf{K})$
$\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_n)$ ind. de $\mathbf{a}$	$\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_n)$ ind. de $\mathbf{a}$
$\Rightarrow \mathbb{E}[\mathbf{y} \mathbf{a}] = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a}$	$\Rightarrow \mathbb{E}[\mathbf{y} \mathbf{a}] = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{a}$
$\Rightarrow \mathbb{E}[\mathbf{y}] = \mathbf{X}\boldsymbol{\beta}$	$\Rightarrow \mathbb{E}[\mathbf{y}] = \mathbf{X}\boldsymbol{\beta}$
$\Rightarrow \text{Var}(y_i) = \sigma_a^2 + \sigma^2$	$\Rightarrow \text{Var}(y_i) = \sigma_a^2 K_{ii} + \sigma^2$
$\Rightarrow \text{cov}(y_i, y_j) = 0$	$\Rightarrow \text{cov}(y_i, y_j) = \sigma_a^2 K_{ij}$
$\Rightarrow \text{Var}(\mathbf{y}) = (\sigma^2 + \sigma_a^2) \mathbf{I}_n$	$\Rightarrow \text{Var}(\mathbf{y}) = \mathbf{I}_n \sigma^2 + \mathbf{K} \sigma_a^2$

Tableau 3.1 – Tableau comparatif entre le modèle généralisé pour des données provenant de familles et le modèle usuel

Remarquons que les deux modèles sont équivalents si  $\mathbf{K} = \mathbf{I}$ .

## 3.2 La matrice $\mathbf{K}$

Dans le cadre de données génétiques,  $\mathbf{K}$  peut être vue comme la matrice de similarité entre les individus. Il y a plusieurs mesures de similarité qui peuvent être utilisées, mais concentrons-nous sur la suivante

$$K_{ij} = \% \text{ moyen d'allèles partagés} \quad (3.1)$$

Lorsque nous écrivons « allèles partagés », nous ne parlons pas de nucléotides (lettre A, C, G ou T) mais bien d'allèles provenant du même ancêtre. De plus, avec cette définition, nous obtenons une matrice avec uniquement des termes positifs ou nuls telle que  $K_{ij} = K_{ji}$  et telle que  $K_{ii} = 1 \forall i$ . En résumé, si nous plaçons tous les membres d'une même famille ensemble, nous obtenons une matrice diagonale par blocs. Voici quelques exemples de la valeur que peut prendre  $K_{ij}$ .

**Exemple 3.1.** *Si les individus  $i$  et  $j$  ont une relation parent-enfant, alors  $K_{ij} = \frac{1}{2}$ , car chaque parent transmet la moitié de ses allèles à l'enfant.*

**Exemple 3.2.** *Si les individus  $i$  et  $j$  sont 2 frères ou sœurs, il est possible de montrer que  $K_{ij} = \frac{1}{2}$ .*

**Exemple 3.3.** *Si les individus  $i$  et  $j$  ont une relation grand-parent-enfant, alors  $K_{ij} = \frac{1}{4}$ .*

**Exemple 3.4** (Calcul de  $\mathbf{K}$ ).

*Si les individus  $i$  et  $j$  ont une relation frère-sœur.*

Soit  $AB$  le génotype de la mère et  $CD$  le génotype du père. Alors, les différents génotypes possibles pour la sœur sont les suivants en supposant (SPDG) que le génotype du frère est  $AC$  :

Génotype frère	Génotype sœur	Nombre d'allèles en commun	% d'allèles partagés
$AC$	$AC$	2	100%
$AC$	$AD$	1	50%
$AC$	$BC$	1	50%
$AC$	$BD$	0	0%

Donc si  $K_{ij}$  désigne le pourcentage moyen d'allèle partagé,  $K_{ij} = 50\% = \frac{1}{2}$ .

Il est donc possible de procéder de façon similaire pour trouver le  $K_{ij}$  de toutes les relations familiales.

Finalement, voici une conjecture sur la matrice  $\mathbf{K}$ . Nous n'avons trouvé aucune preuve dans la littérature, mais la construction de cette matrice nous fait soupçonner cette propriété. De plus, pour chaque exemple tenté, la matrice était définie positive.

**Conjecture 3.1.** *La matrice  $\mathbf{K}$  est définie positive.*

### 3.3 Théorie du nouveau modèle

Voici donc, de façon plus organisée, la théorie du modèle qui est inspirée de Kang et al. [KZW<sup>+</sup>08].

#### 3.3.1 Le modèle

$$\begin{aligned}
 \mathbb{E}[\mathbf{y}|\mathbf{a}] &= \mathbf{X}\boldsymbol{\beta} + \mathbf{a} \\
 \mathbf{y}|\mathbf{a} &\sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta} + \mathbf{a}, \sigma^2\mathbf{I}_n) \\
 \mathbf{a} &\sim \mathcal{N}(0, \sigma_a^2\mathbf{K})
 \end{aligned}
 \tag{3.2}$$

Où

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \beta \end{bmatrix} \quad \mathbf{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

Et  $\mathbf{K}$  est la matrice de similarité comme définie plus haut. De plus,  $\mathbb{E}[y_i] = \mathbb{E}[\mathbb{E}[y_i|a_i]] = \mathbb{E}[\mu + a_i + \beta x_i] = \mu + \beta x_i$ .

### 3.3.2 Les covariances entre les observations

Comme vu plus haut, les variances et covariances sont données par

**Proposition 3.1.**  $\text{Var}(y_i) = \sigma_a^2 K_{ii} + \sigma^2$

*Démonstration.*

$$\begin{aligned} \text{Var}(y_i) &= \text{Var}(\mathbb{E}[y_i|a_i]) + \mathbb{E}[\text{Var}(y_i|a_i)] \\ &= \text{Var}(\mu + a_i + \beta x_i) + \mathbb{E}[\sigma^2] \\ &= \text{Var}(a_i) + \sigma^2 \\ &= \sigma_a^2 K_{ii} + \sigma^2 \end{aligned} \quad \square$$

**Proposition 3.2.**  $\text{cov}(y_i, y_j) = \sigma_a^2 K_{ij}$  pour  $i \neq j$

*Démonstration.*

$$\begin{aligned} \text{cov}(y_i, y_j) &= \text{cov}(\mathbb{E}[y_i|a_i, a_j], \mathbb{E}[y_j|a_i, a_j]) + \mathbb{E}[\text{cov}(y_i, y_j|a_i, a_j)] \\ &= \text{cov}(\mu + a_i + \beta x_i, \mu + a_j + \beta x_j) + 0 \\ &= \text{cov}(a_i, a_j) \\ &= \sigma_a^2 K_{ij} \end{aligned} \quad \square$$

Donc, en général,  $\text{cov}(y_i, y_j) = \sigma_a^2 K_{ij} + \delta_{ij} \sigma^2$ . Où  $\delta_{ij}$  est le symbole de Kronecker, c'est-à-dire que

$$\delta_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

### 3.3.3 Les sommes de carrés

Notation	Expression	Carré moyen	Signification
$SCT$	$\sum_{i=1}^n (y_i - \bar{y})^2$	—	Carrés totaux
$SCC$	$\frac{S_{xy}^2}{S_{xx}}$	$CMC = \frac{SCC}{1}$	Carrés dûs à la covariance
$SCRR$	$\sum_{i=1}^n \left[ y_i - \bar{y} - \frac{S_{xy}}{S_{xx}}(x_i - \bar{x}) \right]^2$	$CMRR = \frac{SCRR}{n-2}$	Carrés résiduels après la régression

Tableau 3.2 – Tableau des sommes de carrés, modèle généralisé aux données génétiques provenant de familles

Où  $S_{xx} = \sum_{i=1}^n (x_i - \bar{x})^2$  et  $S_{xy} = \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$ .

### 3.3.4 La vraisemblance

Pour pouvoir écrire la vraisemblance, trouvons tout d'abord la loi de  $\mathbf{y} = [y_1, \dots, y_n]^t$ , le vecteur de toutes les observations.

**Proposition 3.3.** *La loi de  $\mathbf{y}$  est donnée par*

$$\mathbf{y} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \mathbf{I}_n \sigma^2 + \mathbf{K} \sigma_a^2)$$

*Démonstration.* Avec les espérances et les variances/covariances calculées plus haut et le fait que les erreurs sont i.i.d de loi normale, il en découle que

$$y_i | a_i \sim \mathcal{N}(\mu + a_i + \beta x_i, \sigma^2) \text{ indépendantes}$$

$$y_i \sim \mathcal{N}(\mu + \beta x_i, \sigma_a^2 K_{ii} + \sigma^2)$$

$$\Rightarrow \mathbf{y} | \mathbf{a} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta} + \mathbf{a}, \mathbf{I}_n \sigma^2)$$

$$\Rightarrow \mathbf{y} \sim \mathcal{N}(\mathbf{X}\boldsymbol{\beta}, \mathbf{I}_n \sigma^2 + \mathbf{K} \sigma_a^2) \quad \square$$

Voici donc une proposition sur la vraisemblance.

**Proposition 3.4.**

$$L(\boldsymbol{\beta}, \sigma^2, \sigma_a^2) = \frac{1}{(2\pi)^{n/2} |\mathbf{V}|^{1/2}} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^t \mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\right)$$

$$l(\boldsymbol{\beta}, \sigma^2, \sigma_a^2) = -\frac{1}{2} \left[ n \log(2\pi) + \log(|\mathbf{V}|) + (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^t \mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \right]$$

Où  $\mathbf{V} = \text{Var}(\mathbf{y}) = \mathbf{I}_n \sigma^2 + \mathbf{K} \sigma_a^2$ .

*Démonstration.* Avec la densité de la loi normale multidimensionnelle et le fait que  $\mathbf{y}$  suit une loi normale, il en découle le résultat souhaité. □

### 3.3.5 Les équations du maximum de vraisemblance et leur solution

**Proposition 3.5.** *Les solutions du maximum de vraisemblance sont données par les équations suivantes.*

$$\begin{aligned}\dot{\sigma}_a^2 &= \frac{\dot{R}}{n} \\ \dot{\boldsymbol{\beta}} &= (\mathbf{X}^t \dot{\mathbf{H}}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \dot{\mathbf{H}}^{-1} \mathbf{y} \\ 0 &= \frac{n}{\dot{R}} \frac{\partial \dot{R}}{\partial \delta} + \frac{1}{|\dot{\mathbf{H}}|} \frac{\partial |\dot{\mathbf{H}}|}{\partial \delta}\end{aligned}$$

Où

$$\begin{aligned}\delta &= \frac{\sigma^2}{\sigma_a^2} \\ \dot{R} &= (\mathbf{y} - \mathbf{X}\dot{\boldsymbol{\beta}})^t \dot{\mathbf{H}}^{-1} (\mathbf{y} - \mathbf{X}\dot{\boldsymbol{\beta}}) \\ \dot{\mathbf{H}} &= \dot{\delta} \mathbf{I}_n + \mathbf{K}\end{aligned}$$

*Démonstration.* Posons premièrement

$$\mathbf{H} = \frac{1}{\sigma_a^2} \mathbf{V} = \frac{1}{\sigma_a^2} (\sigma^2 \mathbf{I}_n + \sigma_a^2 \mathbf{K}) = \delta \mathbf{I}_n + \mathbf{K}$$

Où  $\delta = \frac{\sigma^2}{\sigma_a^2}$ . Ainsi,

$$|\mathbf{V}| = |\sigma_a^2 \mathbf{H}| = (\sigma_a^2)^n |\mathbf{H}|$$

Et,

$$\mathbf{V}^{-1} = (\sigma_a^2 \mathbf{H})^{-1} = \frac{1}{\sigma_a^2} \mathbf{H}^{-1}$$

Donc, la vraisemblance devient,



$$\begin{aligned}
l(\boldsymbol{\beta}, \delta, \sigma_a^2) &= -\frac{1}{2} \left[ n \log(2\pi) + \log((\sigma_a^2)^n |\mathbf{H}|) + \frac{1}{\sigma_a^2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^t \mathbf{H}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \right] \\
&= -\frac{1}{2} \left[ n \log(2\pi\sigma_a^2) + \log(|\mathbf{H}|) + \frac{1}{\sigma_a^2} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^t \mathbf{H}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \right]
\end{aligned}$$

Posons  $R = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^t \mathbf{H}^{-1} (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})$ . Dérivons la log-vraisemblance par rapport à  $\sigma_a^2$  et  $\boldsymbol{\beta}$  et égalisons ces expressions à 0 (en supposant que  $\delta$  est connu).

$$\begin{aligned}
\frac{\partial l}{\partial \sigma_a^2} = 0 &\Rightarrow -\frac{n}{2(2\pi\sigma_a^2)}(2\pi) + \frac{R}{2\sigma_a^4} \\
&\Rightarrow -\frac{n}{2\sigma_a^2} + \frac{R}{2\sigma_a^4} = 0 \\
&\Rightarrow R = n\sigma_a^2 \\
&\Rightarrow \dot{\sigma}_a^2 = \frac{\dot{R}}{n} \\
&\Rightarrow \dot{\sigma}_a^2 = \frac{(\mathbf{y} - \mathbf{X}\dot{\boldsymbol{\beta}})^t \mathbf{H}^{-1} (\mathbf{y} - \mathbf{X}\dot{\boldsymbol{\beta}})}{n}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial l}{\partial \boldsymbol{\beta}} = 0 &\Rightarrow -\frac{1}{2\sigma_a^2} \frac{\partial R}{\partial \boldsymbol{\beta}} = 0 \\
&\Rightarrow -\frac{1}{2\sigma_a^2} [2(\mathbf{X}^t \mathbf{H}^{-1} \mathbf{X}\boldsymbol{\beta} - \mathbf{X}^t \mathbf{H}^{-1} \mathbf{y})] = 0 \\
&\Rightarrow \mathbf{X}^t \mathbf{H}^{-1} \mathbf{X}\boldsymbol{\beta} = \mathbf{X}^t \mathbf{H}^{-1} \mathbf{y} \\
&\Rightarrow \dot{\boldsymbol{\beta}} = (\mathbf{X}^t \mathbf{H}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{H}^{-1} \mathbf{y}
\end{aligned}$$

Pour  $\delta$ , il est possible de s'inspirer de la log-vraisemblance modifiée. En effet, remplaçons  $\sigma_a^2$  par  $\dot{\sigma}_a^2$  et  $\boldsymbol{\beta}$  par  $\dot{\boldsymbol{\beta}}$  dans la log-vraisemblance.

$$l^* = l(\boldsymbol{\beta} = \dot{\boldsymbol{\beta}}, \sigma_a^2 = \dot{\sigma}_a^2, \delta) = -\frac{1}{2} \left[ n \log \left( \frac{2\pi\dot{R}}{n} \right) + \log(|\mathbf{H}|) + n \right]$$

Où  $\mathbf{H} = \delta \mathbf{I}_n + \mathbf{K}$ , une fonction de  $\delta$  et

$$\begin{aligned}
\dot{R} &= (\mathbf{y} - \mathbf{X}\dot{\beta})^t \mathbf{H}^{-1} (\mathbf{y} - \mathbf{X}\dot{\beta}) \\
&= (\mathbf{y} - \mathbf{X}(\mathbf{X}^t \mathbf{H}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{H}^{-1} \mathbf{y})^t \mathbf{H}^{-1} (\mathbf{y} - \mathbf{X}(\mathbf{X}^t \mathbf{H}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{H}^{-1} \mathbf{y}) \\
&= \mathbf{y}^t (\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t \mathbf{H}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{H}^{-1})^t \mathbf{H}^{-1} (\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t \mathbf{H}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{H}^{-1}) \mathbf{y} \\
&= \mathbf{y}^t \mathbf{P}^t \mathbf{H}^{-1} \mathbf{P} \mathbf{y}
\end{aligned}$$

Où  $\mathbf{P} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t \mathbf{H}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{H}^{-1}$ . Donc,  $\dot{R}$  est aussi un fonction de  $\delta$ . Donc,  $\dot{\delta}$  est tel que

$$\begin{aligned}
\frac{\partial l^*}{\partial \delta} = 0 &\Rightarrow -\frac{n}{2} \frac{1}{\dot{R}} \frac{\partial \dot{R}}{\partial \delta} - \frac{1}{2|\mathbf{H}|} \frac{\partial |\mathbf{H}|}{\partial \delta} \\
&\Rightarrow \frac{n}{\dot{R}} \frac{\partial \dot{R}}{\partial \delta} + \frac{1}{|\mathbf{H}|} \frac{\partial |\mathbf{H}|}{\partial \delta} = 0
\end{aligned}$$

En considérant que  $\dot{R}$  et  $|\mathbf{H}|$  sont des fonctions de  $\delta$ . □

Essayons maintenant de réécrire  $l^*$  de façon à ce qu'il soit possible de la dériver par rapport à  $\delta$  pour simplifier le résultat précédent. Premièrement, il est possible de montrer que

**Proposition 3.6.**

$$|\mathbf{H}| = \prod_{i=1}^n (\xi_i + \delta)$$

Où  $\xi_i$  sont les valeurs propres de  $\mathbf{K}$ .

*Démonstration.* Puisque  $\mathbf{K}$  est symétrique, elle est diagonalisable. Ainsi,

$$\mathbf{K} = \mathbf{U}_F \{d \xi_i\}_{i=1}^n \mathbf{U}_F^t$$

Où  $\mathbf{U}_F$  est la matrice (orthogonale) des vecteurs propres normés de  $\mathbf{K}$  associés aux valeurs propres  $\xi_i$ .

Ainsi,  $\delta \mathbf{I}_n = \delta \mathbf{U}_F \mathbf{U}_F^t$  et donc,

$$\begin{aligned} \mathbf{H} &= \delta \mathbf{I}_n + \mathbf{K} \\ &= \mathbf{U}_F \delta \mathbf{I}_n \mathbf{U}_F^t + \mathbf{U}_F \{d \xi_i\}_{i=1}^n \mathbf{U}_F^t \\ &= \mathbf{U}_F \{d \xi_i + \delta\}_{i=1}^n \mathbf{U}_F^t \end{aligned}$$

Remarquons aussi que  $\mathbf{U}_F$  et  $\xi_i$  sont indépendants de  $\delta$ . Ensuite, réécrivons  $|\mathbf{H}|$ .

$$\begin{aligned} |\mathbf{H}| &= \det(\mathbf{U}_F) \det(\{d \xi_i + \delta\}) \det(\mathbf{U}_F^t) \\ &= \prod_{i=1}^n (\xi_i + \delta) \end{aligned} \quad \square$$

Voici une proposition concernant  $\mathbf{SHS}$  avec  $\mathbf{S} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t$ .

**Proposition 3.7.**

$$\mathbf{SHS} = \mathbf{V}_R \{d [\lambda_1 + \delta, \dots, \lambda_{n-2} + \delta]\} \mathbf{V}_R^t$$

Où  $\mathbf{S} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t$  et où  $\mathbf{V}_R$  est la matrice des vecteurs propres de  $\mathbf{SKS}$  associés aux valeurs propres non nulles  $\lambda_i$ .

Nous avons besoin de deux propositions pour pouvoir démontrer cette proposition.

**Proposition 3.8.** *Si  $\mathbf{A}$  est une matrice diagonalisable, alors  $\text{rang}(A) = \text{nombre de valeurs propres non nulles}$ .*

*Démonstration.*  $\mathbf{A}$  est diagonalisable, donc il existe  $\mathbf{\Lambda}$  inversible telle que

$$\mathbf{A} = \mathbf{\Lambda} \mathbf{D} \mathbf{\Lambda}^{-1}$$

Où  $\mathbf{D}$  est une matrice diagonale contenant les valeurs propres de  $\mathbf{A}$ .

Il s'ensuit que  $\mathbf{A}$  et  $\mathbf{D}$  sont semblables, elles ont ainsi le même rang. De plus, le rang d'une matrice diagonale est égal au nombre de termes diagonaux non nuls donc ici au nombre de valeurs propres non nulles.  $\square$

**Proposition 3.9.** *Les vecteurs propres de  $\mathbf{SKS}$  sont égaux aux vecteurs propres de  $\mathbf{S}$*

*Démonstration.* Remarquons que  $\mathbf{S}$  et  $\mathbf{SKS}$  sont diagonalisables et qu'elles commutent ( $\mathbf{S}$  est idempotente). Ainsi, elles sont simultanément diagonalisables. Elles ont donc les mêmes vecteurs propres (voir le livre de Horn et Johnson aux pages 49 à 55 [HJ85])  $\square$

*Démonstration pour SHS.* Posons  $\mathbf{S} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t$  une matrice symétrique et idempotente, en effet,

$$\begin{aligned} \mathbf{SS} &= (\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t)(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t) \\ &= \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t + \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \\ &= \mathbf{I}_n - 2\mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t + \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \\ &= \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t \mathbf{X})^{-1} \mathbf{X}^t \\ &= \mathbf{S} \end{aligned}$$

Puisque  $\mathbf{S}$  est idempotente et symétrique, alors elle est diagonalisable avec des valeurs propres égales à 1 ou 0. En utilisant la proposition 3.8, le rang de  $\mathbf{S}$  est égal au nombre de valeurs propres égales à 1. Puisque  $\mathbf{S}$  est idempotente, il en découle que

$$\begin{aligned}
\text{rang}(\mathbf{S}) &= \text{trace}(\mathbf{S}) \\
&= \text{trace}(\mathbf{I}_n) - \text{trace}(\mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t) \\
&= n - \text{trace}((\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{X}) \\
&= n - \text{trace}(\mathbf{I}_2) \\
&= n - 2
\end{aligned}$$

Ainsi,

$$\mathbf{S} = \mathbf{U}_R \underbrace{\{d [1, 1, \dots, 1, 0, 0]\}}_{n-2} \mathbf{U}_R^t$$

Où  $\mathbf{U}_R$  est la matrice (orthogonale) des vecteurs propres de  $\mathbf{S}$  associés aux valeurs propres (0 ou 1) de  $\mathbf{S}$ . Ensuite, puisque  $\mathbf{K}$  ( $= \mathbf{U}_F \{d \xi_i\}_{i=1}^n \mathbf{U}_F^t$ ) est définie positive, il existe une matrice symétrique  $\mathbf{M}$ , elle aussi définie positive (donc inversible) telle que  $\mathbf{K} = \mathbf{M}\mathbf{M}$  avec  $\mathbf{M} = \mathbf{U}_F \{d \sqrt{\xi_i}\}_{i=1}^n \mathbf{U}_F^t$ . Ainsi,

$$\begin{aligned}
\mathbf{SKS} &= \mathbf{SMMS} \\
&= (\mathbf{MS})^t(\mathbf{MS})
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \text{rang}(\mathbf{SKS}) &= \text{rang}((\mathbf{MS})^t(\mathbf{MS})) \\
&= \text{rang}(\mathbf{MS}) \\
&= \text{rang}(\mathbf{S}) \\
&= n - 2
\end{aligned}$$

Utilisons maintenant la proposition 3.9. Il en découle que

$$\mathbf{SKS} = \mathbf{U}_R \{_d [\lambda_1, \dots, \lambda_{n-2}, 0, 0]\} \mathbf{U}_R^t$$

Où les  $\lambda_i$  sont les  $n - 2$  valeurs propres non nulles de  $\mathbf{SKS}$  associées aux  $n - 2$  premiers vecteurs propres dans  $\mathbf{U}_R$ .

Et ainsi,

$$\begin{aligned} \mathbf{SHS} &= \mathbf{SKS} + \mathbf{S}\delta\mathbf{I}_n\mathbf{S} \\ &= \mathbf{U}_R \{_d [\lambda_1, \dots, \lambda_{n-2}, 0, 0]\} \mathbf{U}_R^t + \delta\mathbf{S} \\ &= \mathbf{U}_R \{_d [\lambda_1, \dots, \lambda_{n-2}, 0, 0]\} \mathbf{U}_R^t + \delta \mathbf{U}_R \{_d [1, 1, \dots, 1, 0, 0]\} \mathbf{U}_R^t \\ &= \mathbf{U}_R \{ \{_d [\lambda_1, \dots, \lambda_{n-2}, 0, 0]\} + \delta \{_d [1, 1, \dots, 1, 0, 0]\} \} \mathbf{U}_R^t \\ &= \mathbf{U}_R \{_d [\lambda_1 + \delta, \dots, \lambda_{n-2} + \delta, 0, 0]\} \mathbf{U}_R^t \end{aligned}$$

Il est ensuite possible de décomposer  $\mathbf{U}_R$  en deux sous-matrices  $\mathbf{V}_R$  et  $\mathbf{W}_R$  avec  $\mathbf{V}_R$  contenant les vecteurs propres associés aux valeurs propres non nulles et  $\mathbf{W}_R$  les deux vecteurs propres associés aux valeurs propres 0. Ainsi,

$$\mathbf{SHS} = \mathbf{V}_R \{_d [\lambda_1 + \delta, \dots, \lambda_{n-2} + \delta]\} \mathbf{V}_R^t \quad \square$$

**Proposition 3.10.** Posons  $\dot{R} = (\mathbf{y} - \mathbf{X}\dot{\beta})^t \mathbf{H}^{-1} (\mathbf{y} - \mathbf{X}\dot{\beta}) = \mathbf{y}^t \mathbf{P}^t \mathbf{H}^{-1} \mathbf{P} \mathbf{y}$ , ainsi,

$$\dot{R} = \sum_{i=1}^{n-2} \frac{\eta_i^2}{\lambda_i + \delta}$$

Où  $[\eta_1, \dots, \eta_{n-2}]^t = \mathbf{V}_R^t \mathbf{y}$ .

*Démonstration.* Rappelons que  $\mathbf{P} = \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t \mathbf{H}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{H}^{-1}$ . Montrons premièrement que  $\mathbf{PS} = \mathbf{P}$ ,  $\mathbf{SP} = \mathbf{S}$  et  $\mathbf{PP} = \mathbf{P}$ .

$$\begin{aligned}
\mathbf{PS} &= (\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1})(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t) \\
&= \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} - \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t + \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t \\
&= \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} - \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t + \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t \\
&= \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} \\
&= \mathbf{P}
\end{aligned}$$

$$\begin{aligned}
\mathbf{SP} &= (\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t)(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1}) \\
&= \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} + \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} \\
&= \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} + \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} \\
&= \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t \\
&= \mathbf{S}
\end{aligned}$$

$$\begin{aligned}
\mathbf{PP} &= (\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1})(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1}) \\
&= \mathbf{I}_n - 2\mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} + \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} \\
&= \mathbf{I}_n - 2\mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} + \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} \\
&= \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} \\
&= \mathbf{P}
\end{aligned}$$

Ensuite, montrons que  $(\mathbf{SHS})(\mathbf{P}^t\mathbf{H}^{-1}\mathbf{P})(\mathbf{SHS}) = \mathbf{SHS}$ . Premièrement,

$$\begin{aligned}
\mathbf{HSP}^t\mathbf{H}^{-1}\mathbf{P} &= \mathbf{H}(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t)(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1})^t\mathbf{H}^{-1}(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1}) \\
&= \mathbf{H}(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t)(\mathbf{I}_n - \mathbf{H}^{-1}\mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t)\mathbf{H}^{-1}(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1}) \\
&= \mathbf{H}(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t - \mathbf{H}^{-1}\mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t + \\
&\quad \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t)\mathbf{H}^{-1}(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1}) \\
&= \mathbf{H}(\mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t - \mathbf{H}^{-1}\mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t + \\
&\quad \mathbf{X}(\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t)(\mathbf{H}^{-1} - \mathbf{H}^{-1}\mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1}) \\
&= (\mathbf{H} - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t)(\mathbf{H}^{-1} - \mathbf{H}^{-1}\mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1}) \\
&= \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} + \\
&\quad \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} \\
&= \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} + \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} \\
&= \mathbf{I}_n - \mathbf{X}(\mathbf{X}^t\mathbf{H}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{H}^{-1} \\
&= \mathbf{P}
\end{aligned}$$

Et donc,

$$\begin{aligned}
(\mathbf{SHS})(\mathbf{P}^t\mathbf{H}^{-1}\mathbf{P})(\mathbf{SHS}) &= \mathbf{SPSHS} \\
&= \mathbf{SPHS} \\
&= \mathbf{SHS}
\end{aligned}$$

Ensuite, montrons que  $(\mathbf{P}^t\mathbf{H}^{-1}\mathbf{P})(\mathbf{SHS})(\mathbf{P}^t\mathbf{H}^{-1}\mathbf{P}) = \mathbf{P}^t\mathbf{H}^{-1}\mathbf{P}$ .



$$\begin{aligned}
(\mathbf{P}^t \mathbf{H}^{-1} \mathbf{P})(\mathbf{SHS})(\mathbf{P}^t \mathbf{H}^{-1} \mathbf{P}) &= \mathbf{P}^t \mathbf{H}^{-1} \mathbf{P} \mathbf{S} \mathbf{P} \\
&= \mathbf{P}^t \mathbf{H}^{-1} \mathbf{P} \mathbf{S} \\
&= \mathbf{P}^t \mathbf{H}^{-1} \mathbf{P}
\end{aligned}$$

Aussi,

$$\begin{aligned}
[(\mathbf{SHS})(\mathbf{P}^t \mathbf{H}^{-1} \mathbf{P})]^t &= (\mathbf{S} \mathbf{P})^t \\
&= \mathbf{S}^t \\
&= \mathbf{S} \\
&= (\mathbf{SHS})(\mathbf{P}^t \mathbf{H}^{-1} \mathbf{P})
\end{aligned}$$

Et,

$$\begin{aligned}
[(\mathbf{P}^t \mathbf{H}^{-1} \mathbf{P})(\mathbf{SHS})]^t &= (\mathbf{SHS})(\mathbf{P}^t \mathbf{H}^{-1} \mathbf{P}) \\
&= [(\mathbf{SHS})(\mathbf{P}^t \mathbf{H}^{-1} \mathbf{P})]^t \\
&= (\mathbf{P}^t \mathbf{H}^{-1} \mathbf{P})(\mathbf{SHS})
\end{aligned}$$

Puisque  $(\mathbf{P}^t \mathbf{H}^{-1} \mathbf{P})$  et  $\mathbf{SHS}$  sont symétriques. Ainsi, par la définition et les propriétés du pseudo-inverse de Moore-Penrose (voir la section A.4 des annexes), il en découle que

$$\mathbf{P}^t \mathbf{H}^{-1} \mathbf{P} = (\mathbf{SHS})^+ = \mathbf{V}_R \left\{ \left[ \frac{1}{\lambda_i + \delta}, \dots, \frac{1}{\lambda_{n-2} + \delta} \right] \right\} \mathbf{V}_R^t$$

Donc finalement,

$$\begin{aligned}
\dot{R} &= \mathbf{y}^t (\mathbf{P}^t \mathbf{H}^{-1} \mathbf{P}) \mathbf{y} \\
&= \mathbf{y}^t \mathbf{V}_R \left\{_d \left[ \frac{1}{\lambda_i + \delta}, \dots, \frac{1}{\lambda_{n-2} + \delta} \right] \right\} \mathbf{V}_{R\mathbf{y}}^t \\
&= (\mathbf{V}_{R\mathbf{y}}^t)^t \left\{_d \left[ \frac{1}{\lambda_i + \delta}, \dots, \frac{1}{\lambda_{n-2} + \delta} \right] \right\} (\mathbf{V}_{R\mathbf{y}}^t) \\
&= \sum_{i=1}^{n-2} \frac{\eta_i^2}{\lambda_i + \delta}
\end{aligned}$$

Où  $[\eta_1, \dots, \eta_{n-2}]^t = \mathbf{V}_{R\mathbf{y}}^t$

□

Ensuite, remettons les expressions précédentes dans la log-vraisemblance modifiée, il s'ensuit que

$$\begin{aligned}
l^* &= -\frac{1}{2} \left[ n \log \left( \frac{2\pi \dot{R}}{n} \right) + \log(|\mathbf{H}|) + n \right] \\
&= -\frac{1}{2} \left[ n \log \left( \frac{2\pi}{n} \right) + n \log \left( \sum_{i=1}^{n-2} \frac{\eta_i^2}{\lambda_i + \delta} \right) + \log \left( \prod_{i=1}^n (\xi_i + \delta) \right) + n \right] \quad (3.3) \\
&= -\frac{1}{2} \left[ n \log \left( \frac{2\pi}{n} \right) + n \log \left( \sum_{i=1}^{n-2} \frac{\eta_i^2}{\lambda_i + \delta} \right) + \sum_{i=1}^n \log(\xi_i + \delta) + n \right]
\end{aligned}$$

Où, comme mentionné précédemment,  $[\eta_1, \dots, \eta_{n-2}]^t = \mathbf{V}_{R\mathbf{y}}^t$ , où  $\mathbf{V}_R$  est la matrice contenant les vecteurs propres de  $\mathbf{SKS}$  associés aux  $n - 2$  valeurs propres non nulles  $\lambda_i$  et où les  $\xi_i$  sont les  $n$  valeurs propres de  $\mathbf{K}$ . Notons que  $\mathbf{V}_R$ ,  $\lambda_i$  et  $\xi_i$  sont indépendants de  $\delta$ . Il est maintenant possible de déterminer la dérivée partielle de  $l^*$  par rapport à  $\delta$ .

$$\begin{aligned}
\frac{\partial l^*}{\partial \delta} = 0 &\Rightarrow \frac{n \sum_{i=1}^{n-2} \frac{\eta_i^2}{(\lambda_i + \delta)^2}}{2 \sum_{i=1}^{n-2} \frac{\eta_i^2}{\lambda_i + \delta}} - \frac{1}{2} \sum_{i=1}^n \frac{1}{\xi_i + \delta} = 0 \\
&\Rightarrow n \sum_{i=1}^{n-2} \frac{\eta_i^2}{(\lambda_i + \delta)^2} = \left( \sum_{i=1}^n \frac{1}{\xi_i + \delta} \right) \left( \sum_{i=1}^{n-2} \frac{\eta_i^2}{\lambda_i + \delta} \right)
\end{aligned} \tag{3.4}$$

La dernière équation doit être résolue numériquement pour obtenir  $\hat{\delta}$ . Pour résumer, les équations pour trouver les solutions du maximum de vraisemblance sont les suivantes.

$$\begin{aligned}
\hat{\sigma}_a^2 &= \frac{1}{n} \sum_{i=1}^{n-2} \frac{\eta_i^2}{\lambda_i + \hat{\delta}} \\
\hat{\beta} &= \left( \mathbf{X}^t \mathbf{U}_F \left\{ \frac{1}{\xi_i + \hat{\delta}} \right\}_{i=1}^n \mathbf{U}_F^t \mathbf{X} \right)^{-1} \mathbf{X}^t \mathbf{U}_F \left\{ \frac{1}{\xi_i + \hat{\delta}} \right\}_{i=1}^n \mathbf{U}_F^t \mathbf{y} \\
n \sum_{i=1}^{n-2} \frac{\eta_i^2}{(\lambda_i + \hat{\delta})^2} &= \left( \sum_{i=1}^n \frac{1}{\xi_i + \hat{\delta}} \right) \left( \sum_{i=1}^{n-2} \frac{\eta_i^2}{\lambda_i + \hat{\delta}} \right)
\end{aligned} \tag{3.5}$$

Où  $\hat{\delta} = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2}$ , où  $[\eta_1, \dots, \eta_{n-2}]^t = \mathbf{V}_R^t \mathbf{y}$ , où  $\mathbf{V}_R$  est la matrice contenant les vecteurs propres de **SKS** associés aux  $n-2$  valeurs propres non nulles  $\lambda_i$  et où  $\mathbf{U}_F$  est la matrice contenant les vecteurs propres de **K** associés aux  $n$  valeurs propres  $\xi_i$ .

### 3.3.6 Les estimateurs du maximum de vraisemblance

Puisque la matrice **K** est définie positive, si  $\hat{\delta}$  est positif,  $\hat{\mathbf{H}}$  est aussi définie positive et donc  $\hat{\sigma}_a^2$  est positif lui aussi. Voici donc les estimateurs du maximum de vraisemblance.

**Proposition 3.11.** *Si  $\dot{\delta} > 0$  alors,*

$$\begin{aligned}\hat{\beta} &= \dot{\beta} \\ \hat{\sigma}_a^2 &= \dot{\sigma}_a^2 \\ \hat{\delta} &= \dot{\delta}\end{aligned}$$

Par contre, si nous trouvons  $\dot{\delta}$  négatif et  $\dot{\sigma}_a^2$  positif, ceci indiquerait que  $\dot{\sigma}^2$  est négatif, mais ce n'est pas un modèle intéressant puisque cela signifie que les erreurs sont de variance 0. Nous pensons donc que  $\dot{\sigma}^2$  sera toujours positif comme dans les autres modèles énoncés précédemment. Ensuite, si  $\dot{\delta} < 0$  et  $\dot{\sigma}_a^2 < 0$ , posons  $\hat{\sigma}_a^2 = 0$  et remplaçons  $\sigma_a^2$  par 0 dans la log-vraisemblance. Elle devient

$$l(\mu, \beta, \sigma^2, \sigma_a^2 = 0) = -\frac{1}{2} \left[ n \log(2\pi) + n \log(\sigma^2) + \frac{1}{\sigma^2} \sum_{i=1}^n (y_i - \mu - x_i \beta)^2 \right] \quad (3.6)$$

Et nous retombons sur la régression linéaire classique, donc

$$\begin{aligned}\hat{\mu}_0 &= \bar{y} - \hat{\beta}_0 \bar{x} \\ \hat{\beta}_0 &= \frac{S_{xy}}{S_{xx}} \\ \hat{\sigma}_0^2 &= \frac{1}{n} \sum_{i=1}^n [(y_i - \bar{y}) - \hat{\beta}_0(x_i - \bar{x})]^2 \\ \hat{\sigma}_{a0}^2 &= 0\end{aligned} \quad (3.7)$$

### 3.3.7 Espérance et biais des estimateurs

Les espérances des estimateurs de  $\beta$ ,  $\sigma^2$  et de  $\sigma_a^2$  ne peuvent pas être dérivées analytiquement, car la probabilité que  $\dot{\delta}$  soit négatif n'a pas de forme explicite. Puisque les espérances ne sont pas disponibles, il est aussi impossible de savoir si les estimateurs du maximum de vraisemblance sont sans biais.

### 3.3.8 Variance des estimateurs (et leur loi)

Supposons que  $\mathbf{V}$  soit connue, ainsi,

**Proposition 3.12.**

$$\begin{aligned}\hat{\beta} &= (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} \\ \Rightarrow \hat{\beta} &\sim \mathcal{N}(\beta, (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1})\end{aligned}$$

*Démonstration.* Comme trouvé précédemment,  $\dot{\beta} = (\mathbf{X}^t \dot{\mathbf{H}}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \dot{\mathbf{H}}^{-1} \mathbf{y}$ , mais si nous supposons que les variances  $\sigma^2$  et  $\sigma_a^2$  soient connues, il s'ensuit que

$$\begin{aligned}\dot{\beta} &= (\mathbf{X}^t \mathbf{H}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{H}^{-1} \mathbf{y} \\ &= (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y}\end{aligned}$$

Dans ce cas, l'estimateur de  $\beta$  est la solution du maximum de vraisemblance  $\hat{\beta}$ .

Donc,

$$\begin{aligned}\mathbb{E}[\hat{\beta}] &= \mathbb{E}[(\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y}] \\ &= (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbb{E}[\mathbf{y}] \\ &= (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{X} \beta \\ &= \beta\end{aligned}$$

Et,

$$\begin{aligned}
\text{Var}(\hat{\beta}) &= \text{Var}((\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y}) \\
&= (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \text{Var}(\mathbf{y}) (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1})^t \\
&= (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{V} \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \\
&= (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1}
\end{aligned}
\tag*{$\square$}$$

### 3.3.9 Tests d'hypothèse et intervalles de confiance

Énonçons un test pour l'hypothèse nulle  $H_0 : \beta = 0$  seulement puisqu'il n'est pas intéressant de tester l'effet du groupe, car chaque personne représente un groupe en soi.

Alors, la statistique de Fisher suivante peut servir pour l'hypothèse nulle sur  $\beta$  ([KZW<sup>+</sup>08])

$$F = \frac{(\mathbf{M}\hat{\beta})^t (\mathbf{M}(\mathbf{X}^t \hat{\mathbf{V}}^{-1} \mathbf{X})^{-1} \mathbf{M}^t)^{-1} (\mathbf{M}\hat{\beta})}{1} \approx \mathcal{F}_{n-2}^1 \tag{3.8}$$

Où  $\mathbf{M} = [0, 1]$ .

Ce qui revient presque au même que d'utiliser la statistique suivante (statistique utilisée dans le logiciel R [R C14]).

$$Z = \frac{\hat{\beta}}{\sqrt{\widehat{\text{Var}}(\hat{\beta})}} \approx \mathcal{N}(0, 1) \tag{3.9}$$

# CHAPITRE 4

## Application aux données réelles

### 4.1 Choix du modèle

Puisque nous avons des données génétiques provenant de familles et que nous nous intéressons à la relation entre les variations génétiques et la pression sanguine, nous avons décidé d'utiliser le dernier modèle, c'est-à-dire le modèle à effets mixtes généralisé.

### 4.2 Explication de la procédure

Premièrement, il y a quelques vérifications à faire avant de procéder à une étude d'association pangénomique, c'est-à-dire que nous nous intéressons aux associations entre les variations de tout le génome avec la variable d'intérêt.

1. Nous devons premièrement exclure de l'analyse les variations génétiques (SNP) avec une fréquence d'allèle mineur (allèle ayant la plus petite fréquence dans la population) en dessous de 5 %. Nous les excluons, car ces variations n'apportent

vraiment pas beaucoup d'information puisque leur variabilité est très basse. De plus, elles sont mal représentées dans les plateformes utilisées pour mesurer des variations sur tout le génome. Aussi les erreurs techniques peuvent grandement affecter les résultats sur ces variations génétiques puisqu'elles sont plutôt rares.

2. Ensuite, nous excluons les variations qui ont un taux de valeurs manquantes de plus de 10 % puisque cela peut indiquer un problème technique lors du génotypage.
3. Finalement, nous excluons les variations qui dévient trop de la loi de distribution de Hardy-Weinberg (valeur  $p < 0.01$ ). Cette loi décrit les variations génétiques dans une population hypothétique. Par exemple, par ce test, nous allons exclure les variations pour lesquelles la presque totalité de la population est hétérozygote, ce qui est pratiquement impossible.

La prochaine étape est donc d'appliquer le modèle suivant pour toutes les variations du génome (ici, seulement les chromosomes impairs sont disponibles pour l'analyse pour des raisons propres à T2D-GENES, le consortium à qui les données appartiennent).

$$\mathbb{E}[y_i|a_i] = \mu + \beta x_i + a_i \quad (4.1)$$

Où  $y_i$  représente la pression sanguine (systolique ou diastolique) de la personne  $i$  et où  $x_i$  représente une variation génétique (SNP) de la personne  $i$  (notée 0, 1 ou 2 selon le nombre d'allèles mineurs que possède la personne  $i$ ). Nous utiliserons ce modèle avec toutes les variations génétiques disponibles après les contrôles de qualité.

## 4.3 Résultats

La meilleure façon de représenter les résultats d'une étude d'association pangénomique est de présenter le graphique Manhattan (Manhattan plot en anglais). Dans ce graphique,



nous affichons le négatif du logarithme en base 10 des valeurs  $p$  (« p-value ») sur l'axe des  $y$  et l'axe des  $x$  représente la position génomique (la position de la variation génétique sur l'ensemble de la structure de l'ADN) de chaque variation analysée. La ligne bleu en  $y = 5$  dans le graphique de Manhattan représente une significativité faible. Pour qu'une variation génétique soit considérée comme significative malgré la multitude de tests, il aurait fallu qu'elle dépasse la ligne  $y = 7$  dans le graphique, ce qui équivaut à une valeur  $p$  de moins de  $1 \times 10^{-7}$  ou  $0,05/472049$  qui est le seuil de Bonferroni.

Voici donc le graphique Manhattan pour la pression artérielle systolique.

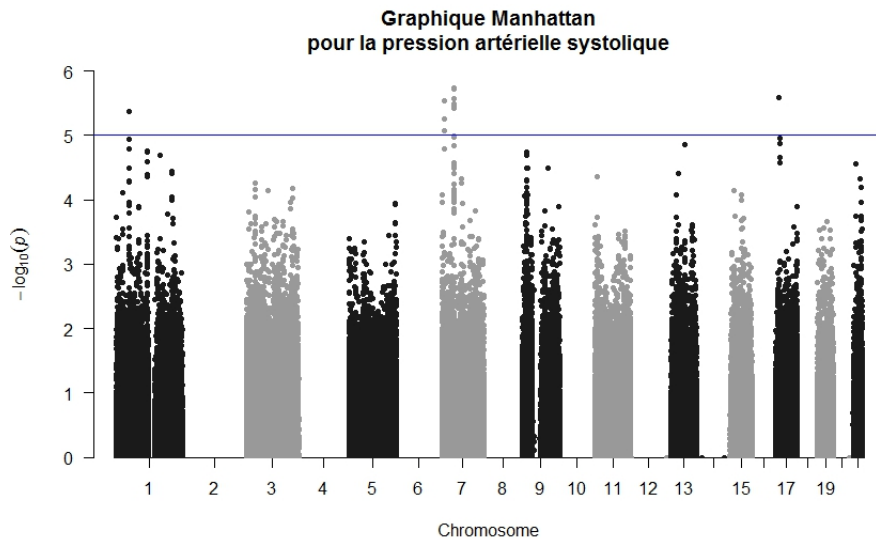


Figure 4.1 – Graphique Manhattan de la pression artérielle systolique

Et voici le graphique Manhattan pour la pression sanguine diastolique.

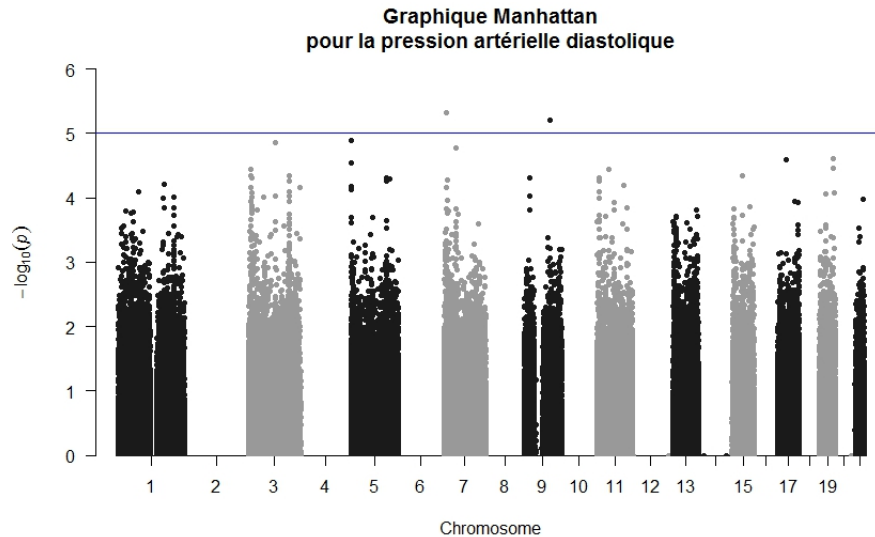


Figure 4.2 – Graphique Manhattan de la pression artérielle diastolique

Nous remarquons qu'il y a plusieurs variations du chromosome 7 (il y en a 12) qui ont une valeur  $p$  en dessous de  $1 \times 10^{-5}$ . Il y a peut-être un gène sur le chromosome 7 qui est fortement relié à la pression artérielle systolique. Par contre, nous n'avons pas trouvé d'autres études d'association pangénomique qui ont trouvé des variations génétiques significatives dans les mêmes régions que nous sur le chromosome 7. Ceci pourrait être expliqué par le fait que nous n'avons mis aucune covariable et que les distributions des pressions artérielles ne sont pas tout à fait celle d'une loi normale puisqu'il y a plusieurs cas d'hypertension. Nous remarquons aussi qu'il y a beaucoup moins de variations qui dépassent le seuil de  $1 \times 10^{-5}$  pour la tension diastolique comparativement à la tension systolique. Ceci pourrait être expliqué par le fait que la tension artérielle diastolique a une plus petite variation et peut-être aussi parce que la tension diastolique serait moins influencée par la génétique.

Il serait intéressant d'ajouter quelques covariables telles que l'âge et le sexe et de comparer

les résultats. L'ajout de covariables vient enlever du bruit et nous donner une meilleure précision donc une meilleure puissance.

Il serait aussi intéressant de voir si une transformation logarithmique des valeurs de la tension artérielle modifierait les résultats puisque l'histogramme des données montre une queue légèrement plus forte à droite. La transformation logarithmique aide à ramener la distribution de la tension artérielle plus près de celle d'une loi normale. Par contre, puisque nous travaillons avec un grand nombre de données, nous ne croyons pas que cela va affecter les résultats à grande échelle ( $n \approx 750$  dans les régressions).

# CHAPITRE 5

## Conclusion

Nous avons présenté la théorie du modèle à effets mixtes généralisé en commençant par détailler la théorie du modèle à effets mixtes usuel sous plusieurs formes. Ensuite, nous avons appliqué le modèle à effets mixtes généralisé à une base de données contenant des données provenant d'une vingtaine de grandes familles comportant en moyenne 40 individus. Nous avons pu trouver quelques variations génétiques (surtout sur le chromosome 7 (12/15) pour la tension artérielle systolique et 2 (chromosome 7 et 9) pour la tension artérielle diastolique) qui pourraient nous permettre de mieux identifier les mécanismes impliqués dans le développement de l'hypertension dans le futur.

Le modèle à effets mixtes généralisé pour des données familiales comporte plusieurs avantages comparativement aux autres tests qui peuvent être appliqués à des données familiales. En effet, plusieurs autres tests ne prennent que des variables binaires comme observations comme le « transmission disequilibrium test » (TDT), le « sibling-TDT » (S-TDT) ou encore la régression logistique conditionnelle. Il y a aussi des tests qui prennent des variables continues comme observations, mais qui négligent de l'information dépendamment du pedigree ou qui ne permettent pas d'ajouter des covariables comme le « Fa-

mily Based Assosiation Test » (FBAT). Finalement, le modèle à effets mixtes généralisé peut être utilisé avec beaucoup d'autres traits reliés à des maladies pour trouver les variations génétiques les plus associées à ces maladies.

# Bibliographie

- [Bar06] J.-B. Bardet. Théorème de Cochran et applications en statistiques. *Agrégation externe mathématiques - Université Rennes 1*, 2006.
- [BF10] N.H. Bingham and J.M. Fry. *Regression : Linear Models in Statistics*. Springer Undergraduate Mathematics Series. Springer, 2010.
- [BH12] João Carlos Alves Barata and Mahir Saleh Hussein. The Moore–Penrose Pseudoinverse : A Tutorial Review of the Theory. *Brazilian Journal of Physics*, 42(1-2) :146–165, 2012.
- [Dau84] Jean Dausset. Centre d’étude du Polymorphisme Humain. <http://www.cephb.fr/>, 1984.
- [HJ85] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge university press, 1985.
- [KZW<sup>+</sup>08] Hyun Min Kang, Noah A Zaitlen, Claire M Wade, Andrew Kirby, David Heckerman, Mark J Daly, and Eleazar Eskin. Efficient Control of Population Structure in Model Organism Association Mapping. *Genetics*, 178(3) :1709–1723, 2008.
- [Lau05] A.J. Laub. *Matrix Analysis for Scientists and Engineers*. Society for Industrial and Applied Mathematics, 2005.

- [MGB73] A.M.F. Mood, F.A. Graybill, and D.C. Boes. *Introduction to the Theory of Statistics*. McGraw-Hill series in probability and statistics. McGraw-Hill, 1973.
- [Mon13] E. Monga. *STT389 - Statistique*. Université de Sherbrooke, 2013.
- [MSN08] C.E. McCulloch, S.R. Searle, and J.M. Neuhaus. *Generalized, Linear and Mixed Models*. Wiley series in probability and mathematical statistics. Hoboken, NJ : Wiley, 2008.
- [Nat48] National Heart, Lung, and Blood Institute and Boston University School of Medicine. Framingham Heart Study. <https://www.framinghamheartstudy.org/>, 1948.
- [R C14] R Core Team. *R : A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2014.
- [SCM92] Shayle R Searle, George Casella, and Charles E McCulloch. *Variance Components*. John Wiley & Sons, 1992.
- [Sea71] S.R. Searle. *Linear Models*. Number vol. 1 in Wiley Publication in Mathematical Statistics. John Wiley & Sons, 1971.
- [Wil08] S.S. Wilks. *Mathematical Statistics*. Read Books, 2008.

# ANNEXE A

## Quelques notions mathématiques

### A.1 Vecteurs et matrices de uns

Posons  $\mathbf{I}_n$  la matrice identité d'ordre  $n$ ,  $\mathbf{J}_n$  la matrice carrée d'ordre  $n$  contenant uniquement des 1 et  $\mathbf{1}_n$  le vecteur de longueur  $n$  contenant uniquement des 1. Voici donc des propriétés qui sont triviales.

**Proposition A.1.**  $\mathbf{J}_n^2 = n\mathbf{J}_n$

**Proposition A.2.**  $(a\mathbf{I}_n + b\mathbf{J}_n)(\alpha\mathbf{I}_n + \beta\mathbf{J}_n) = a\alpha\mathbf{I}_n + (a\beta + b\alpha + b\beta n)\mathbf{J}_n$

Ces propriétés sont utiles pour le théorème suivant qui donne une forme explicite pour l'inverse d'une combinaison linéaire de  $\mathbf{I}_n$  et  $\mathbf{J}_n$ .

**Théorème A.1.**

$$(a\mathbf{I}_n + b\mathbf{J}_n)^{-1} = \frac{1}{a} \left( \mathbf{I}_n - \frac{b}{a + nb} \mathbf{J}_n \right)$$

Avec  $a \neq 0$  et  $a \neq -nb$ .



*Démonstration.* Cherchons  $\alpha$  et  $\beta$  tels que  $(a\mathbf{I}_n + b\mathbf{J}_n)(\alpha\mathbf{I}_n + \beta\mathbf{J}_n) = \mathbf{I}_n + 0\mathbf{J}_n$ . Il en découle le système d'équations suivant.

$$\begin{cases} a\alpha = 1 \\ a\beta + b\alpha + b\beta n = 0 \end{cases}$$

$$\Rightarrow \begin{cases} \alpha = \frac{1}{a} \\ \beta(a + nb) = -\frac{b}{a} \end{cases}$$

$$\Rightarrow \begin{cases} \alpha = \frac{1}{a} \\ \beta = -\frac{b}{a(a + nb)} \end{cases}$$

$$\Rightarrow (a\mathbf{I}_n + b\mathbf{J}_n)^{-1} = (\alpha\mathbf{I}_n + \beta\mathbf{J}_n) = \frac{1}{a} \left( \mathbf{I}_n - \frac{b}{a + nb} \mathbf{J}_n \right) \quad \square$$

Ensuite, cherchons la valeur de  $\det(a\mathbf{I}_n + b\mathbf{J}_n)$ . Voici un lemme pour commencer.

**Lemme A.1.**

$$\det(\mathbf{I}_n + \mathbf{B}\mathbf{A}) = \det(\mathbf{I}_p + \mathbf{A}\mathbf{B})$$

Où  $\mathbf{A}$  est une matrice  $p \times n$  et où  $\mathbf{B}$  est une matrice  $n \times p$ .

*Démonstration.* Posons

$$\mathbf{M}_{(n+p) \times (n+p)} = \begin{pmatrix} \mathbf{I}_p & -\mathbf{A} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix}$$

Ainsi,  $\mathbf{M}$  peut se décomposer de deux façons :

$$\mathbf{M} = \mathbf{L}\mathbf{N} = \begin{pmatrix} \mathbf{I}_p & \mathbf{0}_{p \times n} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{I}_p & -\mathbf{A} \\ \mathbf{0}_{n \times p} & \mathbf{I}_n + \mathbf{B}\mathbf{A} \end{pmatrix}$$

et

$$\mathbf{M} = \mathbf{P}\mathbf{R} = \begin{pmatrix} \mathbf{I}_p + \mathbf{A}\mathbf{B} & -\mathbf{A} \\ \mathbf{0}_{n \times p} & \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{I}_p & \mathbf{0}_{p \times n} \\ \mathbf{B} & \mathbf{I}_n \end{pmatrix}$$

Où  $\mathbf{0}_{a \times b}$  est la matrice  $a \times b$  contenant uniquement des 0.

Donc,

$$\det(\mathbf{M}) = \det(\mathbf{L}) \det(\mathbf{N}) = 1 \times 1 \times 1 \times \det(\mathbf{I}_n + \mathbf{B}\mathbf{A})$$

Et

$$\det(\mathbf{M}) = \det(\mathbf{P}) \det(\mathbf{R}) = \det(\mathbf{I}_p + \mathbf{A}\mathbf{B}) \times 1 \times 1 \times 1$$

$$\Rightarrow \det(\mathbf{I}_p + \mathbf{A}\mathbf{B}) = \det(\mathbf{I}_n + \mathbf{B}\mathbf{A})$$

□

Voici donc le théorème sur  $\det(a\mathbf{I}_n + b\mathbf{J}_n)$ .

**Théorème A.2.**

$$\det(a\mathbf{I}_n + b\mathbf{J}_n) = a^{n-1}(a + bn)$$

*Démonstration.*

$$\begin{aligned} \det(a\mathbf{I}_n + b\mathbf{J}_n) &= a^n \det\left(\mathbf{I}_n + \frac{b}{a}\mathbf{1}_n\mathbf{1}_n^t\right) \\ &= a^n \det\left(\mathbf{I}_1 + \frac{b}{a}\mathbf{1}_n^t\mathbf{1}_n\right) \\ &= a^n \det\left(1 + \frac{bn}{a}\right) \\ &= a^{n-1}(a + bn) \end{aligned}$$

□

## A.2 Produit de Kronecker

**Définition A.1** (Produit de Kronecker). *Le produit de Kronecker entre deux matrices  $\mathbf{A} = (a_{ij})$  et  $\mathbf{B} = (b_{ij})$  est donné par,*

$$\mathbf{A} \otimes \mathbf{B} = (a_{ij}\mathbf{B})$$

**Exemple A.1.**

$$\mathbf{I}_2 \otimes \mathbf{1}_3 = \begin{pmatrix} 1(\mathbf{1}_3) & 0(\mathbf{1}_3) \\ 0(\mathbf{1}_3) & 1(\mathbf{1}_3) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}$$

**Exemple A.2.**

$$\mathbf{1}_2 \otimes \mathbf{I}_3 = \begin{bmatrix} 1(\mathbf{I}_3) \\ 1(\mathbf{I}_3) \end{bmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Cet opérateur a aussi quelques propriétés (en assumant que les dimensions de  $\mathbf{A}$  et  $\mathbf{B}$  sont conformes). En voici deux pour commencer qui n'ont pas besoin de démonstration puisqu'elles sont assez évidentes.

**Proposition A.3.**  $(\mathbf{A} \otimes \mathbf{B})^t = \mathbf{A}^t \otimes \mathbf{B}^t$

**Proposition A.4.**  $(\mathbf{A} \otimes \mathbf{B})(\mathbf{X} \otimes \mathbf{Y}) = \mathbf{A}\mathbf{X} \otimes \mathbf{B}\mathbf{Y}$

Pour la propriété suivante, la démonstration utilise la propriété du produit.

**Proposition A.5.**  $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$

*Démonstration.*

$$\begin{aligned}(\mathbf{A} \otimes \mathbf{B})(\mathbf{A}^{-1} \otimes \mathbf{B}^{-1}) &= \mathbf{A}\mathbf{A}^{-1} \otimes \mathbf{B}\mathbf{B}^{-1} = \mathbf{I} \otimes \mathbf{I} = \mathbf{I} \\ \Rightarrow (\mathbf{A} \otimes \mathbf{B})^{-1} &= \mathbf{A}^{-1} \otimes \mathbf{B}^{-1} \quad \square\end{aligned}$$

Les preuves des prochaines propriétés se retrouvent au chapitre 13 du livre de Laub [Lau05] et elles ne seront pas reprises ici.

**Proposition A.6.**  $Rang(\mathbf{A} \otimes \mathbf{B}) = Rang(\mathbf{A})Rang(\mathbf{B})$

**Proposition A.7.**  $Trace(\mathbf{A} \otimes \mathbf{B}) = Trace(\mathbf{A})Trace(\mathbf{B})$

**Proposition A.8.** *Supposons que  $\mathbf{A}$  est une matrice carrée de dimension  $a$  et  $\mathbf{B}$  est une matrice carrée de dimension  $b$ . Alors,*

$$|\mathbf{A} \otimes \mathbf{B}| = |\mathbf{A}|^b |\mathbf{B}|^a$$

## A.3 Calcul différentiel avec des matrices et des vecteurs

Voici la définition d'une dérivée par rapport à un vecteur.

**Définition A.2** (Dérivée par rapport à un vecteur). *Soit  $\mathbf{x}$  un vecteur de dimension  $k$ , alors*

$$\frac{\partial}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_k} \end{bmatrix}$$

Voici donc quelques propriétés sur la dérivation matricielle.

**Proposition A.9.** *Si  $\mathbf{a}$  est aussi un vecteur de longueur  $k$ , alors  $\frac{\partial}{\partial \mathbf{x}}(\mathbf{a}^t \mathbf{x}) = \frac{\partial}{\partial \mathbf{x}}(\mathbf{x}^t \mathbf{a}) = \mathbf{a}$ .*

*Démonstration.*

$$\begin{aligned} \frac{\partial}{\partial \mathbf{x}}(\mathbf{a}^t \mathbf{x}) &= \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_k} \end{bmatrix} (\mathbf{a}^t \mathbf{x}) \\ &= \begin{bmatrix} \frac{\partial}{\partial x_1}(\mathbf{a}^t \mathbf{x}) \\ \vdots \\ \frac{\partial}{\partial x_k}(\mathbf{a}^t \mathbf{x}) \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial}{\partial x_1}(a_1 x_1 + \cdots + a_k x_k) \\ \vdots \\ \frac{\partial}{\partial x_k}(a_1 x_1 + \cdots + a_k x_k) \end{bmatrix} \\ &= \begin{bmatrix} a_1 \\ \vdots \\ a_k \end{bmatrix} = \mathbf{a} \end{aligned}$$

□

**Proposition A.10.** Posons  $\mathbf{y}^t = [y_1, y_2, \dots, y_p]$ . Alors

$$\frac{\partial \mathbf{y}^t}{\partial \mathbf{x}} = \text{matrice } k \times p \left\{ \frac{\partial y_j}{\partial x_i} \right\}_{i=1, j=1}^{k \quad p}$$

*Démonstration.*

$$\begin{aligned} \frac{\partial}{\partial \mathbf{x}} \mathbf{y}^t &= \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_k} \end{bmatrix} [y_1, \dots, y_p] \\ &= \begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \dots & \frac{\partial y_p}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_k} & \dots & \frac{\partial y_p}{\partial x_k} \end{pmatrix} \\ &= \text{matrice } k \times p \left\{ \frac{\partial y_j}{\partial x_i} \right\}_{i=1, j=1}^{k \quad p} \quad \square \end{aligned}$$

Ainsi,  $\frac{\partial \mathbf{x}^t}{\partial \mathbf{x}} = \mathbf{I}_k$  et pour une matrice  $\mathbf{A}$  de dimension  $k \times n$  qui ne dépend pas de  $\mathbf{x}$ ,  
 $\frac{\partial}{\partial \mathbf{x}} (\mathbf{x}^t \mathbf{A}) = \frac{\partial \mathbf{x}^t}{\partial \mathbf{x}} \mathbf{A} = \mathbf{A}$ .

**Proposition A.11.** Si  $\mathbf{u}$  et  $\mathbf{v}$  sont des vecteurs de la même longueur et qu'ils sont des fonctions de  $\mathbf{x}$ , alors,  $\frac{\partial \mathbf{u}^t \mathbf{v}}{\partial \mathbf{x}} = \frac{\partial \mathbf{u}^t}{\partial \mathbf{x}} \mathbf{v} + \frac{\partial \mathbf{v}^t}{\partial \mathbf{x}} \mathbf{u}$ .

*Démonstration.*

$$\frac{\partial (\mathbf{u}^t \mathbf{v})}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial (\mathbf{u}^t \mathbf{v})}{\partial x_1} \\ \vdots \\ \frac{\partial (\mathbf{u}^t \mathbf{v})}{\partial x_k} \end{bmatrix}$$

$$\begin{aligned}
&= \begin{bmatrix} \frac{\partial(u_1v_1 + \cdots + u_nv_n)}{\partial x_1} \\ \vdots \\ \frac{\partial(u_1v_1 + \cdots + u_nv_n)}{\partial x_k} \end{bmatrix} \\
&= \begin{bmatrix} u_1 \frac{\partial v_1}{\partial x_1} + v_1 \frac{\partial u_1}{\partial x_1} + \cdots + u_n \frac{\partial v_n}{\partial x_1} + v_n \frac{\partial u_n}{\partial x_1} \\ \vdots \\ u_1 \frac{\partial v_1}{\partial x_k} + v_1 \frac{\partial u_1}{\partial x_k} + \cdots + u_n \frac{\partial v_n}{\partial x_k} + v_n \frac{\partial u_n}{\partial x_k} \end{bmatrix} \\
&= \begin{bmatrix} v_1 \frac{\partial u_1}{\partial x_1} + \cdots + v_n \frac{\partial u_n}{\partial x_1} \\ \vdots \\ v_1 \frac{\partial u_1}{\partial x_k} + \cdots + v_n \frac{\partial u_n}{\partial x_k} \end{bmatrix} + \begin{bmatrix} u_1 \frac{\partial v_1}{\partial x_1} + \cdots + u_n \frac{\partial v_n}{\partial x_1} \\ \vdots \\ u_1 \frac{\partial v_1}{\partial x_k} + \cdots + u_n \frac{\partial v_n}{\partial x_k} \end{bmatrix} \\
&= \left\{ \frac{\partial u_j}{\partial x_i} \right\}_{i=1, j=1}^{k, n} \mathbf{v} + \left\{ \frac{\partial v_j}{\partial x_i} \right\}_{i=1, j=1}^{k, n} \mathbf{u} \\
&= \frac{\partial \mathbf{u}^t}{\partial \mathbf{x}} \mathbf{v} + \frac{\partial \mathbf{v}^t}{\partial \mathbf{x}} \mathbf{u} \quad \square
\end{aligned}$$

**Proposition A.12.** Soit la forme quadratique  $\mathbf{x}^t \mathbf{A} \mathbf{x}$ , alors  $\frac{\partial}{\partial \mathbf{x}} \mathbf{x}^t \mathbf{A} \mathbf{x} = \mathbf{A} \mathbf{x} + \mathbf{A}^t \mathbf{x}$  et si  $\mathbf{A}$  est symétrique,  $\frac{\partial}{\partial \mathbf{x}} \mathbf{x}^t \mathbf{A} \mathbf{x} = 2\mathbf{A} \mathbf{x}$ .

*Démonstration.* Inspirons-nous de la dernière proposition en prenant  $\mathbf{u}^t = \mathbf{x}^t$  et  $\mathbf{v} = \mathbf{A} \mathbf{x}$ .

$$\begin{aligned}
\frac{\partial}{\partial \mathbf{x}} \mathbf{x}^t \mathbf{A} \mathbf{x} &= \frac{\partial \mathbf{x}^t}{\partial \mathbf{x}} \mathbf{A} \mathbf{x} + \frac{\partial \mathbf{x}^t \mathbf{A}^t}{\partial \mathbf{x}} \mathbf{x} \\
&= \mathbf{A} \mathbf{x} + \mathbf{A}^t \mathbf{x} \\
&= 2\mathbf{A} \mathbf{x} \text{ si } \mathbf{A} \text{ est symétrique} \quad \square
\end{aligned}$$

## A.4 Pseudo-inverse

Voici premièrement la définition d'un pseudo-inverse

**Définition A.3** (Pseudo-inverse de Moore-Penrose). *Soit  $\mathbf{A}$  une matrice réelle non nécessairement carrée, alors il existe toujours une unique matrice  $\mathbf{A}^+$ , son pseudo-inverse, vérifiant les conditions suivantes*

1.  $\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}$
2.  $\mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+$
3.  $(\mathbf{A}\mathbf{A}^+)^t = \mathbf{A}\mathbf{A}^+$
4.  $(\mathbf{A}^+\mathbf{A})^t = \mathbf{A}^+\mathbf{A}$

*L'existence et l'unicité sont acceptées sans preuve (voir l'article de Barata et Hussein [BH12]).*

Voici maintenant une des propriétés du pseudo-inverse qui est utilisée dans le document.

**Proposition A.13.** *Soit  $\mathbf{A}$  une matrice symétrique singulière  $m \times m$  avec  $\mathbf{B}$ , le pseudo-inverse de  $\mathbf{A}$  ( $\mathbf{B} = \mathbf{A}^+$ ). Si*

$$\begin{aligned}\mathbf{A} &= [\mathbf{U} \mathbf{W}] \{_d [\lambda_1, \dots, \lambda_n, 0, \dots, 0]\} [\mathbf{U} \mathbf{W}]^t \\ &= \mathbf{U} \{_d [\lambda_1, \dots, \lambda_n]\} \mathbf{U}^t\end{aligned}$$

*avec  $[\mathbf{U} \mathbf{W}][\mathbf{U} \mathbf{W}]^t = \mathbf{I}_n$  et  $\mathbf{U}$  la matrice contenant uniquement les  $n$  vecteurs propres de  $\mathbf{A}$  associés aux  $n$  valeurs propres non nulles  $\lambda_i$ , alors*



$$\begin{aligned}
\mathbf{B} &= [\mathbf{U} \mathbf{W}] \left\{ {}_d \left[ \frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}, 0, \dots, 0 \right] \right\} [\mathbf{U} \mathbf{W}]^t \\
&= \mathbf{U} \left\{ {}_d \left[ \frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n} \right] \right\} \mathbf{U}^t
\end{aligned}$$

*Démonstration.* Il suffit de vérifier que  $\mathbf{B}$  satisfait les quatre conditions.

$$\begin{aligned}
\mathbf{ABA} &= [\mathbf{U} \mathbf{W}] \{ {}_d [\lambda_1, \dots, \lambda_n, 0, \dots, 0] \} \left\{ {}_d \left[ \frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}, 0, \dots, 0 \right] \right\} \{ {}_d [\lambda_1, \dots, \lambda_n, 0, \dots, 0] \} [\mathbf{U} \mathbf{W}]^t \\
&= [\mathbf{U} \mathbf{W}] \{ {}_d [\lambda_1, \dots, \lambda_n, 0, \dots, 0] \} [\mathbf{U} \mathbf{W}]^t \\
&= \mathbf{A}
\end{aligned}$$

$$\begin{aligned}
\mathbf{BAB} &= [\mathbf{U} \mathbf{W}] \left\{ {}_d \left[ \frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}, 0, \dots, 0 \right] \right\} \{ {}_d [\lambda_1, \dots, \lambda_n, 0, \dots, 0] \} \left\{ {}_d \left[ \frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}, 0, \dots, 0 \right] \right\} [\mathbf{U} \mathbf{W}]^t \\
&= [\mathbf{U} \mathbf{W}] \left\{ {}_d \left[ \frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}, 0, \dots, 0 \right] \right\} [\mathbf{U} \mathbf{W}]^t \\
&= \mathbf{B}
\end{aligned}$$

$$\begin{aligned}
(\mathbf{AB})^t &= \left( [\mathbf{U} \mathbf{W}] \{ {}_d [\lambda_1, \dots, \lambda_n, 0, \dots, 0] \} \left\{ {}_d \left[ \frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}, 0, \dots, 0 \right] \right\} [\mathbf{U} \mathbf{W}]^t \right)^t \\
&= \left( [\mathbf{U} \mathbf{W}] \{ {}_d [1, \dots, 1, 0, \dots, 0] \} [\mathbf{U} \mathbf{W}]^t \right)^t \\
&= [\mathbf{U} \mathbf{W}] \{ {}_d [1, \dots, 1, 0, \dots, 0] \} [\mathbf{U} \mathbf{W}]^t \\
&= \mathbf{AB}
\end{aligned}$$

$$\begin{aligned}
(\mathbf{BA})^t &= \left( [\mathbf{U} \mathbf{W}] \left\{ {}_d \left[ \frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}, 0, \dots, 0 \right] \right\} \left\{ {}_d [\lambda_1, \dots, \lambda_n, 0, \dots, 0] \right\} [\mathbf{U} \mathbf{W}]^t \right)^t \\
&= \left( [\mathbf{U} \mathbf{W}] \left\{ {}_d [1, \dots, 1, 0, \dots, 0] \right\} [\mathbf{U} \mathbf{W}]^t \right)^t \\
&= [\mathbf{U} \mathbf{W}] \left\{ {}_d [1, \dots, 1, 0, \dots, 0] \right\} [\mathbf{U} \mathbf{W}]^t \\
&= \mathbf{BA}
\end{aligned}$$

Donc  $\mathbf{B} = \mathbf{A}^+$ .

□

# ANNEXE B

## Quelques démonstrations supplémentaires

Commençons avec les lois de *SCE* et *SCR* dans le cas de l'analyse de la variance à un facteur. Nous allons démontrer leur loi dans le cas où le facteur est fixe et dans le cas où le facteur est considéré comme aléatoire.

### B.1 Loi de certaines statistiques

#### B.1.1 SCR avec effet fixe

Commençons par *SCR* pour l'analyse de la variance à un facteur fixe. Pour ce faire, voici une proposition sur la loi de  $s^2 = \sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n-1}$  dans le cas général (les observations suivent toutes une loi normale de moyenne  $\mu$  et de variance  $\sigma^2$  et elles sont toutes indépendantes).

**Proposition B.1.** *Soit*

$$y_i \sim \mathcal{N}(\mu, \sigma^2) \text{ i.i.d } \forall i = 1, \dots, n$$

Alors

$$s^2 = \sum_{i=1}^n \frac{(y_i - \bar{y})^2}{n-1} \sim \chi_{(n-1)}^2$$

*Démonstration.* Cette preuve est inspirée du cours de M. Ernest Monga [Mon13]. écrivons premièrement que

$$\frac{(n-1)s^2}{\sigma^2} = \sum_{i=1}^n \frac{(y_i - \bar{y})^2}{\sigma^2} = \sum_{i=1}^n \frac{(y_i - \mu)^2}{\sigma^2} - \frac{n(\bar{y} - \mu)^2}{\sigma^2}$$

Car,

$$\sum_{i=1}^n (y_i - \mu)^2 = \sum_{i=1}^n (y_i - \bar{y})^2 + n(\bar{y} - \mu)^2 = (n-1)s^2 + n(\bar{y} - \mu)^2$$

Ainsi,  $\frac{(n-1)s^2}{\sigma^2} = W - Z^2$ . Avec

$$W = \sum_{i=1}^n \left( \frac{y_i - \mu}{\sigma} \right)^2 \sim \chi_{(n)}^2$$

$$Z^2 = \left( \frac{\bar{y} - \mu}{\sigma/\sqrt{n}} \right)^2 \sim \chi_{(1)}^2$$

Car,

$$y_i \text{ i.i.d } \mathcal{N}(\mu, \sigma^2) \text{ pour } i = 1, \dots, n$$

et donc,

$$\bar{y} \sim \mathcal{N}(\mu, \sigma^2/n)$$

Aussi,  $\bar{y}$  est indépendant du vecteur  $[y_1 - \bar{y}, \dots, y_n - \bar{y}]^t$  ce qui donne que  $s^2$  est indépendant de  $Z^2$  et donc que

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi_{(n-1)}^2$$

Ce dernier résultat ce démontre très facilement à l'aide de la fonction génératrice des moments. □

Il est maintenant possible de trouver la loi de *SCR*.

**Proposition B.2.** *Soit*

$$y_{ij} \sim \mathcal{N}(\mu + \alpha_i, \sigma^2) \text{ indépendantes} \quad \forall i = 1, \dots, m \text{ et } \forall j = 1, \dots, n_i$$

*Dans le cadre de l'analyse de la variance avec un facteur fixe. Alors,*

$$\frac{SCR}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 \sim \chi_{(N-m)}^2$$

*Démonstration.*

$$\frac{SCR}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 = \sum_{i=1}^m \frac{(n_i - 1)s_i^2}{\sigma^2}$$

Où  $s_i^2$  est la variance échantillonnale du groupe  $i$   $\left( s_i^2 = \sum_{j=1}^{n_i} \frac{(y_{ij} - \bar{y}_i)^2}{n_i - 1} \right)$ . De plus, puisque le vecteur  $[y_{i1}, \dots, y_{in_i}]^t$  est indépendant du vecteur  $[y_{i'1}, \dots, y_{i'n_{i'}}]^t$  pour tout  $i \neq i'$ , les

$\frac{(n_i-1)s_i^2}{\sigma^2}$  sont indépendants et suivent une loi du khi carré avec  $n_i - 1$  degrés de liberté. Il en découle que

$$\begin{aligned} \frac{SCR}{\sigma^2} &\sim \chi^2\left(\sum_{i=1}^m (n_i-1)\right) \\ \Rightarrow \frac{SCR}{\sigma^2} &\sim \chi^2_{(N-m)} \quad \square \end{aligned}$$

De plus, voici la loi de l'estimateur  $\hat{\sigma}^2$ .

$$\hat{\sigma}^2 = \frac{SCR}{N} \Rightarrow \hat{\sigma}^2 = \frac{\sigma^2}{N} \frac{SCR}{\sigma^2} \Rightarrow \frac{N\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{(N-m)} \quad (\text{B.1})$$

Posons  $S^2 = \frac{N}{N-m} \hat{\sigma}^2$  comme autre estimateur de  $\sigma^2$  et voici sa loi.

$$S^2 = \frac{N}{N-m} \hat{\sigma}^2 = \frac{1}{N-m} SCR \Rightarrow S^2 = \frac{\sigma^2}{N-m} \frac{SCR}{\sigma^2} \Rightarrow \frac{(N-m)S^2}{\sigma^2} \sim \chi^2_{(N-m)} \quad (\text{B.2})$$

### B.1.2 SCR avec effet aléatoire

Transportons nous maintenant dans le cas de l'analyse de la variance avec un effet aléatoire. Il n'est pas possible de procéder comme dans la dernière section puisque maintenant, les  $y_{ij}$  ne sont plus tous indépendants. Par contre, il y a deux théorèmes qui relient les formes quadratiques à la loi du khi carré non centrée qui s'énoncent comme suit.

**Théorème B.1.** *Soit  $\mathbf{x} = [x_1, \dots, x_n]^t \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  multidimensionnelle, alors la forme quadratique  $\mathbf{x}^t \mathbf{A} \mathbf{x}$  va suivre une loi du khi carré non centrée de paramètres  $(\text{rang}(\mathbf{A}), \frac{1}{2} \boldsymbol{\mu}^t \mathbf{A} \boldsymbol{\mu})$  si et seulement si  $\mathbf{A} \boldsymbol{\Sigma}$  est idempotente.*

**Théorème B.2.** *De plus,  $\mathbf{x}^t \mathbf{B} \mathbf{x}$  est indépendant de  $\mathbf{x}^t \mathbf{A} \mathbf{x}$  si et seulement si  $\mathbf{A} \boldsymbol{\Sigma} \mathbf{B} = \mathbf{0}_{n \times n}$ .*

Les détails et les preuves se retrouvent dans [Sea71], mais présentons quand même la preuve du théorème B.1. Trois résultats préliminaires sont requis : un lemme, la fonction génératrice des moments (FGM) d'une loi du khi carré décentrée et la FGM d'une forme quadratique. Jusqu'à la fin de la preuve du théorème B.1,  $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  avec  $\mathbf{x} = [x_1, \dots, x_n]^t$ .

**Lemme B.1.** *Pour tout vecteur  $\mathbf{g}$  et toute matrice définie positive et symétrique  $\mathbf{W}$ , l'équation suivante est respectée.*

$$(2\pi)^{\frac{n}{2}} |\mathbf{W}|^{1/2} e^{\frac{1}{2} \mathbf{g}^t \mathbf{W} \mathbf{g}} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2} \mathbf{x}^t \mathbf{W}^{-1} \mathbf{x} + \mathbf{g}^t \mathbf{x} \right] dx_1 \dots dx_n$$

*Démonstration.* La preuve est évidente en considérant la densité de la loi normale multidimensionnelle  $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ . En effet,

$$\begin{aligned} (2\pi)^{\frac{n}{2}} |\boldsymbol{\Sigma}|^{1/2} &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right] dx_1 \dots dx_n \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2} \mathbf{x}^t \boldsymbol{\Sigma}^{-1} \mathbf{x} + \boldsymbol{\mu}^t \boldsymbol{\Sigma}^{-1} \mathbf{x} - \frac{1}{2} \boldsymbol{\mu}^t \boldsymbol{\Sigma} \boldsymbol{\mu} \right] dx_1 \dots dx_n \end{aligned}$$

En posant  $\mathbf{g}^t = \boldsymbol{\mu}^t \boldsymbol{\Sigma}^{-1}$  et  $\mathbf{W} = \boldsymbol{\Sigma}$ , il en découle le résultat souhaité.  $\square$

Utilisons ce lemme (B.1) pour trouver la FGM d'une forme quadratique.

**Proposition B.3.**

$$M_{\mathbf{x}^t \mathbf{A} \mathbf{x}}(t) = |\mathbf{I} - 2t \mathbf{A} \boldsymbol{\Sigma}|^{-1/2} \exp \left[ -\frac{1}{2} \boldsymbol{\mu}^t (\mathbf{I} - (\mathbf{I} - 2t \mathbf{A} \boldsymbol{\Sigma})^{-1}) \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \right]$$

*Démonstration.*

$$\begin{aligned}
M_{\mathbf{x}^t \mathbf{A} \mathbf{x}}(t) &= \mathbb{E} \left[ e^{t \mathbf{x}^t \mathbf{A} \mathbf{x}} \right] \\
&= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{t \mathbf{x}^t \mathbf{A} \mathbf{x}} f_{\mathbf{x}}(\mathbf{x}) dx_1 \dots dx_n \\
&= (2\pi)^{-\frac{n}{2}} |\boldsymbol{\Sigma}|^{-1/2} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left[ t \mathbf{x}^t \mathbf{A} \mathbf{x} - \frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^t \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right] dx_1 \dots dx_n \\
&= \frac{\exp(-\frac{1}{2} \boldsymbol{\mu}^t \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu})}{(2\pi)^{n/2} |\boldsymbol{\Sigma}|^{1/2}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2} \mathbf{x}^t (\mathbf{I} - 2t \mathbf{A} \boldsymbol{\Sigma}) \boldsymbol{\Sigma}^{-1} \mathbf{x} + \boldsymbol{\mu}^t \boldsymbol{\Sigma}^{-1} \mathbf{x} \right] dx_1 \dots dx_n
\end{aligned}$$

En utilisant le lemme B.1 avec  $\mathbf{g}^t = \boldsymbol{\mu}^t \boldsymbol{\Sigma}^{-1}$  et  $\mathbf{W} = [(\mathbf{I} - 2t \mathbf{A} \boldsymbol{\Sigma}) \boldsymbol{\Sigma}^{-1}]^{-1} = \boldsymbol{\Sigma} (\mathbf{I} - 2t \mathbf{A} \boldsymbol{\Sigma})^{-1}$ , il s'ensuit que

$$\begin{aligned}
M_{\mathbf{x}^t \mathbf{A} \mathbf{x}}(t) &= e^{-\frac{1}{2} \boldsymbol{\mu}^t \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}} |\boldsymbol{\Sigma}|^{-1/2} |\boldsymbol{\Sigma} (\mathbf{I} - 2t \mathbf{A} \boldsymbol{\Sigma})^{-1}|^{1/2} \exp \left[ \frac{1}{2} \boldsymbol{\mu}^t \boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma} (\mathbf{I} - 2t \mathbf{A} \boldsymbol{\Sigma})^{-1} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \right] \\
&= |\mathbf{I} - 2t \mathbf{A} \boldsymbol{\Sigma}|^{-1/2} \exp \left[ -\frac{1}{2} \boldsymbol{\mu}^t (\mathbf{I} - (\mathbf{I} - 2t \mathbf{A} \boldsymbol{\Sigma})^{-1}) \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \right] \quad \square
\end{aligned}$$

Ensuite, il ne manque que la FGM d'une loi du khi carré décentrée.

**Proposition B.4.** *Si la variable aléatoire  $U$  suit une loi du khi carré décentrée de paramètres  $n$  et  $\lambda$ , c'est-à-dire que sa densité est la suivante*

$$f(u) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \frac{u^{\frac{n}{2}+k-1} e^{-\frac{u}{2}}}{2^{\frac{n}{2}+k} \Gamma(\frac{n}{2} + k)}$$

Alors, sa FGM est donnée par

$$(1 - 2t)^{-\frac{n}{2}} e^{-\lambda[1-(1-2t)^{-1}]}$$

*Démonstration.* Voir le livre de Searle [Sea71]. □



Revenons maintenant à la preuve du théorème B.1.

*Démonstration du théorème B.1 (suffisance).* Supposons que  $\mathbf{A}\Sigma$  est idempotente, montrons que  $\mathbf{x}^t \mathbf{A}\mathbf{x} \sim \chi^{2r} \left( \text{rang}(\mathbf{A}), \frac{1}{2} \boldsymbol{\mu}^t \mathbf{A}\boldsymbol{\mu} \right)$ .

Par la proposition B.3, la FGM de  $\mathbf{x}^t \mathbf{A}\mathbf{x}$  est donnée par

$$\begin{aligned} M_{\mathbf{x}^t \mathbf{A}\mathbf{x}}(t) &= |\mathbf{I} - 2t\mathbf{A}\Sigma|^{-1/2} \exp \left[ -\frac{1}{2} \boldsymbol{\mu}^t (\mathbf{I} - (\mathbf{I} - 2t\mathbf{A}\Sigma)^{-1}) \Sigma^{-1} \boldsymbol{\mu} \right] \\ &= \prod_{i=1}^n \left[ (1 - 2t\lambda_i)^{-1/2} \right] \exp \left[ -\frac{1}{2} \boldsymbol{\mu}^t \left( -\sum_{k=1}^{\infty} (2t)^k (\mathbf{A}\Sigma)^k \right) \Sigma^{-1} \boldsymbol{\mu} \right] \end{aligned}$$

Où  $\lambda_i$  pour  $i = 1, \dots, n$  sont les valeurs propres de  $\mathbf{A}\Sigma$ . La deuxième égalité provient du fait que le déterminant d'une matrice est égal au produit des valeurs propres de celle-ci et en utilisant la série géométrique sur des matrices. Il faut ajouter la condition que  $\lambda_i < \frac{1}{|2t|} \forall i$  pour que la série géométrique converge. Puisque  $\mathbf{A}\Sigma$  est idempotente,  $\lambda_i = 0$  ou 1 donc la condition précédente devient  $|t| < 1/2$ . Posons  $r$  le rang de  $\mathbf{A}\Sigma$  ce qui donne qu'il y a  $r$  valeurs propres qui sont égales à 1 et  $n - r$  qui sont égales à 0. Posons, sans perte de généralité, que  $\lambda_1 = \dots = \lambda_r = 1$  et que  $\lambda_{r+1} = \dots = \lambda_n = 0$ . De plus,  $(\mathbf{A}\Sigma)^k = \mathbf{A}\Sigma$ . la FGM de  $\mathbf{x}^t \mathbf{A}\mathbf{x}$  devient

$$\begin{aligned} M_{\mathbf{x}^t \mathbf{A}\mathbf{x}}(t) &= \prod_{i=1}^r \left[ (1 - 2t)^{-1/2} \right] \exp \left[ -\frac{1}{2} \boldsymbol{\mu}^t \left( -\sum_{k=1}^{\infty} (2t)^k \right) \mathbf{A}\Sigma \Sigma^{-1} \boldsymbol{\mu} \right] \\ &= (1 - 2t)^{-r/2} \exp \left[ -\frac{1}{2} \boldsymbol{\mu}^t \left( 1 - (1 - 2t)^{-1} \right) \mathbf{A}\boldsymbol{\mu} \right] \\ &= (1 - 2t)^{-r/2} \exp \left[ -\frac{1}{2} \boldsymbol{\mu}^t \mathbf{A}\boldsymbol{\mu} (1 - (1 - 2t)^{-1}) \right] \end{aligned}$$

Par la comparaison avec la FGM de la loi du khi carré décentrée (proposition B.4), il en découle que  $\mathbf{x}^t \mathbf{A}\mathbf{x}$  suit une loi du khi carré décentrée de paramètres  $r = \text{rang}(\mathbf{A}\Sigma)$  et  $\frac{1}{2} \boldsymbol{\mu}^t \mathbf{A}\boldsymbol{\mu}$ . Puisque  $\Sigma$  est inversible,  $\mathbf{x}^t \mathbf{A}\mathbf{x} \sim \chi^{2r}(\text{rang}(\mathbf{A}), \frac{1}{2} \boldsymbol{\mu}^t \mathbf{A}\boldsymbol{\mu})$ .  $\square$

*Démonstration du théorème B.1 (nécessité).* Montrons que si  $\mathbf{x}^t \mathbf{A} \mathbf{x} \sim \chi^{2r}(r, \frac{1}{2} \boldsymbol{\mu}^t \mathbf{A} \boldsymbol{\mu})$ , alors  $\mathbf{A} \boldsymbol{\Sigma}$  est idempotente de rang  $r$ .

Puisque la loi de  $\mathbf{x}^t \mathbf{A} \mathbf{x}$  est connue, sa FGM est celle de la loi du khi carré décentrée (proposition B.4), mais aussi celle d'une forme quadratique (proposition B.3), il s'ensuit l'égalité suivante pour toute valeur de  $\boldsymbol{\mu}$ .

$$(1-2t)^{-r/2} \exp \left[ -\frac{1}{2} \boldsymbol{\mu}^t \mathbf{A} \boldsymbol{\mu} (1 - (1 - 2t^{-1})) \right] = |\mathbf{I} - 2t \mathbf{A} \boldsymbol{\Sigma}|^{-1/2} \exp \left[ -\frac{1}{2} \boldsymbol{\mu}^t (\mathbf{I} - (\mathbf{I} - 2t \mathbf{A} \boldsymbol{\Sigma})^{-1}) \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} \right]$$

Posons  $\boldsymbol{\mu} = \mathbf{0}_{n \times 1}$ . Ainsi,

$$(1 - 2t)^{-r/2} = |\mathbf{I} - 2t \mathbf{A} \boldsymbol{\Sigma}|^{-1/2}$$

Posons  $u = 2t$ ,

$$(1 - u)^r = |\mathbf{I} - u \mathbf{A} \boldsymbol{\Sigma}|$$

Posons  $\lambda_1, \dots, \lambda_n$  les valeurs propres de  $\mathbf{A} \boldsymbol{\Sigma}$ . Donc,

$$(1 - u)^r = \prod_{i=1}^n (1 - u \lambda_i)$$

Par la comparaison des puissances les plus élevées de chaque polynôme en  $u$ , il en découle qu'il doit y avoir exactement  $(n - r)$  valeurs propres égales à 0. Prenons les dernières puisque l'ordre des valeurs propres n'est pas important. Ainsi,

$$(1 - u)^r = \prod_{i=1}^r (1 - u\lambda_i)$$

$$\Rightarrow \lambda_i = 1 \text{ pour } i = 1, \dots, r$$

La dernière égalité vient du fait que lorsqu'il y a une égalité entre deux polynômes, leurs racines doivent être identiques. Le polynôme de gauche a  $r$  fois la racine  $u = 1$  et le polynôme de droite a les racines  $u = 1/\lambda_1, \dots, u = 1/\lambda_r$ . Ce qui donne que  $\lambda_i = 1$  pour  $i = 1, \dots, r$ .

Puisqu'il y a  $r$  valeurs propres égales à 1,  $(n - r)$  valeurs propres égales à 0 et que  $\mathbf{A}$  et  $\mathbf{\Sigma}$  sont symétriques avec  $\mathbf{\Sigma}$  définie positive, alors  $\mathbf{A}\mathbf{\Sigma}$  est idempotente de rang  $r$  (lemme 9 du 2<sup>e</sup> chapitre du livre de Searle [Sea71]).  $\square$

Voici une toute dernière proposition acceptée sans preuve avant de montrer le résultat sur *SCR*.

**Proposition B.5.** *Soit  $\mathbf{A}$  une matrice idempotente, alors  $\text{rang}(\mathbf{A}) = \text{trace}(\mathbf{A})$ .*

Maintenant, il est temps de trouver la loi de *SCR* dans le contexte de l'analyse de la variance à un facteur aléatoire.

**Proposition B.6.** *Soit,*

$$\mathbf{y} \sim \mathcal{N}(\mu \mathbf{1}_N, \mathbf{V})$$

Où  $\mathbf{y}$  est le vecteur de toutes les observations et où  $\mathbf{V} = \{d \sigma^2 \mathbf{I}_{n_i} + \sigma_a^2 \mathbf{J}_{n_i}\}$  est la matrice de variance-covariance de  $\mathbf{y}$ . Tout cela est dans le contexte de l'analyse de la variance avec un facteur aléatoire. Ainsi,

$$\frac{SCR}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i\cdot})^2 \sim \chi_{(N-m)}^2$$

*Démonstration.* Définissons  $\mathbf{A} = \frac{1}{\sigma^2} \left( \mathbf{I}_N - \left\{ \frac{1}{d} \frac{1}{n_i} \mathbf{J}_{n_i} \right\} \right)$ . Ainsi,

$$\begin{aligned} \mathbf{y}^t \mathbf{A} \mathbf{y} &= \frac{1}{\sigma^2} [y_{11} - \bar{y}_{1\cdot}, \dots, y_{1n_1} - \bar{y}_{1\cdot}, y_{21} - \bar{y}_{2\cdot}, \dots, y_{mn_m} - \bar{y}_{m\cdot}] \mathbf{y} \\ &= \frac{1}{\sigma^2} [y_{11}(y_{11} - \bar{y}_{1\cdot}) + y_{12}(y_{12} - \bar{y}_{1\cdot}) + \dots + y_{mn_m}(y_{mn_m} - \bar{y}_{m\cdot})] \\ &= \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij}(y_{ij} - \bar{y}_{i\cdot}) \\ &= \frac{1}{\sigma^2} \left[ \sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^m \bar{y}_{i\cdot} \sum_{j=1}^{n_i} y_{ij} \right] \\ &= \frac{1}{\sigma^2} \left[ \sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^m \sum_{j=1}^{n_i} \bar{y}_{i\cdot}^2 \right] \\ &= \frac{SCR}{\sigma^2} \end{aligned}$$

De plus,

$$\begin{aligned} \mathbf{A} \mathbf{V} &= \frac{1}{\sigma^2} \left( \mathbf{I}_N - \left\{ \frac{1}{d} \frac{1}{n_i} \mathbf{J}_{n_i} \right\} \right) \{ \sigma^2 \mathbf{I}_{n_i} + \sigma_a^2 \mathbf{J}_{n_i} \} \\ &= \frac{1}{\sigma^2} \left\{ \frac{1}{d} \mathbf{I}_{n_i} - \frac{1}{n_i} \mathbf{J}_{n_i} \right\} \{ \sigma^2 \mathbf{I}_{n_i} + \sigma_a^2 \mathbf{J}_{n_i} \} \\ &= \frac{1}{\sigma^2} \left[ \{ \sigma^2 \mathbf{I}_{n_i} \} + \{ \sigma_a^2 \mathbf{J}_{n_i} \} - \left\{ \frac{1}{d} \frac{\sigma^2}{n_i} \mathbf{J}_{n_i} \right\} - \left\{ \frac{\sigma_a^2}{d} \frac{\mathbf{J}_{n_i} \mathbf{J}_{n_i}}{n_i} \right\} \right] \\ &= \left\{ \frac{1}{d} \mathbf{I}_{n_i} - \frac{1}{n_i} \mathbf{J}_{n_i} \right\} + \left\{ \frac{\sigma_a^2}{d} \frac{\mathbf{J}_{n_i}}{\sigma^2} - \frac{\sigma_a^2}{\sigma^2} \mathbf{J}_{n_i} \right\} \\ &= \left\{ \frac{1}{d} \mathbf{I}_{n_i} - \frac{1}{n_i} \mathbf{J}_{n_i} \right\} \\ &= \sigma^2 \mathbf{A} \end{aligned}$$

Et,

$$\begin{aligned}
\mathbf{AVAV} &= \left\{ \begin{matrix} \mathbf{I}_{n_i} - \frac{1}{n_i} \mathbf{J}_{n_i} \\ d \end{matrix} \right\} \left\{ \begin{matrix} \mathbf{I}_{n_i} - \frac{1}{n_i} \mathbf{J}_{n_i} \\ d \end{matrix} \right\} \\
&= \left\{ \begin{matrix} \mathbf{I}_{n_i} \\ d \end{matrix} \right\} - \left\{ \begin{matrix} \frac{1}{n_i} \mathbf{J}_{n_i} \\ d \end{matrix} \right\} - \left\{ \begin{matrix} \frac{1}{n_i} \mathbf{J}_{n_i} \\ d \end{matrix} \right\} + \left\{ \begin{matrix} \frac{1}{n_i^2} \mathbf{J}_{n_i} \mathbf{J}_{n_i} \\ d \end{matrix} \right\} \\
&= \left\{ \begin{matrix} \mathbf{I}_{n_i} \\ d \end{matrix} \right\} - \left\{ \begin{matrix} \frac{1}{n_i} \mathbf{J}_{n_i} \\ d \end{matrix} \right\} - \left\{ \begin{matrix} \frac{1}{n_i} \mathbf{J}_{n_i} \\ d \end{matrix} \right\} + \left\{ \begin{matrix} \frac{1}{n_i} \mathbf{J}_{n_i} \\ d \end{matrix} \right\} \\
&= \left\{ \begin{matrix} \mathbf{I}_{n_i} \\ d \end{matrix} \right\} - \left\{ \begin{matrix} \frac{1}{n_i} \mathbf{J}_{n_i} \\ d \end{matrix} \right\} \\
&= \left\{ \begin{matrix} \mathbf{I}_{n_i} - \frac{1}{n_i} \mathbf{J}_{n_i} \\ d \end{matrix} \right\} \\
&= \mathbf{AV}
\end{aligned}$$

Donc  $\mathbf{AV}$  est idempotente. Par le théorème B.1, il s'ensuit que

$$\mathbf{y}^t \mathbf{A} \mathbf{y} = \frac{SCR}{\sigma^2} \sim \chi^{2'} \left( \text{rang}(\mathbf{A}), \frac{1}{2} (\boldsymbol{\mu} \mathbf{1}_N)^t \mathbf{A} (\boldsymbol{\mu} \mathbf{1}_N) \right)$$

Ensuite, trouvons des expressions simplifiées pour les paramètres de la loi du khi carré non centrée. Par la proposition B.5,  $\text{rang}(\mathbf{A})$  est donné par

$$\text{rang}(\mathbf{A}) = \text{rang}(\sigma^2 \mathbf{A}) = \text{trace}(\sigma^2 \mathbf{A}) = \sum_{i=1}^m \sum_{j=1}^{n_i} \left( 1 - \frac{1}{n_i} \right) = N - m$$

Ensuite,

$$\frac{1}{2} (\boldsymbol{\mu} \mathbf{1}_N)^t \mathbf{A} (\boldsymbol{\mu} \mathbf{1}_N) = \frac{\mu^2}{2} \mathbf{1}_N^t \mathbf{A} \mathbf{1}_N = \frac{\mu^2}{2} \cdot \mathbf{0}_{1 \times N} \cdot \mathbf{1}_N = 0$$

Le résultat final est donc que

$$\frac{SCR}{\sigma^2} \sim \chi_{(N-m)}^2 \quad \square$$

### B.1.3 SCE avec effet fixe

**Proposition B.7.** *Soit*

$$y_{ij} \sim \mathcal{N}(\mu + \alpha_i, \sigma^2) \text{ indépendantes} \quad \forall i = 1, \dots, m \text{ et } \forall j = 1, \dots, n_i$$

Dans le cadre de l'analyse de la variance avec un facteur fixe. Ainsi, sous  $H_0 : \alpha_1 = \dots = \alpha_m = 0$ , la loi de SCE est donnée par

$$\frac{SCE}{\sigma^2} = \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (\bar{y}_{i.} - \bar{y}_{..})^2 \sim \chi_{(m-1)}^2$$

*Démonstration.* Ce qui suit est inspiré de [BF10]. Premièrement, voici une identité.

$$SCT = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2 + \sum_{i=1}^m \sum_{j=1}^{n_i} (\bar{y}_{i.} - \bar{y}_{..})^2 = SCR + SCE$$

Prenons les  $y_{ij}$  tous ensembles avec leur moyenne  $\bar{y}_{..}$ . leur variance échantillonnale  $S^2 = \frac{1}{N-1} \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{..})^2 = \frac{SCT}{(N-1)}$ . La loi de  $S^2$  sous  $H_0 : \alpha_1 = \dots = \alpha_m = 0$  est donc connue, en effet,

$$\frac{(N-1)S^2}{\sigma^2} \sim \chi_{(N-1)}^2 \Rightarrow \frac{SCT}{\sigma^2} = \frac{(N-1)S^2}{\sigma^2} \sim \chi_{(N-1)}^2$$

La contrainte d'être sous  $H_0 : \alpha_1 = \dots = \alpha_m = 0$  est importante pour pouvoir utiliser la proposition sur la variance échantillonnale qui requiert que toutes les variables aléatoires soient i.i.d.

De plus,  $SCR = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)^2 = \sum_{i=1}^m (n_i - 1) s_i^2$  et  $SCE = \sum_{i=1}^m \sum_{j=1}^{n_i} (\bar{y}_i - \bar{y}_{..})^2 = \sum_{i=1}^m n_i (\bar{y}_i - \bar{y}_{..})^2$ . Aussi,  $s_i^2$  est indépendant de  $\bar{y}_i$ . pour tout  $i$  et  $s_i^2$  est indépendant de  $\bar{y}_{i'}$ . pour tout  $i \neq i'$  puisqu'ils sont calculés à partir de deux groupes différents. Donc,  $s_i^2$  est aussi indépendant de  $\bar{y}_{..}$ . Ce qui donne finalement que  $SCE$  est indépendant de  $SCR$ .

Remettons tout cela en perspective.  $SCT/\sigma^2 = SCR/\sigma^2 + SCE/\sigma^2$ ,  $SCR/\sigma^2$  est indépendant de  $SCE/\sigma^2$  et finalement  $SCT/\sigma^2 \sim \chi_{(N-1)}^2$  et  $SCR/\sigma^2 \sim \chi_{(N-m)}^2$ . Tout cela implique, par la FGM de la loi du khi carré, que sous  $H_0 : \alpha_1 = \dots = \alpha_m = 0$ ,

$$\begin{aligned} SCE/\sigma^2 &\sim \chi_{((N-1)-(N-m))}^2 \\ \Rightarrow SCE/\sigma^2 &\sim \chi_{(m-1)}^2 \quad \square \end{aligned}$$

#### B.1.4 SCE avec effet aléatoire

Nous allons présenter deux propositions sur  $SCE$  ici, une pour le cas où les données sont équilibrées et une autre pour le cas où les données sont non équilibrées. Commençons par celle pour le cas où les données sont équilibrées.

**Proposition B.8.** *Soit,*

$$\mathbf{y} \sim \mathcal{N}(\mu \mathbf{1}_N, \mathbf{V})$$

Où  $\mathbf{y}$  est le vecteur de toutes les observations et où  $\mathbf{V} = \{d \sigma^2 \mathbf{I}_n + \sigma_a^2 \mathbf{J}_n\}$  est la matrice de variance-covariance de  $\mathbf{y}$ . Tout cela est dans le contexte de l'analyse de la variance avec un facteur aléatoire dans le cas où les données sont équilibrées. Ainsi,

$$\frac{SCE}{\sigma^2 + n\sigma_a^2} = \frac{1}{\sigma^2 + n\sigma_a^2} \sum_{i=1}^m \sum_{j=1}^n (\bar{y}_{i.} - \bar{y}_{..})^2 \sim \chi_{(m-1)}^2$$

*Démonstration.* Procédons de façon analogue à la démonstration de la loi de  $SCE$  avec effet fixe.

$$SCE = \sum_{i=1}^m \sum_{j=1}^n (\bar{y}_{i.} - \bar{y}_{..})^2 = n \sum_{i=1}^m (\bar{y}_{i.} - \bar{y}_{..})^2$$

Posons,  $\bar{S}^2 = \frac{1}{m-1} \sum_{i=1}^m (\bar{y}_{i.} - \bar{y}_{..})^2$  qui représente la variance échantillonnale des  $\bar{y}_{i.}$ . Ensuite, en utilisant les propriétés de la loi normale multidimensionnelle, il en découle que

$$\bar{y}_{i.} \sim \mathcal{N}\left(\mu, \sigma_a^2 + \frac{1}{n}\sigma^2\right) \text{ indépendantes}$$

Donc, comme montré précédemment pour  $s^2$  et en effectuant quelques substitutions, la loi de  $\bar{S}^2$  est donnée par

$$\frac{(m-1)\bar{S}^2}{\sigma_a^2 + \frac{1}{n}\sigma^2} \sim \chi_{(m-1)}^2$$

Donc,

$$SCE = n(m-1)\bar{S}^2 = (n\sigma_a^2 + \sigma^2) \frac{(m-1)\bar{S}^2}{\sigma_a^2 + \frac{1}{n}\sigma^2} \Rightarrow \frac{SCE}{n\sigma_a^2 + \sigma^2} \sim \chi_{(m-1)}^2 \quad \square$$

Voici maintenant la proposition pour le cas où les données sont non équilibrées.

**Proposition B.9.** *Soit,*

$$\mathbf{y} \sim \mathcal{N}(\mu \mathbf{1}_N, \mathbf{V})$$





$$\begin{aligned}
\mathbf{CVCV} &= \left( \left\{ \frac{1}{d} \mathbf{J}_{n_i} \right\} - \frac{1}{N} \mathbf{J}_N \right) \left( \left\{ \frac{1}{d} \mathbf{J}_{n_i} \right\} - \frac{1}{N} \mathbf{J}_N \right) \\
&= \left\{ \frac{1}{d} \mathbf{J}_{n_i} \right\} \left\{ \frac{1}{d} \mathbf{J}_{n_i} \right\} - \left\{ \frac{1}{d} \mathbf{J}_{n_i} \right\} \frac{1}{N} \mathbf{J}_N - \frac{1}{N} \mathbf{J}_N \left\{ \frac{1}{d} \mathbf{J}_{n_i} \right\} + \frac{1}{N} \mathbf{J}_N \frac{1}{N} \mathbf{J}_N \\
&= \left\{ \frac{1}{d} \frac{1}{n_i} \mathbf{J}_{n_i} \mathbf{J}_{n_i} \right\} - \frac{1}{N} \mathbf{J}_N - \frac{1}{N} \mathbf{J}_N + \frac{1}{N} \mathbf{J}_N \\
&= \left\{ \frac{1}{d} \frac{1}{n_i} \mathbf{J}_{n_i} \right\} - \frac{1}{N} \mathbf{J}_N \\
&= \mathbf{CV}
\end{aligned}$$

Donc  $\mathbf{CV}$  est idempotente. Par le théorème B.1, il en découle que

$$\mathbf{y}^t \mathbf{C} \mathbf{y} = \frac{SCE}{\sigma^2} \sim \chi^{2'} \left( \text{rang}(\mathbf{C}), \frac{1}{2} (\boldsymbol{\mu} \mathbf{1}_N)^t \mathbf{C} (\boldsymbol{\mu} \mathbf{1}_N) \right)$$

Ensuite, trouvons des expressions simplifiées pour les paramètres de la loi du khi carré non centrée. Pour son rang, utilisons la proposition B.5. Ainsi,

$$\begin{aligned}
\text{rang}(\mathbf{C}) &= \text{rang}(\mathbf{CV}) \\
&= \text{trace}(\mathbf{CV}) \\
&= \sum_{i=1}^m \sum_{j=1}^{n_i} \left( \frac{1}{n_i} - \frac{1}{N} \right) \\
&= m - 1
\end{aligned}$$

Et,  $\frac{1}{2} (\boldsymbol{\mu} \mathbf{1}_N)^t \mathbf{C} (\boldsymbol{\mu} \mathbf{1}_N) = 0$ . Le résultat final est donc que

$$\frac{SCE}{\sigma^2} \sim \chi_{(m-1)}^2$$

□

### B.1.5 Indépendance entre SCR et SCE avec effet fixe

Montré dans la section B.1.3 sous  $H_0 : \alpha_1 = \dots = \alpha_m = 0$ .

### B.1.6 Indépendance entre SCR et SCE avec effet aléatoire

Nous allons encore procéder en deux parties. Premièrement voici la proposition dans le cas où les données sont équilibrées.

**Proposition B.10.** *SCE et SCR sont indépendants dans le cadre de l'analyse de la variance avec un facteur aléatoire dans le cas où les données sont équilibrées.*

*Démonstration.* écrivons  $\frac{SCE}{n\sigma_a^2 + \sigma^2} = \mathbf{y}^t \mathbf{B} \mathbf{y}$  avec

$$\mathbf{B} = \frac{1}{n\sigma_a^2 + \sigma^2} \left( \mathbf{I}_m \otimes \left( \frac{1}{n} \mathbf{J}_n \right) - \frac{1}{N} \mathbf{J}_N \right) = \frac{1}{n\sigma_a^2 + \sigma^2} \left( \mathbf{I}_m - \frac{1}{m} \mathbf{J}_m \right) \otimes \left( \frac{1}{n} \mathbf{J}_n \right)$$

Le symbole  $\otimes$  représente le produit de Kronecker, aussi appelé produit direct. Pour la définition et quelques propriétés, voir la section A.2. En effet,

$$\begin{aligned} \mathbf{y}^t \mathbf{B} \mathbf{y} &= \frac{1}{n\sigma_a^2 + \sigma^2} \underbrace{[\bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot}, \dots, \bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot}, \dots]}_{n \text{ fois}} \dots \underbrace{[\bar{y}_{m\cdot} - \bar{y}_{\cdot\cdot}, \dots, \bar{y}_{m\cdot} - \bar{y}_{\cdot\cdot}]}_{n \text{ fois}} \mathbf{y} \\ &= \frac{1}{n\sigma_a^2 + \sigma^2} \sum_{i=1}^m \sum_{j=1}^n y_{ij} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}) \\ &= \frac{1}{n\sigma_a^2 + \sigma^2} \sum_{i=1}^m \sum_{j=1}^n (\bar{y}_{i\cdot}^2 - \bar{y}_{\cdot\cdot}^2) \\ &= \frac{SCE}{n\sigma_a^2 + \sigma^2} \end{aligned}$$

Et

$$\begin{aligned}
\mathbf{AVB} &= \frac{1}{n\sigma_a^2 + \sigma^2} \left( \left\{ \mathbf{I}_n - \frac{1}{n} \mathbf{J}_n \right\} \right) \left( \left( \mathbf{I}_m - \frac{1}{m} \mathbf{J}_m \right) \otimes \left( \frac{1}{n} \mathbf{J}_n \right) \right) \\
&= \frac{1}{n\sigma_a^2 + \sigma^2} \left( \mathbf{I}_m \otimes \left( \mathbf{I}_n - \frac{1}{n} \mathbf{J}_n \right) \right) \left( \left( \mathbf{I}_m - \frac{1}{m} \mathbf{J}_m \right) \otimes \left( \frac{1}{n} \mathbf{J}_n \right) \right) \\
&= \frac{1}{n\sigma_a^2 + \sigma^2} \left( \left( \mathbf{I}_m - \frac{1}{m} \mathbf{J}_m \right) \otimes \left( \frac{1}{n} \mathbf{J}_n - \frac{1}{n} \mathbf{J}_n \right) \right) \\
&= \frac{1}{n\sigma_a^2 + \sigma^2} \left( \left( \mathbf{I}_m - \frac{1}{m} \mathbf{J}_m \right) \otimes \mathbf{0}_{n \times n} \right) \\
&= \mathbf{0}_{mn \times mn}
\end{aligned}$$

Puisque  $\mathbf{AVB} = \mathbf{0}_{mn \times mn}$ ,  $\mathbf{y}^t \mathbf{B} \mathbf{y}$  est indépendant de  $\mathbf{y}^t \mathbf{A} \mathbf{y}$  donc  $SCE$  est indépendant de  $SCR$ .  $\square$

Pour le cas où les données sont non équilibrées, regardons l'indépendance sous  $H_0 : \sigma_a^2 = 0$ .

**Proposition B.11.** *SCE et SCR sont indépendants dans le cadre de l'analyse de la variance avec un facteur aléatoire sous  $H_0 : \sigma_a^2 = 0$ .*

*Démonstration.* Montrons que  $\mathbf{AVC} = \mathbf{0}_{N \times N}$

$$\begin{aligned}
\mathbf{AVC} &= \frac{1}{\sigma^2} \left( \mathbf{I}_N - \left\{ \frac{1}{d n_i} \mathbf{J}_{n_i} \right\} \right) \sigma^2 \mathbf{I}_N \frac{1}{\sigma^2} \left( \left\{ \frac{1}{d n_i} \mathbf{J}_{n_i} \right\} - \frac{1}{N} \mathbf{J}_N \right) \\
&= \frac{1}{\sigma^2} \left( \left\{ \frac{1}{d n_i} \mathbf{J}_{n_i} \right\} - \frac{1}{N} \mathbf{J}_N - \left\{ \frac{1}{d n_i} \mathbf{J}_{n_i} \right\} + \left\{ \frac{1}{d n_i} \mathbf{J}_{n_i} \right\} \frac{1}{N} \mathbf{J}_N \right) \\
&= \frac{1}{\sigma^2} \mathbf{0}_{N \times N} \\
&= \mathbf{0}_{N \times N}
\end{aligned}$$

Donc, sous  $H_0 : \sigma_a^2 = 0$ ,  $SCE$  et  $SCR$  sont indépendants.  $\square$

## B.2 Variance et covariance conditionnelle

**Théorème B.3.**

$$\text{Var}(y) = \text{Var}_u(\mathbb{E}[y|u]) + \mathbb{E}_u[\text{Var}(y|u)]$$

*Démonstration.* Premièrement,

$$\mathbb{E}_u[\text{Var}(y|u)] = \mathbb{E}_u[\mathbb{E}[y^2|u] - \mathbb{E}[y|u]^2] = \mathbb{E}[y^2] - \mathbb{E}_u[\mathbb{E}[y|u]^2]$$

Et

$$\text{Var}_u(\mathbb{E}[y|u]) = \mathbb{E}_u[\mathbb{E}[y|u]^2] - \mathbb{E}_u[\mathbb{E}[y|u]]^2 = \mathbb{E}_u[\mathbb{E}[y|u]^2] - \mathbb{E}[y]^2$$

Donc,

$$\text{Var}_u(\mathbb{E}[y|u]) + \mathbb{E}_u[\text{Var}(y|u)] = \mathbb{E}[y^2] - \mathbb{E}[y]^2 = \text{Var}(y) \quad \square$$

Voici maintenant la preuve de l'équivalent pour la covariance.

**Théorème B.4.**

$$\text{cov}(y, w) = \text{cov}_u(\mathbb{E}[y|u], \mathbb{E}[w|u]) + \mathbb{E}_u[\text{cov}(y, w|u)]$$

*Démonstration.* Premièrement

$$\text{cov}_u(\mathbb{E}[y|u], \mathbb{E}[w|u]) = \mathbb{E}_u[\mathbb{E}[y|u]\mathbb{E}[w|u]] - \mathbb{E}_u[\mathbb{E}[y|u]]\mathbb{E}_u[\mathbb{E}[w|u]] = \mathbb{E}_u[\mathbb{E}[y|u]\mathbb{E}[w|u]] - \mathbb{E}[y]\mathbb{E}[w]$$

et

$$\mathbb{E}_u[\text{cov}(y, w|u)] = \mathbb{E}_u[\mathbb{E}[yw|u] - \mathbb{E}[y|u]\mathbb{E}[w|u]] = \mathbb{E}[yw] - \mathbb{E}_u[\mathbb{E}[y|u]\mathbb{E}[w|u]]$$

Donc,

$$\text{cov}_u(\mathbb{E}[y|u], \mathbb{E}[w|u]) + \mathbb{E}_u[\text{cov}(y, w|u)] = \mathbb{E}[yw] - \mathbb{E}[y]\mathbb{E}[w] = \text{cov}(y, w) \quad \square$$

## B.3 Autres preuves pour l'ANOVA à un facteur aléatoire

Rappelons le modèle en premier lieu :

$$\begin{aligned} \mathbb{E}[y_{ij}|a_i] &= \mu + a_i \\ y_{ij}|a_i &\sim \mathcal{N}(\mu + a_i, \sigma^2) \text{ indépendantes} \\ a_i &\sim \text{i.i.d } \mathcal{N}(0, \sigma_a^2) \end{aligned} \tag{B.3}$$

Où  $i = 1, \dots, m$  et  $j = 1, \dots, n_i$ . Et

$$\begin{aligned} \mathbb{E}[y_{ij}] &= \mu \\ \text{Var}(y_{ij}) &= \sigma^2 + \sigma_a^2 \\ \text{cov}(y_{ij}, y_{ik}) &= \sigma_a^2 \end{aligned}$$

### B.3.1 Quelques espérances et variances

Quelques espérances et variances de certaines fonctions des  $y_{ij}$  sont requises pour commencer.

**Proposition B.12.**  $\mathbb{E}[\bar{y}_{..}^2] = \mu^2 + \frac{1}{N}\sigma^2 + \sum_{i=1}^m \frac{n_i^2}{N^2}\sigma_a^2$

*Démonstration.*

$$\begin{aligned}
\mathbb{E}[\bar{y}_{..}^2] &= \frac{1}{N^2} \mathbb{E} \left[ \left( \sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij} \right) \left( \sum_{k=1}^m \sum_{l=1}^{n_k} y_{kl} \right) \right] \\
&= \frac{1}{N^2} \mathbf{E} \left[ \sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij}^2 \right] + \frac{1}{N^2} \mathbf{E} \left[ \sum_{i=1}^m \sum_{j=1}^{n_i} \sum_{k \neq j}^{n_i} y_{ij} y_{ik} \right] + \frac{1}{N^2} \mathbf{E} \left[ \sum_{i=1}^m \sum_{k \neq i}^m \sum_{j=1}^{n_i} \sum_{l=1}^{n_k} y_{ij} y_{kl} \right] \\
&= \frac{1}{N^2} \sum_{i=1}^m \sum_{j=1}^{n_i} \mathbf{E} [y_{ij}^2] + \frac{1}{N^2} \sum_{i=1}^m \sum_{j=1}^{n_i} \sum_{k \neq j}^{n_i} \mathbf{E} [y_{ij} y_{ik}] + \frac{1}{N^2} \sum_{i=1}^m \sum_{k \neq i}^m \sum_{j=1}^{n_i} \sum_{l=1}^{n_k} \mathbf{E} [y_{ij} y_{kl}] \\
&= \frac{1}{N^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (\mu^2 + \sigma^2 + \sigma_a^2) + \frac{1}{N^2} \sum_{i=1}^m \sum_{j=1}^{n_i} \sum_{k \neq j}^{n_i} (\mu^2 + \sigma_a^2) + \frac{1}{N^2} \sum_{i=1}^m \sum_{k \neq i}^m \sum_{j=1}^{n_i} \sum_{l=1}^{n_k} (\mu^2) \\
&= \frac{N}{N^2} (\mu^2 + \sigma^2 + \sigma_a^2) + \frac{\sum_{i=1}^m (n_i^2) - N}{N^2} (\mu^2 + \sigma_a^2) + \frac{N^2 - \sum_{i=1}^m n_i^2}{N^2} (\mu^2) \\
&= \mu^2 + \frac{1}{N} \sigma^2 + \frac{\sum_{i=1}^m n_i^2}{N^2} \sigma_a^2 \\
&= \mu^2 + \frac{1}{mn} \sigma^2 + \frac{1}{m} \sigma_a^2 \quad \text{si les données sont équilibrées} \quad \square
\end{aligned}$$

**Proposition B.13.**  $\mathbb{E}[\bar{y}_{i.}^2] = \mu^2 + \frac{1}{n_i} \sigma^2 + \sigma_a^2$

*Démonstration.*

$$\begin{aligned}
\mathbb{E}[\bar{y}_{i.}^2] &= \frac{1}{n_i^2} \mathbb{E} \left[ \left( \sum_{j=1}^{n_i} y_{ij} \right) \left( \sum_{k=1}^{n_i} y_{ik} \right) \right] \\
&= \frac{1}{n_i^2} \sum_{j=1}^{n_i} \mathbb{E} [y_{ij}^2] + \frac{1}{n_i^2} \sum_{j=1}^{n_i} \sum_{k \neq j}^{n_i} \mathbb{E} [y_{ij} y_{ik}] \\
&= \frac{1}{n_i^2} \sum_{j=1}^{n_i} (\mu^2 + \sigma^2 + \sigma_a^2) + \frac{1}{n_i^2} \sum_{j=1}^{n_i} \sum_{k \neq j}^{n_i} (\mu^2 + \sigma_a^2) \\
&= \frac{n_i}{n_i^2} (\mu^2 + \sigma^2 + \sigma_a^2) + \frac{n_i(n_i - 1)}{n_i^2} (\mu^2 + \sigma_a^2) \\
&= \mu^2 + \frac{1}{n_i} \sigma^2 + \sigma_a^2 \quad \square
\end{aligned}$$

**Proposition B.14.**  $Var(\bar{y}_{..}) = \frac{1}{N}\sigma^2 + \sum_{i=1}^m \frac{n_i^2}{N^2}\sigma_a^2$

*Démonstration.*

$$\begin{aligned}
Var(\bar{y}_{..}) &= \mathbb{E}[\bar{y}_{..}^2] - \mathbb{E}[\bar{y}_{..}]^2 \\
&= \mu^2 + \frac{1}{N}\sigma^2 + \frac{\sum_{i=1}^m n_i^2}{N^2}\sigma_a^2 - \mu^2 \\
&= \frac{1}{N}\sigma^2 + \frac{\sum_{i=1}^m n_i^2}{N^2}\sigma_a^2 \\
&= \frac{n\sigma_a^2 + \sigma^2}{mn} \quad \text{si les données sont équilibrées} \quad \square
\end{aligned}$$

**Proposition B.15.**  $Var(\bar{y}_{i.}) = \frac{1}{n_i}\sigma^2 + \sigma_a^2$

*Démonstration.*

$$\begin{aligned}
Var(\bar{y}_{i.}) &= \mathbb{E}[\bar{y}_{i.}^2] - \mathbb{E}[\bar{y}_{i.}]^2 \\
&= \mu^2 + \frac{1}{n_i}\sigma^2 + \sigma_a^2 - \mu^2 \\
&= \frac{1}{n_i}\sigma^2 + \sigma_a^2 \quad \square
\end{aligned}$$

**Proposition B.16.**  $\mathbb{E}[CMR] = \sigma^2$

*Démonstration.*

$$\begin{aligned}
\mathbb{E}[CMR] &= \frac{1}{N-m}\mathbb{E}\left[\sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_{i.})^2\right] \\
&= \frac{1}{N-m}\sum_{i=1}^m \sum_{j=1}^{n_i} \mathbb{E}[y_{ij}^2 - 2y_{ij}\bar{y}_{i.} + \bar{y}_{i.}^2] \\
&= \frac{1}{N-m}\sum_{i=1}^m \sum_{j=1}^{n_i} \left(\mathbb{E}[y_{ij}^2] - 2\mathbb{E}[\bar{y}_{i.}^2] + \mathbb{E}[\bar{y}_{i.}^2]\right) \\
&= \frac{1}{N-m}\sum_{i=1}^m \sum_{j=1}^{n_i} \left(\left(\mu^2 + \sigma^2 + \sigma_a^2\right) - \left(\mu^2 + \sigma_a^2 + \frac{1}{n_i}\sigma^2\right)\right)
\end{aligned}$$



$$\begin{aligned}
&= \frac{1}{N-m} \sum_{i=1}^m \sum_{j=1}^{n_i} \left( \left(1 - \frac{1}{n_i}\right) \sigma^2 \right) \\
&= \frac{\sigma^2}{N-m} \sum_{i=1}^m (n_i - 1) \\
&= \frac{(N-m)\sigma^2}{N-m} \\
&= \sigma^2
\end{aligned}$$

□

$\mathbb{E}[CME]$  est seulement nécessaire dans le cas où les données sont équilibrées.

**Proposition B.17.**  $\mathbb{E}[CME] = \sigma^2 + n\sigma_a^2$  lorsque les données sont équilibrées.

*Démonstration.*

$$\begin{aligned}
\mathbb{E}[CME] &= \frac{1}{m-1} \mathbb{E} \left[ \sum_{i=1}^m n(\bar{y}_i - \bar{y}_{..})^2 \right] \\
&= \frac{n}{m-1} \mathbb{E} \left[ \sum_{i=1}^m (\bar{y}_i^2 - 2\bar{y}_i \bar{y}_{..} + \bar{y}_{..}^2) \right] \\
&= \frac{n}{m-1} \mathbb{E} \left[ \sum_{i=1}^m \bar{y}_i^2 \right] - \frac{nm}{m-1} \mathbb{E} [\bar{y}_{..}^2] \\
&= \frac{n}{m-1} \sum_{i=1}^m \left( \mu^2 + \sigma_a^2 + \frac{1}{n} \sigma^2 \right) - \frac{nm}{m-1} \left( \mu^2 + \frac{1}{mn} \sigma^2 + \frac{1}{m} \sigma_a^2 \right) \\
&= \sigma^2 + n\sigma_a^2
\end{aligned}$$

□

### B.3.2 Variance asymptotique des paramètres dans le cas où les données sont non équilibrées

Les dérivées secondes de la log-vraisemblance ont été calculées à la section 2.2.2f), mais les revoici.

$$\begin{aligned}
l_{\mu\mu} &= -\sum_{i=1}^m \frac{n_i}{\lambda_i} \\
l_{\mu\sigma^2} &= -\sum_{i=1}^m \frac{n_i(\bar{y}_{i\cdot} - \mu)}{\lambda_i^2} \\
l_{\mu\sigma_a^2} &= -\sum_{i=1}^m \frac{n_i^2(\bar{y}_{i\cdot} - \mu)}{\lambda_i^2} \\
l_{\sigma^2\sigma_a^2} &= \frac{1}{2} \sum_{i=1}^m \frac{n_i}{\lambda_i^2} - \sum_{i=1}^m \frac{n_i^2(\bar{y}_{i\cdot} - \mu)^2}{\lambda_i^3} \\
l_{\sigma^2\sigma^2} &= \frac{N-m}{2\sigma^4} - \frac{SCR}{\sigma^6} + \frac{1}{2} \sum_{i=1}^m \frac{1}{\lambda_i^2} - \sum_{i=1}^m \frac{n_i(\bar{y}_{i\cdot} - \mu)^2}{\lambda_i^3} \\
l_{\sigma_a^2\sigma_a^2} &= \frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} - \sum_{i=1}^m \frac{n_i^3(\bar{y}_{i\cdot} - \mu)^2}{\lambda_i^3}
\end{aligned} \tag{B.4}$$

Où  $\lambda_i = \sigma^2 + n_i\sigma_a^2$ . Trouvons maintenant la négation des espérances des dérivées partielles mentionnées plus haut.

$$-\mathbb{E}[l_{\mu\mu}] = -\mathbb{E}\left[-\sum_{i=1}^m \frac{n_i}{\lambda_i}\right] = \sum_{i=1}^m \frac{n_i}{\lambda_i} \tag{B.5}$$

$$-\mathbb{E}[l_{\mu\sigma^2}] = -\mathbb{E}\left[-\sum_{i=1}^m \frac{n_i(\bar{y}_{i\cdot} - \mu)}{\lambda_i^2}\right] = \sum_{i=1}^m \frac{n_i(\mathbb{E}[\bar{y}_{i\cdot}] - \mu)}{\lambda_i^2} = 0 \tag{B.6}$$

$$-\mathbb{E}[l_{\mu\sigma_a^2}] = -\mathbb{E}\left[-\sum_{i=1}^m \frac{n_i^2(\bar{y}_{i\cdot} - \mu)}{\lambda_i^2}\right] = \sum_{i=1}^m \frac{n_i^2(\mathbb{E}[\bar{y}_{i\cdot}] - \mu)}{\lambda_i^2} = 0 \tag{B.7}$$

$$\begin{aligned}
-\mathbb{E}[l_{\sigma^2\sigma_a^2}] &= -\mathbb{E}\left[\frac{1}{2}\sum_{i=1}^m \frac{n_i}{\lambda_i^2} - \sum_{i=1}^m \frac{n_i^2(\bar{y}_i - \mu)^2}{\lambda_i^3}\right] \\
&= -\frac{1}{2}\sum_{i=1}^m \frac{n_i}{\lambda_i^2} + \sum_{i=1}^m \frac{n_i^2}{\lambda_i^3}\mathbb{E}[(\bar{y}_i - \mu)^2] \\
&= -\frac{1}{2}\sum_{i=1}^m \frac{n_i}{\lambda_i^2} + \sum_{i=1}^m \frac{n_i^2}{\lambda_i^3}\mathbb{E}[\bar{y}_i^2 - 2\bar{y}_i\mu + \mu^2] \\
&= -\frac{1}{2}\sum_{i=1}^m \frac{n_i}{\lambda_i^2} + \sum_{i=1}^m \frac{n_i^2}{\lambda_i^3}\left(\mu^2 + \frac{1}{n_i}\sigma^2 + \sigma_a^2 - 2\mu^2 + \mu^2\right) \\
&= -\frac{1}{2}\sum_{i=1}^m \frac{n_i}{\lambda_i^2} + \sum_{i=1}^m \frac{n_i^2}{\lambda_i^3}\left(\frac{\lambda_i}{n_i}\right) \\
&= -\frac{1}{2}\sum_{i=1}^m \frac{n_i}{\lambda_i^2} + \sum_{i=1}^m \frac{n_i}{\lambda_i^2} \\
&= \frac{1}{2}\sum_{i=1}^m \frac{n_i}{\lambda_i^2}
\end{aligned} \tag{B.8}$$

$$\begin{aligned}
-\mathbb{E}[l_{\sigma^2\sigma^2}] &= -\mathbb{E}\left[\frac{N-m}{2\sigma^4} - \frac{SCR}{\sigma^6} + \frac{1}{2}\sum_{i=1}^m \frac{1}{\lambda_i^2} - \sum_{i=1}^m \frac{n_i(\bar{y}_i - \mu)^2}{\lambda_i^3}\right] \\
&= -\frac{N-m}{2\sigma^4} + \frac{\mathbb{E}[SCR]}{\sigma^6} - \frac{1}{2}\sum_{i=1}^m \frac{1}{\lambda_i^2} + \sum_{i=1}^m \frac{n_i}{\lambda_i^3}\mathbb{E}[(\bar{y}_i - \mu)^2] \\
&= -\frac{N-m}{2\sigma^4} + \frac{(N-m)\sigma^2}{\sigma^6} - \frac{1}{2}\sum_{i=1}^m \frac{1}{\lambda_i^2} + \sum_{i=1}^m \frac{n_i}{\lambda_i^3}\frac{\lambda_i}{n_i} \\
&= \frac{N-m}{2\sigma^4} - \frac{1}{2}\sum_{i=1}^m \frac{1}{\lambda_i^2} + \sum_{i=1}^m \frac{1}{\lambda_i^2} \\
&= \frac{N-m}{2\sigma^4} + \frac{1}{2}\sum_{i=1}^m \frac{1}{\lambda_i^2}
\end{aligned} \tag{B.9}$$

$$\begin{aligned}
-\mathbb{E}[l_{\sigma_a^2 \sigma_a^2}] &= -\mathbb{E} \left[ \frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} - \sum_{i=1}^m \frac{n_i^3 (\bar{y}_i - \mu)^2}{\lambda_i^3} \right] \\
&= -\frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} + \sum_{i=1}^m \frac{n_i^3}{\lambda_i^3} \mathbb{E}[(\bar{y}_i - \mu)^2] \\
&= -\frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} + \sum_{i=1}^m \frac{n_i^3}{\lambda_i^3} \frac{\lambda_i}{n_i} \\
&= -\frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} + \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} \\
&= \frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2}
\end{aligned} \tag{B.10}$$

### B.3.3 Détails pour le test sur $\sigma_a^2$

Voici la proposition à démontrer.

**Proposition B.18.** *Un test pour tester l'hypothèse nulle  $H_0 : \sigma_a^2 = 0$  utilise la statistique  $\frac{CME}{CMR}$  qui suit une Fisher à  $m - 1$  et  $N - m$  degrés de liberté dans le cas où les données sont équilibrées ou non.*

*Démonstration.* Commençons par le cas où les données sont équilibrées. Montrons que le test associé à  $F = \frac{CME}{CMR}$  est le test du rapport de vraisemblance. La vraisemblance de ce modèle est donnée par

$$\begin{aligned}
l(\mu, \sigma^2, \sigma_a^2) &= -\frac{N}{2} \log(2\pi) - \frac{m}{2} \log(\lambda) - \frac{m(n-1)}{2} \log(\sigma^2) \\
&\quad - \frac{SCR}{2\sigma^2} - \frac{SCE}{2\lambda} - mn \frac{(\bar{y}_i - \mu)^2}{2\lambda}
\end{aligned}$$

Et donc la vraisemblance avec les paramètres égaux aux solutions ( $\hat{\mu} = \bar{y}_i$ ,  $\hat{\sigma}^2 = \frac{SCR}{m(n-1)}$  et  $\hat{\lambda} = \frac{SCE}{m}$ ) est donnée par

$$\begin{aligned}
l(\hat{\mu}, \hat{\sigma}^2, \hat{\sigma}_a^2) &= -\frac{N}{2} \log(2\pi) - \frac{m}{2} \log(\hat{\lambda}) - \frac{m(n-1)}{2} \log(\hat{\sigma}^2) \\
&\quad - \frac{SCR}{2\hat{\sigma}^2} - \frac{SCE}{2\hat{\lambda}} - mn \frac{(\bar{y}_{..} - \hat{\mu})^2}{2\hat{\lambda}} \\
&= -\frac{N}{2} \log(2\pi) - \frac{m}{2} \log\left(\frac{SCE}{m}\right) - \frac{m(n-1)}{2} \log\left(\frac{SCR}{m(n-1)}\right) \\
&\quad - \frac{SCR}{2\frac{SCR}{m(n-1)}} - \frac{SCE}{2\frac{SCE}{m}} - mn \frac{(\bar{y}_{..} - \bar{y}_{..})^2}{2\frac{SCE}{m}} \\
&= -\frac{N}{2} \log(2\pi) - \frac{m}{2} \log\left(\frac{SCE}{m}\right) - \frac{m(n-1)}{2} \log\left(\frac{SCR}{m(n-1)}\right) - \frac{mn}{2}
\end{aligned}$$

Ensuite, trouvons la vraisemblance sous  $H_0$ .  $\hat{\sigma}_0^2 = \frac{SCT}{mn}$  et  $\hat{\mu}_0 = \hat{\mu} = \bar{y}_{..}$ , donc  $\hat{\lambda}_0 = \hat{\sigma}_0^2$ . Il s'ensuit que

$$\begin{aligned}
l(\hat{\mu}_0, \hat{\sigma}_0^2, \sigma_a^2 = 0) &= -\frac{N}{2} \log(2\pi) - \frac{m}{2} \log(\hat{\lambda}_0) - \frac{m(n-1)}{2} \log(\hat{\sigma}_0^2) \\
&\quad - \frac{SCR}{2\hat{\sigma}_0^2} - \frac{SCE}{2\hat{\lambda}_0} - mn \frac{(\bar{y}_{..} - \hat{\mu}_0)^2}{2\hat{\lambda}_0} \\
&= -\frac{N}{2} \log(2\pi) - \frac{m}{2} \log\left(\frac{SCT}{mn}\right) - \frac{m(n-1)}{2} \log\left(\frac{SCT}{mn}\right) \\
&\quad - \frac{SCR}{2\frac{SCT}{mn}} - \frac{SCE}{2\frac{SCT}{mn}} - mn \frac{(\bar{y}_{..} - \bar{y}_{..})^2}{2\frac{SCT}{mn}} \\
&= -\frac{N}{2} \log(2\pi) - \frac{mn}{2} \log\left(\frac{SCT}{mn}\right) - \frac{mn}{2}
\end{aligned}$$

Posons  $q = \frac{SCE}{SCR} = \frac{(m-1)F}{N-m}$  avec  $F \sim \mathcal{F}_{N-m}^{m-1}$  (voir la section B.1 des annexes), alors  $-2 \log(\Lambda)$  est donné par

$$\begin{aligned}
-2 \log(\Lambda) &= -2l(\hat{\mu}_0, \hat{\sigma}_0^2, \sigma_a^2 = 0) + 2l(\hat{\mu}, \hat{\sigma}^2, \hat{\sigma}_a^2) \\
&= N \log(2\pi) + mn \log\left(\frac{SCT}{mn}\right) + mn \\
&\quad - N \log(2\pi) - m \log\left(\frac{SCE}{m}\right) - m(n-1) \log\left(\frac{SCR}{m(n-1)}\right) - mn \\
&= mn \log\left(\frac{SCT}{mn}\right) - mn \log\left(\frac{SCR}{m(n-1)}\right) + m \log\left(\frac{SCR}{m(n-1)}\right) - m \log\left(\frac{SCE}{m}\right) \\
&= mn \log\left(\frac{n-1}{n} \left(1 + \frac{SCE}{SCR}\right)\right) - m \log\left((n-1) \frac{SCE}{SCR}\right) \\
&= mn \log\left(\frac{n-1}{n} (1+q)\right) - m \log((n-1)q)
\end{aligned}$$

Trouvons maintenant la dérivée de  $-2 \log(\Lambda)$  par rapport à  $q$ .

$$\begin{aligned}
\frac{\partial(-2 \log(\Lambda))}{\partial q} &= mn \frac{1}{\frac{n-1}{n}(1+q)} \frac{n-1}{n} - m \frac{1}{(n-1)q} (n-1) \\
&= \frac{mn}{1+q} - \frac{m}{q} \\
&= \frac{m(nq - (1+q))}{q(1+q)} \\
&= \frac{m(q(n-1) - 1)}{(1+q)q}
\end{aligned}$$

Trouvons une expression pour le cas où  $\frac{\partial(-2 \log(\Lambda))}{\partial q} > 0$ .

$$\begin{aligned}
\frac{\partial(-2 \log(\Lambda))}{\partial q} > 0 &\Leftrightarrow q > \frac{1}{n-1} \\
&\Leftrightarrow SCE > \frac{SCR}{n-1} \\
&\Leftrightarrow SCE - \frac{SCR}{n-1} > 0 \\
&\Leftrightarrow \frac{SCE}{mn} - \frac{SCR}{mn(n-1)} > 0
\end{aligned}$$

$$\begin{aligned}
&\Leftrightarrow \frac{\frac{(1-1/m)SCE}{m-1} - \frac{SCR}{m(n-1)}}{n} > 0 \\
&\Leftrightarrow \frac{(1-1/m)CME - CMR}{n} > 0 \\
&\Leftrightarrow \hat{\sigma}_a^2 > 0 \\
&\Leftrightarrow \hat{\sigma}_a^2 > 0
\end{aligned}$$

Et donc, si  $\hat{\sigma}_a^2 > 0$  ( $q > \frac{1}{n-1}$ ),  $\frac{\partial -2\log(\Lambda)}{\partial q} > 0$ . Les deux statistiques ( $-2\log(\Lambda)$  et  $F = aq$ ) varient donc dans le même sens pour tout le domaine de  $\hat{\sigma}_a^2$  sauf en un point (0), mais cela est suffisant pour l'équivalence des statistiques. Ce qui revient à dire que le test du rapport de vraisemblance est équivalent à regarder si la statistique ( $F = \frac{CME}{CMR}$ ) dépasse  $\mathcal{F}_{N-m,1-\alpha}^{m-1}$ .

Pour le cas où les données sont non équilibrées, cette même statistique  $F = \frac{CME}{CMR}$  suit une loi de Fisher sous  $H_0 : \sigma_a^2 = 0$  ( $F \sim \mathcal{F}_{N-m}^{m-1}$ , voir la section B.1). Il est possible d'utiliser cette statistique de Fisher pour tester l'hypothèse nulle  $H_0 : \sigma_a^2 = 0$  dans le cas où les données sont équilibrées ou non équilibrées (voir le livre de Searle [SCM92] autour de la page 77).  $\square$

## B.4 Autres preuves pour la régression linéaire simple classique

Rappelons le modèle en premier lieu :

$$\begin{aligned}
\mathbb{E}[y_{ij}] &= \mu + \beta x_i \\
y_{ij} &\sim \mathcal{N}(\mu + \beta x_i, \sigma^2) \text{ indépendantes}
\end{aligned} \tag{B.11}$$

Où  $i = 1, \dots, N$ . Voici quelques preuves supplémentaires qui ne sont pas présentées dans la section sur la régression linéaire simple classique.

### B.4.1 Loi du couple $\hat{\mu}, \hat{\beta}$

Cherchons la loi des estimateurs du maximum de vraisemblance dans le contexte précédent, c'est-à-dire la régression linéaire simple habituelle. Montrons premièrement la proposition suivante.

**Proposition B.19.**

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \frac{\sigma^2}{S_{xx}} \begin{pmatrix} \sum_{i=1}^N \frac{x_i^2}{N} & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix} \right)$$

*Démonstration.* Premièrement, les estimateurs sont les suivants :

$$\begin{aligned} \hat{\beta} &= \frac{\sum_{i=1}^N x_i y_i - \left( \sum_{i=1}^N x_i \right) \left( \sum_{i=1}^N y_i \right) / N}{\sum_{i=1}^N x_i^2 - \left( \sum_{i=1}^N x_i \right)^2 / N} \\ &= \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2} \\ &= \frac{1}{S_{xx}} [(x_1 - \bar{x}), \dots, (x_N - \bar{x})] \begin{bmatrix} y_1 - \bar{y} \\ \vdots \\ y_N - \bar{y} \end{bmatrix} \end{aligned}$$

et

$$\hat{\mu} = \bar{y} - \hat{\beta} \bar{x}$$



Où  $S_{xx} = \sum_{i=1}^N (x_i - \bar{x})^2$ . Posons  $\mathbf{y} = [y_1, \dots, y_N]^t$ ,  $\mathbf{y}^* = [(y_1 - \bar{y}), \dots, (y_N - \bar{y})]^t = \mathbf{A}_1 \mathbf{y}$  et  $\mathbf{x}^* = [(x_1 - \bar{x}), \dots, (x_N - \bar{x})]^t = \mathbf{A}_1 \mathbf{x}$  avec  $\mathbf{A}_1 = \mathbf{I}_N - \frac{1}{N} \mathbf{J}_N$ . Puisque les  $y_i$  sont indépendantes et suivent une loi normale de paramètres  $\mu + \beta x_i$  et  $\sigma^2$ , alors

$$\mathbf{y} \sim \mathcal{N} \left( \begin{bmatrix} \mu + \beta x_1 \\ \vdots \\ \mu + \beta x_N \end{bmatrix}, \mathbf{I}_N \sigma^2 \right)$$

Et donc, puisque  $\mathbf{y}^* = \mathbf{A}_1 \mathbf{y}$ , il en découle que

$$\begin{aligned} \mathbf{y}^* &\sim \mathcal{N} \left( \mathbf{A}_1 \begin{bmatrix} \mu + \beta x_1 \\ \vdots \\ \mu + \beta x_N \end{bmatrix}, \mathbf{A}_1 \sigma^2 \mathbf{A}_1^t \right) \\ &\sim \mathcal{N} \left( \begin{bmatrix} \beta(x_1 - \bar{x}) \\ \vdots \\ \beta(x_N - \bar{x}) \end{bmatrix}, \sigma^2 \mathbf{A}_1 \right) \\ &\sim \mathcal{N} (\beta \mathbf{x}^*, \sigma^2 \mathbf{A}_1) \end{aligned}$$

Car  $\mathbf{A}_1 = \mathbf{A}_1^t$  et que  $\mathbf{A}_1 \mathbf{A}_1 = (\mathbf{I}_N - \frac{1}{N} \mathbf{J}_N)(\mathbf{I}_N - \frac{1}{N} \mathbf{J}_N) = \mathbf{I}_N - \frac{1}{N} \mathbf{J}_N - \frac{1}{N} \mathbf{J}_N + \frac{1}{N^2} \mathbf{J}_N^2 = \mathbf{I}_N - \frac{1}{N} \mathbf{J}_N = \mathbf{A}_1$ .

Ensuite, posons  $\mathbf{A}_2 = \frac{1}{S_{xx}} [(x_1 - \bar{x}), \dots, (x_N - \bar{x})] = \frac{\mathbf{x}^{*t}}{S_{xx}}$  alors  $\hat{\beta} = \mathbf{A}_2 \mathbf{y}^*$ . Donc,

$$\begin{aligned} \hat{\beta} &\sim \mathcal{N} (\mathbf{A}_2 \beta \mathbf{x}^*, \mathbf{A}_2 \sigma^2 \mathbf{A}_1 \mathbf{A}_2^t) \\ &\sim \mathcal{N} \left( \frac{1}{S_{xx}} \beta \sum_{i=1}^N (x_i - \bar{x})^2, \frac{\sigma^2}{S_{xx}^2} \mathbf{x}^{*t} \mathbf{A}_1 \mathbf{x}^* \right) \\ &\sim \mathcal{N} \left( \beta, \frac{\sigma^2}{S_{xx}} \right) \end{aligned}$$

Car,

$$\begin{aligned}
\mathbf{x}^{*t} \mathbf{A}_1 \mathbf{x}^* &= [(x_1 - \bar{x}), \dots, (x_N - \bar{x})] \begin{pmatrix} 1 - \frac{1}{N} & \frac{-1}{N} & \dots & \frac{-1}{N} \\ \frac{-1}{N} & 1 - \frac{1}{N} & \dots & \frac{-1}{N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{-1}{N} & \frac{-1}{N} & \dots & 1 - \frac{1}{N} \end{pmatrix} \begin{bmatrix} x_1 - \bar{x} \\ \vdots \\ x_N - \bar{x} \end{bmatrix} \\
&= \left[ (x_1 - \bar{x}) - \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}), \dots, (x_N - \bar{x}) - \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}) \right] \begin{bmatrix} x_1 - \bar{x} \\ \vdots \\ x_N - \bar{x} \end{bmatrix} \\
&= [(x_1 - \bar{x}), \dots, (x_N - \bar{x})] \begin{bmatrix} x_1 - \bar{x} \\ \vdots \\ x_N - \bar{x} \end{bmatrix} \\
&= \sum_{i=1}^N (x_i - \bar{x})^2 \\
&= S_{xx}
\end{aligned}$$

Ensuite, cherchons la loi du couple  $(\hat{\mu}, \hat{\beta})$ . Sachant que  $\hat{\mu} = \bar{y} - \hat{\beta}\bar{x}$ , il en découle que

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} = \begin{pmatrix} -\bar{x} & 1 \\ 1 & 0 \end{pmatrix} \begin{bmatrix} \hat{\beta} \\ \bar{y} \end{bmatrix}$$

Donc, avant de trouver la loi du couple  $(\hat{\mu}, \hat{\beta})$ , il faut trouver la loi du couple  $(\hat{\beta}, \bar{y})$ .  $\hat{\beta}$  et  $\bar{y}$  sont indépendants puisque les deux suivent une loi normale et que  $\hat{\beta}$  est une fonction de  $(y_i - \bar{y})$ . La loi de  $\hat{\beta}$  a été trouvée plus haut, il ne reste donc que la loi de  $\bar{y}$ . C'est assez simple puisque les  $y_i$  sont indépendantes. En effet,

$$\bar{y} \sim \mathcal{N} \left( \mu + \beta\bar{x}, \frac{\sigma^2}{N} \right)$$

Donc, pour le couple  $(\hat{\beta}, \bar{y})$ , leur loi est donnée par

$$\begin{bmatrix} \hat{\beta} \\ \bar{y} \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \beta \\ \mu + \beta \bar{x} \end{bmatrix}, \begin{pmatrix} \frac{\sigma^2}{S_{xx}} & 0 \\ 0 & \frac{\sigma^2}{N} \end{pmatrix} \right)$$

Donc, posons

$$\mathbf{A}_3 = \begin{pmatrix} -\bar{x} & 1 \\ 1 & 0 \end{pmatrix}$$

alors

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} = \mathbf{A}_3 \begin{bmatrix} \hat{\beta} \\ \bar{y} \end{bmatrix}$$

Et donc,

$$\begin{aligned} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} &\sim \mathcal{N} \left( \mathbf{A}_3 \begin{bmatrix} \beta \\ \mu + \beta \bar{x} \end{bmatrix}, \mathbf{A}_3 \begin{pmatrix} \frac{\sigma^2}{S_{xx}} & 0 \\ 0 & \frac{\sigma^2}{N} \end{pmatrix} \mathbf{A}_3^t \right) \\ &\sim \mathcal{N} \left( \begin{pmatrix} -\bar{x} & 1 \\ 1 & 0 \end{pmatrix} \begin{bmatrix} \beta \\ \mu + \beta \bar{x} \end{bmatrix}, \begin{pmatrix} -\bar{x} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{\sigma^2}{S_{xx}} & 0 \\ 0 & \frac{\sigma^2}{N} \end{pmatrix} \begin{pmatrix} -\bar{x} & 1 \\ 1 & 0 \end{pmatrix} \right) \\ &\sim \mathcal{N} \left( \begin{bmatrix} -\bar{x}\beta + \mu + \beta \bar{x} \\ \beta \end{bmatrix}, \begin{pmatrix} \frac{-\bar{x}\sigma^2}{S_{xx}} & \frac{\sigma^2}{N} \\ \frac{\sigma^2}{S_{xx}} & 0 \end{pmatrix} \begin{pmatrix} -\bar{x} & 1 \\ 1 & 0 \end{pmatrix} \right) \\ &\sim \mathcal{N} \left( \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \begin{pmatrix} \frac{\bar{x}^2\sigma^2}{S_{xx}} + \frac{\sigma^2}{N} & \frac{-\bar{x}\sigma^2}{S_{xx}} \\ \frac{-\bar{x}\sigma^2}{S_{xx}} & \frac{\sigma^2}{S_{xx}} \end{pmatrix} \right) \\ &\sim \mathcal{N} \left( \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \frac{\sigma^2}{S_{xx}} \begin{pmatrix} \bar{x}^2 + \frac{S_{xx}}{N} & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix} \right) \\ &\sim \mathcal{N} \left( \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \frac{\sigma^2}{S_{xx}} \begin{pmatrix} \sum_{i=1}^N \frac{x_i^2}{N} & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix} \right) \end{aligned}$$

La dernière expression vient du fait que

$$\begin{aligned}
\bar{x}^2 + \frac{S_{xx}}{N} &= \bar{x}^2 + \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 \\
&= \bar{x}^2 + \frac{1}{N} \sum_{i=1}^N (x_i^2 - 2x_i\bar{x} + \bar{x}^2) \\
&= \bar{x}^2 + \frac{1}{N} \left( \sum_{i=1}^N x_i^2 \right) - 2\bar{x}^2 + \bar{x}^2 \\
&= \frac{1}{N} \sum_{i=1}^N x_i^2 \quad \square
\end{aligned}$$

### B.4.2 Loi de $\sigma^2$

Utilisons le théorème de Cochran qui est énoncé ci-dessous (version simplifiée de [Bar06]).

**Théorème B.5** (Cochran). *Soit  $X = [X_1, \dots, X_n]^t$  un vecteur de variables aléatoires de loi normale de paramètres 0 et 1. Pour  $F$  un sous-espace vectoriel de  $\mathbb{R}^n$  de dimension  $p$ , notons  $P_F$  (resp.  $P_{F^\perp}$ ) la matrice de projection orthogonale sur  $F$  (resp.  $F^\perp$ ).*

*Alors, les vecteurs aléatoires  $P_F X$  et  $P_{F^\perp} X$  sont gaussiens et indépendants de lois*

$$P_F X \sim \mathcal{N}(0, P_F) \quad \text{et} \quad P_{F^\perp} X \sim \mathcal{N}(0, P_{F^\perp})$$

*De plus, les variables aléatoires  $\|P_F X\|^2$  et  $\|P_{F^\perp} X\|^2$  sont indépendantes de lois*

$$\|P_F X\|^2 \sim \chi_{(p)}^2 \quad \text{et} \quad \|P_{F^\perp} X\|^2 \sim \chi_{(n-p)}^2$$

*Démonstration.* Accepté sans preuve, voir l'article de Bardet [Bar06]. □

Le théorème précédent nous permet de montrer la prochaine proposition.

**Proposition B.20.**  $\frac{N\hat{\sigma}^2}{\sigma^2} \sim \chi_{(N-2)}^2$

*Démonstration.* Dans le contexte de la régression linéaire simple classique,  $\mathbf{e} = [e_1, \dots, e_N]^t \in \mathbb{R}^N$  avec  $e_i = \frac{1}{\sigma}(y_i - \mu - \beta x_i)$  et les  $e_i$  sont i.i.d de loi normale de paramètres 0 et 1, car ce sont les observations qui sont centrées et réduites. Posons  $F$  le sous-espace vectoriel engendré par  $\mathbf{1}_N$  et  $(\mathbf{x} - \bar{x}\mathbf{1}_N)$  où  $\mathbf{x} = [x_1, \dots, x_N]^t$ . Ces deux vecteurs forment une base orthogonale de  $F$  puisqu'ils engendrent  $F$  et qu'ils sont perpendiculaires. En effet,

$$\langle \mathbf{1}_N, \mathbf{x} - \bar{x}\mathbf{1}_N \rangle = \sum_{i=1}^N (x_i - \bar{x}) = 0$$

Ensuite, notons  $\mathbf{e}^*$  la projection de  $\mathbf{e}$  sur le sous-espace vectoriel  $F$  ( $\mathbf{e}^* = P_F(\mathbf{e})$ ).

$$\begin{aligned} \mathbf{e}^* &= \frac{\langle \mathbf{e}, \mathbf{1}_N \rangle}{\|\mathbf{1}_N\|^2} * \mathbf{1}_N + \frac{\langle \mathbf{e}, \mathbf{x} - \bar{x}\mathbf{1}_N \rangle}{\|\mathbf{x} - \bar{x}\mathbf{1}_N\|^2} * (\mathbf{x} - \bar{x}\mathbf{1}_N) \\ &= \frac{\sum_{i=1}^N e_i}{N} * \mathbf{1}_N + \frac{\sum_{i=1}^N (x_i - \bar{x})e_i}{S_{xx}} * (\mathbf{x} - \bar{x}\mathbf{1}_N) \\ &= \bar{e} * \mathbf{1}_N + b(\mathbf{e}) * (\mathbf{x} - \bar{x}\mathbf{1}_N) \end{aligned}$$

Où  $b(\mathbf{e}) = \frac{\sum_{i=1}^N (x_i - \bar{x})e_i}{S_{xx}}$ . Cette notation sera utilisée pour simplifier les expressions.

Observons que  $\mathbf{e} - \mathbf{e}^*$  est bien la projection orthogonale de  $\mathbf{e}$  sur  $F^\perp$  ( $\mathbf{e} - \mathbf{e}^* = P_{F^\perp}(\mathbf{e})$ ).

En effet,

$$\begin{aligned} \langle \mathbf{e} - \mathbf{e}^*, \mathbf{1}_N \rangle &= \langle \mathbf{e} - \bar{e}\mathbf{1}_N - b(\mathbf{e})(\mathbf{x} - \bar{x}\mathbf{1}_N), \mathbf{1}_N \rangle \\ &= \langle \mathbf{e} - \bar{e}\mathbf{1}_N, \mathbf{1}_N \rangle - b(\mathbf{e})\langle \mathbf{x} - \bar{x}\mathbf{1}_N, \mathbf{1}_N \rangle \\ &= \langle \mathbf{e} - \bar{e}\mathbf{1}_N, \mathbf{1}_N \rangle \\ &= 0 \end{aligned}$$

$$\begin{aligned}
\langle \mathbf{e} - \mathbf{e}^*, \mathbf{x} - \bar{x}\mathbf{1}_N \rangle &= \langle \mathbf{e} - \bar{e}\mathbf{1}_N - b(\mathbf{e})(\mathbf{x} - \bar{x}\mathbf{1}_N), \mathbf{x} - \bar{x}\mathbf{1}_N \rangle \\
&= \langle \mathbf{e}, \mathbf{x} - \bar{x}\mathbf{1}_N \rangle - \bar{e}\langle \mathbf{1}_N, \mathbf{x} - \bar{x}\mathbf{1}_N \rangle - b(\mathbf{e})\langle \mathbf{x} - \bar{x}\mathbf{1}_N, \mathbf{x} - \bar{x}\mathbf{1}_N \rangle \\
&= \langle \mathbf{e}, \mathbf{x} - \bar{x}\mathbf{1}_N \rangle - b(\mathbf{e})\|\mathbf{x} - \bar{x}\mathbf{1}_N\|^2 \\
&= \langle \mathbf{e}, \mathbf{x} - \bar{x}\mathbf{1}_N \rangle - \frac{\sum_{i=1}^N (x_i - \bar{x})e_i}{\sum_{i=1}^N (x_i - \bar{x})^2} \sum_{i=1}^N (x_i - \bar{x})^2 \\
&= \sum_{i=1}^N (x_i - \bar{x})e_i - \sum_{i=1}^N (x_i - \bar{x})e_i \\
&= 0
\end{aligned}$$

Donc,  $\mathbf{e} - \mathbf{e}^*$  est bien dans  $F^\perp$  et  $\mathbf{e} - \mathbf{e}^* = P_{F^\perp}(\mathbf{e})$ .

Maintenant, trouvons une version plus jolie pour  $\mathbf{e}^*$ . Posons premièrement  $\epsilon_i = y_i - \mu - \beta x_i = \sigma e_i$ .

$$\begin{aligned}
\hat{\beta} &= \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^N (x_i - \bar{x})^2} = \frac{\sum_{i=1}^N (x_i - \bar{x})(\mu + \beta x_i + \epsilon_i - (\mu + \beta \bar{x} + \bar{\epsilon}))}{\sum_{i=1}^N (x_i - \bar{x})^2} \\
&= \frac{\sum_{i=1}^N (x_i - \bar{x})(\beta(x_i - \bar{x}) + (\epsilon_i - \bar{\epsilon}))}{\sum_{i=1}^N (x_i - \bar{x})^2} \\
&= \frac{\beta \sum_{i=1}^N (x_i - \bar{x})^2 + \sum_{i=1}^N (x_i - \bar{x})(\epsilon_i - \bar{\epsilon})}{\sum_{i=1}^N (x_i - \bar{x})^2}
\end{aligned}$$

$$\begin{aligned}
&= \beta + \frac{\sum_{i=1}^N (x_i - \bar{x})\epsilon_i}{\sum_{i=1}^N (x_i - \bar{x})^2} \\
&= \beta + \sigma \frac{\sum_{i=1}^N (x_i - \bar{x})e_i}{\sum_{i=1}^N (x_i - \bar{x})^2} \\
&= \beta + \sigma b(\mathbf{e})
\end{aligned}$$

$$\begin{aligned}
\sigma e_i^* &= \sigma \bar{e} + \sigma b(\mathbf{e})(x_i - \bar{x}) \\
&= \frac{\sigma}{\sigma N} \sum_{j=1}^N (y_j - \mu - \beta x_j) + \sigma b(\mathbf{e})(x_i - \bar{x}) \\
&= \bar{y} - \mu - \beta \bar{x} + (\hat{\beta} - \beta)(x_i - \bar{x}) \\
&= \bar{y} - \mu - \beta \bar{x} + \hat{\beta}(x_i - \bar{x}) - \beta x_i + \beta \bar{x} \\
&= \bar{y} - \mu + \hat{\beta} x_i - \hat{\beta} \bar{x} - \beta x_i
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \mu + \beta x_i + \sigma e_i^* &= \mu + \beta x_i + \bar{y} - \mu + \hat{\beta} x_i - \hat{\beta} \bar{x} - \beta x_i \\
&= \bar{y} + \hat{\beta}(x_i - \bar{x}) \\
&= \hat{\mu} + \hat{\beta} \bar{x} + \hat{\beta} x_i - \hat{\beta} \bar{x} \\
&= \hat{\mu} + \hat{\beta} x_i \\
\Rightarrow e^* &= \frac{\hat{\mu} \mathbf{1}_N + \hat{\beta} \mathbf{x} - \mu \mathbf{1}_N - \beta \mathbf{x}}{\sigma}
\end{aligned}$$

Le théorème de Cochran (B.5) dit que  $\|P_{F^\perp} \mathbf{e}\|^2 \sim \chi_{(N-2)}^2$  donc,

$$\begin{aligned}
\|P_{F^\perp} \mathbf{e}\|^2 = \|\mathbf{e} - \mathbf{e}^*\|^2 &= \left\| \frac{\mathbf{y} - \mu \mathbf{1}_N - \beta \mathbf{x}}{\sigma} - \frac{\hat{\mu} \mathbf{1}_N + \hat{\beta} \mathbf{x} - \mu \mathbf{1}_N - \beta \mathbf{x}}{\sigma} \right\|^2 \\
&= \left\| \frac{\mathbf{y} - \hat{\mu} \mathbf{1}_N - \hat{\beta} \mathbf{x}}{\sigma} \right\|^2 \\
&= \frac{1}{\sigma^2} \sum_{i=1}^N (y_i - \hat{\mu} - \hat{\beta} x_i)^2 \\
&= \frac{N \hat{\sigma}^2}{\sigma^2} \\
\Rightarrow \frac{N \hat{\sigma}^2}{\sigma^2} &\sim \chi_{(N-2)}^2
\end{aligned}$$

Où  $\mathbf{y} = [y_1, \dots, y_N]^t$ . □

**Proposition B.21.**  $\hat{\sigma}^2$  est indépendant de  $\hat{\mu}$  et de  $\hat{\beta}$ .

*Démonstration.* Pour montrer l'indépendance entre les différents paramètres, remarquons que

- $\mathbf{e} - \mathbf{e}^*$  est orthogonal à  $\mathbf{e}^*$
- $\Rightarrow \mathbf{e} - \mathbf{e}^*$  est indépendant de  $\mathbf{e}^*$
- $\Rightarrow \mathbf{e} - \mathbf{e}^*$  est indépendant des composantes de  $\mathbf{e}^*$  dans la base  $(\mathbf{1}_N, \mathbf{x} - \bar{x} \mathbf{1}_N)$
- $\Rightarrow \mathbf{e} - \mathbf{e}^*$  est indépendant de  $\bar{e}$  et de  $b(\mathbf{e})$
- $\Rightarrow \hat{\sigma}$ , qui est une fonction de  $\mathbf{e} - \mathbf{e}^*$ , est indépendant de  $\bar{e}$  et de  $b(\mathbf{e})$
- $\Rightarrow \hat{\sigma}$  est indépendant de  $\hat{\beta}$  et  $\bar{y}$ , qui sont des fonctions de  $\bar{e}$  et de  $b(\mathbf{e})$
- $\Rightarrow \hat{\sigma}$  est indépendant de  $\hat{\beta}$  et  $\hat{\mu} = \bar{y} - \hat{\beta} \bar{x}$  □



## B.5 Autres preuves pour l'ANCOVA avec un facteur fixe

### B.5.1 Lois de $\hat{\mu}$ , $\hat{\beta}$ et $\hat{\alpha}_i$

Rappelons la loi de  $\mathbf{y}$  dans le contexte de l'ANCOVA.

$$\mathbf{y} \sim \mathcal{N}(\mu \mathbf{1}_N + \{c \mathbf{1}_{n_i} \alpha_i\} + \beta \mathbf{x}, \mathbf{I}_N \sigma^2) \quad (\text{B.12})$$

Où  $\mathbf{x}$  est le vecteur  $[x_{11}, \dots, x_{mn_m}]^t$ .

**Proposition B.22.**

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \\ \hat{\alpha}_i \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \mu \\ \beta \\ \alpha_i \end{bmatrix}, \begin{pmatrix} \frac{\bar{x}_{..}^2 \sigma^2}{S_{xx}} + \frac{\sigma^2}{N} & -\frac{\bar{x}_{..} \sigma^2}{S_{xx}} & \frac{\bar{x}_{..} (\bar{x}_i - \bar{x}_{..}) \sigma^2}{S_{xx}} \\ -\frac{\bar{x}_{..} \sigma^2}{S_{xx}} & \frac{\sigma^2}{S_{xx}} & -\frac{(\bar{x}_i - \bar{x}_{..}) \sigma^2}{S_{xx}} \\ \frac{\bar{x}_{..} (\bar{x}_i - \bar{x}_{..}) \sigma^2}{S_{xx}} & -\frac{(\bar{x}_i - \bar{x}_{..}) \sigma^2}{S_{xx}} & \frac{(\bar{x}_i - \bar{x}_{..})^2 \sigma^2}{S_{xx}} + \frac{\sigma^2 (N - n_i)}{N n_i} \end{pmatrix} \right)$$

et

$$\begin{bmatrix} \hat{\alpha}_i \\ \hat{\alpha}_j \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \alpha_i \\ \alpha_j \end{bmatrix}, \begin{pmatrix} \frac{(\bar{x}_i - \bar{x}_{..})^2 \sigma^2}{S_{xx}} + \frac{\sigma^2 (N - n_i)}{N n_i} & \frac{(\bar{x}_i - \bar{x}_{..}) (\bar{x}_j - \bar{x}_{..}) \sigma^2}{S_{xx}} - \frac{\sigma^2}{N} \\ \frac{(\bar{x}_i - \bar{x}_{..}) (\bar{x}_j - \bar{x}_{..}) \sigma^2}{S_{xx}} - \frac{\sigma^2}{N} & \frac{(\bar{x}_j - \bar{x}_{..})^2 \sigma^2}{S_{xx}} + \frac{\sigma^2 (N - n_j)}{N n_j} \end{pmatrix} \right)$$

*Démonstration.* Voici premièrement les estimateurs.

$$\hat{\beta} = \frac{S_{xy}}{S_{xx}}$$

$$\hat{\mu} = \bar{y}_{..} - \hat{\beta} \bar{x}_{..}$$

$$\hat{\alpha}_i = (\bar{y}_i - \bar{y}_{..}) - \hat{\beta} (\bar{x}_i - \bar{x}_{..})$$

Voici la loi de  $\mathbf{y}^* = [y_{11} - \bar{y}_{1.}, \dots, y_{mn_m} - \bar{y}_{m.}]^t = \mathbf{A}_1 \mathbf{y}$ , avec  $\mathbf{A}_1 = \left\{ \mathbf{I}_{n_i} - \frac{1}{n_i} \mathbf{J}_{n_i} \right\}$ .

$$\begin{aligned} \mathbf{y}^* &\sim \mathcal{N}(\mathbf{A}_1(\mu \mathbf{1}_N + \{c \mathbf{1}_{n_i} \alpha_i\} + \beta \mathbf{x}), \mathbf{A}_1 \mathbf{I}_N \sigma^2 \mathbf{A}_1^t) \\ &\sim \mathcal{N}(\beta \mathbf{x}^*, \sigma^2 \mathbf{A}_1) \end{aligned}$$

Où  $\mathbf{x}^* = [x_{11} - \bar{x}_{1.}, \dots, x_{mn_m} - \bar{x}_{m.}]^t$ . Le résultat vient du fait que  $\mathbf{A}_1 \mathbf{A}_1^t = \mathbf{A}_1$ .

Posons maintenant  $\mathbf{A}_2 = \frac{\mathbf{x}^{*t}}{S_{xx}}$ , alors  $\hat{\beta} = \mathbf{A}_2 \mathbf{y}^*$  et

$$\begin{aligned} \hat{\beta} &\sim \mathcal{N}(\mathbf{A}_2 \beta \mathbf{x}^*, \sigma^2 \mathbf{A}_2 \mathbf{A}_1 \mathbf{A}_2^t) \\ &\sim \mathcal{N}\left(\beta \frac{S_{xx}}{S_{xx}}, \frac{\sigma^2}{S_{xx}^2} \mathbf{x}^{*t} \mathbf{A}_1 \mathbf{x}^*\right) \\ &\sim \mathcal{N}\left(\beta, \frac{\sigma^2}{S_{xx}^2} \mathbf{x}^{*t} \mathbf{x}^*\right) \\ &\sim \mathcal{N}\left(\beta, \frac{\sigma^2}{S_{xx}}\right) \end{aligned}$$

Fixons  $i$  et trouvons la loi de  $\bar{y}_{i.}$  et de  $\bar{y}_{i.} - \bar{y}_{..}$ . Il est possible de montrer que

$$\bar{y}_{i.} \sim \mathcal{N}\left(\mu + \beta \bar{x}_{i.}, \frac{\sigma^2}{N}\right)$$

Et

$$\bar{y}_{i.} - \bar{y}_{..} \sim \mathcal{N}\left(\beta(\bar{x}_{i.} - \bar{x}_{..}) + \alpha_i, \frac{\sigma^2(N - n_i)}{N n_i}\right)$$

De plus, remarquons que  $\hat{\beta}$ ,  $\bar{y}_{i.}$  et  $\bar{y}_{i.} - \bar{y}_{..}$  sont tous indépendants. Donc,

$$\begin{bmatrix} \hat{\beta} \\ \bar{y}_{..} \\ \bar{y}_{i.} - \bar{y}_{..} \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \beta \\ \mu + \beta \bar{x}_{..} \\ \beta(\bar{x}_{i.} - \bar{x}_{..}) + \alpha_i \end{bmatrix}, \begin{pmatrix} \frac{\sigma^2}{S_{xx}} & 0 & 0 \\ 0 & \frac{\sigma^2}{N} & 0 \\ 0 & 0 & \frac{\sigma^2(N-n_i)}{Nn_i} \end{pmatrix} \right)$$

Et

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \\ \hat{\alpha}_i \end{bmatrix} = \mathbf{A}_3 \begin{bmatrix} \hat{\beta} \\ \bar{y}_{..} \\ \bar{y}_{i.} - \bar{y}_{..} \end{bmatrix}$$

Où

$$\mathbf{A}_3 = \begin{pmatrix} -\bar{x}_{..} & 1 & 0 \\ 1 & 0 & 0 \\ -(\bar{x}_{i.} - \bar{x}_{..}) & 0 & 1 \end{pmatrix}$$

Donc,

$$\begin{aligned} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \\ \hat{\alpha}_i \end{bmatrix} &\sim \mathcal{N} \left( \mathbf{A}_3 \begin{bmatrix} \beta \\ \mu + \beta \bar{x}_{..} \\ \beta(\bar{x}_{i.} - \bar{x}_{..}) + \alpha_i \end{bmatrix}, \mathbf{A}_3 \begin{pmatrix} \frac{\sigma^2}{S_{xx}} & 0 & 0 \\ 0 & \frac{\sigma^2}{N} & 0 \\ 0 & 0 & \frac{\sigma^2(N-n_i)}{Nn_i} \end{pmatrix} \mathbf{A}_3^t \right) \\ &\sim \mathcal{N} \left( \begin{bmatrix} \mu \\ \beta \\ \alpha_i \end{bmatrix}, \begin{pmatrix} -\frac{\bar{x}_{..}\sigma^2}{S_{xx}} & \frac{\sigma^2}{N} & 0 \\ \frac{\sigma^2}{S_{xx}} & 0 & 0 \\ -\frac{(\bar{x}_{i.}-\bar{x}_{..})\sigma^2}{S_{xx}} & 0 & \frac{\sigma^2(N-n_i)}{Nn_i} \end{pmatrix} \begin{pmatrix} -\bar{x}_{..} & 1 & -(\bar{x}_{i.} - \bar{x}_{..}) \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right) \\ &\sim \mathcal{N} \left( \begin{bmatrix} \mu \\ \beta \\ \alpha_i \end{bmatrix}, \begin{pmatrix} \frac{\bar{x}_{..}^2\sigma^2}{S_{xx}} + \frac{\sigma^2}{N} & -\frac{\bar{x}_{..}\sigma^2}{S_{xx}} & \frac{\bar{x}_{..}(\bar{x}_{i.}-\bar{x}_{..})\sigma^2}{S_{xx}} \\ -\frac{\bar{x}_{..}\sigma^2}{S_{xx}} & \frac{\sigma^2}{S_{xx}} & -\frac{(\bar{x}_{i.}-\bar{x}_{..})\sigma^2}{S_{xx}} \\ \frac{\bar{x}_{..}(\bar{x}_{i.}-\bar{x}_{..})\sigma^2}{S_{xx}} & -\frac{(\bar{x}_{i.}-\bar{x}_{..})\sigma^2}{S_{xx}} & \frac{(\bar{x}_{i.}-\bar{x}_{..})^2\sigma^2}{S_{xx}} + \frac{\sigma^2(N-n_i)}{Nn_i} \end{pmatrix} \right) \end{aligned}$$

Ils reste à montrer la loi du vecteur  $[\hat{\alpha}_i, \hat{\alpha}_j]^t$ . Trouvons uniquement leur covariances puisque leur espérance a déjà été calculée (voir l'équation précédente).

$$\begin{aligned}
\text{cov}(\hat{\alpha}_i, \hat{\alpha}_j) &= \text{cov}((\bar{y}_i - \bar{y}_{..}) - \hat{\beta}(\bar{x}_i - \bar{x}_{..}), (\bar{y}_j - \bar{y}_{..}) - \hat{\beta}(\bar{x}_j - \bar{x}_{..})) \\
&= \text{cov}(\bar{y}_i - \bar{y}_{..}, \bar{y}_j - \bar{y}_{..}) - (\bar{x}_i - \bar{x}_{..})\text{cov}(\bar{y}_i - \bar{y}_{..}, \hat{\beta}) \\
&\quad - (\bar{x}_j - \bar{x}_{..})\text{cov}(\bar{y}_j - \bar{y}_{..}, \hat{\beta}) + (\bar{x}_i - \bar{x}_{..})(\bar{x}_j - \bar{x}_{..})\text{Var}(\hat{\beta}) \\
&= -\text{cov}(\bar{y}_i, \bar{y}_{..}) - \text{cov}(\bar{y}_j, \bar{y}_{..}) + \text{Var}(\bar{y}_{..}) + (\bar{x}_i - \bar{x}_{..})(\bar{x}_j - \bar{x}_{..})\frac{\sigma^2}{S_{xx}} \\
&= -\frac{n_i}{N}\text{Var}(\bar{y}_i) - \frac{n_j}{N}\text{Var}(\bar{y}_j) + \frac{\sigma^2}{N} + (\bar{x}_i - \bar{x}_{..})(\bar{x}_j - \bar{x}_{..})\frac{\sigma^2}{S_{xx}} \\
&= -\frac{n_i}{N}\frac{\sigma^2}{n_i} - \frac{n_j}{N}\frac{\sigma^2}{n_j} + \frac{\sigma^2}{N} + (\bar{x}_i - \bar{x}_{..})(\bar{x}_j - \bar{x}_{..})\frac{\sigma^2}{S_{xx}} \\
&= (\bar{x}_i - \bar{x}_{..})(\bar{x}_j - \bar{x}_{..})\frac{\sigma^2}{S_{xx}} - \frac{\sigma^2}{N}
\end{aligned}$$

□

### B.5.2 Loi de $\frac{SCRR}{\sigma^2}$ , $\frac{SCC}{\sigma^2}$ et de $\frac{SCE}{\sigma^2}$

#### a) Loi de $\frac{SCRR}{\sigma^2}$

Montrons tout d'abord que

**Proposition B.23.**  $\frac{SCRR}{\sigma^2} = \mathbf{y}^t \mathbf{A}_5 \mathbf{y}$  avec  $\mathbf{A}_5 = \frac{1}{\sigma^2} \left( \mathbf{I}_N - \left\{ \frac{1}{n_i} \mathbf{J}_{n_i} \right\} - \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \right)$  et  $\mathbf{A}_4 = [x_{11} - \bar{x}_{1.}, \dots, x_{mn_m} - \bar{x}_{m.}] = \mathbf{x}^t - \left\{ l \mathbf{1}_{n_i}^t \bar{x}_i \right\}_{i=1}^m$ .

*Démonstration.* Premièrement, trouvons la forme de  $\mathbf{A}_4^t \mathbf{A}_4$

$$\begin{aligned}
\mathbf{A}_4^t \mathbf{A}_4 &= \begin{bmatrix} x_{11} - \bar{x}_{1.} \\ \vdots \\ x_{mn_m} - \bar{x}_{m.} \end{bmatrix} \begin{bmatrix} x_{11} - \bar{x}_{1.} & \cdots & x_{mn_m} - \bar{x}_{m.} \end{bmatrix} \\
&= \begin{pmatrix} (x_{11} - \bar{x}_{1.})^2 & \cdots & (x_{11} - \bar{x}_{1.})(x_{mn_m} - \bar{x}_{m.}) \\ \vdots & \ddots & \vdots \\ (x_{11} - \bar{x}_{1.})(x_{mn_m} - \bar{x}_{m.}) & \cdots & (x_{mn_m} - \bar{x}_{m.})^2 \end{pmatrix}
\end{aligned}$$

Ainsi,

$$\begin{aligned}
\mathbf{y}^t \mathbf{A}_5 \mathbf{y} &= \frac{1}{\sigma^2} \left[ y_{11} - \sum_{j=1}^{n_1} \frac{1}{n_1} y_{1j} - \frac{(x_{11} - \bar{x}_{1\cdot})}{S_{xx}} \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i\cdot}) y_{ij}, \dots, \right. \\
&\quad \left. y_{mn_m} - \sum_{j=1}^{n_m} \frac{1}{n_m} y_{mj} - \frac{(x_{mn_m} - \bar{x}_{m\cdot})}{S_{xx}} \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i\cdot}) y_{ij} \right] \mathbf{y} \\
&= \frac{1}{\sigma^2} \left[ y_{11} - \bar{y}_{1\cdot} - (x_{11} - \bar{x}_{1\cdot}) \frac{S_{xy}}{S_{xx}}, \dots, y_{mn_m} - \bar{y}_{m\cdot} - (x_{mn_m} - \bar{x}_{m\cdot}) \frac{S_{xy}}{S_{xx}} \right] \mathbf{y} \\
&= \frac{1}{\sigma^2} \left[ \sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij} (y_{ij} - \bar{y}_{i\cdot}) - \frac{S_{xy}}{S_{xx}} \sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij} (x_{ij} - \bar{x}_{i\cdot}) \right] \\
&= \frac{1}{\sigma^2} \left[ S_{yy} - \frac{S_{xy}^2}{S_{xx}} \right] \\
&= \frac{SCRR}{\sigma^2}
\end{aligned}$$

□

Donc il faut montrer que

**Proposition B.24.**  $\mathbf{A}_5 \mathbf{I}_N \sigma^2$  est idempotente.

*Démonstration.*

$$\begin{aligned}
\sigma^2 \mathbf{A}_5 \sigma^2 \mathbf{A}_5 &= \left( \mathbf{I}_N - \left\{ \frac{1}{d} \frac{\mathbf{J}_{n_i}}{n_i} \right\} - \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \right) \left( \mathbf{I}_N - \left\{ \frac{1}{d} \frac{\mathbf{J}_{n_i}}{n_i} \right\} - \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \right) \\
&= \mathbf{I}_N - 2 \left\{ \frac{1}{d} \frac{\mathbf{J}_{n_i}}{n_i} \right\} - 2 \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} + \left\{ \frac{1}{d} \frac{\mathbf{J}_{n_i}}{n_i} \right\} \left\{ \frac{1}{d} \frac{\mathbf{J}_{n_i}}{n_i} \right\} \\
&\quad + 2 \left\{ \frac{1}{d} \frac{\mathbf{J}_{n_i}}{n_i} \right\} \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} + \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \\
&= \mathbf{I}_N - \left\{ \frac{1}{d} \frac{\mathbf{J}_{n_i}}{n_i} \right\} - 2 \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} + \frac{S_{xx} \mathbf{A}_4^t \mathbf{A}_4}{S_{xx}^2} \\
&= \mathbf{I}_N - \left\{ \frac{1}{d} \frac{\mathbf{J}_{n_i}}{n_i} \right\} - \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \\
&= \sigma^2 \mathbf{A}_5
\end{aligned}$$

Car

$$\left\{ \begin{array}{c} 1 \\ d \ n_i \end{array} \mathbf{J}_{n_i} \right\} \left\{ \begin{array}{c} 1 \\ d \ n_i \end{array} \mathbf{J}_{n_i} \right\} = \left\{ \begin{array}{c} 1 \\ d \ n_i \end{array} \mathbf{J}_{n_i} \right\}$$

$$\left\{ \begin{array}{c} 1 \\ d \ n_i \end{array} \mathbf{J}_{n_i} \right\} \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} = \mathbf{0}_{N \times N}$$

$$(\mathbf{A}_4^t \mathbf{A}_4)(\mathbf{A}_4^t \mathbf{A}_4) = S_{xx} \mathbf{A}_4^t \mathbf{A}_4 \quad \square$$

Donc,

$$\frac{SCRR}{\sigma^2} \sim \chi^{2'} \left( \text{rang}(\mathbf{A}_5), \frac{1}{2} (\mu \mathbf{1}_N + \{c\alpha_i \mathbf{1}_{n_i}\} + \beta \mathbf{x})^t \mathbf{A}_5 (\mu \mathbf{1}_N + \{c\alpha_i \mathbf{1}_{n_i}\} + \beta \mathbf{x}) \right)$$

Il est facile de montrer que  $(\mu \mathbf{1}_N + \{c\alpha_i \mathbf{1}_{n_i}\})^t \mathbf{A}_5 (\mu \mathbf{1}_N + \{c\alpha_i \mathbf{1}_{n_i}\}) = 0$ . Pour  $\mathbf{x}^t \mathbf{A}_5 \mathbf{x}$ , voici les calculs,

$$\begin{aligned} \mathbf{x}^t \mathbf{A}_5 \mathbf{x} &= \frac{1}{\sigma^2} \left[ x_{11} - \sum_{j=1}^{n_1} \frac{1}{n_1} x_{1j} - \frac{(x_{11} - \bar{x}_{1\cdot})}{S_{xx}} \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i\cdot}) x_{ij}, \dots, \right. \\ &\quad \left. x_{mn_m} - \sum_{j=1}^{n_m} \frac{1}{n_m} x_{mj} - \frac{(x_{mn_m} - \bar{x}_{m\cdot})}{S_{xx}} \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i\cdot}) x_{ij} \right] \mathbf{x} \\ &= \frac{1}{\sigma^2} \left[ x_{11} - \bar{x}_{1\cdot} - (x_{11} - \bar{x}_{1\cdot}) \frac{S_{xx}}{S_{xx}}, \dots, x_{mn_m} - \bar{x}_{m\cdot} - (x_{mn_m} - \bar{x}_{m\cdot}) \frac{S_{xx}}{S_{xx}} \right] \mathbf{x} \\ &= \frac{1}{\sigma^2} \left[ \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} (x_{ij} - \bar{x}_{i\cdot}) - \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} (x_{ij} - \bar{x}_{i\cdot}) \right] \\ &= 0 \end{aligned}$$

Donc  $\frac{1}{2} (\mu \mathbf{1}_N + \{c\alpha_i \mathbf{1}_{n_i}\} + \beta \mathbf{x})^t \mathbf{A}_5 (\mu \mathbf{1}_N + \{c\alpha_i \mathbf{1}_{n_i}\} + \beta \mathbf{x}) = 0$ . Pour trouver le rang de  $\mathbf{A}_5$ , il suffit d'utiliser la proposition sur le rang des matrices idempotentes (B.5).

Il a été montré que  $\sigma^2 \mathbf{A}_5$  est idempotente, ainsi,

$$\begin{aligned} \text{rang}(\sigma^2 \mathbf{A}_5) &= \text{trace}(\sigma^2 \mathbf{A}_5) \\ &= \sum_{i=1}^m \sum_{j=1}^{n_i} 1 - \sum_{i=1}^m \sum_{j=1}^{n_i} \frac{1}{n_i} - \frac{1}{S_{xx}} \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})^2 \\ &= N - m - 1 \end{aligned}$$

Et  $\text{rang}(\mathbf{A}_5) = \text{rang}(\sigma^2 \mathbf{A}_5) = N - m - 1$

Donc finalement,

$$\frac{SCRR}{\sigma^2} \sim \chi_{(N-m-1)}^2 \quad (\text{B.13})$$

**b) Loi de  $\frac{SCC}{\sigma^2}$  sous  $H_0 : \beta = 0$**

Travaillons sous  $H_0 : \beta = 0$  pour trouver la loi de  $SCC$ . Ainsi,

$$\mathbf{y} \sim \mathcal{N}(\mu \mathbf{1}_N + \{c\alpha_i \mathbf{1}_{n_i}\}, \sigma^2 \mathbf{I}_N) \quad (\text{B.14})$$

Il suffit ensuite de montrer que

**Proposition B.25.**

$$z = S_{xy} = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})(y_{ij} - \bar{y}_{i.}) = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})y_{ij} \sim \mathcal{N}(0, \sigma^2 S_{xx})$$

*Démonstration.*  $z = \mathbf{A}_4 \mathbf{y}$  avec  $\mathbf{A}_4 = [(x_{11} - \bar{x}_{1.}), \dots, (x_{mn_m} - \bar{x}_{m.})]$ . Ainsi,

$$\begin{aligned}
z &\sim \mathcal{N}(\mathbf{A}_4(\mu \mathbf{1}_N + \{\alpha_i \mathbf{1}_{n_i}\}), \mathbf{A}_4 \sigma^2 \mathbf{I}_N \mathbf{A}_4^t) \\
&\sim \mathcal{N}(0, \sigma^2 S_{xx})
\end{aligned}
\quad \square$$

Et donc,

$$\begin{aligned}
z \sim \mathcal{N}(0, \sigma^2 S_{xx}) &\Rightarrow \frac{z - 0}{\sqrt{S_{xx} \sigma^2}} \sim \mathcal{N}(0, 1) \\
&\Rightarrow \left( \frac{z}{\sqrt{S_{xx} \sigma^2}} \right)^2 \sim \chi_{(1)}^2 \\
&\Rightarrow \frac{z^2}{S_{xx} \sigma^2} \sim \chi_{(1)}^2 \\
&\Rightarrow \frac{\left[ \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i) \right]^2}{S_{xx} \sigma^2} \sim \chi_{(1)}^2 \\
&\Rightarrow \frac{SCC}{\sigma^2} \sim \chi_{(1)}^2
\end{aligned}
\quad (\text{B.15})$$

c) **Indépendance de  $SCC$  et  $SCR$  sous  $H_0 : \beta = 0$**

Pour être capable d'effectuer le test de Fisher pour  $H_0 : \beta = 0$  il faut montrer l'indépendance entre  $SCC$  et  $SCR$ .

Posons  $\mathbf{A}_6 = \frac{\mathbf{A}_4^t \mathbf{A}_4}{\sigma^2 S_{xx}}$ , donc  $SCC = \mathbf{y}^t \mathbf{A}_6 \mathbf{y}$  et  $SCR = \mathbf{y}^t \mathbf{A}_5 \mathbf{y}$ . Montrons que

**Proposition B.26.**  $\mathbf{A}_5 \sigma^2 \mathbf{I}_N \mathbf{A}_6 = \mathbf{0}_{N \times N}$

*Démonstration.*

$$\begin{aligned}
\mathbf{A}_5 \sigma^2 \mathbf{I}_N \mathbf{A}_6 &= \frac{1}{\sigma^2} \left( \mathbf{I}_N - \left\{ \frac{1}{d} \mathbf{J}_{n_i} \right\} - \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \right) \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \\
&= \frac{1}{\sigma^2} \left( \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} - \left\{ \frac{1}{d} \mathbf{J}_{n_i} \right\} \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} - \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \right) \\
&= \mathbf{0}_{N \times N}
\end{aligned}$$



Par les mêmes arguments utilisés dans la preuve de la proposition B.24 □

Ce qui démontre la prochaine proposition.

**Proposition B.27.**  $\frac{SCC/1}{SCRR/(N-m-1)} \sim \mathcal{F}_{N-m-1}^1$  sous  $H_0 : \beta = 0$

**d) Loi de  $\frac{SCE}{\sigma^2}$  sous  $H_0 : \alpha_i = 0 \forall i$  et dans le cas où les données sont équilibrées**

Trouvons la loi de  $\frac{SCE}{\sigma^2}$  sous  $H_0 : \alpha_i = 0 \forall i$ , mais seulement dans le cas où les données sont équilibrées (car il est possible de montrer que  $\frac{SCE}{\sigma^2}$  ne suit pas une loi du khi carré centrée dans le cas où les données ne sont pas équilibrées). Ainsi,  $x_{ij} = x_j \forall i$  et  $n_i = n \forall i$ . Dans ce contexte, la loi de  $\mathbf{y}$  est donnée par

$$\mathbf{y} \sim \mathcal{N}(\mu \mathbf{1}_N + \beta \mathbf{x}, \mathbf{I}_N \sigma^2) \quad (\text{B.16})$$

Où  $\mathbf{x} = [x_1, x_2, \dots, x_n, x_1, \dots, x_n]^t$  est le vecteur de longueur  $N$  contenant tout les  $x_j$  correspondants aux  $y_{ij}$ .

**Proposition B.28.**

$$\frac{SCE}{\sigma^2} = \mathbf{y}^t \mathbf{B}_1 \mathbf{y}$$

$$\text{Où } \mathbf{B}_1 = \frac{1}{\sigma^2} \left( \mathbf{I}_m \otimes \frac{1}{n} \mathbf{J}_n - \frac{1}{N} \mathbf{J}_N \right).$$

*Démonstration.*

$$\begin{aligned} \mathbf{y}^t \mathbf{B}_1 \mathbf{y} &= \frac{1}{\sigma^2} [\bar{y}_{1\cdot} - \bar{y}_{\cdot\cdot}, \dots, \bar{y}_{2\cdot} - \bar{y}_{\cdot\cdot}, \dots, \bar{y}_{m\cdot} - \bar{y}_{\cdot\cdot}] \begin{bmatrix} y_{11} \\ \vdots \\ y_{mn} \end{bmatrix} \\ &= \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^n y_{ij} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^n (\bar{y}_i - \bar{y}_{..})^2 \\
&= \frac{SCE}{\sigma^2}
\end{aligned}$$

□

Pour montrer que  $\frac{SCE}{\sigma^2}$  suit une loi du khi carré, il faut montrer que

**Proposition B.29.**  $\mathbf{B}_1 \sigma^2 \mathbf{I}_N = \mathbf{I}_m \otimes \frac{1}{n} \mathbf{J}_n - \frac{1}{N} \mathbf{J}_N$  est idempotente

*Démonstration.*

$$\begin{aligned}
\sigma^2 \mathbf{B}_1 \sigma^2 \mathbf{B}_1 &= \left( \mathbf{I}_m \otimes \frac{1}{n} \mathbf{J}_n - \frac{1}{N} \mathbf{J}_N \right) \left( \mathbf{I}_m \otimes \frac{1}{n} \mathbf{J}_n - \frac{1}{N} \mathbf{J}_N \right) \\
&= \left( \left( \mathbf{I}_m \otimes \frac{1}{n} \mathbf{J}_n \right) - \left( \mathbf{J}_m \otimes \frac{1}{N} \mathbf{J}_n \right) \right) \left( \left( \mathbf{I}_m \otimes \frac{1}{n} \mathbf{J}_n \right) - \left( \mathbf{J}_m \otimes \frac{1}{N} \mathbf{J}_n \right) \right) \\
&= \left( \mathbf{I}_m \otimes \frac{1}{n^2} \mathbf{J}_n^2 \right) - 2 \left( \mathbf{J}_m \otimes \frac{1}{nN} \mathbf{J}_n^2 \right) + \left( \mathbf{J}_m^2 \otimes \frac{1}{N^2} \mathbf{J}_n^2 \right) \\
&= \left( \mathbf{I}_m \otimes \frac{1}{n} \mathbf{J}_n \right) - \frac{2}{N} \left( \mathbf{J}_m \otimes \mathbf{J}_n \right) + \frac{mn}{N^2} \left( \mathbf{J}_m \otimes \mathbf{J}_n \right) \\
&= \left( \mathbf{I}_m \otimes \frac{1}{n} \mathbf{J}_n \right) - \frac{1}{N} \mathbf{J}_N \\
&= \sigma^2 \mathbf{B}_1
\end{aligned}$$

□

Donc,

$$\frac{SCE}{\sigma^2} \sim \chi^{2l} \left( \text{rang}(\mathbf{B}_1), \frac{1}{2} (\mu \mathbf{1}_N + \beta \mathbf{x})^t \mathbf{B}_1 (\mu \mathbf{1}_N + \beta \mathbf{x}) \right)$$

Il est facile de montrer que  $(\mu \mathbf{1}_N + \beta \mathbf{x})^t \mathbf{B}_1 (\mu \mathbf{1}_N + \beta \mathbf{x}) = 0$  puisque les données sont équilibrées. Aussi,  $\text{rang}(\mathbf{B}_1) = m - 1$ , car  $\text{rang}(\mathbf{B}_1) = \text{rang}(\sigma^2 \mathbf{B}_1) = \text{trace}(\sigma^2 \mathbf{B}_1) = \frac{N}{n} - \frac{N}{N} = m - 1$ , car  $\sigma^2 \mathbf{B}_1$  est idempotente.

Donc finalement,

$$\frac{SCE}{\sigma^2} \sim \chi_{(m-1)}^2 \quad (\text{B.17})$$

Si les données sont non équilibrées, il est possible de montrer que

$$\frac{SCE}{\sigma^2} \sim \chi^{2'} \left( m - 1, \beta^2 \sum_{i=1}^m \sum_{j=1}^{n_i} (\bar{x}_{i.} - \bar{x}_{..})^2 \right) \quad (\text{B.18})$$

e) **Indépendance de SCE et SCRR sous  $H_0 : \alpha_i = 0 \forall i$**

Pour être capable d'effectuer le test de Fisher pour  $H_0 : \alpha_i = 0 \forall i$  il faut montrer l'indépendance entre SCE et SCRR. Montrons le de façon générale dans le cas où les données sont non équilibrées

**Proposition B.30.**  $\mathbf{A}_5 \sigma^2 \mathbf{I}_N \mathbf{B}_2 = \mathbf{0}_{N \times N}$

Où  $\mathbf{B}_2$  est l'équivalent de  $\mathbf{B}_1$ , mais dans le cas où les données sont non équilibrées. Donc  $\mathbf{B}_2$  peut s'écrire comme

$$\mathbf{B}_2 = \frac{1}{\sigma^2} \left( \left\{ \frac{1}{d} \mathbf{J}_{n_i} \right\} - \frac{1}{N} \mathbf{J}_N \right)$$

*Démonstration.*

$$\begin{aligned} \mathbf{A}_5 \sigma^2 \mathbf{I}_N \mathbf{B}_2 &= \frac{1}{\sigma^2} \left( \mathbf{I}_N - \left\{ \frac{1}{d} \mathbf{J}_{n_i} \right\} - \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \right) \left( \left\{ \frac{1}{d} \mathbf{J}_{n_i} \right\} - \frac{1}{N} \mathbf{J}_N \right) \\ &= \frac{1}{\sigma^2} \left( \left\{ \frac{1}{d} \mathbf{J}_{n_i} \right\} - \frac{1}{N} \mathbf{J}_N - \left\{ \frac{1}{d} \mathbf{J}_{n_i} \right\} + \frac{1}{N} \mathbf{J}_N \right) \\ &= \mathbf{0}_{N \times N} \end{aligned}$$

Par des arguments semblables à ceux de la preuve de la proposition B.24 □

Ce qui démontre la prochaine proposition.

**Proposition B.31.**  $\frac{SCE/(m-1)}{SCR/(N-m-1)} \sim \mathcal{F}_{N-m-1}^{m-1}$  sous  $H_0 : \alpha_i = 0 \forall i$  mais seulement dans le cas où les données sont équilibrées.

## B.6 Autres preuves pour la régression linéaire simple avec un effet aléatoire et des données équilibrées

### B.6.1 Détails pour $\hat{\mu}$ et $\hat{\beta}$

**Proposition B.32.**

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} = (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} = \begin{bmatrix} \bar{y}_{..} - \frac{S_{xy}}{S_{xx}} \bar{x}_{..} \\ \frac{S_{xy}}{S_{xx}} \end{bmatrix}$$

*Démonstration.* Remplaçons  $\mathbf{V}$ ,  $\mathbf{X}$  et  $\mathbf{y}$  par leur valeur.

$$\begin{aligned} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} &= [(\mathbf{1}_m^t \otimes \mathbf{X}_0^t)(\mathbf{I}_m \otimes \mathbf{V}_0^{-1})(\mathbf{1}_m \otimes \mathbf{X}_0)]^{-1} [(\mathbf{1}_m^t \otimes \mathbf{X}_0^t)(\mathbf{I}_m \otimes \mathbf{V}_0^{-1})] \mathbf{y} \\ &= [m \otimes (\mathbf{X}_0^t \mathbf{V}_0^{-1} \mathbf{X}_0)]^{-1} [\mathbf{1}_m^t \otimes (\mathbf{X}_0^t \mathbf{V}_0^{-1})] \mathbf{y} \\ &= \left[ \frac{1}{m} (\mathbf{X}_0^t \mathbf{V}_0^{-1} \mathbf{X}_0)^{-1} \right] [\mathbf{1}_m^t \otimes (\mathbf{X}_0^t \mathbf{V}_0^{-1})] \mathbf{y} \end{aligned}$$

Ensuite, rappelons quelques notations avant de poursuivre.

$$\tau = \frac{n\sigma_a^2}{\sigma^2 + n\sigma_a^2} = \frac{\sigma_a^2}{\sigma_a^2 + \sigma^2/n}$$

$$1 - \tau = \frac{\sigma^2}{\sigma^2 + n\sigma_a^2} = \frac{\sigma^2}{n\sigma_a^2} \tau$$

$$S_{xx} = \sum_{i=1}^m \sum_{j=1}^n (x_j - \bar{x})^2 = m \sum_{j=1}^n (x_j^2) - mn\bar{x}^2$$

$$S_{xy} = \sum_{i=1}^m \sum_{j=1}^n (x_j - \bar{x})(y_{ij} - \bar{y}_i) = m \sum_{j=1}^n (x_j - \bar{x})(\bar{y}_j - \bar{y}_..) = m \sum_{j=1}^n (x_j \bar{y}_j) - mn\bar{x}\bar{y}_..$$

Maintenant, décortiquons un peu le problème.

$$\begin{aligned} \mathbf{X}_0^t \mathbf{V}_0^{-1} &= \begin{bmatrix} \mathbf{1}_n^t \\ \mathbf{x}_0^t \end{bmatrix} \frac{1}{\sigma^2} (\mathbf{I}_n - \tau \bar{\mathbf{J}}_n) \\ &= \frac{1}{\sigma^2} \begin{pmatrix} 1 & \cdots & 1 \\ x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} 1 - \frac{\tau}{n} & -\frac{\tau}{n} & \cdots & -\frac{\tau}{n} \\ -\frac{\tau}{n} & 1 - \frac{\tau}{n} & \cdots & -\frac{\tau}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{\tau}{n} & -\frac{\tau}{n} & \cdots & 1 - \frac{\tau}{n} \end{pmatrix} \\ &= \frac{1}{\sigma^2} \begin{pmatrix} 1 - \tau & \cdots & 1 - \tau \\ x_1 - \frac{\tau}{n} \sum_{j=1}^n x_j & \cdots & x_n - \frac{\tau}{n} \sum_{j=1}^n x_j \end{pmatrix} \\ &= \frac{1}{\sigma^2} \begin{pmatrix} \frac{\sigma^2 \tau}{n \sigma_a^2} & \cdots & \frac{\sigma^2 \tau}{n \sigma_a^2} \\ x_1 - \tau \bar{x} & \cdots & x_n - \tau \bar{x} \end{pmatrix} \\ &= \begin{bmatrix} \frac{\tau}{n \sigma_a^2} \mathbf{1}_n^t \\ \frac{1}{\sigma^2} (\mathbf{x}_0^t - \tau \bar{x} \mathbf{1}_n^t) \end{bmatrix} \end{aligned}$$

$$\begin{aligned} \mathbf{X}_0^t \mathbf{V}_0^{-1} \mathbf{X}_0 &= \mathbf{X}_0^t \mathbf{V}_0^{-1} [\mathbf{1}_n \quad \mathbf{x}_0] \\ &= \begin{pmatrix} \frac{\tau}{n \sigma_a^2} & \cdots & \frac{\tau}{n \sigma_a^2} \\ \frac{x_1 - \tau \bar{x}}{\sigma^2} & \cdots & \frac{x_n - \tau \bar{x}}{\sigma^2} \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \\ &= \begin{pmatrix} \frac{\tau}{\sigma_a^2} & \frac{\tau \bar{x}}{\sigma_a^2} \\ \frac{1}{\sigma^2} \sum_{j=1}^n (x_j - \tau \bar{x}) & \frac{1}{\sigma^2} \sum_{j=1}^n (x_j - \tau \bar{x}) x_j \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
&= \begin{pmatrix} \frac{\tau}{\sigma_a^2} & \frac{\tau\bar{x}}{\sigma_a^2} \\ \frac{1}{\sigma^2}(n\bar{x} - n\tau\bar{x}) & \frac{1}{\sigma^2} \left( \sum_{j=1}^n (x_j^2) - n\tau\bar{x}^2 \right) \end{pmatrix} \\
&= \begin{pmatrix} \frac{\tau}{\sigma_a^2} & \frac{\tau\bar{x}}{\sigma_a^2} \\ \frac{n\bar{x}}{\sigma^2}(1 - \tau) & \frac{1}{\sigma^2} \left( \frac{S_{xx}}{m} + n\bar{x}^2 - n\tau\bar{x}^2 \right) \end{pmatrix} \\
&= \begin{pmatrix} \frac{\tau}{\sigma_a^2} & \frac{\tau\bar{x}}{\sigma_a^2} \\ \frac{n\bar{x}\sigma^2\tau}{\sigma^2 n\sigma_a^2} & \frac{1}{\sigma^2} \left( \frac{S_{xx}}{m} + \frac{n\bar{x}^2\sigma^2\tau}{n\sigma_a^2} \right) \end{pmatrix} \\
&= \frac{\tau}{\sigma_a^2} \begin{pmatrix} 1 & \bar{x} \\ \bar{x} & \frac{\sigma_a^2 S_{xx}}{m\sigma^2\tau} + \bar{x}^2 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
(\mathbf{1}_m^t \otimes \mathbf{X}_0^t \mathbf{V}_0^{-1})\mathbf{y} &= \begin{pmatrix} \frac{\tau}{\frac{n\sigma_a^2}{x_1 - \tau\bar{x}}} & \cdots & \frac{\tau}{\frac{n\sigma_a^2}{x_n - \tau\bar{x}}} & \frac{\tau}{\frac{n\sigma_a^2}{x_1 - \tau\bar{x}}} & \cdots & \frac{\tau}{\frac{n\sigma_a^2}{x_n - \tau\bar{x}}} \\ \frac{1}{\sigma^2} & \cdots & \frac{1}{\sigma^2} & \frac{1}{\sigma^2} & \cdots & \frac{1}{\sigma^2} \end{pmatrix}_{2 \times mn} \begin{bmatrix} y_{11} \\ \vdots \\ y_{mn} \end{bmatrix}_{mn \times 1} \\
&= \begin{bmatrix} \frac{\tau}{n\sigma_a^2} \sum_{i=1}^m \sum_{j=1}^n y_{ij} \\ \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^n (x_j - \tau\bar{x}) y_{ij} \end{bmatrix} \\
&= \begin{bmatrix} \frac{m\tau}{\sigma_a^2} \bar{y}_{..} \\ \frac{1}{\sigma^2} \left( \sum_{j=1}^n (x_j - \tau\bar{x}) \sum_{i=1}^m y_{ij} \right) \end{bmatrix} \\
&= \begin{bmatrix} \frac{m\tau}{\sigma_a^2} \bar{y}_{..} \\ \frac{m}{\sigma^2} \left( \sum_{j=1}^n (x_j - \tau\bar{x}) \bar{y}_{.j} \right) \end{bmatrix} \\
&= \begin{bmatrix} \frac{m\tau}{\sigma_a^2} \bar{y}_{..} \\ \frac{m}{\sigma^2} \left( \sum_{j=1}^n x_j \bar{y}_{.j} - \tau\bar{x} \sum_{j=1}^n \bar{y}_{.j} - n\bar{x}\bar{y}_{..} + n\bar{x}\bar{y}_{..} \right) \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
&= \left[ \begin{array}{c} \frac{m\tau}{\sigma_a^2} \bar{y}_{..} \\ \frac{m}{\sigma^2} \left( \frac{S_{xy}}{m} + n\bar{x}\bar{y}_{..}(1-\tau) \right) \end{array} \right] \\
&= \left[ \begin{array}{c} \frac{m\tau}{\sigma_a^2} \bar{y}_{..} \\ \frac{m}{\sigma^2} \left( \frac{S_{xy}}{m} + \frac{\sigma^2\tau n\bar{x}\bar{y}_{..}}{n\sigma_a^2} \right) \end{array} \right] \\
&= \frac{m\tau}{\sigma_a^2} \left[ \begin{array}{c} \bar{y}_{..} \\ \frac{\sigma_a^2 S_{xy}}{m\tau\sigma^2} + \bar{x}\bar{y}_{..} \end{array} \right]
\end{aligned}$$

$$\begin{aligned}
(\mathbf{X}_0^t \mathbf{V}_0^{-1} \mathbf{X}_0)^{-1} &= \begin{pmatrix} \frac{\tau}{\sigma_a^2} & \frac{\bar{x}\tau}{\sigma_a^2} \\ \frac{\bar{x}\tau}{\sigma_a^2} & \frac{S_{xx}}{m\sigma^2} + \frac{\bar{x}^2\tau}{\sigma_a^2} \end{pmatrix}^{-1} \\
&= \frac{1}{\star} \begin{pmatrix} \frac{S_{xx}}{m\sigma^2} + \frac{\bar{x}^2\tau}{\sigma_a^2} & -\frac{\bar{x}\tau}{\sigma_a^2} \\ -\frac{\bar{x}\tau}{\sigma_a^2} & \frac{\tau}{\sigma_a^2} \end{pmatrix}
\end{aligned}$$

★ = déterminant

$$\begin{aligned}
&= \frac{\tau}{\sigma_a^2} \left( \frac{S_{xx}}{m\sigma^2} + \frac{\bar{x}^2\tau}{\sigma_a^2} \right) - \frac{\bar{x}^2\tau^2}{\sigma_a^4} \\
&= \frac{\tau S_{xx}}{m\sigma^2\sigma_a^2} + \frac{\bar{x}^2\tau^2}{\sigma_a^4} - \frac{\bar{x}^2\tau^2}{\sigma_a^4} \\
&= \frac{\tau S_{xx}}{m\sigma^2\sigma_a^2}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow (\mathbf{X}_0^t \mathbf{V}_0^{-1} \mathbf{X}_0)^{-1} &= \frac{m\sigma_a^2\sigma^2}{\tau S_{xx}} \begin{pmatrix} \frac{S_{xx}}{m\sigma^2} + \frac{\bar{x}^2\tau}{\sigma_a^2} & -\frac{\bar{x}\tau}{\sigma_a^2} \\ -\frac{\bar{x}\tau}{\sigma_a^2} & \frac{\tau}{\sigma_a^2} \end{pmatrix} \\
&= \frac{m\sigma^2}{S_{xx}} \begin{pmatrix} \frac{\sigma_a^2 S_{xx}}{m\sigma^2\tau} + \bar{x}^2 & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix}
\end{aligned}$$

Avec les résultats précédents, ça va être plus facile de trouver le vecteur  $[\hat{\mu}, \hat{\beta}]^t$ .

$$\begin{aligned}
\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} &= \left[ \frac{1}{m} (\mathbf{X}_0^t \mathbf{V}_0^{-1} \mathbf{X}_0)^{-1} \right] \left[ \mathbf{1}_m^t \otimes (\mathbf{X}_0^t \mathbf{V}_0^{-1}) \right] \mathbf{y} \\
&= \frac{1}{m} \frac{m\sigma^2}{S_{xx}} \begin{pmatrix} \frac{\sigma_a^2 S_{xx}}{m\sigma^2 \tau} + \bar{x}^2 & -\bar{x} \\ -\bar{x} & 1 \end{pmatrix} \left( \frac{m\tau}{\sigma_a^2} \begin{bmatrix} \bar{y} \cdot \\ \frac{\sigma_a^2 S_{xy}}{m\tau\sigma^2} + \bar{x}\bar{y} \cdot \end{bmatrix} \right) \\
&= \frac{m\sigma^2 \tau}{\sigma_a^2 S_{xx}} \begin{bmatrix} \frac{\sigma_a^2 \bar{y} \cdot S_{xx}}{m\sigma^2 \tau} + \bar{x}^2 \bar{y} \cdot - \frac{\sigma_a^2 \bar{x} S_{xy}}{m\tau\sigma^2} - \bar{x}^2 \bar{y} \cdot \\ -\bar{x}\bar{y} \cdot + \frac{\sigma_a^2 S_{xy}}{m\tau\sigma^2} + \bar{x}\bar{y} \cdot \end{bmatrix} \\
&= \frac{m\sigma^2 \tau}{\sigma_a^2 S_{xx}} \begin{bmatrix} \frac{\sigma_a^2 \bar{y} \cdot S_{xx}}{m\sigma^2 \tau} - \frac{\sigma_a^2 \bar{x} S_{xy}}{m\tau\sigma^2} \\ \frac{\sigma_a^2 S_{xy}}{m\tau\sigma^2} \end{bmatrix} \\
&= \begin{bmatrix} \bar{y} \cdot - \bar{x} \frac{S_{xy}}{S_{xx}} \\ \frac{S_{xy}}{S_{xx}} \end{bmatrix}
\end{aligned}$$

□

## B.6.2 Détails pour la vraisemblance modifiée $l^*$

a) Trouver une formule plus simple de la vraisemblance modifiée

Voici premièrement la vraisemblance modifiée.

$$l^* = -\frac{mn}{2} \log(2\pi) - \frac{1}{2} \log(|\mathbf{V}|) - \frac{1}{2} (\mathbf{y} - \boldsymbol{\Theta})^t \mathbf{V}^{-1} (\mathbf{y} - \boldsymbol{\Theta}) \quad (\text{B.19})$$

Où  $\boldsymbol{\Theta} = \mathbb{E}[\mathbf{y}] = \mathbf{X} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix}$ ,  $|\mathbf{V}| = (\sigma^2)^{m(n-1)} (\sigma^2 + n\sigma_a^2)^m$  et où  $\mathbf{V}^{-1} = \mathbf{I}_m \otimes \left[ \frac{1}{\sigma^2} (\mathbf{I}_n - \tau \bar{\mathbf{J}}_n) \right]$ .

**Proposition B.33.**

$$l^* = -\frac{1}{2} \left[ mn \log(2\pi) + m(n-1) \log(\sigma^2) + m \log(\sigma^2 + n\sigma_a^2) + \frac{S_1^*}{\sigma^2} - \frac{n\sigma_a^2 S_2^*}{\sigma^2(\sigma^2 + n\sigma_a^2)} \right]$$

*Démonstration.* La vraisemblance modifiée devient



$$l^* = -\frac{1}{2} \left[ mn \log(2\pi) + m(n-1) \log(\sigma^2) + m \log(\sigma^2 + n\sigma_a^2) \right. \\ \left. + \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \right)^t \left( \mathbf{I}_m \otimes \left[ \frac{1}{\sigma^2} (\mathbf{I}_n - \tau \bar{\mathbf{J}}_n) \right] \right) \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \right) \right]$$

Avec,

$$\begin{aligned} \mathbf{y} - \mathbf{X} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} &= \begin{bmatrix} y_{11} \\ \vdots \\ y_{1n} \\ y_{21} \\ \vdots \\ y_{mn} \end{bmatrix} - \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \\ 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \\ &= \begin{bmatrix} y_{11} \\ \vdots \\ y_{1n} \\ y_{21} \\ \vdots \\ y_{mn} \end{bmatrix} - \begin{bmatrix} \hat{\mu} + x_1 \hat{\beta} \\ \vdots \\ \hat{\mu} + x_n \hat{\beta} \\ \hat{\mu} + x_1 \hat{\beta} \\ \vdots \\ \hat{\mu} + x_n \hat{\beta} \end{bmatrix} \\ &= \begin{bmatrix} y_{11} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{1n} - \hat{\mu} - x_n \hat{\beta} \\ y_{21} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{mn} - \hat{\mu} - x_n \hat{\beta} \end{bmatrix} \end{aligned}$$

Cherchons  $\star = (\mathbf{y} - \Theta)^t \left( \mathbf{I}_m \otimes \left[ \frac{1}{\sigma^2} (\mathbf{I}_n - \tau \bar{\mathbf{J}}_n) \right] \right) (\mathbf{y} - \Theta)$ ,

$$\begin{aligned}
\star &= \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \right)^t \left( \mathbf{I}_m \otimes \frac{1}{\sigma^2} (\mathbf{I}_n - \tau \bar{\mathbf{J}}_n) \right) \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \right) \\
&= \begin{bmatrix} y_{11} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{1n} - \hat{\mu} - x_n \hat{\beta} \\ y_{21} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{mn} - \hat{\mu} - x_n \hat{\beta} \end{bmatrix}^t \left( \mathbf{I}_m \otimes \frac{1}{\sigma^2} (\mathbf{I}_n - \tau \bar{\mathbf{J}}_n) \right) \begin{bmatrix} y_{11} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{1n} - \hat{\mu} - x_n \hat{\beta} \\ y_{21} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{mn} - \hat{\mu} - x_n \hat{\beta} \end{bmatrix} \\
&= \frac{1}{\sigma^2} \begin{bmatrix} y_{11} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{1n} - \hat{\mu} - x_n \hat{\beta} \\ y_{21} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{mn} - \hat{\mu} - x_n \hat{\beta} \end{bmatrix}^t \left( \begin{array}{cccc} 1 - \frac{\tau}{n} & \cdots & -\frac{\tau}{n} & \\ \vdots & \ddots & \vdots & \cdots \\ -\frac{\tau}{n} & \cdots & 1 - \frac{\tau}{n} & \\ \vdots & & \vdots & \ddots \\ \mathbf{0}_{n \times n} & \cdots & \vdots & \ddots \\ & & 1 - \frac{\tau}{n} & \cdots & -\frac{\tau}{n} \\ & & \vdots & \ddots & \vdots \\ & & -\frac{\tau}{n} & \cdots & 1 - \frac{\tau}{n} \end{array} \right) \begin{bmatrix} y_{11} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{1n} - \hat{\mu} - x_n \hat{\beta} \\ y_{21} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{mn} - \hat{\mu} - x_n \hat{\beta} \end{bmatrix} \\
&= \frac{1}{\sigma^2} \left\{ y_{ij} - \hat{\mu} - x_j \hat{\beta} - \frac{\tau}{n} \sum_{k=1}^n (y_{ik} - \hat{\mu} - \hat{\beta} x_k) \right\}_{i=1, j=1}^{m \quad n} \begin{bmatrix} y_{11} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{1n} - \hat{\mu} - x_n \hat{\beta} \\ y_{21} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{mn} - \hat{\mu} - x_n \hat{\beta} \end{bmatrix} \\
&= \frac{1}{\sigma^2} \left\{ y_{ij} - \hat{\mu} - x_j \hat{\beta} - \tau (\bar{y}_{i \cdot} - \hat{\mu} - \hat{\beta} \bar{x}) \right\}_{i=1, j=1}^{m \quad n} \begin{bmatrix} y_{11} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{1n} - \hat{\mu} - x_n \hat{\beta} \\ y_{21} - \hat{\mu} - x_1 \hat{\beta} \\ \vdots \\ y_{mn} - \hat{\mu} - x_n \hat{\beta} \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\sigma^2} \left[ \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \hat{\mu} - \hat{\beta}x_j)^2 - \tau \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \hat{\mu} - \hat{\beta}x_j)(\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x}) \right] \\
&= \frac{1}{\sigma^2} \left[ S_1^* - \tau \sum_{i=1}^m (\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x}) \left( \sum_{j=1}^n (y_{ij} - \hat{\mu} - \hat{\beta}x_j) \right) \right] \\
&= \frac{1}{\sigma^2} \left[ S_1^* - \tau \sum_{i=1}^m (\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x}) n (\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x}) \right] \\
&= \frac{1}{\sigma^2} \left[ S_1^* - \tau \sum_{i=1}^m n (\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x})^2 \right] \\
&= \frac{1}{\sigma^2} [S_1^* - \tau S_2^*]
\end{aligned}$$

Car  $S_1^* = \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \hat{\mu} - \hat{\beta}x_j)^2$  et  $S_2^* = \sum_{i=1}^m n (\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x})^2$ .

Remettons cette expression dans la vraisemblance modifiée.

$$\begin{aligned}
l^* &= -\frac{1}{2} \left[ mn \log(2\pi) + m(n-1) \log(\sigma^2) + m \log(\sigma^2 + n\sigma_a^2) + \frac{1}{\sigma^2} [S_1^* - \tau S_2^*] \right] \\
&= -\frac{1}{2} \left[ mn \log(2\pi) + m(n-1) \log(\sigma^2) + m \log(\sigma^2 + n\sigma_a^2) + \frac{S_1^*}{\sigma^2} - \frac{n\sigma_a^2 S_2^*}{\sigma^2(\sigma^2 + n\sigma_a^2)} \right]
\end{aligned}$$

□

b) **Autre formule pour  $S_1^*$**

Il manque une proposition pour montrer que  $S_1^* = SCE + SCRR$ .

**Proposition B.34.**  $SCRR = SCR - SCC$

*Démonstration.*

$$\begin{aligned}
SCRR &= \sum_{j=1}^n \sum_{i=1}^m \left[ y_{ij} - \bar{y}_{i.} - \frac{S_{xy}}{S_{xx}}(x_j - \bar{x}) \right]^2 \\
&= \sum_{j=1}^n \sum_{i=1}^m \left[ (y_{ij} - \bar{y}_{i.})^2 - 2(y_{ij} - \bar{y}_{i.}) \frac{S_{xy}}{S_{xx}}(x_j - \bar{x}) + \frac{S_{xy}^2}{S_{xx}^2}(x_j - \bar{x})^2 \right] \\
&= \sum_{j=1}^n \sum_{i=1}^m (y_{ij} - \bar{y}_{i.})^2 - 2 \frac{S_{xy}}{S_{xx}} \sum_{i=1}^m \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})(x_j - \bar{x}) + \frac{S_{xy}^2}{S_{xx}^2} \sum_{i=1}^m \sum_{j=1}^n (x_j - \bar{x})^2 \\
&= SCR - 2 \frac{S_{xy}}{S_{xx}} S_{xy} + \frac{S_{xy}^2}{S_{xx}^2} S_{xx} \\
&= SCR - \frac{S_{xy}^2}{S_{xx}} \\
&= SCR - SCC
\end{aligned}$$

□

## B.7 Autres preuves pour la régression linéaire simple avec un effet aléatoire et des données non équilibrées

### B.7.1 Expressions détaillées de $\hat{\mu}$ et $\hat{\beta}$

Nous sommes dans le contexte du modèle 2.29. Cherchons une formule plus détaillée pour le vecteur  $[\hat{\mu}, \hat{\beta}]^t = (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^t \mathbf{V}^{-1}$ .

**Proposition B.35.**

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} \tilde{y}_{..} - \hat{\beta} \tilde{x}_{..} \\ \frac{(\sigma_a^2 / \sigma^2) S_{xy} + W S M_{xy}}{(\sigma_a^2 / \sigma^2) S_{xx} + W S M_{xx}} \end{bmatrix}$$

*Démonstration.* Commençons par trouver  $\mathbf{X}^t \mathbf{V}^{-1}$ .

$$\begin{aligned}
\mathbf{X}^t \mathbf{V}^{-1} &= \begin{pmatrix} 1 & \cdots & 1 \\ x_{11} & \cdots & x_{mn_m} \end{pmatrix} \frac{1}{\sigma^2} \begin{pmatrix} 1 - \frac{\tau_1}{n_1} & \cdots & -\frac{\tau_1}{n_1} & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ -\frac{\tau_1}{n_1} & \cdots & 1 - \frac{\tau_1}{n_1} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 1 - \frac{\tau_m}{n_m} & \cdots & -\frac{\tau_m}{n_m} \\ \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & -\frac{\tau_m}{n_m} & \cdots & 1 - \frac{\tau_m}{n_m} \end{pmatrix} \\
&= \frac{1}{\sigma^2} \begin{pmatrix} 1 - \tau_1 & \cdots & 1 - \tau_1 & \cdots & 1 - \tau_m \\ x_{11} - \frac{\tau_1}{n_1} \sum_{j=1}^{n_1} x_{1j} & \cdots & x_{1n_1} - \frac{\tau_1}{n_1} \sum_{j=1}^{n_1} x_{1j} & \cdots & x_{mn_m} - \frac{\tau_m}{n_m} \sum_{j=1}^{n_m} x_{mj} \end{pmatrix} \\
&= \begin{pmatrix} \frac{\tau_1}{n_1 \sigma_a^2} & \cdots & \frac{\tau_1}{n_1 \sigma_a^2} & \cdots & \frac{\tau_m}{n_m \sigma_a^2} \\ (x_{11} - \tau_1 \bar{x}_{1.}) / \sigma^2 & \cdots & (x_{1n_1} - \tau_1 \bar{x}_{1.}) / \sigma^2 & \cdots & (x_{mn_m} - \tau_m \bar{x}_{m.}) / \sigma^2 \end{pmatrix} \\
&= \left\{ \begin{matrix} \left[ \begin{matrix} \left( \frac{\tau_i}{n_i \sigma_a^2} \right) \mathbf{1}_{n_i}^t \\ (\mathbf{x}_i^t - \tau_i \bar{x}_i \cdot \mathbf{1}_{n_i}^t) / \sigma^2 \end{matrix} \right] \end{matrix} \right\}_{i=1}^m
\end{aligned}$$

Ensuite, trouvons une formule détaillée pour  $\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X}$ .

$$\begin{aligned}
\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X} &= \mathbf{X}^t \mathbf{V}^{-1} \{c [\mathbf{1}_{n_i}, \mathbf{x}_i]\} \\
&= \left\{ \begin{matrix} \left[ \begin{matrix} \left( \frac{\tau_i}{n_i \sigma_a^2} \right) \mathbf{1}_{n_i}^t \\ (\mathbf{x}_i^t - \tau_i \bar{x}_i \cdot \mathbf{1}_{n_i}^t) / \sigma^2 \end{matrix} \right] \end{matrix} \right\}_{i=1}^m \{c [\mathbf{1}_{n_i}, \mathbf{x}_i]\}_{i=1}^m \\
&= \begin{pmatrix} \sum_{i=1}^m \frac{\tau_i}{\sigma_a^2} & \sum_{i=1}^m \sum_{j=1}^{n_i} \frac{\tau_i x_{ij}}{n_i \sigma_a^2} \\ \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \tau_i \bar{x}_i) & \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} (x_{ij} - \tau_i \bar{x}_i) \end{pmatrix} \\
&= \frac{1}{\sigma_a^2} \begin{pmatrix} \sum_{i=1}^m \tau_i & \sum_{i=1}^m \tau_i \bar{x}_i \\ \frac{\sigma_a^2}{\sigma^2} \sum_{i=1}^m (n_i \bar{x}_i - n_i \tau_i \bar{x}_i) & \frac{\sigma_a^2}{\sigma^2} \left( \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}^2 - \sum_{i=1}^m \tau_i \bar{x}_i n_i \bar{x}_i \right) \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\sigma_a^2} \left( \begin{array}{cc} \sum_{i=1}^m \tau_i & \sum_{i=1}^m \tau_i \bar{x}_i \\ \frac{\sigma_a^2}{\sigma^2} \sum_{i=1}^m (1 - \tau_i) n_i \bar{x}_i & \frac{\sigma_a^2}{\sigma^2} \left[ \sum_{i=1}^m \left( \sum_{j=1}^{n_i} x_{ij}^2 - n_i \bar{x}_i^2 \right) + \sum_{i=1}^m n_i \bar{x}_i^2 - \sum_{i=1}^m \tau_i n_i \bar{x}_i^2 \right] \end{array} \right) \\
&= \frac{1}{\sigma_a^2} \left( \begin{array}{cc} \sum_{i=1}^m \tau_i & \sum_{i=1}^m \tau_i \bar{x}_i \\ \sum_{i=1}^m \tau_i \bar{x}_i & \frac{\sigma_a^2}{\sigma^2} S_{xx} + \sum_{i=1}^m \tau_i \bar{x}_i^2 \end{array} \right)
\end{aligned}$$

Où, comme énoncé dans la section 2.3.4 c),  $S_{xx} = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2 = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij}^2 - \bar{x}_i^2)$ .

Trouvons maintenant une expression détaillée pour  $\mathbf{X}^t \mathbf{V}^{-1} \mathbf{y}$ .

$$\begin{aligned}
\mathbf{X}^t \mathbf{V}^{-1} \mathbf{y} &= \left\{ \begin{array}{l} \left[ \begin{array}{c} \left( \frac{\tau_i}{n_i \sigma_a^2} \right) \mathbf{1}_{n_i}^t \\ (\mathbf{x}_i^t - \tau_i \bar{x}_i \mathbf{1}_{n_i}^t) / \sigma^2 \end{array} \right] \end{array} \right\}_{i=1}^m \begin{bmatrix} y_{11} \\ \vdots \\ y_{mn_m} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij} \left( \frac{\tau_i}{n_i \sigma_a^2} \right) \\ \sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij} (x_{ij} - \tau_i \bar{x}_i) / \sigma^2 \end{bmatrix} \\
&= \frac{1}{\sigma_a^2} \begin{bmatrix} \sum_{i=1}^m \tau_i \bar{y}_i \\ \frac{\sigma_a^2}{\sigma^2} \left[ \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} y_{ij} - \sum_{i=1}^m n_i \bar{x}_i \bar{y}_i + \sum_{i=1}^m n_i \bar{x}_i \bar{y}_i - \sum_{i=1}^m \sum_{j=1}^{n_i} y_{ij} \tau_i \bar{x}_i \right] \end{bmatrix} \\
&= \frac{1}{\sigma_a^2} \begin{bmatrix} \sum_{i=1}^m \tau_i \bar{y}_i \\ \frac{\sigma_a^2}{\sigma^2} \left[ S_{xy} + \sum_{i=1}^m n_i \bar{y}_i \bar{x}_i - \sum_{i=1}^m \tau_i \bar{x}_i n_i \bar{y}_i \right] \end{bmatrix} \\
&= \frac{1}{\sigma_a^2} \begin{bmatrix} \sum_{i=1}^m \tau_i \bar{y}_i \\ \frac{\sigma_a^2}{\sigma^2} S_{xy} + \sum_{i=1}^m \tau_i \bar{x}_i \bar{y}_i \end{bmatrix}
\end{aligned}$$

Où  $S_{xy} = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i) = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} y_{ij} - \bar{x}_i \bar{y}_i)$ .

Maintenant, il ne reste qu'à inverser  $\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}$ . Pour ce faire, posons  $\mathbf{B} = \sigma_a^2\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X}$ , donc

$$\mathbf{B} = \begin{pmatrix} p & q \\ q & r \end{pmatrix}$$

Où  $p = \sum_{i=1}^m \tau_i$ ,  $q = \sum_{i=1}^m \tau_i \bar{x}_i$  et  $r = \frac{\sigma_a^2}{\sigma^2} S_{xx} + \sum_{i=1}^m \tau_i \bar{x}_i^2$ . Ensuite,

$$(\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X})^{-1} = \sigma_a^2 \mathbf{B}^{-1} = \frac{\sigma_a^2}{|\mathbf{B}|} \begin{pmatrix} r & -q \\ -q & p \end{pmatrix}$$

Avec

$$\begin{aligned} |\mathbf{B}| &= pr - q^2 = \left( \sum_{i=1}^m \tau_i \right) \left[ \frac{\sigma_a^2}{\sigma^2} S_{xx} + \sum_{i=1}^m \tau_i \bar{x}_i^2 \right] - \left( \sum_{i=1}^m \tau_i \bar{x}_i \right)^2 \\ &= \left( \sum_{i=1}^m \tau_i \right) \left[ \frac{\sigma_a^2}{\sigma^2} S_{xx} + \sum_{i=1}^m \tau_i \bar{x}_i^2 - \frac{\left( \sum_{i=1}^m \tau_i \bar{x}_i \right)^2}{\sum_{i=1}^m \tau_i} \right] \end{aligned}$$

Maintenant, trouvons  $\hat{\mu}$  et  $\hat{\beta}$ .

$$\begin{aligned} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} &= (\mathbf{X}^t\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^t\mathbf{V}^{-1}\mathbf{y} \\ &= \frac{\sigma_a^2}{|\mathbf{B}|} \begin{pmatrix} \frac{\sigma_a^2}{\sigma^2} S_{xx} + \sum_{i=1}^m \tau_i \bar{x}_i^2 & - \sum_{i=1}^m \tau_i \bar{x}_i \\ - \sum_{i=1}^m \tau_i \bar{x}_i & \sum_{i=1}^m \tau_i \end{pmatrix} \frac{1}{\sigma_a^2} \begin{bmatrix} \sum_{i=1}^m \tau_i \bar{y}_i \\ \frac{\sigma_a^2}{\sigma^2} S_{xy} + \sum_{i=1}^m \tau_i \bar{x}_i \bar{y}_i \end{bmatrix} \\ &= \frac{1}{|\mathbf{B}|} \begin{bmatrix} \frac{\sigma_a^2}{\sigma^2} S_{xx} \left( \sum_{i=1}^m \tau_i \bar{y}_i \right) + \left( \sum_{i=1}^m \tau_i \bar{x}_i^2 \right) \left( \sum_{i=1}^m \tau_i \bar{y}_i \right) - \frac{\sigma_a^2}{\sigma^2} S_{xy} \left( \sum_{i=1}^m \tau_i \bar{x}_i \right) - \left( \sum_{i=1}^m \tau_i \bar{x}_i \right) \left( \sum_{i=1}^m \tau_i \bar{x}_i \bar{y}_i \right) \\ - \left( \sum_{i=1}^m \tau_i \bar{x}_i \right) \left( \sum_{i=1}^m \tau_i \bar{y}_i \right) + \frac{\sigma_a^2}{\sigma^2} S_{xy} \left( \sum_{i=1}^m \tau_i \right) + \left( \sum_{i=1}^m \tau_i \bar{x}_i \bar{y}_i \right) \left( \sum_{i=1}^m \tau_i \right) \end{bmatrix} \end{aligned}$$

$$\begin{aligned}
\Rightarrow \hat{\beta} &= \frac{\sum_{i=1}^m \tau_i}{|\mathbf{B}|} \left[ \frac{\sigma_a^2}{\sigma^2} S_{xy} + \sum_{i=1}^m \tau_i \bar{x}_i \bar{y}_i - \frac{\left( \sum_{i=1}^m \tau_i \bar{x}_i \right) \left( \sum_{i=1}^m \tau_i \bar{y}_i \right)}{\sum_{i=1}^m \tau_i} \right] \\
&= \frac{\sum_{i=1}^m \tau_i}{\sum_{i=1}^m \tau_i} \cdot \frac{(\sigma_a^2/\sigma^2) S_{xy} + \sum_{i=1}^m \tau_i \bar{x}_i \bar{y}_i - \left( \sum_{i=1}^m \tau_i \bar{x}_i \right) \left( \sum_{i=1}^m \tau_i \bar{y}_i \right) / \left( \sum_{i=1}^m \tau_i \right)}{(\sigma_a^2/\sigma^2) S_{xx} + \sum_{i=1}^m \tau_i \bar{x}_i^2 - \left( \sum_{i=1}^m \tau_i \bar{x}_i \right)^2 / \left( \sum_{i=1}^m \tau_i \right)} \\
&= \frac{(\sigma_a^2/\sigma^2) S_{xy} + WSM_{xy}}{(\sigma_a^2/\sigma^2) S_{xx} + WSM_{xx}}
\end{aligned}$$

Car,

$$\begin{aligned}
WSM_{xy} &= \sum_{i=1}^m \tau_i (\bar{x}_i - \tilde{x}_{..}) (\bar{y}_i - \tilde{y}_{..}) \\
&= \sum_{i=1}^m \tau_i (\bar{x}_i \bar{y}_i - \tilde{x}_{..} \tilde{y}_{..}) \\
&= \sum_{i=1}^m \tau_i \bar{x}_i \bar{y}_i - \left( \sum_{i=1}^m \tau_i \right) \tilde{x}_{..} \tilde{y}_{..} \\
&= \sum_{i=1}^m \tau_i \bar{x}_i \bar{y}_i - \frac{\left( \sum_{i=1}^m \tau_i \bar{x}_i \right) \left( \sum_{i=1}^m \tau_i \bar{y}_i \right)}{\sum_{i=1}^m \tau_i}
\end{aligned}$$

$$\begin{aligned}
WSM_{xx} &= \sum_{i=1}^m \tau_i (\bar{x}_i - \tilde{x}_{..})^2 \\
&= \sum_{i=1}^m \tau_i (\bar{x}_i^2 - \tilde{x}_{..}^2)
\end{aligned}$$



$$\begin{aligned}
&= \sum_{i=1}^m \tau_i \bar{x}_i^2 - \left( \sum_{i=1}^m \tau_i \right) \tilde{x}^2 \\
&= \sum_{i=1}^m \tau_i \bar{x}_i^2 - \frac{\left( \sum_{i=1}^m \tau_i \bar{x}_i \right)^2}{\sum_{i=1}^m \tau_i}
\end{aligned}$$

Et pour  $\hat{\mu}$ ,

$$\begin{aligned}
\hat{\mu} &= \frac{1}{|\mathbf{B}|} \left[ \frac{\sigma_a^2}{\sigma^2} S_{xx} \left( \sum_{i=1}^m \tau_i \bar{y}_i \right) + \left( \sum_{i=1}^m \tau_i \bar{x}_i^2 \right) \left( \sum_{i=1}^m \tau_i \bar{y}_i \right) - \frac{\sigma_a^2}{\sigma^2} S_{xy} \left( \sum_{i=1}^m \tau_i \bar{x}_i \right) - \left( \sum_{i=1}^m \tau_i \bar{x}_i \right) \left( \sum_{i=1}^m \tau_i \bar{x}_i \bar{y}_i \right) \right] \\
&= \frac{\left( \sum_{i=1}^m \tau_i \bar{y}_i \right) \left[ (\sigma_a^2/\sigma^2) S_{xx} + \sum_{i=1}^m \tau_i \bar{x}_i^2 \right]}{\left( \sum_{i=1}^m \tau_i \right) [(\sigma_a^2/\sigma^2) S_{xx} + WSM_{xx}]} \\
&\quad - \frac{\left( \sum_{i=1}^m \tau_i \bar{x}_i \right) \left[ (\sigma_a^2/\sigma^2) S_{xy} + \sum_{i=1}^m \tau_i \bar{x}_i \bar{y}_i - \frac{\left( \sum_{i=1}^m \tau_i \bar{x}_i \right) \left( \sum_{i=1}^m \tau_i \bar{y}_i \right)}{\sum_{i=1}^m \tau_i} \right] + \frac{\left( \sum_{i=1}^m \tau_i \bar{x}_i \right)^2 \left( \sum_{i=1}^m \tau_i \bar{y}_i \right)}{\sum_{i=1}^m \tau_i}}{\left( \sum_{i=1}^m \tau_i \right) [(\sigma_a^2/\sigma^2) S_{xx} + WSM_{xx}]} \\
&= \frac{\left( \sum_{i=1}^m \tau_i \bar{y}_i \right) \left[ (\sigma_a^2/\sigma^2) S_{xx} + \sum_{i=1}^m \tau_i \bar{x}_i^2 - \frac{\left( \sum_{i=1}^m \tau_i \bar{x}_i \right)^2}{\sum_{i=1}^m \tau_i} \right]}{\left( \sum_{i=1}^m \tau_i \right) [(\sigma_a^2/\sigma^2) S_{xx} + WSM_{xx}]} - \frac{\left( \sum_{i=1}^m \tau_i \bar{x}_i \right) [(\sigma_a^2/\sigma^2) S_{xy} + WSM_{xy}]}{\left( \sum_{i=1}^m \tau_i \right) [(\sigma_a^2/\sigma^2) S_{xx} + WSM_{xx}]} \\
&= \frac{\left( \sum_{i=1}^m \tau_i \bar{y}_i \right) [(\sigma_a^2/\sigma^2) S_{xx} + WSM_{xx}]}{\left( \sum_{i=1}^m \tau_i \right) [(\sigma_a^2/\sigma^2) S_{xx} + WSM_{xx}]} - \frac{\left( \sum_{i=1}^m \tau_i \bar{x}_i \right) [(\sigma_a^2/\sigma^2) S_{xy} + WSM_{xy}]}{\left( \sum_{i=1}^m \tau_i \right) [(\sigma_a^2/\sigma^2) S_{xx} + WSM_{xx}]} \\
&= \frac{\sum_{i=1}^m \tau_i \bar{y}_i}{\sum_{i=1}^m \tau_i} - \hat{\beta} \frac{\sum_{i=1}^m \tau_i \bar{x}_i}{\sum_{i=1}^m \tau_i} \\
&= \tilde{y}^{\cdot} - \hat{\beta} \tilde{x}^{\cdot}
\end{aligned}$$

Avec,

$$\tilde{x}_{..} = \frac{\sum_{i=1}^m \tau_i \bar{x}_i}{\sum_{i=1}^m \tau_i}$$

et

$$\tilde{y}_{..} = \frac{\sum_{i=1}^m \tau_i \bar{y}_i}{\sum_{i=1}^m \tau_i} \quad \square$$

### B.7.2 Vraisemblance modifiée pour trouver $\hat{\sigma}^2$ et $\hat{\sigma}_a^2$

Voici tout d'abord la vraisemblance modifiée.

$$l^* = -\frac{N}{2} \log(2\pi) - \frac{1}{2} \log(|\mathbf{V}|) - \frac{1}{2} (\mathbf{y} - \boldsymbol{\Theta})^t \mathbf{V}^{-1} (\mathbf{y} - \boldsymbol{\Theta}) \quad (\text{B.20})$$

Où  $\boldsymbol{\Theta} = \mathbb{E}[\mathbf{y}] = \mathbf{X} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix}$ ,  $|\mathbf{V}| = (\sigma^2)^{N-m} \prod_{i=1}^m (\sigma^2 + n_i \sigma_a^2)$  et  $\mathbf{V}^{-1} = \frac{1}{\sigma^2} \{d \mathbf{I}_d - \tau_i \bar{\mathbf{J}}_i\}$

**Proposition B.36.**

$$l^* = -\frac{1}{2} \left[ N \log(2\pi) + (N - m) \log(\sigma^2) + \sum_{i=1}^m \log(\sigma^2 + n_i \sigma_a^2) + \frac{S_1^*}{\sigma^2} - \frac{W S_2^*}{\sigma^2} \right]$$

*Démonstration.* La vraisemblance modifiée devient

$$l^* = -\frac{1}{2} \left[ N \log(2\pi) + (N - m) \log(\sigma^2) + \sum_{i=1}^m \log(\sigma^2 + n_i \sigma_a^2) \right. \\ \left. + \frac{1}{\sigma^2} \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \right)^t \left\{ {}_d \mathbf{I}_i - \tau_i \bar{\mathbf{J}}_i \right\} \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \right) \right]$$

Avec,

$$\mathbf{y} - \mathbf{X} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} = \begin{bmatrix} y_{11} - \hat{\mu} - x_{11} \hat{\beta} \\ \vdots \\ y_{1n_1} - \hat{\mu} - x_{1n_1} \hat{\beta} \\ y_{21} - \hat{\mu} - x_{21} \hat{\beta} \\ \vdots \\ y_{mn_m} - \hat{\mu} - x_{mn_m} \hat{\beta} \end{bmatrix} \\ = \{ {}_c y_{ij} - \hat{\mu} - \hat{\beta} x_{ij} \}_{i=1}^m \{ {}_j=1}^{n_i}$$

Cherchons maintenant  $\star = (\mathbf{y} - \Theta)^t \{ {}_d \mathbf{I}_i - \tau_i \bar{\mathbf{J}}_i \} (\mathbf{y} - \Theta)$ ,

$$\star = \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \right)^t \left\{ {}_d \mathbf{I}_i - \tau_i \bar{\mathbf{J}}_i \right\} \left( \mathbf{y} - \mathbf{X} \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \right) \\ = \{ {}_l y_{ij} - \hat{\mu} - \hat{\beta} x_{ij} \} \{ {}_d \mathbf{I}_i - \tau_i \bar{\mathbf{J}}_i \} \{ {}_c y_{ij} - \hat{\mu} - \hat{\beta} x_{ij} \} \\ = \begin{bmatrix} y_{11} - \hat{\mu} - \hat{\beta} x_{11} \\ \vdots \\ y_{mn_m} - \hat{\mu} - \hat{\beta} x_{mn_m} \end{bmatrix}^t \begin{pmatrix} 1 - \frac{\tau_1}{n_1} & \cdots & -\frac{\tau_1}{n_1} & & & \\ \vdots & \ddots & \vdots & \cdots & & \mathbf{0}_{n_i \times n_i} \\ -\frac{\tau_1}{n_1} & \cdots & 1 - \frac{\tau_1}{n_1} & & & \\ & \vdots & & \ddots & & \vdots \\ & & & & 1 - \frac{\tau_m}{n_m} & \cdots & -\frac{\tau_m}{n_m} \\ & & \mathbf{0}_{n_i \times n_i} & \cdots & \vdots & \ddots & \vdots \\ & & & & -\frac{\tau_m}{n_m} & \cdots & 1 - \frac{\tau_m}{n_m} \end{pmatrix} \begin{bmatrix} y_{11} - \hat{\mu} - \hat{\beta} x_{11} \\ \vdots \\ y_{mn_m} - \hat{\mu} - \hat{\beta} x_{mn_m} \end{bmatrix}$$

$$\begin{aligned}
&= \left\{ \left\| y_{ij} - \hat{\mu} - \hat{\beta}x_{ij} - \frac{\tau_i}{n_i} \sum_{k=1}^{n_i} (y_{ik} - \hat{\mu} - \hat{\beta}x_{ik}) \right\|_{i=1, j=1}^{m, n_i} \right\} \begin{bmatrix} y_{11} - \hat{\mu} - \hat{\beta}x_{11} \\ \vdots \\ y_{mn_m} - \hat{\mu} - \hat{\beta}x_{mn_m} \end{bmatrix} \\
&= \left\{ \left\| y_{ij} - \hat{\mu} - \hat{\beta}x_{ij} - \tau_i(\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x}_i) \right\|_{i=1, j=1}^{m, n_i} \right\} \begin{bmatrix} y_{11} - \hat{\mu} - \hat{\beta}x_{11} \\ \vdots \\ y_{mn_m} - \hat{\mu} - \hat{\beta}x_{mn_m} \end{bmatrix} \\
&= \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu} - \hat{\beta}x_{ij})^2 - \sum_{i=1}^m \sum_{j=1}^{n_i} \tau_i (y_{ij} - \hat{\mu} - \hat{\beta}x_{ij})(\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x}_i) \\
&= S_1^* - \sum_{i=1}^m \tau_i (\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x}_i) \left( \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu} - \hat{\beta}x_{ij}) \right) \\
&= S_1^* - \sum_{i=1}^m \tau_i n_i (\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x}_i)^2 \\
&= S_1^* - WS_2^*
\end{aligned}$$

Car  $S_1^* = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \hat{\mu} - \hat{\beta}x_{ij})^2$  et  $WS_2^* = \sum_{i=1}^m \tau_i n_i (\bar{y}_i - \hat{\mu} - \hat{\beta}\bar{x}_i)^2$ .

Remettons cette expression dans la vraisemblance modifiée.

$$l^* = -\frac{1}{2} \left[ N \log(2\pi) + (N - m) \log(\sigma^2) + \sum_{i=1}^m \log(\sigma^2 + n_i \sigma_a^2) + \frac{S_1^*}{\sigma^2} - \frac{WS_2^*}{\sigma^2} \right] \quad \square$$

### B.7.3 Espérance et matrice de covariance de $\hat{\mu}$ et $\hat{\beta}$

Tout découle des deux propositions suivantes.

**Proposition B.37.**  $\mathbb{E}[S_{xy}] = \beta S_{xx}$

*Démonstration.*

$$\begin{aligned}
\mathbb{E}[S_{xy}] &= \mathbb{E} \left[ \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})(y_{ij} - \bar{y}_{i.}) \right] \\
&= \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.}) \mathbb{E}[y_{ij} - \bar{y}_{i.}] \\
&= \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.}) (\mu + \beta x_{ij} - \mu - \beta \bar{x}_{i.}) \\
&= \beta \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})^2 \\
&= \beta S_{xx}
\end{aligned}$$

□

**Proposition B.38.**  $\mathbb{E}[WSM_{xy}] = \beta WSM_{xx}$

*Démonstration.*

$$\begin{aligned}
\mathbb{E}[WSM_{xy}] &= \mathbb{E} \left[ \sum_{i=1}^m \tau_i (\bar{x}_{i.} - \tilde{x}_{..})(\bar{y}_{i.} - \tilde{y}_{..}) \right] \\
&= \sum_{i=1}^m \tau_i (\bar{x}_{i.} - \tilde{x}_{..}) \mathbb{E}[\bar{y}_{i.} - \tilde{y}_{..}] \\
&= \sum_{i=1}^m \tau_i (\bar{x}_{i.} - \tilde{x}_{..}) \left( \mu + \beta \bar{x}_{i.} - \frac{\sum_{k=1}^m \tau_k (\mu + \beta \bar{x}_{k.})}{\sum_{k=1}^m \tau_k} \right) \\
&= \beta \sum_{i=1}^m \tau_i (\bar{x}_{i.} - \tilde{x}_{..})(\bar{x}_{i.} - \tilde{x}_{..}) \\
&= \beta WSM_{xx}
\end{aligned}$$

□

Maintenant, pour la matrice de variance-covariance, procédons comme dans le cas où les données sont équilibrées, c'est-à-dire que  $\text{Var} \left( \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \right) = (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1}$ .

**Proposition B.39.**

$$\begin{aligned} \text{Var}(\hat{\beta}) &= \frac{1}{(S_{xx}/\sigma^2) + (WSM_{xx}/\sigma_a^2)} \\ \text{Var}(\hat{\mu}) &= \frac{\sigma_a^2}{m} + \tilde{x}_{..}^2 \text{Var}(\hat{\beta}) \\ \text{cov}(\hat{\mu}, \hat{\beta}) &= -\tilde{x}_{..} \text{Var}(\hat{\beta}) \end{aligned}$$

*Démonstration.* Reprenons la formule trouvée plus haut pour cette expression.

$$\begin{aligned} \text{Var} \left( \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \right) &= (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} \\ &= \frac{\sigma_a^2}{|\mathbf{B}|} \begin{pmatrix} \frac{\sigma_a^2}{\sigma^2} S_{xx} + \sum_{i=1}^m \tau_i \bar{x}_i^2 & - \sum_{i=1}^m \tau_i \bar{x}_i \\ - \sum_{i=1}^m \tau_i \bar{x}_i & \sum_{i=1}^m \tau_i \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \Rightarrow \text{Var}(\hat{\beta}) &= \frac{\sigma_a^2}{|\mathbf{B}|} \sum_{i=1}^m \tau_i \\ &= \frac{\sigma_a^2 \sum_{i=1}^m \tau_i}{\left( \sum_{i=1}^m \tau_i \right) [(\sigma_a^2/\sigma^2) S_{xx} + WSM_{xx}]} \\ &= \frac{\sigma_a^2}{(\sigma_a^2/\sigma^2) S_{xx} + WSM_{xx}} \\ &= \frac{1}{(S_{xx}/\sigma^2) + (WSM_{xx}/\sigma_a^2)} \end{aligned}$$

Et,

$$\begin{aligned}
\Rightarrow \text{Var}(\hat{\mu}) &= \frac{\sigma_a^2}{|\mathbf{B}|} \left[ (\sigma_a^2/\sigma^2)S_{xx} + \sum_{i=1}^m \tau_i \bar{x}_i^2 \right] \\
&= \frac{\sigma_a^2 \left[ (\sigma_a^2/\sigma^2)S_{xx} + \sum_{i=1}^m \tau_i \bar{x}_i^2 \right]}{\left( \sum_{i=1}^m \tau_i \right) [(\sigma_a^2/\sigma^2)S_{xx} + WSM_{xx}]} \\
&= \frac{\sigma_a^2}{\sum_{i=1}^m \tau_i} \cdot \frac{(\sigma_a^2/\sigma^2)S_{xx} + \sum_{i=1}^m \tau_i \bar{x}_i^2 - \frac{\left( \sum_{i=1}^m \tau_i \bar{x}_i \right)^2}{\sum_{i=1}^m \tau_i} + \frac{\left( \sum_{i=1}^m \tau_i \bar{x}_i \right)^2}{\sum_{i=1}^m \tau_i}}{(\sigma_a^2/\sigma^2)S_{xx} + WSM_{xx}} \\
&= \frac{\sigma_a^2}{\sum_{i=1}^m \tau_i} \left[ 1 + \frac{\tilde{x}_{..}^2 \left( \sum_{i=1}^m \tau_i \right)}{(\sigma_a^2/\sigma^2)S_{xx} + WSM_{xx}} \right] \\
&= \frac{\sigma_a^2}{\sum_{i=1}^m \tau_i} + \tilde{x}_{..}^2 \text{Var}(\hat{\beta})
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \text{cov}(\hat{\mu}, \hat{\beta}) &= \frac{\sigma_a^2}{|\mathbf{B}|} \left( - \sum_{i=1}^m \tau_i \bar{x}_i \right) \\
&= \frac{-\sigma_a^2 \left( \sum_{i=1}^m \tau_i \bar{x}_i \right)}{\left( \sum_{i=1}^m \tau_i \right) [(\sigma_a^2/\sigma^2)S_{xx} + WSM_{xx}]} \\
&= -\tilde{x}_{..} \text{Var}(\hat{\beta}) \quad \square
\end{aligned}$$

## B.7.4 Loi asymptotique des estimateurs du maximum de vraisemblance

La log-vraisemblance non modifiée est donnée par

$$l = -\frac{N}{2} \log(2\pi) - \frac{N-m}{2} \log(\sigma^2) - \frac{1}{2} \sum_{i=1}^m \log(\sigma^2 + n_i \sigma_a^2) - \frac{S_1}{2\sigma^2} + \frac{W S_2}{2\sigma^2} \quad (\text{B.21})$$

$$\text{Où } S_1 = \sum_{i=1}^m \sum_{j=1}^{n_i} (y_{ij} - \mu - \beta x_{ij})^2 \text{ et } S_2 = \sum_{i=1}^m \frac{n_i^2 \sigma_a^2}{\sigma^2 + n_i \sigma_a^2} (\bar{y}_i - \mu - \beta \bar{x}_i)^2.$$

**Proposition B.40.**

$$\begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \approx \mathcal{N} \left( \begin{bmatrix} \mu \\ \beta \end{bmatrix}, \frac{1}{D_2 \sigma^2} \begin{pmatrix} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}^2 - \sum_{i=1}^m n_i \tau_i \bar{x}_i^2 & \sum_{i=1}^m n_i \tau_i \bar{x}_i - N \bar{x}_{..} \\ \sum_{i=1}^m n_i \tau_i \bar{x}_i - N \bar{x}_{..} & N - \sum_{i=1}^m n_i \tau_i \end{pmatrix} \right)$$

Et ils sont indépendants de

$$\begin{bmatrix} \hat{\sigma}^2 \\ \hat{\sigma}_a^2 \end{bmatrix} \approx \mathcal{N} \left( \begin{bmatrix} \sigma^2 \\ \sigma_a^2 \end{bmatrix}, \frac{1}{2D_1} \begin{pmatrix} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} & - \sum_{i=1}^m \frac{n_i}{\lambda_i^2} \\ - \sum_{i=1}^m \frac{n_i}{\lambda_i^2} & \frac{N-m}{\sigma^4} + \sum_{i=1}^m \frac{1}{\lambda_i^2} \end{pmatrix} \right)$$

$$\text{Où } \lambda_i = \sigma^2 + n_i \sigma_a^2, \tau_i = \frac{n_i \sigma_a^2}{\sigma^2 + n_i \sigma_a^2}, D_1 = \frac{N-m}{4\sigma^4} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} + \frac{1}{4} \left( \sum_{i=1}^m \frac{1}{\lambda_i} \right) \left( \sum_{i=1}^m \frac{n_i^2}{\lambda_i} \right) - \frac{1}{4} \left( \sum_{i=1}^m \frac{n_i}{\lambda_i} \right) \left( \sum_{i=1}^m \frac{n_i}{\lambda_i} \right)$$

$$\text{et } D_2 = \left( \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}^2 - \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \bar{x}_i^2 \right) \left( \frac{N}{\sigma^2} - \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \right) - \left( \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \bar{x}_i - \frac{N \bar{x}_{..}}{\sigma^2} \right)^2.$$

*Démonstration.* Les dérivées partielles sont données par

$$\begin{aligned} \frac{\partial l}{\partial \mu} &= \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} 2(y_{ij} - \mu - \beta x_{ij}) - \frac{1}{2\sigma^2} \sum_{i=1}^m \frac{2n_i^2 \sigma_a^2}{\sigma^2 + n_i \sigma_a^2} (\bar{y}_i - \mu - \beta \bar{x}_i) \\ &= \frac{N}{\sigma^2} (\bar{y}_{..} - \mu - \beta \bar{x}_{..}) - \sum_{i=1}^m \frac{n_i^2 \sigma_a^2}{\sigma^2 (\sigma^2 + n_i \sigma_a^2)} (\bar{y}_i - \mu - \beta \bar{x}_i) \end{aligned}$$



$$\begin{aligned}
\frac{\partial l}{\partial \beta} &= \frac{1}{2\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} 2x_{ij}(y_{ij} - \mu - \beta x_{ij}) - \frac{1}{2\sigma^2} \sum_{i=1}^m \frac{2n_i^2 \sigma_a^2}{\sigma^2 + n_i \sigma_a^2} \bar{x}_i (\bar{y}_i - \mu - \beta \bar{x}_i) \\
&= \frac{1}{\sigma^2} \left( \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} y_{ij}) - N \mu \bar{x}_{..} - \beta \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij}^2) \right) - \sum_{i=1}^m \frac{n_i^2 \sigma_a^2 \bar{x}_i}{\sigma^2 (\sigma^2 + n_i \sigma_a^2)} (\bar{y}_i - \mu - \beta \bar{x}_i)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial l}{\partial \sigma^2} &= -\frac{N-m}{2\sigma^2} - \frac{1}{2} \sum_{i=1}^m \frac{1}{\sigma^2 + n_i \sigma_a^2} + \frac{S_1}{2\sigma^4} - \frac{1}{2} \sum_{i=1}^m \frac{n_i^2 \sigma_a^2 (2\sigma^2 + n_i \sigma_a^2)}{(\sigma^2 (\sigma^2 + n_i \sigma_a^2))^2} (\bar{y}_i - \mu - \beta \bar{x}_i)^2 \\
&= -\frac{N-m}{2\sigma^2} - \frac{1}{2} \sum_{i=1}^m \frac{1}{\sigma^2 + n_i \sigma_a^2} + \frac{S_1}{2\sigma^4} - \frac{1}{2} \sum_{i=1}^m n_i \left( \frac{1}{\sigma^4} - \frac{1}{(\sigma^2 + n_i \sigma_a^2)^2} \right) (\bar{y}_i - \mu - \beta \bar{x}_i)^2
\end{aligned}$$

$$\begin{aligned}
\frac{\partial l}{\partial \sigma_a^2} &= -\frac{1}{2} \sum_{i=1}^m \frac{n_i}{\sigma^2 + n_i \sigma_a^2} + \frac{1}{2\sigma^2} \sum_{i=1}^m \frac{n_i^2 (\sigma^2 + n_i \sigma_a^2) - n_i^2 \sigma_a^2 n_i}{(\sigma^2 + n_i \sigma_a^2)^2} (\bar{y}_i - \mu - \beta \bar{x}_i)^2 \\
&= -\frac{1}{2} \sum_{i=1}^m \frac{n_i}{\sigma^2 + n_i \sigma_a^2} + \frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{(\sigma^2 + n_i \sigma_a^2)^2} (\bar{y}_i - \mu - \beta \bar{x}_i)^2
\end{aligned}$$

Et donc les dérivées partielles secondes sont données par

$$\begin{aligned}
\frac{\partial^2 l}{\partial \mu \partial \mu} &= -\frac{N}{\sigma^2} + \sum_{i=1}^m \frac{n_i^2 \sigma_a^2}{\sigma^2 (\sigma^2 + n_i \sigma_a^2)} \\
&= -\frac{N}{\sigma^2} + \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l}{\partial \mu \partial \beta} &= -\frac{N \bar{x}_{..}}{\sigma^2} + \sum_{i=1}^m \frac{n_i^2 \sigma_a^2 \bar{x}_i}{\sigma^2 (\sigma^2 + n_i \sigma_a^2)} \\
&= -\frac{N \bar{x}_{..}}{\sigma^2} + \sum_{i=1}^m \frac{n_i \tau_i \bar{x}_i}{\sigma^2}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l}{\partial \mu \partial \sigma^2} &= -\frac{N}{\sigma^4}(\bar{y}_{..} - \mu - \beta \bar{x}_{..}) + \sum_{i=1}^m \frac{n_i^2 \sigma_a^2 (2\sigma^2 + n_i \sigma_a^2)}{(\sigma^2(\sigma^2 + n_i \sigma_a^2))^2} (\bar{y}_{i.} - \mu - \beta \bar{x}_{i.}) \\
&= -\frac{N}{\sigma^4}(\bar{y}_{..} - \mu - \beta \bar{x}_{..}) + \sum_{i=1}^m \frac{n_i}{\sigma^4} (\bar{y}_{i.} - \mu - \beta \bar{x}_{i.}) - \sum_{i=1}^m \frac{n_i}{(\sigma^2 + n_i \sigma_a^2)^2} (\bar{y}_{i.} - \mu - \beta \bar{x}_{i.}) \\
&= -\sum_{i=1}^m \frac{n_i}{(\sigma^2 + n_i \sigma_a^2)^2} (\bar{y}_{i.} - \mu - \beta \bar{x}_{i.})
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l}{\partial \mu \partial \sigma_a^2} &= -\frac{1}{\sigma^2} \sum_{i=1}^m \frac{n_i^2 (\sigma^2 + n_i \sigma_a^2) - n_i^2 \sigma_a^2 n_i}{(\sigma^2 + n_i \sigma_a^2)^2} (\bar{y}_{i.} - \mu - \beta \bar{x}_{i.}) \\
&= -\sum_{i=1}^m \frac{n_i^2}{(\sigma^2 + n_i \sigma_a^2)^2} (\bar{y}_{i.} - \mu - \beta \bar{x}_{i.})
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l}{\partial \beta \partial \beta} &= -\frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}^2 + \sum_{i=1}^m \frac{n_i^2 \sigma_a^2 \bar{x}_{i.}^2}{\sigma^2 (\sigma^2 + n_i \sigma_a^2)} \\
&= \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}^2 + \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \bar{x}_{i.}^2
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l}{\partial \beta \partial \sigma^2} &= -\frac{1}{\sigma^4} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} (y_{ij} - \mu - \beta x_{ij}) + \sum_{i=1}^m \frac{n_i^2 \sigma_a^2 \bar{x}_{i.} (2\sigma^2 + n_i \sigma_a^2)}{(\sigma^2(\sigma^2 + n_i \sigma_a^2))^2} (\bar{y}_{i.} - \mu - \beta \bar{x}_{i.}) \\
&= -\frac{1}{\sigma^4} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij} (y_{ij} - \mu - \beta x_{ij}) + \sum_{i=1}^m \frac{n_i \bar{x}_{i.}}{\sigma^4} (\bar{y}_{i.} - \mu - \beta \bar{x}_{i.}) \\
&\quad - \sum_{i=1}^m \frac{n_i \bar{x}_{i.}}{(\sigma^2 + n_i \sigma_a^2)^2} (\bar{y}_{i.} - \mu - \beta \bar{x}_{i.})
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 l}{\partial \beta \partial \sigma_a^2} &= -\sum_{i=1}^m \frac{(n_i^2 (\sigma^2 + n_i \sigma_a^2) - n_i^2 \sigma_a^2 n_i) \bar{x}_{i.}}{\sigma^2 (\sigma^2 + n_i \sigma_a^2)^2} (\bar{y}_{i.} - \mu - \beta \bar{x}_{i.}) \\
&= -\sum_{i=1}^m \frac{n_i^2 \bar{x}_{i.}}{(\sigma^2 + n_i \sigma_a^2)^2} (\bar{y}_{i.} - \mu - \beta \bar{x}_{i.})
\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 l}{\partial \sigma^2 \partial \sigma^2} &= \frac{N-m}{2\sigma^4} + \frac{1}{2} \sum_{i=1}^m \frac{1}{(\sigma^2 + n_i \sigma_a^2)^2} - \frac{S_1}{\sigma^6} + \sum_{i=1}^m \frac{n_i}{\sigma^6} (\bar{y}_i - \mu - \bar{x}_i)^2 \\ &\quad - \sum_{i=1}^m \frac{n_i}{(\sigma^2 + n_i \sigma_a^2)^3} (\bar{y}_i - \mu - \bar{x}_i)^2\end{aligned}$$

$$\frac{\partial^2 l}{\partial \sigma^2 \partial \sigma_a^2} = \frac{1}{2} \sum_{i=1}^m \frac{n_i}{(\sigma^2 + n_i \sigma_a^2)^2} - \sum_{i=1}^m \frac{n_i^2}{(\sigma^2 + n_i \sigma_a^2)^3} (\bar{y}_i - \mu - \bar{x}_i)^2$$

$$\frac{\partial^2 l}{\partial \sigma_a^2 \partial \sigma_a^2} = \frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{(\sigma^2 + n_i \sigma_a^2)^2} - \sum_{i=1}^m \frac{n_i^3}{(\sigma^2 + n_i \sigma_a^2)^3} (\bar{y}_i - \mu - \bar{x}_i)^2$$

Donc, la négation de l'espérance des expressions précédentes est donnée par

$$-\mathbb{E}[l_{\mu\mu}] = \frac{N}{\sigma^2} - \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2}$$

$$-\mathbb{E}[l_{\beta\beta}] = \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}^2 - \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \bar{x}_i^2$$

$$-\mathbb{E}[l_{\mu\beta}] = \frac{N \bar{x}_{..}}{\sigma^2} - \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \bar{x}_i$$

$$\begin{aligned}-\mathbb{E}[l_{\sigma^2 \sigma^2}] &= -\frac{N-m}{2\sigma^4} - \sum_{i=1}^m \frac{1}{2\lambda_i^2} + \frac{N(\sigma_a^2 + \sigma^2)}{\sigma^6} - \sum_{i=1}^m \frac{n_i(\sigma^2/n_i + \sigma_a^2)}{\sigma^6} + \sum_{i=1}^m \frac{n_i(\sigma^2/n_i + \sigma_a^2)}{(\sigma^2 + n_i \sigma_a^2)^3} \\ &= -\frac{N-m}{2\sigma^4} - \sum_{i=1}^m \frac{1}{2\lambda_i^2} + \frac{N-m}{\sigma^4} + \sum_{i=1}^m \frac{1}{\lambda_i^2} \\ &= \frac{1}{2} \sum_{i=1}^m \frac{1}{\lambda_i^2} + \frac{N-m}{2\sigma^4}\end{aligned}$$

$$\begin{aligned}
-\mathbb{E}[l_{\sigma^2\sigma_a^2}] &= -\sum_{i=1}^m \frac{n_i}{2\lambda_i^2} + \sum_{i=1}^m \frac{n_i^2(\sigma^2/n_i + \sigma_a^2)}{\lambda_i^3} \\
&= -\sum_{i=1}^m \frac{n_i}{2\lambda_i^2} + \sum_{i=1}^m \frac{n_i}{\lambda_i^2} \\
&= \frac{1}{2} \sum_{i=1}^m \frac{n_i}{\lambda_i^2}
\end{aligned}$$

$$\begin{aligned}
-\mathbb{E}[l_{\sigma_a^2\sigma_a^2}] &= -\sum_{i=1}^m \frac{n_i^2}{2\lambda_i^2} + \sum_{i=1}^m \frac{n_i^3(\sigma^2/n_i + \sigma_a^2)}{\lambda_i^3} \\
&= -\sum_{i=1}^m \frac{n_i^2}{2\lambda_i^2} + \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2} \\
&= \frac{1}{2} \sum_{i=1}^m \frac{n_i^2}{\lambda_i^2}
\end{aligned}$$

$$-\mathbb{E}[l_{\mu\sigma^2}] = 0$$

$$-\mathbb{E}[l_{\mu\sigma_a^2}] = 0$$

$$-\mathbb{E}[l_{\beta\sigma^2}] = 0$$

$$-\mathbb{E}[l_{\beta\sigma_a^2}] = 0$$

Car,

$$\begin{aligned}
\mathbb{E}[S_1] &= \sum_{i=1}^m \sum_{j=1}^{n_i} \mathbb{E}[y_{ij}^2] - 2(\mu + \beta x_{ij})\mathbb{E}[y_{ij}] + (\mu + \beta x_{ij})^2 \\
&= \sum_{i=1}^m \sum_{j=1}^{n_i} \sigma_a^2 + \sigma^2 + (\mu + \beta x_{ij})^2 - 2(\mu + \beta x_{ij})^2 + (\mu + \beta x_{ij})^2 \\
&= N(\sigma_a^2 + \sigma^2)
\end{aligned}$$

et,

$$\begin{aligned}
\mathbb{E}[(\bar{y}_i - \mu - \beta \bar{x}_i)^2] &= (\mu + \beta \bar{x}_i)^2 + \frac{\sigma^2}{n_i} + \sigma_a^2 - 2(\mu + \beta \bar{x}_i)^2 + (\mu + \beta \bar{x}_i)^2 \\
&= \frac{\sigma^2}{n_i} + \sigma_a^2
\end{aligned}$$

Donc,

$$\begin{aligned}
\text{Var} \left( \begin{bmatrix} \hat{\sigma}^2 \\ \hat{\sigma}_a^2 \end{bmatrix} \right) &\approx \begin{pmatrix} \frac{N-m}{2\sigma^4} + \sum_{i=1}^m \frac{1}{2\lambda_i^2} & \sum_{i=1}^m \frac{n_i}{2\lambda_i^2} \\ \sum_{i=1}^m \frac{n_i}{2\lambda_i^2} & \sum_{i=1}^m \frac{n_i^2}{2\lambda_i^2} \end{pmatrix}^{-1} \\
&\approx \frac{1}{D_1} \begin{pmatrix} \sum_{i=1}^m \frac{n_i^2}{2\lambda_i^2} & - \sum_{i=1}^m \frac{n_i}{2\lambda_i^2} \\ - \sum_{i=1}^m \frac{n_i}{2\lambda_i^2} & \frac{N-m}{2\sigma^4} + \sum_{i=1}^m \frac{1}{2\lambda_i^2} \end{pmatrix}
\end{aligned}$$

Où  $D_1$  est le déterminant de la matrice de la première ligne. Aussi,

$$\text{Var} \left( \begin{bmatrix} \hat{\mu} \\ \hat{\beta} \end{bmatrix} \right) \approx \begin{pmatrix} \frac{N}{\sigma^2} - \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} & \frac{N\bar{x}_{..}}{\sigma^2} - \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \bar{x}_i \\ \frac{N\bar{x}_{..}}{\sigma^2} - \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \bar{x}_i & \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}^2 - \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \bar{x}_i^2 \end{pmatrix}^{-1}$$

$$\approx \frac{1}{D_2} \begin{pmatrix} \frac{1}{\sigma^2} \sum_{i=1}^m \sum_{j=1}^{n_i} x_{ij}^2 - \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \bar{x}_i^2 & -\frac{N\bar{x}_{..}}{\sigma^2} + \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \bar{x}_i \\ -\frac{N\bar{x}_{..}}{\sigma^2} + \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \bar{x}_i & \frac{N}{\sigma^2} - \sum_{i=1}^m \frac{n_i \tau_i}{\sigma^2} \end{pmatrix}$$

Où  $D_2$  est le déterminant de la matrice de la première ligne. Avec tout cela, le résultat souhaité est démontré.  $\square$

### B.7.5 Loi de $F = \frac{SCC/1}{SCRR/N-m-1}$ sous $H_0 : \beta = 0$

Commençons par montrer que  $\frac{SCRR}{\sigma^2} \sim \chi_{(N-m-1)}^2$ .

#### a) Loi de $\frac{SCRR}{\sigma^2}$

Montrons tout d'abord que

#### Proposition B.41.

$$\frac{SCRR}{\sigma^2} = \mathbf{y}^t \mathbf{A}_5 \mathbf{y}$$

avec  $\mathbf{A}_5 = \frac{1}{\sigma^2} \left( \mathbf{I}_N - \left\{ \frac{1}{n_i} \mathbf{J}_{n_i} \right\} - \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \right)$  et  $\mathbf{A}_4 = [x_{11} - \bar{x}_{1.}, \dots, x_{mn_m} - \bar{x}_{m.}] = \mathbf{x}^t - \left\{ l \mathbf{1}_{n_i}^t \bar{x}_i \right\}_{i=1}^m$ .

*Démonstration.* Premièrement, trouvons la forme de  $\mathbf{A}_4^t \mathbf{A}_4$ . Par les calculs de la section B.5.2, il en découle que

$$\mathbf{A}_4^t \mathbf{A}_4 = \begin{pmatrix} (x_{11} - \bar{x}_{1.})^2 & \cdots & (x_{11} - \bar{x}_{1.})(x_{mn_m} - \bar{x}_{m.}) \\ \vdots & \ddots & \vdots \\ (x_{11} - \bar{x}_{1.})(x_{mn_m} - \bar{x}_{m.}) & \cdots & (x_{mn_m} - \bar{x}_{m.})^2 \end{pmatrix}$$

Ainsi, (encore grâce aux calculs de la section B.5.2)

$$\mathbf{y}^t \mathbf{A}_5 \mathbf{y} = \frac{SCRR}{\sigma^2}$$

□

Donc il faut montrer que

**Proposition B.42.**  $\mathbf{A}_5 \mathbf{V}$  est idempotente. Où  $\mathbf{V} = \{ {}_d \sigma^2 \mathbf{I}_i + \sigma_a^2 \mathbf{J}_i \}$ .

*Démonstration.*

$$\begin{aligned} \mathbf{A}_5 \mathbf{V} &= \frac{1}{\sigma^2} \left( \mathbf{I}_N - \left\{ {}_d \frac{1}{n_i} \mathbf{J}_i \right\} - \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \right) \{ {}_d \sigma^2 \mathbf{I}_i + \sigma_a^2 \mathbf{J}_i \} \\ &= \frac{1}{\sigma^2} \left( \{ {}_d \sigma^2 \mathbf{I}_i + \sigma_a^2 \mathbf{J}_i \} - \left\{ {}_d \frac{\sigma^2}{n_i} \mathbf{J}_i + \sigma_a^2 \mathbf{J}_i \right\} - \frac{\sigma^2 \mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} - \mathbf{0}_{N \times N} \right) \\ &= \frac{1}{\sigma^2} \left( \sigma^2 \mathbf{I}_N + \{ {}_d \sigma_a^2 \mathbf{J}_i \} - \left\{ {}_d \left( \frac{\sigma^2}{n_i} + \sigma_a^2 \right) \mathbf{J}_i \right\} - \frac{\sigma^2 \mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \right) \\ &= \mathbf{I}_N - \left\{ {}_d \frac{1}{n_i} \mathbf{J}_i \right\} - \frac{\mathbf{A}_4^t \mathbf{A}_4}{S_{xx}} \\ &= \mathbf{A}_5 \sigma^2 \end{aligned}$$

Et  $\sigma^2 \mathbf{A}_5 \sigma^2 \mathbf{A}_5 = \sigma^2 \mathbf{A}_5$  donc  $\mathbf{A}_5 \mathbf{V}$  est idempotente (voir section B.5.2).

□

Donc,

$$\frac{SCRR}{\sigma^2} \sim \chi^{2'} \left( \text{rang}(\mathbf{A}_5), \frac{1}{2} (\mu \mathbf{1}_N + \beta \mathbf{x})^t \mathbf{A}_5 (\mu \mathbf{1}_N + \beta \mathbf{x}) \right)$$

Il est facile de montrer que  $(\mu \mathbf{1}_N)^t \mathbf{A}_5 (\mu \mathbf{1}_N) = 0$ . Et pour  $\mathbf{x}^t \mathbf{A}_5 \mathbf{x} = 0$  voir les calculs de la section B.5.2.

Donc  $\frac{1}{2} (\mu \mathbf{1}_N + \beta \mathbf{x})^t \mathbf{A}_5 (\mu \mathbf{1}_N + \beta \mathbf{x}) = 0$ . Pour le rang de  $\mathbf{A}_5$ , il suffit d'utiliser la proposition qui dit que le rang d'une matrice idempotente est égal à sa trace. Donc  $\text{rang}(\mathbf{A}_5) = N - m - 1$  (voir encore les calculs de la section B.5.2).

Donc finalement,

$$\frac{SCRR}{\sigma^2} \sim \chi_{(N-m-1)}^2$$

b) **Loi de  $\frac{SCC}{\sigma^2}$  sous  $H_0 : \beta = 0$**

Pour la loi de  $SCC$ , ramenons-nous sous  $H_0 : \beta = 0$ . Ainsi,

$$\mathbf{y} \sim \mathcal{N}(\mu \mathbf{1}_N, \mathbf{V}) \tag{B.22}$$

Il suffit ensuite de montrer que

**Proposition B.43.**

$$z = S_{xy} = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})(y_{ij} - \bar{y}_{i.}) = \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_{i.})y_{ij} \sim \mathcal{N}(0, \sigma^2 S_{xx})$$

*Démonstration.*  $z = \mathbf{A}_4 \mathbf{y}$  avec  $\mathbf{A}_4 = [(x_{11} - \bar{x}_{1.}), \dots, (x_{mn_m} - \bar{x}_{m.})]$ . Ainsi,

$$\begin{aligned} z &\sim \mathcal{N}(\mathbf{A}_4 \mu \mathbf{1}_N, \mathbf{A}_4 \mathbf{V} \mathbf{A}_4^t) \\ &\sim \mathcal{N}(0, \mathbf{A}_4 \{ {}_d \sigma^2 \mathbf{I}_i + \sigma_a^2 \mathbf{J}_i \} \mathbf{A}_4^t) \\ &\sim \mathcal{N}(0, \sigma^2 S_{xx} + \sigma_a^2 \mathbf{A}_4 \{ {}_d \mathbf{J}_i \} \mathbf{A}_4^t) \\ &\sim \mathcal{N}(0, \sigma^2 S_{xx}) \end{aligned} \quad \square$$

Et donc,



$$\begin{aligned}
z \sim \mathcal{N}(0, \sigma^2 S_{xx}) &\Rightarrow \frac{z - 0}{\sqrt{S_{xx} \sigma^2}} \sim \mathcal{N}(0, 1) \\
&\Rightarrow \left( \frac{z}{\sqrt{S_{xx} \sigma^2}} \right)^2 \sim \chi_{(1)}^2 \\
&\Rightarrow \frac{z^2}{S_{xx} \sigma^2} \sim \chi_{(1)}^2 \\
&\Rightarrow \frac{\left[ \sum_{i=1}^m \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)(y_{ij} - \bar{y}_i) \right]^2}{S_{xx} \sigma^2} \sim \chi_{(1)}^2 \\
&\Rightarrow \frac{SCC}{\sigma^2} \sim \chi_{(1)}^2
\end{aligned} \tag{B.23}$$

c) **Indépendance de  $SCC$  et  $SCR$  sous  $H_0 : \beta = 0$**

Il faut ensuite montrer l'indépendance entre  $SCC$  et  $SCR$ .

Posons  $\mathbf{A}_6 = \frac{\mathbf{A}_4^t \mathbf{A}_4}{\sigma^2 S_{xx}}$ , donc  $SCC = \mathbf{y}^t \mathbf{A}_6 \mathbf{y}$  et  $SCR = \mathbf{y}^t \mathbf{A}_5 \mathbf{y}$ . Montrons que

**Proposition B.44.**  $\mathbf{A}_5 \mathbf{V} \mathbf{A}_6 = \mathbf{0}_{N \times N}$

*Démonstration.*

$$\begin{aligned}
\mathbf{A}_5 \mathbf{V} \mathbf{A}_6 &= \mathbf{A}_5 \sigma^2 \mathbf{A}_6 \\
&= \mathbf{0}_{N \times N}
\end{aligned}$$

Par les calculs faits dans la preuve de la proposition B.42 et les calculs faits dans la section B.5.2. □

Ce qui démontre la prochaine proposition.

**Proposition B.45.**  $\frac{SCC/1}{SCR/(N-m-1)} \sim \mathcal{F}_{N-m-1}^1$  sous  $H_0 : \beta = 0$

# ANNEXE C

## Tableau des lois des sommes de carrés

Voici deux tableaux résumés des lois des sommes de carrés dans différentes situations.

ANOVA fixe	ANOVA aléatoire
$\frac{SCR}{\sigma^2} \sim \chi^2_{(N-m)}$ Sous $H_0 : \alpha_i = 0$ $\frac{SCE}{\sigma^2} \sim \chi^2_{(m-1)}$	$\frac{SCR}{\sigma^2} \sim \chi^2_{(N-m)}$ Sous $H_0 : \sigma_a^2 = 0$ $\frac{SCE}{\sigma^2} \sim \chi^2_{(m-1)}$ Données équilibrées : $\frac{SCE}{\sigma^2 + n\sigma_a^2} \sim \chi^2_{(m-1)}$
Sous $H_0 : \alpha_i = 0$ ind. $SCE$ et $SCR$	Sous $H_0 : \sigma_a^2 = 0$ ind. $SCE$ et $SCR$ Données équilibrées : ind. $SCE$ et $SCR$

Tableau C.1 – Tableau comparatif des lois des sommes de carrés pour l'ANOVA

ANCOVA fixe	ANCOVA aléatoire
$\frac{SCRR}{\sigma^2} \sim \chi^2_{(N-m-1)}$ Sous $H_0 : \beta = 0$ $\frac{SCC}{\sigma^2} \sim \chi^2_{(1)}$ Sous $H_0 : \alpha_i = 0$ et données équilibrées : $\frac{SCE}{\sigma^2} \sim \chi^2_{(m-1)}$	$\frac{SCRR}{\sigma^2} \sim \chi^2_{(N-m-1)}$ Sous $H_0 : \beta = 0$ $\frac{SCC}{\sigma^2} \sim \chi^2_{(1)}$ Sous $H_0 : \sigma_a^2 = 0$ et données équilibrées : $\frac{SCE}{\sigma^2} \sim \chi^2_{(m-1)}$
Sous $H_0 : \beta = 0$ ind. $SCC$ et $SCRR$ Sous $H_0 : \alpha_i = 0$ ind. $SCE$ et $SCRR$	Sous $H_0 : \beta = 0$ ind. $SCC$ et $SCRR$ Sous $H_0 : \sigma_a^2 = 0$ ind. $SCE$ et $SCRR$

Tableau C.2 – Tableau comparatif des lois des sommes de carrés pour l'ANCOVA