Université de Lausanne
Ecole des Hautes Etudes Commerciales

Extentions en théorie de la crédibilité à classification croisée

Thèse

présentée à l'École des Hautes Études Commerciales de l'Université de Lausanne

par

Vincent Goulet

Bachelier ès sciences (actuariat) de
l'Université Laval, Québec, Canada

Maître ès sciences (mathématiques) de
l'Université Laval, Québec, Canada

pour l'obtention du grade de

Docteur en Sciences Actuarielles

1999

Jury de thèse

Monsieur François Dufresne
Professeur à l'École des HEC de l'Université de Lausanne. Directeur de thèse.

Monsieur André Dubey
Professeur à l'École des HEC de l'Université de Lausanne. Expert interne.

Monsieur Florian De Vylder
Professeur au Département de sciences actuarielles de l'Université d'Amsterdam. Expert externe.

À mes parents

Avant-propos

J'aimerais remercier les personnes sans qui cette thèse ne serait sans doute pas aujourd'hui réalité.

Mes sincères remerciements vont d'abord à mon directeur de recherche, Monsieur le Professeur François Dufresne. Il a su éveiller, puis entretenir chez moi la passion pour la recherche et l'enseignement. Je lui dois beaucoup, tant au plan professionnel que personnel.

Je remercie également Messieurs les Professeurs Florian De Vylder et André Dubey. Leur présence au sein du jury de thèse fut pour moi un grand honneur.

Je garderai un excellent souvenir de mon séjour à l'Université de Lausanne en tant que doctorant, mais aussi en tant qu'assistant. J'ai en effet occupé ce poste pendant quatre années avec un intérêt sans cesse renouvelé. Je m'en voudrais de passer ici sous silence l'influence qu'a eu sur moi Monsieur le Professeur Hans U. Gerber ainsi que tous les membres de l'Institut de sciences actuarielles.

Un grand merci également à mon collègue Gérard Pafumi pour sa disponibilité et son enthousiasme à régler une foule de petits détails. Son aide m'aura été précieuse lors de mon retour au Québec.

Dans de tels remerciements viennent souvent en dernier ceux à qui, pourtant, l'on doit le plus en termes de support moral, encouragements, amitié et amour. Au premier rang de ceux-là viennent évidemment mes parents, ainsi que ma compagne Catherine. Merci pour tout. Je salue, enfin, mes fidèles amis et amies au Québec ainsi tous ceux et celles qui auront fait de mon séjour en Suisse un moment inoubliable.

Montréal
Janvier 1999

Table des matières

Liste des tableaux

Table des figures

Chapitre 1

Introduction

1.1 Théorie de la crédibilité

« Les mathématiques de l'hétérogénéité », telle est parfois la définition donnée de la théorie de la crédibilité. Si cette définition n'éclairera probablement pas beaucoup le profane des sciences actuarielles, elle a néanmoins le mérite de résumer en quelques mots l'essentiel de la théorie et, dans un même temps, de garder plusieurs portes ouvertes vers de possibles champs d'application de l'outil.

La théorie de la crédibilité est la pierre angulaire de la théorie actuarielle des assurance de dommages. Elle origine du début du XX^{e} siècle avec les travaux de Mowbray (1914). Cet article sera le germe de la première grande branche de la théorie de la crédibilité, appelée en anglais *limited fluctuation credibility*. En français, nous proposons l'appellation « crédibilité de stabilité », bien que l'expression « crédibilité américaine » soit aussi parfois rencontrée. La crédibilité de stabilité recherche le point (nombre de sinistres, base d'exposition ou autre) à partir duquel, du point de vue de l'assureur, la prime pure d'un assuré peut être considérée parfaitement *crédible*. Cette approche ne sera pas davantage abordée dans la présente thèse.

La seconde grande branche de la théorie de la crédibilité, celle dans laquelle se situe nos travaux, est l'approche *greatest accuracy*, dont nous proposons la traduction « crédibilité de précision ». On rencontrera aussi l'expression « crédibilité européenne » car, bien que la théorie ait ses racines aux États-Unis, l'essentiel du développement s'est fait sur le Vieux continent. L'approche de précision sert à calculer la meilleure prime — au sens des moindres carrés — à charger à un assuré. On ne s'intéresse alors pas tant à la stabilité (dans le temps) de l'expérience d'un assuré qu'à sa similitude avec celle des autres assurés d'un même groupe. Si l'expérience du groupe est homogène, alors la prime collective constitue une bonne estimation de la prime de chaque assuré. Dans le cas contraire, il convient de prendre compte de l'expérience individuelle dans le calcul des primes, ceci via un *facteur de crédibilité*. La prime résultante est appelée «prime de crédibilité».

Il est généralement reconnu que les pionniers de la crédibilité de précision sont Whitney (1918), Bailey (1945, 1950) et, dans une mesure un peu moindre, Mayerson (1964). Toutefois, ce sont véritablement les deux articles de Bühlmann (1967, 1969) ainsi que celui de Bühlmann & Straub (1970) qui donnèrent un irrésistible essor à une théorie jusqu'alors marginale. Par conséquent, et aussi par souci d'exhaustivité, le désormais célèbre modèle de Bühlmann–Straub sera brièvement présenté ici.

Soit X_{it}, $t = 1, \dots, T_i$ ($T_i > 1$), le ratio d'expérience de l'assuré (ou risque, ou contrat) $i = 1, \dots, I$ d'un portefeuille d'assurance. À chaque assuré correspond une *variable de structure* Θ_i représentant le niveau de risque de l'assuré. Les hypothèses du modèle sont les suivantes.

(BS1) Les assurés $(X_{i1}, \dots, X_{iT_i}, \Theta_i)$, $i = 1, \dots, I$ sont mutuellement indépendants. Les variables $\Theta_1, \dots, \Theta_I$ sont identiquement distribuées et les variables aléatoires X_{it} ont un second moment fini.

(BS2) Pour tout $i = 1, \dots, I$ et $t, u = 1, \dots, T_i$,

$$\begin{aligned} \mathrm{E}[X_{it}|\Theta_i] &= \mu(\Theta_i), \\ \mathrm{Cov}(X_{it}, X_{iu}|\Theta_i) &= \delta_{tu} \frac{\sigma^2(\Theta_i)}{w_{it}}. \end{aligned}$$

Les constantes (connues) w_{it} sont des poids attribués aux observations X_{it} et δ_{tu} est le delta de Kronecker, valant 1 si $t = u$ et zéro sinon. La fonction $\mu(\Theta_i)$ est appelée *prime de risque* de l'assuré i.

Avec la notation $m = \mathrm{E}[\mu(\Theta_i)]$, $s^2 = \mathrm{E}[\sigma^2(\Theta_i)]$ et $a = \mathrm{Var}[\mu(\Theta_i)]$, le meilleur estimateur non homogène linéaire de $\mu(\Theta_i)$ ou de X_{i,T_i+1} au sens des moindres carrés, c'est-à-dire minimisant

$$\mathrm{E}\Big[(\mu(\Theta_i) - \widehat{\mu(\Theta_i)})^2\Big]$$

ou

$$\mathrm{E}\Big[(X_{i,T_i+1} - \widehat{X}_{i,T_i+1})^2\Big]$$

(dans ce second cas on parle généralement d'un *prédicteur*), est

$$\widehat{\mu(\Theta_i)} = \widehat{X}_{i,T_i+1} = z_i X_{iw} + (1 - z_i) m, \tag{1.1}$$

où

$$X_{iw} = \sum_{t=1}^{T_i} \frac{w_{it}}{w_{i\Sigma}} X_{it}, \quad w_{i\Sigma} = \sum_{t=1}^{T_i} w_{it}$$

et

$$z_i = \frac{a w_{i\Sigma}}{a w_{i\Sigma} + s^2}.$$

Le terme z_i est le *facteur de crédibilité*, X_{iw} est la moyenne pondérée par les poids dits «naturels» de l'assuré i et m, s^2 et a sont les *paramètres de structure* du portefeuille représentant, respectivement, la moyenne collective, la variabilité temporelle de l'expérience et la variabilité entre les assurés. En pratique, ces paramètres de structure sont généralement inconnus et doivent par conséquent être estimés à partir des données du portefeuille.

Les estimateurs sans biais classiques des paramètres m, s^2 et a sont les suivants :

$$\widehat{m} = X_{zw}, \tag{1.2}$$

$$\hat{s}^2 = \frac{1}{\sum_{i=1}^{I}(T_i - 1)} \sum_{i=1}^{I} (T_i - 1)\hat{s}_i^2, \tag{1.3}$$

$$\tilde{a} = \frac{1}{I-1} \sum_{i=1}^{I} z_i (X_{iw} - X_{zw})^2 \tag{1.4}$$

$$= \frac{1}{2(I-1)z_\Sigma} \sum_{i=1}^{I} \sum_{j=1}^{I} z_i z_j (X_{iw} - X_{jw})^2, \tag{1.5}$$

où

$$\hat{s}_i^2 = \frac{1}{T_i - 1} \sum_{t=1}^{T_i} w_{it} (X_{it} - X_{iw})^2 \tag{1.6}$$

$$= \frac{1}{2(T_i - 1) w_{i\Sigma}} \sum_{t=1}^{T_i} \sum_{u=1}^{T_i} w_{it} w_{iu} (X_{it} - X_{iu})^2 \tag{1.7}$$

et

$$X_{zw} = \sum_{i=1}^{I} \frac{z_i}{z_\Sigma} X_{iw}, \qquad z_\Sigma = \sum_{i=1}^{I} z_i. \tag{1.8}$$

On remarque que $\hat{s}_i^2$ est un estimateur de s^2 basé sur l'assuré i seulement, et que X_{zw} est une moyenne pondérée du portefeuille où les poids utilisés sont les facteurs de crédibilité. Enfin, on dit de $\tilde{a}$ qu'il s'agit d'un *pseudo-estimateur* parce qu'il est lui-même fonction du paramètre à estimer.

Les quelques lignes qui précèdent ne constituent évidemment qu'un rapide survol de la théorie. De plus amples détails sont donnés dans, entre autres, Norberg (1979), Goovaerts, Kaas, van Heerwaarden & Bauwelinckx (1990) et Goovaerts & Hoogstad (1987).

1.2 Crédibilité à classification croisée

Sans doute comme dans toute activité scientifique impliquant une modélisation de la réalité, la recherche en sciences actuarielles est souvent axée sur l'élaboration d'un modèle décrivant le mieux possible la réalité d'une part, ou étant le plus général possible d'autre part. En théorie de la crédibilité de précision, un grand pas avait été franchi en direction du modèle général avec l'avènement de la crédibilité hiérarchique (Jewell 1975). Ce concept englobait les modèles de crédibilité basés sur une structure de classification simple (on pense ici au modèle de Bühlmann–Straub, mais également à ceux de De Vylder (1976*b*) ou Hachemeister (1975)) et en étendait l'applicabilité à des problèmes plus complexes de portefeuilles divisés en plusieurs niveaux hiérarchisés. La pertinence de la modélisation hiérarchique se limite toutefois aux portefeuilles où il ne saurait y avoir d'interaction entre les différents facteurs de risque.

C'est sur ce problème que s'est penché Dannenburg (1995), pour finalement proposer un modèle de crédibilité à classification croisée (*Crossed Classification Credibility*, ou CCC), où tous les facteurs de risque sont modélisés symétriquement. En plus de trouver des applications dans les cas de grilles de classification présentant des interactions entre les facteurs de risque, ce modèle a la vertu de contenir, et donc de généraliser, le modèle hiérarchique de Jewell ainsi que certains modèles d'estimation de réserves IBNR (*Incurred But Not Reported*). Dans son article original, Dannenburg expose les principales prémisses — applications, notation, hypothèses — et conclusions purement mathématiques de son modèle. Il propose aussi de représenter chaque assuré d'un portefeuille comme une somme de variables aléatoires indépendantes modifiant la moyenne collective, chaque variable exprimant l'apport d'un facteur de risque ou de l'interaction de deux ou plusieurs de ces facteurs à la variance du dit assuré. Cette approche est inspirée des modèles linéaires à effets aléatoires, brièvement étudiés à la section 2.3. On doit d'ailleurs à Dannenburg (1995) cette nouvelle approche de la théo-

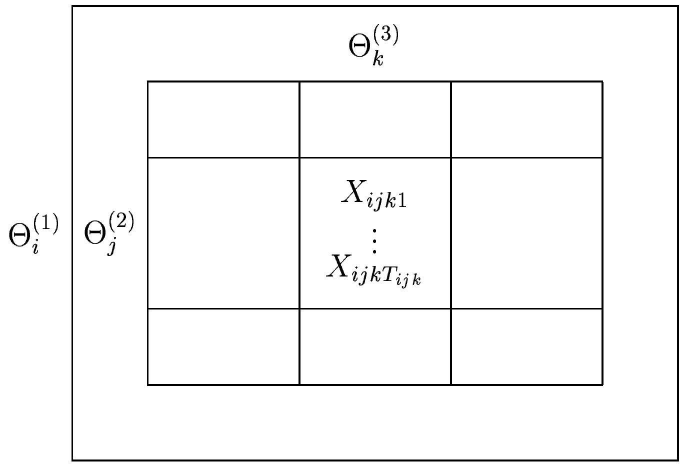

FIGURE 1.1: Représentation graphique d'un modèle de crédibilité à classification croisée à trois facteurs

rie de la crédibilité. Dannenburg, Kaas & Goovaerts (1996) ont par la suite repris l'idée pour présenter plusieurs modèles de crédibilité classiques d'une manière aussi nouvelle qu'originale.

À titre d'exemple, considérons un modèle CCC à trois facteurs, c'est-à-dire que l'assuré X_{ijkt} d'un portefeuille est affecté par trois facteurs de risque et toutes leurs interactions, en plus du temps. La variable de structure $\Theta_i^{(1)}$ représente le niveau de risque du premier facteur de risque, $\Theta_j^{(2)}$ le niveau de risque du second facteur et, enfin, $\Theta_k^{(3)}$ celui du troisième. Ce modèle est représenté schématiquement à la figure 1.1. Dans la version la plus simple du modèle CCC utilisant l'approche des composants de variance, les variables de structure font place à des variables Ξ représentant l'effet d'un facteur de risque ou de l'interaction entre plusieurs facteurs de risque sur un assuré[1]. Le ratio de l'assuré est donc décomposé ainsi :

$$\begin{aligned} X_{ijkt} = m &+ \Xi_i^{(1)} + \Xi_j^{(2)} + \Xi_k^{(3)} + \Xi_{ij}^{(12)} \\ &+ \Xi_{ik}^{(13)} + \Xi_{jk}^{(23)} + \Xi_{ijk}^{(123)} + \Xi_{ijkt}^{(1234)}, \end{aligned} \tag{1.9}$$

où m est la moyenne du portefeuille. Ces variables aléatoires sont toutes mutuellement indépendantes et de moyenne nulle. Les paramètres de structure du modèle sont m et les variances des variables Ξ. Le modèle CCC sera étudié plus en détail au chapitre 3.

Dans le modèle CCC standard proposé par Dannenburg, tout assuré du portefeuille est affecté par chacun des facteurs de risque. Or, on peut imagi-

[1]Le passage de la modélisation avec variables de structure à celle avec composants de variance est présenté plus en détails au chapite 6.

ner, par exemple, que la présence ou l'absence de freins antiblocage ou de moteur turbocompressé soient utilisés comme critère de classification dans une grille de tarification en casco automobile. En assurance-ménage, on pourrait considérer le fait qu'une maison soit ou non équipée d'un système d'alarme et, le cas échéant, de quel type il s'agit (par exemple, rattaché ou non au service de police). De tels exemples démontrent l'intérêt d'une généralisation du modèle CCC où le nombre de facteurs de risque saurait être variable par assuré.

La généralisation proposée dans cette thèse consiste à introduire une nouvelle catégorie pour chaque facteur de risque, numérotée 0 par convention, correspondant au fait de ne *pas* être influencé par un facteur de risque. Le ratio d'un assuré s'écrit alors

$$\begin{aligned} X_{ijkt} = m &+ \overline{\delta}_{0i}\Xi_i^{(1)} + \overline{\delta}_{0j}\Xi_j^{(2)} + \overline{\delta}_{0k}\Xi_k^{(3)} + \overline{\delta}_{0i}\overline{\delta}_{0j}\Xi_{ij}^{(12)} \\ &+ \overline{\delta}_{0i}\overline{\delta}_{0k}\Xi_{ik}^{(13)} + \overline{\delta}_{0i}\overline{\delta}_{0k}\Xi_{jk}^{(23)} + \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k}\Xi_{ijk}^{(123)} + \Xi_{ijkt}^{(1234)}, \end{aligned}$$

où $\overline{\delta}_{0i} = 1 - \delta_{0i}$. Le modèle de crédibilité à classification croisée généralisé (CCCG) est étudié au chapitre 5.

1.3 Estimation optimale de paramètres de structure

On l'a vu à la section 1.1, les paramètres de structure propres à tous les modèles de crédibilité doivent, en pratique, être estimés à partir des données du portefeuille. Pour ce faire, on cherchera évidemment à utiliser les estimateurs les plus performants possibles. Or, depuis l'avènement du modèle de Bühlmann (1967), une large part de la recherche en théorie de la crédibilité a été consacrée au développement d'estimateurs des paramètres de structure ainsi qu'à l'étude de leurs propriétés. Mentionnons, à cet effet, les article de De Vylder (1978, 1981) et Dubey & Gisler (1981), ainsi que la série d'articles de De Vylder & Goovaerts (1991, 1992*a*, 1992*b*), qui nous intéressera plus particulièrement.

Les deux propriétés fréquemment souhaitées pour des estimateurs sont l'absence de biais et la variance minimale. Dans ce texte, un estimateur sera d'ailleurs dit *optimal* dans une certaine classe d'estimateurs sans biais si sa variance est minimale dans cette classe.

La première propriété est évidemment la plus facile à obtenir et à vérifier. Cependant, on notera que les pseudo-estimateurs tels que $\tilde{a}$ et X_{zw} présentés plus haut sont sans biais que si les facteurs de crédibilité sont connus.

Quant à la propriété de variance minimale, elle s'avère plus difficile à établir. Dans le cas du modèle de Bühlmann–Straub, il a été prouvé quelques années après la publication du modèle (De Vylder 1978) que la moyenne des observations pondérée par les facteurs de crédibilité, X_{zw}, a non seulement une plus faible variance que la moyenne pondérée par les poids naturels, X_{ww}, mais même la variance minimale parmi tous les estimateurs linéaires sans biais de la moyenne collective m (on pourra consulter Goulet (1994) pour une interprétation de ce résultat a priori surprenant). Il a par contre fallu attendre plusieurs années et l'article précité de De Vylder & Goovaerts (1992*b*) pour obtenir un résultat similaire pour les estimateurs $\hat{s}^2$ et $\tilde{a}$. Dans cet article, il est démontré que $\hat{s}^2$ et $\tilde{a}$ sont des estimateurs à variance minimale dans de vastes classes d'estimateurs sans biais si, respectivement, les variables conditionnelles $X_{it}|\Theta_i$ et X_{iw} sont un *coefficient d'excès* nul. Goulet (1998) a obtenu des résultats similaires, mais moins restrictifs, avec l'approche des composants de variance. Le coefficient d'excès γ_2 d'une variable aléatoire est défini comme le quotient du quatrième moment central et du carré de la variance, moins trois (voir la section 2.4).

La minimisation d'une variance fait immanquablement intervenir de problématiques quatrièmes moments de variables aléatoires. C'est précisément parce qu'elle fait disparaître ces quatrièmes moments que l'hypothèse d'excès nul se montre très utile dans la recherche d'estimateurs optimaux. La majorité des estimateurs optimaux développés dans la présente thèse l'est avec l'hypothèse d'excès nul.

L'emploi de cette hypothèse est toutefois controversé au sein de la communauté actuarielle du fait que la seule loi de probabilité connue ayant un coefficient d'excès nul est la loi normale. Or, la plupart des résultats d'optimalité peuvent être obtenus avec une hypothèse de normalité, très fréquente en statistique. Les tenants de cet argument devront déjà admettre que l'hypothèse de normalité est plus contraignante que celle d'excès nul. Ceci même davantage avec l'approche des composants de variance telle qu'exposée dans Goulet (1998), où aucune hypothèse n'est faite sur le troisième moment des variables aléatoires impliquées. De plus, que la loi normale soit la seule loi à excès nul connue ne signifie pas qu'elle soit la seule existante. Nous montrerons en effet à la section 7.3 qu'il est possible de construire, assez facilement, une infinité de distributions avec un coefficient d'excès égal à zéro. Enfin, peu importe la ou les hypothèses sous laquelle ou lesquelles sont obtenus des estimateurs optimaux, la situation est toujours préférable à l'ignorance complète des propriétés des estimateurs. Il est d'ailleurs permis de penser que si un estimateur possède de bonnes propriétés sous certaines hypothèses plus ou moins sévères, la situation ne sera pas diamétralement opposée si ces hypothèses ne sont pas satisfaites. Cet aspect sera d'ailleurs étudié numériquement au chapitre 8.

1.4 Sommaire

Le but du présent ouvrage est de contribuer à la théorie de la crédibilité par l'introduction de diverses extensions au modèle de crédibilité à classification croisée. Les nouvelles contributions se trouvent autant dans les éléments proposés que dans la manière de les présenter.

Après un chapitre traitant de quelques notions mathématiques préliminaires, notre étude s'amorce, au chapitre 3, par l'exposé des hypothèses et de l'estimateur de crédibilité du modèle à classification croisée à trois facteurs de risque. L'estimateur de crédibilité y est développé à l'aide de projections dans des espaces de Hilbert. Seul le modèle à deux facteurs de risque fut considéré par Dannenburg (1995) et, de plus, l'aspect géométrique du modèle CCC n'avait encore fait l'objet d'aucune étude.

Le point central de tous les modèles de crédibilité — du moins en vue de leur application pratique — est l'estimation des paramètres de structure. Nous présentons en premier lieu ici les versions, pour un modèle à trois facteurs, des estimateurs de la moyenne collective et des composants de variance proposés initialement par Dannenburg (1995). Une forte proportion d'estimateurs de variance négatifs et un coefficient de variation élevé lors de tests numériques nous ont toutefois incité à rechercher des alternatives aux estimateurs de Dannenburg. La première piste suivie est la généralisation des estimateurs ANOVA utilisés en théorie des modèles linéaires. Les estimateurs « ad hoc » trouvés sont présentés dès le chapitre 3. Au chapitre suivant, nous abordons l'estimation optimale — à variance minimale, donc — des paramètres de structure. L'estimateur à variance minimale de la moyenne collective est d'abord obtenu sans hypothèse additionnelle sur la distribution des ratios ou des effets aléatoires. Avec l'introduction de l'hypothèse d'excès nul pour tous les effets aléatoires, nous montrons ensuite que l'estimateur classique de la variance temporelle est optimal dans une grande classe d'estimateurs sans biais. Enfin, nous présentons les estimateurs optimaux des autres composants de variance, toujours développés sous l'hypothèse d'excès nul. Les propriétés théoriques de tous ces estimateurs sont passées en revue.

Le chapitre 5 est dévolu à la généralisation du modèle de crédibilité à classification croisé. Le modèle généralisé admet un nombre de facteurs de risque variable par assuré. Nous reprenons alors l'étude des deux chapitres précédents, à savoir : calcul de l'estimateur de crédibilité par projections, généralisation des estimateurs de Dannenburg, présentation des estimateurs ad hoc et optimaux.

Les chapitres 3 à 5 se sont concentrés sur les modèles CCC et CCCG à deux et trois facteurs seulement. Or, il nous semble intéressant de pouvoir disposer des formules pour un modèle avec un nombre arbitraire $P > 1$

de facteurs. Du point de vue informatique, par exemple, de telles formules permettent d'écrire un seul programme pour le calcul des estimateurs de crédibilité. Le chapitre 6, aride à cause de la notation lourde qu'on y trouve, présente ces formules générales pour le modèle CCCG.

Enfin, cette thèse se termine par deux chapitres consacrés à l'étude numérique des divers estimateurs des paramètres de structure. Au chapitre 7, on discute de la mise en oeuvre informatique des formules rencontrées précédemment, du modèle de simulation des données et de la construction de distributions ayant un coefficient d'excès nul. Les résultats des simulations, quant à eux, font l'objet du chapitre 8.

Avant de passer au coeur du sujet, nous faisons remarquer que le terme « assuré » doit ici être pris dans son sens le plus large : personne civile, personne morale, contrat de réassurance, etc. De plus, plusieurs conventions d'écriture apparaissent ça et là dans cette thèse. Mentionnons tout de même ici qu'un tilde $\tilde{}$ identifie les fonctions propres au modèle CCC généralisé ainsi que les pseudo-estimateurs (itératifs). Ce double usage ne devrait pas causer de confusion.

Chapitre 2

Préliminaires

Ce chapitre présente quelques résultats, sans véritable lien entre eux, dont il sera fait usage tout au cours de cette thèse. On présente d'abord, dans les deux premières sections, un résumé de la théorie des espaces de Hilbert et leur utilisation en théorie de la crédibilité. Suit un rapide survol des modèles linéaires à composants de variance et de la principale procédure d'estimation de ces composants. Le chapitre se termine enfin sur l'énoncé de deux lemmes et des premières d'une série de conventions d'écriture disséminées dans différents chapitres de cette thèse.

2.1 Théorie des espaces de Hilbert

Retenant plusieurs concepts des espaces euclidiens, notamment celui d'orthogonalité, les espaces de Hilbert constituent une généralisation harmonieuse et intuitive de ceux-ci. La théorie des espaces de Hilbert est aujourd'hui l'outil sinon le plus puissant, certainement le plus utile dans l'application de l'analyse fonctionnelle. On le verra, un concept central de la théorie est celui de projection, consistant à identifier, dans un espace de Hilbert, le point le plus rapproché (selon une métrique préalablement définie) d'un autre point. C'est cette notion de minimisation de distance qui constitue l'attrait de la technique en théorie de la crédibilité.

On s'en tiendra, dans cet ouvrage, aux nombres réels, dont l'ensemble sera noté $\mathbb{R}$. L'ensemble des nombres naturels sera quant à lui noté $\mathbb{N}$.

Un espace pré-hilbertien $\mathcal{H}$ est un espace vectoriel muni d'un produit scalaire $\langle \cdot, \cdot \rangle$, soit une fonction satisfaisant les propriétés suivantes : pour tout x, y et $z \in \mathcal{H}$ et $\alpha \in \mathbb{R}$,

i) $\langle x, y \rangle = \langle y, x \rangle$

ii) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$

iii) $\langle\, \alpha\, x, y \,\rangle = \alpha \,\langle\, x, y \,\rangle$
iv) $\langle\, x, x \,\rangle \geq 0$
v) $\langle\, x, x \,\rangle = 0 \iff x = 0.$

Ce produit scalaire induit une norme, définie ainsi :

$$\|x\| = \langle\, x, x \,\rangle^{1/2}.$$

Enfin, on dit qu'un élément x d'un espace pré-hilbertien $\mathcal{H}$ est orthogonal (ou « perpendiculaire ») à $y \in \mathcal{H}$ si et seulement si $\langle\, x, y \,\rangle = 0$.

Un espace de Hilbert $\mathcal{H}$ est un espace pré-hilbertien complet, c'est-à-dire que toute suite de Cauchy $\{x_n\}$ dans $\mathcal{H}$ converge dans $\mathcal{H}$. Puisque tout sous-espace fermé (contenant tous ses points limites) d'un ensemble complet est lui-même complet, tout sous-espace fermé $\mathcal{M}$ d'un espace de Hilbert est lui-même un espace de Hilbert et passe par conséquent par l'origine. En revanche, le sous-espace fermé *translaté* $X + \mathcal{M} = \{X + Y, Y \in \mathcal{M}\}$, ne passe pas par l'origine si $X \notin \mathcal{M}$ est différent de 0. Enfin, on définit $\mathcal{M}^{\perp}$ comme le complément orthogonal de $\mathcal{M}$, soit l'ensemble des points de $\mathcal{H}$ orthogonaux à tous les éléments de $\mathcal{M}$. L'espace des nombres réels, $\mathbb{R}$, est un exemple d'espace de Hilbert.

Le *span* fermé $\overline{\mathrm{sp}}\{x_n, n \in \mathbb{N}\}$ d'un ensemble de points $\{x_n, n \in \mathbb{N}\}$ d'un espace de Hilbert $\mathcal{H}$ est le plus petit sous-espace fermé de $\mathcal{H}$ contenant tous les éléments x_n. Pour un ensemble fini de points $\{x_1, \ldots, x_n\}$, le *span* fermé est l'ensemble des combinaisons linéaires $y = \alpha_1 x_1 + \cdots + \alpha_n x_n$ de $x_1, \ldots, x_n$ (où $\alpha_1, \ldots, \alpha_n$ sont des nombres réels quelconques). Pour cette raison, nous dirons que $\mathcal{M} = \overline{\mathrm{sp}}\{x_n, n \in \mathbb{N}\}$ est un sous-espace *linéaire* de $\mathcal{H}$.

La définition centrale dans l'utilisation des espaces de Hilbert est celle de *projection*. Soit $\mathcal{M}$ un sous-espace fermé d'un espace de Hilbert $\mathcal{H}$ et x un élément de $\mathcal{H}$. On définit la projection de x sur $\mathcal{M}$, $\widehat{x} = \mathrm{pro}(x|\mathcal{M})$, comme l'élément de $\mathcal{M}$ tel que $(x - \widehat{x})$ est orthogonal à tous les éléments de $\mathcal{M}$:

$$\widehat{x} = \mathrm{pro}(x|\mathcal{M}) \iff \langle\, x - \widehat{x}, y \,\rangle = 0 \quad \forall y \in \mathcal{M}. \tag{2.1}$$

L'expression de droite dans (2.1) définit les *équations normales* à résoudre dans un calcul de projection.

Le lemme suivant, énoncé sans preuve, présente les principales propriétés de l'opérateur projection.

Lemme 2.1. *Soient* $\mathcal{M}$ *et* $\mathcal{M}_1 \subseteq \mathcal{M}_2$ *des sous-espaces fermés d'un espace de Hilbert* $\mathcal{H}$, x *et* y *des éléments de* $\mathcal{H}$ *et* a *et* b *des éléments de* $\mathbb{R}$. *Alors,*

(P1) $\mathrm{pro}(ax + by|\mathcal{M}) = a\,\mathrm{pro}(x|\mathcal{M}) + b\,\mathrm{pro}(y|\mathcal{M})$

(P2) $\mathrm{pro}(x|\mathcal{M}_1) = \mathrm{pro}(\mathrm{pro}(x|\mathcal{M}_2)|\mathcal{M}_1)$

(P3) $\mathrm{pro}(x|\mathcal{M}) = x \iff x \in \mathcal{M}$

(P4) $\mathrm{pro}(x|\mathcal{M}) = 0 \iff x \in \mathcal{M}^{\perp}$.

On notera que la propriété de linéarité (P1) ne s'applique pas aux sous-espaces fermés translatés.

Le résultat suivant est présenté sous forme de théorème puisqu'il est au coeur de la théorie et de l'utilisation des espaces de Hilbert. Puisqu'il s'agit d'un résultat classique dans le domaine, ce théorème ne sera pas prouvé (se référer, par exemple, à Kreyszig (1989), section 3.3 pour une preuve).

Théorème 2.1 (Théorème de projection). *Soit $\mathcal{M}$ un sous-espace fermé de l'espace de Hilbert $\mathcal{H}$ et $x \in \mathcal{H}$. Alors il existe un et un seul élément $\widehat{x} \in \mathcal{M}$ tel que*

$$\|x - \widehat{x}\| = \min_{y \in \mathcal{M}} \|x - y\|. \tag{2.2}$$

Cet élément $\widehat{x}$ est la projection de x sur $\mathcal{M}$, c'est-à-dire $\widehat{x} = \mathrm{pro}(x|\mathcal{M})$.

Un seul espace de Hilbert sera utilisé dans cet ouvrage, il s'agit de l'espace communément appelé $\mathcal{L}^2$. L'espace de Hilbert $\mathcal{L}^2$ sur un espace de Kolmogorov (Ω, F, P) est l'ensemble des variables aléatoires de second moment fini. Deux variables aléatoires X et Y de cet espace sont considérées équivalentes si elles ne diffèrent que sur un ensemble de mesure nulle (c'est-à-dire que $P[X = Y] = 1$). Le produit scalaire y est défini ainsi :

$$\langle X, Y \rangle = \mathrm{E}[XY]\,. \tag{2.3}$$

Puisque toute constante peut être considérée comme une variable aléatoire, $\mathbb{R}$ est un sous-espace de $\mathcal{L}^2$.

On remarque donc que $\|X\|^2 = \mathrm{E}[X^2]$ dans l'espace $\mathcal{L}^2$. Si $X \in \mathcal{L}^2$, alors le théorème de projection précise que $\widehat{X} = \mathrm{pro}(X|\mathcal{M})$ est la meilleure prédiction de X dans $\mathcal{M}$ au sens des moindres carrés, c'est-à-dire l'élément $Y \in \mathcal{M}$ minimisant $\mathrm{E}[(X - Y)^2]$. $\widehat{X}$ est donc l'espérance conditionnelle de X sachant $\mathcal{M}$. Si $\mathcal{M} = \mathcal{M}(Y_1, \ldots, Y_n)$ est l'ensemble de toutes les fonctions à valeurs réelles de $Y_1, \ldots, Y_n$, alors $\widehat{X} = \mathrm{pro}(X|\mathcal{M}(Y_1, \ldots Y_n)) = \mathrm{E}[X|Y_1, \ldots, Y_n]$, soit l'estimateur bayesien pur. Comme cette dernière projection s'avère fréquemment difficile à calculer, on se contente habituellement de la meilleure prédiction *linéaire* de X sur $\overline{\mathrm{sp}}\{1, Y_1, \ldots Y_n\}$, plus facile à calculer. Par la propriété (P2) du lemme 2.1,

$$\begin{multline} \mathrm{pro}(X|\overline{\mathrm{sp}}\{1, Y_1, \ldots, Y_n\}) = \\ \mathrm{pro}(\mathrm{pro}(X|\mathcal{M}(Y_1, \ldots, Y_n))|\overline{\mathrm{sp}}\{1, Y_1, \ldots Y_n\}). \end{multline} \tag{2.4}$$

La projection de X peut donc se décomposer en deux parties, soit d'abord la projection sur l'espace des fonctions des observations (estimateur bayesien pur), puis la projection sur l'espace des fonctions linéaires des observations (estimateur bayesien linéaire). Il arrive que $\text{pro}(X|\mathcal{M}(Y_1,\dots,Y_n)) \in \overline{\text{sp}}\{1, Y_1, \dots Y_n\}$, c'est-à-dire que l'estimateur bayesien pur est linéaire.

Les sous-espaces linéaires de $\mathcal{L}^2$ étant les plus fréquents dans cet ouvrage, $\overline{\text{sp}}\{X_1, \dots X_n\}$ sera dorénavant noté $\mathcal{L}\{X_1, \dots X_n\}$. De même, on définit $\mathcal{L}_m\{X_1, \dots X_n\}$ comme étant le sous-espace translaté linéaire de $\mathcal{L}^2$ dont l'espérance de chaque élément est égale à m, c'est-à-dire

$$\mathrm{E}[\alpha_1 X_1 + \cdots + \alpha_n X_n] = m.$$

On suppose toujours qu'au moins une espérance $\mathrm{E}[X_i]$, $i = 1, \dots, n$ est non nulle, ceci afin d'assurer que le sous-espace $\mathcal{L}_m$ n'est pas vide. $\mathcal{L}_m$ est un sous-espace linéaire translaté de $\mathcal{L}^2$ dont un sous-espace linéaire parallèle est $\mathcal{L}_0 = \mathcal{L}_m - m$ (De Vylder & Goovaerts 1992*a*).

2.2 Espaces de Hilbert et crédibilité

Depuis que De Vylder (1976*a*) en a fait l'éclatante démonstration, il est reconnu que la théorie de la crédibilité (de précision) se prête fort bien à une modélisation et une interprétation géométrique. Le calcul de l'estimateur de crédibilité peut être grandement facilité en ayant recours à la théorie des espaces de Hilbert, ceci tout particulièrement dans le cas des modèles hiérarchiques (Bühlmann & Jewell 1987). Cette technique sera utilisée plus loin pour développer les formules générales des modèles CCC et CCC généralisé.

Prenons le cas d'un seul assuré. Le problème de la théorie de la crédibilité de précision consiste à estimer la variable aléatoire X_{n+1} à partir du vecteur aléatoire des observations disponibles $\underline{X} = (X_1, \dots X_n)$. On suppose en général que toutes les variables aléatoires sous étude ont un second moment fini et, par conséquent, l'ensemble de ces variables aléatoires forme un sous-espace de $\mathcal{L}^2$. Tel que mentionné ci-dessus, le meilleur estimateur de X_{n+1} au sens des moindres carrés est donc $\mathrm{E}[X_{n+1}|X_1, \dots, X_n]$. Cet estimateur se présentant sous forme explicite dans certains cas seulement, on se limite plutôt à des estimateurs linéaires de la forme

$$\widehat{X}_{n+1} = \alpha_0 + \sum_{j=1}^{n} \alpha_j X_j. \tag{2.5}$$

Par le théorème de projection, on sait que cet *estimateur de crédibilité* existe, est unique et est donné par la projection de X_{n+1} sur $\mathcal{L}\{1, X_1, \dots, X_n\} \equiv$

$\overline{\mathrm{sp}}\{1, X_1, \ldots X_n\}$:

$$\widehat{X}_{n+1} = \mathrm{pro}(X|\mathcal{L}\{1, X_1, \ldots, X_n\}). \tag{2.6}$$

La constante 1 (ou tout autre constante) est incluse dans l'espace afin que l'estimateur de crédibilité soit sans biais.

Les $n+1$ *équations normales* dans un tel cas sont par conséquent :

$$\begin{gathered}
\langle\, X_{n+1} - \widehat{X}_{n+1}, 1 \,\rangle = 0 \overset{(\mathrm{P1})}{\Longleftrightarrow} \langle\, X_{n+1}, 1 \,\rangle = \langle\, \widehat{X}_{n+1}, 1 \,\rangle \\
\langle\, X_{n+1} - \widehat{X}_{n+1}, X_k \,\rangle = 0 \overset{(\mathrm{P1})}{\Longleftrightarrow} \langle\, X_{n+1}, X_k \,\rangle = \langle\, \widehat{X}_{n+1}, X_k \,\rangle, \quad k = 1, \ldots, n.
\end{gathered}$$

En utilisant la définition du produit scalaire dans l'espace $\mathcal{L}^2$ et en remplaçant $\widehat{X}_{n+1}$ par (2.5), on peut réécrire le système de $n+1$ équations normales sous cette forme :

$$\begin{aligned}
\alpha_0 + \sum_{j=1}^{n} \alpha_j \mathrm{E}[X_j] &= \mathrm{E}[X_{n+1}] \\
\sum_{j=1}^{n} \alpha_j \mathrm{Cov}(X_j, X_k) &= \mathrm{Cov}(X_{n+1}, X_k), \quad k = 1, \ldots, n.
\end{aligned} \tag{2.7}$$

Il suffit de résoudre ce système d'équations, explicitement ou numériquement, pour obtenir l'estimateur de crédibilité.

D'autre part, l'on pourrait aussi chercher à obtenir un estimateur (linéaire) d'un paramètre constant p comme, par exemple, la moyenne ou la variance. Si l'on suit la procédure tout juste exposée et projette ledit paramètre sur l'espace $\mathcal{L}\{1, X_1, \ldots, X_n\}$, alors $\hat{p} = p$, puisque les constantes font partie de l'espace. Cette projection ne s'avère donc d'aucune utilité.

La solution, proposée à l'origine par De Vylder & Goovaerts (1992*a*), consiste plutôt à projeter le paramètre p sur $\mathcal{L}_p\{X_1, \ldots, X_n\}$. L'estimateur $\hat{p}$ ainsi obtenu est une fonction linéaire des variables aléatoires $X_1, \ldots, X_n$ dont l'espérance est égale à p. De plus, par la définition de projection et de la norme dans $\mathcal{L}^2$, on sait que $\hat{p}$ est l'unique élément de $\mathcal{L}_p\{X_1, \ldots, X_n\}$ qui minimise

$$||p - \hat{p}||^2 = \mathrm{E}\big[(\hat{p} - p)^2\big] = \mathrm{Var}[\hat{p}]\,.$$

L'estimateur $\hat{p}$ de p est par conséquent l'estimateur sans biais à variance minimale dans la classe définie par $\mathcal{L}_p\{X_1, \ldots X_n\}$ ou, selon la définition du chapitre 1, l'estimateur *optimal* de p. Il est démontré dans De Vylder

& Goovaerts (1992a) que $\hat{p}$ est la solution des $n+1$ équations normales suivantes :

$$\begin{aligned} \mathrm{E}[\hat{p}] &= p \\ \mathrm{Cov}(\hat{p}, X_j) &= c\,\mathrm{E}[X_j]\,, \quad j = 1, \dots, n, \end{aligned} \tag{2.8}$$

où c est une constante quelconque.

L'estimation optimale de paramètres de structure sera la seconde fréquente utilisation des espaces de Hilbert dans cette thèse.

2.3 Modèles linéaires

L'analyse de variance, dont l'invention remonte au début du siècle avec R. A. Fischer, est une technique aussi utile que répandue en statistique. Au fil des décennies, la formulation théorique de l'analyse de variance s'est de plus en plus faite en termes de *modèles linéaires* (Searle (1971), Scheffé (1959), Rao (1965) pour ne nommer que trois références classiques dans le domaine). Les modèles linéaires sont de trois types : à effets fixes, aléatoires ou mixtes. Dans ce bref résumé, nous ne nous attarderons qu'aux modèles à effets aléatoires.

Un exemple classique de modèle à effets aléatoire est le modèle à classification croisée à deux facteurs (concept introduit à la section 1.2). Dans un tel modèle, la variable aléatoire X_{ijt} ($i = 1, \dots, I$; $j = 1, \dots, J$; $t = 1, \dots, T_{ij}$) représentant le résultat d'une expérience est décomposée ainsi[1] :

$$X_{ijt} = m + \Xi_i^{(1)} + \Xi_j^{(2)} + \Xi_{ij}^{(12)} + \Xi_{ijt}^{(123)},$$

où m est une constante et toutes les variables aléatoires Ξ — les *effets aléatoires* — sont indépendantes, de moyenne nulle et de variance $b^{(1)}$, $b^{(2)}$, $b^{(12)}$ et s^2, dans l'ordre. La variable $\Xi_{ijt}^{(123)}$ représente l'erreur. Ce modèle recèle quatre sources de variabilité : la variance de chacun des deux facteurs, la variance de leur interaction et l'erreur. Ces variances sont appelées les *composants de variance*, leur somme étant égale à la variance de la variable observée X_{ijt}.

Si $T_{ij} = T$, c'est-à-dire si le nombre d'observations est le même dans chaque « cellule », alors le modèle est dit « équilibré » (*balanced*). Dans le cas contraire, le modèle est « non équilibré » (*unbalanced*).

Le problème central ici est l'estimation des composants de variance à partir des données expérimentales. La procédure d'estimation de base repose sur le tableau d'analyse de variance, ou tableau ANOVA, dans lequel on

[1] La notation utilisée ici est celle ayant cours en théorie de la crédibilité et n'est pas du tout standard dans le domaine des modèles linéaires.

retrouve, rappelons-le, diverses sommes de carrés (SS) ainsi que le nombre de degrés de liberté leur étant associés. Pour un modèle équilibré, les principaux éléments d'un tel tableau sont les suivants :

$$\begin{aligned}
\mathrm{SSA} &= JT\sum_{i=1}^{I}(\bar{X}_i^{(1)} - \bar{X})^2, & \mathrm{MSA} &= \frac{\mathrm{SSA}}{I-1}, \\
\mathrm{SSB} &= IT\sum_{j=1}^{J}(\bar{X}_j^{(2)} - \bar{X})^2, & \mathrm{MSB} &= \frac{\mathrm{SSB}}{I-1}, \\
\mathrm{SSAB} &= T\sum_{i=1}^{I}\sum_{j=1}^{J}(\bar{X}_{ij}^{(12)} - \bar{X}_i^{(1)} - \bar{X}_j^{(2)} + \bar{X})^2, & \mathrm{MSAB} &= \frac{\mathrm{SSAB}}{(I-1)(J-1)}, \\
\mathrm{SSE} &= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T}(X_{ijt} - \bar{X}_{ij}^{(12)})^2, & \mathrm{MSE} &= \frac{\mathrm{SSE}}{IJ(T-1)},
\end{aligned}$$

où $\bar{X}_{ij}^{(12)} = T^{-1}\sum_t X_{ijt}$, etc., et A représente le premier facteur, B le second, AB leur interaction et E l'erreur. Il est par la suite aisé de vérifier que

$$\begin{aligned}
\mathrm{E}[\mathrm{MSA}] &= JTb^{(1)} + Tb^{(12)} + s^2, \\
\mathrm{E}[\mathrm{MSB}] &= ITb^{(2)} + Tb^{(12)} + s^2, \\
\mathrm{E}[\mathrm{MSAB}] &= Tb^{(12)} + s^2, \\
\mathrm{E}[\mathrm{MSE}] &= s^2.
\end{aligned}$$

En laissant tomber les opérateurs d'espérance et en isolant les composants de variance, on obtient les estimateurs sans biais suivants :

$$\begin{aligned}
\hat{b}^{(1)} &= \frac{\mathrm{MSA} - \mathrm{MSAB}}{JT}, \\
\hat{b}^{(2)} &= \frac{\mathrm{MSB} - \mathrm{MSAB}}{IT}, \\
\hat{b}^{(12)} &= \frac{\mathrm{MSAB} - \mathrm{MSE}}{T}, \\
\hat{s}^2 &= \mathrm{MSE}.
\end{aligned}$$

Ces estimateurs sont appelés les « estimateurs ANOVA ». On trouve des règles pour développer les estimateurs ANOVA pour des modèles à plus de deux facteurs à la section 4.1 de Searle, Casella & McCulloch (1992).

Remarque. En reprenant l'idée de la section 3.4 de De Vylder & Goovaerts (1992*a*), on peut réécrire les fonctions MS sous cette forme :

$$\mathrm{MSA} = \frac{JT}{2I(I-1)}\sum_{i=1}^{I}\sum_{k=1}^{I}(\bar{X}_i^{(1)} - \bar{X}_k^{(1)})^2,$$

$$\begin{aligned}
\text{MSB} &= \frac{IT}{2J(J-1)} \sum_{j=1}^{J} \sum_{l=1}^{J} (\bar{X}_j^{(2)} - \bar{X}_l^{(2)})^2, \\
\text{MSAB} &= \frac{T}{2J(I-1)(J-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{l=1}^{J} (\bar{X}_{ij}^{(12)} - \bar{X}_{il}^{(12)})^2 \\
&\quad + \frac{T}{2I(I-1)(J-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{I} (\bar{X}_{ij}^{(12)} - \bar{X}_{kj}^{(12)})^2 \\
&\quad - \frac{T}{2IJ(I-1)(J-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{I} \sum_{l=1}^{J} (\bar{X}_{ij}^{(12)} - \bar{X}_{kl}^{(12)})^2, \\
\text{MSE} &= \frac{1}{2IJT(T-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{t=1}^{T} \sum_{u=1}^{T} (X_{ijt} - X_{iju})^2.
\end{aligned}$$

En général, nous favoriserons dorénavant ce type d'écriture pour les estimateurs de variance.

Pour les modèles équilibrés, Graybill & Hultquist (1961) ont démontré que les estimateurs ANOVA sont à variance minimale dans la classe des estimateurs sans biais à fonction quadratique des observations. Si, au surplus, une hypothèse de normalité est imposée aux variables aléatoires Ξ, alors les estimateurs ANOVA sont optimaux dans toute la classe des estimateurs sans biais (Graybill 1954). Il importe toutefois de noter qu'aucun résultat similaire n'est disponible pour les modèles non équilibrés.

Un des principaux inconvénients des estimateurs de type ANOVA est leur probabilité non nulle d'être négatifs. Il existe d'autres techniques d'estimation des composants de variance résultant en des estimateurs toujours positifs, mais celles-ci reposent souvent sur une hypothèse de normalité. La théorie de la crédibilité étant principalement développée dans un contexte non paramétrique, ces techniques ne seront pas abordées ici.

2.4 Définition, lemmes et conventions

Avant de passer véritablement au sujet principal de cette thèse, il convient d'énoncer encore trois lemmes et quelques conventions d'écriture, en plus de définir formellement le coefficient d'excès d'une variable aléatoire.

Le coefficient d'excès γ_2 d'une variable aléatoire X est défini ainsi :

$$\gamma_2(X) = \frac{\mathrm{E}[(X - \mathrm{E}[X])^4]}{\mathrm{E}[(X - \mathrm{E}[X])^2]^2} - 3. \tag{2.9}$$

Par exemple, la loi normale n'a pas d'excès, ou $\gamma_2 = 0$.

Deux résultats fréquemment utilisés par la suite sont ici présentés sous forme de lemme. Une preuve ne s'avère toutefois pas nécessaire.

Lemme 2.2. *Soient X_i, X_j, X_k, X_l, X_t et X_u des variables aléatoires d'espérance commune m et de quatrième moment fini. Alors*

$$\mathrm{E}[(X_i - X_j)(X_k - X_l)] = \sigma_{ik} - \sigma_{il} - \sigma_{jk} + \sigma_{jl} \tag{2.10}$$

et

$$\begin{aligned}\mathrm{E}\big[(X_t - X_u)^2(X_i - X_j)(X_k - X_l)\big] = \\ &\mathrm{E}[X_tX_tX_iX_k - X_tX_tX_iX_l - X_tX_tX_jX_k + X_tX_tX_jX_l \\ &- 2\,X_tX_uX_iX_k + 2\,X_tX_uX_iX_l + 2\,X_tX_uX_jX_k - 2\,X_tX_uX_jX_l \\ &+X_uX_uX_iX_k - X_uX_uX_iX_l - X_uX_uX_jX_k + X_uX_uX_jX_l]\,,\end{aligned} \tag{2.11}$$

où $\sigma_{ij} = \mathrm{Cov}(X_i, X_j)$.

Le lemme suivant, déjà présenté dans Goulet (1998), sera essentiel pour travailler avec l'hypothèse d'excès nul dans le modèle CCC.

Lemme 2.3. *Soit $Y = X_1 + \cdots + X_n$, où les X_i $(i = 1, \dots, n)$ sont des variables aléatoires indépendantes sans excès. Alors $\gamma_2(Y) = 0$.*

Preuve. Ce résultat est une conséquence directe de la définition du coefficient d'excès en termes des quatrième et deuxième cumulants de la distribution sous-jacente. Voir Johnson, Kotz & Kemp (1992, chapitre 1). □

Le second lemme, utilisé abondamment dans la suite, est tiré de De Vylder & Goovaerts (1992a).

Lemme 2.4. *Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes, de moyenne nulle, variance σ_i^2 $(i = 1, \dots, n)$ et quatrième moment fini. Alors, pour $i, j, k, l = 1, \dots, n$:*

$$\mathrm{E}[X_iX_j] = \delta_{ij}\sigma_i^2 \tag{2.12}$$

$$\begin{aligned}\mathrm{E}[X_iX_jX_kX_l] &= \delta_{ij}\delta_{kl}\sigma_i^2\sigma_k^2 + \delta_{ik}\delta_{jl}\sigma_i^2\sigma_j^2 + \delta_{il}\delta_{jk}\sigma_i^2\sigma_j^2 \\ &\quad + \delta_{ijkl}(\mathrm{E}\big[X_i^4\big] - 3\,\sigma_i^4).\end{aligned} \tag{2.13}$$

Si toutes les variables X_i ont un coefficient d'excès nul, alors le dernier terme à la droite de (2.13) tombe.

Preuve. Il suffit de considérer toutes les combinaisons $i = j = k = l$, $i = j = k \neq l$, etc., pour vérifier la validité des formules. □

Dans le lemme précédent, la fonction δ est le delta de Kronecker, valant 1 si tous ses indices sont égaux et zéro sinon. Une redéfinition légèrement plus générale de cette fonction fait partie des conventions d'écriture suivantes.

Conventions. Les conventions d'écriture ci-dessous auront cours dans cette thèse.

(C1) Une variable soulignée représente un vecteur de variables avec un niveau supplémentaire d'indices. Ainsi, par exemple : $\underline{X}_{ij} = (X_{ij1}, \dots, X_{ijK})'$.

(C2) Un Σ à la fin d'une suite d'indices d'une variable signifie que l'on a sommé cette variable sur tous les indices possibles ne se trouvant pas dans la suite précitée. D'autre part, un w signifie que la variable est une moyenne pondérée par des poids naturels, toujours sur les indices ne se trouvant pas dans la liste qui précède. Par exemple, si X est une variable dont les indices possibles sont i, j, k, l, alors

$$X_{ij\Sigma} = \sum_k \sum_l X_{ijkl}$$

et

$$X_{iw} = \sum_j \sum_k \sum_l \frac{w_{ijkl}}{w_{i\Sigma}} X_{ijkl}.$$

De manière analogue, un z en indice caractérise une moyenne pondérée par les facteurs de crédibilité.

(C3) Soient $\mathcal{A} = \{a_1, \dots, a_m\}$ et $\mathcal{B} = \{b_1, \dots, b_n\}$ deux ensembles ordonnés de nombres. On dit que $\mathcal{A} = \mathcal{B}$ si et seulement si $m = n$ et $a_1 = b_1, \dots, a_n = b_n$. On utilise ainsi le delta de Kronecker :

$$\delta_{\mathcal{A}\mathcal{B}} = \begin{cases} 1 & \text{si } \mathcal{A} = \mathcal{B} \\ 0 & \text{sinon.} \end{cases}$$

De nouvelles conventions d'écriture apparaîtront ponctuellement dans le texte au fur et à mesure que leur besoin s'en fera sentir.

Chapitre 3

Modèles à deux et trois facteurs

En théorie de la crédibilité de précision, il est généralement supposé que la classification des assurés est sous forme hiérarchique à un ou plusieurs niveaux (Jewell (1975), Bühlmann & Jewell (1987)). Or, la classification hiérarchique peut s'avérer inappropriée lorsque, dans les faits, des assurés de différentes divisions d'un même niveau partagent des caractéristiques communes. On a qu'à penser, par exemple, à un portefeuille d'assurance automobile qui serait classifié hiérarchiquement d'abord selon le sexe, puis selon l'âge. Il semble évident que les jeunes femmes partagent des caractéristiques communes avec les jeunes hommes, qui à leur tour ont des habitudes de conduite semblables à quelque égard à celle des hommes plus âgés. La classification hiérarchique néglige donc l'existante interaction entre les facteurs de risque âge et sexe.

C'est pour pallier à cette lacune que Dannenburg (1995) a introduit sous une forme générale le modèle de crédibilité à classification croisée (CCC), où tous les facteurs de risque sont modélisés symétriquement.

L'idée maîtresse de ce modèle — à savoir la modélisation par effets aléatoires — a déjà été exposée au chapitre 1. Dans le présent chapitre ainsi qu'au suivant, nous entrerons davantage au coeur du sujet avec la présentation de l'estimateur de crédibilité et des estimateurs des paramètres de structure. Afin de se concentrer sur les principes et la «mécanique» du modèle CCC, on se contentera, dans un premier temps, d'étudier principalement le modèle à trois facteurs. Toutefois, lorsqu'un souci de simplicité et de brièveté l'exigera, on considérera également le modèle à deux facteurs, celui présenté dans Dannenburg (1995). La connaissance des principales formules des modèles CCC à deux et trois facteurs pavera ensuite la voie vers les formules générales d'un modèle à P facteurs, objet du chapitre 6.

À noter également que nous entamons l'étude géométrique du modèle CCC, c'est-à-dire que l'estimateur de crédibilité ainsi que les estimateurs optimaux des paramètres de structure seront obtenus à l'aide de projections dans des espaces de Hilbert.

3.1 Hypothèses et estimateur de crédibilité

Dans un modèle CCC à trois facteurs, tout assuré d'un portefeuille d'assurance est affecté par trois facteurs de risque ainsi que par toutes les interactions entre ces facteurs (on n'aborde pas, ici, des modèles tels que celui à risques additifs de Dannenburg (1995)). Le temps est toujours considéré comme un facteur de risque additionnel n'apparaissant que sous forme d'interaction avec tous les autres facteurs. La catégorie $i = 1, \dots, I$ du premier facteur de risque est caractérisée par une variable de structure $\Theta_i^{(1)}$, la catégorie $j = 1, \dots, J$ du second par $\Theta_j^{(2)}$ et la catégorie $k = 1, \dots, K$ du troisième par $\Theta_k^{(3)}$. La variable aléatoire X_{ijkt} représente l'expérience observée dans la cellule (ijk) au temps t ($t = 1, \dots, T_{ijk} > 1$) et, par conséquent, l'ensemble des données $\underline{X}$ forme un vecteur à quatre dimensions. Enfin, à chaque donnée X_{ijkt} est rattaché un poids w_{ijkt}. Les hypothèses du modèle sont les suivantes :

(CCC1) Pour chaque facteur de risque, les variables aléatoires $\Theta_i^{(1)}$, $\Theta_j^{(2)}$ et $\Theta_k^{(3)}$ sont i.i.d.. Les variables Θ sont indépendantes.

(CCC2) Il existe des fonctions $\mu.(\cdot)$ telles que

$$\begin{aligned}
\mathrm{E}\Big[X_{ijkt}|\Theta_i^{(1)}, \Theta_j^{(2)}, \Theta_k^{(3)}\Big] &= \mu_{123}(\Theta_i^{(1)}, \Theta_j^{(2)}, \Theta_k^{(3)}), \\
\mathrm{E}\Big[X_{ijkt}|\Theta_i^{(1)}, \Theta_j^{(2)}\Big] &= \mu_{12}(\Theta_i^{(1)}, \Theta_j^{(2)}), \\
\mathrm{E}\Big[X_{ijkt}|\Theta_i^{(1)}\Big] &= \mu_1(\Theta_i^{(1)}),
\end{aligned}$$

et ainsi de suite.

(CCC3) $\mathrm{E}\Big[\mathrm{Cov}(X_{ijkt}, X_{fghu}|\underline{\Theta}^{(1)}, \underline{\Theta}^{(2)}, \underline{\Theta}^{(3)})\Big] = \delta_{ijkt,fghu}\, s^2/w_{ijkt}$.

Ci-dessus, $\underline{\Theta}^{(1)} = (\Theta_1^{(1)}, \dots, \Theta_I^{(1)})$ et $\delta_{ijkt,fghu}$ est toujours le delta de Kronecker, valant 1 si $i = f$, $j = g$, $k = h$, $t = u$ et zéro sinon.

L'hypothèse (CCC3), classique en théorie de la crédibilité, statue sur l'indépendance conditionnelle des assurés ou, autrement dit, entre les ensembles $(X_{ijkt}, \underline{\Theta}^{(1)}, \underline{\Theta}^{(2)}, \underline{\Theta}^{(3)})$. La fonction $\mu_{123}(\Theta_i^{(1)}, \Theta_j^{(2)}, \Theta_k^{(3)})$ est la prime de risque de l'assuré (ijk).

Dannenburg (1995) utilise par la suite l'approche des composants de variance (Searle et al. 1992) et propose de représenter le ratio X_{ijkt} de chaque assuré comme une somme de variables aléatoires indépendantes exprimant chacune la contribution d'un facteur de risque ou de l'interaction entre plusieurs facteurs sur la variance de ces assurés. Ce faisant, l'interprétation et l'utilisation du modèle en sont grandement simplifiées. Dans la version la plus simple[1] du modèle, on écrit, tel qu'introduit au chapitre 1 :

$$\begin{aligned} X_{ijkt} = m &+ \Xi_i^{(1)} + \Xi_j^{(2)} + \Xi_k^{(3)} + \Xi_{ij}^{(12)} \\ &+ \Xi_{ik}^{(13)} + \Xi_{jk}^{(23)} + \Xi_{ijk}^{(123)} + \Xi_{ijkt}^{(1234)}, \end{aligned} \tag{3.1}$$

où les effets aléatoires $\Xi_i^{(1)}, \Xi_j^{(2)}, \dots, \Xi_{ijk}^{(123)}$ ont une espérance nulle et une variance $b^{(1)}, b^{(2)}, \dots, b^{(123)}$, respectivement. Le paramètre

$$m = \mathrm{E}[X_{ijkt}] = \mathrm{E}\Big[\mu_{123}(\Theta_i^{(1)}, \Theta_j^{(2)}, \Theta_k^{(3)})\Big]$$

est la moyenne collective du portefeuille. La variance de $\Xi_{ijkt}^{(1234)}$ est s^2/w_{ijkt}.

L'assuré représenté par la variable aléatoire X_{ijkt} est donc à la fois affecté par trois facteurs de risque séparément (variables aléatoires $\Xi_i^{(1)}, \Xi_j^{(2)}, \Xi_k^{(3)}$) ainsi que par l'interaction entre deux de ces facteurs (variables $\Xi_{ij}^{(12)}, \Xi_{ik}^{(13)}$ et $\Xi_{jk}^{(23)}$) et l'interaction entre les trois facteurs (variable $\Xi_{ijk}^{(123)}$). La variable aléatoire $\Xi_{ijkt}^{(1234)}$ ne fait qu'incorporer la dimension temporelle au modèle. Il importe de remarquer ici que la cellule (ijk) identifie uniquement un assuré ou, autrement dit, que la cellule ne contient qu'un seul assuré. Si d'aventure une même cellule devait contenir plusieurs assurés, une modélisation que l'on pourrait qualifier de non symétrique s'avérerait nécessaire. Un quatrième facteur de risque n'interagissant pas isolément avec les trois premiers devrait en effet être introduit :

$$\begin{aligned} X_{ijklt} = m &+ \Xi_i^{(1)} + \Xi_j^{(2)} + \Xi_k^{(3)} + \Xi_{ij}^{(12)} \\ &+ \Xi_{ik}^{(13)} + \Xi_{jk}^{(23)} + \Xi_{ijk}^{(123)} + \Xi_{ijkl}^{(1234)} + \Xi_{ijklt}^{(12345)}. \end{aligned}$$

On pourrait dire que cette formulation présente un mélange de modèle à classification croisée et de modèle hiérarchique. La présente étude ne s'attarde toutefois pas à de tels cas.

Le point central en théorie de la crédibilité consiste à estimer la prime de risque d'un assuré — la prime idéale à lui charger — ou encore, ce qui revient mathématiquement au même, à estimer la sinistralité de l'année qui suit à partir de l'expérience observée jusqu'à ce jour. Le théorème suivant expose cet *estimateur de crédibilité* pour le modèle CCC.

Théorème 3.1. *Soit X_{ijkt} le ratio d'expérience d'un assuré tel que défini en* (3.1). *Le meilleur estimateur — au sens des moindres carrés — de $X_{ijk,T_{ijk}+1}$*

[1]La connaissance du modèle « complet » n'est pas essentielle dans ce qui va suivre. Celui-ci ne sera donc présenté qu'avec les formules générales, au chapitre 6.

est alors

$$\begin{aligned}\widehat{X}_{ijk,T_{ijk}+1} &= m + z_{ijk}^{(123)}(X_{ijkw} - m) \\ &\quad + (1 - z_{ijk}^{(123)})\left(\widehat{\Xi}_i^{(1)} + \widehat{\Xi}_j^{(2)} + \widehat{\Xi}_k^{(3)} + \widehat{\Xi}_{ij}^{(12)} + \widehat{\Xi}_{ik}^{(13)} + \widehat{\Xi}_{jk}^{(23)}\right), \qquad (3.2)\end{aligned}$$

où

$$\begin{aligned}\widehat{\Xi}_i^{(1)} &= z_i^{(1)}(X_{izw} - m) \\ &\quad - z_i^{(1)} \sum_{j=1}^{J}\sum_{k=1}^{K} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} \frac{z_{ijk}^{(123)}}{z_{ij\Sigma}^{(123)}} \left(\widehat{\Xi}_j^{(2)} + \widehat{\Xi}_k^{(3)} + \widehat{\Xi}_{ik}^{(13)} + \widehat{\Xi}_{jk}^{(23)}\right) \\ &= z_i^{(1)}(X_{izw} - m) - z_i^{(1)} \sum_{j=1}^{J} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} \widehat{\Xi}_j^{(2)} \\ &\quad - z_i^{(1)} \sum_{j=1}^{J}\sum_{k=1}^{K} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} \frac{z_{ijk}^{(123)}}{z_{ij\Sigma}^{(123)}} \left(\widehat{\Xi}_k^{(3)} + \widehat{\Xi}_{ik}^{(13)} + \widehat{\Xi}_{jk}^{(23)}\right), \qquad (3.3) \\ \widehat{\Xi}_{ij}^{(12)} &= z_{ij}^{(12)}(X_{ijzw} - m) \\ &\quad - z_{ij}^{(12)} \sum_{k=1}^{K} \frac{z_{ijk}^{(123)}}{z_{ij\Sigma}^{(123)}} \left(\widehat{\Xi}_i^{(1)} + \widehat{\Xi}_j^{(2)} + \widehat{\Xi}_k^{(3)} + \widehat{\Xi}_{ik}^{(13)} + \widehat{\Xi}_{jk}^{(23)}\right) \\ &= z_{ij}^{(12)}(X_{ijzw} - m) - z_{ij}^{(12)}\left(\widehat{\Xi}_i^{(1)} + \widehat{\Xi}_j^{(2)}\right) \\ &\quad - z_{ij}^{(12)} \sum_{k=1}^{K} \frac{z_{ijk}^{(123)}}{z_{ij\Sigma}^{(123)}} \left(\widehat{\Xi}_k^{(3)} + \widehat{\Xi}_{ik}^{(13)} + \widehat{\Xi}_{jk}^{(23)}\right), \qquad (3.4)\end{aligned}$$

avec des expressions similaires à (3.3) *et* (3.4) *pour* $\widehat{\Xi}_j^{(2)}$, $\widehat{\Xi}_k^{(3)}$, $\widehat{\Xi}_{ik}^{(13)}$, *et* $\widehat{\Xi}_{jk}^{(23)}$. *Enfin,*

$$z_i^{(1)} = \frac{z_{i\Sigma}^{(12)}\, b^{(1)}}{z_{i\Sigma}^{(12)}\, b^{(1)} + b^{(12)}}, \qquad z_{ij}^{(12)} = \frac{z_{ij\Sigma}^{(123)}\, b^{(12)}}{z_{ij\Sigma}^{(123)}\, b^{(12)} + b^{(123)}},$$

$$z_{ijk}^{(123)} = \frac{w_{ijk\Sigma}\, b^{(123)}}{w_{ijk\Sigma}\, b^{(123)} + s^2}$$

et

$$X_{izw} = \sum_{j=1}^{J} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} X_{ijzw}, \qquad X_{ijzw} = \sum_{k=1}^{K} \frac{z_{ijk}^{(123)}}{z_{ij\Sigma}^{(123)}} X_{ijkw},$$

$$X_{ijkw} = \sum_{t=1}^{T_{ijk}} \frac{w_{ijkt}}{w_{ijk\Sigma}} X_{ijkt}.$$

L'estimateur de crédibilité (3.2) est donc égal à la moyenne collective m plus deux termes d'ajustement. Le premier terme d'ajustement correspond à la différence entre la moyenne (pondérée) de l'assuré et la moyenne collective, différence pondérée par un facteur de crédibilité $0 \leq z_{ijk}^{(123)} \leq 1$. Le complément du facteur de crédibilité est affecté au second terme d'ajustement, constitué de la somme des estimateurs de crédibilité des effets aléatoires $\Xi_i^{(1)}, \dots, \Xi_{jk}^{(23)}$. Les estimateurs de crédibilité de ces variables ne se présentent pas sous forme explicite, mais bien comme la solution d'un système d'équations linéaires dont (3.3) et (3.4) sont deux des six composants. (On remarquera que la variable $\Xi_{ijk}^{(123)}$ n'apparaît pas dans l'expression de l'estimateur de crédibilité.)

Dans les parties droites de (3.3) et (3.4), on retrouve la somme de tous les effets aléatoires, à l'exception, d'une part, de l'effet se trouvant déjà à gauche de l'équation et, d'autre part, de tous les effets des niveaux d'interaction « immédiatement supérieurs ». Par convention, le niveau d'interaction immédiatement supérieur est celui impliquant le plus petit facteur (dans sa représentation numérique) non représenté dans le niveau d'interaction précédent. Par exemple, dans l'expression de $\widehat{\Xi}_{jk}^{(23)}$ sera absente la variable $\widehat{\Xi}_{ijk}^{(123)}$, alors que de celle de $\widehat{\Xi}_k^{(3)}$ seront absentes $\widehat{\Xi}_{ik}^{(13)}$ et $\widehat{\Xi}_{ijk}^{(123)}$.

La preuve de ce théorème étant plutôt longue, elle fera l'objet de toute la section suivante.

3.2 Preuve du théorème 3.1

Tel que mentionné précédemment, les formules seront démontrées à l'aide de projections dans des espaces de Hilbert. Les variables aléatoires Ξ étant par hypothèse de carré intégrables, tout ensemble de fonctions (mesurables) de ces variables forme un sous-espace de $\mathcal{L}^2$. Il en va de même des variables X_{ijkt}.

L'algorithme de projection est inspiré de Bühlmann & Jewell (1987) et fait appel à la propriété (P2) de l'opérateur projection (page 11). Sommairement, la procédure consiste à projeter la variable que l'on cherche à estimer d'abord sur un espace restreint où les calculs s'avèrent relativement simples, puis ensuite sur l'espace désiré.

On cherche ici à calculer $\widehat{X}_{ijk,T_{ijk}+1} = \text{pro}(X_{ijk,T_{ijk}+1} | \mathcal{L}\{\underline{X}, m\})$. Par la propriété (P2) de l'opérateur projection, on peut écrire :

$$\widehat{X}_{ijk,T_{ijk}+1} = \text{pro}(\text{pro}(X_{ijk,T_{ijk}+1} | \mathcal{L}\{\underline{X}_{ijk}, \Xi_i^{(1)}, \Xi_j^{(2)}, \Xi_k^{(3)}, \Xi_{ij}^{(12)}, \Xi_{ik}^{(13)}, \Xi_{jk}^{(23)}, \Xi_{ijk}^{(123)}, m\}) | \mathcal{L}\{\underline{X}, m\}).$$

C'est par l'hypothèse (CCC3) d'indépendance conditionnelle entre les assurés que l'on peut se limiter, à chaque niveau, aux variables de la même « filière » que la variable à projeter, c'est-à-dire aux variables portant le ou les mêmes indices (ijk dans le cas présent). La projection intérieure est de la forme

$$\begin{aligned}\alpha m &+ \beta^{(1)} \Xi_i^{(1)} + \beta^{(2)} \Xi_j^{(2)} + \beta^{(3)} \Xi_k^{(3)} + \beta^{(12)} \Xi_{ij}^{(12)} \\ &+ \beta^{(13)} \Xi_{ik}^{(13)} + \beta^{(23)} \Xi_{jk}^{(23)} + \beta^{(123)} \Xi_{ijk}^{(123)} + \sum_{t=1}^{T_{ijk}} \kappa_t X_{ijkt},\end{aligned}$$

où α, les $\beta^{(\cdot)}$ et les κ_t sont des constantes à déterminer. Les équations normales (2.7) sont donc très simples :

$$\begin{aligned}&\alpha + \kappa_\Sigma = 1 \\ &\beta^{(i_1 \dots i_q)} + \kappa_\Sigma = 1, && q = 1,2,3,\ i_1 < \cdots < i_q \in (1,2,3), \\ &\kappa_u \frac{s^2}{w_{ijku}} = 0, && u = 1, \dots, T_{ijk}.\end{aligned}$$

De ce système, on déduit aisément $\kappa_u = 0$ pour tout u et $\alpha = \beta^{(i_1 \dots i_q)} = 1$, d'où

$$\begin{aligned}\widehat{X}_{ijk,t+1} = m &+ \widehat{\Xi}_i^{(1)} + \widehat{\Xi}_j^{(2)} + \widehat{\Xi}_k^{(3)} + \widehat{\Xi}_{ij}^{(12)} \\ &+ \widehat{\Xi}_{ik}^{(13)} + \widehat{\Xi}_{jk}^{(23)} + \widehat{\Xi}_{ijk}^{(123)},\end{aligned} \tag{3.5}$$

où $\widehat{\Xi}_i^{(1)} = \text{pro}(\Xi_i^{(1)} | \mathcal{L}\{\underline{X}, m\})$, etc., est l'estimateur de crédibilité de l'effet aléatoire $\Xi_i^{(1)}$.

Ce résultat pourrait aussi être obtenu en réalisant que l'estimateur de crédibilité de $X_{ijk,T_{ijk}+1}$ est égal à celui de $m + \Xi_i^{(1)} + \Xi_j^{(2)} + \Xi_k^{(3)} + \Xi_{ij}^{(12)} + \Xi_{ik}^{(13)} + \Xi_{jk}^{(23)} + \Xi_{ijk}^{(123)}$, puis en utilisant la propriété (P1) de l'opérateur projection.

Prouver le théorème requiert donc de calculer l'estimateur de crédibilité de chaque variable Ξ, soit leur projection sur $\mathcal{L}\{\underline{X}, m\}$. C'est là le propos des sous-sections qui suivent.

3.2.1 Projection de $\Xi_{ijk}^{(123)}$

Supposons que les indices i, j et k représentent ici des valeurs particulières des indices de $\Xi^{(123)}$. Par la propriété (P2),

$$\begin{aligned}\widehat{\Xi}_{ijk}^{(123)} &= \text{pro}(\Xi_{ijk}^{(123)} | \mathcal{L}\{\underline{X}, m\}) \\ &= \text{pro}(\text{pro}(\Xi_{ijk}^{(123)} | \mathcal{L}\{\underline{X}_{ijk}, \Xi_i^{(1)}, \Xi_j^{(2)}, \Xi_k^{(3)}, \\ &\qquad \Xi_{ij}^{(12)}, \Xi_{ik}^{(13)}, \Xi_{jk}^{(23)}, m\}) | \mathcal{L}\{\underline{X}, m\}).\end{aligned} \tag{3.6}$$

On doit donc d'abord calculer la projection intérieure

$$\mathrm{pro}(\Xi_{ijk}^{(123)}|\mathcal{L}\{\underline{X}_{ijk}, \Xi_i^{(1)}, \Xi_j^{(2)}, \Xi_k^{(3)}, \Xi_{ij}^{(12)}, \Xi_{ik}^{(13)}, \Xi_{jk}^{(23)}, m\}),$$

qui est de la forme

$$\begin{aligned}&\alpha m + \beta^{(1)}\Xi_i^{(1)} + \beta^{(2)}\Xi_j^{(2)} + \beta^{(3)}\Xi_k^{(3)} \\ &\qquad + \beta^{(12)}\Xi_{ij}^{(12)} + \beta^{(13)}\Xi_{ik}^{(13)} + \beta^{(23)}\Xi_{jk}^{(23)} + \sum_{t=1}^{T_{ijk}} \kappa_t X_{ijkt}.\end{aligned}$$

Les équations normales sont alors les suivantes :

$$\begin{aligned}\alpha + \kappa_\Sigma &= 0, \\ \beta^{(i_1\ldots i_q)} + \kappa_\Sigma &= 0, \qquad q = 1, 2,\ i_1 < \cdots < i_q \in (1,2,3), \\ \kappa_\Sigma b^{(123)} + \kappa_u \frac{s^2}{w_{ijku}} &= b^{(123)}, \qquad u = 1, \ldots, T_{ijk}.\end{aligned} \tag{3.7}$$

Les systèmes d'équations de ce type interviennent fréquemment en théorie de la crédibilité, et la technique pour les résoudre est toujours sensiblement la même. Se référer, par exemple, à Goovaerts & Hoogstad (1987) à cet effet. Ainsi, à partir de (3.7), on obtient :

$$\begin{aligned}\kappa_t &= \frac{w_{ijkt}}{w_{ijk\Sigma}} \frac{w_{ijk\Sigma} b^{(123)}}{w_{ijk\Sigma} b^{(123)} + s^2} = \frac{w_{ijkt}}{w_{ijk\Sigma}} z_{ijk}^{(123)}, \quad t = 1, \ldots, T_{ijk}, \\ \kappa_\Sigma &= z_{ijk}^{(123)} = -\beta^{(1)} = -\beta^{(2)} = -\beta^{(3)} = -\beta^{(13)} = -\beta^{(12)} = -\beta^{(23)},\end{aligned}$$

d'où,

$$\begin{aligned}&\mathrm{pro}(\Xi_{ijk}^{(123)}|\mathcal{L}\{\underline{X}_{ijk}, \Xi_i^{(1)}, \Xi_j^{(2)}, \Xi_k^{(3)}, \Xi_{ij}^{(12)}, \Xi_{ik}^{(13)}, \Xi_{jk}^{(23)}, m\}) = \\ &\quad z_{ijk}^{(123)}(X_{ijkw} - m) - z_{ijk}^{(123)} \left(\Xi_i^{(1)} + \Xi_j^{(2)} + \Xi_k^{(3)} + \Xi_{ij}^{(12)} + \Xi_{ik}^{(13)} + \Xi_{jk}^{(23)}\right)\end{aligned}$$

et enfin, par (3.6),

$$\begin{aligned}\widehat{\Xi}_{ijk}^{(123)} &= z_{ijk}^{(123)}(X_{ijkw} - m) \\ &\quad - z_{ijk}^{(123)} \left(\widehat{\Xi}_i^{(1)} + \widehat{\Xi}_j^{(2)} + \widehat{\Xi}_k^{(3)} + \widehat{\Xi}_{ij}^{(12)} + \widehat{\Xi}_{ik}^{(13)} + \widehat{\Xi}_{jk}^{(23)}\right).\end{aligned} \tag{3.8}$$

En insérant (3.8) dans (3.5), on obtient (3.2), l'expression de l'estimateur de crédibilité.

3.2.2 Projection de $\varXi_{ij}^{(12)}$

La première partie de ce calcul est tout à fait similaire au précédent :

$$\begin{aligned}\widehat{\varXi}_{ij}^{(12)} &= \mathrm{pro}(\varXi_{ij}^{(12)}|\mathcal{L}\{\underline{X}, m\}) \\ &= \mathrm{pro}(\mathrm{pro}(\varXi_{ij}^{(12)}|\mathcal{L}\{\underline{X}_{ij}, \varXi_i^{(1)}, \varXi_j^{(2)}, \underline{\varXi}^{(3)}, \\ &\qquad \underline{\varXi}_i^{(13)}, \underline{\varXi}_j^{(23)}, \underline{\varXi}_{ij}^{(123)}, m\})|\mathcal{L}\{\underline{X}, m\}).\end{aligned} \tag{3.9}$$

Comme précédemment, on calcule d'abord la projection intérieure de (3.9), qui sera de la forme

$$\begin{aligned}\alpha m + \beta^{(1)}\varXi_i^{(1)} + \beta^{(2)}\varXi_j^{(2)} + \sum_{k=1}^{K}\beta_k^{(3)}\varXi_k^{(3)} + \sum_{k=1}^{K}\beta_k^{(13)}\varXi_{ik}^{(13)} \\ + \sum_{k=1}^{K}\beta_k^{(23)}\varXi_{jk}^{(23)} + \sum_{k=1}^{K}\beta_k^{(123)}\varXi_{ijk}^{(123)} + \sum_{k=1}^{K}\sum_{t=1}^{T_{ijk}}\kappa_{kt}X_{ijkt}.\end{aligned}$$

Les équations normales sont alors, pour $h = 1, \ldots, K$ et $u = 1, \ldots, T_{ijh}$,

$$\begin{aligned} \alpha + \kappa_\Sigma &= 0, & \beta_h^{(13)} + \kappa_{h\Sigma} &= 0, \\ \beta^{(1)} + \kappa_\Sigma &= 0, & \beta_h^{(23)} + \kappa_{h\Sigma} &= 0, \\ \beta^{(2)} + \kappa_\Sigma &= 0, & \beta_h^{(123)} + \kappa_{h\Sigma} &= 0, \\ \beta_h^{(3)} + \kappa_{h\Sigma} &= 0, & \kappa_\Sigma b^{(12)} + \kappa_{hu}\frac{s^2}{w_{ijru}} &= b^{(12)}. \end{aligned}$$

Toujours par les mêmes manipulations, on trouve aisément

$$\begin{aligned}\kappa_{kt} &= w_{ijkt}\frac{b^{(12)}}{w_{ij\Sigma}b^{(12)} + s^2}, \\ \kappa_{k\Sigma} &= w_{ijk\Sigma}\frac{b^{(12)}}{w_{ij\Sigma}b^{(12)} + s^2} = -\beta_k^{(3)} = -\beta_k^{(13)} = -\beta_k^{(23)} = -\beta_k^{(123)}, \\ \kappa_\Sigma &= w_{ij\Sigma}\frac{b^{(12)}}{w_{ij\Sigma}b^{(12)} + s^2} = -\alpha = -\beta^{(1)} = -\beta^{(2)},\end{aligned}$$

d'où, par (3.9),

$$\begin{aligned}\widehat{\varXi}_{ij}^{(12)} = \frac{b^{(12)}}{w_{ij\Sigma}b^{(12)} + s^2}\sum_{k=1}^{K} w_{ijk\Sigma}\Big(X_{ijkw} - m \\ - \widehat{\varXi}_i^{(1)} - \widehat{\varXi}_j^{(2)} - \widehat{\varXi}_k^{(3)} - \widehat{\varXi}_{ik}^{(13)} - \widehat{\varXi}_{jk}^{(23)} - \widehat{\varXi}_{ijk}^{(123)}\Big).\end{aligned}$$

Afin d'obtenir le résultat souhaité, il s'agit maintenant de remplacer $\widehat{\Xi}_{ijk}^{(123)}$ par (3.8), sa valeur calculée auparavant. Cette substitution faite et en réarrangeant un peu les termes, on obtient

$$\begin{aligned}\widehat{\Xi}_{ij}^{(12)} = \frac{b^{(12)}}{w_{ij\Sigma}b^{(12)}+s^2}\sum_{k=1}^{K} w_{ijk\Sigma}\Big[(1-z_{ijk}^{(123)})\Big(X_{ijkw}-m \\ -\widehat{\Xi}_i^{(1)}-\widehat{\Xi}_j^{(2)}-\widehat{\Xi}_k^{(3)}-\widehat{\Xi}_{ik}^{(13)}-\widehat{\Xi}_{jk}^{(23)}\Big)+z_{ijk}^{(123)}\widehat{\Xi}_{ij}^{(12)}\Big],\end{aligned}$$

d'où

$$\begin{aligned}\widehat{\Xi}_{ij}^{(12)}\left[1-\frac{b^{(12)}}{w_{ij\Sigma}b^{(12)}+s^2}\sum_{k=1}^{K} w_{ijk\Sigma}z_{ijk}^{(123)}\right] = \\ = \frac{b^{(12)}}{w_{ij\Sigma}b^{(12)}+s^2}\sum_{k=1}^{K} w_{ijk\Sigma}(1-z_{ijk}^{(123)})\Big(X_{ijkw}-m \\ -\widehat{\Xi}_i^{(1)}-\widehat{\Xi}_j^{(2)}-\widehat{\Xi}_k^{(3)}-\widehat{\Xi}_{ik}^{(13)}-\widehat{\Xi}_{jk}^{(23)}\Big).\end{aligned}$$

En multipliant de part et d'autre par $w_{ij\Sigma}b^{(12)}+s^2$ et en isolant de nouveau $\widehat{\Xi}_{ij}^{(12)}$, on obtient :

$$\begin{aligned}\widehat{\Xi}_{ij}^{(12)} = \frac{b^{(12)}}{b^{(12)}\sum_k w_{ijk\Sigma}(1-z_{ijk}^{(123)})+s^2}\sum_{k=1}^{K} w_{ijk\Sigma}(1-z_{ijk}^{(123)})\Big(X_{ijkw}-m \\ -\Xi_i^{(1)}-\Xi_j^{(2)}-\Xi_k^{(3)}-\Xi_{ik}^{(13)}-\Xi_{jk}^{(23)}\Big).\end{aligned}$$

Or, on vérifie par quelques manipulations algébriques que

$$\frac{b^{(12)}w_{ijk\Sigma}(1-z_{ijk}^{(123)})}{b^{(12)}\sum_k w_{ijk\Sigma}(1-z_{ijk}^{(123)})+s^2} = \frac{z_{ij\Sigma}^{(123)}b^{(12)}}{z_{ij\Sigma}^{(123)}b^{(12)}+b^{(123)}}\frac{z_{ijk}^{(123)}}{z_{ij\Sigma}^{(123)}},$$

d'où, finalement,

$$\begin{aligned}\widehat{\Xi}_{ij}^{(12)} &= z_{ij}^{(12)}(X_{ijzw}-m) \\ &\quad - z_{ij}^{(12)}\sum_{k=1}^{K}\frac{z_{ijk}^{(123)}}{z_{ij\Sigma}^{(123)}}\left(\widehat{\Xi}_i^{(1)}+\widehat{\Xi}_j^{(2)}+\widehat{\Xi}_k^{(3)}+\widehat{\Xi}_{ik}^{(13)}+\widehat{\Xi}_{jk}^{(23)}\right),\end{aligned}$$

avec

$$z_{ij}^{(12)} = \frac{z_{ij\Sigma}^{(123)}b^{(12)}}{z_{ij\Sigma}^{(123)}b^{(12)}+b^{(123)}}, \qquad X_{ijzw} = \sum_{k=1}^{K}\frac{z_{ijk}^{(123)}}{z_{ij\Sigma}^{(123)}}X_{ijkw}.$$

3.2.3 Projection de $\Xi_i^{(1)}$

La procédure est toujours sensiblement la même :

$$\begin{aligned}\widehat{\Xi}_i^{(1)} &= \mathrm{pro}(\Xi_i^{(1)}|\mathcal{L}\{\underline{X}, m\}) \\ &= \mathrm{pro}(\mathrm{pro}(\Xi_i^{(1)}|\mathcal{L}\{\underline{X}_i, \underline{\Xi}^{(2)}, \underline{\Xi}^{(3)}, \\ &\qquad \underline{\Xi}_i^{(12)}, \underline{\Xi}_i^{(13)}, \underline{\Xi}^{(23)}, \underline{\Xi}_i^{(123)}\})|\mathcal{L}\{\underline{X}, m\})\end{aligned} \tag{3.10}$$

et la projection intérieure est de la forme

$$\begin{aligned}\alpha m &+ \sum_{j=1}^{J} \beta_j^{(2)} \Xi_j^{(2)} + \sum_{k=1}^{K} \beta_k^{(3)} \Xi_k^{(3)} + \sum_{j=1}^{J} \beta_j^{(12)} \Xi_{ij}^{(12)} + \sum_{k=1}^{K} \beta_k^{(13)} \Xi_{ik}^{(13)} \\ &+ \sum_{j=1}^{J}\sum_{k=1}^{K} \beta_{jk}^{(23)} \Xi_{jk}^{(23)} + \sum_{j=1}^{J}\sum_{k=1}^{K} \beta_{jk}^{(123)} \Xi_{ijk}^{(123)} + \sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{t=1}^{T_{ijk}} \kappa_{jkt} X_{ijkt}.\end{aligned}$$

Les équations normales sont les suivantes, pour $g = 1, \ldots, J$, $h = 1, \ldots, K$ et $u = 1, \ldots, T_{igh}$:

$$\begin{aligned}\alpha + \kappa_\Sigma &= 0, & \beta_h^{(13)} + \kappa_{h\Sigma} &= 0 \\ \beta_g^{(2)} + \kappa_{g\Sigma} &= 0, & \beta_{gh}^{(23)} + \kappa_{gh\Sigma} &= 0 \\ \beta_h^{(3)} + \kappa_{h\Sigma} &= 0, & \beta_{gh}^{(123)} + \kappa_{gh\Sigma} &= 0 \\ \beta_g^{(12)} + \kappa_{g\Sigma} &= 0, & \kappa_\Sigma b^{(1)} + \kappa_{ghu}\frac{s^2}{w_{ighu}} &= b^{(1)}.\end{aligned}$$

On obtient donc

$$\begin{aligned}\kappa_{jkt} &= w_{ijkt}\frac{b^{(1)}}{w_{i\Sigma}b^{(1)} + s^2}, \\ \kappa_{jk\Sigma} &= w_{ijk\Sigma}\frac{b^{(1)}}{w_{i\Sigma}b^{(1)} + s^2} = -\beta_{jk}^{(123)} = -\beta_{jk}^{(23)}, \\ \kappa_{j\Sigma} &= w_{ij\Sigma}\frac{b^{(1)}}{w_{i\Sigma}b^{(1)} + s^2} = -\beta_j^{(12)} = -\beta_j^{(2)}, \\ \kappa_{k\Sigma} &= w_{ik\Sigma}\frac{b^{(1)}}{w_{i\Sigma}b^{(1)} + s^2} = -\beta_k^{(13)} = -\beta_k^{(3)},\end{aligned}$$

d'où

$$\begin{aligned}\widehat{\Xi}_i^{(1)} = \frac{b^{(1)}}{w_{i\Sigma}b^{(1)} + s^2} \sum_{j=1}^{J}\sum_{k=1}^{K} w_{ijk\Sigma}\Big(X_{ijkw} - m \\ - \widehat{\Xi}_j^{(2)} - \widehat{\Xi}_k^{(3)} - \widehat{\Xi}_{ij}^{(12)} - \widehat{\Xi}_{ik}^{(13)} - \widehat{\Xi}_{jk}^{(23)} - \widehat{\Xi}_{ijk}^{(123)}\Big).\end{aligned} \tag{3.11}$$

En remplaçant comme ci-dessus $\widehat{\Xi}^{(123)}_{ijk}$ par (3.8) et en réarrangeant un peu les termes, on trouve

$$\begin{aligned}\widehat{\Xi}^{(1)}_i = \frac{b^{(1)}}{w_{i\Sigma}b^{(1)} + s^2} \sum_{j=1}^{J}\sum_{k=1}^{K} w_{ijk\Sigma} \Big[(1 - z^{(123)}_{ijk})\Big(X_{ijkw} - m \\ - \widehat{\Xi}^{(2)}_j - \widehat{\Xi}^{(3)}_k - \widehat{\Xi}^{(12)}_{ij} - \widehat{\Xi}^{(13)}_{ik} - \widehat{\Xi}^{(23)}_{jk}\Big) + z^{(123)}_{ijk}\widehat{\Xi}^{(1)}_i\Big],\end{aligned}$$

d'où

$$\begin{aligned}\widehat{\Xi}^{(1)}_i \left[1 - \frac{b^{(1)}}{w_{i\Sigma}b^{(1)} + s^2} \sum_{j=1}^{J}\sum_{k=1}^{K} w_{ijk\Sigma} z^{(123)}_{ijk}\right] = \\ \frac{b^{(1)}}{w_{i\Sigma}b^{(1)} + s^2} \sum_{j=1}^{J}\sum_{k=1}^{K} w_{ijk\Sigma}(1 - z^{(123)}_{ijk})\Big(X_{ijkw} - m \\ - \widehat{\Xi}^{(2)}_j - \widehat{\Xi}^{(3)}_k - \widehat{\Xi}^{(12)}_{ij} - \widehat{\Xi}^{(13)}_{ik} - \widehat{\Xi}^{(23)}_{jk}\Big).\end{aligned}$$

Par les mêmes manipulations que précédemment et en notant que

$$\frac{b^{(1)} w_{ijk\Sigma}(1 - z^{(123)}_{ijk})}{b^{(1)} \sum_j \sum_k w_{ijk\Sigma}(1 - z^{(123)}_{ijk}) + s^2} = \frac{z^{(123)}_{i\Sigma} b^{(1)}}{z^{(123)}_{i\Sigma} b^{(1)} + b^{(123)}} \frac{z^{(123)}_{ijk}}{z^{(123)}_{i\Sigma}},$$

on obtient

$$\begin{aligned}\widehat{\Xi}^{(1)}_i = \frac{b^{(1)}}{z^{(123)}_{i\Sigma} b^{(1)} + b^{(123)}} \sum_{j=1}^{J}\sum_{k=1}^{K} z^{(123)}_{ijk}\Big(X_{ijkw} - m \\ - \widehat{\Xi}^{(2)}_j - \widehat{\Xi}^{(3)}_k - \widehat{\Xi}^{(12)}_{ij} - \widehat{\Xi}^{(13)}_{ik} - \widehat{\Xi}^{(23)}_{jk}\Big). \quad (3.12)\end{aligned}$$

Contrairement à la projection de $\Xi^{(12)}$, cette forme n'est pas encore celle cherchée. Pour obtenir le résultat souhaité, il s'agit de remplacer maintenant $\widehat{\Xi}^{(12)}_{ij}$ par sa valeur calculée précédemment. Ce sont ces remplacements successifs des estimateurs de crédibilité des interactions entre les facteurs de risque par leurs valeurs respectives qui font que les variables $\widehat{\Xi}^{(12)}_{ij}$ et $\widehat{\Xi}^{(123)}_{ijk}$ n'apparaissent pas dans (3.3) et que $\widehat{\Xi}^{(123)}_{ijk}$ n'apparaît pas dans (3.4). On notera qu'il eut été possible de faire d'autres remplacements, seulement on a choisi par convention de le faire dans l'ordre croissant d'apparition des nouveaux facteurs de risque. Cette convention interviendra à plusieurs reprises dans la suite.

Le remplacement de $\widehat{\Xi}_{ij}^{(12)}$ par sa valeur suit exactement les mêmes lignes que ce qui a été fait jusqu'à maintenant. Ainsi, on a

$$\begin{aligned}\widehat{\Xi}_i^{(1)} = \frac{b^{(1)}}{z_{i\Sigma}^{(123)} b^{(1)} + b^{(123)}} \Bigg[&\sum_{j=1}^{J} z_{ij\Sigma}^{(123)} (1 - z_{ij}^{(12)})(X_{ijzw} - m) \\ &- \sum_{j=1}^{J} \sum_{k=1}^{K} z_{ijk}^{(123)} (1 - z_{ij}^{(12)}) \left(\widehat{\Xi}_j^{(2)} - \widehat{\Xi}_k^{(3)} - \widehat{\Xi}_{ik}^{(13)} - \widehat{\Xi}_{jk}^{(23)} \right) \\ &+ \sum_{j=1}^{J} z_{ij\Sigma}^{(123)} z_{ij}^{(12)} \widehat{\Xi}_i^{(1)} \Bigg],\end{aligned}$$

puis on peut vérifier que

$$\frac{b^{(1)} z_{ij\Sigma}^{(123)} (1 - z_{ij}^{(12)})}{b^{(1)} \sum_{j=1}^{J} z_{ij\Sigma}^{(123)} (1 - z_{ij}^{(12)}) + b^{(123)}} = \frac{z_{i\Sigma}^{(12)} b^{(1)}}{z_{i\Sigma}^{(12)} b^{(1)} + b^{(12)}} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}},$$

d'où, finalement, on peut écrire $\widehat{\Xi}_i^{(1)}$ sous la forme

$$\begin{aligned}\widehat{\Xi}_i^{(1)} &= z_i^{(1)} (X_{izw} - m) \\ &\quad - z_i^{(1)} \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} \frac{z_{ijk}^{(123)}}{z_{ij\Sigma}^{(123)}} \left(\widehat{\Xi}_j^{(2)} + \widehat{\Xi}_k^{(3)} + \widehat{\Xi}_{ik}^{(13)} + \widehat{\Xi}_{jk}^{(23)} \right),\end{aligned}$$

avec

$$z_i^{(1)} = \frac{z_{i\Sigma}^{(12)} b^{(1)}}{z_{i\Sigma}^{(12)} b^{(1)} + b^{(12)}}, \qquad X_{izw} = \sum_{j=1}^{J} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} X_{ijzw}.$$

Les projections de $\Xi_{ik}^{(13)}, \Xi_{jk}^{(23)}, \Xi_j^{(2)}$ et $\Xi_k^{(3)}$ sont similaires à ce qui vient d'être fait. Ceci complète donc la preuve du théorème. □

Afin de pouvoir utiliser ces formules dans une application pratique, il faut disposer d'estimateurs pour les différents paramètres de structure. Avant d'aborder, au chapitre suivant, l'estimation optimale de ces paramètres, on présente d'abord les estimateurs proposés par Dannenburg (1995), puis des pseudo-estimateurs des variances $b^{(\cdot)}$ pour des modèles à deux et trois facteurs.

La présentation de ces estimateurs requiert toutefois quelques relations de covariance entre certaines moyennes pondérées des observations. Ces relations seront fréquemment réutilisées par la suite.

3.3 Relations de covariance

Les relations de covariance qui nous intéresseront seront celles entre deux moyennes des observations pondérées par les facteurs de crédibilité. Ce choix sera justifié plus loin. On présente ici les relations de covariance du modèle à deux facteurs en plus de celles du modèle à trois facteurs.

Tout d'abord, étant donné les hypothèses de l'approche des composants de variance, on voit aisément que, pour un modèle à deux facteurs,

$$\operatorname{Cov}(X_{ijt}, X_{klu}) = \delta_{ik} b^{(1)} + \delta_{jl} b^{(2)} + \delta_{ij,kl} b^{(12)} + \delta_{ijt,klu} \frac{s^2}{w_{ijt}} \tag{3.13}$$

et, pour un modèle à trois facteurs,

$$\begin{aligned}\operatorname{Cov}(X_{ijkt}, X_{fghu}) &= \delta_{if} b^{(1)} + \delta_{jg} b^{(2)} + \delta_{kh} b^{(3)} + \delta_{ij,fg} b^{(12)} \\ &\quad + \delta_{ik,fh} b^{(13)} + \delta_{jk,gh} b^{(23)} + \delta_{ijk,fgh} b^{(123)} + \delta_{ijkt,fghu} \frac{s^2}{w_{ijkt}}.\end{aligned} \tag{3.14}$$

Les deux théorèmes suivants exposent les principales relations de covariances pour les modèles à deux et trois facteurs et introduisent la notation qui sera dorénavant utilisée pour identifier ces covariances.

Théorème 3.2 (Modèle à deux facteurs). *Dans un modèle CCC à deux facteurs,*

$$\begin{aligned}\operatorname{Cov}(X_{ijw}, X_{klw}) &= \sigma_{ijkl}^{(12)} \\ &= \delta_{ik} b^{(1)} + \delta_{jl} b^{(2)} + \delta_{ij,kl} \frac{b^{(12)}}{z_{ij}^{(12)}},\end{aligned} \tag{3.15}$$

$$\begin{aligned}\operatorname{Cov}(X_{izw}, X_{kzw}) &= \sigma_{ik}^{(1)} \\ &= \delta_{ik} \frac{b^{(1)}}{z_i^{(1)}} + b^{(2)} S_{ik}^{(1)},\end{aligned} \tag{3.16}$$

$$\begin{aligned}\operatorname{Cov}(X_{zjw}, X_{zlw}) &= \sigma_{jl}^{(2)} \\ &= b^{(1)} S_{jl}^{(2)} + \delta_{jl} \frac{b^{(2)}}{z_j^{(2)}},\end{aligned} \tag{3.17}$$

avec

$$S_{ik}^{(1)} = \sum_{j=1}^{J} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} \frac{z_{kj}^{(12)}}{z_{k\Sigma}^{(12)}},$$

$$S_{jl}^{(2)} = \sum_{i=1}^{I} \frac{z_{ij}^{(12)}}{z_{j\Sigma}^{(12)}} \frac{z_{il}^{(12)}}{z_{l\Sigma}^{(12)}}.$$

Preuve. On obtient (3.15) à partir de (3.13) en remarquant que $b^{(12)} + s^2/w_{ij\Sigma} = b^{(12)}/z_{ij}^{(12)}$. Quant à (3.16), on a

$$\begin{aligned}
\sigma_i^{(1)} &= \sum_{j=1}^{J}\sum_{l=1}^{J} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} \frac{z_{kl}^{(12)}}{z_{k\Sigma}^{(12)}} \sigma_{ijkl}^{(12)} \\
&= \delta_{ik} b^{(1)} + b^{(2)} \sum_{j=1}^{J} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} \frac{z_{kj}^{(12)}}{z_{k\Sigma}^{(12)}} + \delta_{ik} \frac{b^{(12)}}{z_{i\Sigma}^{(12)}} \\
&= \delta_{ik} \frac{b^{(1)}}{z_i^{(1)}} + b^{(2)} S_{ik}^{(1)}
\end{aligned}$$

puisque $b^{(1)} + b^{(12)}/z_{i\Sigma}^{(12)} = b^{(1)}/z_i^{(1)}$. La preuve de (3.17) est tout à fait symétrique. □

Théorème 3.3 (Modèle à trois facteurs). *Dans un modèle CCC à trois facteurs,*

$$\begin{aligned}
\mathrm{Cov}(X_{ijkw}, X_{fghw}) &= \sigma_{ijkfgh}^{(123)} \\
&= \delta_{if} b^{(1)} + \delta_{jg} b^{(2)} + \delta_{kh} b^{(3)} + \delta_{ij,fg} b^{(12)} \\
&\quad + \delta_{ik,fh} b^{(13)} + \delta_{jk,gh} b^{(23)} + \delta_{ijk,fgh} \frac{b^{(123)}}{z_{ijk}^{(123)}}
\end{aligned} \tag{3.18}$$

$$\begin{aligned}
\mathrm{Cov}(X_{ijzw}, X_{fgzw}) &= \sigma_{ijfg}^{(12)} \\
&= \delta_{if} b^{(1)} + \delta_{jg} b^{(2)} + \delta_{ij,fg} \frac{b^{(12)}}{z_{ij}^{(12)}} \\
&\quad + (b^{(3)} + \delta_{if} b^{(13)} + \delta_{jg} b^{(23)}) S_{ijfg}^{(12)}
\end{aligned} \tag{3.19}$$

$$\begin{aligned}
\mathrm{Cov}(X_{izw}, X_{fzw}) &= \sigma_{if}^{(1)} \\
&= \delta_{if} \frac{b^{(1)}}{z_i^{(1)}} + b^{(2)} S_{if}^{(1)} \\
&\quad + (b^{(3)} + \delta_{if} b^{(13)}) S_{izfz}^{(12)} + b^{(23)} S_{izf\star}^{(12)}
\end{aligned} \tag{3.20}$$

avec

$$\begin{aligned}
S_{ijfg}^{(12)} &= \sum_{k=1}^{K} \frac{z_{ijk}^{(123)}}{z_{ij\Sigma}^{(123)}} \frac{z_{fgk}^{(123)}}{z_{fg\Sigma}^{(123)}}, \\
S_{izfz}^{(12)} &= \sum_{j=1}^{J}\sum_{g=1}^{J} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} \frac{z_{fg}^{(12)}}{z_{f\Sigma}^{(12)}} S_{ijfg}^{(12)},
\end{aligned}$$

$$S^{(12)}_{izf\star} = \sum_{j=1}^{J} \frac{z^{(12)}_{ij}}{z^{(12)}_{i\Sigma}} \frac{z^{(12)}_{fj}}{z^{(12)}_{f\Sigma}} S^{(12)}_{ijfj},$$
$$S^{(1)}_{if} = \sum_{j=1}^{J} \frac{z^{(12)}_{ij}}{z^{(12)}_{i\Sigma}} \frac{z^{(12)}_{fj}}{z^{(12)}_{f\Sigma}}$$

et des relations similaires pour $\sigma^{(13)}_{ikfh}$, $\sigma^{(23)}_{jkgh}$, $\sigma^{(2)}_{jg}$ *et* $\sigma^{(3)}_{kh}$.

Preuve. En tous points similaire à celle du théorème précédent. □

3.4 Estimateurs de Dannenburg

Dans les formules du théorème 3.1, les paramètres m, s^2 et $b^{(1)}, \dots, b^{(123)}$ forment l'ensemble des paramètres de structure du modèle de crédibilité à classification croisée. Inconnus en pratique, ces paramètres doivent être estimés à partir du vecteur des données $\underline{X}$ pour rendre lesdites formules opérationnelles. Les premiers estimateurs présentés ci-dessous sont les versions pour un modèle à trois facteurs de ceux proposés à l'origine par Dannenburg (1995).

Un estimateur sans biais intuitif de la moyenne collective m est la moyenne pondérée par les poids naturels des ratios moyens de tous les assurés :

$$\widehat{m} = X_w = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \frac{w_{ijk\Sigma}}{w_\Sigma} X_{ijkw}. \tag{3.21}$$

C'est l'estimateur de la moyenne collective retenu par Dannenburg (1995).

Le paramètre s^2 est facile à estimer sans biais en généralisant simplement l'estimateur classique du modèle de Bühlmann–Straub :

$$\hat{s}^2 = \frac{1}{\sum_{ijk}(T_{ijk}-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{t=1}^{T_{ijk}} w_{ijkt}(X_{ijkt} - X_{ijkw})^2 \tag{3.22}$$
$$= \frac{1}{2\sum_{ijk}(T_{ijk}-1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \sum_{t=1}^{T_{ijk}} \sum_{u=1}^{T_{ijk}} \frac{w_{ijkt} w_{ijku}}{w_{ijk\Sigma}} (X_{ijkt} - X_{ijku})^2. \tag{3.23}$$

On verra au chapitre suivant que cet estimateur est optimal (à variance minimale) sous certaines hypothèses. Nous ne reviendrons donc pas d'ici là sur l'estimation du paramètre s^2 et il sera désormais sous-entendu que $\hat{s}^2$ est exclu de toute discussion sur la performance ou l'optimalité des estimateurs du modèle CCC. De plus, il sera pratique de réserver l'appellation « composants de variance » au seul ensemble de variances $b^{(1)}, \ldots, b^{(123)}$, même si s^2 est, formellement, un composant de variance.

L'estimation des composants de variance pose davantage problème. C'est d'ailleurs cette étape du processus d'estimation qui nous intéressera principalement par la suite. Les estimateurs ANOVA de la section 2.3 ne peuvent être utilisés directement vu la présence de poids dans notre modèle. Dannenburg (1995) a néanmoins repris l'idée de construire un système d'équations linéaires en $b^{(1)}, \ldots, b^{(123)}$ dont la solution constitue un ensemble d'estimateurs des composants de variance. Parmi de multiples possibilités, Dannenburg suggère d'utiliser des équations de ce type :

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\tilde{g}_{ij}^{(12)}\left(\sum_{k=1}^{K}\frac{w_{ijk\Sigma}}{w_{ij\Sigma}}(X_{ijkw}-X_{ijw})^2-\frac{(K-1)}{w_{ij\Sigma}}\hat{s}^2\right)= \\ (b^{(3)}+b^{(13)}+b^{(23)}+b^{(123)})\left[1-\sum_{i=1}^{I}\sum_{j=1}^{J}\tilde{g}_{ij}^{(12)}\sum_{k=1}^{K}\left(\frac{w_{ijk\Sigma}}{w_{ij\Sigma}}\right)^2\right], \tag{3.24}$$

$$\sum_{i=1}^{I}\tilde{g}_{i}^{(1)}\left(\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{w_{ijk\Sigma}}{w_{i\Sigma}}(X_{ijkw}-X_{iw})^2-\frac{(JK-1)}{w_{i\Sigma}}\hat{s}^2\right)= \\ (b^{(2)}+b^{(12)})\left[1-\sum_{i=1}^{I}\tilde{g}_{i}^{(1)}\sum_{j=1}^{J}\left(\frac{w_{ij\Sigma}}{w_{i\Sigma}}\right)^2\right] \\ +(b^{(3)}+b^{(13)})\left[1-\sum_{i=1}^{I}\tilde{g}_{i}^{(1)}\sum_{k=1}^{K}\left(\frac{w_{ik\Sigma}}{w_{i\Sigma}}\right)^2\right] \\ +(b^{(23)}+b^{(123)})\left[1-\sum_{i=1}^{I}\tilde{g}_{i}^{(1)}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\frac{w_{ijk\Sigma}}{w_{i\Sigma}}\right)^2\right] \tag{3.25}$$

et

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\frac{w_{ijk\Sigma}}{w_{\Sigma}}(X_{ijkw}-X_{w})^2-\frac{(IJK-1)}{w_{\Sigma}}\hat{s}^2= \\ b^{(1)}\left[1-\sum_{i=1}^{I}\left(\frac{w_{i\Sigma}}{w_{\Sigma}}\right)^2\right]+b^{(2)}\left[1-\sum_{j=1}^{J}\left(\frac{w_{j\Sigma}}{w_{\Sigma}}\right)^2\right] \\ +b^{(3)}\left[1-\sum_{k=1}^{K}\left(\frac{w_{k\Sigma}}{w_{\Sigma}}\right)^2\right]+b^{(12)}\left[1-\sum_{i=1}^{I}\sum_{j=1}^{J}\left(\frac{w_{ij\Sigma}}{w_{\Sigma}}\right)^2\right]$$

$$+ b^{(13)} \left[1 - \sum_{i=1}^{I} \sum_{k=1}^{K} \left(\frac{w_{ik\Sigma}}{w_{\Sigma}} \right)^2 \right] + b^{(23)} \left[1 - \sum_{j=1}^{J} \sum_{k=1}^{K} \left(\frac{w_{jk\Sigma}}{w_{\Sigma}} \right)^2 \right]$$
$$+ b^{(123)} \left[1 - \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \left(\frac{w_{ijk\Sigma}}{w_{\Sigma}} \right)^2 \right], \quad (3.26)$$

où les $\tilde{g}$ sont des poids quelconques dont la somme vaut 1. En pratique, le jugement de l'utilisateur entre en jeu au moment de choisir ces poids. Il convient, en effet, d'attribuer ces poids avec mesure afin qu'ils reflètent l'importance relative des différents groupements entre eux. On pourra employer des poids égaux si, a priori, aucun groupement ne semble avoir plus d'importance que les autres sur l'estimation des paramètres, ou encore les poids naturels dans le cas contraire.

En permutant adéquatement l'ordre de sommation, on obtient, pour chacune des deux premières équations, deux nouvelles équations semblables. Un exemple de permutations adéquates serait, pour la première des équations ci-dessus, de sommer d'abord dans l'ordre ijk, puis dans l'ordre ikj et enfin dans l'ordre jki. Cela porte à sept le nombre total d'équations, soit le même nombre que les inconnues. Des estimateurs des composants de variance $b^{(\cdot)}$ sont obtenus en résolvant ce système d'équations linéaires. Outre l'absence de biais, aisément vérifiable en prenant l'espérance de part et d'autre des équations ci-dessus, les estimateurs résultant de cette procédure d'estimation n'ont aucune propriété d'optimalité connue. On notera, de plus, que les estimateurs peuvent être négatifs.

On appellera dorénavant « estimateurs de Dannenburg » l'ensemble formé de ces estimateurs des composants de variance et de l'estimateur X_w de la moyenne collective. L'estimateur $\hat{s}^2$, rappelons-le, est traité à part.

3.5 Pseudo-estimateurs ad hoc

Au peu de connaissances théoriques relatives à la performance des estimateurs de Dannenburg s'ajoute, on le verra au chapitre 8, de décevants résultats numériques. Brièvement, les estimateurs des composants de variance sont fréquemment négatifs lorsque les vraies valeurs des paramètres sont faibles (autour de 1) et leur coefficient de variation est élevé (parfois plus de 5). Ceci incite à rechercher de nouveaux estimateurs aux performances accrues, voire des estimateurs optimaux sous certaines conditions.

Les pseudo-estimateurs[2] dits « ad hoc » développés dans cette section sont

[2]Rappelons qu'un pseudo-estimateur est un estimateur lui-même fonction du paramètre à estimer.

issus d'une recherche non systématique d'estimateurs optimaux. Ils se veulent une généralisation — ou une adaptation au contexte de la théorie de la crédibilité — des estimateurs ANOVA, dans le cas des estimateurs des composants de variance, et une généralisation directe de l'estimateur du modèle de Bühlmann–Straub dans le cas de la moyenne collective. Si leurs propriétés théoriques connues ne s'avèrent pas plus nombreuses que celles des estimateurs de Dannenburg, les résultats numériques tranchent, en revanche, nettement en leur faveur.

La présentation des pseudo-estimateurs ad hoc sera facilitée par l'étude préalable et détaillée du cas à deux facteurs. Le modèle à trois facteurs sera abordé par la suite.

3.5.1 Estimation de la moyenne collective

Dans le modèle de Bühlmann–Straub, l'estimateur de la moyenne collective m à variance minimale est X_{zw}, la moyenne pondérée par les facteurs de crédibilité. Il semble donc naturel d'estimer la moyenne collective du modèle CCC à l'aide de telles moyennes pondérées et d'espérer obtenir ainsi un estimateur performant.

Attardons-nous, dans un premier temps, au cas à deux facteurs. La moyenne pondérée par les poids naturels peut s'écrire, de manière équivalente, comme la moyenne pondérée de toutes les observations du portefeuille ou encore comme une suite de moyennes pondérées :

$$\begin{aligned}
X_w &= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T_{ij}} \frac{w_{ijt}}{w_\Sigma} X_{ijt} \\
&= \sum_{i=1}^{I} \frac{w_{i\Sigma}}{w_\Sigma} X_{iw} = \sum_{i=1}^{I}\sum_{j=1}^{J} \frac{w_{i\Sigma}}{w_\Sigma}\frac{w_{ij\Sigma}}{w_{i\Sigma}} X_{ijw} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{t=1}^{T_{ij}} \frac{w_{i\Sigma}}{w_\Sigma}\frac{w_{ij\Sigma}}{w_{i\Sigma}}\frac{w_{ijt}}{w_{ij\Sigma}} X_{ijt}.
\end{aligned}$$

Ce type de relation n'est toutefois plus vérifié lorsque les facteurs de crédibilité sont utilisés comme poids. Par exemple,

$$\sum_{i=1}^{I}\sum_{j=1}^{J} \frac{z_i^{(1)}}{z_\Sigma^{(1)}}\frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} X_{ijw} \neq \sum_{i=1}^{I}\sum_{j=1}^{J} \frac{z_{ij}^{(12)}}{z_\Sigma^{(12)}} X_{ijw}$$

car, en général, $z_{i\Sigma}^{(12)} \neq z_i^{(1)}$ et $z_\Sigma^{(1)} \neq z_\Sigma^{(12)}$. Cela permet donc plusieurs définitions différentes de la moyenne collective pondérée par des facteurs de crédibilité, soit,

$$
\begin{aligned}
X_{zw} &= \sum_{i=1}^{I} \frac{z_i^{(1)}}{z_\Sigma^{(1)}} X_{izw} = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{z_i^{(1)}}{z_\Sigma^{(1)}} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} X_{ijw}, \\
X_{zw}^{\diamond} &= \sum_{i=1}^{I} \frac{z_j^{(2)}}{z_\Sigma^{(2)}} X_{zjw} = \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{z_j^{(2)}}{z_\Sigma^{(2)}} \frac{z_{ij}^{(12)}}{z_{j\Sigma}^{(12)}} X_{ijw}, \\
X_{zw}^{\circ} &= \sum_{i=1}^{I} \sum_{j=1}^{J} \frac{z_{ij}^{(12)}}{z_\Sigma^{(12)}} X_{ijw}.
\end{aligned}
$$

Ces définitions étant en principe toutes aussi valables les unes que les autres, il serait avisé de choisir simplement celle débouchant sur le meilleur estimateur. Or, les tests numériques ont démontré (voir le chapitre 8) que si toutes les variantes ont une variance moindre que X_w, elles ont également à peu de choses près la même variance entre elles. En l'absence d'une claire indication de supériorité d'un estimateur par rapport aux autres, nous suggérons donc d'utiliser $\widetilde{m} = X_{zw}$ comme estimateur du paramètre m. En effet, cet estimateur s'inscrit mieux dans la logique des formules du modèle CCC telles que présentées au théorème 3.1.

Pour un modèle à trois facteurs, on a évidemment

$$
\widetilde{m} = X_{zw} = \sum_{i=1}^{I} \frac{z_i^{(1)}}{z_\Sigma^{(1)}} X_{izw},
$$

où X_{izw} est défini dans l'énoncé du théorème 3.1.

3.5.2 Modèle à deux facteurs

On peut maintenant aborder l'estimation des composants de variance. Tel que mentionné au chapitre 2, les estimateurs ANOVA sont à variance minimale parmi tous les estimateurs quadratiques sans biais pour un modèle à composants de variance équilibré. Ce type d'estimateurs apparaît donc naturellement comme un bon point de départ lors d'une recherche d'estimateurs performants pour le modèle CCC. Rappelons toutefois que la situation en crédibilité est très différente de celle rencontrée en statistique, notamment, d'une part, de par la présence de poids naturels dans le modèle et, d'autre part, de par le souhait d'éviter toute hypothèse de distribution pour les données, ce qui élimine d'office toute procédure d'estimation telle que le maximum de vraisemblance.

Nous généralisons donc, ici, les estimateurs ANOVA pour les adapter au contexte de notre modèle de crédibilité. Ceci passe, principalement, par l'introduction explicite de poids dans les formules. Le premier réflexe serait alors d'utiliser les poids naturels w. Les estimateurs en résultant seraient toutefois de proches cousins des estimateurs de Dannenburg, ce qui d'emblée soulève quelques doutes sur leur éventuelle supériorité.

Comme pour l'estimation de la moyenne collective, les nombreux travaux menés sur l'estimation des paramètres de structure dans le modèle de Bühlmann–Straub nous incitent plutôt à utiliser les facteurs de crédibilité en tant que poids. En effet, De Vylder & Goovaerts (1992*b*) ont démontré certaines propriétés d'optimalité du pseudo-estimateur $\tilde{a}$. L'idée consiste donc, ici, à développer des fonctions MSAz, MSBz et MSABz (on l'aura compris, « z » pour « pondération par les facteurs de crédibilité »), similaires à leurs équivalents de l'estimation ANOVA classique, dont on ajustera l'espérance afin d'obtenir des estimateurs sans biais de $b^{(1)}$, $b^{(2)}$ et $b^{(12)}$. Ceci peut évidemment se faire de plusieurs manières différentes. Celle proposée ici se transpose bien au cas à trois facteurs.

La première étape consiste à définir les fonctions MS et à en calculer l'espérance. Ainsi, soit

$$\begin{aligned}\mathrm{MSABz} = c_1 \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{J} z_{ij}^{(12)} z_{il}^{(12)} (X_{ijw} - X_{ilw})^2 \\ + c_2 \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{I} z_{ij}^{(12)} z_{kj}^{(12)} (X_{ijw} - X_{kjw})^2 \\ - c_3 \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{I}\sum_{l=1}^{J} z_{ij}^{(12)} z_{kl}^{(12)} (X_{ijw} - X_{klw})^2,\end{aligned} \tag{3.27}$$

où les constantes c_1, c_2 et c_3 sont les solutions du système d'équations

$$\begin{aligned}(Z_{(1)}^{(12)} - Z_{(12)}^{(12)})\, c_1 - (Z^{(12)} - Z_{(2)}^{(12)})\, c_3 = 0 \\ (Z_{(2)}^{(12)} - Z_{(12)}^{(12)})\, c_2 - (Z^{(12)} - Z_{(1)}^{(12)})\, c_3 = 0 \\ 2\, z_{\Sigma}^{(12)} \left[(J-1)\, c_1 + (I-1)\, c_2 - (IJ-1)\, c_3\right] = 1,\end{aligned} \tag{3.28}$$

avec

$$\begin{aligned} Z^{(12)} &= (z_{\Sigma}^{(12)})^2, & Z_{(1)}^{(12)} &= \sum_{i=1}^{I} (z_{i\Sigma}^{(12)})^2, \\ Z_{(2)}^{(12)} &= \sum_{j=1}^{J} (z_{j\Sigma}^{(12)})^2, & Z_{(12)}^{(12)} &= \sum_{i=1}^{I}\sum_{j=1}^{J} (z_{ij}^{(12)})^2. \end{aligned}$$

Les définitions de MSAz et MSBz sont un peu plus simples. En effet, on définit

$$\mathrm{MSAz} = \frac{1}{2\,(I-1)z_\Sigma^{(1)}} \sum_{i=1}^{I} \sum_{k=1}^{I} z_i^{(1)} z_k^{(1)} (X_{izw} - X_{kzw})^2 \tag{3.29}$$

et

$$\mathrm{MSBz} = \frac{1}{2\,(J-1)z_\Sigma^{(2)}} \sum_{j=1}^{J} \sum_{l=1}^{J} z_j^{(2)} z_k^{(2)} (X_{zjw} - X_{zlw})^2. \tag{3.30}$$

Les espérances de ces trois fonctions sont données dans le théorème suivant.

Théorème 3.4. *Les espérances des fonctions* MSABz, MSAz *et* MSBz *ci-dessus sont :*

$$\mathrm{E}[\mathrm{MSABz}] = b^{(12)}, \tag{3.31}$$

$$\mathrm{E}[\mathrm{MSAz}] = b^{(1)} + b^{(2)} K^{(1)}, \tag{3.32}$$

$$\mathrm{E}[\mathrm{MSBz}] = b^{(1)} K^{(2)} + b^{(2)}, \tag{3.33}$$

où

$$K^{(1)} = \frac{1}{(I-1)z_\Sigma^{(1)}} \sum_{i=1}^{I} \sum_{k=1}^{I} z_i^{(1)} z_k^{(1)} (S_{ii}^{(1)} - S_{ik}^{(1)}), \tag{3.34}$$

$$K^{(2)} = \frac{1}{(J-1)z_\Sigma^{(2)}} \sum_{j=1}^{J} \sum_{l=1}^{J} z_j^{(2)} z_l^{(2)} (S_{jj}^{(2)} - S_{jl}^{(2)}). \tag{3.35}$$

Preuve. Commençons par (3.31). On a, en utilisant les résultats du théorème 3.2,

$$\begin{aligned}
\mathrm{E}\big[(X_{ijw} - X_{klw})^2\big] &= \sigma_{ijij}^{(12)} - 2\,\sigma_{ijkl}^{(12)} + \sigma_{klkl}^{(12)} \\
&= 2\,(1-\delta_{ik})\,b^{(1)} + 2\,(1-\delta_{jl})\,b^{(2)} \\
&\quad + (1-\delta_{ij,kl})\frac{b^{(12)}}{z_{ij}^{(12)}} + (1-\delta_{ij,kl})\frac{b^{(12)}}{z_{kl}^{(12)}},
\end{aligned}$$

d'où (en abrégeant l'écriture des sommes)

$$\begin{aligned}
\mathrm{E}[\mathrm{MSABz}] = 2\,c_1 \sum_{ijl} z_{ij}^{(12)} z_{il}^{(12)} &\left[(1-\delta_{jl})\,b^{(2)} + (1-\delta_{jl})\frac{b^{(12)}}{z_{ij}^{(12)}}\right] \\
+ 2\,c_2 \sum_{ijk} z_{ij}^{(12)} z_{kj}^{(12)} &\left[(1-\delta_{ik})\,b^{(1)} + (1-\delta_{ik})\frac{b^{(12)}}{z_{ij}^{(12)}}\right]
\end{aligned}$$

$$
\begin{aligned}
&- 2\, c_3 \sum_{ijkl} z_{ij}^{(12)} z_{kl}^{(12)} \Bigg[(1-\delta_{ik})\, b^{(1)} + (1-\delta_{jl})\, b^{(2)} \\
&+ (1-\delta_{ij,kl}) \frac{b^{(12)}}{z_{ij}^{(12)}} \Bigg] \\
= {}& 2\, c_1 \Big[(Z_{(1)}^{(12)} - Z_{(12)}^{(12)})\, b^{(2)} + (J-1) z_\Sigma^{(12)} b^{(12)} \Big] \\
&+ 2\, c_2 \Big[(Z_{(2)}^{(12)} - Z_{(12)}^{(12)})\, b^{(1)} + (I-1) z_\Sigma^{(12)} b^{(12)} \Big] \\
&- 2\, c_3 \Big[(Z^{(12)} - Z_{(1)}^{(12)})\, b^{(1)} + (Z^{(12)} - Z_{(2)}^{(12)})\, b^{(2)} \\
&+ (IJ-1) z_\Sigma^{(12)} b^{(12)} \Big],
\end{aligned}
$$

avec les fonctions $Z_{(\cdot)}^{(\cdot)}$ telles que définies précédemment. Or, en définissant les constantes c_1, c_2 et c_3 comme les solutions du système (3.28), on obtient le résultat (3.31). Quant aux résultats (3.32) et (3.33), ils s'obtiennent aisément et de façon similaire à partir des identités

$$
\begin{aligned}
\mathrm{E}\big[(X_{izw} - X_{kzw})^2\big] &= \sigma_{ii}^{(1)} - 2\, \sigma_{ik}^{(1)} + \sigma_{kk}^{(1)} \\
&= (1-\delta_{ik}) \frac{b^{(1)}}{z_i^{(1)}} + (1-\delta_{ik}) \frac{b^{(1)}}{z_k^{(1)}} \\
&\quad + b^{(2)} (S_{ii}^{(1)} - 2\, S_{ik}^{(1)} + S_{kk}^{(1)}), \\
\mathrm{E}\big[(X_{zjw} - X_{zlw})^2\big] &= \sigma_{jj}^{(2)} - 2\, \sigma_{jl}^{(2)} + \sigma_{ll}^{(2)} \\
&= b^{(1)} (S_{jj}^{(2)} - 2\, S_{jl}^{(2)} + S_{ll}^{(2)}) \\
&\quad + (1-\delta_{jl}) \frac{b^{(2)}}{z_j^{(2)}} + (1-\delta_{jl}) \frac{b^{(2)}}{z_l^{(2)}}.
\end{aligned}
$$

□

Les espérances du théorème précédent découlent des définitions choisies pour les constantes c_1, c_2, c_3, $K^{(1)}$ et $K^{(2)}$. Ce choix est évidemment purement arbitraire.

Les pseudo-estimateurs sans biais de $b^{(1)}$, $b^{(2)}$ et $b^{(12)}$ sont par conséquent les solutions du système d'équations

$$
\begin{bmatrix} \mathrm{MSAz} \\ \mathrm{MSBz} \\ \mathrm{MSABz} \end{bmatrix} = \begin{bmatrix} 1 & K^{(1)} & 0 \\ K^{(2)} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{b}^{(1)} \\ \tilde{b}^{(2)} \\ \tilde{b}^{(12)} \end{bmatrix},
$$

soit, dans ce cas simple du modèle à deux facteurs,

$$\begin{bmatrix} \tilde{b}^{(1)} \\ \tilde{b}^{(2)} \\ \tilde{b}^{(12)} \end{bmatrix} = \frac{1}{1 - K^{(1)}K^{(2)}} \begin{bmatrix} 1 & -K^{(1)} & 0 \\ -K^{(2)} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \text{MSAz} \\ \text{MSBz} \\ \text{MSABz} \end{bmatrix}. \tag{3.36}$$

On notera que les estimateurs $\tilde{b}^{(1)}$, $\tilde{b}^{(2)}$ et $\tilde{b}^{(12)}$ ne forment pas à proprement parler la solution d'un système d'équations linéaires, comme l'égalité ci-dessus pourrait le laisser croire. Les estimateurs résultent plutôt d'une procédure de point fixe, amorcée avec des valeurs arbitraires de $\tilde{b}^{(1)}$, $\tilde{b}^{(2)}$ et $\tilde{b}^{(12)}$ et lors de laquelle (3.36) est évaluée itérativement jusqu'à convergence vers des valeurs vérifiant approximativement l'égalité. Nous revenons plus longuement sur le calcul des pseudo-estimateurs par la méthode du point fixe à la section 4.6.

Théorème 3.5. *Les pseudo-estimateurs* (3.36) *sont équivalents aux estimateurs ANOVA si tous les poids sont égaux et si $T_{ij} = T$ pour tout i, j.*

Preuve. Si tous les poids naturels sont égaux, alors $w_{ijt} \equiv w$ et, puisque $T_{ij} = T$, $\hat{s}^2 = w\,\text{MSE}$. Les facteurs de crédibilité sont également tous égaux :

$$z_{ij}^{(12)} \equiv z^{(12)} = \frac{Twb^{(12)}}{Twb^{(12)} + s^2}.$$

Ceci entraîne

$$\begin{aligned} Z_{(12)}^{(12)} &= IJ(z^{(12)})^2, & Z_{(1)}^{(12)} &= IJ^2(z^{(12)})^2, \\ Z_{(2)}^{(12)} &= I^2J(z^{(12)})^2, & Z^{(12)} &= I^2J^2(z^{(12)})^2 \end{aligned}$$

et la solution du système d'équations (3.28) est

$$\begin{aligned} c_1 &= \frac{1}{2\,z^{(12)}J(I-1)(J-1)}, \\ c_2 &= \frac{1}{2\,z^{(12)}I(I-1)(J-1)}, \\ c_3 &= \frac{1}{2\,z^{(12)}IJ(I-1)(J-1)}. \end{aligned}$$

Par conséquent,

$$\text{MSABz} = \frac{z^{(12)}}{T}\text{MSAB}.$$

Or, comme $b^{(12)}$ est estimé itérativement en remplaçant à chaque fois les paramètres $b^{(12)}$ et s^2 par leurs estimateurs et que, en fait, $\tilde{b}^{(12)} = \text{MSABz}$, on a

$$\begin{aligned}\text{MSABz} &= \frac{w\,\text{MSABz}}{Tw\,\text{MSABz} + w\,\text{MSE}}\text{MSAB} \\ &= \frac{\text{MSAB} - \text{MSE}}{T},\end{aligned}$$

ce qui est l'estimateur ANOVA de $b^{(12)}$. Pour les deux autres estimateurs, il s'agit de remarquer que $K^{(1)} = K^{(2)} = 0$ et le reste de la preuve est similaire à ce qui précède. □

3.5.3 Modèle à trois facteurs

La procédure pour développer des pseudo-estimateurs basés sur les estimateurs ANOVA est évidemment la même pour le modèle à trois facteurs. Pour cette raison, le théorème sur les espérances des fonctions MS ne sera pas prouvé dans cette section.

Le troisième facteur de risque est identifié par la lettre C. Comme précédemment, on commence par définir les fonctions MSABCz, MSABz et MSAz. Ainsi, soit

$$\begin{aligned}\text{MSABCz} = c_1 &\sum_{ijkh} z_{ijk}^{(123)} z_{ijh}^{(123)} (X_{ijkw} - X_{ijhw})^2 \\ &+ c_2 \sum_{ijkg} z_{ijk}^{(123)} z_{igk}^{(123)} (X_{ijkw} - X_{igkw})^2 \\ &+ c_3 \sum_{ijkf} z_{ijk}^{(123)} z_{fjk}^{(123)} (X_{ijkw} - X_{fjkw})^2 \\ &- c_4 \sum_{ijkgh} z_{ijk}^{(123)} z_{igh}^{(123)} (X_{ijkw} - X_{ighw})^2 \\ &- c_5 \sum_{ijkfh} z_{ijk}^{(123)} z_{fjh}^{(123)} (X_{ijkw} - X_{fjhw})^2 \\ &- c_6 \sum_{ijkfg} z_{ijk}^{(123)} z_{fgk}^{(123)} (X_{ijkw} - X_{fgkw})^2 \\ &+ c_7 \sum_{ijkfgh} z_{ijk}^{(123)} z_{fgh}^{(123)} (X_{ijkw} - X_{fghw})^2,\end{aligned} \tag{3.37}$$

où les constantes $c_1, \ldots, c_7$ sont les solutions du système d'équations suivant :

$$
\begin{aligned}
{}_2Z_{13}c_2 + {}_1Z_{23}c_3 - {}_2Z_1c_4 - {}_1Z_2c_5 - {}_{12}Z_3c_6 + {}_{12}Zc_7 &= 0 \\
{}_3Z_{12}c_1 + {}_1Z_{23}c_3 - {}_3Z_1c_4 - {}_{13}Z_2c_5 - {}_1Z_3c_6 + {}_{13}Zc_7 &= 0 \\
{}_3Z_{12}c_1 + {}_2Z_{13}c_2 - {}_{23}Z_1c_4 - {}_3Z_2c_5 - {}_2Z_3c_6 + {}_{23}Zc_7 &= 0 \\
{}_1Z_{23}c_3 - {}_1Z_2c_5 - {}_1Z_3c_6 + {}_1Zc_7 &= 0 \\
{}_2Z_{13}c_2 - {}_2Z_1c_4 - {}_2Z_3c_6 + {}_2Zc_7 &= 0 \\
{}_3Z_{12}c_1 - {}_3Z_1c_4 - {}_3Z_2c_5 + {}_3Zc_7 &= 0 \\
2\, z_\Sigma^{(123)} \left[(K-1)c_1 + (J-1)c_2 + (I-1)c_3 - (JK-1)c_4 \right.& \\
\left. -(IK-1)c_5 - (IJ-1)c_6 + (IJK-1)c_7\right] &= 1,
\end{aligned}
$$

avec ${}_2Z_1 = Z_{(1)}^{(123)} - Z_{(12)}^{(123)}$ et

$$
Z^{(123)} = (z_\Sigma^{(123)})^2, \qquad Z_{(1)}^{(123)} = \sum_{i=1}^{I} (z_{i\Sigma}^{(123)})^2,
$$

$$
Z_{(12)}^{(123)} = \sum_{i=1}^{I}\sum_{j=1}^{J} (z_{ij\Sigma}^{(123)})^2, \qquad Z_{(123)}^{(123)} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} (z_{ijk}^{(123)})^2,
$$

ainsi de suite.

De plus,

$$
\begin{aligned}
\mathrm{MSABz} = d_1 \sum_{ijg} z_{ij}^{(12)} z_{ig}^{(12)} (X_{ijzw} - X_{igzw})^2 & \\
+ d_2 \sum_{ijf} z_{ij}^{(12)} z_{fj}^{(12)} (X_{ijzw} - X_{fjzw})^2 & \\
- d_3 \sum_{ijfg} z_{ij}^{(12)} z_{fg}^{(12)} (X_{ijzw} - X_{fgzw})^2, &
\end{aligned}
\tag{3.38}
$$

où les constantes d_1, d_2 et d_3 sont les solutions de

$$
\begin{aligned}
(Z_{(2)}^{(12)} - Z_{(12)}^{(12)})\, d_2 - (Z^{(12)} - Z_{(1)}^{(12)})\, d_3 &= 0 \\
(Z_{(1)}^{(12)} - Z_{(12)}^{(12)})\, d_1 - (Z^{(12)} - Z_{(2)}^{(12)})\, d_3 &= 0 \\
2\, z_\Sigma^{(12)} \left[(J-1)\, d_1 + (I-1)\, d_2 - (IJ-1)\, d_3\right] &= 1,
\end{aligned}
$$

avec

$$
Z^{(12)} = (z_\Sigma^{(12)})^2, \qquad Z_{(1)}^{(12)} = \sum_{i=1}^{I} (z_{i\Sigma}^{(12)})^2,
$$

$$
Z_{(2)}^{(12)} = \sum_{j=1}^{J} (z_{j\Sigma}^{(12)})^2, \qquad Z_{(12)}^{(12)} = \sum_{i=1}^{I}\sum_{j=1}^{J} (z_{ij}^{(12)})^2.
$$

Les expressions de MSACz et MSBCz sont similaires à celle de MSABz.

Enfin, on définit

$$\mathrm{MSAz} = \frac{1}{2\,(I-1)z_\Sigma^{(1)}} \sum_{i=1}^{I} \sum_{f=1}^{I} z_i^{(1)} z_f^{(1)} (X_{izw} - X_{fzw})^2, \tag{3.39}$$

avec des expressions similaires pour MSBz et MSCz.

Les constantes $c_1, \ldots, c_7$ et $d_1, \ldots, d_3$ de même que la constante à la droite de l'expression de MSAz, ci-dessus, ont été choisies de manière à obtenir les espérances du théorème suivant.

Théorème 3.6. *Les espérances des fonctions* MSABCz, MSABz *et* MSAz *ci-dessus sont :*

$$\mathrm{E[MSABCz]} = b^{(123)} \tag{3.40}$$

$$\mathrm{E[MSABz]} = b^{(12)} + b^{(3)} K_{(3)}^{(12)} + b^{(13)} K_{(13)}^{(12)} + b^{(23)} K_{(23)}^{(12)} \tag{3.41}$$

$$\mathrm{E[MSAz]} = b^{(1)} + b^{(2)} K_{(2)}^{(1)} + b^{(3)} K_{(3)}^{(1)} + b^{(13)} K_{(13)}^{(1)} + b^{(23)} K_{(23)}^{(1)} \tag{3.42}$$

avec

$$K_{(2)}^{(1)} = \frac{1}{(I-1)z_\Sigma^{(1)}} \sum_{i=1}^{I} \sum_{f=1}^{I} z_i^{(1)} z_f^{(1)} (S_{ii}^{(1)} - S_{if}^{(1)}),$$

$$K_{(3)}^{(1)} = \frac{1}{(I-1)z_\Sigma^{(1)}} \sum_{i=1}^{I} \sum_{f=1}^{I} z_i^{(1)} z_f^{(1)} (S_{iziz}^{(12)} - S_{izfz}^{(12)}),$$

$$K_{(13)}^{(1)} = \frac{1}{(I-1)z_\Sigma^{(1)}} \sum_{i=1}^{I} \sum_{f=1}^{I} z_i^{(1)} z_f^{(1)} (S_{iziz}^{(12)} - \delta_{if} S_{izfz}^{(12)}),$$

$$K_{(23)}^{(1)} = \frac{1}{(I-1)z_\Sigma^{(1)}} \sum_{i=1}^{I} \sum_{f=1}^{I} z_i^{(1)} z_f^{(1)} (S_{izi\star}^{(12)} - S_{izf\star}^{(12)})$$

et

$$\begin{aligned} K_{(3)}^{(12)} = {} & d_1 \sum_{ijg} z_{ij}^{(12)} z_{ig}^{(12)} (S_{ijij}^{(12)} - S_{ijig}^{(12)}) \\ & + d_2 \sum_{ijf} z_{ij}^{(12)} z_{fj}^{(12)} (S_{ijij}^{(12)} - S_{ijfj}^{(12)}) \\ & - d_3 \sum_{ijfg} z_{ij}^{(12)} z_{fg}^{(12)} (S_{ijij}^{(12)} - S_{ijfg}^{(12)}), \end{aligned}$$

$$\begin{aligned}
K_{(13)}^{(12)} &= d_1 \sum_{ijg} z_{ij}^{(12)} z_{ig}^{(12)} (S_{ijij}^{(12)} - S_{ijig}^{(12)}) \\
&\quad + d_2 \sum_{ijf} z_{ij}^{(12)} z_{fj}^{(12)} (S_{ijij}^{(12)} - \delta_{if} S_{ijfj}^{(12)}) \\
&\quad - d_3 \sum_{ijfg} z_{ij}^{(12)} z_{fg}^{(12)} (S_{ijij}^{(12)} - \delta_{if} S_{ijfg}^{(12)}), \\
K_{(23)}^{(12)} &= d_1 \sum_{ijg} z_{ij}^{(12)} z_{ig}^{(12)} (S_{ijij}^{(12)} - \delta_{jg} S_{ijig}^{(12)}) \\
&\quad + d_2 \sum_{ijf} z_{ij}^{(12)} z_{fj}^{(12)} (S_{ijij}^{(12)} - S_{ijfj}^{(12)}) \\
&\quad - d_3 \sum_{ijfg} z_{ij}^{(12)} z_{fg}^{(12)} (S_{ijij}^{(12)} - \delta_{jg} S_{ijfg}^{(12)}).
\end{aligned}$$

Les pseudo-estimateurs ad hoc des composants de variance d'un modèle CCC à trois facteurs constituent donc la solution du système d'équations (en apparence linéaire) suivant :

$$\begin{bmatrix} \text{MSAz} \\ \text{MSBz} \\ \text{MSCz} \\ \text{MSABz} \\ \text{MSACz} \\ \text{MSBCz} \\ \text{MSABCz} \end{bmatrix} = \begin{bmatrix} 1 & K_{(2)}^{(1)} & K_{(3)}^{(1)} & 0 & K_{(13)}^{(1)} & K_{(23)}^{(1)} & 0 \\ K_{(1)}^{(2)} & 1 & K_{(3)}^{(2)} & 0 & K_{(13)}^{(2)} & K_{(23)}^{(2)} & 0 \\ K_{(1)}^{(3)} & K_{(2)}^{(3)} & 1 & K_{(12)}^{(3)} & 0 & K_{(23)}^{(3)} & 0 \\ 0 & 0 & K_{(3)}^{(12)} & 1 & K_{(13)}^{(12)} & K_{(23)}^{(12)} & 0 \\ 0 & K_{(2)}^{(13)} & 0 & K_{(12)}^{(13)} & 1 & K_{(23)}^{(13)} & 0 \\ K_{(1)}^{(23)} & 0 & 0 & K_{(12)}^{(23)} & K_{(13)}^{(23)} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{b}^{(1)} \\ \tilde{b}^{(2)} \\ \tilde{b}^{(3)} \\ \tilde{b}^{(12)} \\ \tilde{b}^{(13)} \\ \tilde{b}^{(23)} \\ \tilde{b}^{(123)} \end{bmatrix}. \tag{3.43}$$

Pour un modèle CCC à plus de trois facteurs, il est possible de dériver des pseudo-estimateurs de type « ad hoc » à partir des règles de construction des estimateurs ANOVA de la section 4.1 de Searle et al. (1992). Il est toutefois clair que les calculs deviennent rapidement fastidieux quand le nombre de facteurs augmente. Il s'agit bien là du principal inconvénient de cette technique d'estimation.

Un autre inconvénient de tous les pseudo-estimateurs est que leur absence de biais (le cas échéant) n'est vérifiée théoriquement que si les paramètres que l'on souhaite précisément estimer sont considérés connus. Par exemple, X_{zw} est un (pseudo-)estimateur sans biais de m si les facteurs de crédibilité sont vus comme des constantes fixes. Dans le cas contraire, il devient très difficile, voire impossible, de calculer l'espérance de X_{zw}. Cet inconvénient est toutefois davantage de nature théorique puisque, en pratique, le biais des pseudo-estimateurs rencontrés ici se révèle être tout à fait négligeable.

Chapitre 4

Estimateurs optimaux

Nous abordons, dans ce chapitre, l'un des éléments clé de cette thèse, à savoir Nous abordons, dans ce chapitre, l'un des éléments clé de cette thèse, à savoir l'introduction d'estimateurs optimaux des paramètres de structure pour le modèle de crédibilité à classification croisée. Cette recherche s'inscrit évidemment dans le cadre des travaux de Dannenburg (1995), mais s'inspire principalement des travaux de De Vylder & Goovaerts (1991, 1992*a*, 1992*b*). En effet, c'est à ces auteurs que l'on doit les premières études sur l'optimalité des estimateurs du modèle de Bühlmann–Straub sous des hypothèses d'excès nul.

Le présent auteur a ajouté sa brique à l'édifice avec Goulet (1998), en couplant l'hypothèse d'excès nul à l'approche des effets aléatoires. Les conclusions concernant l'optimalité des estimateurs $\hat{s}^2$ et $\tilde{a}$ du modèle de Bühlmann–Straub n'en sont pas changées, mais les hypothèses nécessaires pour y arriver sont alors moins contraignantes. C'est cette approche qui sera reprise ici. L'idée consiste simplement à supposer que, dans un modèle CCC, toutes les variables Ξ ont un coefficient d'excès nul. Du lemme 2.3, il découle alors que la distribution du ratio d'expérience est également sans excès. Nous reviendrons sur les détails en temps et lieu.

Les estimateurs optimaux de ce chapitre seront exprimés sous forme de matrices représentées par des lettres majuscules en caractère gras. La notation

$$\mathbf{A} = [a_{ij}]_{I \times J}$$

signifie que $\mathbf{A}$ est une matrice de dimension $I \times J$ dont l'élément en position (i, j) est a_{ij}. Si $J = 1$, on a un vecteur. On rencontrera également la notation

$$\mathbf{A} = [\mathbf{A}_{ij}]_{I \times J},$$

qui définit alors $\mathbf{A}$ comme une matrice par blocs, c'est-à-dire dont les éléments $\mathbf{A}_{ij}$ sont eux-mêmes des matrices. L'utilisation de matrices par blocs peut aussi être évité avec la notation

$$\mathbf{A} = [a_{ijkl}]_{IJ \times IJ},$$

où les indices i et j identifient alors les lignes de la matrice et les indices k et l, les colonnes. La transposée de $\mathbf{A}$ est notée $\mathbf{A}'$ et l'inverse, $\mathbf{A}^{-1}$. La matrice identité est notée $\mathbf{I}$ et on définit la «matrice unitée» (carrée) $\mathbf{U}$, ne contenant que des 1. Enfin, si $\mathbf{V} = [v_i]_{I \times 1}$ est un vecteur, alors $\operatorname{diag}(\mathbf{V})$ est une matrice diagonale $I \times I$ dont la diagonale est formée par les éléments v_i.

Les dernières sections du chapitre précédent ont bien montré qu'il s'avère parfois plus simple d'étudier d'abord le modèle CCC à deux facteurs, pour passer ensuite au modèle à trois facteurs. C'est donc ce qui sera fait ici. Entre les deux, nous aborderons le calcul et les propriétés des estimateurs optimaux. Le premier paramètre estimé de manière optimale est la moyenne collective m. Par la suite, on démontre sous quelles hypothèses et dans quel espace l'estimateur $\hat{s}^2$ est optimal. Enfin, on développe des estimateurs optimaux des composants de variance. Les estimateurs optimaux seront identifiés par un astérisque en indice.

4.1 Estimateur de la moyenne collective

Dans le modèle de Bühlmann–Straub, l'estimateur à variance minimale de la moyenne collective m dans la classe des estimateurs linéaires sans biais, X_{zw}, peut être obtenu facilement par la méthode des multiplicateurs de Lagrange ou encore par projection. Dans le modèle CCC, cependant, des équations complexes apparaissent si l'on souhaite obtenir une expression explicite pour l'estimateur. En adoptant une écriture matricielle, il est toutefois possible d'obtenir une expression qui, à défaut d'être véritablement explicite, demeure néanmoins assez simple.

Théorème 4.1. *Dans un modèle CCC à deux facteurs, l'estimateur sans biais à variance minimale de m dans la classe des fonctions linéaires des observations est*

$$\widehat{m}_* = \frac{\mathbf{B}'\mathbf{C}^{-1}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{X}, \tag{4.1}$$

où

$$\begin{aligned} \mathbf{X} &= [\mathbf{X}_i]_{I \times 1}, & \mathbf{X}_i &= [X_{ijw}]_{J \times 1}, \\ \mathbf{B} &= [\mathbf{B}_i]_{I \times 1}, & \mathbf{B}_i &= [1]_{J \times 1}, \\ \mathbf{C} &= [\mathbf{C}_{ik}]_{I \times I}, & \mathbf{C}_{ik} &= \delta_{ik} b^{(1)}\mathbf{U} + b^{(2)}\mathbf{I} + \delta_{ik} b^{(12)}\mathbf{Z}_i \end{aligned}$$

et $\mathbf{Z}_i$ *est une matrice diagonale* $J \times J$ *formée des réciproques des facteurs de crédibilité* $z_{i1}^{(12)}, \dots, z_{iJ}^{(12)}$ *:*

$$\mathbf{Z}_i = \mathrm{diag}([(z_{ij}^{(12)})^{-1}]_{J\times 1}).$$

Preuve. L'espace des fonctions linéaires des observations d'espérance m est $\mathcal{L}_m\{X_{ijt}\}$. On cherche donc la projection de m sur cet espace. On peut toutefois facilement démontrer que celle-ci est égale à la projection de m sur $\mathcal{L}_m\{X_{ijw}\}$. Par conséquent,

$$\begin{aligned}\widehat{m}_* &= \mathrm{pro}(m|\mathcal{L}_m\{X_{ijw}\}) \\ &= \sum_{i=1}^{I}\sum_{j=1}^{J} \alpha_{ij} X_{ijw}. \end{aligned} \tag{4.2}$$

Les équations normales (2.8) sont donc, pour $k = 1, \dots, I$ et $l = 1, \dots, J$,

$$\begin{aligned}\mathrm{E}[\widehat{m}_*] &= m, \\ \mathrm{Cov}(\widehat{m}_*, X_{klw}) &= c\,\mathrm{E}[X_{klw}]\,,\end{aligned}$$

où c est une constante quelconque. Or, $\mathrm{E}[X_{ijw}] = m$ et, en utilisant le résultat (3.15) du théorème 3.2,

$$\begin{aligned}\mathrm{Cov}(\widehat{m}_*, X_{klw}) &= \sum_{i=1}^{I}\sum_{j=1}^{J} \alpha_{ij}\sigma_{ijkl}^{(12)} \\ &= \sum_{i=1}^{I}\sum_{j=1}^{J} \alpha_{ij}\left[\delta_{ik}b^{(1)} + \delta_{jl}b^{(2)} + \delta_{ij,kl}\frac{b^{(12)}}{z_{ij}^{(12)}}\right] \\ &= \alpha_{k\Sigma}b^{(1)} + \alpha_{\Sigma l}b^{(2)} + \alpha_{kl}\frac{b^{(12)}}{z_{kl}^{(12)}}.\end{aligned}$$

Les équations normales à résoudre pour $\alpha_{11}, \dots, \alpha_{IJ}$ sont par conséquent

$$\sum_{i=1}^{I}\sum_{j=1}^{J} \alpha_{ij} = 1, \tag{4.3}$$

$$\alpha_{i\Sigma}b^{(1)} + \alpha_{\Sigma j}b^{(2)} + \alpha_{ij}\frac{b^{(12)}}{z_{ij}^{(12)}} = \tilde{c}, \tag{4.4}$$

ou encore, sous forme matricielle,

$$\begin{aligned}\mathbf{B}'\mathbf{A} &= 1 \\ \mathbf{C}\mathbf{A} &= \tilde{c}\,\mathbf{B},\end{aligned}$$

où $\tilde{c}$ est une nouvelle constante, $\mathbf{A} = [\mathbf{A}_i]_{I\times 1}$, $\mathbf{A}_i = [\alpha_{ij}]_{J\times 1}$ et $\mathbf{B}$, $\mathbf{C}$ sont telles que définies dans l'énoncé du théorème. Or, la solution de ce système est donnée par

$$\mathbf{A} = \frac{\mathbf{C}^{-1}\mathbf{B}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}$$

et (4.2) peut s'écrire $\widehat{m}_* = \mathbf{A}'\mathbf{X}$. □

Remarque. Dans un modèle avec $I = 2$ et $J = 3$, la matrice $\mathbf{C}$ ci-dessus a cette forme (les pointillés délimitent les blocs $\mathbf{C}_{ik}$) :

$$\left[\begin{array}{ccc:ccc}
b + \frac{b^{(12)}}{z_{11}^{(12)}} & b^{(1)} & b^{(1)} & b^{(2)} & 0 & 0 \\
b^{(1)} & b + \frac{b^{(12)}}{z_{12}^{(12)}} & b^{(1)} & 0 & b^{(2)} & 0 \\
b^{(1)} & b^{(1)} & b + \frac{b^{(12)}}{z_{13}^{(12)}} & 0 & 0 & b^{(2)} \\
\hdashline
b^{(2)} & 0 & 0 & b + \frac{b^{(12)}}{z_{21}^{(12)}} & b^{(1)} & b^{(1)} \\
0 & b^{(2)} & 0 & b^{(1)} & b + \frac{b^{(12)}}{z_{22}^{(12)}} & b^{(1)} \\
0 & 0 & b^{(2)} & b^{(1)} & b^{(1)} & b + \frac{b^{(12)}}{z_{23}^{(12)}}
\end{array}\right],$$

où $b \equiv b^{(1)} + b^{(2)}$. On vérifie alors que $\mathbf{CA}$ correspond bel et bien à la seconde équation normale (4.4).

On remarquera également que l'estimateur optimal de la moyenne collective est obtenu sans aucune hypothèse additionnelle sur la distribution des ratios X_{ijt}. L'hypothèse d'excès nul, dont il a déjà été fait mention quelques fois jusqu'ici, n'intervient que dans l'estimation de s^2 et des composants de variance, ci-dessous.

4.2 Estimation de s^2

Lors de la présentation de l'estimateur du paramètre s^2 pour le modèle à trois facteurs, au chapitre 3, nous avions déjà mentionné que $\hat{s}^2$ s'avérait être optimal dans une certaine classe d'estimateurs. Il convient maintenant d'étudier les détails de cette proposition pour un modèle à deux facteurs.

L'étude de l'optimalité de $\hat{s}^2$ marque d'abord l'entrée en scène de l'hypothèse d'excès nul. Tel que mentionné en introduction de ce chapitre, la stratégie employée ici consiste à supposer que tous les effets aléatoires ont un coefficient d'excès nul. La conséquence de ceci est formalisée dans le lemme suivant.

Lemme 4.1. *Dans un modèle CCC à deux facteurs, si les variables aléatoires $\Xi_i^{(1)}$, $\Xi_j^{(2)}$, $\Xi_{ij}^{(12)}$ et $\Xi_{ijt}^{(123)}$ sont sans excès, alors X_{ijt}, X_{ijw} et toute autre moyenne pondérée sont aussi sans excès.*

Preuve. La variable X_{ijt} est sans excès par le lemme 2.3. Quant à la moyenne pondérée individuelle, on peut écrire

$$X_{ijw} = m + \Xi_i^{(1)} + \Xi_j^{(2)} + \Xi_{ij}^{(12)} + \Xi_{ijw}^{(123)},$$

où

$$\Xi_{ijw}^{(123)} = \sum_{t=1}^{T_{ij}} \frac{w_{ijt}}{w_{ij\Sigma}} \Xi_{ijt}^{(123)}.$$

Toujours par le lemme 2.3, $\Xi_{ijw}^{(123)}$ est sans excès et, par conséquent, X_{ijw} l'est aussi. Le raisonnement est le même pour d'autres moyennes pondérées comme, par exemple, X_{iw}, X_{izw} ou X_w. □

On remarquera qu'il n'est fait aucune hypothèse sur le troisième moment de quelque variable aléatoire. Il n'est donc pas nécessaire que la distribution de la variable X_{ijt}, par exemple, soit symétrique.

Le théorème suivant énonce les conditions sous lesquelles l'estimateur $\hat{s}^2$ est optimal.

Théorème 4.2. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCC à deux facteurs, alors l'estimateur du paramètre de structure s^2 à variance minimale parmi toutes les combinaisons du type*

$$\sum_{ijturs} \alpha_{ijturs}(X_{ijt} - X_{iju})(X_{ijr} - X_{ijs})$$

avec moyenne s^2 est

$$\hat{s}_*^2 = \frac{1}{\sum_{ij}(T_{ij} - 1)} \sum_{i=1}^{I} \sum_{j=1}^{J} (T_{ij} - 1)\hat{s}_{ij}^2, \tag{4.5}$$

avec

$$\hat{s}_{ij}^2 = \frac{1}{(T_{ij} - 1)} \sum_{t=1}^{T_{ij}} w_{ijt}(X_{ijt} - X_{ijw})^2 \tag{4.6}$$

$$= \frac{1}{2(T_{ij} - 1)w_{ij\Sigma}} \sum_{t=1}^{T_{ij}} \sum_{u=1}^{T_{ij}} w_{ijt}w_{iju}(X_{ijt} - X_{iju})^2. \tag{4.7}$$

Preuve. Il faut démontrer que $\hat{s}^2_*$ est la projection de s^2 sur $\mathcal{L}_{s^2}\{(X_{ijt} - X_{iju})(X_{ijr} - X_{ijs})\}$ ou, autrement dit, que $\hat{s}^2_*$ satisfait, pour $k = 1, \dots, I$ et $l = 1, \dots, J$, les équations normales

$$\begin{aligned}\mathrm{E}\big[\hat{s}^2_*\big] &= s^2, \\ \mathrm{Cov}(\hat{s}^2_*, Q_{kl}) &= c\,\mathrm{E}[Q_{kl}]\,,\end{aligned}$$

où $Q_{kl} \in \{(X_{klt} - X_{klu})(X_{klr} - X_{kls}); t, u, r, s = 1, \dots, T_{kl}\}$. L'absence de biais de $\hat{s}^2_*$ et $\hat{s}^2_{ij}$ est bien connue. Pour prouver la seconde équation normale, on commence avec l'identité

$$\begin{aligned}\mathrm{Cov}(\hat{s}^2_*, Q_{kl}) &= \frac{1}{\sum_{ij}(T_{ij}-1)} \sum_{i=1}^{I}\sum_{j=1}^{J}(T_{ij}-1)\left(\mathrm{E}\big[\hat{s}^2_{ij}Q_{kl}\big] - \mathrm{E}\big[\hat{s}^2_{ij}\big]\,\mathrm{E}[Q_{kl}]\right) \\ &= \frac{T_{kl}-1}{\sum_{ij}(T_{ij}-1)}\left(\mathrm{E}\big[\hat{s}^2_{kl}Q_{kl}\big] - \mathrm{E}\big[\hat{s}^2_{kl}\big]\,\mathrm{E}[Q_{kl}]\right),\end{aligned}$$

par indépendance entre les assurés. Or, de (2.12) du lemme 2.4 et (2.10) du lemme 2.2, on obtient

$$\begin{aligned}\mathrm{E}[Q_{ij}] &= \mathrm{E}[(X_{ijt} - X_{iju})(X_{ijr} - X_{ijs})] \\ &= \mathrm{E}\Big[(\Xi^{(123)}_{ijt} - \Xi^{(123)}_{iju})(\Xi^{(123)}_{ijr} - \Xi^{(123)}_{ijs})\Big] \\ &= s^2\left[w^{-1}_{ijt}(\delta_{tr} - \delta_{ts}) - w^{-1}_{iju}(\delta_{ur} - \delta_{us})\right] \\ &= s^2\, e(Q_{ij}),\end{aligned}$$

où

$$e(Q_{ij}) = w^{-1}_{ijt}(\delta_{tr} - \delta_{ts}) - w^{-1}_{iju}(\delta_{ur} - \delta_{us}).$$

De plus, en utilisant (2.11) du lemme 2.2, (2.13) du lemme 2.4 et le fait que les variables X_{ijt} sont sans excès (lemme 4.1 ci-dessus), on a

$$\begin{aligned}&2\,w_{ij\Sigma}(T_{ij}-1)\mathrm{E}\big[\hat{s}^2_{ij}Q_{ij}\big] = \\ &\quad= \sum_{pq} w_{ijp}w_{ijq}\mathrm{E}\Big[(\Xi^{(123)}_{ijp} - \Xi^{(123)}_{ijq})^2(\Xi^{(123)}_{ijt} - \Xi^{(123)}_{iju})(\Xi^{(123)}_{ijr} - \Xi^{(123)}_{ijs})\Big] \\ &\quad= 2\,s^4\sum_{pq} w_{ijp}w_{ijq}\left[w^{-1}_{ijp}w^{-1}_{ijt}(\delta_{tr} - \delta_{ts}) - w^{-1}_{ijp}w^{-1}_{iju}(\delta_{ur} - \delta_{us})\right. \\ &\qquad - w^{-1}_{ijp}w^{-1}_{ijt}\delta_{pq}(\delta_{tr} - \delta_{ts}) + w^{-1}_{ijp}w^{-1}_{iju}\delta_{pq}(\delta_{ur} - \delta_{us}) \\ &\qquad + 2\,w^{-1}_{ijp}w^{-1}_{ijp}(\delta_{pt} - \delta_{pu})(\delta_{pr} - \delta_{ps}) \\ &\qquad \left. - w^{-1}_{ijp}w^{-1}_{ijq}(\delta_{pt} - \delta_{pu})(\delta_{qr} - \delta_{qs}) - w^{-1}_{ijp}w^{-1}_{ijq}(\delta_{pr} - \delta_{ps})(\delta_{qt} - \delta_{qu})\right]\end{aligned}$$

$$
\begin{aligned}
&= 2\, s^4 \sum_{pq} w_{ijq} \big[(1-\delta_{pq}) \big(w_{ijt}^{-1}(\delta_{tr}-\delta_{ts}) - w_{iju}^{-1}(\delta_{ur}-\delta_{us}) \big) \\
&\quad + (\delta_{pt}-\delta_{pu}) \big(w_{ijp}^{-1}(\delta_{pr}-\delta_{ps}) - w_{ijq}^{-1}(\delta_{qr}-\delta_{qs}) \big) \\
&\quad + (\delta_{pr}-\delta_{ps}) \big(w_{ijp}^{-1}(\delta_{pt}-\delta_{pu}) - w_{ijq}^{-1}(\delta_{qt}-\delta_{qu}) \big) \big] \\
&= 2\, s^4 (T_{ij}+1) w_{ij\Sigma} \big[w_{ijt}^{-1}(\delta_{tr}-\delta_{ts}) - w_{iju}^{-1}(\delta_{ur}-\delta_{us}) \big] \\
&= 2\, s^4 (T_{ij}+1) w_{ij\Sigma}\, e(Q_{ij}).
\end{aligned}
$$

Par conséquent,

$$
\begin{aligned}
\mathrm{Cov}(\hat{s}_*^2, Q_{kl}) &= \frac{T_{kl}-1}{\sum_{ij}(T_{ij}-1)} \left(\frac{T_{kl}+1}{T_{kl}-1}\, s^4\, e(Q_{kl}) - s^4\, e(Q_{kl}) \right) \\
&= \frac{2}{\sum_i \sum_j (T_{ij}-1)} s^4\, e(Q_{kl}),
\end{aligned}
$$

ce qui vérifie la seconde équation normale. □

On notera que l'estimateur $\hat{s}^2$ est optimal même si le modèle n'est pas équilibré ($T_{ij} \equiv T$).

4.3 Estimation de $b^{(12)}$

La recherche d'estimateurs optimaux pour les composants de variance $b^{(\cdot)}$ est principalement inspirée des travaux de De Vylder & Goovaerts (1992*b*) et Cossette (1996). Lorsque adaptée à l'estimation du paramètre $b^{(12)}$ de notre modèle CCC à deux facteurs, l'approche privilégiée par ces auteurs consisterait à développer un estimateur de la forme

$$
\sum_{ij} \sum_{kl} \sum_{gh} \sum_{pq} \alpha_{ijklghpq} (X_{ijw} - X_{klw})(X_{ghw} - X_{pqw}).
$$

Or, on observe que, dans un tel cas, le nombre d'équations normales à résoudre est de l'ordre de $(IJ)^4$, où I et J sont les nombres de catégories pour chaque facteur ! Si cela n'a aucune véritable incidence au plan théorique, au plan pratique, cependant, les calculs deviennent rapidement excessivement lourds et longs.

Pour cette raison, nous avons préféré limiter l'espace sur lequel notre estimateur de $b^{(12)}$ est optimal en posant $\{gh\} = \{ij\}$ et $\{pq\} = \{kl\}$ ci-dessus. Le nombre d'équations normales s'en trouve alors ramené à un ordre de $(IJ)^2$, ce qui est encore considérable mais davantage acceptable. Les détails sont présentés dans le théorème suivant, de même que dans sa preuve.

Théorème 4.3. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCC à deux facteurs, alors l'estimateur du paramètre de structure $b^{(12)}$ à variance minimale parmi les estimateurs de la forme*

$$\sum_{ij}\sum_{kl>ij}\alpha_{ijkl}(X_{ijw}-X_{klw})^2 \tag{4.8}$$

avec espérance $b^{(12)}$ est

$$\tilde{b}_*^{(12)} = \mathbf{X}'\mathrm{diag}(\mathbf{A})\mathbf{X}, \tag{4.9}$$

où

$$\mathbf{A} = \frac{b^{(12)}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{C}^{-1}\mathbf{B}, \tag{4.10}$$

et

$$\begin{aligned}
\mathbf{B} &= [\sigma_{ijij}^{(12)} - 2\,\sigma_{ijkl}^{(12)} + \sigma_{klkl}^{(12)}]_{\frac{IJ(IJ-1)}{2}\times 1},\\
\mathbf{C} &= [(\sigma_{ijgh}^{(12)} - \sigma_{ijpq}^{(12)} - \sigma_{klgh}^{(12)} + \sigma_{klpq}^{(12)})^2]_{\frac{IJ(IJ-1)}{2}\times\frac{IJ(IJ-1)}{2}},\\
\mathbf{X} &= [X_{ijw} - X_{klw}]_{\frac{IJ(IJ-1)}{2}\times 1}.
\end{aligned}$$

Dans l'expression de $\mathbf{C}$, les indices i, j, k et l identifient les lignes de la matrice et les indices g, h, p et q, les colonnes. La notation $kl > ij$ est définie ainsi, pour $k \geq i = 1, \ldots, I$ et $j = 1, \ldots, J$:

$$kl > ij \Leftrightarrow \begin{cases} l = 1, \ldots, J & \text{si } k > i,\\ l > j & \text{si } k = i.\end{cases}$$

Preuve. On cherche la projection de $b^{(12)}$ sur $\mathfrak{L}_{b^{(12)}}\{(X_{ijw} - X_{klw})^2\}$. Les équations normales usuelles sont donc, pour $g, p = 1, \ldots, I$ et $h, q = 1, \ldots, J$,

$$\sum_{ij}\sum_{kl}\alpha_{ijkl}\mathrm{E}\left[(X_{ijw}-X_{klw})^2\right] = b^{(12)},$$

$$\sum_{ij}\sum_{kl}\alpha_{ijkl}\mathrm{Cov}((X_{ijw}-X_{klw})^2, (X_{ghw}-X_{pqw})^2) = c\,\mathrm{E}\left[(X_{ghw}-X_{pqw})^2\right].$$

Certaines de ces équations sont toutefois inintéressantes. D'une part, on peut laisser tomber les cas où $\{ij\} = \{kl\}$ ou $\{gh\} = \{pq\}$. D'autre part, une équation où intervient, par exemple, $(X_{12w} - X_{11w})^2$ s'avère redondante avec celle ou intervient $(X_{11w} - X_{12w})^2$. Ce sont précisément ces cas qui se trouvent éliminés lorsque les valeurs des indices sont restreintes à $kl > ij$ et $gh >$

pq tel que défini dans le théorème. Ce faisant, la matrice $\mathbf{C}$, plus bas, est non singulière et le nombre d'équations normales est ramené de $(IJ)^2+1$ à $IJ(IJ-1)/2+1$.

On sait du théorème 3.2 que

$$\sigma_{ijkl}^{(12)} = \delta_{ik}b^{(1)} + \delta_{jl}b^{(2)} + \delta_{ij,kl}\frac{b^{(12)}}{z_{ij}^{(12)}}. \tag{4.11}$$

En utilisant le résultat (2.13) du lemme 2.4 et le fait que les variables $X_{ijw}^0 = X_{ijw} - m$ sont sans excès (lemme 4.1), on peut vérifier que

$$\mathrm{E}\left[X_{ijw}^0 X_{klw}^0 X_{ghw}^0 X_{pqw}^0\right] = \sigma_{ijkl}^{(12)}\sigma_{ghpq}^{(12)} + \sigma_{ijgh}^{(12)}\sigma_{klpq}^{(12)} + \sigma_{ijpq}^{(12)}\sigma_{klgh}^{(12)}. \tag{4.12}$$

À partir des identités du lemme 2.2, on obtient

$$\mathrm{E}\left[(X_{ijw} - X_{klw})^2\right] = \sigma_{ijij}^{(12)} - 2\,\sigma_{ijkl}^{(12)} + \sigma_{klkl}^{(12)}$$

et

$$\begin{aligned}\mathrm{E}\left[(X_{ijw} - X_{klw})^2(X_{ghw} - X_{pqw})^2\right] = {}&\\ &\left(\sigma_{ijij}^{(12)} - 2\,\sigma_{ijkl}^{(12)} + \sigma_{klkl}^{(12)}\right)\left(\sigma_{ghgh}^{(12)} - 2\,\sigma_{ghpq}^{(12)} + \sigma_{pqpq}^{(12)}\right)\\ &\qquad + 2\left(\sigma_{ijgh}^{(12)} - \sigma_{ijpq}^{(12)} - \sigma_{klgh}^{(12)} + \sigma_{klpq}^{(12)}\right)^2.\end{aligned}$$

Avec ces résultats, les équations normales peuvent donc être réécrites ainsi, pour $gh > pq$:

$$\begin{aligned}\sum_{ij}\sum_{kl>ij}\alpha_{ijkl}\left(\sigma_{ijij}^{(12)} - 2\,\sigma_{ijkl}^{(12)} + \sigma_{klkl}^{(12)}\right) &= b^{(12)},\\ \sum_{ij}\sum_{kl>ij}\alpha_{ijkl}\left(\sigma_{ijgh}^{(12)} - \sigma_{ijpq}^{(12)} - \sigma_{klgh}^{(12)} + \sigma_{klpq}^{(12)}\right)^2 &= \frac{c}{2}\left(\sigma_{ghgh}^{(12)} - 2\,\sigma_{ghpq}^{(12)} + \sigma_{pqpq}^{(12)}\right)\end{aligned}$$

soit, sous forme matricielle,

$$\begin{aligned}\mathbf{B}'\mathbf{A} &= b^{(12)},\\ \mathbf{C}\mathbf{A} &= \frac{c}{2}\,\mathbf{B},\end{aligned}$$

avec $\mathbf{A} = [\alpha_{ijkl}]_{IJ(IJ-1)/2\times 1}$. Or, la solution de ce système est (4.10) et les écritures (4.8) et (4.9) sont équivalentes. $\square$

En isolant le terme constant $b^{(12)}/(\mathbf{B}'\mathbf{C}^{-1}\mathbf{B})$ de $\mathbf{A}$ dans (4.9), on remarque que l'estimateur $\tilde{b}_*^{(12)}$ pourrait aussi être défini comme la valeur de $b^{(12)}$ satisfaisant l'équation

$$\mathbf{X}'\mathrm{diag}(\mathbf{C}^{-1}\mathbf{B})\mathbf{X} = \mathbf{B}'\mathbf{C}^{-1}\mathbf{B}. \tag{4.13}$$

Parce qu'elle est, en apparence du moins, plus conventionnelle, nous préférons la formulation (4.9).

Par la présence — on ne peut plus explicite — du paramètre $b^{(12)}$ du côté droit de (4.10), $\tilde{b}_*^{(12)}$ est donc un pseudo-estimateur. Celui-ci dépend également des paramètres $b^{(1)}$ et $b^{(2)}$ par les covariances (4.11). Passons donc immédiatement à l'estimation de ces deux paramètres de structure, avant de s'étendre sur leur calcul.

4.4 Estimation de $b^{(1)}$

Il est tentant, au moment de choisir l'espace sur lequel sera projeté le paramètre $b^{(1)}$ afin d'en déterminer l'estimateur à variance minimale, de jeter son dévolu sur $\mathcal{L}_{b^{(1)}}\{(X_{ijw} - X_{klw})^2\}$, un espace translaté parallèle à $\mathcal{L}_{b^{(12)}}\{(X_{ijw} - X_{klw})^2\}$ assez vaste pour contenir un grand nombre d'estimateurs. Malheureusement, un tel choix se révèle tout à fait inapproprié dans un contexte où tous les paramètres de structure doivent être estimés simultanément. En effet, l'estimateur découlant d'une projection sur cet espace serait identique à $\tilde{b}_*^{(12)}$ à une différence près : $b^{(1)}$ apparaîtrait dans (4.10) en lieu et place de $b^{(12)}$. L'estimateur serait donc la valeur de $b^{(1)}$ satisfaisant (4.13). L'estimation de $b^{(2)}$ étant tout à fait symétrique à celle de $b^{(1)}$, on voit que l'on se retrouverait avec trois paramètres à estimer et une seule équation pour ce faire.

Afin d'éviter une telle situation, on projette plutôt le paramètre $b^{(1)}$ sur un autre espace intuitivement pertinent : $\mathcal{L}_{b^{(1)}}\{(X_{izw} - X_{kzw})^2\}$. Cette approche pose toutefois de nouveaux problèmes.

Lors de la construction de l'estimateur optimal de $b^{(12)}$, le lemme 2.4 a pu être utilisé sans problème en définissant $X_{ijw}^0 = X_{ijw} - m$. Les nouvelles variables aléatoires ainsi créées sont, comme les hypothèses du lemme l'exigent, indépendantes (si $\{ij\} \neq \{kl\}$) et de moyenne nulle. Malheureusement, pour un même facteur de risque, les moyennes pondérées par les facteurs de crédibilité ne peuvent être indépendantes, et en soustraire la moyenne collective m n'arrange évidemment rien. En effet, le théorème 3.2 nous apprend que, par exemple,

$$\mathrm{Cov}(X_{izw}, X_{kzw}) = \sigma_{ik}^{(1)} = \delta_{ik}\frac{b^{(1)}}{z_i^{(1)}} + b^{(2)}S_{ik}^{(1)},$$

et cette covariance n'est jamais nulle tant que les facteurs de crédibilité ne sont pas eux-mêmes tous nuls.

Une solution consisterait à définir de nouvelles moyennes pondérées utilisant aussi les facteurs de crédibilité du type

$$\check{X}_{izw} = m + \Xi_i^{(1)} + \check{\Xi}_z^{(2)} + \check{\Xi}_{iz}^{(12)} + \check{\Xi}_{izw}^{(123)},$$

où

$$\begin{aligned}
\check{\Xi}_j^{(2)} &= \sum_{j=1}^{J} \frac{z_j^{(2)}}{z_\Sigma^{(2)}} \Xi_j^{(2)}, \\
\check{\Xi}_{iz}^{(12)} &= \sum_{j=1}^{J} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} \Xi_{ij}^{(12)}, \\
\check{\Xi}_{izw}^{(123)} &= \sum_{j=1}^{J} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} \Xi_{ijw}^{(123)}.
\end{aligned}$$

Ainsi, en définissant, par exemple, $\dot{X}_{izw} = \check{X}_{izw} - m - \check{\Xi}_z^{(2)}$, on obtient $\operatorname{Cov}(\dot{X}_{izw}, \dot{X}_{kzw}) = 0$ si $i \neq k$. Le charme de cette solution est toutefois rompu par un élément majeur : il est impossible d'exprimer la nouvelle moyenne $\check{X}_{izw}$ comme une combinaison linéaire des X_{ijw} et, ce faisant, de la calculer en pratique ! Retour à la case départ...

En fait, armé d'un peu de patience, la bonne solution est nettement plus simple. Pour obtenir, pour un facteur donné, l'espérance du produit de quatre moyennes pondérées par les facteurs de crédibilité « usuelles » — l'analogue du second résultat du lemme 2.4, donc — il suffit de la calculer au long. Et le résultat, présenté au lemme suivant, surprend agréablement.

Lemme 4.2. *Dans un modèle CCC à deux facteurs, si les moyennes pondérées X_{izw}, $i = 1, \dots, I$, sont sans excès et $X_{izw}^0 = X_{izw} - m$, alors*

$$\mathrm{E}\big[X_{izw}^0 X_{kzw}^0 X_{gzw}^0 X_{pzw}^0\big] = \sigma_{ik}^{(1)} \sigma_{gp}^{(1)} + \sigma_{ig}^{(1)} \sigma_{kp}^{(1)} + \sigma_{ip}^{(1)} \sigma_{kg}^{(1)}.$$

Preuve. On sait que, par définition, $X_{izw} = (z_{i\Sigma}^{(12)})^{-1} \sum_j z_{ij}^{(12)} X_{ijw}$. Par conséquent,

$$\begin{aligned}
&\mathrm{E}\big[X_{izw}^0 X_{kzw}^0 X_{gzw}^0 X_{pzw}^0\big] = \\
&\quad = \sum_{jlhq} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} \frac{z_{kl}^{(12)}}{z_{k\Sigma}^{(12)}} \frac{z_{gh}^{(12)}}{z_{g\Sigma}^{(12)}} \frac{z_{pq}^{(12)}}{z_{p\Sigma}^{(12)}} \mathrm{E}\big[X_{ijw}^0 X_{klw}^0 X_{ghw}^0 X_{pqw}^0\big] \\
&\quad = \sum_{jlhq} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} \frac{z_{kl}^{(12)}}{z_{k\Sigma}^{(12)}} \frac{z_{gh}^{(12)}}{z_{g\Sigma}^{(12)}} \frac{z_{pq}^{(12)}}{z_{p\Sigma}^{(12)}} \left(\sigma_{ijkl}^{(12)} \sigma_{ghpq}^{(12)} + \sigma_{ijgh}^{(12)} \sigma_{klpq}^{(12)} + \sigma_{ijpq}^{(12)} \sigma_{klgh}^{(12)}\right).
\end{aligned}$$

On peut maintenant remplacer les covariances ci-dessus par leurs valeurs (4.11). Par de longues mais élémentaires manipulations, on peut réorganiser les termes pour obtenir

$$
\begin{aligned}
&\mathrm{E}\left[X^0_{izw}X^0_{kzw}X^0_{gzw}X^0_{pzw}\right] = \\
&= \sum_{jlhq} \frac{z^{(12)}_{ij}}{z^{(12)}_{i\Sigma}} \frac{z^{(12)}_{kl}}{z^{(12)}_{k\Sigma}} \frac{z^{(12)}_{gh}}{z^{(12)}_{g\Sigma}} \frac{z^{(12)}_{pq}}{z^{(12)}_{p\Sigma}} \Bigg[b^{(1)}b^{(1)}(\delta_{ik}\delta_{gp} + \delta_{ig}\delta_{kp} + \delta_{ip}\delta_{kg}) \\
&\quad + b^{(1)}b^{(2)}(\delta_{ik}\delta_{hq} + \delta_{gp}\delta_{jl} + \delta_{ig}\delta_{lq} + \delta_{kp}\delta_{jh} + \delta_{ip}\delta_{lh} + \delta_{kg}\delta_{jq}) \\
&\quad + \frac{b^{(1)}b^{(12)}}{z^{(12)}_{ij}}(\delta_{ik}\delta_{gp}\delta_{jl} + \delta_{ig}\delta_{kp}\delta_{jh} + \delta_{ip}\delta_{kg}\delta_{jq}) \\
&\quad + \frac{b^{(1)}b^{(12)}}{z^{(12)}_{kl}}(\delta_{ig}\delta_{kp}\delta_{lq} + \delta_{ip}\delta_{kg}\delta_{lh}) + \frac{b^{(1)}b^{(12)}}{z^{(12)}_{gh}}\delta_{ik}\delta_{gp}\delta_{hq} \\
&\quad + b^{(2)}b^{(2)}(\delta_{jl}\delta_{hq} + \delta_{jh}\delta_{lq} + \delta_{jq}\delta_{lh}) \\
&\quad + \frac{b^{(2)}b^{(12)}}{z^{(12)}_{ij}}(\delta_{ik}\delta_{jl}\delta_{hq} + \delta_{ig}\delta_{jh}\delta_{lq} + \delta_{ip}\delta_{jq}\delta_{lh}) \\
&\quad + \frac{b^{(2)}b^{(12)}}{z^{(12)}_{kl}}(\delta_{kp}\delta_{jh}\delta_{lq} + \delta_{kg}\delta_{jq}\delta_{lh}) + \frac{b^{(2)}b^{(12)}}{z^{(12)}_{gh}}\delta_{gp}\delta_{jl}\delta_{hq} \\
&\quad + b^{(12)}b^{(12)}\left(\delta_{ik}\delta_{gp}\frac{\delta_{jl}\delta_{hq}}{z^{(12)}_{ij}z^{(12)}_{gh}} + \delta_{ig}\delta_{kp}\frac{\delta_{jh}\delta_{lq}}{z^{(12)}_{ij}z^{(12)}_{kl}} + \delta_{ip}\delta_{kg}\frac{\delta_{jq}\delta_{lh}}{z^{(12)}_{ij}z^{(12)}_{kl}}\right)\Bigg] \\
&= b^{(1)}b^{(1)}(\delta_{ik}\delta_{gp} + \delta_{ig}\delta_{kp} + \delta_{ip}\delta_{kg}) \\
&\quad + b^{(1)}b^{(2)}(\delta_{ik}S^{(1)}_{gp} + \delta_{gp}S^{(1)}_{ik} + \delta_{ig}S^{(1)}_{kp} + \delta_{kp}S^{(1)}_{ig} + \delta_{ip}S^{(1)}_{kg} + \delta_{kg}S^{(1)}_{ip}) \\
&\quad + b^{(1)}\frac{b^{(12)}}{z^{(12)}_{i\Sigma}}(\delta_{ik}\delta_{gp} + \delta_{ig}\delta_{kp} + \delta_{ip}\delta_{kg}) \\
&\quad + b^{(1)}\frac{b^{(12)}}{z^{(12)}_{k\Sigma}}(\delta_{ig}\delta_{kp} + \delta_{ip}\delta_{kg}) + b^{(1)}\frac{b^{(12)}}{z^{(12)}_{g\Sigma}}\delta_{ik}\delta_{gp} \\
&\quad + b^{(2)}b^{(2)}(S^{(1)}_{ik}S^{(1)}_{gp} + S^{(1)}_{ig}S^{(1)}_{kp} + S^{(1)}_{ip}S^{(1)}_{kg}) \\
&\quad + b^{(2)}\frac{b^{(12)}}{z^{(12)}_{i\Sigma}}(\delta_{ik}S^{(1)}_{gp} + \delta_{ig}S^{(1)}_{kp} + \delta_{ip}S^{(1)}_{kg}) \\
&\quad + b^{(2)}\frac{b^{(12)}}{z^{(12)}_{k\Sigma}}(\delta_{kp}S^{(1)}_{ig} + \delta_{kg}S^{(1)}_{ip}) + b^{(2)}\frac{b^{(12)}}{z^{(12)}_{g\Sigma}}\delta_{gp}S^{(1)}_{ik} \\
&\quad + b^{(12)}\frac{b^{(12)}}{z^{(12)}_{i\Sigma}}\left(\frac{\delta_{ik}\delta_{gp}}{z^{(12)}_{g\Sigma}} + \frac{\delta_{ig}\delta_{kp}}{z^{(12)}_{k\Sigma}} + \frac{\delta_{ip}\delta_{kg}}{z^{(12)}_{k\Sigma}}\right)\Bigg],
\end{aligned}
$$

où, tel que défini précédemment,

$$S_{ik}^{(1)} = \sum_{j=1}^{J} \frac{z_{ij}^{(12)}}{z_{i\Sigma}^{(12)}} \frac{z_{kj}^{(12)}}{z_{k\Sigma}^{(12)}}.$$

En notant que $b^{(1)} + b^{(12)}/z_{i\Sigma}^{(12)} = b^{(1)}/z_i^{(1)}$ et $\sigma_{ik} = \delta_{ik} b^{(1)}/z_i^{(1)} + b^{(2)} S_{ik}^{(1)}$, on obtient enfin

$$\begin{aligned}
\mathrm{E}\big[X_{izw}^0 X_{kzw}^0 X_{gzw}^0 X_{pzw}^0\big] &= b^{(1)} \frac{b^{(1)}}{z_i^{(1)}} \left(\frac{\delta_{ik}\delta_{gp}}{z_g^{(1)}} + \frac{\delta_{ig}\delta_{kp}}{z_k^{(1)}} + \frac{\delta_{ip}\delta_{kg}}{z_k^{(1)}} \right) \\
&\quad + b^{(2)} \frac{b^{(1)}}{z_i^{(1)}} (\delta_{ik} S_{gp}^{(1)} + \delta_{ig} S_{kp}^{(1)} + \delta_{ip} S_{kg}^{(1)}) \\
&\quad + b^{(2)} \frac{b^{(1)}}{z_k^{(1)}} (\delta_{kp} S_{ig}^{(1)} + \delta_{kg} S_{ip}^{(1)}) + b^{(2)} \frac{b^{(1)}}{z_g^{(1)}} \delta_{gp} S_{ik}^{(1)} \\
&\quad + b^{(2)} b^{(2)} (S_{ik}^{(1)} S_{gp}^{(1)} + S_{ig}^{(1)} S_{kp}^{(1)} + S_{ip}^{(1)} S_{kg}^{(1)}) \\
&= \sigma_{ik}^{(1)} \sigma_{gp}^{(1)} + \sigma_{ig}^{(1)} \sigma_{kp}^{(1)} + \sigma_{ip}^{(1)} \sigma_{kg}^{(1)}.
\end{aligned}$$

□

Si ce résultat surprend un peu, c'est qu'il est exactement du même type que (4.12) alors que les hypothèses nécessaires pour obtenir cette dernière identité ne sont, ici, pas rencontrées. On remarquera toutefois que

$$\mathrm{E}\big[X_{izw}^0 X_{kzw}^0 X_{gzw}^0 X_{pzw}^0\big] \neq 0$$

même si $i \neq k \neq g \neq p$. On a évidemment un résultat similaire pour $\mathrm{E}\big[X_{zjw}^0 X_{zlw}^0 X_{zhw}^0 X_{zqw}^0\big]$.

Nous disposons maintenant de tous les éléments nécessaires à l'élaboration d'un (pseudo-)estimateur de $b^{(1)}$.

Théorème 4.4. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCC à deux facteurs, alors l'estimateur du paramètre de structure $b^{(1)}$ à variance minimale parmi les estimateurs de la forme*

$$\sum_i \sum_{k>i} \alpha_{ik} (X_{izw} - X_{kzw})^2$$

avec espérance $b^{(1)}$ est

$$\tilde{b}_*^{(1)} = \mathbf{X}' \mathrm{diag}(\mathbf{A}) \mathbf{X}, \tag{4.14}$$

où

$$\mathbf{A} = \frac{b^{(1)}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{C}^{-1}\mathbf{B}, \tag{4.15}$$

et

$$\begin{aligned}\mathbf{B} &= [\sigma_{ii}^{(1)} - 2\,\sigma_{ik}^{(1)} + \sigma_{kk}^{(1)}]_{\frac{I(I-1)}{2}\times 1},\\ \mathbf{C} &= [(\sigma_{ig}^{(1)} - \sigma_{ip}^{(1)} - \sigma_{kg}^{(1)} + \sigma_{kp}^{(1)})^2]_{\frac{I(I-1)}{2}\times\frac{I(I-1)}{2}},\\ \mathbf{X} &= [X_{izw} - X_{kzw}]_{\frac{I(I-1)}{2}\times 1}.\end{aligned}$$

Dans l'expression de $\mathbf{C}$*, les indices* i *et* k *identifient les lignes de la matrice et les indices* g *et* p*, les colonnes.*

Preuve. Cette preuve est très similaire à celle du théorème 4.3. On cherche la projection de $b^{(1)}$ sur $\mathcal{L}_{b^{(1)}}\{(X_{izw} - X_{kzw})^2\}$. Dans les équations normales, on se limite aux cas $k > i$ pour les raisons invoquées plus haut. Ces équations sont, pour $g = 1, \dots, I$ et $p > g$,

$$\sum_i \sum_{k>i} \alpha_{ik} \mathrm{E}\big[(X_{izw} - X_{kzw})^2\big] = b^{(1)},$$

$$\sum_i \sum_{k>i} \alpha_{ik} \mathrm{Cov}((X_{izw} - X_{kzw})^2, (X_{gzw} - X_{pzw})^2) = c\,\mathrm{E}\big[(X_{gzw} - X_{pzw})^2\big]\,.$$

Les variables X_{izw} sont sans excès, toujours par le lemme 4.1. À partir des résultats des lemmes 2.2 et 4.2, on trouve

$$\begin{aligned}\mathrm{E}\big[(X_{izw} - X_{kzw})^2(X_{gzw} - X_{pzw})^2\big] &= \\ &= \Big(\sigma_{ii}^{(1)} - 2\,\sigma_{ik}^{(1)} + \sigma_{kk}^{(1)}\Big)\Big(\sigma_{gg}^{(1)} - 2\,\sigma_{gp}^{(1)} + \sigma_{pp}^{(1)}\Big)\\ &\quad + 2\Big(\sigma_{ig}^{(1)} - \sigma_{ip}^{(1)} - \sigma_{kg}^{(1)} + \sigma_{kp}^{(1)}\Big)^2\end{aligned}$$

et l'on réécrit les équations normales ainsi :

$$\sum_i \sum_{k>i} \alpha_{ik} \left(\sigma_{ii}^{(1)} - 2\,\sigma_{ik}^{(1)} + \sigma_{kk}^{(1)}\right) = b^{(1)},$$

$$\sum_i \sum_{k>i} \alpha_{ik} \left(\sigma_{ig}^{(1)} - \sigma_{ip}^{(1)} - \sigma_{kg}^{(1)} + \sigma_{kp}^{(1)}\right)^2 = \frac{c}{2}\Big(\sigma_{gg}^{(1)} - 2\,\sigma_{gp}^{(1)} + \sigma_{pp}^{(1)}\Big)$$

soit, sous forme matricielle,

$$\begin{aligned}\mathbf{B}'\mathbf{A} &= b^{(1)},\\ \mathbf{C}\mathbf{A} &= \frac{c}{2}\mathbf{B},\end{aligned}$$

avec $\mathbf{A} = [\alpha_{ik}]_{I(I-1)/2\times 1}$. La solution de ce système est (4.15), avec de nouvelles définitions pour $\mathbf{B}$ et $\mathbf{C}$ par rapport au théorème 4.3. □

4.5 Estimation de $b^{(2)}$

Le pseudo-estimateur optimal du paramètre $b^{(2)}$ est tout à fait symétrique à celui de $b^{(1)}$. Il serait donc superflu d'en discuter plus longuement que l'énoncé du théorème qui suit.

Théorème 4.5. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCC à deux facteurs, alors l'estimateur du paramètre de structure $b^{(2)}$ à variance minimale parmi les estimateurs de la forme*

$$\sum_j \sum_{l>j} \alpha_{jl}(X_{zjw} - X_{zlw})^2$$

avec espérance $b^{(2)}$ est

$$\tilde{b}_*^{(2)} = \mathbf{X}'\mathrm{diag}(\mathbf{A})\mathbf{X}, \tag{4.16}$$

où

$$\mathbf{A} = \frac{b^{(2)}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{C}^{-1}\mathbf{B}, \tag{4.17}$$

et

$$\begin{aligned}
\mathbf{B} &= [\sigma_{jj}^{(1)} - 2\,\sigma_{jl}^{(1)} + \sigma_{ll}^{(1)}]_{\frac{J(J-1)}{2}\times 1},\\
\mathbf{C} &= [(\sigma_{jh}^{(1)} - \sigma_{jq}^{(1)} - \sigma_{lh}^{(1)} + \sigma_{lq}^{(1)})^2]_{\frac{J(J-1)}{2}\times\frac{J(J-1)}{2}},\\
\mathbf{X} &= [X_{zjw} - X_{zlw}]_{\frac{J(J-1)}{2}\times 1}.
\end{aligned}$$

Dans l'expression de $\mathbf{C}$, les indices j et l identifient les lignes de la matrice et les indices h et q, les colonnes.

Preuve. Identique à celle du théorème 4.4. □

4.6 Calcul et propriétés des estimateurs optimaux

Au-delà des formules, l'estimation des paramètres de structure procède d'une manière précise sur laquelle il convient de s'attarder quelque peu. On s'intéresse également, dans cette section, au signe et à la convergence des pseudo-estimateurs des composants de variance.

La procédure d'estimation dans le modèle CCC est évidemment calquée sur celle du modèle de Bühlmann–Straub. Ainsi, bien que l'estimateur de la moyenne collective soit le premier présenté dans le texte, il est le dernier calculé en pratique. En effet, le calcul de $\widehat{m}_*$ est repoussé à la fin de la procédure d'estimation, lorsque les facteurs de crédibilité sont connus, annihilant de ce fait l'aspect a priori itératif (ou « pseudo ») de l'estimateur.

Le premier paramètre de structure estimé est s^2, le calcul de $\hat{s}^2$ ne posant en soi aucun problème particulier. Viens ensuite le calcul des pseudo-estimateurs optimaux des composants de variance. On l'a vu, les estimateurs $\tilde{b}_*^{(1)}$, $\tilde{b}_*^{(2)}$ et $\tilde{b}_*^{(12)}$ sont tous des fonctions à la fois du paramètre à estimer et des deux autres paramètres. Les estimateurs constituent donc le point fixe d'un système d'équations de la forme

$$\begin{aligned} b^{(12)} &= g_{12}(b^{(12)}, b^{(1)}, b^{(2)}) \\ b^{(1)} &= g_1(b^{(12)}, b^{(1)}, b^{(2)}) \\ b^{(2)} &= g_2(b^{(12)}, b^{(1)}, b^{(2)}). \end{aligned}$$

La technique pour calculer ce point fixe multidimensionnel est bien connue (voir par exemple Burden & Faires (1988, chapitre 10)) : on se donne des valeurs de départ $b_0^{(12)}$, $b_0^{(1)}$ et $b_0^{(2)}$ puis on calcule itérativement de nouvelles valeurs $b_n^{(12)}$, $b_n^{(1)}$ et $b_n^{(2)}$ $(n = 1, 2, \dots)$ à partir de

$$\begin{aligned} b_n^{(12)} &= g_{12}(b_{n-1}^{(12)}, b_{n-1}^{(1)}, b_{n-1}^{(2)}) \\ b_n^{(1)} &= g_1(b_{n-1}^{(12)}, b_{n-1}^{(1)}, b_{n-1}^{(2)}) \\ b_n^{(2)} &= g_2(b_{n-1}^{(12)}, b_{n-1}^{(1)}, b_{n-1}^{(2)}), \end{aligned}$$

jusqu'à convergence vers un point $(\tilde{b}_*^{(12)}, \tilde{b}_*^{(1)}, \tilde{b}_*^{(2)})$. Un critère d'arrêt fréquemment utilisé est le suivant :

$$\max(|b_n^{(12)} - b_{n-1}^{(12)}|, |b_n^{(1)} - b_{n-1}^{(1)}|, |b_n^{(2)} - b_{n-1}^{(2)}|) < \varepsilon,$$

où ε est une « petite » valeur, par exemple 10^{-5}. La convergence peut être accélérée — mais pas nécessairement — avec la méthode de Seidel, qui consiste à utiliser, à chaque étape, la plus récente estimation des paramètres. Les équations ci-dessus deviennent alors

$$\begin{aligned} b_n^{(12)} &= g_{12}(b_{n-1}^{(12)}, b_{n-1}^{(1)}, b_{n-1}^{(2)}) \\ b_n^{(1)} &= g_1(b_n^{(12)}, b_{n-1}^{(1)}, b_{n-1}^{(2)}) \\ b_n^{(2)} &= g_2(b_n^{(12)}, b_n^{(1)}, b_{n-1}^{(2)}). \end{aligned}$$

Les deux principales questions d'intérêt sont maintenant : la procédure ci-dessus converge-t-elle vers un point fixe unique et les coordonnées de ce point fixe sont-elles toutes positives ? Malheureusement, nous n'arrivons pas à donner de réponses théoriques à ces deux questions. Considérons d'abord le cas de la convergence et de l'unicité. La preuve analytique de ces deux propriétés à l'aide du théorème du point fixe multivarié requiert, entre autres choses, de calculer les dérivées partielles par rapport à $b^{(1)}$, $b^{(2)}$ et $b^{(12)}$ des formules (4.9), (4.14) et (4.16) des trois pseudo-estimateurs. Étant donnée la très complexe structure de ces formules, la tâche nous apparaît difficilement réalisable.

D'autre part, montrer que les pseudo-estimateurs sont toujours positifs consiste à montrer que la matrice $\mathrm{diag}(\mathbf{A})$ apparaissant dans chaque formule est définie positive ou, de manière équivalente, que les éléments du vecteur $\mathbf{A}$ sont positifs. Or, dans tous les cas, la matrice de variance-covariance $\mathbf{C}$ est composée d'éléments positifs et est définie positive par définition. Son inverse, la matrice $\mathbf{C}^{-1}$, est également définie positive, d'où le scalaire $\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}$ est non négatif. Le fait que le vecteur $\mathbf{B}$ soit composé de covariances positives[1] ne garantit toutefois pas que le produit $\mathbf{C}^{-1}\mathbf{B}$ résulte en des valeurs positives, d'où la possibilité de coefficients α négatifs.

Des tests numériques peuvent heureusement pallier aux difficultés rencontrées avec l'approche analytique et répondre par l'affirmative avec un satisfaisant niveau de certitude aux deux questions ci-dessus. En effet, lors des nombreux tests numériques menés sur les pseudo-estimateurs optimaux, la procédure d'estimation a toujours convergé vers un point fixe unique indépendant des valeurs de départ. De plus, cette convergence s'est toujours faite vers des valeurs positive ou vers zéro pour chacune des coordonnées du point fixe — pour les différents pseudo-estimateurs, donc. La convergence vers une valeur positive est en général assez rapide : moins d'une vingtaine d'itérations, selon le niveau du critère d'arrêt. La convergence vers zéro est en revanche un peu plus lente, mais elle est aisément décelable. Passé un certain point, 10^{-10} par exemple, on peut être certain que la convergence se fait vers zéro et il convient alors d'arrêter la procédure et de fixer la valeur de l'estimateur à zéro.

Dans le modèle de Bühlmann–Straub, Dubey & Gisler (1981) ont démontré que l'estimateur $\tilde{a}$ converge nécessairement vers zéro lorsqu'un autre estimateur $\hat{a}$ du paramètre a est négatif. Nous n'avons toutefois pas pu établir une telle correspondance entre nos estimateurs optimaux et les estimateurs de Dannenburg ou nos pseudo-estimateurs ad hoc de la section 3.5.

Enfin, toujours par des tests numériques, on parvient à une conclusion similaire quant au signe de l'estimateur optimal de la moyenne collective. En

[1]On remarquera au passage que, dans tous les cas, la diagonale de la matrice $\mathbf{C}$ est formée des éléments du vecteur $\mathbf{B}$ au carré.

effet, on observe que l'estimateur $\widehat{m}_*$ est toujours positif.

On trouvera une plus complète étude numérique des différents estimateurs du modèle CCC au chapitre 8.

4.7 Le modèle à trois facteurs

Dans les sections précédentes, nous avons étudié en détail les estimateurs optimaux des paramètres de structure dans un modèle CCC à deux facteurs. Conformément à ce qui se fait depuis le chapitre 3, on présente maintenant dans cette section les versions de ces estimateurs pour un modèle à trois facteurs. On se contentera cependant d'exposer les différents estimateurs dans une suite de théorèmes. Leur justification et leur développement sont en effet identiques à ceux du modèle à deux facteurs.

On débute, comme d'habitude, avec l'estimateur de la moyenne collective.

Théorème 4.6. *Dans un modèle CCC à trois facteurs, l'estimateur sans biais à variance minimale de m dans la classe des fonctions linéaires des observations est*

$$\widehat{m}_* = \frac{\mathbf{B}'\mathbf{C}^{-1}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{X}, \tag{4.18}$$

où

$$\begin{aligned}
&\mathbf{X} = [\mathbf{X}_i]_{I\times 1}, && \mathbf{X}_i = [\mathbf{X}_{ij}]_{J\times 1}, && \mathbf{X}_{ij} = [X_{ijkw}]_{K\times 1},\\
&\mathbf{B} = [\mathbf{B}_i]_{I\times 1}, && \mathbf{B}_i = [\mathbf{B}_{ij}]_{J\times 1}, && \mathbf{B}_{ij} = [1]_{K\times 1},\\
&\mathbf{C} = [\mathbf{C}_{if}]_{I\times I}, && \mathbf{C}_{if} = [\mathbf{C}_{ifjg}]_{J\times J},
\end{aligned}$$

avec

$$\begin{aligned}
\mathbf{C}_{ifjg} = \delta_{if} b^{(1)}\mathbf{U} + \delta_{jg} b^{(2)}\mathbf{U} + b^{(3)}\mathbf{I} + \delta_{ij,fg} b^{(12)}\mathbf{U}\\
+ \delta_{if} b^{(13)}\mathbf{I} + \delta_{jg} b^{(23)}\mathbf{I} + \delta_{ij,fg} b^{(123)}\mathbf{Z}_{ij}
\end{aligned}$$

et $\mathbf{Z}_{ij}$ une matrice diagonale $K \times K$ formée des réciproques des facteurs de crédibilité $z_{ij1}^{(123)}, \dots, z_{ijK}^{(123)}$:

$$\mathbf{Z}_{ij} = \mathrm{diag}([(z_{ijk}^{(123)})^{-1}]_{K\times 1}).$$

Pour l'estimation optimale des autres paramètres, on fait intervenir l'hypothèse d'excès nul pour les variables aléatoires $\Xi_i^{(1)}, \dots, \Xi_{ijkt}^{(1234)}$. Chaque ratio X_{ijkt} est donc sans excès, de même que toute combinaison linéaire de ceux-ci.

L'estimateur optimal de s^2 a la forme usuelle, tel qu'indiqué dans le théorème suivant.

Théorème 4.7. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCC à trois facteurs, alors l'estimateur du paramètre de structure s^2 à variance minimale parmi toutes les combinaisons du type*

$$\sum_{ijkturs} \alpha_{ijkturs}(X_{ijkt} - X_{ijku})(X_{ijkr} - X_{ijks})$$

avec moyenne s^2 est

$$\hat{s}^2_* = \frac{1}{\sum_{ijk}(T_{ijk} - 1)} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} (T_{ijk} - 1)\hat{s}^2_{ijk}, \tag{4.19}$$

avec

$$\hat{s}^2_{ijk} = \frac{1}{(T_{ijk} - 1)} \sum_{t=1}^{T_{ijk}} w_{ijkt}(X_{ijkt} - X_{ijkw})^2 \tag{4.20}$$

$$= \frac{1}{2(T_{ijk} - 1)w_{ijk\Sigma}} \sum_{t=1}^{T_{ijk}} \sum_{u=1}^{T_{ijk}} w_{ijkt}w_{ijku}(X_{ijkt} - X_{ijku})^2. \tag{4.21}$$

Les raisons invoquées à la section 4.3 pour restreindre l'espace sur lequel les estimateurs des composants de variance sont optimaux demeurent encore ici valables — et même encore plus pertinentes vu l'ajout d'un facteur de risque. De même, le problème de l'utilisation des moyennes pondérées par les facteurs de crédibilité trouve ici une solution similaire. Les calculs sont encore plus longs que dans le modèle à deux facteurs, mais l'on peut néanmoins prouver le lemme suivant.

Lemme 4.3. *Dans un modèle CCC à trois facteurs, si les moyennes pondérées X_{ijzw} et X_{izw}, $i = 1, \dots, I$, $j = 1, \dots, J$, ont un coefficient d'excès nul et $X^0_{ijzw} = X_{ijzw} - m$, $X^0_{izw} = X_{izw} - m$, alors*

$$\mathrm{E}\big[X^0_{ijzw}X^0_{fgzw}X^0_{pqzw}X^0_{stzw}\big] = \sigma^{(12)}_{ijfg}\sigma^{(12)}_{pqst} + \sigma^{(12)}_{ijpq}\sigma^{(12)}_{fgst} + \sigma^{(12)}_{ijst}\sigma^{(12)}_{fgpq}$$

et

$$\mathrm{E}\big[X^0_{izw}X^0_{fzw}X^0_{pzw}X^0_{szw}\big] = \sigma^{(1)}_{if}\sigma^{(1)}_{ps} + \sigma^{(1)}_{ip}\sigma^{(1)}_{fs} + \sigma^{(1)}_{is}\sigma^{(1)}_{fp}.$$

On a évidemment un résultat similaire pour les moyennes pondérées X_{izkw}, X_{zjkw}, X_{jzw} et X_{kzw}. Rappelons de plus que les formules des covariances $\sigma^{(12)}_{ijfg}$ et $\sigma^{(1)}_{if}$ sont présentées au lemme 3.3 à la page 33.

Nous définissons maintenant les pseudo-estimateurs optimaux des paramètres $b^{(123)}$, $b^{(12)}$ et $b^{(1)}$. Les formules des estimateurs des autres composants de variance sont similaires.

Théorème 4.8. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCC à trois facteurs, alors l'estimateur du paramètre de structure $b^{(123)}$ à variance minimale parmi les estimateurs de la forme*

$$\sum_{ijk} \sum_{fgh>ijk} \alpha_{ijkfgh}(X_{ijkw} - X_{fghw})^2 \tag{4.22}$$

avec espérance $b^{(123)}$ est

$$\tilde{b}_*^{(123)} = \mathbf{X}'\mathrm{diag}(\mathbf{A})\mathbf{X}, \tag{4.23}$$

où

$$\mathbf{A} = \frac{b^{(123)}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{C}^{-1}\mathbf{B}, \tag{4.24}$$

et

$$\begin{aligned}
\mathbf{B} &= [\sigma_{ijkijk}^{(123)} - 2\,\sigma_{ijkfgh}^{(123)} + \sigma_{fghfgh}^{(123)}]_{\frac{IJK(IJK-1)}{2}\times 1},\\
\mathbf{C} &= [(\sigma_{ijkpqr}^{(123)} - \sigma_{ijkstu}^{(123)} - \sigma_{fghpqr}^{(123)} + \sigma_{fghstu}^{(123)})^2]_{\frac{IJK(IJK-1)}{2}\times\frac{IJK(IJK-1)}{2}},\\
\mathbf{X} &= [X_{ijkw} - X_{fghw}]_{\frac{IJK(IJK-1)}{2}\times 1}.
\end{aligned}$$

Dans l'expression de $\mathbf{C}$, les indices i, j, k, f, g et h identifient les lignes de la matrice et les indices p, q, r, s, t et u, les colonnes. La notation $fgh > ijk$ est définie ainsi, pour $f \geq i = 1, \ldots, I$, $j = 1, \ldots, J$ et $k = 1, \ldots, K$:

$$fgh > ijk \Leftrightarrow \begin{cases} g = 1, \ldots, J;\ h = 1, \ldots, K & \text{si } f > i, \\ h = 1, \ldots, K & \text{si } f = i \text{ et } g > j, \\ h > k & \text{si } f = i \text{ et } g = j. \end{cases}$$

Théorème 4.9. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCC à trois facteurs, alors l'estimateur du paramètre de structure $b^{(12)}$ à variance minimale parmi les estimateurs de la forme*

$$\sum_{ij} \sum_{fg>ij} \alpha_{ijfg}(X_{ijzw} - X_{fgzw})^2 \tag{4.25}$$

avec espérance $b^{(12)}$ est

$$\tilde{b}_*^{(12)} = \mathbf{X}'\mathrm{diag}(\mathbf{A})\mathbf{X}, \tag{4.26}$$

où

$$\mathbf{A} = \frac{b^{(12)}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{C}^{-1}\mathbf{B}, \tag{4.27}$$

et

$$\begin{aligned}
\mathbf{B} &= [\sigma^{(12)}_{ijij} - 2\,\sigma^{(12)}_{ijfg} + \sigma^{(12)}_{fgfg}]_{\frac{IJ(IJ-1)}{2}\times 1},\\
\mathbf{C} &= [(\sigma^{(12)}_{ijpq} - \sigma^{(12)}_{ijst} - \sigma^{(12)}_{fgpq} + \sigma^{(12)}_{fgst})^2]_{\frac{IJ(IJ-1)}{2}\times\frac{IJ(IJ-1)}{2}},\\
\mathbf{X} &= [X_{ijzw} - X_{fgzw}]_{\frac{IJ(IJ-1)}{2}\times 1}.
\end{aligned}$$

Dans l'expression de $\mathbf{C}$*, les indices* i, j, f *et* g *identifient les lignes de la matrice et les indices* p, q, s *et* t*, les colonnes. La notation* $fg > ij$ *est définie comme au théorème 4.3.*

Théorème 4.10. *Si toutes les variables* Ξ *ont un coefficient d'excès nul dans un modèle CCC à trois facteurs, alors l'estimateur du paramètre de structure* $b^{(1)}$ *à variance minimale parmi les estimateurs de la forme*

$$\sum_i \sum_{f>i} \alpha_{if}(X_{izw} - X_{fzw})^2$$

avec espérance $b^{(1)}$ *est*

$$\tilde{b}^{(1)}_* = \mathbf{X}'\mathrm{diag}(\mathbf{A})\mathbf{X}, \tag{4.28}$$

où

$$\mathbf{A} = \frac{b^{(1)}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{C}^{-1}\mathbf{B}, \tag{4.29}$$

et

$$\begin{aligned}
\mathbf{B} &= [\sigma^{(1)}_{ii} - 2\,\sigma^{(1)}_{if} + \sigma^{(1)}_{ff}]_{\frac{I(I-1)}{2}\times 1},\\
\mathbf{C} &= [(\sigma^{(1)}_{ip} - \sigma^{(1)}_{is} - \sigma^{(1)}_{fp} + \sigma^{(1)}_{fs})^2]_{\frac{I(I-1)}{2}\times\frac{I(I-1)}{2}},\\
\mathbf{X} &= [X_{izw} - X_{fzw}]_{\frac{I(I-1)}{2}\times 1}.
\end{aligned}$$

Dans l'expression de $\mathbf{C}$*, les indices* i *et* f *identifient les lignes de la matrice et les indices* p *et* s*, les colonnes.*

Ceci complète la présentation des estimateurs optimaux des paramètres de structure pour le modèle CCC. Le chapitre suivant traite du modèle CCC généralisé admettant un nombre de facteurs variable par assuré.

Chapitre 5

Modèle à nombre de facteurs variable

Ce chapitre traite d'une généralisation du modèle de crédibilité à classification croisée où le nombre de facteurs de risque peut varier d'un assuré à un autre. Des exemples d'application de ce modèle généralisé ainsi que son idée maîtresse — l'introduction d'une nouvelle catégorie pour chaque facteur de risque — ont déjà été mentionnés à la section 1.2.

Le modèle de base demeure évidemment le modèle CCC étudié jusqu'ici. Au départ, notre modèle CCC généralisé (CCCG) diffère peu du modèle standard en ce qu'il s'agit toujours d'un modèle à P facteurs. La différence se situe dans le nombre de catégories disponibles pour chaque facteur, qui sera dorénavant augmenté de un. Cette catégorie additionnelle représente le fait, pour un assuré du portefeuille, de ne *pas* être influencé par un facteur de risque donné. Par convention, cette catégorie sera numérotée 0. À noter que le $P+1^{\text{e}}$ facteur de risque — le temps — ne possède évidemment pas de catégorie 0.

Comme au chapitre 3, seuls les cas à trois et, parfois, deux facteurs sont étudiés ici, les formules générales faisant l'objet du chapitre 6. Le chapitre débute par un exposé des hypothèses et de l'estimateur de crédibilité du modèle CCCG à trois facteurs. La preuve de ce dernier résultat, similaire à celle de son équivalent dans le modèle CCC, fera encore appel aux projections dans des espaces de Hilbert et prendra à elle seule toute une section. Enfin, trois séries d'estimateurs des paramètres de structure, toutes dérivées des estimateurs du modèle standard, seront proposées : des estimateurs de type Dannenburg, des pseudo-estimateurs ad hoc et des pseudo-estimateurs optimaux.

5.1 Hypothèses et estimateur de crédibilité

Dans le modèle CCCG à trois facteurs, tout assuré d'un portefeuille d'assurance *peut* être affecté par *au plus* trois facteurs de risque ainsi que par toutes les interactions entre les facteurs de risque présents pour cet assuré. S'il n'est pas affecté par un facteur de risque, un assuré est classé dans la catégorie 0 de ce facteur. Le temps est toujours considéré comme un facteur de risque additionnel. La catégorie $i = 0, 1, \dots, I$ du premier facteur de risque est caractérisée par une variable de structure $\Theta_i^{(1)}$, la catégorie $j = 0, 1, \dots, J$ du second par $\Theta_j^{(2)}$ et la catégorie $k = 0, 1, \dots, K$ du troisième par $\Theta_k^{(3)}$. La variable aléatoire X_{ijkt} représente l'expérience observée dans la cellule (ijk) au temps t $(t = 1, \dots, T_{ijk} > 1)$ et, par conséquent, l'ensemble des données $\underline{X}$ forme un vecteur à quatre dimensions. Enfin, à chaque donnée X_{ijkt} est rattaché un poids dit naturel w_{ijkt}. Les hypothèses du modèle sont exactement les mêmes que celles du modèle standard, à la page 21 ; elles ne seront par conséquent pas répétées ici.

Jusqu'ici, rien ne distingue le modèle généralisé du modèle standard, hormis le nombre augmenté de catégories pour chaque facteur. La distinction importante consiste à limiter l'influence sur la moyenne d'un assuré aux seuls facteurs pertinents — c'est-à-dire ceux ne se trouvant pas dans la catégorie 0. Dans la version la plus simple[1] du modèle, on écrit alors le ratio d'expérience d'un assuré ainsi :

$$\begin{aligned} X_{ijkt} = m &+ \overline{\delta}_{0i}\Xi_i^{(1)} + \overline{\delta}_{0j}\Xi_j^{(2)} + \overline{\delta}_{0k}\Xi_k^{(3)} + \overline{\delta}_{0i}\overline{\delta}_{0j}\Xi_{ij}^{(12)} \\ &+ \overline{\delta}_{0i}\overline{\delta}_{0k}\Xi_{ik}^{(13)} + \overline{\delta}_{0j}\overline{\delta}_{0k}\Xi_{jk}^{(23)} + \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k}\Xi_{ijk}^{(123)} + \Xi_{ijkt}^{(1234)}, \end{aligned} \tag{5.1}$$

où on définit, pour abréger et alléger l'écriture, $\overline{\delta}_{0i} = 1 - \delta_{0i}$. Ce delta de Kronecker modifié vaut donc zéro si $i = 0$ et 1 sinon.

Dès qu'un assuré n'est pas affecté par l'un ou l'autre des facteurs de risque, les effets aléatoires relatifs à ce facteur se trouvent donc éliminés de l'expression du ratio de cet assuré. Par exemple, on a

$$\begin{aligned} X_{351t} = m &+ \Xi_3^{(1)} + \Xi_5^{(2)} + \Xi_1^{(3)} + \Xi_{35}^{(12)} \\ &+ \Xi_{31}^{(13)} + \Xi_{51}^{(23)} + \Xi_{351}^{(123)} + \Xi_{351t}^{(1234)}, \end{aligned}$$

mais

$$X_{301t} = m + \Xi_3^{(1)} + \Xi_1^{(3)} + \Xi_{31}^{(13)} + \Xi_{301t}^{(1234)}$$

[1] Comme précédemment, le modèle « complet » ne sera présenté qu'avec les formules générales.

et

$$X_{300t} = m + \Xi_3^{(1)} + \Xi_{300t}^{(1234)}.$$

Soulignons également que la catégorie 0 représente le fait de ne pas être influencé par un facteur de risque et que, par conséquent, ce facteur ne doit pas être considéré comme présent. Concrètement, cela signifie par exemple que les assurés X_{102t} et X_{103u}, $t \neq u$, n'appartiennent pas à un même groupe d'assurés n'ayant pas de deuxième facteur. Leur seul point commun est plutôt de se trouver dans la même catégorie pour le premier facteur. En termes de covariances, on a donc

$$\mathrm{Cov}(X_{102t}, X_{103u}) = b^{(1)}$$

et non

$$\mathrm{Cov}(X_{102t}, X_{103u}) = b^{(1)} + b^{(2)} + b^{(12)}.$$

L'intuition commande que, le modèle généralisé différant peu du modèle standard, les estimateurs de crédibilité dans les deux modèles soient semblables. Le théorème suivant montre que c'est bien le cas, à la différence près que le facteur de crédibilité est en certains endroits remplacé par une nouvelle fonction ζ.

Théorème 5.1. *Soit X_{ijkt} le ratio d'expérience d'un assuré tel que défini en* (5.1). *Le meilleur estimateur — au sens des moindres carrés — de $X_{ijk,T_{ijk}+1}$ est alors*

$$\begin{aligned}\widehat{X}_{ijk,T_{ijk}+1} = m &+ \tilde{z}_{ijk}^{(123)}(X_{ijkw} - m) \\ &+ (1 - \tilde{z}_{ijk}^{(123)}) \Big(\overline{\delta}_{0i} \widehat{\Xi}_i^{(1)} + \overline{\delta}_{0j} \widehat{\Xi}_j^{(2)} + \overline{\delta}_{0k} \widehat{\Xi}_k^{(3)} \\ &+ \overline{\delta}_{0i}\overline{\delta}_{0j} \widehat{\Xi}_{ij}^{(12)} + \overline{\delta}_{0i}\overline{\delta}_{0k} \widehat{\Xi}_{ik}^{(13)} + \overline{\delta}_{0j}\overline{\delta}_{0k} \widehat{\Xi}_{jk}^{(23)} \Big),\end{aligned} \tag{5.2}$$

où

$$\begin{aligned}\widehat{\Xi}_i^{(1)} &= \tilde{z}_i^{(1)} (\widetilde{X}_{izw} - m) \\ &\quad - \tilde{z}_i^{(1)} \sum_{j=0}^{J} \sum_{k=0}^{K} \frac{\zeta_{ij}^{(12)}}{\zeta_{i\Sigma}^{(12)}} \frac{\zeta_{ijk}^{(123)}}{\zeta_{ij\Sigma}^{(123)}} \left(\widehat{\Xi}_j^{(2)} + \widehat{\Xi}_k^{(3)} + \widehat{\Xi}_{ik}^{(13)} + \widehat{\Xi}_{jk}^{(23)} \right) \\ &= \tilde{z}_i^{(1)} (\widetilde{X}_{izw} - m) - \tilde{z}_i^{(1)} \sum_{j=0}^{J} \frac{\zeta_{ij}^{(12)}}{\zeta_{i\Sigma}^{(12)}} \widehat{\Xi}_j^{(2)} \\ &\quad - \tilde{z}_i^{(1)} \sum_{j=0}^{J} \sum_{k=0}^{K} \frac{\zeta_{ij}^{(12)}}{\zeta_{i\Sigma}^{(12)}} \frac{\zeta_{ijk}^{(123)}}{\zeta_{ij\Sigma}^{(123)}} \left(\widehat{\Xi}_k^{(3)} + \widehat{\Xi}_{ik}^{(13)} + \widehat{\Xi}_{jk}^{(23)} \right),\end{aligned} \tag{5.3}$$

$$
\begin{aligned}
\widehat{\Xi}_{ij}^{(12)} &= \tilde{z}_{ij}^{(12)}(\widetilde{X}_{ijzw} - m) \\
&\quad - \tilde{z}_{ij}^{(12)} \sum_{k=0}^{K} \frac{\zeta_{ijk}^{(123)}}{\zeta_{ij\Sigma}^{(123)}} \left(\widehat{\Xi}_i^{(1)} + \widehat{\Xi}_j^{(2)} + \widehat{\Xi}_k^{(3)} + \widehat{\Xi}_{ik}^{(13)} + \widehat{\Xi}_{jk}^{(23)} \right) \\
&= \tilde{z}_{ij}^{(12)}(\widetilde{X}_{ijzw} - m) - \tilde{z}_{ij}^{(12)} \left(\widehat{\Xi}_i^{(1)} + \widehat{\Xi}_j^{(2)} \right) \\
&\quad - \tilde{z}_{ij}^{(12)} \sum_{k=0}^{K} \frac{\zeta_{ijk}^{(123)}}{\zeta_{ij\Sigma}^{(123)}} \left(\widehat{\Xi}_k^{(3)} + \widehat{\Xi}_{ik}^{(13)} + \widehat{\Xi}_{jk}^{(23)} \right),
\end{aligned} \tag{5.4}
$$

avec des expressions similaires à (5.3) *et* (5.4) *pour* $\widehat{\Xi}_j^{(2)}$, $\widehat{\Xi}_k^{(3)}$, $\widehat{\Xi}_{ik}^{(13)}$ *et* $\widehat{\Xi}_{jk}^{(23)}$. *Les nouveaux facteurs de crédibilité sont donnés par*

$$\tilde{z}_i^{(1)} = \overline{\delta}_{0i} \frac{\zeta_{i\Sigma}^{(12)}\, b^{(1)}}{\zeta_{i\Sigma}^{(12)}\, b^{(1)} + b^{(12)}}, \tag{5.5}$$

$$\tilde{z}_{ij}^{(12)} = \overline{\delta}_{0i}\overline{\delta}_{0j} \frac{\zeta_{ij\Sigma}^{(123)}\, b^{(12)}}{\zeta_{ij\Sigma}^{(123)}\, b^{(12)} + b^{(123)}}, \tag{5.6}$$

$$\tilde{z}_{ijk}^{(123)} = \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k} \frac{w_{ijk\Sigma}\, b^{(123)}}{w_{ijk\Sigma}\, b^{(123)} + s^2}, \tag{5.7}$$

avec

$$\zeta_i^{(1)} = (1 - \overline{\delta}_{0i}) \frac{b^{(1)}}{s^2} w_{i\Sigma} + \tilde{z}_i^{(1)}, \tag{5.8}$$

$$\zeta_{ij}^{(12)} = (1 - \overline{\delta}_{0i}\overline{\delta}_{0j}) \frac{b^{(12)}}{s^2} w_{ij\Sigma} + \tilde{z}_{ij}^{(12)}, \tag{5.9}$$

$$\zeta_{ijk}^{(123)} = (1 - \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k}) \frac{b^{(123)}}{s^2} w_{ijk\Sigma} + \tilde{z}_{ijk}^{(123)}. \tag{5.10}$$

Enfin,

$$
\begin{aligned}
\widetilde{X}_{izw} &= \sum_{j=0}^{J} \frac{\zeta_{ij}^{(12)}}{\zeta_{i\Sigma}^{(12)}} \widetilde{X}_{ijzw}, \\
\widetilde{X}_{ijzw} &= \sum_{k=0}^{K} \frac{\zeta_{ijk}^{(123)}}{\zeta_{ij\Sigma}^{(123)}} X_{ijkw}.
\end{aligned}
$$

Toute variable surmontée d'un tilde ~ (autre que les pseudo-estimateurs, bien sûr) fait appel à une particularité du modèle CCCG.

L'estimateur de crédibilité dans un modèle CCC généralisé n'est donc pas très différent a priori de celui du modèle standard. Le facteur de crédibilité est toujours attribué à la différence entre la moyenne individuelle et la moyenne collective, alors que son complément multiplie la somme des estimateurs de crédibilité des effets aléatoires $\Xi_i^{(1)}, \dots, \Xi_{jk}^{(23)}$ pertinents pour chaque assuré. Ces six estimateurs de crédibilité sont toujours calculés en résolvant un système d'équations linéaires dont (5.3) et (5.4) ne sont que deux des six composants. Rappelons qu'il n'est pas nécessaire de calculer la valeur numérique de $\widehat{\Xi}_{ijk}^{(123)}$.

Les grandes différences entre les résultats des deux modèles se situent dans l'expression des facteurs de crédibilité de même que dans l'apparition de la fonction ζ. Tout d'abord, on remarque que, de par les premiers termes à la droite de (5.5)–(5.7), tout facteur de crédibilité est nul dès que l'une des catégories auxquelles il se rapporte est la catégorie 0. Quant aux fonctions ζ, elles remplacent les facteurs de crédibilité lorsque ceux-ci servent de poids. Chacune est la somme d'un facteur de crédibilité et d'une fonction des poids naturels faisant contrepoids aux facteurs de crédibilité nuls. En effet, à l'inverse des facteurs de crédibilité, le premier terme à droite de (5.8)–(5.10) est non nul que si au moins une des catégories considérées est la catégorie 0. Notons que les définitions des fonctions ζ sous la forme (5.8)–(5.10) n'ont pas tant été retenues pour leur simplicité que pour leur interprétation aisée. D'autres formulations sont possibles, comme on le verra dans la preuve du théorème 5.2.

Enfin, on remarque également que les sommes commencent maintenant à zéro plutôt qu'à 1. C'est d'ailleurs ce qui sera implicite lorsque les bornes de sommation ne seront pas indiquées pour alléger la notation.

Avant de passer à la preuve de ce théorème, remarquons que le modèle généralisé se ramène, comme il se doit, au modèle standard lorsque tous les assurés du portefeuille sont influencés par chacun des facteurs de risque.

5.2 Preuve du théorème 5.1

La preuve du théorème 5.1 est en tous points similaire à celle du théorème 3.1 du modèle standard. Comme pour ce dernier, les résultats du théorème seront démontrés à l'aide de projections dans des espaces de Hilbert. Quelques détails des calculs de projections et autres simplifications pourront être omis, étant donné leur similitude avec ceux de la section 3.2.

Tout ensemble de fonctions (mesurables) des variables X_{ijkt} et Ξ forme un sous-espace de $\mathcal{L}^2$. L'estimateur de crédibilité de $X_{ijk,T_{ijk}+1}$ étant égal à l'estimateur de crédibilité de

$$\begin{aligned} m &+ \overline{\delta}_{0i}\Xi_i^{(1)} + \overline{\delta}_{0j}\Xi_j^{(2)} + \overline{\delta}_{0k}\Xi_k^{(3)} + \overline{\delta}_{0i}\overline{\delta}_{0j}\Xi_{ij}^{(12)} \\ &+ \overline{\delta}_{0i}\overline{\delta}_{0k}\Xi_{ik}^{(13)} + \overline{\delta}_{0i}\overline{\delta}_{0k}\Xi_{jk}^{(23)} + \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k}\Xi_{ijk}^{(123)}, \end{aligned}$$

on cherche par conséquent à calculer

$$\begin{aligned}\widehat{X}_{ijk,T_{ijk}+1} = m &+ \overline{\delta}_{0i}\widehat{\Xi}_i^{(1)} + \overline{\delta}_{0j}\widehat{\Xi}_j^{(2)} + \overline{\delta}_{0k}\widehat{\Xi}_k^{(3)}\overline{\delta}_{0i}\overline{\delta}_{0j}\widehat{\Xi}_{ij}^{(12)} \\ &+ \overline{\delta}_{0i}\overline{\delta}_{0k}\widehat{\Xi}_{ik}^{(13)} + \overline{\delta}_{0i}\overline{\delta}_{0k}\widehat{\Xi}_{jk}^{(23)} + \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k}\widehat{\Xi}_{ijk}^{(123)}.\end{aligned} \tag{5.11}$$

Les estimateurs de crédibilité à droite de la précédente égalité sont obtenus par projection des variables correspondantes dans l'espace de Hilbert approprié. Ces calculs font l'objet des sous-sections suivantes.

5.2.1 Projection de $\Xi_{ijk}^{(123)}$

Supposons dorénavant que les indices i, j et k ci-dessous représentent des valeurs particulières des indices de $\Xi^{(123)}$. On a tout d'abord

$$\begin{aligned}\widehat{\Xi}_{ijk}^{(123)} &= \mathrm{pro}(\Xi_{ijk}^{(123)}|\mathcal{L}\{\underline{X}, m\}) \\ &= \mathrm{pro}(\mathrm{pro}(\Xi_{ijk}^{(123)}|\mathcal{L}\{\underline{X}_{ijk}, \Xi_i^{(1)}, \Xi_j^{(2)}, \Xi_k^{(3)}, \\ &\quad \Xi_{ij}^{(12)}, \Xi_{ik}^{(13)}, \Xi_{jk}^{(23)}, m\})|\mathcal{L}\{\underline{X}, m\}).\end{aligned} \tag{5.12}$$

On doit donc d'abord calculer la projection intérieure de (5.12), qui est de la forme

$$\begin{aligned}&\alpha m + \beta^{(1)}\Xi_i^{(1)} + \beta^{(2)}\Xi_j^{(2)} + \beta^{(3)}\Xi_k^{(3)} \\ &\quad + \beta^{(12)}\Xi_{ij}^{(12)} + \beta^{(13)}\Xi_{ik}^{(13)} + \beta^{(23)}\Xi_{jk}^{(23)} + \sum_{t=1}^{T_{ijk}} \kappa_t X_{ijkt}.\end{aligned}$$

Les équations normales sont alors les suivantes, pour $u = 1, \ldots, T_{ijk}$:

$$\begin{aligned}\alpha + \kappa_\Sigma &= 0, & \beta^{(12)} + \overline{\delta}_{0i}\overline{\delta}_{0j}\kappa_\Sigma &= 0, \\ \beta^{(1)} + \overline{\delta}_{0i}\kappa_\Sigma &= 0, & \beta^{(13)} + \overline{\delta}_{0i}\overline{\delta}_{0k}\kappa_\Sigma &= 0, \\ \beta^{(2)} + \overline{\delta}_{0j}\kappa_\Sigma &= 0, & \beta^{(23)} + \overline{\delta}_{0j}\overline{\delta}_{0k}\kappa_\Sigma &= 0, \\ \beta^{(3)} + \overline{\delta}_{0k}\kappa_\Sigma &= 0, & \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k}\kappa_\Sigma b^{(123)} + \kappa_u \frac{s^2}{w_{ijku}} &= \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k} b^{(123)}.\end{aligned}$$

Ainsi, à partir de la dernière équation, on trouve, pour $t = 1, \ldots, T_{ijk}$,

$$\begin{aligned}\kappa_t &= w_{ijkt}\frac{\overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k} b^{(123)}}{\overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k} w_{ijk\Sigma} b^{(123)} + s^2} \\ &= w_{ijkt}\,\overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k}\frac{b^{(123)}}{w_{ijk\Sigma} b^{(123)} + s^2} \\ &= \frac{w_{ijkt}}{w_{ijk\Sigma}}\tilde{z}_{ijk}^{(123)},\end{aligned}$$

d'où, par (5.12),

$$\begin{aligned}\widehat{\Xi}_{ijk}^{(123)} &= \tilde{z}_{ijk}^{(123)}(X_{ijkw}-m) - \tilde{z}_{ijk}^{(123)}\Big(\overline{\delta}_{0i}\widehat{\Xi}_i^{(1)} + \overline{\delta}_{0j}\widehat{\Xi}_j^{(2)} + \overline{\delta}_{0k}\widehat{\Xi}_k^{(3)} \\ &\quad + \overline{\delta}_{0i}\overline{\delta}_{0j}\widehat{\Xi}_{ij}^{(12)} + \overline{\delta}_{0i}\delta_{0k}\widehat{\Xi}_{ik}^{(13)} + \overline{\delta}_{0j}\overline{\delta}_{0k}\widehat{\Xi}_{jk}^{(23)}\Big).\end{aligned}$$

De plus, en remarquant que, par exemple, $\overline{\delta}_{0i}\tilde{z}_{ijk}^{(123)} = \tilde{z}_{ijk}^{(123)}$, on peut écrire

$$\begin{aligned}\widehat{\Xi}_{ijk}^{(123)} &= \tilde{z}_{ijk}^{(123)}(X_{ijkw}-m) \\ &\quad - \tilde{z}_{ijk}^{(123)}\Big(\widehat{\Xi}_i^{(1)} + \widehat{\Xi}_j^{(2)} + \widehat{\Xi}_k^{(3)} + \widehat{\Xi}_{ij}^{(12)} + \widehat{\Xi}_{ik}^{(13)} + \widehat{\Xi}_{jk}^{(23)}\Big).\end{aligned} \tag{5.13}$$

En insérant (5.13) dans (5.11), on obtient (5.2), l'expression de l'estimateur de crédibilité.

5.2.2 Projection de $\Xi_{ij}^{(12)}$

Comme précédemment,

$$\begin{aligned}\widehat{\Xi}_{ij}^{(12)} &= \mathrm{pro}(\Xi_{ij}^{(12)}|\mathcal{L}\{\underline{X}, m\}) \\ &= \mathrm{pro}(\mathrm{pro}(\Xi_{ij}^{(12)}|\mathcal{L}\{\underline{X}_{ij}, \Xi_i^{(1)}, \Xi_j^{(2)}, \underline{\Xi}^{(3)}, \\ &\qquad \underline{\Xi}_i^{(13)}, \underline{\Xi}_j^{(23)}, \underline{\Xi}_{ij}^{(123)}, m\})|\mathcal{L}\{\underline{X}, m\})\end{aligned} \tag{5.14}$$

et la projection intérieure de (5.14), sera de la forme

$$\begin{aligned}&\alpha m + \beta^{(1)}\Xi_i^{(1)} + \beta^{(2)}\Xi_j^{(2)} + \sum_{k=0}^{K}\beta_k^{(3)}\Xi_k^{(3)} + \sum_{k=0}^{K}\beta_k^{(13)}\Xi_{ik}^{(13)} \\ &\quad + \sum_{k=0}^{K}\beta_k^{(23)}\Xi_{jk}^{(23)} + \sum_{k=0}^{K}\beta_k^{(123)}\Xi_{ijk}^{(123)} + \sum_{k=0}^{K}\sum_{t=1}^{T_{ijk}}\kappa_{kt}X_{ijkt}.\end{aligned}$$

Les équations normales sont alors, pour $h = 0, \ldots, K$ et $u = 1, \ldots, T_{ijh}$:

$$\begin{aligned}\alpha + \kappa_{\Sigma} &= 0 & \beta_h^{(13)} + \overline{\delta}_{0i}\overline{\delta}_{0h}\kappa_{h\Sigma} &= 0, \\ \beta^{(1)} + \overline{\delta}_{0i}\kappa_{\Sigma} &= 0 & \beta_h^{(23)} + \overline{\delta}_{0j}\overline{\delta}_{0h}\kappa_{h\Sigma} &= 0, \\ \beta^{(2)} + \overline{\delta}_{0j}\kappa_{\Sigma} &= 0 & \beta_h^{(123)} + \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0h}\kappa_{h\Sigma} &= 0, \\ \beta_h^{(3)} + \overline{\delta}_{0h}\kappa_{h\Sigma} &= 0, & \overline{\delta}_{0i}\overline{\delta}_{0j}\kappa_{\Sigma}b^{(12)} + \kappa_{hu}\frac{s^2}{w_{ijhu}} &= \overline{\delta}_{0i}\overline{\delta}_{0j}b^{(12)}.\end{aligned}$$

Toujours par les mêmes manipulations, on trouve

$$\kappa_{kt} = w_{ijkt}\frac{\overline{\delta}_{0i}\overline{\delta}_{0j}b^{(12)}}{w_{ij\Sigma}b^{(12)} + s^2},$$

d'où, par (5.14),

$$\widehat{\Xi}_{ij}^{(12)} = \frac{\overline{\delta}_{0i}\overline{\delta}_{0j}b^{(12)}}{w_{ij\Sigma}b^{(12)} + s^2}\sum_{k=0}^{K} w_{ijk\Sigma}\Big(X_{ijkw} - m - \overline{\delta}_{0i}\widehat{\Xi}_i^{(1)} - \overline{\delta}_{0j}\widehat{\Xi}_j^{(2)} \\ - \overline{\delta}_{0k}\widehat{\Xi}_k^{(3)} - \overline{\delta}_{0i}\overline{\delta}_{0k}\widehat{\Xi}_{ik}^{(13)} - \overline{\delta}_{0j}\overline{\delta}_{0k}\widehat{\Xi}_{jk}^{(23)} - \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k}\widehat{\Xi}_{ijk}^{(123)}\Big).$$

En remplaçant $\widehat{\Xi}_{ijk}^{(123)}$ par (5.13) dans l'équation précédente et en réarrangeant un peu les termes, on obtient

$$\widehat{\Xi}_{ij}^{(12)} = \frac{\overline{\delta}_{0i}\overline{\delta}_{0j}b^{(12)}}{w_{ij\Sigma}b^{(12)} + s^2}\sum_{k=0}^{K} w_{ijk\Sigma}\Big[(1 - \tilde{z}_{ijk}^{(123)})\Big(X_{ijkw} - m - \overline{\delta}_{0i}\widehat{\Xi}_i^{(1)} \\ - \overline{\delta}_{0j}\widehat{\Xi}_j^{(2)} - \overline{\delta}_{0k}\widehat{\Xi}_k^{(3)} - \overline{\delta}_{0i}\overline{\delta}_{0k}\widehat{\Xi}_{ik}^{(13)} - \overline{\delta}_{0j}\overline{\delta}_{0k}\widehat{\Xi}_{jk}^{(23)}\Big) + \tilde{z}_{ijk}^{(123)}\,\widehat{\Xi}_{ij}^{(12)}\Big],$$

d'où

$$\widehat{\Xi}_{ij}^{(12)}\left[1 - \frac{\overline{\delta}_{0i}\overline{\delta}_{0j}b^{(12)}\sum_{k=0}^{K} w_{ijk\Sigma}\tilde{z}_{ijk}^{(123)}}{w_{ij\Sigma}b^{(12)} + s^2}\right] = \\ \frac{\overline{\delta}_{0i}\overline{\delta}_{0j}b^{(12)}}{w_{ij\Sigma}b^{(12)} + s^2}\sum_{k=0}^{K} w_{ijk\Sigma}(1 - \tilde{z}_{ijk}^{(123)})\Big(X_{ijkw} - m - \overline{\delta}_{0i}\widehat{\Xi}_i^{(1)} \\ - \overline{\delta}_{0j}\widehat{\Xi}_j^{(2)} - \overline{\delta}_{0k}\widehat{\Xi}_k^{(3)} - \overline{\delta}_{0i}\overline{\delta}_{0k}\widehat{\Xi}_{ik}^{(13)} - \overline{\delta}_{0j}\overline{\delta}_{0k}\widehat{\Xi}_{jk}^{(23)}\Big). \quad (5.15)$$

Puisque $\overline{\delta}_{0i}\overline{\delta}_{0j}\tilde{z}_{ijk}^{(123)} = \tilde{z}_{ijk}^{(123)}$, le facteur de $\widehat{\Xi}_{ij}^{(12)}$ est

$$1 - \frac{b^{(12)}\sum_{k=0}^{K} w_{ijk\Sigma}\tilde{z}_{ijk}^{(123)}}{w_{ij\Sigma}b^{(12)} + s^2} = \frac{b^{(12)}\sum_{k=0}^{K} w_{ijk\Sigma}(1 - \tilde{z}_{ijk}^{(123)}) + s^2}{w_{ij\Sigma}b^{(12)} + s^2}. \quad (5.16)$$

Or, comme $\tilde{z}_{ijk}^{(123)} = 0$ dès que i, j ou k vaut 0,

$$\begin{aligned} w_{ijk\Sigma}(1 - \tilde{z}_{ijk}^{(123)}) &= (1 - \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k})w_{ijk\Sigma} + \frac{s^2}{b^{(123)}}\tilde{z}_{ijk}^{(123)} \\ &= \frac{s^2}{b^{(123)}}\,\zeta_{ijk}^{(123)}, \qquad (5.17) \\ \sum_{k=0}^{K} w_{ijk\Sigma}(1 - \tilde{z}_{ijk}^{(123)}) &= \frac{s^2}{b^{(123)}}\,\zeta_{ij\Sigma}^{(123)}, \end{aligned}$$

où $\zeta_{ijk}^{(123)} = (1 - \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k})(b^{(123)}/s^2)w_{ijk\Sigma} + \tilde{z}_{ijk}^{(123)}$. En insérant (5.17) dans (5.16) puis en isolant de nouveau $\widehat{\Xi}_{ij}^{(12)}$ dans (5.15), on obtient finalement

$$
\begin{aligned}
\widehat{\Xi}_{ij}^{(12)} &= \frac{\overline{\delta}_{0i}\overline{\delta}_{0j}b^{(12)}}{\zeta_{ij\Sigma}^{(123)}\, b^{(12)} + s^2} \sum_{k=0}^{K} \zeta_{ijk}^{(123)} \Big(X_{ijkw} - m - \overline{\delta}_{0i}\widehat{\Xi}_i^{(1)} - \overline{\delta}_{0j}\widehat{\Xi}_j^{(2)} \\
&\qquad - \overline{\delta}_{0k}\widehat{\Xi}_k^{(3)} - \overline{\delta}_{0i}\overline{\delta}_{0k}\widehat{\Xi}_{ik}^{(13)} - \overline{\delta}_{0j}\overline{\delta}_{0k}\widehat{\Xi}_{jk}^{(23)} \Big) \\
&= \tilde{z}_{ij}^{(12)} (\widetilde{X}_{ijzw} - m) \\
&\qquad - \tilde{z}_{ij}^{(12)} \sum_{k=0}^{K} \frac{\zeta_{ijk}^{(123)}}{\zeta_{ij\Sigma}^{(123)}} \Big(\widehat{\Xi}_i^{(1)} + \widehat{\Xi}_j^{(2)} + \overline{\delta}_{0k}\widehat{\Xi}_k^{(3)} + \overline{\delta}_{0k}\widehat{\Xi}_{ik}^{(13)} + \overline{\delta}_{0k}\widehat{\Xi}_{jk}^{(23)} \Big),
\end{aligned}
$$

avec

$$
\tilde{z}_{ij}^{(12)} = \overline{\delta}_{0i}\overline{\delta}_{0j} \frac{\zeta_{ij\Sigma}^{(123)}\, b^{(12)}}{\zeta_{ij\Sigma}^{(123)}\, b^{(12)} + s^2}
$$

et

$$
\widetilde{X}_{ijzw} = \sum_{k=0}^{K} \frac{\zeta_{ijk}^{(123)}}{\zeta_{ij\Sigma}^{(123)}} X_{ijkw}.
$$

5.2.3 Projection de $\Xi_i^{(1)}$

La procédure est toujours sensiblement la même :

$$
\begin{aligned}
\widehat{\Xi}_i^{(1)} &= \mathrm{pro}(\Xi_i^{(1)} | \mathcal{L}\{\underline{X}, m\}) \\
&= \mathrm{pro}(\mathrm{pro}(\Xi_i^{(1)} | \mathcal{L}\{\underline{X}_i, \underline{\Xi}^{(2)}, \underline{\Xi}^{(3)}, \\
&\qquad \underline{\Xi}_i^{(12)}, \underline{\Xi}_i^{(13)}, \underline{\Xi}^{(23)}, \underline{\Xi}_i^{(123)}\}) | \mathcal{L}\{\underline{X}, m\})
\end{aligned}
\tag{5.18}
$$

et la projection intérieure est de la forme

$$
\begin{aligned}
&\alpha m + \sum_{j=0}^{J} \beta_j^{(2)} \Xi_j^{(2)} + \sum_{k=0}^{K} \beta_k^{(3)} \Xi_k^{(3)} + \sum_{j=0}^{J} \beta_j^{(12)} \Xi_{ij}^{(12)} + \sum_{k=0}^{K} \beta_{ik}^{(13)} \Xi_{ik}^{(13)} \\
&\quad + \sum_{j=0}^{J} \sum_{k=0}^{K} \beta_{jk}^{(23)} \Xi_{jk}^{(23)} + \sum_{j=0}^{J} \sum_{k=0}^{K} \beta_{ijk}^{(123)} \Xi_{ijk}^{(123)} + \sum_{j=0}^{J} \sum_{k=0}^{K} \sum_{t=1}^{T_{ijk}} \kappa_{jkt} X_{ijkt}.
\end{aligned}
$$

Les équations normales sont les suivantes, pour $g = 0, \ldots, J$, $h = 0, \ldots, K$ et $u = 1, \ldots, T_{igh}$:

$$\begin{aligned} \alpha + \kappa_\Sigma &= 0 & \beta^{(13)}_{ih} + \overline{\delta}_{0i}\overline{\delta}_{0h}\kappa_{h\Sigma} &= 0, \\ \beta^{(2)}_g + \overline{\delta}_{0g}\kappa_{g\Sigma} &= 0, & \beta^{(23)}_{gh} + \overline{\delta}_{0g}\overline{\delta}_{0h}\kappa_{gh\Sigma} &= 0, \\ \beta^{(3)}_h + \overline{\delta}_{0h}\kappa_{h\Sigma} &= 0, & \beta^{(123)}_{igh} + \overline{\delta}_{0i}\overline{\delta}_{0g}\overline{\delta}_{0h}\kappa_{gh\Sigma} &= 0, \\ \beta^{(12)}_g + \overline{\delta}_{0i}\overline{\delta}_{0g}\kappa_{g\Sigma} &= 0, & \overline{\delta}_{0i}\kappa_\Sigma b^{(1)} + \kappa_{ghu}\frac{s^2}{w_{ighu}} &= \overline{\delta}_{0i} b^{(1)}. \end{aligned}$$

On trouve alors aisément

$$\kappa_{jkt} = w_{ijkt}\frac{\overline{\delta}_{0i} b^{(1)}}{w_\Sigma b^{(1)} + s^2},$$

d'où, par (5.18),

$$\begin{aligned} \widehat{\Xi}^{(1)}_i = \frac{\overline{\delta}_{0i} b^{(1)}}{w_{i\Sigma} b^{(1)} + s^2} \sum_{j=0}^{J}\sum_{k=0}^{K} w_{ijk\Sigma}\Big(& X_{ijkw} - m - \overline{\delta}_{0j}\widehat{\Xi}^{(2)}_j - \overline{\delta}_{0k}\widehat{\Xi}^{(3)}_k \\ & - \overline{\delta}_{0i}\overline{\delta}_{0j}\widehat{\Xi}^{(12)}_{ij} - \overline{\delta}_{0i}\overline{\delta}_{0k}\widehat{\Xi}^{(13)}_{ik} - \overline{\delta}_{0j}\overline{\delta}_{0k}\widehat{\Xi}^{(23)}_{jk} - \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k}\widehat{\Xi}^{(123)}_{ijk}\Big). \end{aligned}$$

Comme on l'a fait plus haut, on remplace $\widehat{\Xi}^{(123)}_{ijk}$ ci-dessus par (5.13) et on réunit les termes $\widehat{\Xi}^{(1)}_i$ à la gauche de la nouvelle égalité. Le facteur de $\widehat{\Xi}^{(1)}_i$ est alors

$$1 - \frac{\overline{\delta}_{0i} b^{(1)} \sum_{j=0}^{J}\sum_{k=0}^{K} w_{ijk\Sigma}\tilde{z}^{(123)}_{ijk}}{w_{i\Sigma} b^{(1)} + s^2} = \frac{b^{(1)} \sum_{j=0}^{J}\sum_{k=0}^{K} w_{ijk\Sigma}(1 - \tilde{z}^{(123)}_{ijk}) + s^2}{w_{i\Sigma} b^{(1)} + s^2}.$$

Or,

$$\sum_{j=0}^{J}\sum_{k=0}^{K} w_{ijk\Sigma}(1 - \tilde{z}^{(123)}_{ijk}) = \sum_{j=0}^{J}\sum_{k=0}^{K} \frac{s^2}{b^{(123)}}\zeta^{(123)}_{ijk} = \frac{s^2}{b^{(123)}}\zeta^{(123)}_{i\Sigma}.$$

Par les manipulations maintenant connues, on obtient

$$\begin{aligned} \widehat{\Xi}^{(1)}_i = \frac{\overline{\delta}_{0i} b^{(1)}}{\zeta^{(123)}_{i\Sigma}\, b^{(1)} + b^{(123)}} \sum_{j=0}^{J} \zeta^{(123)}_{ij\Sigma}\Bigg[& \widetilde{X}_{ijzw} - m - \overline{\delta}_{0j}\widehat{\Xi}^{(2)}_j - \overline{\delta}_{0i}\overline{\delta}_{0j}\widehat{\Xi}^{(12)}_{ij} \\ & - \sum_{k=0}^{K} \frac{\zeta^{(123)}_{ijk}}{\zeta^{(123)}_{ij\Sigma}}\Big(\overline{\delta}_{0k}\widehat{\Xi}^{(3)}_k + \overline{\delta}_{0i}\overline{\delta}_{0k}\widehat{\Xi}^{(13)}_{ik} + \overline{\delta}_{0j}\overline{\delta}_{0k}\widehat{\Xi}^{(23)}_{jk}\Big)\Bigg]. \end{aligned}$$

Afin d'obtenir la forme désirée pour $\widehat{\Xi}_i^{(1)}$, on remplace maintenant $\widehat{\Xi}_{ij}^{(12)}$ par sa valeur calculée auparavant. Rappelons que le choix de remplacer par leurs valeurs les variables des niveaux dits « immédiatement supérieurs » est purement arbitraire. Les remplacements pourraient se faire selon une logique différente.

En réunissant ensuite les termes $\widehat{\Xi}_i^{(1)}$, le nouveau facteur de $\widehat{\Xi}_i^{(1)}$ devient

$$1 - \frac{\overline{\delta}_{0i} b^{(1)} \sum_{j=0}^{J} \zeta_{ij\Sigma}^{(123)} \tilde{z}_{ij}^{(12)}}{\zeta_{i\Sigma}^{(123)} b^{(1)} + b^{(123)}} = \frac{b^{(1)} \sum_{j=0}^{J} \zeta_{ij\Sigma}^{(123)} (1 - \tilde{z}_{ij}^{(12)}) + b^{(123)}}{\zeta_{i\Sigma}^{(123)} b^{(1)} + b^{(123)}}.$$

En posant $\zeta_{ij}^{(12)} = (1 - \overline{\delta}_{0i}\overline{\delta}_{0j})(b^{(12)}/b^{(123)})\zeta_{ij\Sigma}^{(123)} + \tilde{z}_{ij}^{(12)}$, on a que

$$\begin{aligned}
\zeta_{ij\Sigma}^{(123)}(1 - \tilde{z}_{ij}^{(12)}) &= (1 - \overline{\delta}_{0i}\overline{\delta}_{0j})\zeta_{ij\Sigma}^{(123)} + \frac{b^{(123)}}{b^{(12)}} \tilde{z}_{ij}^{(12)} \\
&= \frac{b^{(123)}}{b^{(12)}} \zeta_{ij}^{(12)}, \\
\sum_{j=0}^{J} \zeta_{ij\Sigma}^{(123)}(1 - \tilde{z}_{ij}^{(12)}) &= \frac{b^{(123)}}{b^{(12)}} \zeta_{i\Sigma}^{(12)}.
\end{aligned}$$

On remarque toutefois que $(1 - \overline{\delta}_{0i}\overline{\delta}_{0j})\tilde{z}_{ij\Sigma}^{(123)} = 0$ puisque le facteur $\overline{\delta}_{0i}\overline{\delta}_{0j}$ apparaît dans $\tilde{z}_{ijk}^{(123)}$. En vérifiant de plus que

$$\zeta_{ij\Sigma}^{(123)} = \frac{b^{(123)}}{s^2}(w_{ij\Sigma} - \overline{\delta}_{0i}\overline{\delta}_{0j}(w_{ij\Sigma} - w_{ij0})) + \tilde{z}_{ij\Sigma}^{(123)},$$

on voit que l'on peut écrire $\zeta_{ij}^{(12)}$ sous la forme

$$\zeta_{ij}^{(12)} = (1 - \overline{\delta}_{0i}\overline{\delta}_{0j}) \frac{b^{(12)}}{s^2} w_{ij\Sigma} + \tilde{z}_{ij}^{(12)}.$$

Ainsi,

$$\begin{aligned}
\widehat{\Xi}_i^{(1)} &= \frac{\overline{\delta}_{0i} b^{(1)}}{\zeta_{i\Sigma}^{(12)} b^{(1)} + b^{(12)}} \sum_{j=0}^{J} \zeta_{ij}^{(12)} \Bigg[\widetilde{X}_{ijzw} - m \\
&\quad - \sum_{k=0}^{K} \frac{\zeta_{ijk}^{(123)}}{\zeta_{ij\Sigma}^{(123)}} \Big(\overline{\delta}_{0j} \widehat{\Xi}_j^{(2)} + \overline{\delta}_{0k} \widehat{\Xi}_k^{(3)} + \overline{\delta}_{0i}\overline{\delta}_{0k} \widehat{\Xi}_{ik}^{(13)} + \overline{\delta}_{0j}\overline{\delta}_{0k} \widehat{\Xi}_{jk}^{(23)} \Big) \Bigg] \\
&= \tilde{z}_i^{(1)} (\widetilde{X}_{izw} - m) \\
&\quad - \tilde{z}_i^{(1)} \sum_{j=0}^{J} \sum_{k=0}^{K} \frac{\zeta_{ij}^{(12)}}{\zeta_{i\Sigma}^{(12)}} \frac{\zeta_{ijk}^{(123)}}{\zeta_{ij\Sigma}^{(123)}} \Big(\overline{\delta}_{0j} \widehat{\Xi}_j^{(2)} + \overline{\delta}_{0k} \widehat{\Xi}_k^{(3)} + \overline{\delta}_{0i}\overline{\delta}_{0k} \widehat{\Xi}_{ik}^{(13)} + \overline{\delta}_{0j}\overline{\delta}_{0k} \widehat{\Xi}_{jk}^{(23)} \Big),
\end{aligned}$$

avec

$$\tilde{z}_i^{(1)} = \overline{\delta}_{0i} \frac{\zeta_{i\Sigma}^{(12)} b^{(1)}}{\zeta_{i\Sigma}^{(12)} b^{(1)} + b^{(12)}}$$

et

$$\widetilde{X}_{izw} = \sum_{k=0}^{K} \frac{\zeta_{ij}^{(12)}}{\zeta_{i\Sigma}^{(12)}} X_{ijzw}.$$

Dernier détail pour en arriver aux expressions du théorème : on constate que, par symétrie, tous les termes $\widehat{\Xi}$ sont nuls dès qu'au moins un de leurs indices est égal à 0. Par conséquent, on peut omettre les deltas de Kronecker modifiés dans les parties droites des expressions pour $\widehat{\Xi}_{ij}^{(12)}$ et $\widehat{\Xi}_i^{(1)}$ ci-dessus.

Ceci complète la preuve du théorème. □

La prochaine étape consiste à estimer les paramètres de structure de ce modèle. Tel que mentionné en introduction de ce chapitre, nous proposons, dans les sections qui suivent, trois série d'estimateurs pour les composants de variance et la moyenne collective : des estimateurs dérivés de ceux de Dannenburg du modèle CCC, des pseudo-estimateurs ad hoc et des estimateurs optimaux. Comme ce fut le cas précédemment, l'estimateur du paramètre s^2 est traité un peu à part puisqu'il est le même dans chacune de ces trois séries d'estimateurs.

Avant de passer aux formules des estimateurs, on introduit une nouvelle convention d'écriture relative au modèle CCCG, puis énonce les usuelles et toujours indispensables relations de covariance.

Conventions. La convention ci-dessous sera nécessaire pour l'écriture des estimateurs du modèle CCCG et servira à nouveau dans les formules générales.

(C4) Il pourra parfois s'avérer nécessaire de spécifier à quel facteur un ou plusieurs indices 0 se rapportent. Dans de tels cas, les indices porteront en sous-indice le numéro du facteur ou encore la lettre le représentant habituellement. Par exemple, dans un modèle à trois facteurs, $w_{i0_3\Sigma}$ ou $w_{i0_k\Sigma}$ représente la somme des poids naturels pour la catégorie i du premier facteur et 0 du troisième.

5.3 Relations de covariance

Nous nous concentrons sur le même type de relation de covariance qu'à la section 3.3, et ce toujours pour les modèles à deux et trois facteurs. La pondération des moyennes, dans le modèle CCCG, se fait toutefois par les fonctions ζ plutôt que par les facteurs de crédibilité.

Étant donné l'écriture du ratio d'expérience dans le modèle CCC généralisé, on a, pour un modèle à deux facteurs,

$$\operatorname{Cov}(X_{ijt}, X_{klu}) = \overline{\delta}_{0i}\delta_{ik}b^{(1)} + \overline{\delta}_{0j}\delta_{jl}b^{(2)} + \overline{\delta}_{0i}\overline{\delta}_{0j}\delta_{ij,kl}b^{(12)} + \delta_{ijt,klu}\frac{s^2}{w_{ijt}} \quad (5.19)$$

et, pour un modèle à trois facteurs,

$$\begin{aligned}\operatorname{Cov}(X_{ijkt}, X_{fghu}) = {} & \overline{\delta}_{0i}\delta_{if}b^{(1)} + \overline{\delta}_{0j}\delta_{jg}b^{(2)} + \overline{\delta}_{0k}\delta_{kh}b^{(3)} \\ & + \overline{\delta}_{0i}\overline{\delta}_{0j}\delta_{ij,fg}b^{(12)} + \overline{\delta}_{0i}\overline{\delta}_{0k}\delta_{ik,fh}b^{(13)} + \overline{\delta}_{0j}\overline{\delta}_{0k}\delta_{jk,gh}b^{(23)} \\ & + \overline{\delta}_{0i}\overline{\delta}_{0j}\overline{\delta}_{0k}\delta_{ijk,fgh}b^{(123)} + \delta_{ijkt,fghu}\frac{s^2}{w_{ijkt}}. \end{aligned} \quad (5.20)$$

Les principales relations de covariance entre deux moyennes pondérées par les fonctions ζ pour les modèles à deux et trois facteurs sont présentées dans les deux théorèmes qui suivent. La notation de ces covariances est similaire à celle utilisée jusqu'à maintenant.

Théorème 5.2 (Modèle à deux facteurs). *Dans un modèle CCC généralisé à deux facteurs,*

$$\begin{aligned}\operatorname{Cov}(X_{ijw}, X_{klw}) &= \tilde{\sigma}_{ijkl}^{(12)} \\ &= \overline{\delta}_{0i}\delta_{ik}b^{(1)} + \overline{\delta}_{0j}\delta_{jl}b^{(2)} + \delta_{ij,kl}\frac{b^{(12)}}{\zeta_{ij}^{(12)}}, \end{aligned} \quad (5.21)$$

$$\begin{aligned}\operatorname{Cov}(\widetilde{X}_{izw}, \widetilde{X}_{kzw}) &= \tilde{\sigma}_{ik}^{(1)} \\ &= \delta_{ik}\frac{b^{(1)}}{\zeta_i^{(1)}} + b^{(2)}\widetilde{S}_{ik}^{(1)}, \end{aligned} \quad (5.22)$$

$$\begin{aligned}\operatorname{Cov}(\widetilde{X}_{zjw}, \widetilde{X}_{zlw}) &= \tilde{\sigma}_{jl}^{(2)} \\ &= b^{(1)}\widetilde{S}_{jl}^{(2)} + \delta_{jl}\frac{b^{(2)}}{\zeta_j^{(2)}}, \end{aligned} \quad (5.23)$$

avec

$$\widetilde{S}_{ik}^{(1)} = \sum_{j=1}^{J} \frac{\zeta_{ij}^{(12)}}{\zeta_{i\Sigma}^{(12)}}\frac{\zeta_{kj}^{(12)}}{\zeta_{k\Sigma}^{(12)}},$$

$$\widetilde{S}_{jl}^{(2)} = \sum_{i=1}^{I} \frac{\zeta_{ij}^{(12)}}{\zeta_{j\Sigma}^{(12)}}\frac{\zeta_{il}^{(12)}}{\zeta_{l\Sigma}^{(12)}}.$$

On notera que les sommes ci-dessus commencent à 1.

Preuve. Pour obtenir (5.21) à partir de (5.19), il faut montrer que $\overline{\delta}_{0i}\overline{\delta}_{0j}b^{(12)} + s^2/w_{ij\Sigma} = b^{(12)}/\zeta_{ij}^{(12)}$. Or, la version de (5.17) pour un modèle à deux facteurs est

$$\begin{aligned}\zeta_{ij}^{(12)} &= (1-\overline{\delta}_{0i}\overline{\delta}_{0j})\frac{b^{(12)}}{s^2}w_{ij\Sigma} + \tilde{z}_{ij}^{(12)}\\ &= \frac{b^{(12)}}{s^2}w_{ij\Sigma}(1-\tilde{z}_{ij}^{(12)}).\end{aligned}$$

En remarquant que l'on peut, si souhaité et sans rien changer, ajouter $\overline{\delta}_{0i}\overline{\delta}_{0j}$ au dénominateur de $\tilde{z}_{ij}^{(12)}$, on obtient

$$\begin{aligned}\zeta_{ij}^{(12)} &= \frac{b^{(12)}w_{ij\Sigma}}{s^2}\frac{s^2}{\overline{\delta}_{0i}\overline{\delta}_{0j}b^{(12)}w_{ij\Sigma}+s^2}\\ &= \frac{b^{(12)}w_{ij\Sigma}}{\overline{\delta}_{0i}\overline{\delta}_{0j}b^{(12)}w_{ij\Sigma}+s^2},\end{aligned}$$

d'où le résultat recherché. De même, on montre que $\overline{\delta}_{0i}b^{(1)} + b^{(12)}/\zeta_{i\Sigma}^{(12)} = b^{(1)}/\zeta_i^{(1)}$, soit

$$\zeta_i^{(1)} = \frac{\zeta_{i\Sigma}^{(12)}b^{(1)}}{\overline{\delta}_{0i}\zeta_{i\Sigma}^{(12)}b^{(1)}+b^{(12)}}.$$

Pour (5.22), on a donc

$$\begin{aligned}\mathrm{Cov}(\widetilde{X}_{izw},\widetilde{X}_{kzw}) &= \sum_{j=0}^{J}\sum_{l=0}^{J}\frac{\zeta_{ij}^{(12)}}{\zeta_{i\Sigma}^{(12)}}\frac{\zeta_{kl}^{(12)}}{\zeta_{k\Sigma}^{(12)}}\tilde{\sigma}_{ijkl}^{(12)}\\ &= \overline{\delta}_{0i}\delta_{ik}b^{(1)} + b^{(2)}\sum_{j=0}^{J}\frac{\zeta_{ij}^{(12)}}{\zeta_{i\Sigma}^{(12)}}\frac{\zeta_{kj}^{(12)}}{\zeta_{k\Sigma}^{(12)}}\overline{\delta}_{0j} + \delta_{ik}\frac{b^{(12)}}{\zeta_{i\Sigma}^{(12)}}\\ &= \delta_{ik}\frac{b^{(1)}}{\zeta_i^{(1)}} + b^{(2)}\widetilde{S}_{ik}^{(1)}.\end{aligned}$$

La preuve de (5.23) est identique. $\square$

Théorème 5.3 (Modèle à trois facteurs). *Dans un modèle CCC généralisé à trois facteurs,*

$$\begin{aligned}\mathrm{Cov}(\widetilde{X}_{ijkw},\widetilde{X}_{fghw}) &= \tilde{\sigma}_{ijkfgh}^{(123)}\\ &= \overline{\delta}_{0i}\delta_{if}b^{(1)} + \overline{\delta}_{0j}\delta_{jg}b^{(2)} + \overline{\delta}_{0k}\delta_{kh}b^{(3)}\\ &\quad + \overline{\delta}_{0i}\overline{\delta}_{0j}\delta_{ij,fg}b^{(12)} + \overline{\delta}_{0i}\overline{\delta}_{0k}\delta_{ik,fh}b^{(13)}\\ &\quad + \overline{\delta}_{0j}\overline{\delta}_{0k}\delta_{jk,gh}b^{(23)} + \delta_{ijk,fgh}\frac{b^{(123)}}{\zeta_{ijk}^{(123)}},\end{aligned}\tag{5.24}$$

$$
\begin{aligned}
\operatorname{Cov}(\widetilde{X}_{ijzw}, \widetilde{X}_{fgzw}) &= \tilde{\sigma}^{(12)}_{ijfg} \\
&= \overline{\delta}_{0i}\delta_{if}b^{(1)} + \overline{\delta}_{0j}\delta_{jg}b^{(2)} + \delta_{ij,fg}\frac{b^{(12)}}{\zeta^{(12)}_{ij}} \\
&\quad + (b^{(3)} + \overline{\delta}_{0i}\delta_{if}b^{(13)} + \overline{\delta}_{0j}\delta_{jg}b^{(23)})\widetilde{S}^{(12)}_{ijfg},
\end{aligned} \tag{5.25}
$$

$$
\begin{aligned}
\operatorname{Cov}(\widetilde{X}_{izw}, \widetilde{X}_{fzw}) &= \tilde{\sigma}^{(1)}_{if} \\
&= \delta_{if}\frac{b^{(1)}}{\zeta^{(1)}_{i}} + b^{(2)}\widetilde{S}^{(1)}_{if} \\
&\quad + (b^{(3)} + \overline{\delta}_{0i}\delta_{if}b^{(13)})\widetilde{S}^{(12)}_{izfz} + b^{(23)}\widetilde{S}^{(12)}_{izf\star},
\end{aligned} \tag{5.26}
$$

avec

$$
\begin{aligned}
\widetilde{S}^{(12)}_{ijfg} &= \sum_{k=1}^{K} \frac{\zeta^{(123)}_{ijk}}{\zeta^{(123)}_{ij\Sigma}} \frac{\zeta^{(123)}_{fgk}}{\zeta^{(123)}_{fg\Sigma}}, \\
\widetilde{S}^{(12)}_{izfz} &= \sum_{j=0}^{J}\sum_{g=0}^{J} \frac{\zeta^{(12)}_{ij}}{\zeta^{(12)}_{i\Sigma}} \frac{\zeta^{(12)}_{fg}}{\zeta^{(12)}_{f\Sigma}} \widetilde{S}^{(12)}_{ijfg}, \\
\widetilde{S}^{(12)}_{izf\star} &= \sum_{j=1}^{J} \frac{\zeta^{(12)}_{ij}}{\zeta^{(12)}_{i\Sigma}} \frac{\zeta^{(12)}_{fj}}{\zeta^{(12)}_{f\Sigma}} \widetilde{S}^{(12)}_{ijfj}, \\
\widetilde{S}^{(1)}_{if} &= \sum_{j=1}^{J} \frac{\zeta^{(12)}_{ij}}{\zeta^{(12)}_{i\Sigma}} \frac{\zeta^{(12)}_{fj}}{\zeta^{(12)}_{f\Sigma}}
\end{aligned}
$$

et des relations similaires pour $\tilde{\sigma}^{(13)}_{ikfh}$, $\tilde{\sigma}^{(23)}_{jkgh}$, $\tilde{\sigma}^{(2)}_{jg}$ et $\tilde{\sigma}^{(3)}_{kh}$. On notera les valeurs de départ des sommes ci-dessus.

5.4 Estimateurs de Dannenburg généralisés

Les paramètres de structure du modèle CCC généralisé sont évidemment les mêmes que ceux du modèle standard, à savoir la moyenne collective m, la variance temporelle s^2 et les composants de variance $b^{(1)}, \dots, b^{(123)}$. La première série d'estimateurs proposée dans cette section est une généralisation des estimateurs de Dannenburg du modèle CCC. Ces estimateurs apparaissent ici davantage par souci d'exhaustivité que par réel besoin, puisque nous leur préférerons immédiatement les pseudo-estimateurs ad hoc ou optimaux.

Nous adaptons ici les estimateurs de Dannenburg de la section 3.4 à notre modèle CCCG à trois facteurs. Il n'est pas difficile, à partir des formules cidessous, de dériver les estimateurs d'un modèle à deux facteurs.

L'estimateur sans biais de la moyenne collective m demeure la moyenne pondérée par les poids naturels des ratios moyens de tous les assurés :

$$\widehat{m} = X_w = \sum_{i=0}^{I} \sum_{j=0}^{J} \sum_{k=0}^{K} \frac{w_{ijk\Sigma}}{w_\Sigma} X_{ijkw}. \tag{5.27}$$

De même, l'estimateur du paramètre s^2 est facilement généralisé :

$$\begin{aligned}
\hat{s}^2 &= \frac{1}{\sum_{ijk}(T_{ijk}-1)} \sum_{i=0}^{I} \sum_{j=0}^{J} \sum_{k=0}^{K} \sum_{t=1}^{T_{ijk}} w_{ijkt}(X_{ijkt} - X_{ijkw})^2 & (5.28)\\
&= \frac{1}{2\sum_{ijk}(T_{ijk}-1)} \sum_{i=0}^{I} \sum_{j=0}^{J} \sum_{k=0}^{K} \sum_{t=1}^{T_{ijk}} \sum_{u=1}^{T_{ijk}} \frac{w_{ijkt} w_{ijku}}{w_{ijk\Sigma}} (X_{ijkt} - X_{ijku})^2. & (5.29)
\end{aligned}$$

Cet estimateur est toujours à variance minimale sous les hypothèses usuelles d'excès nul ; ce résultat sera confirmé plus loin.

Pour l'estimation sans biais des paramètres $b^{(1)}, \dots, b^{(123)}$, on dispose de trois équations types et de quatre autres équations dérivées des deux premières, pour un total de sept équations. Les trois équations type sont :

$$\begin{aligned}
&\sum_{i=0}^{I} \sum_{j=0}^{J} \tilde{g}_{ij}^{(12)} \left(\sum_{k=0}^{K} \frac{w_{ijk\Sigma}}{w_{ij\Sigma}} (X_{ijkw} - X_{ijw})^2 - K \frac{\hat{s}^2}{w_{ij\Sigma}} \right) = \\
&\qquad \sum_{i=0}^{I} \sum_{j=0}^{J} \tilde{g}_{ij}^{(12)} (b^{(3)} + \overline{\delta}_{0i} b^{(13)} + \overline{\delta}_{0j} b^{(23)} \\
&\qquad\qquad + \overline{\delta}_{0i} \overline{\delta}_{0j} b^{(123)}) \left[1 - \frac{w_{ij0\Sigma}}{w_{ij\Sigma}} - \sum_{k=1}^{K} \left(\frac{w_{ijk\Sigma}}{w_{ij\Sigma}} \right)^2 \right], \qquad (5.30)
\end{aligned}$$

$$\begin{aligned}
&\sum_{i=0}^{I} \tilde{g}_{i}^{(1)} \left(\sum_{j=0}^{J} \sum_{k=0}^{K} \frac{w_{ijk\Sigma}}{w_{i\Sigma}} (X_{ijkw} - X_{iw})^2 - (JK + J + K) \frac{\hat{s}^2}{w_{i\Sigma}} \right) = \\
&\qquad \sum_{i=0}^{I} \tilde{g}_{i}^{(1)} (b^{(2)} + \overline{\delta}_{0i} b^{(12)}) \left[1 - \frac{w_{i0_j\Sigma}}{w_{i\Sigma}} - \sum_{j=1}^{J} \left(\frac{w_{ij\Sigma}}{w_{i\Sigma}} \right)^2 \right] \\
&\qquad + \sum_{i=0}^{I} \tilde{g}_{i}^{(1)} (b^{(3)} + \overline{\delta}_{0i} b^{(13)}) \left[1 - \frac{w_{i0_k\Sigma}}{w_{i\Sigma}} - \sum_{k=1}^{K} \left(\frac{w_{ik\Sigma}}{w_{i\Sigma}} \right)^2 \right]
\end{aligned}$$

$$+\sum_{i=0}^{I}\tilde{g}_i^{(1)}(b^{(23)}+\overline{\delta}_{0i}b^{(123)})\left[1-\frac{w_{i0_j\Sigma}}{w_{i\Sigma}}-\frac{w_{i0_k\Sigma}}{w_{i\Sigma}}\right.$$
$$\left.+\frac{w_{i0\Sigma}}{w_{i\Sigma}}-\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\frac{w_{ijk\Sigma}}{w_{i\Sigma}}\right)^2\right] \quad (5.31)$$

et

$$\sum_{i=0}^{I}\sum_{j=0}^{J}\sum_{k=0}^{K}\frac{w_{ijk\Sigma}}{w_\Sigma}(X_{ijkw}-X_w)^2-(IJK+IJ+IK+JK+I+J+K)\frac{\hat{s}^2}{w_\Sigma}=$$
$$b^{(1)}\left[1-\frac{w_{0_i\Sigma}}{w_\Sigma}-\sum_{i=1}^{I}\left(\frac{w_{i\Sigma}}{w_\Sigma}\right)^2\right]$$
$$+b^{(2)}\left[1-\frac{w_{0_j\Sigma}}{w_\Sigma}-\sum_{j=1}^{J}\left(\frac{w_{j\Sigma}}{w_\Sigma}\right)^2\right]$$
$$+b^{(3)}\left[1-\frac{w_{0_k\Sigma}}{w_\Sigma}-\sum_{k=1}^{K}\left(\frac{w_{k\Sigma}}{w_\Sigma}\right)^2\right]$$
$$+b^{(12)}\left[1-\frac{w_{0_i\Sigma}}{w_\Sigma}-\frac{w_{0_j\Sigma}}{w_\Sigma}+\frac{w_{0_i0_j\Sigma}}{w_\Sigma}-\sum_{i=1}^{I}\sum_{j=1}^{J}\left(\frac{w_{ij\Sigma}}{w_\Sigma}\right)^2\right]$$
$$+b^{(13)}\left[1-\frac{w_{0_i\Sigma}}{w_\Sigma}-\frac{w_{0_k\Sigma}}{w_\Sigma}+\frac{w_{0_i0_k\Sigma}}{w_\Sigma}-\sum_{i=1}^{I}\sum_{k=1}^{K}\left(\frac{w_{ik\Sigma}}{w_\Sigma}\right)^2\right]$$
$$+b^{(23)}\left[1-\frac{w_{0_j\Sigma}}{w_\Sigma}-\frac{w_{0_k\Sigma}}{w_\Sigma}+\frac{w_{0_j0_k\Sigma}}{w_\Sigma}-\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\frac{w_{jk\Sigma}}{w_\Sigma}\right)^2\right]$$
$$+b^{(123)}\left[1-\frac{w_{0_i\Sigma}}{w_\Sigma}-\frac{w_{0_j\Sigma}}{w_\Sigma}-\frac{w_{0_k\Sigma}}{w_\Sigma}+\frac{w_{0_i0_j\Sigma}}{w_\Sigma}+\frac{w_{0_i0_k\Sigma}}{w_\Sigma}\right.$$
$$\left.+\frac{w_{0_j0_k\Sigma}}{w_\Sigma}-\frac{w_{000\Sigma}}{w_\Sigma}-\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\left(\frac{w_{ijk\Sigma}}{w_\Sigma}\right)^2\right], \quad (5.32)$$

où les $\tilde{g}$ sont toujours des poids quelconques dont la somme vaut 1. Les deux premières équations en génèrent chacune deux nouvelles lorsque l'on permute l'ordre de sommation, ce qui donne au total sept équations linéaires pour sept inconnues.

5.5 Pseudo-estimateurs ad hoc

Dans Goulet (1997), l'auteur a étudié par simulation la performance des estimateurs de Dannenburg généralisés ci-dessus. Les résultats se sont montrés plutôt passables pour les raisons déjà évoquées au début de la section 3.5 : coefficient de variation élevé et forte proportion d'estimateurs de variance négatifs. Les pseudo-estimateurs ad hoc qui suivent, dont la performance s'avère meilleure, sont essentiellement identiques à ceux du modèle CCC dans la mesure où les données des catégories 0 ne sont pas prises en compte ailleurs que dans le calcul des moyennes pondérées. Le bien fondé de ceci est expliqué plus bas.

Comme précédemment, l'estimateur de la moyenne collective est traité en premier, suivi des estimateurs des composants de variance du modèle à deux facteurs, puis de ceux du modèle à trois facteurs.

5.5.1 Estimation de la moyenne collective

Nous ne reprenons pas ici la discussion sur les différentes possibilités de pondération par les facteurs de crédibilité dans le modèle CCC, puisque la situation est bien sûr la même dans le modèle CCCG, à l'exception près que la pondération des moyennes individuelles se fait maintenant par les fonctions ζ. Nous optons donc pour le même type d'estimateur de la moyenne collective m, à savoir

$$\widetilde{m} = \widetilde{X}_{zw} = \sum_{i=0}^{I} \frac{\zeta_i^{(1)}}{\zeta_\Sigma^{(1)}} \widetilde{X}_{izw},$$

où $\widetilde{X}_{izw}$ est défini dans l'énoncé du théorème 5.1.

5.5.2 Modèle à deux facteurs

L'estimation ad hoc des composants de variance se fait encore ici en construisant un système d'équations à partir de formes quadratiques des observations. Nous définissons donc ci-dessous les fonctions $\widetilde{\mathrm{MSA}}\mathrm{z}$, $\widetilde{\mathrm{MSB}}\mathrm{z}$ et $\widetilde{\mathrm{MSAB}}\mathrm{z}$ et en calculons les espérances afin d'établir le système d'équations dont les solutions estimeront les paramètres inconnus. La procédure est donc identique à celle étudiée dans le modèle CCC standard.

Le premier réflexe de l'auteur fut de définir des fonctions MS utilisant toutes les données disponibles à un niveau donné. En d'autres termes, les sommes apparaissant dans ces fonctions commençaient à zéro. Il en résultait des formules certes longues et lourdes, mais néanmoins valables d'un strict point de vue algébrique. C'est toutefois lors de la mise en oeuvre informatique de ces formules que les problèmes sont apparus : pléthore d'estimateurs négatifs, convergence vers des valeurs tout à fait improbables ou alors pas de convergence du tout.

Cette distorsion des estimateurs était évidemment une conséquence de l'utilisation fautive des données des catégories 0. En effet, il est inapproprié d'estimer un composant de variance ($b^{(12)}$, par exemple) à partir de données qui ne sont précisément pas affectées par ce composant de variance (X_{ijw} avec $i = 0$ ou $j = 0$). Les formules présentées dans cette sous-section ainsi qu'à la suivante sont par conséquent très semblables à celles du modèle CCC, puisque les différentes sommes commencent à 1 plutôt qu'à zéro. Notons toutefois que les données des catégories 0 n'en sont pas pour autant entièrement écartées ; elles apparaissent en effet toujours dans le calcul des différentes moyennes pondérées par les fonctions ζ.

Les formules ci-dessous sont clairement équivalentes à celles de la section 3.5 si le nombre de facteurs est le même pour tous les assurés.

En premier lieu, on définit donc

$$\begin{aligned}\widetilde{\mathrm{MSABz}} = c_1 \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{l=1}^{J} \zeta_{ij}^{(12)} \zeta_{il}^{(12)} (X_{ijw} - X_{ilw})^2 \\ + c_2 \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{I} \zeta_{ij}^{(12)} \zeta_{kj}^{(12)} (X_{ijw} - X_{kjw})^2 \\ - c_3 \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{I} \sum_{l=1}^{J} \zeta_{ij}^{(12)} \zeta_{kl}^{(12)} (X_{ijw} - X_{klw})^2,\end{aligned} \tag{5.33}$$

où les constantes c_1, c_2 et c_3 sont les solutions du système d'équations usuel

$$\begin{aligned}(Z_{(1)}^{(12)} - Z_{(12)}^{(12)})\, c_1 - (Z^{(12)} - Z_{(2)}^{(12)})\, c_3 = 0 \\ (Z_{(2)}^{(12)} - Z_{(12)}^{(12)})\, c_2 - (Z^{(12)} - Z_{(1)}^{(12)})\, c_3 = 0 \\ 2\, z_{\Sigma}^{(12)} \left[(J-1)\, c_1 + (I-1)\, c_2 - (IJ-1)\, c_3 \right] = 1,\end{aligned}$$

avec, cependant,

$$Z^{(12)} = \Big(\sum_{i=1}^{I} \sum_{j=1}^{I} \zeta_{ij}^{(12)} \Big)^2, \qquad Z_{(1)}^{(12)} = \sum_{i=1}^{I} \Big(\sum_{j=1}^{J} \zeta_{ij}^{(12)} \Big)^2,$$

$$Z_{(2)}^{(12)} = \sum_{j=1}^{J} \Big(\sum_{i=1}^{I} \zeta_{ij}^{(12)} \Big)^2, \qquad Z_{(12)}^{(12)} = \sum_{i=1}^{I} \sum_{j=1}^{J} \Big(\zeta_{ij}^{(12)} \Big)^2.$$

(On remarquera, en effet, que $\zeta_{i\Sigma}^{(12)} = \sum_{j=0}^{J} \zeta_{ij}^{(12)}$, etc.)

De plus, on définit

$$\widetilde{\mathrm{MSAz}} = \frac{1}{2\,(I-1)\sum_{i=1}^{I} \zeta_i^{(1)}} \sum_{i=1}^{I}\sum_{k=1}^{I} \zeta_i^{(1)}\zeta_k^{(1)}(\widetilde{X}_{izw} - \widetilde{X}_{kzw})^2 \tag{5.34}$$

et

$$\widetilde{\mathrm{MSBz}} = \frac{1}{2\,(J-1)\sum_{j=1}^{J} \zeta_j^{(2)}} \sum_{j=1}^{J}\sum_{l=1}^{J} \zeta_j^{(2)}\zeta_l^{(2)}(\widetilde{X}_{zjw} - \widetilde{X}_{zlw})^2. \tag{5.35}$$

Les espérances de ces trois fonctions sont données dans le théorème suivant.

Théorème 5.4. *Les espérances des fonctions* $\widetilde{\mathrm{MSABz}}$, $\widetilde{\mathrm{MSAz}}$ *et* $\widetilde{\mathrm{MSBz}}$ *ci-dessus sont :*

$$\mathrm{E}\Big[\widetilde{\mathrm{MSABz}}\Big] = b^{(12)}, \tag{5.36}$$

$$\mathrm{E}\Big[\widetilde{\mathrm{MSAz}}\Big] = b^{(1)} + b^{(2)}\widetilde{K}^{(1)}, \tag{5.37}$$

$$\mathrm{E}\Big[\widetilde{\mathrm{MSBz}}\Big] = b^{(1)}\widetilde{K}^{(2)} + b^{(2)}, \tag{5.38}$$

où

$$\widetilde{K}^{(1)} = \frac{1}{(I-1)\sum_{i=1}^{I} \zeta_i^{(1)}} \sum_{i=1}^{I}\sum_{k=1}^{I} \zeta_i^{(1)}\zeta_k^{(1)}(\widetilde{S}_{ii}^{(1)} - \widetilde{S}_{ik}^{(1)}), \tag{5.39}$$

$$\widetilde{K}^{(2)} = \frac{1}{(J-1)\sum_{j=1}^{J} \zeta_j^{(2)}} \sum_{j=1}^{J}\sum_{l=1}^{J} \zeta_j^{(2)}\zeta_l^{(2)}(\widetilde{S}_{jj}^{(2)} - \widetilde{S}_{jl}^{(2)}). \tag{5.40}$$

Preuve. La preuve de ce théorème est en tous points similaire à celle du théorème 3.4. □

Les paramètres $b^{(1)}$, $b^{(2)}$ et $b^{(12)}$ peuvent donc être estimés sans biais en résolvant le système d'équations

$$\begin{bmatrix} \widetilde{\mathrm{MSAz}} \\ \widetilde{\mathrm{MSBz}} \\ \widetilde{\mathrm{MSABz}} \end{bmatrix} = \begin{bmatrix} 1 & \widetilde{K}^{(1)} & 0 \\ \widetilde{K}^{(2)} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{b}^{(1)} \\ \tilde{b}^{(2)} \\ \tilde{b}^{(12)} \end{bmatrix}.$$

Dans le cas simple du modèle à deux facteurs, cette solution est

$$\begin{bmatrix} \tilde{b}^{(1)} \\ \tilde{b}^{(2)} \\ \tilde{b}^{(12)} \end{bmatrix} = \frac{1}{1 - \widetilde{K}^{(1)}\widetilde{K}^{(2)}} \begin{bmatrix} 1 & -\widetilde{K}^{(1)} & 0 \\ -\widetilde{K}^{(2)} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \widetilde{\mathrm{MSAz}} \\ \widetilde{\mathrm{MSBz}} \\ \widetilde{\mathrm{MSABz}} \end{bmatrix}. \tag{5.41}$$

De par la présence des paramètres à estimer dans les formules de $\mathrm{MS\widetilde{A}z}$, $\mathrm{M\widetilde{SB}z}$ et $\mathrm{M\widetilde{SAB}z}$, $\tilde{b}^{(1)}$, $\tilde{b}^{(2)}$ et $\tilde{b}^{(12)}$ sont des pseudo-estimateurs obtenus par l'évaluation itérative de (5.41) jusqu'à convergence vers des valeurs vérifiant approximativement l'égalité.

5.5.3 Modèle à trois facteurs

On se contente, comme cela en est devenu l'habitude avec le modèle à trois facteurs, de définir une seule fonction MS par niveau d'interaction entre les facteurs, les autres fonctions leur étant similaires. En se basant sur les formules correspondantes dans le modèle CCC et en se permettant un intelligible abus de notation pour alléger l'écriture des sommes, on a

$$\begin{aligned}
\mathrm{MS\widetilde{ABC}z} = c_1 &\sum_{ijkh=1} \zeta_{ijk}^{(123)}\zeta_{ijh}^{(123)}(X_{ijkw} - X_{ijhw})^2 \\
&+ c_2 \sum_{ijkg=1} \zeta_{ijk}^{(123)}\zeta_{igk}^{(123)}(X_{ijkw} - X_{igkw})^2 \\
&+ c_3 \sum_{ijkf=1} \zeta_{ijk}^{(123)}\zeta_{fjk}^{(123)}(X_{ijkw} - X_{fjkw})^2 \\
&- c_4 \sum_{ijkgh=1} \zeta_{ijk}^{(123)}\zeta_{igh}^{(123)}(X_{ijkw} - X_{ighw})^2 \\
&- c_5 \sum_{ijkfh=1} \zeta_{ijk}^{(123)}\zeta_{fjh}^{(123)}(X_{ijkw} - X_{fjhw})^2 \\
&- c_6 \sum_{ijkfg=1} \zeta_{ijk}^{(123)}\zeta_{fgk}^{(123)}(X_{ijkw} - X_{fgkw})^2 \\
&+ c_7 \sum_{ijkfgh=1} \zeta_{ijk}^{(123)}\zeta_{fgh}^{(123)}(X_{ijkw} - X_{fghw})^2,
\end{aligned} \tag{5.42}$$

où les constantes $c_1, \ldots, c_7$ sont les solutions du système d'équations rencontré auparavant :

$$\begin{aligned}
{}_2Z_{13}c_2 + {}_1Z_{23}c_3 - {}_2Z_1c_4 - {}_1Z_2c_5 - {}_{12}Z_3c_6 + {}_{12}Zc_7 &= 0 \\
{}_3Z_{12}c_1 + {}_1Z_{23}c_3 - {}_3Z_1c_4 - {}_{13}Z_2c_5 - {}_1Z_3c_6 + {}_{13}Zc_7 &= 0 \\
{}_3Z_{12}c_1 + {}_2Z_{13}c_2 - {}_{23}Z_1c_4 - {}_3Z_2c_5 - {}_2Z_3c_6 + {}_{23}Zc_7 &= 0 \\
{}_1Z_{23}c_3 - {}_1Z_2c_5 - {}_1Z_3c_6 + {}_1Zc_7 &= 0 \\
{}_2Z_{13}c_2 - {}_2Z_1c_4 - {}_2Z_3c_6 + {}_2Zc_7 &= 0 \\
{}_3Z_{12}c_1 - {}_3Z_1c_4 - {}_3Z_2c_5 + {}_3Zc_7 &= 0
\end{aligned}$$

$$\begin{aligned}
2\, z_\Sigma^{(123)} \big[&(K-1)c_1 + (J-1)c_2 + (I-1)c_3 - (JK-1)c_4 \\
&-(IK-1)c_5 - (IJ-1)c_6 + (IJK-1)c_7\big] = 1,
\end{aligned}$$

avec ${}_2Z_1 = Z^{(123)}_{(1)} - Z^{(123)}_{(12)}$ et

$$Z^{(123)} = \Big(\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\zeta^{(123)}_{ijk}\Big)^2, \qquad Z^{(123)}_{(1)} = \sum_{i=1}^{I}\Big(\sum_{j=1}^{J}\sum_{k=1}^{K}\zeta^{(123)}_{ijk}\Big)^2,$$
$$Z^{(123)}_{(12)} = \sum_{i=1}^{I}\sum_{j=1}^{J}\Big(\sum_{k=1}^{K}\zeta^{(123)}_{ijk}\Big)^2, \qquad Z^{(123)}_{(123)} = \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(\zeta^{(123)}_{ijk})^2,$$

ainsi de suite.

On définit par la suite

$$\begin{aligned}\mathrm{MS\widetilde{AB}z} = {}& d_1 \sum_{ijg=1} \zeta^{(12)}_{ij}\zeta^{(12)}_{ig}(\widetilde{X}_{ijzw} - \widetilde{X}_{igzw})^2 \\ & + d_2 \sum_{ijf=1} \zeta^{(12)}_{ij}\zeta^{(12)}_{fj}(\widetilde{X}_{ijzw} - \widetilde{X}_{fjzw})^2 \\ & - d_3 \sum_{ijfg=1} \zeta^{(12)}_{ij}\zeta^{(12)}_{fg}(\widetilde{X}_{ijzw} - \widetilde{X}_{fgzw})^2,\end{aligned} \tag{5.43}$$

où les constantes d_1, d_2 et d_3 sont les solutions de

$$\begin{aligned}(Z^{(12)}_{(1)} - Z^{(12)}_{(12)})\, d_1 - (Z^{(12)} - Z^{(12)}_{(2)})\, d_3 &= 0 \\ (Z^{(12)}_{(2)} - Z^{(12)}_{(12)})\, d_2 - (Z^{(12)} - Z^{(12)}_{(1)})\, d_3 &= 0 \\ 2\, z^{(12)}_{\Sigma}\, [(J-1)\, d_1 + (I-1)\, d_2 - (IJ-1)\, d_3] &= 1,\end{aligned}$$

avec

$$Z^{(12)} = \Big(\sum_{i=1}^{I}\sum_{j=1}^{I}\zeta^{(12)}_{ij}\Big)^2, \qquad Z^{(12)}_{(1)} = \sum_{i=1}^{I}\Big(\sum_{j=1}^{J}\zeta^{(12)}_{ij}\Big)^2,$$
$$Z^{(12)}_{(2)} = \sum_{j=1}^{J}\Big(\sum_{i=1}^{I}\zeta^{(12)}_{ij}\Big)^2, \qquad Z^{(12)}_{(12)} = \sum_{i=1}^{I}\sum_{j=1}^{J}(\zeta^{(12)}_{ij})^2.$$

Enfin, on a

$$\mathrm{MS\widetilde{A}z} = \frac{1}{2\,(I-1)\sum_{i=1}^{I}\zeta^{(1)}_i}\sum_{i=1}^{I}\sum_{f=1}^{I}\zeta^{(1)}_i\zeta^{(1)}_f(\widetilde{X}_{izw} - \widetilde{X}_{fzw})^2. \tag{5.44}$$

Théorème 5.5. *Les espérances des fonctions* $\mathrm{MS\widetilde{ABC}z}$, $\mathrm{MS\widetilde{AB}z}$ *et* $\mathrm{MS\widetilde{A}z}$

ci-dessus sont :

$$\mathrm{E}\Big[\mathrm{MS}\widetilde{\mathrm{ABCz}}\Big] = b^{(123)} \tag{5.45}$$

$$\mathrm{E}\Big[\mathrm{MS}\widetilde{\mathrm{ABz}}\Big] = b^{(12)} + b^{(3)}\widetilde{K}^{(12)}_{(3)} + b^{(13)}\widetilde{K}^{(12)}_{(13)} + b^{(23)}\widetilde{K}^{(12)}_{(23)} \tag{5.46}$$

$$\mathrm{E}\Big[\mathrm{MS}\widetilde{\mathrm{Az}}\Big] = b^{(1)} + b^{(2)}\widetilde{K}^{(1)}_{(2)} + b^{(3)}\widetilde{K}^{(1)}_{(3)} + b^{(13)}\widetilde{K}^{(1)}_{(13)} + b^{(23)}\widetilde{K}^{(1)}_{(23)} \tag{5.47}$$

avec

$$\begin{aligned}
\widetilde{K}^{(1)}_{(2)} &= \frac{1}{(I-1)\sum_{i=1}^{I}\zeta_i^{(1)}} \sum_{i=1}^{I}\sum_{f=1}^{I} \zeta_i^{(1)}\zeta_f^{(1)}(\widetilde{S}^{(1)}_{ii} - \widetilde{S}^{(1)}_{if}),\\
\widetilde{K}^{(1)}_{(3)} &= \frac{1}{(I-1)\sum_{i=1}^{I}\zeta_i^{(1)}} \sum_{i=1}^{I}\sum_{f=1}^{I} \zeta_i^{(1)}\zeta_f^{(1)}(\widetilde{S}^{(12)}_{iziz} - \widetilde{S}^{(12)}_{izfz}),\\
\widetilde{K}^{(1)}_{(13)} &= \frac{1}{(I-1)\sum_{i=1}^{I}\zeta_i^{(1)}} \sum_{i=1}^{I}\sum_{f=1}^{I} \zeta_i^{(1)}\zeta_f^{(1)}(\widetilde{S}^{(12)}_{iziz} - \delta_{if}\widetilde{S}^{(12)}_{izfz}),\\
\widetilde{K}^{(1)}_{(23)} &= \frac{1}{(I-1)\sum_{i=1}^{I}\zeta_i^{(1)}} \sum_{i=1}^{I}\sum_{f=1}^{I} \zeta_i^{(1)}\zeta_f^{(1)}(\widetilde{S}^{(12)}_{izi\star} - \widetilde{S}^{(12)}_{izf\star})
\end{aligned}$$

et

$$\begin{aligned}
\widetilde{K}^{(12)}_{(3)} &= d_1 \sum_{ijg=1} \zeta^{(12)}_{ij}\zeta^{(12)}_{ig}(\widetilde{S}^{(12)}_{ijij} - \widetilde{S}^{(12)}_{ijig})\\
&\quad + d_2 \sum_{ijf=1} \zeta^{(12)}_{ij}\zeta^{(12)}_{fj}(\widetilde{S}^{(12)}_{ijij} - \widetilde{S}^{(12)}_{ijfj})\\
&\quad - d_3 \sum_{ijfg=1} \zeta^{(12)}_{ij}\zeta^{(12)}_{fg}(\widetilde{S}^{(12)}_{ijij} - \widetilde{S}^{(12)}_{ijfg}),\\
\widetilde{K}^{(12)}_{(13)} &= d_1 \sum_{ijg=1} \zeta^{(12)}_{ij}\zeta^{(12)}_{ig}(\widetilde{S}^{(12)}_{ijij} - \widetilde{S}^{(12)}_{ijig})\\
&\quad + d_2 \sum_{ijf=1} \zeta^{(12)}_{ij}\zeta^{(12)}_{fj}(\widetilde{S}^{(12)}_{ijij} - \delta_{if}\widetilde{S}^{(12)}_{ijfj})\\
&\quad - d_3 \sum_{ijfg=1} \zeta^{(12)}_{ij}\zeta^{(12)}_{fg}(\widetilde{S}^{(12)}_{ijij} - \delta_{if}\widetilde{S}^{(12)}_{ijfg}),
\end{aligned}$$

$$\begin{aligned}
\widetilde{K}_{(23)}^{(12)} &= d_1 \sum_{ijg=1} \zeta_{ij}^{(12)} \zeta_{ig}^{(12)} (\widetilde{S}_{ijij}^{(12)} - \delta_{jg} \widetilde{S}_{ijig}^{(12)}) \\
&\quad + d_2 \sum_{ijf=1} \zeta_{ij}^{(12)} \zeta_{fj}^{(12)} (\widetilde{S}_{ijij}^{(12)} - \widetilde{S}_{ijfj}^{(12)}) \\
&\quad - d_3 \sum_{ijfg=1} \zeta_{ij}^{(12)} \zeta_{fg}^{(12)} (\widetilde{S}_{ijij}^{(12)} - \delta_{jg} \widetilde{S}_{ijfg}^{(12)}).
\end{aligned}$$

On notera que la fonction $\widetilde{S}_{izfz}^{(12)}$, dont la définition est donnée au théorème 5.3, est une moyenne pondérée par les fonctions ζ sur *toutes* les catégories des fonctions $\widetilde{S}_{ijfg}^{(12)}$.

Les pseudo-estimateurs ad hoc des composants de variance d'un modèle CCCG à trois facteurs sont donc donnés par la solution du système d'équations (en apparence linéaire) suivant :

$$\begin{bmatrix} \widetilde{\mathrm{MSAz}} \\ \widetilde{\mathrm{MSBz}} \\ \widetilde{\mathrm{MSCz}} \\ \widetilde{\mathrm{MSABz}} \\ \widetilde{\mathrm{MSACz}} \\ \widetilde{\mathrm{MSBCz}} \\ \widetilde{\mathrm{MSABCz}} \end{bmatrix} = \begin{bmatrix} 1 & \widetilde{K}_{(2)}^{(1)} & \widetilde{K}_{(3)}^{(1)} & 0 & \widetilde{K}_{(13)}^{(1)} & \widetilde{K}_{(23)}^{(1)} & 0 \\ \widetilde{K}_{(1)}^{(2)} & 1 & \widetilde{K}_{(3)}^{(2)} & 0 & \widetilde{K}_{(13)}^{(2)} & \widetilde{K}_{(23)}^{(2)} & 0 \\ \widetilde{K}_{(1)}^{(3)} & \widetilde{K}_{(2)}^{(3)} & 1 & \widetilde{K}_{(12)}^{(3)} & 0 & \widetilde{K}_{(23)}^{(3)} & 0 \\ 0 & 0 & \widetilde{K}_{(3)}^{(12)} & 1 & \widetilde{K}_{(13)}^{(12)} & \widetilde{K}_{(23)}^{(12)} & 0 \\ 0 & \widetilde{K}_{(2)}^{(13)} & 0 & \widetilde{K}_{(12)}^{(13)} & 1 & \widetilde{K}_{(23)}^{(13)} & 0 \\ \widetilde{K}_{(1)}^{(23)} & 0 & 0 & \widetilde{K}_{(12)}^{(23)} & \widetilde{K}_{(13)}^{(23)} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{b}^{(1)} \\ \tilde{b}^{(2)} \\ \tilde{b}^{(3)} \\ \tilde{b}^{(12)} \\ \tilde{b}^{(13)} \\ \tilde{b}^{(23)} \\ \tilde{b}^{(123)} \end{bmatrix}. \tag{5.48}$$

5.6 Estimateurs optimaux

Les estimateurs optimaux — ou à variance minimale — des paramètres de structure du modèle CCC généralisé diffèrent eux aussi en très peu de points de ceux du modèle standard. Nous ne nous étendrons donc pas sur les principes sous-jacents à ces estimateurs. Un simple exposé des formules pour les modèles à deux et trois facteurs devrait suffire. Une exception, cependant : l'estimateur optimal de la moyenne collective, qui diffère suffisamment de son équivalent dans le modèle CCC pour que l'on en discute plus longuement.

Sommairement, les modifications à apporter aux formules du chapitre 4 sont les suivantes :

- remplacer les facteurs de crédibilité z par les fonctions ζ ;
- pour les estimateurs des composants de variance, remplacer les covariances σ par les covariances $\tilde{\sigma}$ du modèle CCCG et les moyennes pondérées par les facteurs de crédibilité par les moyennes pondérées par les fonctions ζ ;

- pour l'estimateur de la moyenne collective seulement, modifier les dimensions des matrices afin de refléter le nombre accru de catégories pour chaque facteur de risque.

L'hypothèse d'excès nul intervient toujours de la même manière en s'appliquant à toutes les variables Ξ du modèle.

5.6.1 Modèle à deux facteurs

Nous présentons les estimateurs des paramètres de structure dans une suite de théorèmes. Rappelons que la matrice $\mathbf{U}$ est une matrice carrée ne contenant que des 1 et que $\mathbf{I}$ est la matrice identité. Nous introduisons également la matrice carrée $\mathbf{J}$, dont seul l'élément j_{11} vaut 1, les autres éléments de la matrice étant des zéros. Par exemple,

$$\mathbf{J}_{3\times 3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Le premier paramètre à estimer est la moyenne collective.

Théorème 5.6. *Dans un modèle CCCG à deux facteurs, l'estimateur sans biais à variance minimale de m dans la classe des fonctions linéaires des observations est*

$$\widehat{m}_* = \frac{\mathbf{B}'\mathbf{C}^{-1}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{X}, \tag{5.49}$$

où

$$\begin{aligned} \mathbf{X} &= [\mathbf{X}_i]_{(I+1)\times 1}, & \mathbf{X}_i &= [X_{ijw}]_{(J+1)\times 1}, \\ \mathbf{B} &= [\mathbf{B}_i]_{(I+1)\times 1}, & \mathbf{B}_i &= [1]_{(J+1)\times 1}, \\ \mathbf{C} &= [\mathbf{C}_{ik}]_{(I+1)\times(I+1)}, & \mathbf{C}_{ik} &= \overline{\delta}_{0i}\delta_{ik}b^{(1)}\mathbf{U} + b^{(2)}(\mathbf{I}-\mathbf{J}) + \delta_{ik}b^{(12)}\mathbf{Z}_i \end{aligned}$$

et $\mathbf{Z}_i$ est une matrice diagonale $(J+1)\times(J+1)$ formée des réciproques des fonctions $\zeta_{i0}^{(12)}, \dots, \zeta_{iJ}^{(12)}$:

$$\mathbf{Z}_i = \operatorname{diag}([(\zeta_{ij}^{(12)})^{-1}]_{(J+1)\times 1}).$$

L'ensemble complet des données disponibles est donc utilisé pour estimer la moyenne collective.

Preuve. Le début de la preuve est identique à celle du théorème 4.1. Les différences entre le résultat ci-dessus et celui du chapitre 4 proviennent de la formule de covariance entre deux moyennes pondérées individuelles. En effet, on a désormais

$$\begin{aligned}
\operatorname{Cov}(\widehat{m}_*, X_{klw}) &= \sum_{i=0}^{I} \sum_{j=0}^{J} \alpha_{ij} \tilde{\sigma}_{ijkl}^{(12)} \\
&= \sum_{i=0}^{I} \sum_{j=0}^{J} \alpha_{ij} \left[\overline{\delta}_{0i} \delta_{ik} b^{(1)} + \overline{\delta}_{0j} \delta_{jl} b^{(2)} + \delta_{ij,kl} \frac{b^{(12)}}{\zeta_{ij}^{(12)}} \right] \\
&= \overline{\delta}_{0k} \alpha_{k\Sigma} b^{(1)} + \overline{\delta}_{0l} \alpha_{\Sigma l} b^{(2)} + \alpha_{kl} \frac{b^{(12)}}{\zeta_{kl}^{(12)}}.
\end{aligned}$$

Les équations normales à résoudre pour $\alpha_{00}, \ldots, \alpha_{IJ}$ sont par conséquent

$$\begin{aligned}
\sum_{i=0}^{I} \sum_{j=0}^{J} \alpha_{ij} &= 1, \\
\overline{\delta}_{0i} \alpha_{i\Sigma} b^{(1)} + \overline{\delta}_{0j} \alpha_{\Sigma j} b^{(2)} + \alpha_{ij} \frac{b^{(12)}}{z_{ij}^{(12)}} &= \tilde{c}.
\end{aligned}$$

Ces nouvelles équations s'écrivent comme auparavant sous forme matricielle, à condition d'éliminer de la matrice $\mathbf{C}$ les facteurs $b^{(1)}$ du premier bloc de $(J+1) \times (J+1)$ éléments et les facteurs $b^{(2)}$ des lignes $1, J+2, 2J+3, \ldots, (I+1)(J+1)-J$. C'est là le rôle des éléments $\overline{\delta}_{0i}$ et $\mathbf{J}$ dans l'expression de $\mathbf{C}_{ik}$. □

Remarque. Dans un modèle CCCG avec $I = 1$ et $J = 2$, la matrice $\mathbf{C}$ ci-dessus a cette forme (les pointillés délimitent toujours les blocs $\mathbf{C}_{ik}$) :

$$\left[\begin{array}{ccc:ccc}
\frac{b^{(12)}}{\zeta_{00}^{(12)}} & 0 & 0 & 0 & 0 & 0 \\
0 & b^{(2)} + \frac{b^{(12)}}{\zeta_{01}^{(12)}} & 0 & 0 & b^{(2)} & 0 \\
0 & 0 & b^{(2)} + \frac{b^{(12)}}{\zeta_{02}^{(12)}} & 0 & 0 & b^{(2)} \\
\hdashline
0 & 0 & 0 & b^{(1)} + \frac{b^{(12)}}{\zeta_{10}^{(12)}} & b^{(1)} & b^{(1)} \\
0 & b^{(2)} & 0 & b^{(1)} & b + \frac{b^{(12)}}{\zeta_{11}^{(12)}} & b^{(1)} \\
0 & 0 & b^{(2)} & b^{(1)} & b^{(1)} & b + \frac{b^{(12)}}{\zeta_{12}^{(12)}}
\end{array}\right],$$

où $b \equiv b^{(1)} + b^{(2)}$. On pourra comparer cette matrice avec celle de la page 50. Il faut relever ici un point pratique important. Formellement, on peut imaginer un assuré qui ne serait affecté par aucun facteur de risque ; un tel assuré serait donc classé, dans un modèle à deux facteurs, dans la catégorie $(i = 0, j = 0)$. En pratique, toutefois, cette situation est peu vraisemblable, aussi tous les ratios et les poids naturels de la catégorie $(0, 0)$ sont-ils fort susceptibles d'être nuls. Par conséquent, $\zeta_{00}^{(12)} = 0$ et le premier élément de la matrice ci-dessus est non défini. En fait, la chose à faire dans une telle situation consiste simplement à éliminer la première ligne et la première colonne de la matrice $\mathbf{C}$ ainsi que le premier élément du vecteur $\mathbf{X}$ (qui est égal à zéro de toute façon).

Tel que mentionné précédemment, l'estimateur $\hat{s}^2$ du paramètre s^2 est optimal dans un certain ensemble de fonctions des observations. Les détails sont donnés dans le théorème qui suit, identique à quelques indices de sommation près au théorème 4.2. Une preuve n'est donc pas nécessaire.

Théorème 5.7. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCCG à deux facteurs, alors l'estimateur du paramètre de structure s^2 à variance minimale parmi toutes les combinaisons du type*

$$\sum_{ijturs} \alpha_{ijturs}(X_{ijt} - X_{iju})(X_{ijr} - X_{ijs})$$

avec moyenne s^2 est

$$\hat{s}_*^2 = \frac{1}{\sum_{ij}(T_{ij} - 1)} \sum_{i=0}^{I} \sum_{j=0}^{J} (T_{ij} - 1)\hat{s}_{ij}^2, \tag{5.50}$$

avec

$$\begin{aligned} \hat{s}_{ij}^2 &= \frac{1}{(T_{ij} - 1)} \sum_{t=1}^{T_{ij}} w_{ijt}(X_{ijt} - X_{ijw})^2 & (5.51) \\ &= \frac{1}{2(T_{ij} - 1)w_{ij\Sigma}} \sum_{t=1}^{T_{ij}} \sum_{u=1}^{T_{ij}} w_{ijt}w_{iju}(X_{ijt} - X_{iju})^2. & (5.52) \end{aligned}$$

Les trois théorèmes qui suivent exposent les (pseudo-)estimateurs optimaux des autres paramètres de structure d'un modèle à deux facteurs, soit les composants de variance $b^{(12)}$, $b^{(1)}$ et $b^{(2)}$.

Théorème 5.8. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCCG à deux facteurs, alors l'estimateur du paramètre de structure $b^{(12)}$ à variance minimale parmi les estimateurs de la forme*

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{kl>ij} \alpha_{ijkl}(X_{ijw} - X_{klw})^2 \tag{5.53}$$

avec espérance $b^{(12)}$ est

$$\tilde{b}_*^{(12)} = \mathbf{X}'\mathrm{diag}(\mathbf{A})\mathbf{X}, \tag{5.54}$$

où

$$\mathbf{A} = \frac{b^{(12)}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{C}^{-1}\mathbf{B}, \tag{5.55}$$

et

$$\begin{aligned}
\mathbf{B} &= [\tilde{\sigma}_{ijij}^{(12)} - 2\,\tilde{\sigma}_{ijkl}^{(12)} + \tilde{\sigma}_{klkl}^{(12)}]_{\frac{IJ(IJ-1)}{2}\times 1},\\
\mathbf{C} &= [(\tilde{\sigma}_{ijgh}^{(12)} - \tilde{\sigma}_{ijpq}^{(12)} - \tilde{\sigma}_{klgh}^{(12)} + \tilde{\sigma}_{klpq}^{(12)})^2]_{\frac{IJ(IJ-1)}{2}\times\frac{IJ(IJ-1)}{2}},\\
\mathbf{X} &= [X_{ijw} - X_{klw}]_{\frac{IJ(IJ-1)}{2}\times 1}.
\end{aligned}$$

Dans l'expression de $\mathbf{C}$*, les indices i, j, k et l identifient les lignes de la matrice et les indices g, h, p et q, les colonnes. La notation $kl > ij$ est définie ainsi, pour $k \geq i = 1, \ldots, I$ et $j = 1, \ldots, J$:*

$$kl > ij \Leftrightarrow \begin{cases} l = 1, \ldots, J & \text{si } k > i,\\ l > j & \text{si } k = i. \end{cases}$$

Théorème 5.9. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCCG à deux facteurs, alors l'estimateur du paramètre de structure $b^{(1)}$ à variance minimale parmi les estimateurs de la forme*

$$\sum_{i=1}^{I}\sum_{k>i}\alpha_{ik}(\widetilde{X}_{izw} - \widetilde{X}_{kzw})^2$$

avec espérance $b^{(1)}$ est

$$\tilde{b}_*^{(1)} = \mathbf{X}'\mathrm{diag}(\mathbf{A})\mathbf{X}, \tag{5.56}$$

où

$$\mathbf{A} = \frac{b^{(1)}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{C}^{-1}\mathbf{B}, \tag{5.57}$$

et

$$\begin{aligned}
\mathbf{B} &= [\tilde{\sigma}_{ii}^{(1)} - 2\,\tilde{\sigma}_{ik}^{(1)} + \tilde{\sigma}_{kk}^{(1)}]_{\frac{I(I-1)}{2}\times 1},\\
\mathbf{C} &= [(\tilde{\sigma}_{ig}^{(1)} - \tilde{\sigma}_{ip}^{(1)} - \tilde{\sigma}_{kg}^{(1)} + \tilde{\sigma}_{kp}^{(1)})^2]_{\frac{I(I-1)}{2}\times\frac{I(I-1)}{2}},\\
\mathbf{X} &= [\widetilde{X}_{izw} - \widetilde{X}_{kzw}]_{\frac{I(I-1)}{2}\times 1}.
\end{aligned}$$

Dans l'expression de $\mathbf{C}$*, les indices i et k identifient les lignes de la matrice et les indices g et p, les colonnes.*

Théorème 5.10. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCCG à deux facteurs, alors l'estimateur du paramètre de structure $b^{(2)}$ à variance minimale parmi les estimateurs de la forme*

$$\sum_{j=1}^{J} \sum_{l>j} \alpha_{jl} (\widetilde{X}_{zjw} - \widetilde{X}_{zlw})^2$$

avec espérance $b^{(2)}$ est

$$\tilde{b}_*^{(2)} = \mathbf{X}' \mathrm{diag}(\mathbf{A}) \mathbf{X}, \tag{5.58}$$

où

$$\mathbf{A} = \frac{b^{(2)}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}} \mathbf{C}^{-1}\mathbf{B}, \tag{5.59}$$

et

$$\begin{aligned}
\mathbf{B} &= [\tilde{\sigma}_{jj}^{(1)} - 2\,\tilde{\sigma}_{jl}^{(1)} + \tilde{\sigma}_{ll}^{(1)}]_{\frac{J(J-1)}{2} \times 1}, \\
\mathbf{C} &= [(\tilde{\sigma}_{jh}^{(1)} - \tilde{\sigma}_{jq}^{(1)} - \tilde{\sigma}_{lh}^{(1)} + \tilde{\sigma}_{lq}^{(1)})^2]_{\frac{J(J-1)}{2} \times \frac{J(J-1)}{2}}, \\
\mathbf{X} &= [\widetilde{X}_{zjw} - \widetilde{X}_{zlw}]_{\frac{J(J-1)}{2} \times 1}.
\end{aligned}$$

Dans l'expression de $\mathbf{C}$, les indices j et l identifient les lignes de la matrice et les indices h et q, les colonnes.

Pour la forme, rappelons que $\tilde{b}_*^{(1)}$, $\tilde{b}_*^{(2)}$ et $\tilde{b}_*^{(12)}$ sont des pseudo-estimateurs, tout comme, bien sûr, les estimateurs des composants de variance du modèle à trois facteurs, présentés à la sous-section suivante.

5.6.2 Modèle à trois facteurs

Pour le modèle CCCG à trois facteurs, nous ne présentons que l'estimateur de la moyenne collective et trois pseudo-estimateurs types des composants de variance.

Théorème 5.11. *Dans un modèle CCCG à trois facteurs, l'estimateur sans biais à variance minimale de m dans la classe des fonctions linéaires des observations est*

$$\widehat{m}_* = \frac{\mathbf{B}'\mathbf{C}^{-1}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}} \mathbf{X}, \tag{5.60}$$

où

$$\begin{aligned}
&\mathbf{X} = [\mathbf{X}_i]_{(I+1)\times 1}, &&\mathbf{X}_i = [\mathbf{X}_{ij}]_{(J+1)\times 1}, &&\mathbf{X}_{ij} = [X_{ijkw}]_{(K+1)\times 1},\\
&\mathbf{B} = [\mathbf{B}_i]_{(I+1)\times 1}, &&\mathbf{B}_i = [\mathbf{B}_{ij}]_{(J+1)\times 1}, &&\mathbf{B}_{ij} = [1]_{(K+1)\times 1},\\
&\mathbf{C} = [\mathbf{C}_{if}]_{(I+1)\times(I+1)}, &&\mathbf{C}_{if} = [\mathbf{C}_{ifjg}]_{(J+1)\times(J+1)},
\end{aligned}$$

avec

$$\begin{aligned}
\mathbf{C}_{ifjg} = {}& \overline{\delta}_{0i}\delta_{if}b^{(1)}\mathbf{U} + \overline{\delta}_{0j}\delta_{jg}b^{(2)}\mathbf{U} + b^{(3)}(\mathbf{I}-\mathbf{J}) + \overline{\delta}_{0i}\overline{\delta}_{0j}\delta_{ij,fg}b^{(12)}\mathbf{U}\\
&+ \overline{\delta}_{0i}\delta_{if}b^{(13)}(\mathbf{I}-\mathbf{J}) + \overline{\delta}_{0j}\delta_{jg}b^{(23)}(\mathbf{I}-\mathbf{J}) + \delta_{ij,fg}b^{(123)}\mathbf{Z}_{ij}
\end{aligned}$$

et $\mathbf{Z}_{ij}$ une matrice diagonale $(K+1)\times(K+1)$ formée des réciproques des fonctions $\zeta^{(123)}_{ij0},\ldots,\zeta^{(123)}_{ijK}$:

$$\mathbf{Z}_{ij} = \mathrm{diag}([(\zeta^{(123)}_{ijk})^{-1}]_{(K+1)\times 1}).$$

La remarque suivant le théorème 5.6 demeure valable ici.

Il n'apparaît pas essentiel d'énoncer un nouveau théorème statuant sur l'optimalité de l'estimateur $\hat{s}^2$ du modèle à trois facteurs déjà présenté à la section 5.4. Mentionnons donc simplement que, sous les hypothèses usuelles d'excès nul, cet estimateur est à variance minimale dans la classe des estimateurs sans biais de la forme

$$\sum_{ijkturs} \alpha_{ijkturs}(X_{ijkt} - X_{ijku})(X_{ijkr} - X_{ijks}).$$

Les pseudo-estimateurs des composants de variance $b^{(123)}$, $b^{(12)}$ et $b^{(1)}$ sont les suivants.

Théorème 5.12. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCCG à trois facteurs, alors l'estimateur du paramètre de structure $b^{(123)}$ à variance minimale parmi les estimateurs de la forme*

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}\sum_{fgh>ijk} \alpha_{ijkfgh}(X_{ijkw} - X_{fghw})^2 \tag{5.61}$$

avec espérance $b^{(123)}$ est

$$\tilde{b}^{(123)}_{*} = \mathbf{X}'\mathrm{diag}(\mathbf{A})\mathbf{X}, \tag{5.62}$$

où

$$\mathbf{A} = \frac{b^{(123)}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{C}^{-1}\mathbf{B}, \tag{5.63}$$

et

$$\begin{aligned}
\mathbf{B} &= [\tilde{\sigma}^{(123)}_{ijkijk} - 2\,\tilde{\sigma}^{(123)}_{ijkfgh} + \tilde{\sigma}^{(123)}_{fghfgh}]_{\frac{IJK(IJK-1)}{2}\times 1},\\
\mathbf{C} &= [(\tilde{\sigma}^{(123)}_{ijkpqr} - \tilde{\sigma}^{(123)}_{ijkstu} - \tilde{\sigma}^{(123)}_{fghpqr} + \tilde{\sigma}^{(123)}_{fghstu})^2]_{\frac{IJK(IJK-1)}{2}\times\frac{IJK(IJK-1)}{2}},\\
\mathbf{X} &= [X_{ijkw} - X_{fghw}]_{\frac{IJK(IJK-1)}{2}\times 1}.
\end{aligned}$$

Dans l'expression de $\mathbf{C}$*, les indices* i, j, k, f, g *et* h *identifient les lignes de la matrice et les indices* p, q, r, s, t *et* u*, les colonnes. La notation* $fgh > ijk$ *est définie ainsi, pour* $f \geq i = 1, \ldots, I$*,* $j = 1, \ldots, J$ *et* $k = 1, \ldots, K$ *:*

$$fgh > ijk \Leftrightarrow \begin{cases} g = 1, \ldots, J;\ h = 1, \ldots, K & \text{si } f > i,\\ h = 1, \ldots, K & \text{si } f = i \text{ et } g > j,\\ h > k & \text{si } f = i \text{ et } g = j. \end{cases}$$

Théorème 5.13. *Si toutes les variables* Ξ *ont un coefficient d'excès nul dans un modèle CCCG à trois facteurs, alors l'estimateur du paramètre de structure* $b^{(12)}$ *à variance minimale parmi les estimateurs de la forme*

$$\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{fg>ij} \alpha_{ijfg}(\widetilde{X}_{ijzw} - \widetilde{X}_{fgzw})^2 \tag{5.64}$$

avec espérance $b^{(12)}$ *est*

$$\tilde{b}^{(12)}_* = \mathbf{X}'\mathrm{diag}(\mathbf{A})\mathbf{X}, \tag{5.65}$$

où

$$\mathbf{A} = \frac{b^{(12)}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{C}^{-1}\mathbf{B}, \tag{5.66}$$

et

$$\begin{aligned}
\mathbf{B} &= [\tilde{\sigma}^{(12)}_{ijij} - 2\,\tilde{\sigma}^{(12)}_{ijfg} + \tilde{\sigma}^{(12)}_{fgfg}]_{\frac{IJ(IJ-1)}{2}\times 1},\\
\mathbf{C} &= [(\tilde{\sigma}^{(12)}_{ijpq} - \tilde{\sigma}^{(12)}_{ijst} - \tilde{\sigma}^{(12)}_{fgpq} + \tilde{\sigma}^{(12)}_{fgst})^2]_{\frac{IJ(IJ-1)}{2}\times\frac{IJ(IJ-1)}{2}},\\
\mathbf{X} &= [\widetilde{X}_{ijzw} - \widetilde{X}_{fgzw}]_{\frac{IJ(IJ-1)}{2}\times 1}.
\end{aligned}$$

Dans l'expression de $\mathbf{C}$*, les indices* i, j, f *et* g *identifient les lignes de la matrice et les indices* p, q, s *et* t*, les colonnes. La notation* $fg > ij$ *est définie comme au théorème 5.8.*

Théorème 5.14. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCCG à trois facteurs, alors l'estimateur du paramètre de structure $b^{(1)}$ à variance minimale parmi les estimateurs de la forme*

$$\sum_{i=1}^{I} \sum_{f>i} \alpha_{if} (\widetilde{X}_{izw} - \widetilde{X}_{fzw})^2$$

avec espérance $b^{(1)}$ est

$$\tilde{b}_*^{(1)} = \mathbf{X}' \mathrm{diag}(\mathbf{A}) \mathbf{X}, \tag{5.67}$$

où

$$\mathbf{A} = \frac{b^{(1)}}{\mathbf{B}' \mathbf{C}^{-1} \mathbf{B}} \mathbf{C}^{-1} \mathbf{B}, \tag{5.68}$$

et

$$\begin{aligned}
\mathbf{B} &= [\tilde{\sigma}_{ii}^{(1)} - 2\,\tilde{\sigma}_{if}^{(1)} + \tilde{\sigma}_{ff}^{(1)}]_{\frac{I(I-1)}{2} \times 1}, \\
\mathbf{C} &= [(\tilde{\sigma}_{ip}^{(1)} - \tilde{\sigma}_{is}^{(1)} - \tilde{\sigma}_{fp}^{(1)} + \tilde{\sigma}_{fs}^{(1)})^2]_{\frac{I(I-1)}{2} \times \frac{I(I-1)}{2}}, \\
\mathbf{X} &= [\widetilde{X}_{izw} - \widetilde{X}_{fzw}]_{\frac{I(I-1)}{2} \times 1}.
\end{aligned}$$

Dans l'expression de $\mathbf{C}$, les indices i et f identifient les lignes de la matrice et les indices p et s, les colonnes.

Les pseudo-estimateurs des autres composants de variance sont, comme toujours, similaires.

Toute la théorie de base des modèles CCC et CCCG a maintenant été couverte : hypothèses, estimateurs de crédibilité, estimateurs des paramètres de structure. La dernière étape, avant de passer à l'étude numérique de la performance des estimateurs, consiste à développer les formules d'un modèle CCCG avec un nombre arbitraire P de facteurs risque. C'est là l'objet du chapitre suivant.

Chapitre 6

Formules générales

Jusqu'à maintenant, seuls les modèles de crédibilité à classification croisée (généralisés ou non) à deux et trois facteurs ont été étudiés, ceci afin de ne considérer, dans un premier temps, que les cas les plus simples. La notation en est, par le fait même, maintenue à un niveau de complexité acceptable. Il est toutefois bien sûr possible d'imaginer des applications des modèles CCC ou CCCG comportant un nombre accru de facteurs de risque. Des formules générales pour un nombre arbitraire P de facteurs de risque s'avèrent donc utiles, voire essentielles. C'est là le propos de ce chapitre.

Afin de traiter le cas le plus général possible, les formules du modèle CCCG à P facteurs seront présentées ici. On sait qu'il suffit de poser égaux à zéro tous les poids et les données des catégories 0 de chaque facteur de risque pour obtenir les formules du modèle CCC standard. Ce chapitre ne comporte que deux sections. Dans la première, on présente les hypothèses et l'écriture générales du modèle CCCG «complet» — auquel il est fait référence aux chapitres 3 et 5 — ainsi que l'estimateur de crédibilité. Dans la seconde section, on s'attarde aux relations de covariance et à l'estimation des paramètres de structure. Deux séries d'estimateurs sont proposées sous forme générale, soit les estimateurs de Dannenburg et les pseudo-estimateurs optimaux.

6.1 Hypothèses et estimateur de crédibilité

Soit P le nombre de facteurs de risque dans un modèle CCCG et $J_p + 1$ le nombre de catégories pour le p^{e} facteur de risque ($p \in \{1, \dots, P\}$). Le niveau de risque de la catégorie $j_p \in \{0, 1, \dots, J_p\}$ du p^{e} facteur est caractérisé par la variable aléatoire $\Theta_{j_p}^{(p)}$. La variable aléatoire $X_{j_1 \dots j_P t}$ représente l'expérience observée dans la cellule $(j_1, \dots, j_P)$ au temps t ($t = 1, 2, \dots, T_{j_1, \dots j_P} > 1$)

et, par conséquent, l'ensemble des données $\underline{X}$ forme un vecteur à $P+1$ dimensions. Enfin, à chaque donnée $X_{j_1\dots j_P t}$ est rattaché un poids $w_{j_1\dots j_P t}$. Les hypothèses du modèle CCC généralisé sont les suivantes :

(CCCG1) Pour p fixe, les variables $\Theta^{(p)}_{j_p}$ sont i.i.d.. Les variables Θ sont mutuellement indépendantes.

(CCCG2) Pour $q, i_1, \dots, i_q \in \{1, \dots, P\}$ et $j_{i_r} \in \{0, \dots, J_{i_r}\}$, $r = 1, \dots, q$, il existe des fonctions $\mu_\cdot(\cdot)$ telles que

$$\mathrm{E}\Big[X_{j_1\dots j_P t}|\Theta^{(i_1)}_{j_{i_1}}, \dots, \Theta^{(i_q)}_{j_{i_q}}\Big] = \mu_{i_1\dots i_q}(\Theta^{(i_1)}_{j_{i_1}}, \dots, \Theta^{(i_q)}_{j_{i_q}}).$$

(CCCG3) On a

$$\mathrm{E}\Big[\mathrm{Cov}(X_{j_1\dots j_P t}, X_{l_1\dots l_P u}|\underline{\Theta}^{(1)}, \dots, \underline{\Theta}^{(P)})\Big] = \frac{\delta_{j_1\dots j_P t, l_1\dots l_P u}}{w_{j_1\dots j_P t}}\, s^2.$$

Le théorème ci-dessous présente la version dite « complète » du modèle CCCG reposant sur les hypothèses précédentes. Il s'agit d'une version d'un théorème de Dannenburg (1995) adaptée pour tenir compte des catégories 0 du modèle généralisé. Ce théorème ne sera donc pas prouvé. Le théorème s'en tient aux p premiers facteurs de risque, mais il est clair que le résultat demeure valable pour toute autre combinaison de ces facteurs.

Théorème 6.1. *Pour tout $p \in \{1, \dots, P\}$, l'espérance conditionnelle de l'assuré $X_{j_1\dots j_P t}$ sachant les p premiers facteurs de risque $\Theta^{(1)}_{j_1}, \dots, \Theta^{(p)}_{j_p}$ est donnée par*

$$\begin{aligned}\mu_{1\dots p}(\Theta^{(1)}_{j_1}, \dots, \Theta^{(p)}_{j_p}) &= m \\ &+ \sum_{q=1}^{p}\left[\sum_{i_1=1}^{p-q+1} \cdots \sum_{i_q=i_{q-1}+1}^{p}\left(\prod_{k=i_1}^{i_q}\overline{\delta}_{0j_k}\right)\nu_{i_1\dots i_q}(\Theta^{(i_1)}_{j_{i_1}}, \dots, \Theta^{(i_q)}_{j_{i_q}})\right], \quad (6.1)\end{aligned}$$

où $\overline{\delta}_{0j} = 1 - \delta_{0j}$ et avec les termes $\nu_{i_1\dots i_q}(\Theta^{(i_1)}_{j_{i_1}}, \dots, \Theta^{(i_q)}_{j_{i_q}})$, $i_1 < i_2 < \dots < i_q$, ainsi définis par récurrence :

$$\begin{aligned}\nu_{i_1\dots i_q}(\Theta^{(i_1)}_{j_{i_1}}, \dots, \Theta^{(i_q)}_{j_{i_q}}) &= \mu_{i_1\dots i_q}(\Theta^{(i_1)}_{j_{i_1}}, \dots, \Theta^{(i_q)}_{j_{i_q}}) - m \\ &- \sum_{r=1}^{q-1}\left[\sum_{k_1=1}^{q-r+1} \cdots \sum_{k_r=k_{r-1}+1}^{q} \nu_{i_{k_1}\dots i_{k_r}}(\Theta^{(i_{k_1})}_{j_{i_{k_1}}}, \dots, \Theta^{(i_{k_r})}_{j_{i_{k_r}}})\right], \quad (6.2)\end{aligned}$$

si et seulement si $(j_{i_1}, \dots, j_{i_q}) \neq (0, \dots, 0)$ et avec comme premier terme $\nu_{i_1}(\Theta^{(i_1)}_{j_{i_1}}) = \mu_{i_1}(\Theta^{(i_1)}_{j_{i_1}}) - m$, $j_{i_1} \neq 0$. De même, on peut représenter chaque risque $X_{j_1\dots j_P t}$ sous la forme

$$X_{j_1\dots j_P t} = \nu_{1\dots P t}(\Theta^{(1)}_{j_1}, \dots, \Theta^{(P)}_{j_P}, X_{j_1\dots j_P t}) + \mu_{1\dots P}(\Theta^{(1)}_{j_1}, \dots, \Theta^{(P)}_{j_P}), \quad (6.3)$$

avec

$$\nu_{1\dots Pt}(\Theta_{j_1}^{(1)},\dots,\Theta_{j_P}^{(P)},X_{j_1\dots j_Pt}) = X_{j_1\dots j_Pt} - \mu_{1\dots P}(\Theta_{j_1}^{(1)},\dots,\Theta_{j_P}^{(P)}). \tag{6.4}$$

Or, $\nu_{1\dots Pt}(\Theta_{j_1}^{(1)},\dots,\Theta_{j_P}^{(P)},X_{j_1\dots j_Pt})$ *et tous les termes* $\nu_{i_1\dots i_q}(\Theta_{j_{i_1}}^{(i_1)},\dots,\Theta_{j_{i_q}}^{(i_q)})$ *sont non corrélés et ont une espérance nulle.*

Du théorème précédent, on déduit que le paramètre m constitue la moyenne collective du portefeuille, c'est-à-dire que

$$m = \mathrm{E}[X_{j_1\dots j_Pt}] = \mathrm{E}\Big[\mu_{1\dots P}(\Theta_{j_1}^{(1)},\dots,\Theta_{j_P}^{(P)})\Big]. \tag{6.5}$$

On définit par la suite

$$b^{(i_1\dots i_q)} = \mathrm{Var}\Big[\nu_{i_1\dots i_q}(\Theta_{j_{i_1}}^{(i_1)},\dots,\Theta_{j_{i_q}}^{(i_q)})\Big], \quad (j_{i_1},\dots,j_{i_q}) \neq (0,\dots,0), \tag{6.6}$$

ce qui, avec le paramètre s^2 et les poids $w_{j_1\dots j_Pt}$, détermine entièrement la structure de covariance du modèle.

Dorénavant, on peut se concentrer sur la version simplifiée du modèle, qui comporte tout de même les mêmes premier et second moments. Pour ce faire, on remplace les termes $\nu_{i_1\dots i_q}(\Theta_{j_{i_1}}^{(i_1)},\dots,\Theta_{j_{i_q}}^{(i_q)})$ par des variables aléatoires indépendantes $\Xi_{j_{i_1}\dots j_{i_q}}^{(i_1\dots i_q)}$ d'espérance nulle et de variance égale à $b^{(i_1\dots i_q)}$. Ce faisant, tout assuré $X_{j_1\dots j_Pt}$ est représenté ainsi :

$$\begin{aligned} X_{j_1\dots j_Pt} = m + \sum_{q=1}^{P}\left[\sum_{i_1=1}^{P-q+1}\cdots\sum_{i_q=i_{q-1}+1}^{P}\left(\prod_{k=i_1}^{i_q}\overline{\delta}_{0j_k}\right)\Xi_{j_{i_1}\dots j_{i_q}}^{(i_1\dots i_q)}\right] \\ + \Xi_{j_1\dots j_Pt}^{(1\dots P,P+1)}. \end{aligned} \tag{6.7}$$

La variance du dernier terme est égale à $s^2/w_{j_1\dots j_Pt}$.

L'essentiel de la notation qui sera rencontrée dans ce chapitre a dès lors été utilisée. Quelques conventions d'écriture supplémentaires seront néanmoins nécessaires pour la suite.

Conventions. Les conventions ci-dessous se rapportent plus spécifiquement aux formules générales de ce chapitre et à la notation y ayant cours.

(C5) Sans perte de généralité, on peut supposer que $i_1 < \cdots < i_q$. De plus, l'indice i_{q+1} est toujours le plus petit indice différent de $i_1,\dots,i_q$.

(C6) Dans un modèle à classification croisée à P facteurs, il est clair que $(i_1\dots i_P) \equiv (1\dots P)$.

(C7) On pose

$$b^{(i_1\dots i_P, i_{P+1})} \equiv b^{(1\dots P, P+1)} \equiv s^2$$

ainsi que

$$\tilde{z}^{(i_1\dots i_P, i_{P+1})}_{j_{i_1}\dots j_{i_P}\Sigma} \equiv \tilde{z}^{(1\dots P, P+1)}_{j_1\dots j_P\Sigma} \equiv w_{j_1\dots j_P\Sigma}$$

et

$$\zeta^{(1\dots P, P+1)}_{j_1\dots j_P\Sigma} \equiv w_{j_1\dots j_P\Sigma}.$$

Les termes $\tilde{z}$ et ζ sont définis dans le théorème 6.2, plus bas.

(C8) On définit les ensembles de nombres suivants :

$$\begin{aligned}\mathfrak{I}(q,r) &= \{i_k, 1 \leq k \leq q | i_k \in \{1,\dots,r\}\},\\ \bar{\mathfrak{I}}(q,r) &= \{i_k, 1 \leq k \leq q | i_k \notin \{1,\dots,r\}\}.\end{aligned}$$

Pour une valeur de q donnée, $\mathfrak{I}(q,r)$ représente donc l'ensemble des indices i_k dont la valeur se trouve entre 1 et r au moment où l'on fait référence au dit ensemble. Quant à $\bar{\mathfrak{I}}(q,r)$, il s'agit simplement du complément du précédent. Si $q = 0$, les deux ensembles sont vides.

(C9) On supposera qu'à une cellule ne comportant aucune donnée sera attribuée un poids w nul.

(C10) On définit $N^{(i_1\dots i_q)}$ comme le nombre total de données dans les catégories j_{i_1} à j_{i_q}. Étant donnée la convention précédente, $N^{(i_1\dots i_q)}$ se définit mathématiquement par

$$N^{(i_1\dots i_q)} = \sum_{j_{i_1}} \cdots \sum_{j_{i_q}} I_{\{w_{j_{i_1}\dots j_{i_q}\Sigma} > 0\}},$$

où $I_{\{A\}}$ est une fonction indicatrice valant 1 si A est vraie et zéro sinon.

(C11) Soit $\mathcal{A}$ un ensemble de nombres. Alors l'écriture $j_{\mathcal{A}}$ signifie que j est indicé sur toutes les valeurs de $\mathcal{A}$. Par exemple, si $\mathcal{A} = \{a_1,\dots,a_n\}$, alors $j_{\mathcal{A}} = \{j_{a_1},\dots,j_{a_n}\}$.

(C12) La notation précédente peut également être employée pour écrire succinctement une somme multiple :

$$\sum_{j_{\mathcal{A}}=0}^{J_{\mathcal{A}}} x_{j_1\dots j_q j_{\mathcal{A}}} \equiv \sum_{j_{a_1}=0}^{J_{a_1}} \cdots \sum_{j_{a_n}=0}^{J_{a_n}} x_{j_1\dots j_q j_{a_1}\dots j_{a_n}}.$$

Si $\mathcal{A} = \emptyset$, on ne somme pas la quantité $x_{j_1\dots j_q j_{\mathcal{A}}}$, qui se ramène alors simplement à $x_{j_1\dots j_q}$.

L'estimateur de crédibilité de $X_{j_1\dots,J_P,T_{j_1\dots,j_P}+1}$ d'un modèle CCCG à P facteurs est défini dans le théorème suivant. Ce théorème est présenté sans preuve, celle-ci étant de la même eau que la preuve du théorème 5.1.

Théorème 6.2. *Soit $X_{j_1\dots j_Pt}$ le ratio d'expérience d'un assuré tel que défini en (6.7). Le meilleur estimateur — au sens des moindres carrés — de $X_{j_1\dots,J_P,T_{j_1\dots,j_P}+1}$ est alors*

$$\begin{aligned}\widehat{X}_{j_1\dots j_P,T_{j_1\dots j_P}+1} &= m + \tilde{z}^{(1\dots P)}_{j_1\dots j_P}(X_{j_1\dots j_Pw} - m)\\ &\quad + (1-\tilde{z}^{(1\dots P)}_{j_1\dots j_P})\sum_{q=1}^{P-1}\left[\sum_{i_1=1}^{P-q+1}\cdots\sum_{i_q=i_{q-1}+1}^{P}\left(\prod_{k=i_1}^{i_q}\overline{\delta}_{0j_k}\right)\widehat{\Xi}^{(i_1\dots i_q)}_{j_{i_1}\dots j_{i_q}}\right],\end{aligned}\tag{6.8}$$

où, pour $r = 1,\dots,P$,

$$\begin{aligned}\widehat{\Xi}^{(i_1\dots i_r)}_{j_{i_1}\dots j_{i_r}} &= \tilde{z}^{(i_1\dots i_r)}_{j_{i_1}\dots j_{i_r}}(\widetilde{X}_{j_{i_1}\dots j_{i_r}zw} - m)\\ &\quad - \tilde{z}^{(i_1\dots i_r)}_{j_{i_1}\dots j_{i_r}}\sum_{j_{i_{r+1}}=0}^{J_{i_{r+1}}}\cdots\sum_{j_{i_P}=0}^{J_{i_P}}\left(\prod_{k=r+1}^{P}\frac{\zeta^{(i_1\dots i_k)}_{j_{i_1}\dots j_{i_k}}}{\zeta^{(i_1\dots i_k)}_{j_{i_1}\dots j_{i_{k-1}}\Sigma}}\right)\\ &\quad\times\sum_{q=1}^{P-1}\left(\sum_{l_1=1}^{P-q+1}\cdots\sum_{l_q=l_{q-1}+1}^{P}h(l_q,i_r)\,\widehat{\Xi}^{(l_1\dots l_q)}_{j_{l_1}\dots j_{l_q}}\right),\end{aligned}\tag{6.9}$$

avec

$$h(l_q,i_r) = (1-\delta_{l_1\dots l_q,i_1\dots i_r})(1-\delta_{l_1\dots l_q,i_1\dots i_{r+1}})\cdots(1-\delta_{l_1\dots l_q,i_1\dots i_{P-1}}),\tag{6.10}$$

$$\tilde{z}^{(i_1\dots i_r)}_{j_{i_1}\dots j_{i_r}} = \left(\prod_{k=i_1}^{i_r}\overline{\delta}_{0j_k}\right)\frac{\zeta^{(i_1\dots i_{r+1})}_{j_{i_1}\dots j_{i_r}\Sigma}\,b^{(i_1\dots i_r)}}{\zeta^{(i_1\dots i_{r+1})}_{j_{i_1}\dots j_{i_r}\Sigma}\,b^{(i_1\dots i_r)} + b^{(i_1\dots i_{r+1})}},\tag{6.11}$$

$$\begin{aligned}\zeta^{(i_1\dots i_r)}_{j_{i_1}\dots j_{i_r}} &= \left(1-\prod_{k=i_1}^{i_r}\overline{\delta}_{0j_k}\right)\frac{b^{(i_1\dots i_r)}}{s^2}w_{j_{i_1}\dots j_{i_r}\Sigma} + \tilde{z}^{(i_1\dots i_r)}_{j_{i_1}\dots j_{i_r}} &(6.12)\\ &= \frac{\zeta^{(i_1\dots i_{r+1})}_{j_{i_1}\dots j_{i_r}\Sigma}\,b^{(i_1\dots i_r)}}{\left(\prod_{k=i_1}^{i_r}\overline{\delta}_{0j_k}\right)\zeta^{(i_1\dots i_{r+1})}_{j_{i_1}\dots j_{i_r}\Sigma}\,b^{(i_1\dots i_r)} + b^{(i_1\dots i_{r+1})}} &(6.13)\end{aligned}$$

et, enfin,

$$\widetilde{X}_{j_{i_1}\dots j_{i_r}zw} = \sum_{j_{i_{r+1}}=0}^{J_{i_{r+1}}}\frac{\zeta^{(i_1\dots i_{r+1})}_{j_{i_1}\dots j_{i_{r+1}}}}{\zeta^{(i_1\dots i_{r+1})}_{j_{i_1}\dots j_{i_r}\Sigma}}\widetilde{X}_{j_1\dots j_{r+1}zw}.\tag{6.14}$$

On reconnaîtra ci-dessus la forme maintenant familière de l'estimateur de crédibilité. Le système d'équations à résoudre pour obtenir les estimateurs de crédibilité des variables Ξ est donné par (6.9). On remarquera toutefois que la variable $\Xi_{j_1 \dots j_P}^{(1 \dots P)}$ fait partie de ce système, alors qu'il n'est pas nécessaire de calculer sa valeur pour obtenir l'estimateur de crédibilité (6.8). De plus, le rôle, a priori obscur, de la fonction $h(l_q, i_r)$ est d'exclure de la somme à la droite de (6.9) la variable $\widehat{\Xi}$ se trouvant déjà à gauche de l'équation, ainsi que toutes les variables des niveaux d'interaction dits « immédiatement supérieurs ». On notera, enfin, la définition alternative des fonctions ζ donnée en (6.13), qui peut s'avérer utile à des fins de calcul.

L'interprétation des résultats de ce théorème ayant déjà fait l'objet des chapitres 3 et 5, on peut sans plus tarder passer à l'estimation des paramètres de structure.

6.2 Estimation des paramètres de structure

Les paramètres de structure du modèle CCCG à P facteurs de risque sont la moyenne collective m, la variance temporelle s^2 et les $\sum_{n=1}^{P} \binom{P}{n} = 2^P - 1$ composants de variance $b^{(i_1 \dots, i_q)}$, $q, i_1, \dots, i_q \in \{1, \dots P\}$. Nous présentons ici deux séries d'estimateurs de ces paramètres sous forme générale : les estimateurs de Dannenburg et les pseudo-estimateurs optimaux. L'absence de biais des estimateurs de Dannenburg (généralisés) sera également démontrée. Les pseudo-estimateurs ad hoc sont écartés à cause de la trop grande complexité (ou lourdeur, c'est selon) des formules. Cet argument défavorable aux estimateurs ad hoc avait d'ailleurs déjà été avancé à la fin du chapitre 3.

Comme c'en est devenu l'habitude, l'exposé des estimateurs est précédé des relations de covariance pertinentes. Dans le cas présent, deux types de relations de covariance se révèlent fondamentales, soit une pour chaque série d'estimateurs.

6.2.1 Relations de covariance

Les deux théorèmes suivants se concentrent sur les $r \geq 1$ premiers indices de la variable X afin de ne pas davantage surcharger la notation. Les résultats pour les autres combinaisons des indices sont similaires.

Les premières relations de covariance seront nécessaires pour démontrer l'absence de biais des estimateurs de Dannenburg.

Théorème 6.3. *Soient* $j_1, \dots, j_P \in \{0, \dots, P\}$ *et* $l_1, \dots, l_P \in \{0, \dots, P\}$

deux ensembles d'indices pouvant être différents. Alors, pour $r = 0, 1, \ldots, P$,

$$\begin{aligned}\mathrm{Cov}(X_{j_1\ldots j_P w}, X_{l_1\ldots l_r w}) = \sum_{q=1}^{P} \Bigg[\sum_{i_1=1}^{P-q+1} \cdots \sum_{i_q=i_{q-1}+1}^{P} \left(\prod_{k=i_1}^{i_r} \overline{\delta}_{0 j_k} \right) \\ \times \, \delta_{j_{\mathcal{I}(q,r)}, l_{\mathcal{I}(q,r)}} \frac{w_{l_1\ldots l_r j_{\bar{\mathcal{I}}(q,r)} \Sigma}}{w_{l_1\ldots l_r \Sigma}} b^{(i_1\ldots i_q)} \Bigg] \\ + \, \delta_{j_1\ldots j_r, l_1\ldots l_r} \frac{s^2}{w_{j_1\ldots j_r \Sigma}}.\end{aligned} \tag{6.15}$$

Preuve par induction. De par la construction du modèle, on sait déjà que

$$\begin{aligned}\mathrm{Cov}(X_{j_1\ldots j_P t}, X_{l_1\ldots l_P u}) = \sum_{q=1}^{P} \Bigg[\sum_{i_1=1}^{P-q+1} \cdots \sum_{i_q=i_{q-1}+1}^{P} \left(\prod_{k=i_1}^{i_r} \overline{\delta}_{0 j_k} \right) \\ \times \, \delta_{j_{i_1}\ldots j_{i_q}, l_{i_1}\ldots l_{i_q}} \, b^{(i_1\ldots i_q)} \Bigg] \\ + \, \delta_{j_1\ldots j_P t, l_1\ldots l_P u} \frac{s^2}{w_{j_1\ldots j_P t}},\end{aligned}$$

donc

$$\begin{aligned}\mathrm{Cov}(X_{j_1\ldots j_P w}, X_{l_1\ldots l_P w}) &= \sum_{t=1}^{T_{j_1\ldots j_P}} \sum_{u=1}^{T_{l_1\ldots l_P}} \frac{w_{j_1\ldots j_P t}}{w_{j_1\ldots j_P \Sigma}} \frac{w_{l_1\ldots l_P u}}{w_{l_1\ldots l_P \Sigma}} \mathrm{Cov}(X_{j_1\ldots j_P t}, X_{l_1\ldots l_P u}) \\ &= \sum_{q=1}^{P} \Bigg[\sum_{i_1=1}^{P-q+1} \cdots \sum_{i_q=i_{q-1}+1}^{P} \left(\prod_{k=i_1}^{i_r} \overline{\delta}_{0 j_k} \right) \\ &\quad \times \, \delta_{j_{i_1}\ldots j_{i_q}, l_{i_1}\ldots l_{i_q}} \, b^{(i_1\ldots i_q)} \Bigg] \\ &\quad + \, \delta_{j_1\ldots j_P, l_1\ldots l_P} \frac{s^2}{w_{j_1\ldots j_P \Sigma}}.\end{aligned}$$

Or, comme $\mathcal{I}(q, P) = \{i_1, \ldots, i_q\}$, la dernière égalité prouve que (6.15) est vraie pour $r = P$. En supposant la formule correcte pour $r = h + 1$, on prouve maintenant qu'elle est vraie pour $r = h$:

$$\begin{aligned}
\mathrm{Cov}(X_{j_1\ldots j_P w}, X_{l_1\ldots l_h w}) &= \sum_{l_{h+1}=0}^{J_{h+1}} \frac{w_{l_1\ldots l_{h+1}\Sigma}}{w_{l_1\ldots l_h\Sigma}} \mathrm{Cov}(X_{j_1\ldots j_P w}, X_{l_1\ldots l_{h+1} w}) \\
&= \sum_{q=1}^{P} \left[\sum_{i_1=1}^{P-q+1} \cdots \sum_{i_q=i_{q-1}+1}^{P} \sum_{l_{h+1}=0}^{J_{h+1}} \left(\prod_{k=i_1}^{i_r} \overline{\delta}_{0j_k} \right) \right. \\
&\qquad \left. \times\, \delta_{j_{\mathfrak{I}(q,h+1)}, l_{\mathfrak{I}(q,h+1)}} \frac{w_{l_1\ldots l_{h+1} j_{\bar{\mathfrak{I}}(q,h+1)}\Sigma}}{w_{l_1\ldots l_h\Sigma}}\, b^{(i_1\ldots i_q)} \right] \\
&\qquad + \sum_{l_{h+1}=0}^{J_{h+1}} \delta_{j_1\ldots j_{h+1}, l_1\ldots l_{h+1}} \frac{s^2}{w_{j_1\ldots j_h\Sigma}} \\
&= \sum_{q=1}^{P} \left[\sum_{i_1=1}^{P-q+1} \cdots \sum_{i_q=i_{q-1}+1}^{P} \left(\prod_{k=i_1}^{i_r} \overline{\delta}_{0j_k} \right) \right. \\
&\qquad \left. \times\, \delta_{j_{\mathfrak{I}(q,h)}, l_{\mathfrak{I}(q,h)}} \frac{w_{l_1\ldots l_h j_{\bar{\mathfrak{I}}(q,h)}\Sigma}}{w_{l_1\ldots l_h\Sigma}}\, b^{(i_1\ldots i_q)} \right] \\
&\qquad + \delta_{j_1\ldots j_h, l_1\ldots l_h} \frac{s^2}{w_{j_1\ldots j_h\Sigma}}.
\end{aligned}$$

En effet,

$$\sum_{l_{h+1}=0}^{J_{h+1}} \delta_{j_{\mathfrak{I}(q,h+1)}, l_{\mathfrak{I}(q,h+1)}} w_{l_1\ldots l_{h+1} j_{\bar{\mathfrak{I}}(q,h+1)}\Sigma} = \delta_{j_{\mathfrak{I}(q,h)}, l_{\mathfrak{I}(q,h)}} w_{l_1\ldots l_h j_{\bar{\mathfrak{I}}(q,h)}\Sigma},$$

ce qui prouve le résultat. □

Par exemple, dans un modèle CCCG à deux facteurs, on a

$$\begin{aligned}
\mathrm{Cov}(X_{j_1 j_2 w}, X_{l_1 w}) = \overline{\delta}_{0j_1} \delta_{j_1 l_1} b^{(1)} + \overline{\delta}_{0j_2} \frac{w_{l_1 j_2\Sigma}}{w_{l_1\Sigma}}\, b^{(2)} \\
+ \overline{\delta}_{0j_1} \overline{\delta}_{0j_2} \delta_{j_1 l_1} \frac{w_{l_1 j_2\Sigma}}{w_{l_1\Sigma}}\, b^{(12)} + \frac{s^2}{w_{l_1\Sigma}}
\end{aligned}$$

et

$$\begin{aligned}
\mathrm{Cov}(X_{j_1 j_2 w}, X_w) = \overline{\delta}_{0j_1} \frac{w_{j_1\Sigma}}{w_\Sigma}\, b^{(1)} + \overline{\delta}_{0j_2} \frac{w_{j_2\Sigma}}{w_\Sigma}\, b^{(2)} \\
+ \overline{\delta}_{0j_1} \overline{\delta}_{0j_2} \frac{w_{j_1 j_2\Sigma}}{w_\Sigma}\, b^{(12)} + \frac{s^2}{w_\Sigma}.
\end{aligned}$$

Les secondes relations de covariance retenues pour ce chapitre sont les versions générales des relations des théorèmes 5.2 et 5.3. Le théorème suivant ne sera donc pas prouvé, puisque la preuve serait identique à celle des deux théorèmes sus mentionnés.

Théorème 6.4. *Soient $j_1, \dots, j_P \in \{0, \dots, P\}$ et $l_1, \dots, l_P \in \{0, \dots, P\}$ deux ensembles d'indices pouvant être différents. Alors, pour $r = 1, \dots, P$,*

$$\begin{aligned}\operatorname{Cov}(\widetilde{X}_{j_1\dots j_r zw}, \widetilde{X}_{l_1\dots l_r zw}) &= \tilde{\sigma}^{(1\dots r)}_{j_1\dots j_r l_1\dots l_r} \\ &= \sum_{q=1}^{P}\Bigg[\sum_{i_1=1}^{P-q+1}\cdots\sum_{i_q=i_{q-1}+1}^{P} h(i_q, r)\Bigg(\prod_{k=i_1}^{i_r}\overline{\delta}_{0j_k}\Bigg) \\ &\qquad \times \delta_{j_{\jmath(q,r)}, l_{\jmath(q,r)}}\, b^{(i_1\dots i_q)}\, \widetilde{S}^{(1\dots i_q-1)}_{j_1\dots j_r z l_1\dots l_r \star_{\jmath(q-1,r)} z}\Bigg] \\ &\qquad + \delta_{j_1\dots j_r, l_1\dots l_r}\frac{b^{(1\dots r)}}{\zeta^{(1\dots r)}_{j_1\dots j_r}},\end{aligned} \tag{6.16}$$

où

$$h(i_q, r) = (1-\delta_{i_1\dots i_q, 1\dots r})(1-\delta_{i_1\dots i_q, 1\dots r+1})\cdots(1-\delta_{i_1\dots i_q, 1\dots P})$$

et

$$\widetilde{S}^{(1\dots q)}_{j_1\dots j_r l_1\dots l_r} = \begin{cases} 1, & si\ q < r \\ \displaystyle\sum_{j_{r+1}=1}^{J_{r+1}} \frac{\zeta^{(1\dots r+1)}_{j_1\dots j_{r+1}}}{\zeta^{(1\dots r+1)}_{j_1\dots j_r \Sigma}} \frac{\zeta^{(1\dots r+1)}_{l_1\dots l_r j_{r+1}}}{\zeta^{(1\dots r+1)}_{l_1\dots l_r \Sigma}}, & si\ q = r \\ 0, & si\ q > r, \end{cases}$$

$$\widetilde{S}^{(1\dots q)}_{j_1\dots j_r z l_1\dots l_r z} = \sum_{j_{r+1}=0}^{J_{r+1}}\sum_{l_{r+1}=0}^{J_{r+1}} \frac{\zeta^{(1\dots r+1)}_{j_1\dots j_{r+1}}}{\zeta^{(1\dots r+1)}_{j_1\dots j_r \Sigma}} \frac{\zeta^{(1\dots r+1)}_{l_1\dots l_{r+1}}}{\zeta^{(1\dots r+1)}_{l_1\dots l_r \Sigma}} \widetilde{S}^{(1\dots q)}_{j_1\dots j_{r+1} z l_1\dots l_{r+1} z},$$

$$\widetilde{S}^{(1\dots q)}_{j_1\dots j_r z l_1\dots l_r \star_{r+1} z} = \sum_{j_{r+1}=0}^{J_{r+1}}\sum_{l_{r+1}=0}^{J_{r+1}} \frac{\zeta^{(1\dots r+1)}_{j_1\dots j_{r+1}}}{\zeta^{(1\dots r+1)}_{j_1\dots j_r \Sigma}} \frac{\zeta^{(1\dots r+1)}_{l_1\dots l_r j_{r+1}}}{\zeta^{(1\dots r+1)}_{l_1\dots l_r \Sigma}} \widetilde{S}^{(1\dots q)}_{j_1\dots j_{r+1} z l_1\dots l_r j_{r+1} z}.$$

Si $i_q = 1$ dans (6.16), l'exposant de la fonction $\widetilde{S}$ est (0).

Il peut être utile de donner un exemple supplémentaire afin de bien illustrer les formules ci-dessus. Ainsi, dans modèle à quatre facteurs de risque, on a par exemple

$$\begin{aligned}\operatorname{Cov}(\widetilde{X}_{j_1 zw}, \widetilde{X}_{l_1 zw}) &= \delta_{j_1 l_1}\frac{b^{(1)}}{\zeta^{(1)}_{j_1}} + b^{(2)}\widetilde{S}^{(1)}_{j_1 l_1} + b^{(23)}\widetilde{S}^{(12)}_{j_1 z l_1 \star} + b^{(234)}\widetilde{S}^{(123)}_{j_1 z l_1 \star} \\ &\qquad + (b^{(3)} + \overline{\delta}_{0i}\delta_{j_1 l_1} b^{(13)})\widetilde{S}^{(12)}_{j_1 z l_1 z}\end{aligned}$$

$$+ (b^{(4)} + \overline{\delta}_{0i}\delta_{j_1l_1}b^{(14)})\widetilde{S}^{(123)}_{j_1zl_1z}$$
$$+ (b^{(24)} + \overline{\delta}_{0i}\delta_{j_1l_1}b^{(124)})\widetilde{S}^{(123)}_{j_1zzl_1\star z}$$
$$+ (b^{(34)} + \overline{\delta}_{0i}\delta_{j_1l_1}b^{(134)})\widetilde{S}^{(123)}_{j_1zzl_1z\star}.$$

On peut maintenant passer aux formules générales des estimateurs de Dannenburg et des pseudo-estimateurs optimaux.

6.2.2 Estimateurs de Dannenburg

La définition des estimateurs de Dannenburg des composants de variance $b^{(\cdot)}$ pour un modèle CCCG à P facteurs sera l'occasion de démontrer leur absence de biais. Ce point a, en effet, été escamoté dans les chapitres précédents. Nous démontrerons également l'absence de biais de l'estimateur s^2 sous sa forme générale. Mais, en premier lieu, nous considérons deux estimateurs de la moyenne collective.

L'estimateur de Dannenburg usuel de la moyenne collective m est la moyenne pondérée par les poids naturels :

$$\widehat{m} = X_w = \sum_{j_1=0}^{J_1} \cdots \sum_{j_P=0}^{J_P} \frac{w_{j_1\ldots j_P}}{w_\Sigma} X_{j_1\ldots j_P w}. \tag{6.17}$$

Néanmoins, nous croyons utile de mentionner ici la moyenne pondérée par les fonctions ζ, qui fut jusqu'ici classée dans les estimateurs ad hoc :

$$\widetilde{m} = \widetilde{X}_{zw} = \sum_{j_1=0}^{J_1} \cdots \sum_{j_P=0}^{J_P} \frac{\zeta^{(1)}_{j_1}}{\zeta^{(1)}_\Sigma} \cdots \frac{\zeta^{(1\ldots P)}_{j_1\ldots j_P}}{\zeta^{(1\ldots P)}_{j_1\ldots j_{P-1}\Sigma}} X_{j_1\ldots j_P w}. \tag{6.18}$$

Cet estimateur semble en effet avoir une variance inférieure à X_w, selon les tests du chapitre 8, et son calcul n'est guère plus complexe. Nous voyons donc peu de raisons de lui préférer la moyenne pondérée par les poids naturels comme estimateur de la moyenne collective.

Des chapitres précédents, l'on sait que l'estimateur classique du paramètre s^2 est optimal dans une certaine classe d'estimateurs. Son absence de biais n'a toutefois pas été formellement démontrée. Cette omission est réparée dans la preuve du théorème suivant, qui expose la version générale de l'estimateur $\hat{s}^2$.

Théorème 6.5. *L'expression ci-dessous est un estimateur sans biais du paramètre s^2 :*

$$\hat{s}^2 = \frac{1}{N^{(1\ldots P+1)} - N^{(1\ldots P)}} \sum_{j_1=0}^{J_1} \cdots \sum_{j_P=0}^{J_P} \sum_{t=1}^{T_{j_1\ldots j_P}} w_{j_1\ldots j_P t}(X_{j_1\ldots j_P t} - X_{j_1\ldots j_P w})^2. \tag{6.19}$$

Preuve. On vérifie l'absence de biais de l'estimateur s^2 :

$$\begin{aligned}
&\mathrm{E}\left[\sum_{j_1=0}^{J_1}\cdots\sum_{j_P=0}^{J_P}\sum_{t=1}^{T_{j_1\ldots j_P}} w_{j_1\ldots j_Pt}(X_{j_1\ldots j_Pt}-X_{j_1\ldots j_Pw})^2\right] = \\
&\quad= \sum_{j_1=0}^{J_1}\cdots\sum_{j_P=0}^{J_P}\sum_{t=1}^{T_{j_1\ldots j_P}} w_{j_1\ldots j_Pt}\left(\frac{s^2}{w_{j_1\ldots j_Pt}}-\frac{s^2}{w_{j_1\ldots j_P\Sigma}}\right) \\
&\quad= \left(\sum_{j_1=0}^{J_1}\cdots\sum_{j_P=0}^{J_P}\sum_{t=1}^{T_{j_1\ldots j_P}} I_{\{w_{j_1\ldots j_Pt}>0\}}-\sum_{j_1=0}^{J_1}\cdots\sum_{j_P=0}^{J_P} I_{\{w_{j_1\ldots j_P\Sigma}>0\}}\right)s^2 \\
&\quad= (N^{(1\ldots P+1)}-N^{(1\ldots P)})s^2,
\end{aligned}$$

car, en effet,

$$\begin{aligned}
\mathrm{Var}[X_{j_1\ldots j_Pt}] &- 2\,\mathrm{Cov}(X_{j_1\ldots j_Pt}, X_{j_1\ldots j_Pw}) \\
&+ \mathrm{Var}[X_{j_1\ldots j_Pw}] = \frac{s^2}{w_{j_1\ldots j_Pt}}-\frac{s^2}{w_{j_1\ldots j_P\Sigma}}.
\end{aligned}$$

On remarque que $N^{(1\ldots P+1)}-N^{(1\ldots P)} = \sum_{j_1=0}^{J_1}\cdots\sum_{j_P=0}^{J_P}(T_{j_1\ldots j_P}-1)$, avec $T_{j_1\ldots j_P}>1$. □

Les estimateurs de Dannenburg des composants de variance $b^{(\cdot)}$ sont définis comme les solutions d'un système d'équations linéaires basées sur certaines formes quadratiques des observations. Le théorème suivant permet de construire ces équations pour un modèle CCCG à P facteurs et montre que les estimateurs en résultant sont sans biais.

Théorème 6.6. *L'expression suivante est vérifiée pour* $r = 0,\ldots,P-1$ *:*

$$\begin{aligned}
&\mathrm{E}\Bigg[\sum_{j_1=0}^{J_1}\cdots\sum_{j_r=0}^{J_r}\tilde{g}^{(1\ldots r)}_{j_1\ldots j_r}\Bigg(\sum_{j_{r+1}=0}^{J_{r+1}}\cdots\sum_{j_P=0}^{J_P}\frac{w_{j_1\ldots j_P\Sigma}}{w_{j_1\ldots j_r\Sigma}}(X_{j_1\ldots j_Pw}-X_{j_1\ldots j_rw})^2 \\
&\qquad -(N^{(r+1\ldots P)}-1)\frac{s^2}{w_{j_1\ldots j_r\Sigma}}\Bigg)\Bigg] = \\
&= \sum_{j_1=0}^{J_1}\cdots\sum_{j_r=0}^{J_r}\tilde{g}^{(1\ldots r)}_{j_1\ldots j_r}\sum_{q=1}^{P}\Bigg[\sum_{i_1=1}^{P-q+1}\cdots\sum_{i_q=i_{q-1}+1}^{P}\left(\prod_{k\in\mathfrak{I}(q,r)}\bar{\delta}_{0j_k}\right)b^{(i_1\ldots i_q)} \\
&\quad\times\left(\sum_{j_{r+1}=0}^{J_{r+1}}\cdots\sum_{j_P=0}^{J_P}\left(\prod_{k\in\bar{\mathfrak{I}}(q,r)}\bar{\delta}_{0j_k}\right)\frac{w_{j_1\ldots j_P\Sigma}}{w_{j_1\ldots j_r\Sigma}}-\sum_{j_{\bar{\mathfrak{I}}(q,r)}=1}^{J_{\bar{\mathfrak{I}}(q,r)}}\left(\frac{w_{j_1\ldots j_rj_{\bar{\mathfrak{I}}(q,r)}\Sigma}}{w_{j_1\ldots j_r\Sigma}}\right)^2\right)\Bigg],
\end{aligned} \tag{6.20}$$

où les $\tilde{g}$ sont des poids quelconques dont la somme vaut 1. Lorsque $\bar{\mathfrak{I}}(q,r)=\emptyset$, le dernier produit et la dernière somme à la droite de (6.20) valent 1 et le facteur de $b^{(i_1\dots i_q)}$ est par conséquent nul pour cette valeur de q.

Preuve.

$$\mathrm{E}\left[\sum_{j_1=0}^{J_1}\cdots\sum_{j_r=0}^{J_r}\tilde{g}_{j_1\dots j_r}^{(1\dots r)}\sum_{j_{r+1}=0}^{J_{r+1}}\cdots\sum_{j_P=0}^{J_P}\frac{w_{j_1\dots j_P\Sigma}}{w_{j_1\dots j_r\Sigma}}(X_{j_1\dots j_Pw}-X_{j_1\dots j_rw})^2\right]=$$

$$\begin{aligned}
&=\sum_{j_1=0}^{J_1}\cdots\sum_{j_r=0}^{J_r}\tilde{g}_{j_1\dots j_r}^{(1\dots r)}\sum_{j_{r+1}=0}^{J_{r+1}}\cdots\sum_{j_P=0}^{J_P}\frac{w_{j_1\dots j_P\Sigma}}{w_{j_1\dots j_r\Sigma}}\\
&\quad\times\left(\mathrm{Var}[X_{j_1\dots j_Pw}]-2\,\mathrm{Cov}(X_{j_1\dots j_Pw},X_{j_1\dots j_rw})+\mathrm{Var}[X_{j_1\dots j_rw}]\right)\\
&=\sum_{j_1=0}^{J_1}\cdots\sum_{j_r=0}^{J_r}\tilde{g}_{j_1\dots j_r}^{(1\dots r)}\sum_{j_{r+1}=0}^{J_{r+1}}\cdots\sum_{j_P=0}^{J_P}\frac{w_{j_1\dots j_P\Sigma}}{w_{j_1\dots j_r\Sigma}}\\
&\quad\times\left\{\sum_{q=1}^{P}\left[\sum_{i_1=1}^{P-q+1}\cdots\sum_{i_q=i_{q-1}+1}^{P}b^{(i_1\dots i_q)}\left(\prod_{k\in\mathfrak{I}(q,P)}\bar{\delta}_{0j_k}\right.\right.\right.\\
&\quad-2\left(\prod_{k\in\mathfrak{I}(q,P)}\bar{\delta}_{0j_k}\right)\frac{w_{j_1\dots j_rj_{\bar{\mathfrak{I}}(q,r)}\Sigma}}{w_{j_1\dots j_r\Sigma}}+\prod_{k\in\mathfrak{I}(q,r)}\bar{\delta}_{0j_k}\left.\left.\sum_{j_{\bar{\mathfrak{I}}(q,r)}=1}^{J_{\bar{\mathfrak{I}}(q,r)}}\left(\frac{w_{j_1\dots j_rj_{\bar{\mathfrak{I}}(q,r)}\Sigma}}{w_{j_1\dots j_r\Sigma}}\right)^2\right)\right]\\
&\quad\left.+\frac{s^2}{w_{j_1\dots j_P\Sigma}}-\frac{s^2}{w_{j_1\dots j_r\Sigma}}\right\}\\
&=\sum_{j_1=0}^{J_1}\cdots\sum_{j_r=0}^{J_r}\tilde{g}_{j_1\dots j_r}^{(1\dots r)}\left\{\sum_{q=1}^{P}\left[\sum_{i_1=1}^{P-q+1}\cdots\sum_{i_q=i_{q-1}+1}^{P}\left(\prod_{k\in\mathfrak{I}(q,r)}\bar{\delta}_{0j_k}\right)b^{(i_1\dots i_q)}\right.\right.\\
&\quad\times\left(\sum_{j_{r+1}=0}^{J_{r+1}}\cdots\sum_{j_P=0}^{J_P}\left(\prod_{k\in\bar{\mathfrak{I}}(q,r)}\delta_{0j_k}\right)\frac{w_{j_1\dots j_P\Sigma}}{w_{j_1\dots j_r\Sigma}}\right.\\
&\quad-2\sum_{j_{r+1}=0}^{J_{r+1}}\cdots\sum_{j_P=0}^{J_P}\left(\prod_{k\in\bar{\mathfrak{I}}(q,r)}\delta_{0j_k}\right)\frac{w_{j_1\dots j_P\Sigma}}{w_{j_1\dots j_r\Sigma}}\frac{w_{j_1\dots j_rj_{\bar{\mathfrak{I}}(q,r)}\Sigma}}{w_{j_1\dots j_r\Sigma}}\\
&\quad\left.\left.+\sum_{j_{r+1}=0}^{J_{r+1}}\cdots\sum_{j_P=0}^{J_P}\frac{w_{j_1\dots j_P\Sigma}}{w_{j_1\dots j_r\Sigma}}\sum_{j_{\bar{\mathfrak{I}}(q,r)}=1}^{J_{\bar{\mathfrak{I}}(q,r)}}\left(\frac{w_{j_1\dots j_rj_{\bar{\mathfrak{I}}(q,r)}\Sigma}}{w_{j_1\dots j_r\Sigma}}\right)^2\right)\right]\\
&\quad\left.+\sum_{j_{r+1}=0}^{J_{r+1}}\cdots\sum_{j_P=0}^{J_P}\frac{w_{j_1\dots j_P\Sigma}}{w_{j_1\dots j_r\Sigma}}\left(\frac{s^2}{w_{j_1\dots j_P\Sigma}}-\frac{s^2}{w_{j_1\dots j_r\Sigma}}\right)\right\}
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{j_1=0}^{J_1} \cdots \sum_{j_r=0}^{J_r} \tilde{g}_{j_1\ldots j_r}^{(1\ldots r)} \sum_{q=1}^{P} \left[\sum_{i_1=1}^{P-q+1} \cdots \sum_{i_q=i_{q-1}+1}^{P} \left(\prod_{k \in \bar{\mathfrak{I}}(q,r)} \delta_{0j_k} \right) b^{(i_1\ldots i_q)} \right. \\
&\quad \times \left. \left(\sum_{j_{r+1}=0}^{J_{r+1}} \cdots \sum_{j_P=0}^{J_P} \left(\prod_{k \in \bar{\mathfrak{I}}(q,r)} \delta_{0j_k} \right) \frac{w_{j_1\ldots j_P\Sigma}}{w_{j_1\ldots j_r\Sigma}} - \sum_{j_{\mathfrak{I}(q,r)}=1}^{J_{\mathfrak{I}(q,r)}} \left(\frac{w_{j_1\ldots j_r j_{\mathfrak{I}(q,r)}\Sigma}}{w_{j_1\ldots j_r\Sigma}} \right)^2 \right) \right] \\
&\quad + \sum_{j_1=0}^{J_1} \cdots \sum_{j_r=0}^{J_r} \tilde{g}_{j_1\ldots j_r}^{(1\ldots r)} (N^{(r+1\ldots P)} - 1) \frac{s^2}{w_{j_1\ldots j_r\Sigma}}.
\end{aligned}
$$

□

Les composants de variance $b^{(1)}, \ldots, b^{(1\ldots P)}$ peuvent donc être estimés sans biais à l'aide des expressions du théorème 6.6 en enlevant l'opérateur espérance. Pour chaque valeur de r, il existe $\binom{P}{r}$ combinaisons différentes de l'ordre de sommation, d'où le total de $2^P - 1$ équations nécessaires à l'estimation de tous les paramètres $b^{(\cdot)}$.

6.2.3 Estimateurs optimaux

Un point favorable aux estimateurs optimaux introduits aux chapitres 4 et 5 — par rapport aux estimateurs ad hoc, notamment — est la facilité avec laquelle leurs formules s'étendent au cas à P facteurs de risque. En fait, seule la formule générale de covariance (6.16) cause un tant soit peu problème. Une fois celle-ci en main, les résultats des trois théorèmes suivants ne sont que de simples généralisations des théorèmes du chapitre 5. À ce titre, il n'apparaît pas essentiel que les théorèmes ci-dessous soient prouvés.

Le premier estimateur optimal proposé est toujours celui de la moyenne collective. On rappelle qu'aucune hypothèse d'excès nul n'intervient dans la définition de cet estimateur.

Théorème 6.7. *Dans un modèle CCCG à P facteurs, l'estimateur sans biais à variance minimale de m dans la classe des fonctions linéaires des observations est*

$$
\widehat{m}_* = \frac{\mathbf{B}'\mathbf{C}^{-1}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{X}, \tag{6.21}
$$

où

$$
\begin{aligned}
&\mathbf{X} = [\mathbf{X}_{j_1}]_{(J_1+1)\times 1}, && \cdots && \mathbf{X}_{j_1\ldots j_{P-1}} = [X_{j_1\ldots j_P w}]_{(J_P+1)\times 1}, \\
&\mathbf{B} = [\mathbf{B}_{j_1}]_{(J_1+1)\times 1}, && \cdots && \mathbf{B}_{j_1\ldots j_{P-1}} = [1]_{(J_P+1)\times 1}, \\
&\mathbf{C} = [\mathbf{C}_{j_1 l_1}]_{(J_1+1)\times(J_1+1)}, && \cdots &&
\end{aligned}
$$

$$\begin{aligned}\mathbf{C}_{j_1 l_1 \ldots j_{P-1} l_{P-1}} &= \sum_{q=1}^{P-1}\left[\sum_{i_1=1}^{P-q}\cdots\sum_{i_q=i_{q-1}+1}^{P-1}\left(\prod_{k=i_1}^{i_q}\overline{\delta}_{0j_k}\right)\delta_{j_{i_1}\ldots j_{i_q},l_{i_1}\ldots l_{i_q}} b^{(i_1\ldots i_q)}\mathbf{U}\right]\\ &+\sum_{q=0}^{P-2}\left[\sum_{i_1=1}^{P-q}\cdots\sum_{i_q=i_{q-1}+1}^{P-1}\left(\prod_{k=i_1}^{i_q}\overline{\delta}_{0j_k}\right)\delta_{j_{i_1}\ldots j_{i_q},l_{i_1}\ldots l_{i_q}} b^{(i_1\ldots i_q P)}(\mathbf{I}-\mathbf{J})\right]\\ &+\delta_{j_1\ldots j_{P-1},l_1\ldots l_{P-1}} b^{(1\ldots P)}\mathbf{Z}_{j_1\ldots j_{P-1}},\end{aligned}$$

avec $\mathbf{Z}_{j_1\ldots j_{P-1}}$ *une matrice diagonale* $(J_P+1)\times(J_P+1)$ *formée des réciproques des fonctions* $\zeta^{(1\ldots P)}_{j_1\ldots j_{P-1}0},\ldots,\zeta^{(1\ldots P)}_{j_1\ldots j_{P-1}J_P}$ *:*

$$\mathbf{Z}_{j_1\ldots j_{P-1}} = \operatorname{diag}([(\zeta^{(123)}_{j_1\ldots j_P})^{-1}]_{(J_P+1)\times 1}).$$

Rappelons que $\mathbf{U}$ est une matrice carrée ne contenant que des 1, $\mathbf{I}$ est la matrice identité et $\mathbf{J}$ est une matrice carrée contenant un 1 en position (1,1) et des zéros partout ailleurs.

L'hypothèse d'excès nul pour tous les effets aléatoires Ξ est introduite à partir d'ici afin de statuer sur l'optimalité de l'estimateur $\hat{s}^2$ et pour définir les estimateurs optimaux des composants de variance $b^{(\cdot)}$.

Théorème 6.8. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCCG à P facteurs, alors l'estimateur du paramètre de structure s^2 à variance minimale parmi toutes les combinaisons du type*

$$\sum_{j_1\ldots j_P turs}\alpha_{j_1\ldots j_P turs}(X_{j_1\ldots j_P t}-X_{j_1\ldots j_P u})(X_{j_1\ldots j_P r}-X_{j_1\ldots j_P s})$$

avec moyenne s^2 est $\hat{s}^2$, défini en (6.19).

Le théorème suivant couvre l'estimation optimale des composants de variance $b^{(1)}, b^{(12)}, b^{(123)},\ldots,b^{(1\ldots P)}$. Les estimateurs des autres composants de variance sont évidemment similaires.

Théorème 6.9. *Si toutes les variables Ξ ont un coefficient d'excès nul dans un modèle CCCG à P facteurs, alors l'estimateur du composant de variance $b^{(1\ldots r)}$, $r=1,\ldots P$, à variance minimale parmi les estimateurs de la forme*

$$\sum_{j_1=1}^{J_1}\cdots\sum_{j_r=1}^{J_r}\sum_{l_1\ldots l_r>j_1\ldots j_r}\alpha_{j_1\ldots j_r l_1\ldots l_r}(\widetilde{X}_{j_1\ldots j_r zw}-\widetilde{X}_{l_1\ldots l_r zw})^2 \tag{6.22}$$

avec espérance $b^{(1\ldots r)}$ est

$$\tilde{b}_*^{(1\ldots r)} = \mathbf{X}'\operatorname{diag}(\mathbf{A})\mathbf{X}, \tag{6.23}$$

où

$$\mathbf{A} = \frac{b^{(1\ldots r)}}{\mathbf{B}'\mathbf{C}^{-1}\mathbf{B}}\mathbf{C}^{-1}\mathbf{B}, \tag{6.24}$$

et

$$\begin{aligned}
\mathbf{B} &= [\tilde{\sigma}^{(1\ldots r)}_{j_1\ldots j_r j_1\ldots j_r} - 2\,\tilde{\sigma}^{(1\ldots r)}_{j_1\ldots j_r l_1\ldots l_r} + \tilde{\sigma}^{(1\ldots r)}_{l_1\ldots l_r l_1\ldots l_r}]_{\frac{N(N-1)}{2}\times 1},\\
\mathbf{C} &= [(\tilde{\sigma}^{(1\ldots r)}_{j_1\ldots j_r h_1\ldots h_r} - \tilde{\sigma}^{(1\ldots r)}_{j_1\ldots j_r q_1\ldots q_r} - \tilde{\sigma}^{(1\ldots r)}_{l_1\ldots l_r h_1\ldots h_r} + \tilde{\sigma}^{(1\ldots r)}_{l_1\ldots l_r q_1\ldots q_r})^2]_{\frac{N(N-1)}{2}\times\frac{N(N-1)}{2}},\\
\mathbf{X} &= [\widetilde{X}_{j_1\ldots j_r zw} - \widetilde{X}_{l_1\ldots l_r zw}]_{\frac{N(N-1)}{2}\times 1},\\
N &= J_1\cdots J_r
\end{aligned}$$

Dans l'expression de $\mathbf{C}$*, les indices* j *et* l *identifient les lignes de la matrice et les indices* h *et* q*, les colonnes. La notation* $l_1\ldots l_r > j_1\ldots j_r$ *est définie ainsi, pour* $l_1 \geq j_1 = 1,\ldots,J_1,\ \ldots,\ j_r = 1,\ldots,J_r$ *et* $s = 1,\ldots,r-1$ *:*

$$l_1\ldots l_r > j_1\ldots j_r \Leftrightarrow \begin{cases} l_{s+1} = 1,\ldots,J_{s+1}, \\ \quad\vdots \\ l_r = 1,\ldots,J_r, \end{cases} \quad \text{si} \quad \begin{array}{c} l_1 = j_1, \\ \vdots \\ l_{s-1} = j_{s-1}, \\ l_s > j_s. \end{array}$$

Ceci complète la partie théorique de cette thèse. Les deux derniers chapitres qui suivent se concentrent sur l'aspect plus appliqué de la performance des divers estimateurs des paramètres de structure lors de simulations. Nous tâchons par le fait même d'établir une hiérarchie entre les estimateurs selon des critères de précision, variance et robustesse.

Chapitre 7

Mise en oeuvre informatique

Ce bref chapitre amorce l'étude numérique des modèles de crédibilité à classification croisée standard et généralisé. Par diverses simulations, nous étudierons la performance des différents estimateurs des paramètres de structure présentés dans les chapitres précédents. Un exemple numérique comprenant les valeurs des estimateurs des effets aléatoires et des estimateurs de crédibilité de quelques assurés est également proposé.

Avant de s'intéresser aux résultats numériques, nous nous proposons toutefois de discuter sommairement de la programmation informatique des formules des modèles CCC et CCCG, ainsi que du modèle de simulation des données. Ce dernier point sera aussi l'occasion de traiter plus en détail la construction de distributions au coefficient d'excès nul autres que la loi normale.

7.1 Programmation des formules

La programmation des formules de l'estimateur de crédibilité et des estimateurs des paramètres de structure est faite en APL pour la rapidité avec laquelle ce langage permet de mettre en oeuvre des formules reposant sur un emploi intensif des vecteurs et des matrices. Le manque d'efficacité de ce langage interprété lors des simulations est ainsi largement récupéré lors de la programmation.

La programmation des formules nous semblant d'intérêt tant du point de vue algorithmique que méthodologique, la partie centrale du code source est présentée à l'annexe A. Cela est précédé d'une représentation schématique de la « hiérarchie » de toutes les fonctions impliquées dans le calcul de l'estimateur de crédibilité. Sommairement, la fonction principale `CCCG` prend en arguments les données observées et les poids naturels leur étant rattachés, demande à l'utilisateur quels estimateurs des composants de variance et de la moyenne collective il souhaite utiliser et retourne les estimateurs de crédibilité de l'expérience à venir. Cette fonction fait appel à plusieurs autres, certaines purement utilitaires, d'autres plus fondamentales. Dans cette dernière catégorie, pensons, entre autres, aux fonctions d'estimation des paramètres de structure, à `gen_matrice` — qui construit la matrice de calcul des $\widehat{\Xi}$ — et à `extrait_Xi` — qui extrait du vecteur des $\widehat{\Xi}$ ceux s'appliquant à un assuré donné du portefeuille.

Parce qu'elles sont codées à partir des formules générales du chapitre 6 la fonction `CCCG` et ses « descendantes » s'appliquent aussi bien à un modèle standard que généralisé, et ce pour un nombre arbitraire P de facteurs. Font toutefois exception à cette règle les fonctions calculant les estimateurs ad hoc et optimaux (et leurs descendantes). Ces fonctions ne sont en effet adaptées qu'aux cas à deux et trois facteurs. De plus, les fonctions pour le modèle généralisé sont distinctes de celles du modèle standard.

Plusieurs fonctions autres que `CCCG` utilisent le paramètre s^2 — ou plutôt son estimateur. Afin d'éviter de recalculer chaque fois cet estimateur ou de toujours le passer en argument, la valeur de l'estimateur $\hat{s}^2$ est conservée en variable globale dans l'espace de travail.

On sait que les pseudo-estimateurs ad hoc et optimaux des composants de variance sont obtenus par convergence d'un processus itératif. Or, le critère d'arrêt du processus choisi est le suivant :

$$\max |\mathbf{b}_{n+1} - \mathbf{b}_n| < \mathtt{TOL},$$

où $\mathbf{b}_n$ est un vecteur contenant les estimations des composants de variance de l'itération n. Le processus est également arrêté si la plus petite valeur de $\mathbf{b}_n$ passe sous le point `ZERO`, le composant de variance correspondant étant alors considéré comme nul. Les valeurs de `TOL` et `ZERO` — respectivement 10^{-6} et 10^{-12}, dans nos tests — sont stockées dans des variables globales utilisées par les fonctions de calcul des pseudo-estimateurs (`calculbt`, `calculbtG`, `calculbo` et `calculboG`). Afin d'accélérer les calculs lors de longues simulations, nous avons également limité le nombre d'itérations à `MAX_ITER` $= 50$. Cette limite n'a pas d'impact marqué sur les résultats puisqu'elle ne s'applique généralement qu'à de lentes convergences vers zéro.

La mise en oeuvre informatique des formules des pseudo-estimateurs optimaux des composants de variance est compliquée par la grande dimension de la matrice $\mathbf{C}$. Par exemple, pour un simple modèle CCC à deux facteurs ne comptant que $I = 3$ catégories pour le premier facteur de risque et $J = 4$ pour le second, la matrice $\mathbf{C}$ compte déjà 66 lignes et autant de colonnes, ce qui représente un total de 4 356 éléments ! Stocker en mémoire de telles matrices exige évidemment de grands espaces de travail et calculer tous leurs éléments prend beaucoup de temps. Ajoutons à cela que c'est d'abord et avant tout l'inverse de la matrice $\mathbf{C}$ qui est nécessaire pour calculer les pseudo-estimateurs et il devient clair qu'une certaine optimisation du code informatique s'impose.

La stratégie que nous avons retenue pour la fonction `construitCinv` est la suivante : puisqu'il est inutile de connaître explicitement la matrice $\mathbf{C}$, seul son inverse sera calculé au complet. Cela économise de l'espace mémoire et allège les calculs. Cette échappatoire repose sur la définition de l'inverse d'une matrice par blocs. Soit $\mathbf{A}$ une matrice $n \times n$ définie par blocs :

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix},$$

où $\mathbf{A}_{11}$ et $\mathbf{A}_{22}$ sont des matrices carrées de dimensions $r \times r$ et $(n-r)\times(n-r)$, respectivement. Définissons l'inverse de cette matrice de manière analogue :

$$\mathbf{A}^{-1} = \begin{bmatrix} \mathbf{A}^{11} & \mathbf{A}^{12} \\ \mathbf{A}^{21} & \mathbf{A}^{22} \end{bmatrix}.$$

En supposant que tous les inverses existent, les éléments de $\mathbf{A}^{-1}$ sont donnés par les expressions suivantes (Healy 1986) :

$$\begin{aligned} \mathbf{A}^{11} &= (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}, \\ \mathbf{A}^{12} &= -\mathbf{A}^{11}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}, \\ \mathbf{A}^{21} &= -\mathbf{A}_{22}^{-1}\mathbf{A}_{21}\mathbf{A}^{11}, \\ \mathbf{A}^{22} &= \mathbf{A}_{22}^{-1} - \mathbf{A}^{21}\mathbf{A}_{12}\mathbf{A}_{22}^{-1}. \end{aligned}$$

Supposons maintenant que l'inverse de la matrice $\mathbf{A}_{22}$ est connu et que $\mathbf{A}_{11}$ est de dimension 1×1. Calculer l'inverse de $\mathbf{A}$ ne requiert alors que l'inversion d'un scalaire et quelques produits matriciels. De plus, si $\mathbf{A}$ est symétrique — comme c'est le cas de la matrice $\mathbf{C}$ —, on peut même laisser tomber la troisième équation ci-dessus. En fait, cette situation correspond à ajouter une ligne et une colonne (identiques) à une matrice $\mathbf{A}_{22}$ dont on connaît l'inverse. On peut donc calculer directement et assez efficacement l'inverse de la matrice $\mathbf{C}$ en commençant avec $\mathbf{C} = c_{NN}$, $N = IJ(IJ-1)/2$, puis en ajoutant successivement une ligne et une colonne à la matrice précédente jusqu'à obtenir la matrice $\mathbf{C}^{-1}$ complète.

L'inverse d'une matrice s'obtient en APL avec la primitive ⌹, que l'on devine hautement optimisée. Le gain de temps réalisé par la procédure ci-dessus — comparativement au calcul de la matrice $\mathbf{C}$ suivi de son inversion

— n'est donc pas énorme. Par contre, le calcul direct de l'inverse permet de traiter de plus gros portefeuilles, à espace de travail égal.

Des détails supplémentaires sur les autres fonctions APL se trouvent à même le code source, à l'annexe A.

7.2 Simulation des données

Nous souhaitons généralement simuler des portefeuilles conformes aux hypothèses d'excès nul avec lesquelles les estimateurs optimaux ont été développés. Considérons, par exemple, le modèle CCC à deux facteurs de risque. On a

$$X_{ijt} = m + \Xi_i^{(1)} + \Xi_j^{(2)} + \Xi_{ij}^{(12)} + \Xi_{ijt}^{(123)}.$$

Une des manières les plus simples de se conformer aux hypothèses d'excès nul est d'adopter le modèle de simulation suivant :

$$\begin{aligned} \Xi_i^{(1)} &\sim \mathrm{N}(0, b^{(1)}), \\ \Xi_j^{(2)} &\sim \mathrm{N}(0, b^{(2)}), \\ \Xi_{ij}^{(12)} &\sim \mathrm{N}(0, b^{(12)}), \\ \Xi_{ijt}^{(123)} &\sim \mathrm{N}(0, s^2/w_{ijt}). \end{aligned}$$

Des lois normales sont utilisées pour la simplicité de leur simulation. Il existe néanmoins une infinité de distributions ayant un coefficient d'excès nul pouvant remplacer la loi normale. On montre à la section 7.3, ci-dessous, comment il est possible de construire de telles distributions.

Les poids naturels, quant à eux, sont simulés ainsi : en premier lieu, un poids moyen $\bar{w}_{ij}$ est tiré d'une loi uniforme sur $[2, 10]$ pour chaque cellule du portefeuille, puis les poids w_{ijt} sont tirés d'une loi uniforme sur $[0{,}5\,\bar{w}_{ij}, 1{,}5\,\bar{w}_{ij}]$. Nous préférons cette simulation en deux temps à une simple simulation de poids uniforme afin d'éviter de trop grandes variations de poids entre les années pour un même assuré.

Certaines fonctions n'étant pas adaptées aux poids nuls, tous les portefeuilles simulés comptent un nombre égal de données dans chaque cellule ($T_{ij} = T$, par exemple).

Si l'hypothèse d'excès nul ne s'avère pas une limitation dans le choix d'un modèle de simulation, nous proposons — pour mémoire, dirons-nous — le modèle suivant, faisant appel aux hypothèses du modèle « complet » présenté au chapitre 6. Dans un modèle CCCG à trois facteurs, soit S_{ijkt} le montant total des sinistres de l'assuré (ijk) pour la période t, w_{ijkt} le poids accordé à ces sinistres — nombre de sinistres, masse salariale ou autre, selon le contexte — et X_{ijkt} le ratio de l'assuré. On pose

$$\begin{aligned}
S_{ijkt}|\Theta_i^{(1)}, \Theta_j^{(2)}, \Theta_k^{(3)} &\sim \text{Poisson}(w_{ijkt}\, \tilde{\mu}_{123}(\Theta_i^{(1)}, \Theta_j^{(2)}, \Theta_k^{(3)})) \\
\Theta_i^{(1)} &\sim \text{N}(0, \sigma_1^2) \\
\Theta_j^{(2)} &\sim \text{N}(0, \sigma_2^2) \\
\Theta_k^{(3)} &\sim \text{N}(0, \sigma_3^2)
\end{aligned}$$

et $X_{ijkt} = S_{ijkt}/w_{ijkt}$, de sorte que

$$\text{E}\Big[X_{ijkt}|\Theta_i^{(1)}, \Theta_j^{(2)}, \Theta_k^{(3)}\Big] = \tilde{\mu}_{123}(\Theta_i^{(1)}, \Theta_j^{(2)}, \Theta_k^{(3)}).$$

Toutes les variables Θ sont indépendantes et les poids w_{ijkt} sont simulés comme ci-dessus.

Le choix de la prime de risque $\tilde{\mu}_{123}(\Theta_i^{(1)}, \Theta_j^{(2)}, \Theta_k^{(3)})$ s'avère délicat dans un modèle à nombre de facteurs variable puisque cette prime de risque doit tenir compte de possibles facteurs absents tout en satisfaisant les hypothèses de base du modèle CCC généralisé. Pour ce faire, on peut fixer une prime de risque pour un modèle CCC standard, déterminer les fonctions $\nu_1(\Theta_i^{(1)}), \dots, \nu_{123}(\Theta_{ijk}^{(123)})$ correspondantes, puis reconstruire une prime de risque $\tilde{\mu}_{\cdot}(\cdot)$ avec la relation (6.1). Ainsi, en partant d'une prime de risque du type

$$\begin{aligned}
\mu_{123}(\Theta_i^{(1)}, \Theta_j^{(2)}, \Theta_k^{(3)}) &= (\Theta_i^{(1)} + \Theta_j^{(2)} + \Theta_i^{(1)}\Theta_k^{(3)} + \Theta_j^{(2)}\Theta_k^{(3)})^2 \\
&= \Big[(\Theta_i^{(1)} + \Theta_j^{(2)})(\Theta_k^{(3)} + 1)\Big]^2,
\end{aligned}$$

on obtient la prime de risque à nombre de facteurs variable suivante :

$$\begin{aligned}
\tilde{\mu}_{123}(\Theta_i^{(1)}, \Theta_j^{(2)}, \Theta_k^{(3)}) = \Big[(&\overline{\delta}_{0i}\Theta_i^{(1)} + \overline{\delta}_{0j}\Theta_j^{(2)})(\overline{\delta}_{0k}\Theta_k^{(3)} + 1)\Big]^2 \\
&+ \delta_{0i}\sigma_1^2 \Big[\delta_{0k}(1 + \sigma_3^2) + \overline{\delta}_{0k}(\Theta_k^{(3)} + 1)^2\Big] \\
&+ \delta_{0j}\sigma_2^2 \Big[\delta_{0k}(1 + \sigma_3^2) + \overline{\delta}_{0k}(\Theta_k^{(3)} + 1)^2\Big] \\
&+ \delta_{0k} \Big[\overline{\delta}_{0i}\Theta_i^{(1)} + \overline{\delta}_{0j}\Theta_j^{(2)}\Big].
\end{aligned}$$

Cette prime de risque satisfait toutes les hypothèses du modèle, y compris l'absence de corrélation entre les fonctions $\nu_{\cdot}(\cdot)$ correspondantes. On peut alors démontrer que

$$\begin{aligned} m &= (\sigma_1^2+\sigma_2^2)(\sigma_3^2+1), & b^{(1)} &= 2\sigma_1^4(\sigma_3^2+1)^2, \\ b^{(2)} &= 2\sigma_2^4(\sigma_3^2+1)^2, & b^{(3)} &= 2(\sigma_1^2+\sigma_2^2)^2\sigma_3^2(\sigma_3^2+2), \\ b^{(12)} &= 4\sigma_1^2\sigma_2^2(\sigma_3^2+1)^2, & b^{(13)} &= 4\sigma_1^4\sigma_3^2(\sigma_3^2+2), \\ b^{(23)} &= 4\sigma_2^4\sigma_3^2(\sigma_3^2+2), & b^{(123)} &= 8\sigma_1^2\sigma_2^2\sigma_3^2(\sigma_3^2+2), \\ s^2 &= (\sigma_3^2+1). \end{aligned}$$

7.3 Autres distributions sans excès

Les détracteurs de l'hypothèse d'excès nul avancent souvent comme argument que la seule loi connue ayant un coefficient d'excès nul est la loi normale et que les travaux menés avec cette hypothèse auraient tout aussi bien pu l'être avec une usuelle hypothèse de normalité. Rappelons d'abord que l'hypothèse d'excès nul, telle qu'utilisée dans cette thèse, ne nécessite aucune hypothèse additionnelle sur le troisième moment des variables aléatoires. Par conséquent, les estimateurs optimaux dérivés ici conservent leurs propriétés même si les distributions des variables Ξ ne sont pas symétriques. Voilà qui est déjà plus général qu'une hypothèse de normalité.

De plus, que la loi normale soit la seule loi connue ayant un coefficient d'excès nul ne signifie évidemment pas qu'elle soit la seule existante. Nous montrons ici une façon — il en existe sûrement plusieurs autres — de construire une distribution sans excès. De plus, par le lemme 2.3, disposer d'une telle distribution équivaut à en avoir une infinité puisqu'il suffit de calculer autant de fois que désiré la convolution de cette distribution avec elle-même pour obtenir de nouvelles distributions sans excès.

La technique que nous proposons ici pour obtenir une distribution avec coefficient d'excès nul est finalement assez simple : il s'agit de mélanger adéquatement une distribution ayant un coefficient inférieur à zéro avec une autre ayant un coefficient supérieur à zéro. En d'autres termes, on mélange une distribution avec des queues moins lourdes que la loi normale avec une ayant des queues plus lourdes. Par exemple, la loi uniforme sur $[0, a]$ et la loi exponentielle de moyenne λ^{-1} ont, respectivement, un coefficient d'excès γ_2 de $-1{,}2$ et 3. Soit $f(x)$ la fonction de densité de la distribution X à excès nul cherchée, $u(x) = a^{-1}$ celle de la loi uniforme et $e(x) = \lambda e^{-\lambda x}$ celle de la loi exponentielle. On pose donc

$$f(x) = we(x)I_{\{x>0\}} + (1-w)u(x)I_{\{0\leq x\leq a\}}$$

Paramètres du mélange	E[X]	Var[X]
$\lambda = 1$ $a = 3{,}945058535$ $w = 0{,}5$	1,486264634	1,384931912
$\lambda = 1$ $a = 7{,}794576297$ $w = 0{,}6$	2,158915260	4,639807524

TABLEAU 7.1: Paramètres résultant en un mélange exponentielle/uniforme sans excès.

et on cherche les valeurs de w, a et λ telles que

$$\int_0^\infty (x-\mu)^4 f(x)\,dx = 3\left(\int_0^\infty (x-\mu)^2 f(x)\,dx\right)^2 \tag{7.1}$$

avec

$$\mu = \mathrm{E}[X] = \int_0^\infty x f(x)\,dx.$$

Il est évidemment préférable de rechercher une solution à l'aide d'un logiciel d'algèbre symbolique tel que Maple. En fixant, par exemple, la valeur de λ à 1, on obtient un ensemble de combinaisons possibles de w et a vérifiant (7.1). Deux de ces combinaisons sont présentées au tableau 7.1.

On trouve à la figure 7.1 le graphe du premier mélange (celui du second est tout à fait similaire) comparé avec celui d'une loi normale de même variance. On observe que le graphe du mélange comporte une discontinuité en $x = 3,945058535$, soit à l'endroit où cesse la contribution de la loi uniforme. Ce point correspond également à l'endroit où la queue de la loi normale commence à s'éteindre. L'intérêt du mélange réside toutefois principalement dans son héritage partiel de la queue de la loi exponentielle, plus longue que celle de la loi normale. On peut déjà déceler ce fait à la figure 7.1, mais il est illustré plus clairement à la figure 7.2. Les mélanges exponentielle/uniforme peuvent donc s'avérer utiles pour générer des ratios d'expérience relativement élevés tout en respectant l'hypothèse d'excès nul.

Le mélange exponentielle/uniforme n'est défini que sur les réels positifs. Or, la modélisation des variables Ξ requiert des distributions à excès nul de moyenne zéro définies préférablement pour des valeurs négatives et positives. Ceci est obtenu aisément par convolution des mélanges présentés ci-dessus. Soient X_1 et X_2 des variables aléatoires indépendantes dont la distribution

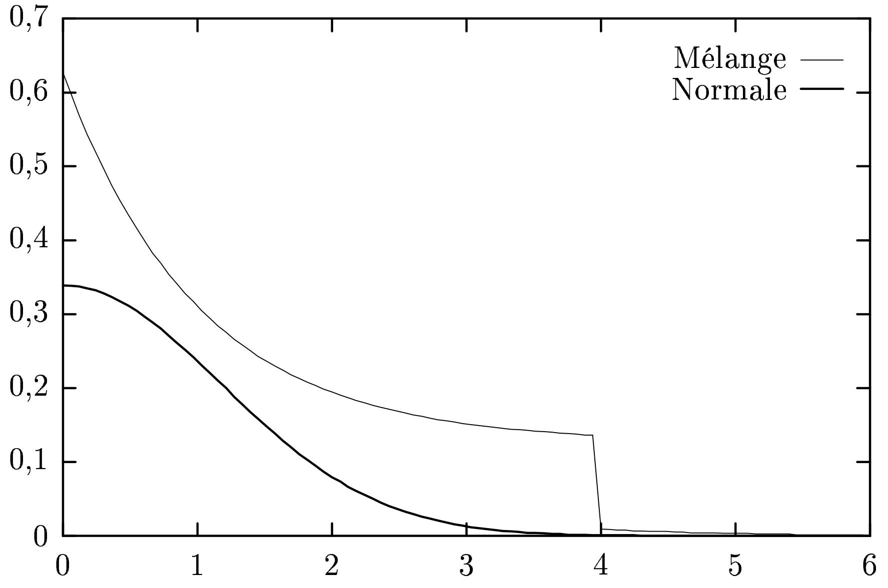

FIGURE 7.1: Comparaison entre la distribution du mélange exponentielle/uniforme et celle de la loi normale sur l'intervalle $[0,6]$.

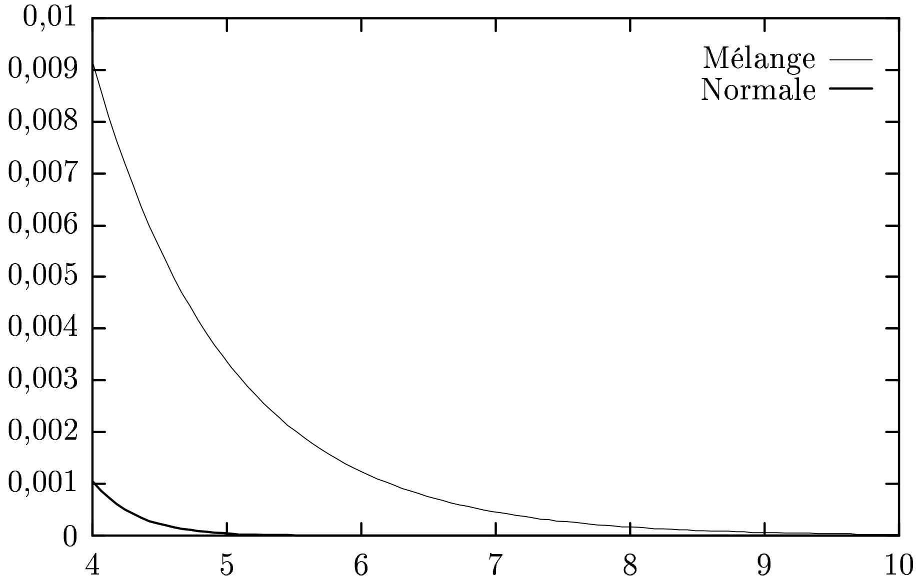

FIGURE 7.2: Comparaison entre la queue de la distribution du mélange exponentielle/uniforme et celle de la loi normale.

est donnée par le premier mélange du tableau 7.1 et X_3 une variable aléatoire indépendante de X_1 et X_2 et dont la distribution est donnée par le second mélange. Alors $Y_1 = X_1 - X_2$ est une variable aléatoire symétrique sans excès d'espérance nulle. Si l'on souhaite plutôt disposer d'une distribution non symétrique, on peut opter pour la convolution $Y_2 = X_3 - X_1 - 0{,}672650626$, qui a toujours une espérance nulle et est asymétrique vers la droite.

Il est clair que des lois autres que l'exponentielle peuvent être mélangées à une loi uniforme pour obtenir une distribution sans excès. Le choix fait ici est toutefois celui de la simplicité, tant de par la formule de la fonction de densité de probabilité que par le nombre réduit de paramètres (un seul).

Chapitre 8

Étude numérique des estimateurs

Nous avons présenté, dans les chapitres 3 à 5, trois grandes séries d'estimateurs des paramètres de structure pour les modèles de crédibilité à classification croisée standard et généralisé : les estimateurs de Dannenburg, les pseudo-estimateurs ad hoc et les pseudo-estimateurs optimaux (rappelons que l'estimateur du paramètre s^2 est le même dans les trois séries). Il nous a toutefois été impossible de statuer, par des arguments théoriques, sur la supériorité (ou non) des estimateurs ad hoc par rapport aux estimateurs de Dannenburg. Tout au plus sait-on que les deux séries sont sans biais et que les estimateurs ad hoc proviennent des estimateurs ANOVA, ayant eux-mêmes des propriétés d'optimalité connues. On aimerait également vérifier empiriquement que les estimateurs optimaux ont bel et bien une variance inférieure aux autres. Leur performance se dégrade-t-elle rapidement si les distributions des effets aléatoires ont un coefficient d'excès non nul ? C'est là tout le propos de ce chapitre.

La grande majorité des résultats de ce chapitre est obtenue avec le modèle CCC standard à deux facteurs. Le modèle généralisé est peu traité puisque, comme on l'a vu, les formules des estimateurs sont sensiblement les mêmes que celles du modèle standard. On ne vérifiera que rapidement que leur comportement l'est aussi. Les modèles à trois facteurs sont, pour leur part, souvent laissés de côté parce qu'il n'est raisonnable de calculer les estimateurs optimaux un grand nombre de fois que pour des portefeuilles si petits qu'il s'avère difficile de véritablement inférer à partir des résultats. Par exemple, pour un portefeuille modeste ne comportant que trois catégories pour chaque facteur de risque, la matrice $\mathbf{C}$ intervenant dans le calcul de $\tilde{b}_*^{(123)}$ compte déjà 351 lignes et autant de colonnes.

Dorénavant, on référera à un portefeuille comportant I catégories pour le premier facteur de risque, J pour le second et T données par cellule simplement par la formule « un portefeuille $I \times J \times T$ » (et ainsi de suite pour le modèle à trois facteurs).

Nombre de répétitions : 50 000
Dimensions du portefeuille : $6 \times 6 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation
$m = 5$	X_w	4,995301	0,994798	0,1991467
	X_{zw}	4,995167	0,989292	0,1980498
	$X_{zw}^{\diamond}$	4,995191	0,989299	0,1980502
	X_{zw}°	4,995185	0,989299	0,1980505
	$\widehat{m}_*$	4,995173	0,989292	0,1980495

TABLEAU 8.1: Comparaison des estimateurs de la moyenne collective pour un modèle à deux facteurs. Valeurs théoriques des facteurs de crédibilité.

8.1 Estimateurs de la moyenne collective

Notre étude numérique débute par les estimateurs de la moyenne collective afin de respecter l'ordre établi dans les chapitres précédents. C'est néanmoins dans cette section que les résultats se révèlent les moins concluants.

Au chapitre 3, trois moyennes pondérées par les facteurs de crédibilité différentes ont été introduites en tant qu'estimateurs ad hoc de la moyenne collective (voir page 38). Nous avons justifié le choix de la variante notée X_{zw} par le fait que la variance des trois alternatives était sensiblement la même lors de tests numériques, tout en étant inférieure à celle de X_w. On peut observer ces faits, ainsi que la performance de l'estimateur optimal, au tableau 8.1, qui présente les résultats de 50 000 répétitions d'un portefeuille $6 \times 6 \times 5$. Notons que, dans cette simulation, les valeurs théoriques des facteurs de crédibilité ont été utilisées pour calculer les estimateurs.

L'absence de biais de chaque estimateur est clairement vérifiée. Formellement, l'estimateur optimal $\widehat{m}_*$ présente effectivement ici le plus faible coefficient de variation. Il ne doit toutefois ce statut qu'à la septième décimale de la statistique ! En substance, le constat est donc plutôt que les moyennes pondérées X_{zw}, $X_{zw}^{\diamond}$ et X_{zw}° et l'estimateur optimal ont des performances virtuellement équivalentes, alors que la moyenne pondérée par les poids naturels, X_w, se montre très légèrement en retrait. Cet état de fait s'est vu confirmé lors d'autres simulations dont les résultats ne sont pas présentés ici. Les graphes des densités empiriques des cinq estimateurs, à la figure 8.1, renforce d'ailleurs notre conclusion : les tracés des quatre derniers estimateurs sont essentiellement confondus, seul le tracé de X_w se démarquant des autres.

Le portrait se trouble un peu plus lorsque, comme ce serait le cas en pratique, les facteurs de crédibilité estimés sont utilisés pour calculer les estimateurs ad hoc et optimal de la moyenne collective. Dans deux simulations

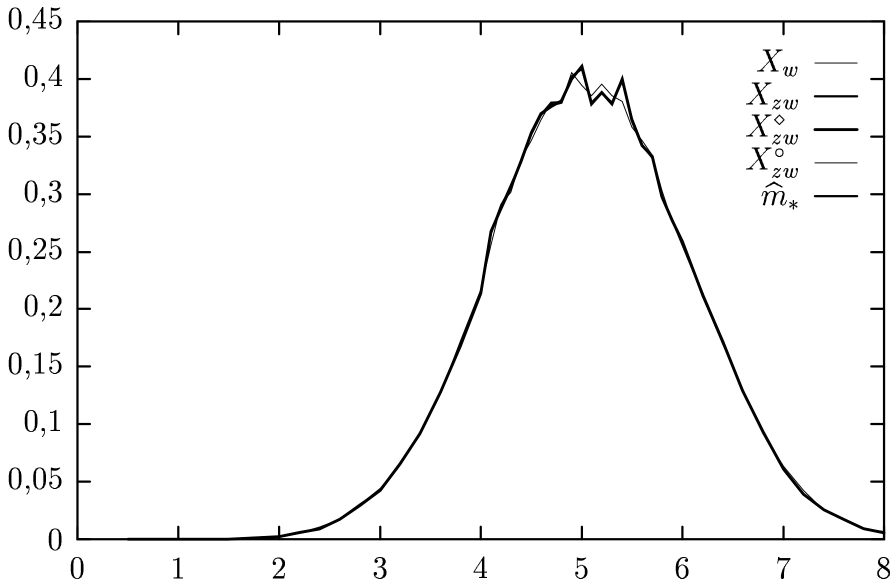

FIGURE 8.1: Densités empiriques des estimateurs de la moyenne collective.

nettement moins significatives que celle du tableau 8.1 — tant en nombre de répétitions qu'en taille du portefeuille — nous avons calculé les estimateurs X_{zw} et $\widehat{m}_*$ avec deux séries de facteurs de crédibilité : la première calculée avec les estimateurs ad hoc des composants de variance ($\tilde{b}$) et la seconde avec les estimateurs optimaux ($\tilde{b}_*$). Sans anticiper sur la section 8.3, mentionnons que les estimateurs optimaux des composants de variance ont une plus faible variance empirique dans les deux cas. Les résultats de ces deux simulations se trouvent aux tableaux 8.2 et 8.3. (On remarquera qu'une répétition a été enlevée dans ce dernier tableau, à cause d'un résultat complètement aberrant et inexplicable, les données des simulations n'étant pas sauvegardées.) On observe dans ces tableaux que, dans les deux cas, c'est l'estimateur ad hoc de la moyenne combiné avec les pseudo-estimateurs optimaux des composants de variance qui présente la plus faible variance et le plus faible coefficient de variation. Dans le modèle à deux facteurs, cette combinaison est toutefois presque à égalité avec la combinaison de tous les estimateurs optimaux. D'un autre côté, l'estimateur optimal de la moyenne a la plus forte variance dans le modèle à trois facteurs, peu importe les estimateurs des composants de variance utilisés. L'estimateur X_w performe donc mieux que $\widehat{m}_*$ dans certains cas. La portée de ces résultats reste néanmoins limitée étant donné la taille très restreinte des portefeuilles.

Des tests numériques menés sur les estimateurs de la moyenne collective, il ressort donc que l'estimateur ad hoc X_{zw} et l'estimateur optimal semblent

Nombre de répétitions : 500
Dimensions du portefeuille : $3 \times 3 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation
$m = 5$	X_w	4,9737	1,2507	0,2515
	$X_{zw}/\tilde{b}$	4,9748	1,3317	0,2677
	$X_{zw}/\tilde{b}_*$	4,9903	1,2318	0,2468
	$\widehat{m}_*/\tilde{b}$	4,9792	1,2662	0,2543
	$\widehat{m}_*/\tilde{b}_*$	4,9908	1,2321	0,2469

TABLEAU 8.2: Comparaison des estimateurs de la moyenne collective pour un modèle à deux facteurs. Facteurs de crédibilité estimés.

Nombre de répétitions : 499
Dimensions du portefeuille : $2 \times 2 \times 3 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation
$m = 5$	X_w	5,0735	2,1086	0,4156
	$X_{zw}/\tilde{b}$	5,0623	2,3147	0,4572
	$X_{zw}/\tilde{b}_*$	5,0666	2,1014	0,4148
	$\widehat{m}_*/\tilde{b}$	5,0204	2,4657	0,4911
	$\widehat{m}_*/\tilde{b}_*$	4,9999	2,3887	0,4777

TABLEAU 8.3: Comparaison des estimateurs de la moyenne collective pour un modèle à trois facteurs. Facteurs de crédibilité estimés.

généralement avoir une variance inférieure à la moyenne pondérée par les poids naturels. Il serait toutefois aventureux, dans l'état actuel des choses, de départager définitivement l'estimateur ad hoc et l'estimateur optimal. Ajoutons néanmoins que l'estimateur optimal semble mieux adapté aux situations où un ou quelques composants de variance sont beaucoup plus grands (ou petits) que les autres.

8.2 Estimateur du paramètre s^2

Cette section est très brève, l'estimateur $\hat{s}^2$ ne posant pas de problèmes particuliers dans des situations normales. De plus, cet estimateur est tout à fait équivalent à celui rencontré dans le modèle de Bühlmann–Straub. Les études parues dans ce contexte s'appliquent donc au modèle sous étude ici.

On trouvera aux tableaux 8.4–8.6 les résultats de quelques simulations. Il est intéressant d'y constater que le coefficient de variation se maintient aux environs de 0,175.

Mentionnons que, comme tout estimateur de variance, $\hat{s}^2$ est très peu robuste aux données extrêmes. Ceci peu avoir de fâcheuses conséquences sur l'estimation des autres paramètres de structure puisque le paramètre s^2 apparaît dans les formules de leurs estimateurs (Goulet 1994, chapitre 8).

8.3 Estimateurs des composants de variance

Les tests menés sur les estimateurs des composants de variance sont les plus exhaustifs. En effet, en plus des usuelles comparaisons entre les estimateurs de Dannenburg et les pseudo-estimateurs ad hoc et optimaux, cette section comporte des études sur le comportement de ces estimateurs dans certaines situations spéciales : distribution des ratios d'expérience avec un coefficient d'excès non nul, facteurs de risque dépendants, données extrêmes.

Toutes les simulations de cette sections comptent 300 répétitions. À ce stade, il s'est avéré que la variance empirique des estimateurs est généralement suffisamment stable. La taille des portefeuilles est modeste afin de ne pas davantage rallonger des simulations prenant déjà plusieurs heures. Rappelons, en effet, que le calcul des pseudo-estimateurs optimaux demande beaucoup de temps, même pour de petits portefeuilles.

8.3.1 Modèle de simulation standard

Les résultats de diverses simulations faites selon le modèle de la section 7.2 sont présentés aux tableaux 8.7–8.10. Les trois premiers tableaux pré-

Nombre de répétitions : 500
Dimensions du portefeuille : $4 \times 4 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation
$s^2 = 5$	$\hat{s}^2$	5,051	0,879	0,174

TABLEAU 8.4: Performance de l'estimateur du paramètre s^2 pour un modèle à deux facteurs.

Nombre de répétitions : 500
Dimensions du portefeuille : $4 \times 4 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation
$s^2 = 25$	$\hat{s}^2$	24,747	4,622	0,187

TABLEAU 8.5: Performance de l'estimateur du paramètre s^2 pour un modèle à deux facteurs.

Nombre de répétitions : 500
Dimensions du portefeuille : $3 \times 3 \times 2 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation
$s^2 = 10$	$\hat{s}^2$	10,111	1,739	0,172

TABLEAU 8.6: Performance de l'estimateur du paramètre s^2 pour un modèle à trois facteurs.

Nombre de répétitions : 300
Dimensions du portefeuille : $4 \times 4 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation	< 0
$b^{(1)} = 0{,}5$	$\hat{b}^{(1)}$	0,516	0,668	1,294	68
	$\tilde{b}^{(1)}$	0,558	0,586	1,050	0
	$\tilde{b}_*^{(1)}$	0,563	0,588	1,044	0
$b^{(2)} = 0{,}2$	$\hat{b}^{(2)}$	0,199	0,428	2,149	104
	$\tilde{b}^{(2)}$	0,248	0,320	1,290	0
	$\tilde{b}_*^{(2)}$	0,249	0,317	1,273	0
$b^{(12)} = 0{,}7$	$\hat{b}^{(12)}$	0,718	0,495	0,689	10
	$\tilde{b}^{(12)}$	0,702	0,418	0,595	1
	$\tilde{b}_*^{(12)}$	0,651	0,366	0,562	0

TABLEAU 8.7: Comparaison des estimateurs des composants de variance pour un modèle à deux facteurs. Petites valeurs des paramètres.

sentent les statistiques des simulations d'un modèle à deux facteurs, mais avec diverses valeurs des composants de variance. Les statistiques du quatrième tableau se rapportent à un modèle à trois facteurs. On remarquera, dans ces tableaux, la colonne additionnelle donnant le nombre total d'estimateurs négatifs lors de la simulation. (Les statistiques sont calculées avec les valeurs négatives des estimateurs.)

Ces simulations donnent exactement les résultats espérés et attendus. En premier lieu, tous les estimateurs apparaissent sans biais. Les estimateurs de Dannenburg, $\hat{b}^{(\cdot)}$, montreraient cependant sûrement un biais si les valeurs négatives des estimateurs étaient remplacées par des valeurs nulles. Ensuite, la variance — ou le coefficient de variation — des pseudo-estimateurs optimaux $\tilde{b}_*^{(\cdot)}$ est légèrement inférieure à celle des pseudo-estimateurs ad hoc $\tilde{b}^{(\cdot)}$, qui est à son tour davantage inférieure à celle des estimateurs de Dannenburg. Cette assertion est valide autant pour le modèle à deux qu'à trois facteurs et ce, peu importe la magnitude des paramètres. La réduction du coefficient de variation avec les pseudo-estimateurs ad hoc et optimaux atteint jusqu'à 40 %, dans certains cas, par rapport aux estimateurs de Dannenburg.

L'une des raisons évoquées précédemment pour rechercher de nouveaux estimateurs des composants de variance dans le modèle CCC était la possibilité, avec les estimateurs de Dannenburg, d'obtenir des valeurs négatives. La dernière colonne des tableaux 8.7–8.10 montre à quel point cela s'avère justifié : dans certains cas, plus du tiers des estimateurs de Dannenburg sont

Nombre de répétitions: 300
Dimensions du portefeuille: $4 \times 4 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation	< 0
$b^{(1)} = 2$	$\hat{b}^{(1)}$	1,825	2,249	1,232	62
	$\tilde{b}^{(1)}$	1,898	2,037	1,073	0
	$\tilde{b}_*^{(1)}$	1,906	2,035	1,067	0
$b^{(2)} = 1{,}5$	$\hat{b}^{(2)}$	1,478	2,216	1,499	78
	$\tilde{b}^{(2)}$	1,556	1,838	1,181	0
	$\tilde{b}_*^{(2)}$	1,563	1,825	1,168	0
$b^{(12)} = 3$	$\hat{b}^{(12)}$	2,783	1,579	0,568	9
	$\tilde{b}^{(12)}$	2,846	1,327	0,466	2
	$\tilde{b}_*^{(12)}$	2,700	1,179	0,437	0

TABLEAU 8.8: Comparaison des estimateurs des composants de variance pour un modèle à deux facteurs. Valeurs modérées des paramètres.

Nombre de répétitions: 300
Dimensions du portefeuille: $4 \times 4 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation	< 0
$b^{(1)} = 10$	$\hat{b}^{(1)}$	9,709	15,081	1,553	71
	$\tilde{b}^{(1)}$	10,880	13,050	1,200	0
	$\tilde{b}_*^{(1)}$	10,945	13,068	1,194	0
$b^{(2)} = 15$	$\hat{b}^{(2)}$	16,126	19,264	1,195	56
	$\tilde{b}^{(2)}$	16,946	17,729	1,046	0
	$\tilde{b}_*^{(2)}$	17,094	17,766	1,039	0
$b^{(12)} = 20$	$\hat{b}^{(12)}$	20,028	11,837	0,591	6
	$\tilde{b}^{(12)}$	19,599	9,819	0,501	1
	$\tilde{b}_*^{(12)}$	18,580	8,783	0,473	0

TABLEAU 8.9: Comparaison des estimateurs des composants de variance pour un modèle à deux facteurs. Grandes valeurs des paramètres.

Nombre de répétitions : 300
Dimensions du portefeuille : $3 \times 3 \times 2 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation	< 0
$b^{(1)} = 2$	$\hat{b}^{(1)}$	1,793	6,777	3,779	118
	$\tilde{b}^{(1)}$	3,015	6,623	2,197	70
	$\tilde{b}^{(1)}_*$	2,860	4,420	1,546	0
$b^{(2)} = 2$	$\hat{b}^{(2)}$	1,946	7,205	3,703	120
	$\tilde{b}^{(2)}$	3,391	8,328	2,456	57
	$\tilde{b}^{(2)}_*$	3,039	4,631	1,524	0
$b^{(3)} = 2$	$\hat{b}^{(3)}$	1,689	8,536	5,054	147
	$\tilde{b}^{(3)}$	2,648	8,519	3,218	101
	$\tilde{b}^{(3)}_*$	3,661	6,646	1,815	0
$b^{(12)} = 3$	$\hat{b}^{(12)}$	2,944	6,715	2,281	96
	$\tilde{b}^{(12)}$	3,196	8,759	2,741	4
	$\tilde{b}^{(12)}_*$	2,216	2,518	1,136	0
$b^{(13)} = 3$	$\hat{b}^{(13)}$	3,746	7,387	1,972	93
	$\tilde{b}^{(13)}$	3,745	5,322	1,421	6
	$\tilde{b}^{(13)}_*$	3,251	4,660	1,433	0
$b^{(23)} = 3$	$\hat{b}^{(23)}$	3,501	6,815	1,946	102
	$\tilde{b}^{(23)}$	2,297	6,103	2,092	8
	$\tilde{b}^{(23)}_*$	2,915	3,845	1,319	0
$b^{(123)} = 4$	$\hat{b}^{(123)}$	3,456	7,260	2,101	78
	$\tilde{b}^{(123)}$	3,626	3,141	0,866	26
	$\tilde{b}^{(123)}_*$	3,889	2,552	0,656	0

TABLEAU 8.10: Comparaison des estimateurs des composants de variance pour un modèle à trois facteurs. Valeurs modérées des paramètres.

négatifs ! Les pseudo-estimateurs ad hoc prennent parfois aussi des valeurs négatives — surtout lorsque le volume de données est limité, comme au tableau 8.10 —, mais moins fréquemment que les estimateurs de Dannenburg. Il ne faut pas se surprendre de ce fait, étant donnée la parenté des estimateurs ad hoc et des estimateurs ANOVA. En revanche, on remarque avec intérêt que les pseudo-estimateurs optimaux ne sont jamais négatifs.

La performance des estimateurs s'améliore — la variance diminue, moins de valeurs négatives — lorsque l'on descend dans les tableaux simplement parce que le nombre de données (ou de « points ») augmente. Par exemple, dans le tableau 8.7, les estimateurs de $b^{(1)}$ ne disposent que de quatre points, alors que ceux de $b^{(12)}$ bénéficient de seize.

Remplacer les lois normales du modèle de simulation par la convolution $Y_2 = X_3 - X_1 - 0,672650626$, présentée à la toute fin du chapitre précédent, n'a pas d'impact notable sur les estimateurs. Cela vérifie au moins que la propriété d'optimalité des estimateurs $b_*^{(\cdot)}$ ne dépend pas de la symétrie des distributions des variables Ξ.

Tel que mentionné au début de ce chapitre, nous nous attarderons peu sur l'estimation des composants de variance dans le modèle généralisé. On a, en effet, toutes les raisons de croire que la performance des estimateurs y sera semblable à celle des estimateurs du modèle standard. Diverses simulations ont d'ailleurs corroboré cette intuition. Nous ne présentons ici, au tableau 8.11, que les résultats relatifs à un modèle à deux facteurs. Tout au plus fera-t-on remarquer le nombre nettement réduit d'estimateurs de Dannenburg négatifs et leur performance parfois meilleure que celle des pseudo-estimateurs ad hoc. Cela est sans doute dû à la présence d'une catégorie additionnelle — la catégorie 0 — pour chaque facteur de risque. Les données des catégories 0 sont utilisées dans le calcul de ces estimateurs, mais non dans celui des pseudo-estimateurs ad hoc et optimaux (sauf pour les moyennes pondérées).

Les résultats obtenus jusqu'ici l'ont été avec un modèle de simulation aux « belles » propriétés : toutes les hypothèses du modèle CCC sont respectées et les distributions des effets aléatoires sont des lois normales, sans excès et ne générant pas de données extrêmes. Voyons maintenant l'effet sur les estimateurs d'une perturbation de ce portrait.

8.3.2 Hypothèse d'excès nul non respectée

La perturbation sans doute la plus intéressante à imposer au modèle de simulation consiste à ne pas respecter l'hypothèse d'excès nul sous laquelle sont développés les pseudo-estimateurs optimaux. Pour ce faire, la loi normale est remplacée, dans le modèle de simulation, par la distribution de Laplace,

Nombre de répétitions : 300
Dimensions du portefeuille : $5 \times 5 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation	< 0
$b^{(1)} = 2$	$\hat{b}^{(1)}$	1,942	2,104	1,083	43
	$\tilde{b}^{(1)}$	1,785	3,327	1,864	0
	$\tilde{b}_*^{(1)}$	1,917	1,720	0,898	0
$b^{(2)} = 2{,}5$	$\hat{b}^{(2)}$	2,860	2,899	1,014	25
	$\tilde{b}^{(2)}$	3,142	6,549	2,085	0
	$\tilde{b}_*^{(2)}$	2,760	2,445	0,886	0
$b^{(12)} = 3$	$\hat{b}^{(12)}$	3,106	1,809	0,582	5
	$\tilde{b}^{(12)}$	3,182	1,561	0,491	0
	$\tilde{b}_*^{(12)}$	3,226	1,495	0,464	0

TABLEAU 8.11: Comparaison des estimateurs des composants de variance pour un modèle généralisé à deux facteurs. Valeurs modérées des paramètres.

dont la fonction de densité est

$$f(x) = \frac{1}{2}\lambda\, e^{-\lambda|x|}, \quad -\infty < x < \infty.$$

Le coefficient d'excès de cette distribution symétrique autour de l'origine vaut 3 ; ses queues sont donc plus lourdes que celles d'une normale.

Nous avons refait les simulations des tableaux 8.7 et 8.8 avec des distributions de Laplace. Or, le bilan se révèle positif pour les pseudo-estimateurs optimaux puisque leur variance (empirique) reste peu affectée par le viol de l'hypothèse d'excès nul. La variance augmente, certes, mais dans les mêmes proportions que celle des autres estimateurs des paramètres. Cette augmentation est attribuable aux queues lourdes de la distribution de Laplace. La hiérarchie des estimateurs établie à la sous-section précédente n'est donc pas modifiée.

Le tableau 8.12 présente les résultats de la simulation avec de petites valeurs des paramètres.

8.3.3 Facteurs de risque dépendants

Une hypothèse fondamentale du modèle CCC est l'indépendance mutuelle[1] de tous les effets aléatoires. En pratique, cette hypothèse est très

[1] L'indépendance complète est requise pour le développement d'estimateurs optimaux. Sinon, l'absence de corrélation suffit.

Nombre de répétitions : 300
Dimensions du portefeuille : $4 \times 4 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation	< 0
$b^{(1)} = 0{,}5$	$\hat{b}^{(1)}$	0,491	0,763	1,553	64
	$\tilde{b}^{(1)}$	0,552	0,741	1,343	0
	$\tilde{b}^{(1)}_*$	0,560	0,743	1,327	0
$b^{(2)} = 0{,}2$	$\hat{b}^{(2)}$	0,144	0,399	2,775	118
	$\tilde{b}^{(2)}$	0,215	0,337	1,567	0
	$\tilde{b}^{(2)}_*$	0,215	0,332	1,545	0
$b^{(12)} = 0{,}7$	$\hat{b}^{(12)}$	0,753	0,624	0,829	8
	$\tilde{b}^{(12)}$	0,745	0,588	0,790	2
	$\tilde{b}^{(12)}_*$	0,669	0,488	0,729	0

TABLEAU 8.12: Comparaison des estimateurs des composants de variance lorsque l'hypothèse d'excès nul n'est pas respectée. Petites valeurs des paramètres.

susceptible d'être peu ou pas du tout vérifiée. L'usage consiste toutefois généralement à supposer que l'hypothèse est presque valable et que cette approximation n'a pas d'effet trop néfaste sur les calculs. Mais qu'en est-il si les facteurs de risque sont très corrélés ?

Nous avons testé l'effet d'une dépendance des facteurs de risque sur les estimateurs des composants de variance avec deux modèles différents, tous deux pour un modèle CCC à deux facteurs. Dans le premier, toutes les variables $\Xi_i^{(1)}$ et $\Xi_j^{(2)}$ demeurent mutuellement indépendantes, mais on pose

$$\Xi_{ij}^{(12)} = \Xi_i^{(1)} + \Xi_j^{(2)}.$$

Le coefficient de corrélation entre $\Xi_{ij}^{(12)}$ et $\Xi_i^{(1)}$, par exemple, est alors donné par

$$\rho(\Xi_i^{(1)}, \Xi_{ij}^{(12)}) = \left(\frac{b^{(1)}}{b^{(1)} + b^{(2)}} \right)^{1/2}.$$

Avec les paramètres $b^{(1)} = 2$ et $b^{(2)} = 1$, on a $b^{(12)} = 3$ et le coefficient de corrélation est donc égal à $\sqrt{2/3} \approx 0{,}82$. Il s'agit là d'une forte corrélation entre des variables qui devraient être indépendantes. Comme on peut le constater au tableau 8.13, l'effet sur les estimateurs est d'ailleurs assez marqué, mais pas tant au niveau de la variance (les coefficients de corrélation sont

Nombre de répétitions : 300
Dimensions du portefeuille : $4 \times 4 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation	< 0
$b^{(1)} = 2$	$\hat{b}^{(1)}$	7,462	5,863	0,786	2
	$\tilde{b}^{(1)}$	3,820	47,739	12,498	12
	$\tilde{b}_*^{(1)}$	7,430	5,716	0,769	0
$b^{(2)} = 1$	$\hat{b}^{(2)}$	3,839	4,126	1,075	16
	$\tilde{b}^{(2)}$	5,589	28,123	5,032	14
	$\tilde{b}_*^{(2)}$	3,784	3,724	0,984	0
$b^{(12)} = 3$	$\hat{b}^{(12)}$	−0,047	1,623	−34,421	141
	$\tilde{b}^{(12)}$	0,086	0,354	4,114	132
	$\tilde{b}_*^{(12)}$	0,053	0,082	1,539	0

TABLEAU 8.13: Comparaison des estimateurs des composants de variance lorsque l'hypothèse d'indépendance entre les effets aléatoires n'est pas respectée. Effet aléatoire de l'interaction dépendant des deux autres.

similaires à ceux obtenus auparavant, à l'exception des estimateurs ad hoc) qu'au niveau du biais. En effet, les estimateurs de $b^{(1)}$ et $b^{(2)}$ sont environ quatre fois plus élevés que les valeurs théoriques, alors que ceux de $b^{(12)}$ sont beaucoup trop petits. La moyenne des estimateurs de Dannenburg est même négative dans ce dernier cas. On note que les pseudo-estimateurs ad hoc sont davantage affectés que les autres.

Le second modèle de simulation avec facteurs de risque dépendants a eu un impact nettement moindre sur les estimateurs. Dans ce modèle, le second effet aléatoire est défini comme la somme du premier effet aléatoire et d'une loi normale centrée à zéro :

$$\Xi_j^{(2)} = \Xi_i^{(1)} + \mathrm{N}(0, \sigma^2).$$

Les autres effets aléatoires ne sont pas modifiés par rapport au modèle de simulation usuel. Le coefficient de corrélation entre $\Xi_i^{(1)}$ et $\Xi_j^{(2)}$ est :

$$\rho(\Xi_i^{(1)}, \Xi_j^{(2)}) = \left(\frac{b^{(1)}}{b^{(1)} + \sigma^2} \right)^{1/2}.$$

Si $b^{(1)} = 2$ et $\sigma^2 = 1$, alors $b^{(2)} = 3$ et, comme précédemment, $\rho = \sqrt{2/3} \approx 0{,}82$.

Les résultats du tableau 8.14 montrent que les différents estimateurs des composants de variance ne sont, pour ainsi dire, pas affectés par la forte

Nombre de répétitions : 300
Dimensions du portefeuille : $4 \times 4 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation	< 0
$b^{(1)} = 2$	$\hat{b}^{(1)}$	1,828	2,075	1,135	30
	$\tilde{b}^{(1)}$	2,437	13,725	5,632	2
	$\tilde{b}_*^{(1)}$	1,839	1,880	1,022	0
$b^{(2)} = 3$	$\hat{b}^{(2)}$	2,900	2,679	0,924	16
	$\tilde{b}^{(2)}$	2,878	4,477	1,555	2
	$\tilde{b}_*^{(2)}$	2,853	2,413	0,846	0
$b^{(12)} = 3$	$\hat{b}^{(12)}$	0,970	0,995	1,026	39
	$\tilde{b}^{(12)}$	0,947	0,641	0,677	25
	$\tilde{b}_*^{(12)}$	0,986	0,554	0,562	0

TABLEAU 8.14: Comparaison des estimateurs des composants de variance lorsque l'hypothèse d'indépendance entre les effets aléatoires n'est pas respectée. Effet aléatoire du second facteur de risque dépendant du premier.

corrélation entre les variables $\Xi_i^{(1)}$ et $\Xi_j^{(2)}$ — si l'on excepte l'estimateur $\tilde{b}^{(1)}$. Il semble donc que les dépendances entre les facteurs de risque ne soient pas toutes égales au regard de leur effet sur l'estimation des paramètres de structure. Nous ne nous avancerons cependant pas jusqu'à affirmer que, dans tous les cas, une corrélation du second type ci-dessus est moins dommageable qu'une corrélation du premier type. Nos tests sont trop sommaires pour cela.

8.3.4 Présence de données extrêmes

Les estimateurs de variance sont par essence très sensibles aux données extrêmes. La définition d'estimateurs robustes pour le modèle de Bühlmann–Straub a d'ailleurs été une préoccupation au début des années quatre-vingt-dix (Künsch 1992, Gisler & Reinhard 1993). Les estimateurs introduits dans cette thèse n'ont évidemment aucune prétention à la robustesse. Il peut néanmoins être intéressant de comparer leur comportement en présence d'une seule donnée extrême dans un portefeuille. Pour ce faire, nous avons simplement choisi un ratio d'expérience au hasard dans un portefeuille simulé avec le modèle usuel et multiplié celui-ci par 10. L'effet sur les estimateurs fut alors négligeable.

Nous avons donc repris l'expérience en multipliant cette fois le ratio choisi au hasard par 50. On constate au tableau 8.15 que tous les estimateurs ont à peu près aussi mal réagi à cette donnée extrême, particulièrement au niveau

Nombre de répétitions: 300
Dimensions du portefeuille: $4 \times 4 \times 5$

Paramètre	Estimateur	Moyenne	Écart-type	Coefficient de variation	< 0
$b^{(1)} = 2$	$\hat{b}^{(1)}$	27.869	112.389	4.033	116
	$\tilde{b}^{(1)}$	14.142	39.586	2.799	1
	$\tilde{b}_*^{(1)}$	17.049	42.600	2.499	0
$b^{(2)} = 1{,}5$	$\hat{b}^{(2)}$	26.919	122.209	4.540	122
	$\tilde{b}^{(2)}$	11.665	36.854	3.159	0
	$\tilde{b}_*^{(2)}$	13.624	38.997	2.862	0
$b^{(12)} = 3$	$\hat{b}^{(12)}$	−18.590	203.283	−10.935	165
	$\tilde{b}^{(12)}$	42.920	88.940	2.072	1
	$\tilde{b}_*^{(12)}$	31.739	79.748	2.513	0

TABLEAU 8.15: Comparaison des estimateurs des composants de variance lorsque une donnée extrême est insérée dans le portefeuille.

du biais. Les pseudo-estimateurs optimaux conservent néanmoins généralement le plus faible coefficient de variation.

8.4 Estimateurs de crédibilité

À titre d'exemple numérique, nous présentons ici les estimateurs de crédibilité de quelques assurés d'un portefeuille $4 \times 4 \times 4 \times 4$. Celui-ci contenait les données d'un modèle généralisé à trois facteurs de risque dont les paramètres de structure théoriques sont $m = 5$, $b^{(1)} = b^{(2)} = b^{(3)} = 2$, $b^{(12)} = b^{(13)} = b^{(23)} = 3$, $b^{(123)} = 4$ et $s^2 = 5$. Les estimateurs optimaux des paramètres de structure ont donné les résultats suivants:

$$
\begin{aligned}
\widehat{m}_* &= 6{,}013 & \tilde{b}_*^{(1)} &= 1{,}216 \\
\tilde{b}_*^{(2)} &= 0{,}472 & \tilde{b}_*^{(3)} &= 0{,}209 \\
\tilde{b}_*^{(12)} &= 3{,}396 & \tilde{b}_*^{(13)} &= 4{,}437 \\
\tilde{b}_*^{(23)} &= 0{,}866 & \tilde{b}_*^{(123)} &= 5{,}049 \\
\hat{s}^2 &= 5{,}129.
\end{aligned}
$$

De plus, puisque $\widehat{\Xi}_1^{(1)} = -0{,}798$, $\widehat{\Xi}_1^{(2)} = -2{,}647$, $\widehat{\Xi}_1^{(3)} = -0{,}103$, $\widehat{\Xi}_{11}^{(12)} = 1{,}545$, $\widehat{\Xi}_{11}^{(13)} = 1{,}488$, $\widehat{\Xi}_{11}^{(23)} = -0{,}923$ et sachant que $X_{100w} = 4{,}991$, $X_{011w} =$

6,135, $X_{111w} = 4{,}054$, on obtient, de (5.2) :

$$\begin{aligned}
\widehat{X}_{1005} &= 6{,}013 + (0)(4{,}968 - 6{,}013) \\
&\quad + (1)(-0{,}798) = 5{,}215, \\
\widehat{X}_{0115} &= 6{,}013 + (0)(2{,}319 - 6{,}013) \\
&\quad + (1)(-2{,}647 - 0{,}103 - 0{,}923) = 2{,}339, \\
\widehat{X}_{1115} &= 6{,}013 + (0{,}967)(4{,}084 - 6{,}013) \\
&\quad + (0{,}033)(-0{,}798 - 2{,}647 - 0{,}103 \\
&\quad + 1{,}545 + 1{,}488 - 0{,}923) = 4{,}100.
\end{aligned}$$

Malgré qu'elle soit réduite par rapport à celle des autres estimateurs, la variance des pseudo-estimateurs optimaux demeure relativement importante. Dans cet exemple précis, les estimateurs des paramètres sont, en effet, assez éloignés des valeurs théoriques.

8.5 Quels estimateurs utiliser ?

Après cette avalanche de chiffres, il convient sans doute de résumer les propriétés numériques des estimateurs des paramètres de structure étudiés ici et de tirer une conclusion. Nous aimerions évidemment en arriver à plébisciter l'emploi des estimateurs optimaux pour leurs propriétés théoriques. Or, si le bilan de l'estimateur optimal de la moyenne est plutôt contrasté, celui des pseudo-estimateurs des composants de variance est par contre excellent. Ainsi, ces estimateurs se sont révélés demeurer toujours positifs (ou non négatifs) et avoir la plus faible variance des estimateurs testés. De plus, ils ne semblent pas trop sensibles au non respect de l'hypothèse d'excès nul sous laquelle ils sont développés et il est aisé de généraliser leurs formules aux cas à plus de trois facteurs de risque. Leur principal désavantage, cependant, est le temps requis pour leur calcul. À titre indicatif, il faut compter environ quatre heures pour obtenir les résultats du tableau 8.7 sur un PC muni d'un processeur PENTIUM cadencé à 200 MHz. De ces quatre heures, le calcul des estimateurs de Dannenburg et des pseudo-estimateurs ad hoc occupe plus ou moins dix minutes...

Nous pensons que, dans l'état actuel des recherches, seuls les pseudo-estimateurs ad hoc se présentent comme une alternative intéressante aux estimateurs optimaux. La variance des premiers est en effet généralement similaire à celle des seconds, alors que leur calcul est très rapide. Par contre, la procédure de calcul des pseudo-estimateurs ad hoc résulte parfois en des valeurs négatives. La transposition des formules de ces estimateurs à un cas à plus de trois facteurs n'est également pas chose aisée, quoique réalisable.

Ajoutons, en terminant, que les pseudo-estimateurs optimaux semblent plus fiables lorsque le volume de données est très limité, comme c'était souvent le cas lors de nos simulations. Pour avoir une meilleure idée de la performance des pseudo-estimateurs ad hoc, nous avons dû, à quelques reprises, simuler de plus grands portefeuilles.

Conclusion

La thèse qui s'achève fut consacrée à trois éléments principaux, tous relatifs au modèle de crédibilité à classification croisée de Dannenburg (1995) : introduction d'estimateurs des paramètres de structure à variance minimale (ou optimaux), définition d'un modèle généralisé permettant un nombre de facteurs de risque variable par assuré et, enfin, écriture générale des formules des modèles standard et généralisé pour un nombre quelconque de facteurs de risque. Ces développements se sont faits à l'aide de la théorie des espaces de Hilbert, un puissant outil non utilisé jusqu'ici en crédibilité à classification croisée.

Parmi les quelques chemins de traverse empruntés lors de notre recherche des estimateurs optimaux, l'un se sera révélé particulièrement opportun puisqu'il aura mené vers la famille des estimateurs ad hoc des paramètres de structure. Dans le cas des composants de variance, ces estimateurs généralisent les estimateurs ANOVA, bien connus dans la théorie des modèles linéaires. Cette famille comprend également la moyenne des ratios pondérée par les facteurs de crédibilité, un estimateur de la moyenne collective inspiré de celui du modèle de Bühlmann–Straub.

Enfin, un peu en périphérie de ces travaux, nous trouvons, au chapitre 7, la description d'une technique pour développer des distributions ayant un coefficient d'excès nul autres que la normale.

Après une rapide introduction à la théorie de la crédibilité en général et au modèle à classification croisée en particulier, nous avons présenté, au chapitre 2, les principes de base de la théorie des espaces de Hilbert et des modèles à composants de variance, trois lemmes abondamment employés par la suite et les premières conventions d'écriture. Au chapitre suivant, nous avons calculé l'estimateur de crédibilité du modèle CCC à trois facteurs à l'aide de projections, présenté la version à trois facteurs de risque des estimateurs de Dannenburg des paramètres de structure et introduit les pseudo-estimateurs ad hoc. Le développement des pseudo-estimateurs optimaux a fait l'objet de tout le chapitre 4. Le modèle CCC généralisé a par la suite été présenté en suivant les lignes des deux chapitres précédents. Enfin, la partie théorique de la thèse s'est achevée au chapitre 6, avec l'exposé des formules du modèle CCCG à $P > 1$ facteurs de risque. Ne restait alors plus qu'à discuter sommairement de la mise en oeuvre informatique des formules étudiées jusque là et du modèle de simulation des données, puis d'étudier les résultats de simulation. Il en est ressorti, au chapitre 8, que les pseudo-estimateurs optimaux des paramètres de structure ont une performance généralement meilleure que les autres estimateurs sous étude, mais qu'ils peuvent être très longs à calculer.

Nous n'avons pas traité, dans ce texte, des applications pratiques des modèles de crédibilité à classification croisée. Entre autres possibilités, ces modèles pourraient certainement se montrer utiles dans le domaine de la tarification géographique. Les modèles de crédibilité n'ont, semble-t-il, pas encore été employés dans ce domaine. La généralisation du modèle CCC standard admettant un nombre de facteurs de risque variable par assuré a d'ailleurs été motivée par une telle application. Les recherches préliminaires se sont toutefois révélées infructueuses.

Si les pseudo-estimateurs optimaux des composants de variance se sont montrés intéressants de par leur variance réduite par rapport aux alternatives et, surtout, de par leur propriété de ne jamais tomber sous zéro, leur calcul pose, en revanche, un réel problème. Il y a certaine là place à amélioration et, donc, à de futures recherches.

Finalement, dans une perspective plus large, l'approche de la théorie de la crédibilité par les modèles à composants de variance initiée par Dannenburg (1995) et poursuivie par Dannenburg et al. (1996) nous a semblé assez prometteuse. L'extension à des modèles à composants de covariance pourrait sans doute ouvrir de nouvelles possibilités comme, par exemple, la redéfinition du modèle à contrats dépendants de Cossette (1996).

Annexe A

Code source APL

On trouvera dans cette annexe le code source APL de la fonction `CCCG` présentée au chapitre 7 et d'une partie de ses descendantes. La hiérarchie de toutes les fonctions est d'abord illustrée dans les figures ci-dessous.

Les boîtes aux lignes épaisses identifient les fonctions ayant des descendantes.

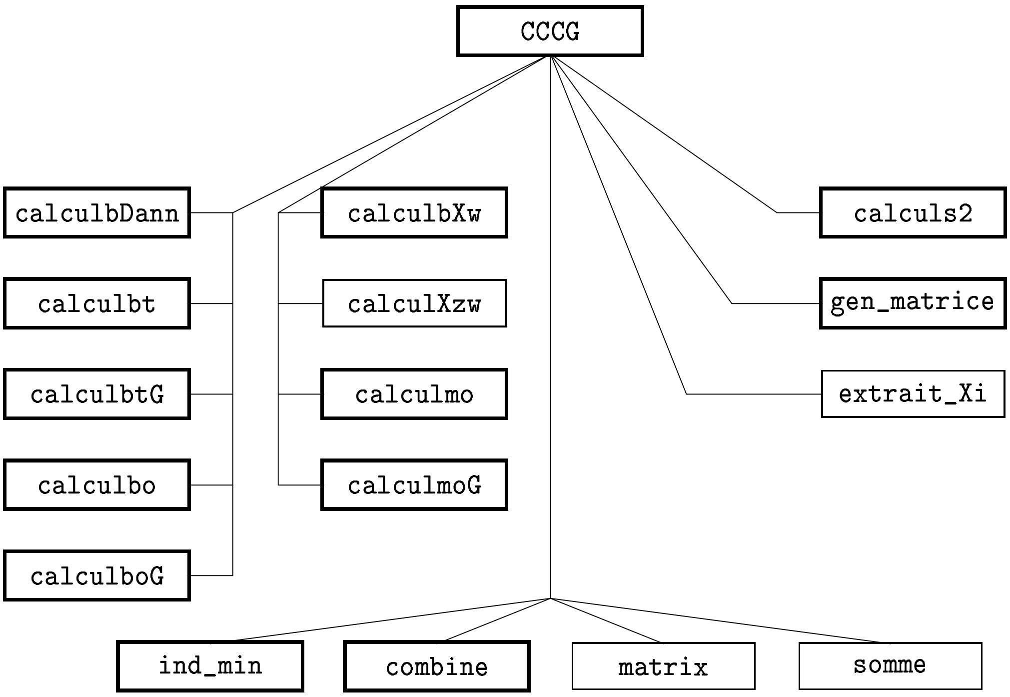

FIGURE A.1: Fonctions appelées par `CCCG`.

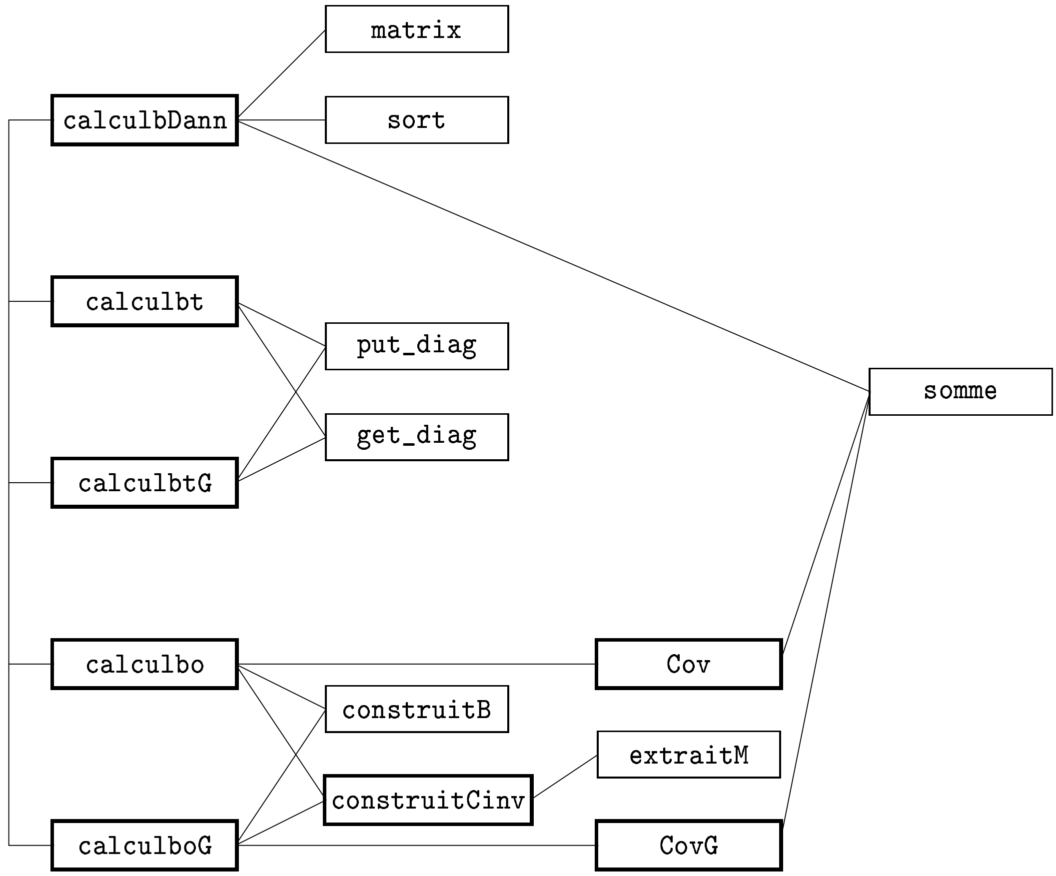

FIGURE A.2: Fonctions appelées par les différentes fonctions de calcul des estimateurs des composants de variance.

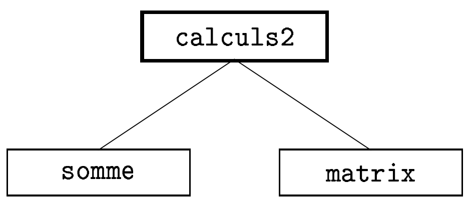

FIGURE A.3: Fonctions appelées par `calculs2`.

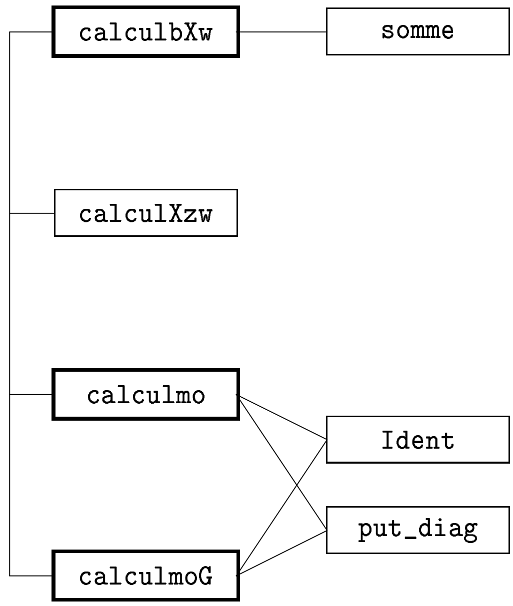

FIGURE A.4: Fonctions appelées par les différentes fonctions de calcul de l'estimateur de la moyenne collective.

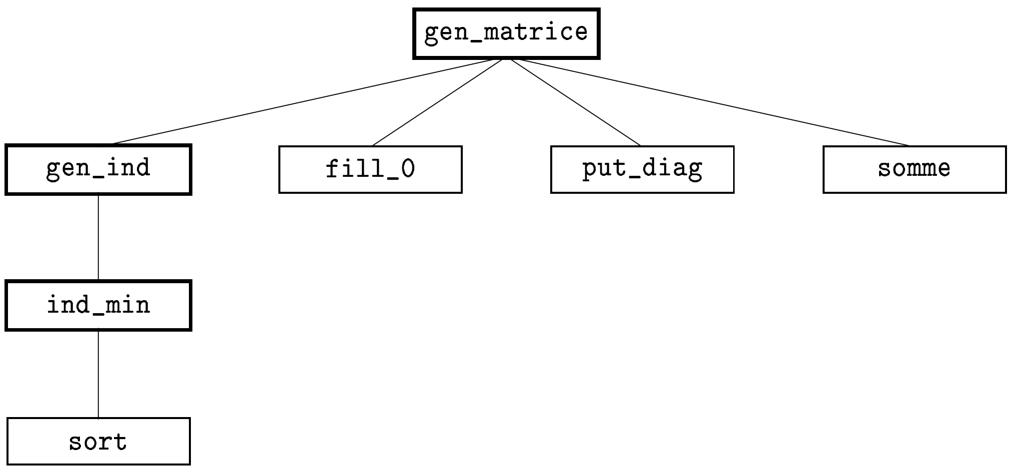

FIGURE A.5: Fonctions appelées par `gen_matrice` et `ind_min`.

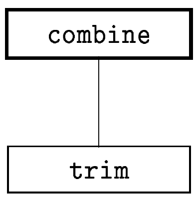

FIGURE A.6: Fonction appelée par `combine`.

Seul le code source complet de CCCG et celui, partiel, de ses principales fonctions descendantes est présenté ici.

Les différentes fonctions font appel, en plus des primitives de l'APL de base, à des tableaux imbriqués (*nested arrays*) introduits dans l'APL II, ainsi qu'à des structures de contrôle (IF, WHILE, REPEAT, etc.) propres à l'APL III.

Les données et les poids naturels d'un modèle CCCG contiennent généralement des valeurs nulles, correspondant à des cellules vides. On remarquera, dans le code des fonctions appelées à traiter de telles données, les mesures prises pour éviter les divisions par zéro.

Dans la plupart des fonctions, la variable locale t est une variable temporaire servant à de multiples usages. Il ne faut donc pas s'attacher à la première définition de la variable apparaissant dans le code d'une fonction, mais bien à la dernière.

CCCG

```
      ∇ z ← X CCCG w;⎕io;dim;P;nb_est;c;i;j;X_;w_;Xzw;m;cred;ze
        ta;Xi;poids_z;s;b;t;ci;cj;d;Q1;Q2;G
[1]
[2]     ⍝ Calculs pour un modèle CCC généralisé avec un nombre
[3]     ⍝ de facteurs quelconque. Le nombre de données par
[4]     ⍝ assuré est le même pour tous les assurés. Pour un
[5]     ⍝ modèle CCC standard, ajouter des zéros aux P
[6]     ⍝ premières dimensions des données et des poids.
[7]     ⍝
[8]     ⍝ Quelques résultats intermédiaires sont affichés à
[9]     ⍝ l'écran.
[10]
[11]    ⍝ Options pour l'utilisateur
[12]
[13]    'Quels estimateurs des composants de variance utiliser?
        '
[14]    '     1. Estimateurs de Dannenburg'
[15]    '     2. Estimateurs ad hoc'
[16]    '     3. Estimateurs optimaux'
[17]    ⍞ ← 'Entrer une option: ' ⋄ Q1←⍞
[18]
[19]    ''
[20]    'Quel estimateur de la moyenne colective utiliser?'
[21]    '     1. Moyenne pondérée par les poids naturels'
[22]    '     2. Moyenne pondérée par les fonctions zeta'
[23]    '     3. Estimateur optimal'
[24]    ⍞ ← 'Entrer une option: ' ⋄ Q2 ← ⍞
[25]
```

```
⍝ Initialisation des variables

⎕io ← 1

nb_est ← +/(⍳P) ! P ← ¯1 + ⍴dim ← ⍴X
cred ← zeta ← Xzw ← poids_z ← nb_est ⍴ 0
z ← ⍬

c ← (combine P), ⊂⍳P + 1

X_ ← (w_ > 0) × (+/w × X) ÷ w_ ← +/w

⍝ Estimation de s2 (variable globale!)

'---- s2 ----'
+s2 ← X calculs2 w

⍝ Estimation des composants de variance

⍝ Si tous les poids de la première dimension sont nuls,
⍝ on a un modèle standard

    ⍝ On calcule toujours les estimateurs de Dannenburg
    ⍝ pour avoir des valeurs de départ
    b ← c calculbDann X_ w_

'---- b  ----'

:IF G ← ∧/0 = ,(1, 1 ↓ dim) ↑ w
'std'
    ⍎ ⊃('+b ← b' '+b ← b calculbt_q (t ↓ X) ((t ← (P ⍴ 1), 0) ↓ w)' '+b ← b calculbo_q (t ↓ X) ((t ← (P ⍴ 1), 0) ↓ w)')[⍎Q1]
:ELSE
'gen'
    ⍎ ⊃('+b ← b' '+b ← b calculbtG X w' '+b ← b calculboG X w')[⍎Q1]
:END

⍝ La dernière dimension n'est plus nécessaire

dim ← ¯1 ↓ dim

⍝ Calcul des facteurs de crédibilité, des Xzw
⍝ et des poids de la matrice de calcul des Xi

    ⍝ La matrice t contient des 0 où les facteurs
    ⍝ de crédibilité sont nuls et des 1 ailleurs
    d ← (-dim) ↑ (dim - 1) ⍴ 1

cred[nb_est] ← ⊂d × s ÷ s2 + s ← b[nb_est] × w_
zeta[nb_est] ← ⊂s ÷ s2 + d × s
Xzw[nb_est] ← ⊂X_
```

```
poids_z[nb_est] ← ⊂dim ⍴ 1

:FOR i :IN ⌽⍳nb_est - 1
    j ← c ⍳ ⊂P ind_min ci ← ⊃c[i]
    d ← (-dim[ci]) ↑ (dim[ci] - 1) ⍴ 1
    cred[i] ← ⊂d × s ÷ b[j] + s ← b[i] × t ← ((cj ← ⊃c[j]) ~ ci) somme ⊃zeta[j]
    zeta[i] ← ⊂s ÷ b[j] + d × s ← b[i] × t
    Xzw[i] ← ⊂(t > 0) × ((cj ~ ci) somme ⊃Xzw[j] × zeta[j]) ÷ t
    poids_z[i] ← ⊂(⊃poids_z[j]) × (dim cj matrix ⊃zeta[j]) ÷ dim ci matrix t
:ENDFOR

⍝ Estimation de m

:IF G
    ⍎ ⊃('m ← X_ calculXw w_' 'm ← X_ calculXzw zeta[c ⍳ P gen_ind 1]' 'm ← (t ↓ X_) calculmo b ((t ← P ⍴ 1) ↓ nb_est ⊃ zeta)')[⍎Q2]
:ELSE
    ⍎ ⊃('m ← X_ calculXw w_' 'm ← X_ calculXzw zeta[c ⍳ P gen_ind 1]' 'm ← X_ calculmoG b (nb_est ⊃ zeta)')[⍎Q2]
:END

⍝ Calcul des Xi

Xi ← (∊cred[⍳nb_est - 1] × Xzw[⍳nb_est - 1] - m) ⌹ c[⍳nb_est] gen_matrice cred poids_z

⍝ Mettre certains Xi à 0 (qui ne le sont pas pour
⍝ cause d'arrondi)

t ← 0
:FOR i :IN ⍳nb_est - 1
    Xi[1 + t] ← 0
    t ← t + ×/dim[⊃c[i]]
:ENDFOR

'---- Xi ----'
Xi

'---- zi ----'
+cred ← ⊃cred[nb_est]

'---- m ----'
m

'--- Résultats ---'

cred ← ,cred ⋄ X_ ← ,X_
:FOR i :IN ⍳×/dim
    t ← dim ⊤ ¯1 + i
```

```
    z ← z, m + (cred[i] × X_[i] - m) + (1 - cred[i]) ×
+/t dim (c[⍳nb_est]) extrait_Xi Xi
:ENDFOR

z ← dim ⍴ z
∇
```

calculbDann

```
∇ z ← c calculbDann Xw;⎕io;X_;w_;ww;ww0;nb_est;P;nb_boucl
es;ordre;ind_f;ind_d;Ib;e;E;S;r;i;j;t;s;f

⍝ Calcule les estimateurs de Dannenburg des composants
⍝ de variance dans un modèle CCCG à P facteurs.
⍝
⍝ Pour modèle CCC, ajouter une ligne et une colonne
⍝ de zéros aux données.
⍝
⍝ Note: les poids g utilisés dans le calcul sont tous
⍝ égaux à 1. De plus, Ti == T pour tout i.
⍝
⍝   Xw[⎕io] == ratios
⍝ Xw[⎕io+1] == poids naturels
⍝         c == combinaisons de tous les facteurs de
⍝              risque

⎕io ← 1

X_ ← (+/(2 ⊃ Xw) × 1 ⊃ Xw) ÷ w_ ← +/2 ⊃ Xw

nb_est ← +/(⍳P) ! P ← ⍴⍴X_
S ← E ← ⍬

⍝ Calcul des «espérances»

:FOR r :IN ⍳P

    nb_boucles ← ⍴ordre ← ⌽(r = ∊⍴¨ c)/c

    :FOR i :IN ⍳nb_boucles
        ind_f ← ⊃ordre[i] ⋄ ind_d ← (⍳P) ~ ind_f

        ⍝ Calcul des Xw et ww appropries
        Xw ← (ww0 ← ww > 0) × (ind_f somme w_ × X_) ÷ w
w ← ind_f somme w_

        e ← ind_f somme w_ × (X_ - (⍴X_) ind_d matrix X
w) ⋆ 2
        e ← ww0 × (e - s2 × (ind_f somme w_ > 0) - ww0)
 ÷ ww

        E ← E, ((⍳⍴ind_d) somme e) ÷ +/,ww0

```

```
         ⍝ Calcul des coefficients des b^(i)
         :FOR j :IN ⍳nb_est
             :IF ⍬ ≡ Ib ← (⊃c[j]) ~ ind_d
                 S ← S, 0
             :ELSE
                 t ← P ⍴ 0 ⋄ t[Ib] ← 1
                 s ← (ind_f ~ Ib) somme t ↓ w_
                 s ← ww0 × (((sort ind_d, Ib) ⍳ Ib) somme s ⋆ 2) ÷ ww ⋆ 2
                 f ← ww0 × (ind_f somme t ↓ w_) ÷ ww

                 S ← S, ((⍳⍴ind_d) somme (ind_d ∊ ⊃c[j]) ↓ f - s) ÷ +/,ww0
             :ENDIF
         :ENDFOR
     :ENDFOR
 :ENDFOR

 ⍝ Résolution du système d'équations linéaires

 z ← E ⌹ (2 ⍴ nb_est) ⍴ S
∇
```

calculbtG

```
∇ b ← b calculbtG Xw;⎕io;X;X_;X∆;X_∆;w_;X12_;X12_∆;X13_;X13_∆;X23_;X23_∆;zeta123;zeta123_;zeta12;zeta12_;zeta13;zeta13_;zeta23;zeta23_;zeta123∆;zeta123_∆;zeta12∆;zeta12_∆;zeta13∆;zeta13_∆;zeta23∆;zeta23_∆;zeta∆;Z123;Z12;Z13;Z23;Z1;Z2;Z3;Z;K1;K2;K3;K12;K13;K23;S12;S13;S23;S12∆;S13∆;S23∆;S∆;S_∆;MS;I;J;K;i;j;bt;t;t12;t13;t23;t123;d123;d12;d13;d23;d1;d2;d3;N;c;Nb_iter

 ⍝ Calcule les estimateurs (itératifs) «ad hoc» des
 ⍝ composants de variance dans un modèle CCCG à deux
 ⍝ ou trois facteurs.
 ⍝
 ⍝ Xw[1] == ratios Xijt ou Xijkt
 ⍝ Xw[2] == poids wijt ou wijkt
 ⍝     b == composants de variance (valeurs de départ)
 ⍝
 ⍝ La valeur de s2 est prise en variable globale.
 ⍝
 ⍝ Fonction NON ADAPTÉE aux poids nuls (sauf pour la première ligne)!

 ⎕io ← 1

 X_ ← (+/(2 ⊃ Xw) × 1 ⊃ Xw) ÷ w_ ← +/2 ⊃ Xw

 Nb_iter ← 0

 ⍝ Début du calcul itératif
```

```
bt ← b

    ⍝ On n'admet pas de valeurs de départ négatives
    bt[(bt ≤ 0)/⍳ρbt] ← 0.1


⍝ Cas à deux facteurs
:IF 2 = ρρX_

    X_∆ ← 1 1 ↓ X_
    I ← (ρX_∆)[1] ⋄ J ← (ρX_∆)[2]
    d12 ← (¯1 - I, J) ↑ (I, J) ρ 1
    d1 ← (¯1 - I) ↑ I ρ 1 ⋄ d2 ← (¯1 - J) ↑ J ρ 1
    MS ← 2 ρ 0

    :REPEAT
        Nb_iter ← Nb_iter + 1
        b ← bt

            ⍝ Calcul des crédibilités de niveau (12)
            zeta12∆ ← 1 1 ↓ zeta12 ← t ÷ s2 + d12 × t ← bt[3] × w_

        ⍝ Calcul de MSABz ≡ b12 (le signe moins devant
        ⍝ c3 est intégré dans son calcul, ci-dessous)
        t12 ← 2 ρ 0
        :FOR i :IN ⍳I
            t12[1] ← t12[1] + +/+/ (zeta12∆[i;] ∘.× zeta12∆[i;]) × (X_∆[i;] ∘.- X_∆[i;]) ⋆ 2
        :END
        :FOR i :IN ⍳J
            t12[2] ← t12[2] + +/+/ (zeta12∆[;i] ∘.× zeta12∆[;i]) × (X_∆[;i] ∘.- X_∆[;i]) ⋆ 2
        :END

        Z12 ← +/+/zeta12∆ ⋆ 2 ⋄ Z ← (+/+/zeta12∆) ⋆ 2
        Z1 ← +/(+/zeta12∆) ⋆ 2 ⋄  Z2 ← +/(+⌿zeta12∆) ⋆ 2
        c ← (Z1 - Z12), 0, (Z - Z2)
        c ← c, 0, (Z2 - Z12), (Z - Z1)
        c ← c, 2 × (+/+/zeta12∆) × (J - 1), (I - 1), ¯1 + I × J
        c ← 0 0 1 ⌹ 3 3 ρ c

        bt[3] ← +/c × t12, +/+/+/+/ (zeta12∆ ∘.× zeta12∆) × (X_∆ ∘.- X_∆) ⋆ 2

            ⍝ Mise à jour des crédibilités de niveau (12)
            zeta12 ← t ÷ s2 + d12 × t ← bt[3] × w_

        ⍝ Calcul de MSAz et K1
        zeta12_∆ ← 1 ↓ zeta12_ ← +/[2] zeta12
        zeta∆ ← 1 ↓ ÷ d1 + bt[3] ÷ bt[1] × zeta12_
```

```
[68]          N ← (I - 1) × +/zeta∆
[69]          zeta∆ ← zeta∆ ∘.× zeta∆
[70]          K1 ← (+/+/zeta∆ × (⍉(I, I) ⍴ (+/[2] zeta12∆ ⋆ 2
     ) ÷ zeta12_∆ ⋆ 2) - (zeta12∆ +.× ⍉zeta12∆) ÷ zeta12_∆ ∘
     .× zeta12_∆) ÷ N
[71]          X∆ ← 1 ↓ (+/[2] zeta12 × X_) ÷ zeta12_
[72]          MS[1] ← (+/+/zeta∆ × (X∆ ∘.- X∆) ⋆ 2) ÷ 2 × N
[73]
[74]          ⍝ Calcul de MSBz et K2
[75]          zeta12_∆ ← 1 ↓ zeta12_ ← +/[1] zeta12
[76]          zeta∆ ← 1 ↓ ÷ d2 + bt[3] ÷ bt[2] × zeta12_
[77]          N ← (J - 1) × +/zeta∆
[78]          zeta∆ ← zeta∆ ∘.× zeta∆
[79]          K2 ← (+/+/zeta∆ × (⍉(J, J) ⍴ (+/[1] zeta12∆ ⋆ 2
     ) ÷ zeta12_∆ ⋆ 2) - ((⍉zeta12∆) +.× zeta12∆) ÷ zeta12_∆
      ∘.× zeta12_∆) ÷ N
[80]          X∆ ← 1 ↓ (+/[1] zeta12 × X_) ÷ zeta12_
[81]          MS[2] ← (+/+/zeta∆ × (X∆ ∘.- X∆) ⋆ 2) ÷ 2 × N
[82]
[83]          ⍝ Calcul des estimateurs
[84]          bt[1 2] ← MS ⌹ 2 2 ⍴ 1 K1 K2 1
[85]
[86]      :UNTIL (TOL > ⌈/|bt - b) ∨ (ZERO > ⌊/bt) ∨ MAX_ITER
      < Nb_iter

[...] Cas à trois facteurs

[287] b ← bt, Nb_iter > MAX_ITER
    ∇
```

calculboG

```
    ∇ b ← b calculboG Xw;⎕io;I;J;K;Xw∆;X∆;X_;X_∆;X12;X12∆;X13
     ;X13∆;w_;zeta123;zeta12;zeta13;S∆;S123;S123∆;S12;S12∆;S
     13;S13∆;t;bt;B;N;d123;d12;d13;Nb_iter
[1]
[2]  ⍝ Calcule les estimateurs (itératifs) optimaux des
[3]  ⍝ composants de variance dans un modèle CCCG à deux
[4]  ⍝ ou trois facteurs.
[5]  ⍝
[6]  ⍝ Les ∆ identifient les variables où la première ligne
[7]  ⍝ et la première colonne des données sont enlevées.
[8]  ⍝
[9]  ⍝ Xw[1] == ratios Xijt ou Xijkt
[10] ⍝ Xw[2] == poids wijt ou wijkt
[11] ⍝     b == composants de variance (valeurs de départ)
[12] ⍝
[13] ⍝ La valeur de s2 est prise en variable globale.
[14] ⍝
[15] ⍝ Fonction NON ADAPTÉE aux poids nuls (sauf pour la
[16] ⍝ première ligne)!
[17]
[18] ⎕io ← 0
```

```

X_ ← (+/(1 ⊃ Xw) × 0 ⊃ Xw) ÷ w_ ← +/1 ⊃ Xw

Nb_iter ← 0

⍝ Début du calcul itératif
bt ← b

    ⍝ On n'admet pas de valeurs de départ négatives
    bt[(bt ≤ 0)/⍳⍴bt] ← 0.1


⍝ Cas à deux facteurs
:IF 2 = ⍴⍴X_

    X_∆ ← 1 1 ↓ X_
    N ← ×/⍴X_∆ ◇ I ← (⍴X_∆)[0] ◇ J ← (⍴X_∆)[1]
    Xw∆ ← (,X_∆ ∘.- X_∆)[∊1 + ((N + 1) × ⍳N) + ⍳¨⌽⍳N]
    d12 ← (¯1 - I, J) ↑ (I, J) ⍴ 1

    :REPEAT
        Nb_iter ← Nb_iter + 1
        b ← bt

        ⍝ Calcul de b12
        zeta12 ← t ÷ s2 + d12 × t ← bt[2] × w_
        B ← construitB S12∆ ← 1 1 1 1 ↓ S12 ← 1 2 CovG bt zeta12
        bt[2] ← +/(b[2] × t ÷ +/B × t ← B +.× construitCinv S12∆) × Xw∆ ⋆ 2

            ⍝ Mises à jour
            zeta12 ← t ÷ s2 + d12 × t ← bt[2] × w_
            S12∆ ← 1 1 1 1 ↓ S12 ← 1 2 CovG bt zeta12

        ⍝ Calcul de b1
        X∆ ← 1 ↓ (+/zeta12 × X_) ÷ +/zeta12
        X∆ ← (,X∆ ∘.- X∆)[∊1 + ((1 + I) × ⍳I) + ⍳¨⌽⍳I]
        B ← construitB S∆ ← 1 1 ↓ 1 CovG S12 zeta12
        bt[0] ← +/(bt[0] × t ÷ +/B × t ← B +.× construitCinv S∆) × X∆ ⋆ 2

        ⍝ Calcul de b2
        X∆ ← 1 ↓ (+⌿zeta12 × X_) ÷ +⌿zeta12
        X∆ ← (,X∆ ∘.- X∆)[∊1 + ((1 + J) × ⍳J) + ⍳¨⌽⍳J]
        B ← construitB S∆ ← 1 1 ↓ 2 CovG S12 zeta12
        bt[1] ← +/(bt[1] × t ÷ +/B × t ← B +.× construitCinv S∆) × X∆ ⋆ 2

    :UNTIL (TOL > ⌈/|bt - b) ∨ (ZERO > ⌊/bt) ∨ MAX_ITER < Nb_iter

```

```
[...]  Cas à trois facteurs

[138] b ← b, Nb_iter > MAX_ITER
    ∇
```

construitB

```
    ∇ z ← construitB S;⎕io;dimS;M;N;P;t
[1]
[2]   ⍝ Construit le vecteur B pour l'estimation optimale
[3]   ⍝ des composants de variance dans un modèle CCC ou
[4]   ⍝ CCCG.
[5]   ⍝
[6]   ⍝ S == matrice des covariances
[7]
[8]   ⎕io ← 0
[9]
[10]  P ← .5 × ⍴⍴S ⋄ N ← ×/P ↑ dimS ← ⍴S
[11]
[12]  S ← ,S
[13]
[14]  M ← dimS ⊤ t ← ∊1 + ((N + 1) × ⍳N) + ⍳¨⌽⍳N
[15]
[16]  z ← S[dimS ⊥ M[t;]] + S[dimS ⊥ M[(P + t ← t, t ← ⍳P);]]
       - 2 × S[t]
    ∇
```

construitCinv

```
    ∇ z ← construitCinv S;⎕io;N;dimC;P;PP;i;j;M;t;v;A11;A21;L
      1;L2;L3;L4
[1]
[2]   ⍝ Construit la matrice inverse de la matrice C pour
[3]   ⍝ l'estimation optimale des composants de variance
[4]   ⍝ dans un modèle CCC ou CCCG.
[5]   ⍝
[6]   ⍝ Utilise l'algorithme d'inversion d'une matrice par
[7]   ⍝ blocs
[8]   ⍝
[9]   ⍝ S == matrice des covariances
[10]
[11]  ⎕io ← 0
[12]
[13]  P ← 0.5 × PP ← ⍴dimC ← ⍴S ⋄ dimC ← dimC, dimC
[14]  N ← ×/P ↑ dimC
[15]
[16]  ⍝ Combinaisons à calculer
[17]  v ← ∊1 + ((N + 1) × ⍳N) + ⍳¨⌽⍳N
[18]
[19]  N ← N × N
[20]
```

```
⍝ Lignes à sélectionner pour obtenir les covariances
L1 ← (⍳P), PP + ⍳P
L2 ← (⍳P), PP + P + ⍳P
L3 ← P + ⍳PP
L4 ← P + L1


⍝ Calcul récursif des éléments de la matrice inverse
z ← 0 0 ⍴ 0 ⋄ j ← 0
:FOR i :IN ⌽⍳⍴v
    M ← dimC ⊤ v[i] + N × (j ← j - 1) ↑ v
    A21 ← ∊(-/(M[L1;] extraitM S) (M[L2;] extraitM S) (
M[L4;] extraitM S) (M[L3;] extraitM S)) ⋆ 2
    A11 ← 1 ↑ A21 ⋄ A21 ← 1 ↓ A21
    A11 ← ÷A11 - +/A21 × t ← z +.× A21
    z ← (A11, t)⍪ t, z - (t ← -A11 × t) ∘.× t
:END
∇
```

CovG

```
∇ z ← n CovG Szeta;⎕io;S;zeta;zeta_;dim;dim_;t;t1;t2;t3;t
1_;t2_;t3_

⍝ Fonction à deux vocations:
⍝ 1. si ⍴n = nombre de dimensions de Sz[2], calcule la
⍝    covariance entre deux moyennes pondérées dans un
⍝    modèle CCCG à deux ou trois facteurs; dans ce cas
⍝
⍝    Szeta[1] == composants de variance
⍝    Szeta[2] == fonctions zeta du dernier niveau
⍝
⍝ 2. si ⍴n < nombre de dimensions de Sz[2], calcule la
⍝    covariance entre deux moyennes pondérées par les
⍝    fonctions zeta dans un modèle CCCG; dans ce cas
⍝
⍝    Szeta[1] == covariances du niveau inférieur
⍝    Szeta[2] == fonctions zeta du niveau inférieur
⍝
⍝ Dans tous les cas
⍝
⍝ n == niveau de covariance souhaité

⎕io ← 1

S ← 1 ⊃ Szeta ⋄ zeta ← 2 ⊃ Szeta
dim ← ⍴zeta

:IF (⍴,n) = ⍴dim
    :IF 2 = ⍴n
        dim_ ← dim - 1 0
        z ← (S[1] × (-dim, dim) ↑ (-dim_, dim_) ↑ t1 ←
t ∘.= t ← ⍉(⌽dim) ⍴ ⍳dim[1])
```

```
[30]          dim_ ← dim - 0 1
[31]          z ← z + (S[2] × (-dim, dim) ↑ (-dim_, dim_) ↑ t
     2 ← t ∘.= t ← dim ⍴ ⍳dim[2])
[32]          z ← z + (dim, dim) ⍴ (,t1 ∧ t2) \ S[3] × ,t × (
     t ← 0 < zeta) ÷ zeta
[33]       :ELSEIF 3 = ⍴n

[...]         Cas à trois facteurs

[44]       :END
[45]   :ELSE
[46]       z ← ((n, n + ⍴dim) somme S × zeta ∘.× zeta) ÷ zeta_
      ∘.× zeta_ ← +/[n ← (⍳⍴dim) ~ n] zeta
[47]   :END
    ∇
```

calculs2

```
    ∇ z ← X calculs2 w;⎕io;dim;P;w_;X_
[1]
[2]    ⍝ Calcule l'estimateur optimal de s2 dans un modèle
[3]    ⍝ CCC ou CCCG à P facteurs
[4]
[5]    ⎕io ← 1
[6]
[7]    P ← ¯1 + ⍴dim ← ⍴X
[8]    X_ ← (w_ > 0) × (+/w × X) ÷ w_ ← +/w
[9]    z ← ((⍳P+1) somme w × (X - dim (⍳P) matrix X_) ⋆ 2) ÷ (
       +/,w > 0) - +/,w_ > 0
    ∇
```

calculXw

```
    ∇ z ← X calculXw w
[1]
[2]    ⍝ Calcule la moyenne pondérée par les poids naturels
[3]
[4]    z ← ((⍳⍴⍴X) somme w × X) ÷ (⍳⍴⍴X) somme w
    ∇
```

calculXzw

```
    ∇ z ← X calculXzw Z;P;i;t
[1]
[2]    ⍝ Calcule l'estimateur «ad hoc» de la moyenne
[3]    ⍝ collective dans un modèle CCC ou CCCG, soit la
[4]    ⍝ moyenne pondérée par les facteurs de crédibilité ou
[5]    ⍝ les fonctions zeta en partant du dernier niveau
[6]    ⍝ d'interaction en montant vers le niveau 1.
[7]    ⍝
[8]    ⍝ Z[1] == facteurs de crédibilité ou fonctions zeta du
```

```
⍝         niveau 1
⍝ ...
⍝ Z[P] == facteurs de crédibilité ou fonctions zeta du
⍝         dernier niveau d'interaction
⍝    X == moyennes pondérées par les poids naturels
P ← ⍴Z
z ← X
:FOR i :IN ⌽⍳P
    z ← (t > 0) × (+/z × i ⊃ Z) ÷ t ← +/i ⊃ Z
:END
∇
```

calculmoG

```
∇ z ← X calculmoG bzeta;⎕io;b;zeta;C;I;J;K;IJ;JK;i;j;iJK;
  iJ;bU;bI;Id;U;t;d_if;d_jg
⍝ Calcule l'estimateur optimal de la moyenne
⍝ dans un modèle CCCG à deux ou trois facteurs.
⍝
⍝ bzeta[0] == composants de variance
⍝ bzeta[1] == fonctions zeta
⍝        X == moyennes pondérées par assuré
⎕io ← 0
X ← ,X ⋄ b ← 0 ⊃ bzeta ⋄ zeta ← 1 ⊃ bzeta
⍝ Cas à deux facteurs
:IF 2 = ⍴⍴zeta
    I ← 0 ⊃ ⍴zeta ⋄ J ← 1 ⊃ ⍴zeta
    Id ← Ident J
    bU ← (J, J) ⍴ b[0]
    bI ← I ⍴ ⊂b[1] × (-J, J) ↑ Ident J - 1
    C ← (2 ⍴ I × J) ⍴ 0
    :FOR i :IN ⍳I
        C[(i × J) + ⍳J;] ← ⊃,/bI + (-i) ⌽ I ↑ ⊂(bU × i
≠ 0) + (b[2] × (zeta[i;] > 0) ÷ zeta[i;]) put_diag Id
    :END
    ⍝ Si tous les poids de la catégorie (0,0)
    ⍝ sont nuls
    :IF 0 = 1 ↑ ,zeta
        X ← 1 ↓ X
        C ← 1 1 ↓ C
    :END
```

```
[34]      z ← +/X × t ÷ +/t ← +/⍟C
[35]

[...]  Cas à 3 facteurs

    ∇
```

gen_matrice

```
    ∇ z ← c gen_matrice m;⎕io;cred;poids_z;P;dim;dimI;I;i;J;i
      nd;cI;cJ;indsomme;poids_I;poids_Ii;t;t2;k
[1]
[2]   ⍝ Construit la matrice d'estimation des Xi pour un
[3]   ⍝ modèle CCCG
[4]   ⍝
[5]   ⍝   m[⎕io] == facteurs de crédibilité
[6]   ⍝ m[⎕io+1] == poids (fonctions zeta, en fait)
[7]   ⍝        c == combinaisons de tous les facteurs de
[8]   ⍝             risque
[9]
[10]  ⎕io ← 0
[11]
[12]  cred ← ⊃m[0] ⋄ poids_z ← ⊃m[1]
[13]  P ← ⍴dim ← ⍴ (¯1 + (⍴c)[0]) ⊃ cred
[14]
[15]  z ← ⍬
[16]  :FOR I :IN ⍳¯1 + ⍴c
[17]      cI ← ⊃c[I]
[18]      dimI ← dim[¯1 + cI]
[19]      ind ← P gen_ind cI
[20]      indsomme ← (1 + ⍳P) ~ cI
[21]      poids_I ← ⊃poids_z[I]
[22]
[23]      :FOR i :IN ⍳×/dimI
[24]          t ← dim ⋄ t2 ← P ⍴ 0 ⋄ k ← dimI ⊤ i
[25]          t[¯1 + cI] ← 1 + k ⋄ t2[¯1 + cI] ← k
[26]          poids_Ii ← t2 ↓ t ↑ poids_I
[27]
[28]          :FOR J :IN ⍳¯1 + ⍴c
[29]              cJ ← ⊃c[J]
[30]              :IF c[J] ∊ ind
[31]                  z ← z, (×/dim[¯1 + cJ]) ⍴ 0
[32]              :ELSE
[33]                  t ← (,⊃cred[I])[i] × ,(indsomme ~ cJ) s
      omme poids_Ii
[34]                  z ← z, cI cJ i dim fill_0 t
[35]              :ENDIF
[36]          :ENDFOR
[37]      :ENDFOR
[38]  :ENDFOR
[39]
[40]  z ← 1 put_diag (2 ⍴ (⍴z) * .5) ⍴ z
    ∇
```

extrait_Xi

```
∇ z ← v extrait_Xi Xi;⎕io;pos;dim;c;I;cI

⍝ Lors du calcul final de la prime de crédibilité
⍝ dans le modèle CCCG, cette fonction sélectionne
⍝ les Xi appropriés pour chaque assuré.
⍝ Travaille par "catégorie" de Xi.

⎕io ← 0

pos ← ⊃v[0] ⋄ dim ← ⊃v[1] ⋄ c ← ⊃v[2]
z ← ⍬

:FOR I :IN ⍳¯1 + ⍴c
    z ← z, Xi[dim[cI] ⊥ pos[cI ← ¯1 + ⊃c[I]]]
    Xi ← (×/dim[cI]) ↓ Xi
:ENDFOR
∇
```

Bibliographie

Bailey, A. L. (1945), A generalized theory of credibility, *Proceedings of the Casualty Actuarial Society* **32**, 13–20.

Bailey, A. L. (1950), Credibility procedures, Laplace's generalization of Bayes' rule and the combination of collateral knowledge with observed data, *Proceedings of the Casualty Actuarial Society* **37**, 7–23.

Bühlmann, H. (1967), Experience rating and credibility, *ASTIN Bulletin* **4**, 199–207.

Bühlmann, H. (1969), Experience rating and credibility, *ASTIN Bulletin* **5**, 157–165.

Bühlmann, H. & Jewell, W. S. (1987), Hierarchical credibility revisited, *Bulletin de l'Association suisse des actuaires* **87**, 35–54.

Bühlmann, H. & Straub, E. (1970), Glaubwürdigkeit für Schadensätze, *Bulletin de l'Association suisse des actuaires* **70**, 111–133. Traduction anglaise par C.E. Brooks.

Burden, R. L. & Faires, J. D. (1988), *Numerical analysis*, quatrième édition, PWS-Kent, Boston.

Cossette, H. (1996), Dependent contracts in credibility models and parameter estimation, Thèse de doctorat, Université Catholique de Louvain.

Dannenburg, D. (1995), Crossed classification credibility models, *Comptes rendus du 25^e^ congrès International d'actuaires* **4**, 1–35.

Dannenburg, D. R., Kaas, R. & Goovaerts, M. J. (1996), *Practical actuarial credibility models*, Institute of Actuarial Science and Econometrics, University of Amsterdam.

De Vylder, F. (1976*a*), Geometrical credibility, *Scandinavian Actuarial Journal* **1976**, 121–149.

De Vylder, F. (1976*b*), Optimal semilinear credibility, *Bulletin of the Swiss Association of Actuaries* **76**, 27–40.

De Vylder, F. (1978), Parameter estimation in credibility theory, *ASTIN Bulletin* **10**, 99–112.

De Vylder, F. (1981), Practical credibility theory with emphasis on parameter estimation, *ASTIN Bulletin* **12**, 115–131.

De Vylder, F. & Goovaerts, M. J. (1991), Estimation of the heterogeneity parameter in the Bühlmann–Straub credibility theory model, *Insurance : Mathematics and Economics* **10**, 233–238.

De Vylder, F. & Goovaerts, M. J. (1992*a*), Optimal parameter estimation under zero-excess assumptions in a classical model, *Insurance : Mathematics and Economics* **11**, 1–6.

De Vylder, F. & Goovaerts, M. J. (1992*b*), Optimal parameter estimation under zero-excess assumptions in the Bühlmann–Straub model, *Insurance : Mathematics and Economics* **11**, 167–171.

Dubey, A. & Gisler, A. (1981), On parameter estimation in credibility, *Bulletin de l'Association suisse des actuaires* **81**, 187–211.

Gisler, A. & Reinhard, P. (1993), Robust credibility, *ASTIN Bulletin* **23**, 117–143.

Goovaerts, M. J. & Hoogstad, W. J. (1987), *Credibility theory, Surveys of actuarial studies, no. 4*, Nationale-Nederlanden N.V., Pays-Bas.

Goovaerts, M. J., Kaas, R., van Heerwaarden, A. E. & Bauwelinckx, T. (1990), *Effective actuarial methods*, North-Holland, Amsterdam.

Goulet, V. (1994), Théorie de la crédibilité : histoire, principes et applications, Mémoire de maîtrise, Université Laval.

Goulet, V. (1997), Crédibilité à classification croisée avec nombre de facteurs variable. Cahier de recherche 97-09, Institut de sciences actuarielles, Université de Lausanne.

Goulet, V. (1998), A note on optimal parameter estimation under zero-excess assumptions, à paraître dans *Insurance : Mathematics and Economics*.

Graybill, F. A. & Hultquist, R. A. (1961), Theorems concerning Eisenhart's model II, *Annals of Mathematical Statistics* **32**, 261–269.

Graybill, R. A. (1954), On quadratic estimates of variance components, *Annals of Mathematical Statistics* **25**, 367–372.

Hachemeister, C. A. (1975), Credibility for regression models with application to trend, *dans* Credibility, theory and applications, Proceedings of the Berkeley actuarial research conference on credibility, Academic Press, New York, pp. 129–163.

Healy, M. J. R. (1986), *Matrices for statistics*, Oxford University Press, Oxford.

Jewell, W. S. (1975), The use of collateral data in credibility theory : a hierarchical model, *Giornale dell'Istituto Italiano degli Attuari* **38**, 1–16.

Johnson, N. L., Kotz, S. & Kemp, A. W. (1992), *Univariate discrete distributions*, seconde édition, Wiley, New York.

Kreyszig, E. (1989), *Introductory functional analysis with applications*, Wiley, New York.

Künsch, H. R. (1992), Robust methods for credibility, *ASTIN Bulletin* **22**, 33–49.

Mayerson, A. L. (1964), A bayesian view of credibility, *Proceedings of the Casualty Actuarial Society* **51**, 85–104.

Mowbray, A. H. (1914), How extensive a payroll exposure is necessary to give a dependable pure premium ?, *Proceedings of the Casualty Actuarial Society* **1**, 25–30.

Norberg, R. (1979), The credibility approach to ratemaking, *Scandinavian Actuarial Journal* **1979**, 181–221.

Rao, C. R. (1965), *Linear statistical inference and its applications*, Wiley, New York.

Scheffé, H. (1959), *The analysis of variance*, Wiley, New York.

Searle, S. R. (1971), *Linear models*, Wiley, New York.

Searle, S. R., Casella, G. & McCulloch, C. E. (1992), *Variance Components*, Wiley, New York.

Whitney, A. W. (1918), The theory of experience rating, *Proceedings of the Casualty Actuarial Society* **4**, 275–293.