



Incertitude et mesure de performance

Une nouvelle approche à travers la théorie mathématique des indications

Thèse

présentée à la Faculté des sciences
économiques et sociales de l'Université de Fribourg (Suisse)

par

Elitza Ouzounova

de Fribourg / FR

pour l'obtention du grade de
docteur ès sciences économiques et sociales.

Acceptée par la Faculté des sciences économiques et sociales
le 18 octobre 2005, sur proposition de
Monsieur le Professeur H. W. Brachinger (premier rapporteur) et de
Monsieur le Professeur J. Kohlas (second rapporteur).

Fribourg 2005

La Faculté des sciences économiques et sociales de l'Université de Fribourg (Suisse) n'entend ni approuver, ni désapprouver les opinions émises dans une thèse : elles doivent être considérées comme propres à l'auteur (Décision du Conseil de Faculté du 23 janvier 1990).

A ma soeur, Théodora

Remerciements

Le présent travail est le fruit de recherches accomplies au cours de mon engagement à la Chaire de Statistique de l'Université de Fribourg (Suisse).

Je tiens à remercier tout particulièrement mon maître de thèse, le Professeur Dr Hans Wolfgang Brachinger, pour avoir accepté de diriger mon projet. Ses conseils avisés, tout au long de mes recherches, ont très largement contribué au succès de ce travail. Ma reconnaissance s'adresse également au Professeur Dr Jürg Kohlas, qui a assumé la tâche de second rapporteur, ainsi qu'à son collaborateur, le Dr Norbert Lehmann ; tous deux m'ont apporté une aide précieuse dans le domaine de la théorie mathématique des indications.

Je dois l'idée de ce projet au Dr Paul-André Monney, grâce à qui j'ai eu l'occasion de découvrir le raisonnement sous incertitude en général et la théorie mathématique des indications en particulier ; qu'il en soit chaleureusement remercié.

Je tiens également à remercier mes collègues de la Chaire de Statistique de l'Université de Fribourg et tout particulièrement Diane Karusisi et Michael Beer, pour leur soutien tant professionnel que personnel.

De tout coeur, je remercie Jérôme Levrat pour la relecture du texte, pour ses remarques constructives et pour le soutien inconditionnel dont il a fait preuve durant l'accomplissement de ce projet.

A mes parents, j'adresse enfin ma gratitude pour leurs encouragements et leur amour.

Fribourg, automne 2005

Elitza Ouzounova

Table des matières

Introduction	1
I Théorie mathématique de l'évidence	5
1 Origine et applications	7
1.1 Historique	7
1.2 Applications de la théorie mathématique de l'évidence	7
2 Notions de base	11
2.1 Probabilités objectives et subjectives	11
2.2 Distinction entre chance et crédibilité	13
2.3 Fonction de masse	14
2.4 Fonction de crédibilité	18
2.5 Fonction de plausibilité	22
2.6 Fonction de crédibilité bayésienne	24
3 Règle de combinaison de Dempster	27
3.1 Combinaison de deux fonctions de crédibilité	27
3.2 Combinaison de plusieurs fonctions de crédibilité	31
3.3 Poids du conflit	33
4 Conditionnement	35
5 Critiques	37

6	Problème du casino	39
II	Théorie mathématique des indications	41
7	Concepts de base	45
7.1	Imprécision, incertitude et contradiction	45
7.2	Interprétations et mesure de probabilité	46
7.3	Ensembles focaux	46
7.4	Indications	48
8	Support et plausibilité	51
8.1	Famille des propositions impliquées	51
8.2	Famille des propositions possibles	52
8.3	Fonctions de support et de plausibilité	54
8.4	Propriétés élémentaires	55
8.5	Fonction de masse	57
8.6	Indications et fonctions de crédibilité	59
9	Combinaison selon la règle de Dempster	61
9.1	Règle de combinaison de Dempster	61
9.2	Combinaison de deux indications	62
9.3	Combinaison de plusieurs indications	66
9.4	Cas spéciaux	67
10	Conditionnement des indications	71
III	Combinaison d'indications en chaîne	77
11	Problème du casino	79
12	Chaînes d'indications	85

12.1	Construction d'une chaîne d'indications	85
12.2	Indications de confirmation	89
12.3	Meta-indications	98
13	Règle de combinaison en chaîne	99
13.1	Combinaison d'une \mathcal{H}_1 avec une Con_2	99
13.2	Combinaison d'une \mathcal{H}_1 avec plusieurs Con_i	103
13.3	Combinaison en chaîne de plusieurs indications	105
14	Propriétés de la combinaison en chaîne	111
14.1	Commutativité	111
14.2	Associativité	112
14.3	Distributivité	113
14.4	Distributivité	115
15	Combinaison d'indications de confirmation	119
15.1	Contexte	119
15.2	Stratégie 1	120
15.3	Stratégie 2	123
15.4	Stratégie 3	127
15.5	Règle de Dempster et meta-indications	127
16	Adaptation de la règle de Dempster	129
16.1	Combinaison d'indications de confirmation	129
16.2	Combinaison d'indications de confirmation	131
IV	Nouvelle mesure de performance	135
17	Théorie Moderne du Portefeuille	139
17.1	Introduction	139

17.2	Relation rendement-risque	140
17.3	Mesures de performance dans la TMP	141
18	Théorie Post-Moderne du Portefeuille	143
18.1	Origine	143
18.2	Moments partiels inférieurs comme mesure du risque	145
18.3	Mesures de performance dans la TPMP	145
19	Nouvelle approche	147
19.1	Modèle de performance et incertitude	147
19.2	Nouvelle source de risque	155
19.3	Classement final	157
19.4	Classement définitif basé sur les degrés de support	158
19.5	Classement définitif basé sur les degrés de plausibilité	160
20	Exemple d'application	163
20.1	Classements d'après les mesures de performance de la TMP	164
20.2	Classements d'après la TPMP	165
20.3	Construction des meta-indications	167
	Conclusion	175

Table des figures

- 2.1 Connaissance et ambiguïté 12
- 11.1 Exemple de la roulette avec \mathcal{H}_1 80
- 11.2 Arbre de probabilité de \mathcal{H}_1 81
- 12.1 Exemple de la roulette avec \mathcal{H}_1 et \mathcal{H}_2 86
- 12.2 Arbre de probabilité de \mathcal{H}_1 et \mathcal{H}_2 87
- 12.3 Exemple de la roulette avec \mathcal{H}_1 , Con_2 et Con_3 92
- 12.4 Arbre de probabilité de \mathcal{H}_1 , Con_2 et Con_3 94
- 13.1 Combinaison en chaîne de Con_2 confirmant \mathcal{H}_1 100
- 13.2 Arbre de probabilité de Con_2 confirmant \mathcal{H}_1 102
- 15.1 Exemple de la roulette avec deux indications de confirmation 121
- 16.1 Combinaison de $(n - 1) Con_i$ confirmant toutes \mathcal{H}_1 132
- 19.1 Schéma graphique de la meta-indication \mathcal{M}'_{H_1} 152
- 19.2 Schéma graphique de la meta-indication \mathcal{M}''_{H_2} 154

Liste des tableaux

3.1	Masses des intersections $A_i \cap B_j$	28
3.2	Exemple : masses des intersections $A_i \cap B_j$	28
8.1	Exemple : interprétations impliquant H	52
8.2	Exemple : interprétations rendant H possible	53
8.3	Exemple : degrés de support	55
8.4	Exemple : degrés de plausibilité	55
8.5	Exemple : masses	59
9.1	Exemple : interprétations combinées	65
9.2	Exemple : probabilités jointes	65
9.3	Exemple : ensembles focaux joints	65
9.4	Exemple : degrés de support d'une indication combinée	66
9.5	Exemple : degrés de plausibilité d'une indication combinée	66
9.6	Exemple : masses d'une indication combinée	66
10.1	Exemple : masses conditionnelles avec contradiction	72
10.2	Exemple : degrés de support conditionnels avec contradiction	73
10.3	Exemple : degrés de plausibilité conditionnels	73
10.4	Exemple : masses conditionnelles sans contradiction	73
10.5	Exemple : degrés de support conditionnels sans contradiction	74
10.6	Exemple : degrés de plausibilité conditionnels sans contradiction	74
20.1	Paramètres estimés pour les 10 actions et le SMI	164

20.2	Indices de la TMP	165
20.3	Classement d'après les trois indices de la TMP	165
20.4	Paramètres estimés pour les 10 actions et le SMI	166
20.5	Indices de la TPMP	166
20.6	Classement d'après les trois indices de la TPMP	167
20.7	Paramètres estimés pour chaque action	168
20.8	Probabilités estimées de $p_3(\omega_{3,f})$ et $p_4(\omega_{4,f})$	169
20.9	Fonctions de probabilités de H_1 et H_2	169
20.10	Fonctions de probabilités (non ajustées) des 10 \mathcal{M}_P	169
20.11	Ensembles focaux des 10 \mathcal{M}_P	170
20.12	Fonctions de probabilités ajustées des 10 \mathcal{M}_P	171
20.13	Fonctions de masse des 10 \mathcal{M}_P	171
20.14	Degrés de support des 10 \mathcal{M}_P	172
20.15	Degrés de plausibilité des 10 \mathcal{M}_P	172
20.16	Indices de support et de plausibilité des 10 \mathcal{M}_P	172
20.17	Classement des 10 actions P d'après les indices de support	173
20.18	Classement des 10 actions P d'après les indices de plausibilité	174

Introduction

Raisonnement sous incertitude

Raisonner dans un contexte d'incertitude se révèle complexe et difficile, cependant, l'ambiguïté et le manque d'information complète sont des caractéristiques inhérentes au monde qui nous entoure. En effet, il s'agit d'un problème essentiel pour de nombreuses disciplines scientifiques, telles que la médecine, le droit ou encore l'économie.

Imaginons que l'on ait une décision à prendre. Bien évidemment, on désire que cette décision soit prise de manière rationnelle, sur la base de toutes les informations disponibles. Celles-ci peuvent cependant se présenter de différentes manières selon la situation. Il n'est pas à exclure qu'elles soient floues, ambiguës, ou encore incomplètes. Il est également nécessaire d'être en mesure de combiner toutes les informations entre elles en tenant compte même des informations contradictoires. Enfin, certaines informations peuvent se révéler pertinentes pour la décision à prendre, non pas directement, mais en confirmant ou en infirmant d'autres informations relatives, quant à elles, directement à la décision. Il serait de ce fait indispensable d'en tenir également compte. Comment intégrer toutes ces considérations dans un processus de prise de décision ?

Ainsi, la question essentielle que traitera ce travail est comment construire un raisonnement logique et rationnel lorsqu'on ne dispose que d'informations incomplètes ou ambiguës, d'informations confirmantes ou infirmantes d'autres informations et comment les combiner entre elles, tout en tenant compte des éventuelles contradictions.

La théorie mathématique de l'évidence

De nombreuses approches ont traité cette question. Parmi elles, la théorie mathématique de l'évidence de Dempster et Shafer [95], exposée dans la Partie 1 de ce travail, donne une réponse à ce problème. Cette théorie propose une distinction entre les probabilités fréquentistes, propres au monde qui nous entoure, et les probabilités subjectives, propres à notre jugement. De plus, elle permet de traiter des cas d'information partielle, voire inexistante. Ainsi, elle arrive à modéliser et à différencier des situations qui, jusque là, étaient traitées de manière identique. En effet, dans la théorie classique des probabilités, comme on le verra en début de Partie 1, les cas d'information partielle ou inexistante sont modélisés comme ceux d'information complète ; or, la différence est évidente. Par ailleurs,

la règle de combinaison de Dempster, que l'on trouve dans la théorie mathématique de l'évidence, permet de combiner différentes informations, actualisant ainsi à tout moment les données de départ. Il s'agit d'une alternative au théorème de Bayes, qui parvient à modéliser des combinaisons de sources d'information que la théorie des probabilités ne pouvait évaluer. Enfin, on mentionnera quelques applications de la théorie mathématique de l'évidence à des problèmes économiques, ainsi que les critiques formulées à son encontre.

La théorie mathématique des indications

La théorie mathématique de l'évidence de Dempster et Shafer a été à l'origine de plusieurs nouvelles théories, telle la théorie mathématique des indications de Kohlas et Monney [62], présentée dans la Partie 2. D'autres théories, également basées sur la théorie mathématique de l'évidence, comme celle de Smets [106], ont aussi tenté de répondre au problème du raisonnement sous incertitude, cependant elles ne seront pas abordées dans ce travail. La théorie mathématique des indications reprend en grande partie les concepts de base de la théorie mathématique de l'évidence. Elle développe le concept d'indication comme modèle explicite d'une masse d'information incertaine et imprécise. Elle dépasse la théorie de Dempster et Shafer en l'étendant à des cadres de discernement infinis. Elle réussit enfin à éliminer le jugement personnel, en distinguant le concept de crédibilité de Dempster et Shafer du concept de support, introduit par Kohlas et Monney. Ainsi, dans la théorie mathématique des indications, il ne s'agit plus d'évaluer soi-même le degré de crédibilité que l'on accorde à une certaine hypothèse, mais d'évaluer les arguments en faveur ou contre cette hypothèse, arguments fournis par les différentes interprétations possibles d'une indication. Enfin, cette théorie généralise la règle de combinaison de Dempster à la combinaison d'informations stochastiquement dépendantes ou indépendantes, présentant ainsi une solution à une des critiques majeures de la théorie mathématique de l'évidence.

Les indications de confirmation et les meta-indications

On utilisera la théorie mathématique des indications comme point de départ de tous les développements de la Partie 3. On y présentera deux nouvelles opérations de combinaison des indications, ainsi que deux nouveaux concepts d'indications : les indications de confirmations et les meta-indications. Il s'agit de résultats absents jusqu'à présent dans toutes les théories basées sur la théorie mathématique de l'évidence, aussi bien de la théorie mathématique des indications que de celle de Smets par exemple. L'importance de ces résultats est illustrée par des exemples simples, mais néanmoins réalistes. Ces résultats permettent en effet de modéliser certaines situations courantes, chose impossible auparavant. Il s'agit de toutes les situations où une partie de l'information dont on dispose est mise en doute. Les théories mathématiques de l'évidence et des indications traitaient déjà des situations d'information partielle ou floue ; cependant, elles n'arrivaient pas à modéliser les remises en question de cette information. On imagine aisément une situation où cela serait pourtant nécessaire.

Supposons le client d'un casino qui aimerait jouer à la roulette et qui aurait le choix entre dix tables. On peut facilement estimer ses chances de gain et de perte. Cela dit, supposons en plus qu'une des dix tables soit truquée. Nous sommes clairement dans une situation d'information complète. Nous connaissons en effet la loi de probabilité, d'habitude uniforme, que suivent les numéros gagnants, pourtant, une information supplémentaire vient perturber cet état de fait. Elle remet en question, avec une certaine probabilité, cette distribution uniforme. En ce sens, un parallèle sera fait dans la Partie 3 avec le concept de «probabilités de second ordre».

Une telle remise en question peut également être imaginée dans des situations d'information partielle. Supposons, parmi dix tables de roulette, une table dont les numéros gagnants ne suivent pas une distribution uniforme, contrairement aux neuf autres tables, mais dont on sait uniquement que le numéro 13 est favorisé de manière à sortir une fois sur dix. N'ayant aucune information quant aux probabilités des autres numéros de sortir, nous sommes obligés de faire appel à la théorie mathématique des indications afin de modéliser cette situation. Cela dit, si l'on ignore laquelle des dix tables est celle qui favorise le numéro 13, on est alors obligé de passer par les concepts d'indication de confirmation, de meta-indication et par les nouvelles règles de combinaison pour réussir à évaluer cette situation.

Vers une nouvelle mesure de performance

La Partie 4 exposera, enfin, un nouveau modèle de mesure de performance des actifs financiers, basé sur la théorie mathématique des indications et sur les concepts développés dans la Partie 3. Ce modèle introduit un nouveau concept de risque, le risque d'incertitude, complémentaire aux risques déjà connus des financiers. Celui-ci représente le risque que les rendements des actifs ne remplissent pas certaines hypothèses à la base des mesures de performance utilisées. Concrètement, nous nous sommes basés sur les six mesures de performance les plus employées par les financiers : le ratio de Sharpe, l'indice de Treynor, l'alpha de Jensen, tous étant basés sur la Théorie Moderne du Portefeuille [75] et leurs trois correspondants de la Théorie Post-Moderne du Portefeuille [8], [30], [10]. Notre modèle combine toute l'information contenue dans ces six mesures de performance, tout en y intégrant le risque d'incertitude en deux nouveaux indices de performance permettant un classement des actifs. Finalement, on présentera dans la Partie 4 une application concrète de ce modèle à dix actions cotées en bourse.

Première partie

Théorie mathématique de l'évidence

Chapitre 1

Origine et applications

*On mesure l'intelligence
d'un individu à la quantité
d'incertitudes qu'il est
capable de supporter.*

EMMANUEL KANT (1724-1804)

1.1 Historique

Les concepts de base de la théorie de l'évidence trouvent leur origine au XVII^e siècle déjà, dans les travaux du mathématicien suisse Jacob Bernoulli (1654-1705), puis dans ceux de Johann Heinrich Lambert (1728-1777). Cependant, ce n'est qu'en 1967 qu'Arthur Dempster posa les fondements de ce que l'on appelle **la théorie de l'évidence**, ou encore **la théorie de Dempster et Shafer** [21], Shafer ayant été le collaborateur de Dempster. Une dizaine d'années plus tard, la théorie mathématique de l'évidence vit le jour par la publication de Glenn Shafer en 1976 [95]. Celui-ci reconnut en 1979 [98] que les travaux de Bernoulli et Lambert peuvent être considérés comme précurseurs de sa théorie de l'évidence.

1.2 Applications de la théorie mathématique de l'évidence

La théorie de Dempster et Shafer a connu de nombreuses applications dans différents domaines de l'intelligence artificielle, tels que les diagnostics médicaux, la localisation de sous-marins, la classification d'images-radar, etc.

Nous nous sommes intéressés plus particulièrement aux applications économiques de cette théorie. On en trouve notamment des applications dans le domaine de l'audit, de l'analyse financière ou encore de l'économie politique.

Dans ces différentes applications économiques de la théorie mathématique de l'évidence, de nombreux chercheurs ont relevé des avantages considérables, tant au niveau de la modélisation de situations d'incertitude qu'au niveau de la possibilité d'intégrer facilement des informations nouvelles ou de supprimer des informations qui ne sont plus pertinentes. Ainsi, Monroe et Ng [79], Jung et Fink [47], ou encore Harrison, Srivastava et Plumlee [40], qui ont appliqué cette théorie à l'audit, ont relevé les avantages suivants :

- possibilité de chercher de l'évidence additionnelle ou de déduire de l'évidence existante ;
- approche la mieux appropriée pour la conduite d'audits financiers ;
- faible risque d'erreur d'acceptation incorrecte d'une hypothèse ;
- meilleure représentation de l'incertitude contenue dans l'évidence de l'audit ;
- représentation de l'ignorance comme composante explicite séparée de l'évaluation ;
- robustesse par rapport aux erreurs dans l'interprétation de l'information fournie par les procédures analytiques.

Srivastava et Datta [111], Jirousek, Vejnarova et Gemela [46], ainsi que Shenoy et Shenoy, [104] font partie des auteurs qui ont appliqué la théorie mathématique de l'évidence à des problèmes d'analyse financière. Ils estiment que cette théorie se prête très bien à ce type d'applications et ceci pour plusieurs raisons :

- implications tant au niveau de la théorie que de la gestion des prises de décision stratégiques ;
- possibilité pour les gestionnaires d'incorporer les jugements incertains associés à de nombreux facteurs d'une manière intuitive ;
- possibilité de chercher de l'évidence additionnelle ou de déduire de l'évidence existante ;
- meilleure compréhension du processus de prise de décision sous incertitude ;
- description des dépendances parmi les variables basée sur la connaissance des relations entre ces variables ;
- intégration facile de toute nouvelle information ou évidence pour la mise à jour des prédictions, grâce à la règle de combinaison de Dempster.

Enfin, McBurney et Parsons [77], ainsi que Lien [71], se sont basés sur la théorie mathématique de l'évidence pour traiter de sujets classiques de l'économie politique, tels que l'utilité et la demande. Ils arrivent à la conclusion que cette théorie présente des qualités non négligeables, comme :

- possibilité de représenter l'ignorance et d'éviter d'insérer ou de percevoir une précision injustifiée dans les prévisions du marché ;

- mécanisme cohérent pour la combinaison de nouvelle évidence avec les prévisions déjà existantes, grâce à la règle de combinaison de Dempster ;
- possibilité d'attribuer une portion de la masse totale de crédibilité à des paramètres de prévision considérés fortement improbables, mais perçus par l'expert comme possibles ;
- potentiel considérable pour les applications dans les domaines économiques de grande incertitude, tels que la prévision de la demande de nouveaux produits et de services.

Chapitre 2

Notions de base

2.1 Probabilités objectives et subjectives

Afin de faciliter d'exposition des concepts de base de la théorie mathématique de l'évidence, commençons par un exemple simple proposé par Ellsberg [26] et illustré par la Figure 2.1 :

Exemple 2.1. Supposons une urne contenant 100 boules, dont 50 rouges et 50 noires. Considérons l'expérience qui consiste à tirer une boule de cette urne. Quel sera le résultat de cette expérience ou, autrement dit, la boule tirée sera-t-elle rouge ou noire ? L'issue est incertaine car l'expérience est aléatoire ; cependant elle est quantifiable. En effet, connaissant la proportion de chaque couleur de boules, on peut conclure que dans 50% des cas, la boule tirée sera rouge et que dans 50% des cas, elle sera noire. Imaginons à présent une seconde urne, contenant également 100 boules, rouges et noires, mais de proportion inconnue. Il se peut que toutes les boules soient rouges, qu'elles soient au contraire toutes noires ou encore de n'importe quelle autre proportion. Einhorn et Hogarth [25] appellent la situation de la première urne une **situation de connaissance complète** et celle de la seconde urne une **situation d'ambiguïté complète**.

Dans le cadre de la théorie classique des probabilités, on supposera souvent des distributions uniformes pour des situations ambiguës comme celle de la seconde urne. Ce faisant, on attribuera la probabilité de 0.5 à chacune des deux issues possibles, c'est-à-dire à ce que la boule tirée soit rouge ou qu'elle soit noire.

Il est dès lors aisé de saisir tout l'intérêt des théories traitant du raisonnement sous incertitude. En effet, sur la seule base de la théorie classique des probabilités, on considère les situations des deux urnes de manière identique, sans les distinguer, malgré leur différence

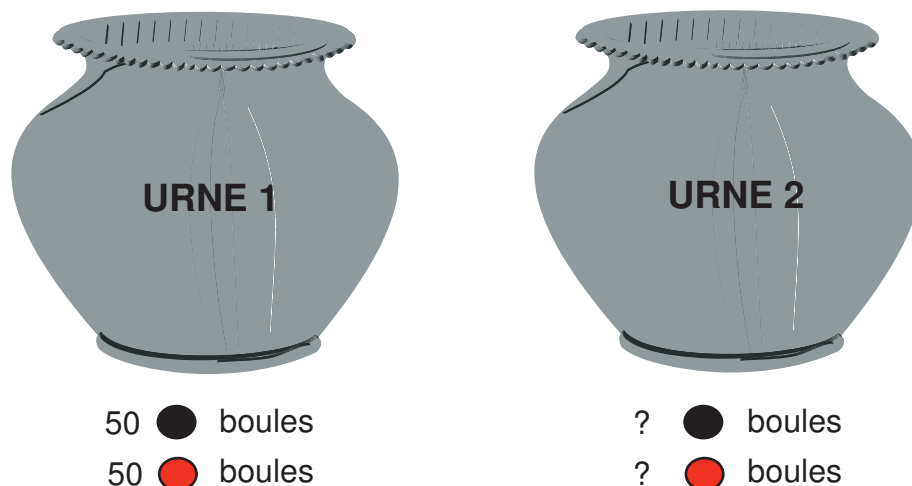


FIG. 2.1 – Connaissance et ambiguïté

fondamentale quant à l'information véhiculée. A ce propos, notons que ce problème avait déjà été relevé en 1921 par Knight [51]. Celui-ci distingue la notion de risque de celle d'incertitude. Selon lui, la notion de risque se réfère aux situations où l'on peut assigner des probabilités mathématiques au phénomène aléatoire auquel on est confronté. Par opposition, sa notion d'incertitude se réfère aux situations dont le comportement aléatoire ne peut être exprimé en termes de probabilités mathématiques spécifiques.

Depuis des siècles, les idées du **degré numérique de crédibilité** et de **chance** ont été confondues et exprimées sous une même notion, celle de **probabilité**. Le concept de chance est utilisé uniquement pour décrire une **expérience aléatoire**, comme le sont par exemple le lancer d'un dé ou d'une pièce de monnaie. Dans ce cas, on parle de probabilité **fréquentiste** ou **objective**. Le degré numérique de crédibilité exprime le degré que l'on doit accorder à une proposition en se basant sur l'évidence qui relève de l'information dont on dispose. Il est également représenté par un nombre compris entre 0 et 1. Il s'agit non plus d'un concept aléatoire, mais **épistémique**. Dans ce cas, on parle de probabilité **subjective** (voir [94]). Celle-ci représente la crédibilité que l'on accorde à la réalisation d'un événement. En effet, on peut croire en quelque chose sans que pour autant cette chose soit vérifiable, comme ce serait le cas pour les expériences aléatoires. Supposons une personne convaincue d'avoir vu un OVNI dans le ciel. Elle supportera de ce fait l'hypothèse selon laquelle une vie extra-terrestre existe. Elle lui accordera un degré de crédibilité en se basant sur l'évidence résultant de ce qu'elle a cru voir, de sa propre expérience. Le degré de crédibilité accordé n'est que la représentation numérique du niveau auquel elle croit en cette hypothèse. Il sera nul, lorsqu'aucune évidence ne permet de supporter l'hypothèse, et de 1, lorsque l'évidence sur laquelle on se base nous mène à la certitude.

Les chances régissant une expérience aléatoire ne coïncident pas obligatoirement avec les degrés de crédibilité qu'on accorde aux résultats possibles de cette expérience. Il est important de noter que les chances sont des caractéristiques du monde qui nous entoure

et non pas de notre savoir ou expérience. A l'opposé, les degrés de crédibilité accordés à une proposition sont issus d'un acte de jugement personnel, donc d'un acte subjectif.

Il existe trois fonctions de base importantes pour la compréhension de l'utilisation de la théorie mathématique de l'évidence :

1. les **fonctions de masse** (an anglais «basic belief mass functions» ou «basic probability assignment functions» dans les travaux de Shafer [95]);
2. les **fonctions de crédibilité** (en anglais «belief functions»);
3. les **fonctions de plausibilité** (an anglais «plausibility functions»).

Mais avant d'introduire ces trois fonctions de base, il est important de distinguer clairement les idées de chance et de crédibilité.

2.2 Distinction entre l'idée de chance et l'idée de crédibilité

La théorie des probabilités, ou **théorie des chances** comme on l'appellera dans la Partie 1, afin d'éviter toute confusion entre les deux notions de probabilité, repose sur une idée simple, celle de la densité $q : \mathcal{X} \rightarrow [0, 1]$ qui décrit une expérience aléatoire avec des issues possibles dans \mathcal{X} . Cette théorie mathématique porte le nom de doctrine des chances de Abraham de Moivre (1667-1754). On doit ses fondements axiomatiques complets à Kolmogoroff [64].

Soit \mathcal{X} un ensemble fini qui dénote l'ensemble de tous les résultats possibles d'une expérience aléatoire. Si l'on spécifie la chance $q(x)$ de chaque résultat possible $x \in \mathcal{X}$, tel que $0 \leq q(x) \leq 1$ pour tout $x \in \mathcal{X}$ et

$$\sum_{x \in \mathcal{X}} q(x) = 1, \quad (2.1)$$

soit rempli, alors, on a spécifié la fonction $q : \mathcal{X} \rightarrow [0, 1]$. Ainsi, q spécifie complètement les chances impliquées dans l'expérience et est appelée la **densité de chance** qui régit l'expérience. On pourrait également s'intéresser à la chance qu'aurait un résultat de ne pas être un élément particulier x de \mathcal{X} , mais un sous-ensemble particulier U de \mathcal{X} . Ceci s'appelle la chance que U survienne. Cette chance est notée par $Ch(U)$ et est calculée en additionnant les chances des différents éléments de U :

$$Ch(U) = \sum_{x \in U} q(x). \quad (2.2)$$

Il est évident que la fonction Ch véhicule exactement la même information que q . Elle porte le nom de **fonction de chance** correspondant à q .

Une fonction $Ch : 2^{\mathcal{X}} \rightarrow [0, 1]$ est une fonction de chance (c'est-à-dire qu'elle peut être obtenue à partir d'une densité de chance q de \mathcal{X}) si et seulement si les règles suivantes sont respectées :

1. $Ch(\emptyset) = 0$.
2. $Ch(\mathcal{X}) = 1$.
3. Si $U, V \subset \mathcal{X}$ et $U \cap V = \emptyset$, alors $Ch(U \cup V) = Ch(U) + Ch(V)$.

Ces trois règles sont appelées **règles de base pour les chances** ; la troisième porte en plus le nom de **règle d'additivité pour les chances**. En comparaison avec les règles pour les fonctions de crédibilité, cette règle sera rejetée comme règle pour les degrés de crédibilité.

2.3 Fonction de masse

Imaginons que l'on s'intéresse à la vraie valeur d'une certaine quantité. Notons cette quantité par θ et l'ensemble de toutes ses valeurs possibles par Θ . On s'intéresse en particulier aux propositions de la forme «la vraie valeur de θ est dans T », où T est un sous-ensemble de Θ . Dans la théorie mathématique de l'évidence, cet ensemble est supposé **fini**. Ainsi, les propositions sur lesquelles porte notre intérêt correspondent une à une aux sous-ensembles de Θ , et l'ensemble de toutes les propositions correspond à l'ensemble de tous les sous-ensembles de Θ , qui est noté par le symbole 2^{Θ} . A présent, si θ est un paramètre arbitraire et prend éventuellement des valeurs non numériques, alors θ et Θ peuvent être choisis de manière à ce que 2^{Θ} inclut chaque ensemble fixé et particulier que nous aimerions qu'il contienne. En effet, dans la plupart des cas, il suffit de choisir Θ de manière à ce qu'il soit l'ensemble de toutes les différentes possibilités considérées. Considérons un événement avec n issues ou états possibles $\{a_1, a_2, \dots, a_n\}$ qui sont exhaustifs et mutuellement exclusifs. Ces états forment un ensemble représenté par Θ . Celui-ci est désigné par le terme **cadre** (en anglais «frame»). Rappelons que dans la théorie classique des probabilités, à chaque état a_i , avec $i = 1, \dots, n$, est attribuée une chance $Ch(a_i) \geq 0$ et la somme de toutes ces chances est égale à 1, c'est-à-dire $\sum_{i=1}^n Ch(a_i) = 1$.

Dans le cadre des fonctions de crédibilité, les chances sont remplacées par des masses (en anglais «basic belief masses» ou «basic probability numbers» dans les travaux de Shafer) et plutôt que d'être attribuées à chaque état, elles le sont à tous les sous-ensembles possibles d'états. Rappelons que le nombre de sous-ensembles possibles de n états est de 2^n . Ainsi, les masses indiquent les degrés numériques de crédibilité que l'on accorde aux sous-ensembles. En résumé, il faut attribuer une masse à chaque état isolé, à chaque sous-

ensemble de deux états, à chaque sous-ensemble de trois états, et ainsi de suite. Comme pour les chances, toutes les masses attribuées doivent sommer à 1, c'est-à-dire

$$\sum_{A \subset \Theta} m(A) = 1, \quad (2.3)$$

où la somme porte sur tous les sous-ensembles du cadre Θ et $m(A)$ représente les masses qui leur sont associées. L'ensemble vide fait bien évidemment partie des 2^n sous-ensembles possibles ; cependant, il est important de noter que $m(\emptyset) = 0$ par définition. Dans le *Transferable Belief Model* de Smets [107], [108], il est possible d'attribuer une masse non nulle à l'ensemble vide, mais nous ne traiterons pas ce cas ici. Il est nécessaire de souligner qu'une portion de la crédibilité de n'importe quel sous-ensemble de A implique non seulement A , mais peut éventuellement impliquer d'autres sous-ensembles. En effet, la portion de crédibilité attribuée à un sous-ensemble est également attribuée à n'importe quel sous-ensemble qui le contient. Donc, de la crédibilité totale attribuée à un sous-ensemble donné de A du cadre Θ , une partie peut être attribuée à un ou plusieurs sous-ensembles de A , tandis que la partie restante est attribuée exactement à A et à aucun autre sous-ensemble plus petit. C'est cette partie-là justement que l'on appelle la masse de A .

Afin de souligner la nature épistémique de l'ensemble de possibilités Θ , Shafer [95] le nomme **cadre de discernement** (en anglais «frame of discernment»). De même, lorsqu'une proposition correspond à un sous-ensemble du cadre de discernement, il dit que le cadre **discerne** cette proposition. Shafer explique en ces termes l'utilité de la terminologie et des notions employées :

«Une des raisons pour lesquelles la correspondance entre les propositions et les sous-ensembles est utile est qu'elle traduit les notions logiques de conjugaison, disjonction, implication et négation en notions d'avantage graphiques de la théorie des ensembles comme celles d'intersection, union, inclusion et complémentation.»

Si Θ est un cadre de discernement, alors une fonction $m : 2^\Theta \rightarrow [0, 1]$ est appelée une **fonction de masse** si et seulement si

1. $m(\emptyset) = 0$,
2. $\sum_{A \subset \Theta} m(A) = 1$.

La quantité $m(A)$ est appelée **masse** de A et est entendue comme la mesure de la crédibilité allant exactement dans le sens de A . De plus, chaque $A \subset \Theta$ tel que $m(A) > 0$ est appelé **élément focal**.

Pour éviter d'alourdir la notation, les masses des singletons éléments de Θ , seront dorénavant notées :

$$m(x) := m(\{x\}), \quad (2.4)$$

pour tout $x \in \Theta$. Reprenons ces concepts à travers un exemple.

Exemple 2.2. Imaginons que l'on s'intéresse à l'évolution sur 24 heures du cours d'une action cotée en bourse. Pour faciliter la compréhension, simplifions cette évolution à uniquement deux issues possibles. Ainsi, au bout de 24 heures, soit le cours aura changé, événement a_1 , soit se sera resté stable, événement a_2 . Le cadre de discernement de cet exemple est évidemment représenté par $\Theta = \{a_1, a_2\}$. Nous devons attribuer à présent une masse à chacun des sous-ensembles possibles du cadre et, vu que celui-ci est composé de deux éléments, nous aurons $2^2 = 4$ masses à déterminer, c'est-à-dire $m(\emptyset)$, $m(a_1)$, $m(a_2)$ et enfin $m(\{a_1, a_2\})$. Rappelons que $m(\emptyset) = 0$ par définition. Pour la détermination des trois autres masses, nous devons nous baser sur notre expérience et notre savoir, autrement dit, sur l'information que nous avons à disposition. Cette information représentera l'évidence allant dans le sens de l'une ou l'autre des possibilités. Elle n'est pas forcément complète, elle peut également être floue ou ambiguë. En effet, il ne peut exister que trois situations possibles, celle de l'information complète, celle de l'information partielle et enfin celle de l'information nulle. Présentons les fonctions de masse dans chacune de ces trois situations.

1. Information complète

Tout en sachant qu'il est impossible de connaître les chances exactes de l'évolution du cours d'une action, supposons que ce soit le cas aux fins d'illustrer de cet exemple. Supposons en effet que nous sachions de manière sûre que la chance que le cours reste stable est de 0.4 et celle que le cours change est de 0.6. Si on utilise le mot **chance** c'est justement car en situation d'information complète, on retrouve le cadre de la théorie des chances. La masse attribuée à chaque état possible est alors égale à la chance de cet état ; on a ainsi :

- $m(\emptyset) = 0$,
- $m(a_1) = 0.6$,
- $m(a_2) = 0.4$,
- $m(\{a_1, a_2\}) = 0$.

Il est évident que la masse de l'ensemble vide soit nulle, cela résulte en effet des hypothèses de base. Les masses des deux états possibles sont égales aux chances de ces mêmes états, comme il a déjà été expliqué. Pour comprendre pourquoi $m(\{a_1, a_2\}) = 0$, il faut revenir aux explications données au début de cette section. Rappelons que nous appelons **masse** la crédibilité attribuée à A , exactement à A et à aucun autre de ses sous-ensembles plus petits. Comme on dispose d'une information parfaite ou complète, la totalité de la crédibilité est partagée entre a_1 et a_2 , qui sont des sous-ensembles de $\{a_1, a_2\}$. Autrement dit, aucune portion de la crédibilité ne va spécifiquement vers $\{a_1, a_2\}$ et de ce fait, il a une masse

nulle. Enfin, $m(a_1) = 0.6$ et $m(a_2) = 0.4$ car, rappelons-le, ces masses indiquent le degré numérique de crédibilité que l'on accorde à a_1 et à a_2 . Puisqu'on dispose de toute l'information, il est logique de leur attribuer des masses identiques à leurs chances respectives de réalisation.

Enfin, lorsque l'information est complète, il n'est pas nécessaire de recourir à la théorie de l'évidence, car, rappelons-le, celle-ci traite avant tout du raisonnement sous incertitude, c'est-à-dire du raisonnement en situations d'information partielle ou nulle. Même si, comme on vient de le constater, elle peut également s'appliquer aux situations d'information complète, il est possible, dans ces cas-là, de se limiter à la théorie des chances. Ainsi, celle-ci peut être interprétée comme un cas spécial de la théorie mathématique de l'évidence, le cas justement où on n'est pas en présence d'ambiguïté, mais où on dispose d'une information complète.

2. Information partielle

Supposons à présent que nous disposons d'information partielle concernant l'évolution du cours de l'action. Par exemple, une analyse financière publiée la veille nous pousse à croire que le cours restera stable, mais ceci sans grande certitude. Nous en déduisons que $m(a_2) = 0.3$. D'un autre côté, le fait que le marché boursier soit très volatile en ce moment, nous indique que le cours aura tendance à fluctuer, mais il ne s'agit, là aussi, que d'une supposition. Nous en concluons que $m(a_1) = 0.2$. Nous sommes en présence d'une évidence partielle pour chacune des deux éventualités. En effet, certains éléments nous laissent penser que l'une ou l'autre éventualité se produira ; cependant, du fait que l'évidence n'est que partielle, notre crédibilité n'est pas totalement partagée entre les deux possibilités. L'évidence restante, c'est-à-dire celle qui ne va dans le sens d'aucune des deux issues possibles, sera considérée comme évidence allant vers l'ensemble du cadre de discernement. Par conséquent, une partie de la crédibilité sera attribuée à $\Theta = \{a_1, a_2\}$, ce qui en termes intuitifs se traduit par : «Nous ne pouvons pas nous prononcer». Il s'agit de la masse restante qui n'a pas encore été attribuée. Il s'en suit que $m(\{a_1, a_2\}) = 0.5$. Pour reprendre l'exemple 2.1 cité au début du chapitre, nous sommes à mi-chemin entre le cas de la première urne et celui de la seconde. En résumé, notre crédibilité sera partagée entre une partie de connaissance et une partie d'ignorance. Nous aurons

- $m(\emptyset) = 0$,
- $m(a_1) = 0.2$,
- $m(a_2) = 0.3$,
- $m(\{a_1, a_2\}) = 0.5$.

3. Information nulle

Supposons pour illustrer cette troisième situation que l'on ne dispose d'aucune information, ni sur l'action considérée, ni sur le marché boursier. Par conséquent, nous sommes en présence d'une ignorance totale concernant l'évolution du cours de l'action pendant les 24 prochaines heures (cas de la deuxième urne). Aucun élément d'évidence ne va dans le sens d'une des deux issues possibles ; notre degré numérique de crédibilité les concernant sera donc nul et nous leur attribuons les masses suivantes : $m(a_1) = 0$, $m(a_2) = 0$. De ce fait la totalité de l'évidence sera dirigée vers l'ensemble du cadre de discernement Θ et $m(\{a_1, a_2\}) = 1$. En résumé, aucun élément ne permettant de s'exprimer en faveur de l'une ou l'autre éventualité, nous ne sommes capables de dire qu'une seule chose : «Nous ne pouvons pas nous prononcer». On aura :

- $m(\emptyset) = 0$,
- $m(a_1) = 0$,
- $m(a_2) = 0$,
- $m(\{a_1, a_2\}) = 1$.

2.4 Fonction de crédibilité

La crédibilité d'un ensemble d'éléments A , dans un cadre de discernement Θ , est définie comme la crédibilité totale de A . Autrement dit, il s'agit de l'évidence qui va précisément dans le sens de A ou de l'ensemble de l'évidence qui prouve ou supporte cette proposition. Cela représente la somme de toutes les masses attribuées aux éléments contenus dans A , plus la masse attribuée à A lui-même. Donc, si nous voulons mesurer la crédibilité totale attribuée à A , nous devons ajouter à $m(A)$ les quantités $m(B)$ pour tous les sous-ensembles B de A , étant donné que A est impliqué par tous les sous-ensembles $B \subset A$. On obtient :

$$Bel(A) = \sum_{B \subset A} m(B). \quad (2.5)$$

$Bel(A)$ est appelé le degré de crédibilité de A . La fonction $Bel : 2^\Theta \rightarrow [0, 1]$, caractérisée par la fonction de masse $m : 2^\Theta \rightarrow [0, 1]$, est appelée **fonction de crédibilité** sur Θ . De même, $Bel(A)$ est appelé **degré de crédibilité** de A .

Pour éviter d'alourdir la notation, les degrés de crédibilité des singletons éléments de Θ , seront dorénavant notées :

$$Bel(x) := Bel(\{x\}), \quad (2.6)$$

pour tout $x \in \Theta$.

Reprenons l'exemple 2.2 pour illustrer le concept de fonction de crédibilité.

Exemple 2.3. Rappelons les trois situations :

1. Information complète

Nous avons les masses suivantes :

- $m(\emptyset) = 0$,
- $m(a_1) = 0.6$,
- $m(a_2) = 0.4$,
- $m(\{a_1, a_2\}) = 0$.

A partir des masses, il est aisé de calculer les fonctions de crédibilité correspondantes. Pour obtenir la fonction de crédibilité d'un des sous-ensembles, il suffit d'additionner les masses de tous les sous-ensembles qu'il contient.

Les degrés de crédibilité qui en résultent sont :

- $Bel(\emptyset) = 0$,
- $Bel(a_1) = 0.6$,
- $Bel(a_2) = 0.4$,
- $Bel(\{a_1, a_2\}) = 1$.

Le degré de crédibilité de $\{a_1, a_2\} = \Theta$ est de 1, car il équivaut à la somme des masses de tous ses sous-ensembles. Comme Θ contient tous les sous-ensembles considérés et que la somme de leurs masses respectives est égale à 1 par définition, le degré de crédibilité de Θ sera toujours de 1, comme on le verra dans le théorème 2.1.

2. Information partielle

Nous avons les masses suivantes :

- $m(\emptyset) = 0$,
- $m(a_1) = 0.2$,
- $m(a_2) = 0.3$,
- $m(\{a_1, a_2\}) = 0.5$.

Les degrés de crédibilité qui en résultent sont :

- $Bel(\emptyset) = 0$,
- $Bel(a_1) = 0.2$,
- $Bel(a_2) = 0.3$,
- $Bel(\{a_1, a_2\}) = 1$.

Comme l'ensemble nul ne contient que lui-même, il est clair que $Bel(\emptyset) = 0$, propriété également mentionnée au théorème 2.1. La fonction de crédibilité de a_1 est donnée par la somme de $m(\emptyset)$ et $m(a_1)$, les masses des

deux sous-ensembles de a_1 , donc nous obtenons $Bel(a_1) = 0.2$. Etant donné que la masse de l'ensemble vide est par définition nulle, on en conclut évidemment que la crédibilité totale d'un sous-ensemble composé d'un seul et unique élément est égale à sa masse. De ce fait, nous pouvons directement déterminer la valeur de la fonction de crédibilité pour a_2 : $Bel(a_2) = 0.3$. Enfin, avant de déterminer la valeur de la fonction de crédibilité pour $\{a_1, a_2\}$, rappelons que la somme des masses de tous les sous-ensembles du cadre de discernement est toujours égale à 1 et comme $\Theta = \{a_1, a_2\}$, $Bel(\Theta) = Bel(\{a_1, a_2\}) = 1$.

3. Information nulle

Nous avons les masses suivantes :

- $m(\emptyset) = 0$,
- $m(a_1) = 0$,
- $m(a_2) = 0$,
- $m(\{a_1, a_2\}) = 1$.

Les degrés de crédibilité qui en résultent sont :

- $Bel(\emptyset) = 0$,
- $Bel(a_1) = 0$,
- $Bel(a_2) = 0$,
- $Bel(\{a_1, a_2\}) = 1$.

Notons au passage que dans la situation d'information nulle, c'est-à-dire lorsque $m(A) = 0$ et $m(\Theta) = 1$ pour tout $A \neq \Theta$, on a $Bel(A) = 0$ pour tout $A \neq \Theta$ et $Bel(\Theta) = 1$. Une telle fonction est appelée **fonction de crédibilité vide**.

Il est également possible de caractériser la classe des fonctions de crédibilité sans se référer aux fonctions de masse.

Théorème 2.1. *Si Θ est un cadre de discernement, alors une fonction $Bel : 2^\Theta \rightarrow [0, 1]$ est une fonction de crédibilité si et seulement si elle satisfait les conditions suivantes :*

1. $Bel(\emptyset) = 0$.
2. $Bel(\Theta) = 1$.
3. Pour chaque entier positif n et chaque collection A_1, \dots, A_n de sous-ensembles de Θ ,

$$Bel(A_1 \cup \dots \cup A_n) \geq \sum_{\substack{I \subset \{1, \dots, n\} \\ I \neq \emptyset}} (-1)^{|I|+1} Bel\left(\bigcap_{i \in I} A_i\right). \quad (2.7)$$

Ce théorème et sa preuve ont été présentés par Shafer [95].

De plus, la masse qui génère une certaine fonction de crédibilité est unique est peut être obtenue à partir de la fonction de crédibilité.

Théorème 2.2. *Supposons que $Bel : 2^\Theta \rightarrow [0, 1]$ est une fonction de crédibilité donnée par la fonction de masse $m : 2^\Theta \rightarrow [0, 1]$. Alors*

$$m(A) = \sum_{B \subset A} (-1)^{|A-B|} Bel(B) \quad (2.8)$$

pour tout $A \subset \Theta$.

Ce théorème et sa preuve ont été présentés par Shafer [95].

Reprenons l'exemple 2.3 en situation d'information partielle pour illustrer le théorème 2.2.

Exemple 2.4. Rappelons qu'une analyse publiée la veille nous poussait à croire que le cours resterait stable, mais sans grande certitude. Nous disposons des valeurs suivantes :

- $Bel(\emptyset) = 0$,
- $Bel(a_1) = 0.2$,
- $Bel(a_2) = 0.3$,
- $Bel(\{a_1, a_2\}) = 1$.

Par définition $m(\emptyset) = 0$. Pour obtenir $m(a_1)$, appliquons la formule :

$$m(a_1) = \sum_{B \subset a_1} (-1)^{|a_1-B|} Bel(B) = (-1)^0 Bel(a_1) = 0.2.$$

De même,

$$m(a_2) = (-1)^0 Bel(a_2) = 0.3.$$

Enfin,

$$m(\{a_1, a_2\}) = (-1)^1 Bel(a_1) + (-1)^1 Bel(a_2) + (-1)^0 Bel(\{a_1, a_2\}) = 0.5.$$

Rappelons qu'un sous-ensemble A du cadre de discernement Θ , tel que $m(A) > 0$, est appelé un élément focal. Maintenant que les fonctions de crédibilité ont été introduites, on peut ajouter qu'il s'agit d'un élément focal d'une fonction de crédibilité sur Θ . L'union de tous les éléments focaux d'une fonction de crédibilité s'appelle le **coeur** de la fonction de crédibilité.

Théorème 2.3. *Supposons que \mathcal{C} soit le coeur d'une fonction de crédibilité Bel sur Θ . Alors, un sous-ensemble $B \subset \Theta$ satisfait $Bel(B) = 1$ si et seulement si $\mathcal{C} \subset B$.*

Ce théorème et sa preuve ont été présentés par Shafer [95].

En effet, la valeur de la fonction de crédibilité pour un sous-ensemble B est égale à la somme des masses de tous les sous-ensembles contenus dans B , plus celle de B . Comme cette somme est égale à 1, cela implique que tous les éléments focaux doivent être contenus dans B et que tous les sous-ensembles de Θ que B ne contient pas ont une masse nulle. Cela équivaut à dire que si $Bel(B) = 1$, alors B contient le coeur de la fonction de crédibilité. De même, si B contient le coeur de la fonction de crédibilité, cela signifie qu'il contient tous les éléments focaux. La somme des masses de tous les éléments focaux étant égale à 1 par définition, il s'en suit que $Bel(B) = 1$. Reprenons l'exemple 2.3 en situation d'information partielle.

Exemple 2.5. Le coeur de la fonction de crédibilité correspondante est l'union de tous les sous-ensembles ayant une masse non nulle; donc $\mathcal{C} = (\{a_1\} \cup \{a_1\} \cup \{a_1, a_2\}) = \{a_1, a_2\}$. Comme le seul sous-ensemble qui contienne \mathcal{C} est le cadre de discernement Θ lui-même et il est évident que $Bel(\Theta) = 1$.

2.5 Fonction de plausibilité

La plausibilité d'un ensemble d'éléments A , dans un cadre Θ , est définie comme la crédibilité maximale possible qui pourrait être attribuée à A si toute l'évidence future supportait A . Autrement dit, il s'agit de toute l'évidence qui ne contredit pas A . Mathématiquement, la fonction de plausibilité est définie comme :

$$Pl(A) = \sum_{B \cap A \neq \emptyset} m(B). \quad (2.9)$$

Une fonction $Pl : 2^\Theta \rightarrow [0, 1]$, caractérisée par la fonction de masse $m : 2^\Theta \rightarrow [0, 1]$, est appelée une **fonction de plausibilité** sur Θ . De même, $Pl(A)$ est appelé **degré de plausibilité** de A .

Pour éviter d'alourdir la notation, les degrés de plausibilité des singletons éléments de Θ , seront dorénavant notées :

$$Pl(x) := Pl(\{x\}), \quad (2.10)$$

pour tout $x \in \Theta$.

A noter que le nom donné à cette fonction par Shafer [95] est **fonction de probabilité supérieure**, nom qu'il justifie par la définition même de la fonction. On privilégiera cependant, dans la suite de ce travail, l'expression **fonction de plausibilité**, car ce terme est le plus fréquemment utilisé dans la littérature (voir [114]).

La fonction $Pl : 2^\Theta \rightarrow [0, 1]$ peut également être définie à travers la relation suivante :

$$Pl(A) = 1 - Bel(\bar{A}). \quad (2.11)$$

En effet, le degré de plausibilité de A représente tout le degré de crédibilité qui n'est pas attribué à \bar{A} .

Présentons cela à travers l'exemple 2.2 en situation d'information partielle.

Exemple 2.6. Rappelons que nous avons les masses suivantes :

- $m(\emptyset) = 0$,
- $m(a_1) = 0.2$,
- $m(a_2) = 0.3$,
- $m(\{a_1, a_2\}) = 0.5$.

Sur la base de l'équation (2.9), on obtient

$$Pl(\emptyset) = \sum_{B \cap \emptyset \neq \emptyset} m(B) = 0,$$

car toute intersection entre l'ensemble vide et n'importe quel autre sous-ensemble donne forcément un ensemble vide. Il est aisé d'arriver au même résultat en utilisant tout simplement la relation entre une fonction de crédibilité et une fonction de plausibilité :

$$Pl(\emptyset) = 1 - Bel(\bar{\emptyset}) = 1 - Bel(\Theta) = 1 - 1 = 0.$$

En effet, le complément de l'ensemble vide est le cadre de discernement lui-même. Celui-ci bénéficie d'un degré de crédibilité par définition égal à 1, par conséquent, le degré de plausibilité de l'ensemble vide est de 0. Pour a_1 , de manière identique, on obtient

$$Pl(a_1) = \sum_{B \cap a_1 \neq \emptyset} m(B) = m(a_1) + m(\{a_1, a_2\}) = 0.2 + 0.5 = 0.7.$$

En effet, dans ce cadre de discernement, le seul sous-ensemble qui soit disjoint avec a_1 est a_2 , c'est donc le seul dont la masse ne sera pas considérée dans cette somme. Il est également possible d'obtenir ce résultat en appliquant l'équation (2.11) :

$$Pl(a_1) = 1 - Bel(\bar{a}_1) = 1 - Bel(a_2) = 1 - 0.3 = 0.7.$$

De la même manière,

$$Pl(a_2) = 1 - Bel(a_1) = 1 - 0.2 = 0.8$$

et enfin

$$Pl(\{a_1, a_2\}) = Pl(\Theta) = 1 - Bel(\emptyset) = 1 - 0 = 1.$$

Notons que, dans un cadre de discernement Θ , $Pl(\emptyset)$ est toujours égal à 0 et $Pl(\Theta)$ à 1. On peut en effet le constater à partir de l'équation (2.11).

Il est important d'attirer l'attention sur le fait qu'une fonction de plausibilité transmet exactement la même information que la fonction de crédibilité correspondante et que chacune des deux peut être obtenue à partir de l'autre.

Enfin, si nous comparons la fonction de crédibilité avec la fonction de plausibilité, il apparaît que la fonction de plausibilité domine la fonction de crédibilité. En effet, dans un cadre Θ ,

$$Bel(A) \leq Pl(A), \quad (2.12)$$

pour tout $A \subset \Theta$.

2.6 Fonction de crédibilité bayésienne

La **théorie bayésienne** doit son nom à Thomas Bayes (1702-1761). Selon cette théorie, tous les degrés de crédibilité respectent les règles pour les chances. En effet, la théorie bayésienne adopte les trois règles pour les chances (voir section 2.2) comme règles pour le degré de crédibilité étant donné un certain niveau d'évidence. De même, elle adopte la règle de conditionnement des chances comme règle générale pour le changement du degré de crédibilité lorsque l'évidence augmente par la connaissance d'une proposition particulière.

Les trois premières règles de la théorie bayésienne peuvent être exprimées comme règles régissant une fonction portant sur un cadre Θ . Le degré de crédibilité doit être donné par une fonction $Bel : 2^\Theta \rightarrow [0, 1]$ qui respecte les trois règles suivantes :

1. $Bel(\emptyset) = 0$.
2. $Bel(\Theta) = 1$.
3. Si $A \cap B = \emptyset$, alors $Bel(A \cup B) = Bel(A) + Bel(B)$.

Seule la troisième règle diffère des règles pour les fonctions de crédibilité. Toutes les fonctions qui respectent ces trois règles sont des fonctions de crédibilité, mais le contraire n'est pas forcément vrai. En effet, les fonctions qui respectent les trois règles de Bayes forment une sous-classe de la classe des fonctions de crédibilité et sont appelées **fonctions de crédibilité bayésiennes**.

Théorème 2.4. *Une fonction de crédibilité bayésienne est une fonction de crédibilité.*

Ce théorème et sa preuve ont été présentés par Shafer [95].

En effet, si l'on commence avec une fonction de crédibilité bayésienne $Bel : 2^\Theta \rightarrow [0, 1]$, puis on apprend que $A \subset \Theta$ est vrai, Bel sera remplacée par une nouvelle fonction de crédibilité bayésienne, notée par $Bel_A : 2^\Theta \rightarrow [0, 1]$ et qui sera donnée par

$$Bel_A = \frac{Bel(B \cap A)}{Bel(A)} \quad (2.13)$$

pour tout $B \subset \Theta$ et $Bel(A) > 0$. Cette quatrième règle est appelée **règle de Bayes pour le conditionnement**. La plupart du temps, cette expression est notée par $Bel(B|A)$ et est lue «le degré conditionnel de crédibilité de B étant donné A ». Notons au passage que cette règle de conditionnement est un cas spécial de la règle de combinaison de Dempster, qu'on présentera plus loin.

Le problème de base de la théorie bayésienne par rapport à celle de l'évidence est qu'elle ne peut faire la distinction entre le manque de crédibilité et l'incrédibilité. Par exemple, on peut n'avoir aucune idée du temps qu'il fera demain. On ne soutient ni l'hypothèse qu'il fera beau temps, ni celle qu'il pleuvra. Ne disposant d'aucune information allant dans le sens de l'une ou de l'autre de ces deux hypothèses, on est en manque de crédibilité concernant chacune d'entre elles. Cependant, on avait consulté les prévisions météorologiques, annonçant du beau temps pour le lendemain, alors, on aurait accordé à cette hypothèse un degré de crédibilité élevé, tandis que l'hypothèse selon laquelle il pleuvra nous aurait laissée incrédules (degré de crédibilité très faible ou nul). Il s'agit à l'évidence de deux situations différentes que la théorie bayésienne ne distingue justement pas.

La règle de Bayes pour l'additivité requiert cependant que $Bel(A) + Bel(\bar{A}) = 1$, car $A \cup \bar{A} = \Theta$ et $Bel(\Theta) = 1$. Ceci implique qu'à défaut de croire en A , il faut croire en \bar{A} , mais le fait qu'on ne soutienne pas A ne doit pas obligatoirement impliquer qu'on soutient \bar{A} .

Chapitre 3

Règle de combinaison de Dempster

La règle de combinaison de Dempster donne une méthode pour changer les opinions antérieures à la lumière de nouvelle évidence. Il s'agit d'une règle simple permettant de calculer, à partir de deux ou plusieurs fonctions de crédibilité portant sur le même ensemble Θ , une nouvelle fonction de crédibilité appelée leur somme orthogonale et basée sur l'évidence combinée.

Le cas spécial d'un cadre de discernement ne contenant que deux éléments a été traité pour la première fois par Johann Heinrich Lambert dans l'ouvrage *Neues Organon* publié en 1764. Cependant, la formulation générale de cette règle de calcul n'a été achevée que deux cents ans plus tard et publiée par Arthur Dempster en 1967 [21].

3.1 Combinaison de deux fonctions de crédibilité

Supposons que m_1 soit la fonction de masse associée à la fonction de crédibilité Bel_1 sur un cadre de discernement Θ . Notons par A_1, \dots, A_k les éléments focaux de Bel_1 . Dans ce cas, les masses seront mesurées par $m_1(A_1), \dots, m_1(A_k)$. Supposons également l'existence d'une deuxième fonction de crédibilité Bel_2 sur Θ avec la fonction de masse m_2 , les éléments focaux B_1, \dots, B_l et dont les masses sont mesurées par $m_2(B_1), \dots, m_2(B_l)$. Comment mesurer la crédibilité de l'intersection de deux éléments, par exemple la crédibilité de $A_i \cap B_j$?

Graphiquement, on peut représenter ces deux fonctions de masse sous la forme d'un tableau où $m_1(A_i)$ avec $i = 1, \dots, k$ correspond aux colonnes et $m_2(B_j)$ avec $j = 1, \dots, l$ respectivement aux lignes. D'après Shafer, la masse de l'intersection de A_i et B_j est donnée par le croisement de la colonne i avec la ligne j , donc par le produit $m_1(A_i)m_2(B_j)$. Ainsi, Shafer suppose l'indépendance des fonctions de masse.

	$m_1(A_1)$	$m_1(A_2)$	\dots	$m_1(A_i)$	\dots	$m_1(A_k)$
$m_2(B_1)$	$m_1(A_1) \cdot m_2(B_1)$	$m_1(A_2) \cdot m_2(B_1)$	\dots	$m_1(A_i) \cdot m_2(B_1)$	\dots	$m_1(A_k) \cdot m_2(B_1)$
$m_2(B_2)$	$m_1(A_1) \cdot m_2(B_2)$	$m_1(A_2) \cdot m_2(B_2)$	\dots	$m_1(A_i) \cdot m_2(B_2)$	\dots	$m_1(A_k) \cdot m_2(B_2)$
\vdots	\vdots	\vdots		\vdots		\vdots
$m_2(B_j)$	$m_1(A_1) \cdot m_2(B_j)$	$m_1(A_2) \cdot m_2(B_j)$	\dots	$m_1(A_i) \cdot m_2(B_j)$	\dots	$m_1(A_k) \cdot m_2(B_j)$
\vdots	\vdots	\vdots		\vdots		\vdots
$m_2(B_l)$	$m_1(A_1) \cdot m_2(B_l)$	$m_1(A_2) \cdot m_2(B_l)$	\dots	$m_1(A_i) \cdot m_2(B_l)$	\dots	$m_1(A_k) \cdot m_2(B_l)$

TAB. 3.1 – Masses des intersections $A_i \cap B_j$

Reprenons cela à travers l'exemple 2.2 en situation d'information partielle.

Exemple 3.1. Rappelons que nous avons la fonction de masse suivante, notée cette fois-ci par m_1 :

- $m_1(\emptyset) = 0$,
- $m_1(a_1) = 0.2$,
- $m_1(a_2) = 0.3$,
- $m_1(\Theta) = 0.5$.

Supposons à présent que votre conseiller financier vous informe que, de son côté, il suppose une autre fonction de masse, m_2 , telle que :

- $m_2(\emptyset) = 0$,
- $m_2(a_1) = 0.1$,
- $m_2(a_2) = 0.6$,
- $m_2(\Theta) = 0.3$.

Si l'on veut représenter les masses des intersections des éléments focaux, on obtient le tableau suivant :

	$m_1(a_1)$	$m_1(a_2)$	$m_1(\Theta)$
$m_2(a_1)$	$m_1(a_1) \cdot m_2(a_1)$	$m_1(a_2) \cdot m_2(a_1)$	$m_1(\Theta) \cdot m_2(a_1)$
$m_2(a_2)$	$m_1(a_1) \cdot m_2(a_2)$	$m_1(a_2) \cdot m_2(a_2)$	$m_1(\Theta) \cdot m_2(a_2)$
$m_2(\Theta)$	$m_1(a_1) \cdot m_2(\Theta)$	$m_1(a_2) \cdot m_2(\Theta)$	$m_1(\Theta) \cdot m_2(\Theta)$

TAB. 3.2 – Exemple : masses des intersections $A_i \cap B_j$

Supposons à présent qu'on s'intéresse à un sous-ensemble D de Θ . La masse de ce sous-ensemble peut correspondre à plusieurs intersections de lignes et de colonnes ou, autrement dit, à plusieurs cellules du tableau. La masse de D sera alors donnée par :

$$m(D) = \sum_{\substack{i,j \\ A_i \cap B_j = D}} m_1(A_i) m_2(B_j). \quad (3.1)$$

Le problème de ce schéma est qu'il peut attribuer une crédibilité non nulle à l'ensemble vide, ce qui contredit les règles de base pour les fonctions de crédibilité. En effet, A_i et B_j sont des éléments focaux, donc $m_1(A_i) > 0$ et $m_2(B_j) > 0$. Leur produit sera par conséquent forcément non nul et cela même lorsque $A_i \cap B_j = \emptyset$, c'est-à-dire lorsque leur réalisation simultanée est impossible. Nous aurons dans ce cas :

$$\sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}} m_1(A_i)m_2(B_j) > 0. \quad (3.2)$$

Pour remédier à ce problème, il faut écarter tous les produits $m_1(A_i)m_2(B_j)$ tels que $A_i \cap B_j = \emptyset$. En effet, si $A_i \cap B_j = \emptyset$, cela implique que A_i et B_j sont contradictoires, c'est-à-dire qu'ils ne peuvent se réaliser simultanément. Quant aux produits restants $m_1(A_i)m_2(B_j)$ tels que $A_i \cap B_j \neq \emptyset$, ils doivent être réajustés ou normalisés. En effet, n'oublions pas que ces produits représentent une nouvelle fonction de masse ; leur somme doit par conséquent être égale à 1. Le réajustement s'opère en multipliant chaque produit $m_1(A_i)m_2(B_j)$ par la constante

$$K = \left(1 - \sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}} m_1(A_i)m_2(B_j) \right)^{-1}, \quad (3.3)$$

appelée **facteur d'ajustement**.

Revenons à l'exemple.

Exemple 3.2. Dans l'exemple 3.1, on constate que l'intersection de certains ensembles focaux est effectivement vide, tandis que la masse qui est attribuée à cette intersection est non nulle. Il s'agit en particulier des intersections de $\{a_1\}$ et $\{a_2\}$. Nous obtenons d'après la Table 3.2,

$$m_1(a_1) \cdot m_2(a_2) = 0.2 \cdot 0.6 = 0.12$$

et

$$m_1(a_2) \cdot m_2(a_1) = 0.3 \cdot 0.1 = 0.03.$$

Ainsi,

$$\sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}} m_1(A_i)m_2(B_j) = 0.12 + 0.03 = 0.15.$$

Par conséquent, le facteur d'ajustement par lequel doit être multiplié chacun des 7 produits de masse restants est de

$$K = \left(1 - \sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}} m_1(A_i)m_2(B_j) \right)^{-1} = (1 - 0.15)^{-1} = \frac{100}{85}.$$

Théorème 3.1. *Supposons que Bel_1 et Bel_2 soient des fonctions de crédibilité portant sur le même cadre de discernement Θ , avec les fonctions de masse m_1 et m_2 et les éléments focaux A_1, \dots, A_k et B_1, \dots, B_l respectivement.*

Supposons que

$$\sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}} m_1(A_i)m_2(B_j) < 1 \quad (3.4)$$

Alors la fonction $m : 2^\Theta \rightarrow [0, 1]$ qui est définie par $m(\emptyset) = 0$ et

$$m(D) = \frac{\sum_{\substack{i,j \\ A_i \cap B_j = D}} m_1(A_i)m_2(B_j)}{1 - \sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}} m_1(A_i)m_2(B_j)} \quad (3.5)$$

pour tous les sous-ensembles non vides $A \subset \Theta$ est une fonction de masse. Le coeur de la fonction de crédibilité donnée par m est égal à l'intersection des coeurs de Bel_1 et Bel_2 .

Ce théorème et sa preuve ont été présentés par Shafer [95].

Reprenons l'exemple 3.1 et observons la Table 3.2.

Exemple 3.3. Des neuf intersections possibles, nous en avons éliminé deux, car elle correspondaient à des événements impossibles. Observons à présent les sept intersections restantes. Nous avons en effet

- $\{a_1\} \cap \{a_1\} = \{a_1\}$,
- $\{a_1\} \cap \Theta = \{a_1\}$,
- $\{a_2\} \cap \{a_2\} = \{a_2\}$,
- $\{a_2\} \cap \Theta = \{a_2\}$,
- $\Theta \cap \{a_1\} = \{a_1\}$,
- $\Theta \cap \{a_2\} = \{a_2\}$,
- $\Theta \cap \Theta = \Theta$.

On remarque que trois intersections aboutissent à $\{a_1\}$, trois à $\{a_2\}$ et une à Θ .

En appliquant à notre exemple l'équation (3.4) du théorème 3.1, nous obtenons la fonction de masse m suivante :

$$\begin{aligned} m(a_1) &= (100)/(85) \cdot (m_1(a_1) \cdot m_2(a_1) + m_1(a_1) \cdot m_2(\Theta) + m_1(\Theta) \cdot m_2(a_1)) \\ &= (100)/(85) \cdot (0.2 \cdot 0.1 + 0.2 \cdot 0.3 + 0.5 \cdot 0.1) \\ &= (13)/(85), \end{aligned}$$

$$\begin{aligned} m(a_2) &= (100)/(85) \cdot (m_1(a_2) \cdot m_2(a_2) + m_1(a_2) \cdot m_2(\Theta) + m_1(\Theta) \cdot m_2(a_2)) \\ &= (100)/(85) \cdot (0.3 \cdot 0.6 + 0.3 \cdot 0.3 + 0.5 \cdot 0.6) \\ &= (57)/(85), \end{aligned}$$

$$\begin{aligned} m(\Theta) &= (100)/(85) \cdot (m_1(\Theta) \cdot m_2(\Theta)) \\ &= (100)/(85) \cdot (0.5 \cdot 0.3) \\ &= (15)/(85). \end{aligned}$$

La somme de toutes les masses de la fonction m donne bien évidemment 1.

La fonction de crédibilité donnée par m est appelée **la somme orthogonale** de Bel_1 et Bel_2 ; elle est notée par $Bel_1 \oplus Bel_2$. En présence d'un cas tel que

$$\sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}} m_1(A_i)m_2(B_j) = 1, \quad (3.6)$$

on dit que la somme orthogonale $Bel_1 \oplus Bel_2$ **n'existe pas** et Bel_1 et Bel_2 se contredisent clairement.

3.2 Combinaison de plusieurs fonctions de crédibilité

Examinons comment combiner non plus deux, mais plusieurs fonctions de crédibilité. Soit la collection Bel_1, \dots, Bel_n de fonctions de crédibilité. Pour trouver la somme orthogonale de cette collection, il suffit de combiner les fonctions de crédibilité deux par deux, puis de combiner chaque nouvelle fonction de crédibilité ainsi obtenue avec une autre fonction de crédibilité jusqu'à ce que toutes les fonctions de crédibilité de la collection soient incluses :

$$\begin{aligned} &Bel_1 \oplus Bel_2, \\ &(Bel_1 \oplus Bel_2) \oplus Bel_3, \\ &((Bel_1 \oplus Bel_2) \oplus Bel_3) \oplus Bel_4, \\ &\vdots \end{aligned}$$

$$(((Bel_1 \oplus Bel_2) \cdots) \oplus Bel_{n-1}) \oplus Bel_n.$$

En effet, la règle de combinaison de Dempster est **commutative** et **associative**. Dans le calcul de cette somme orthogonale, chaque nouvelle entrée de Bel_i , avec $i = 1, \dots, n$, correspond à l'addition de l'évidence véhiculée par Bel_i à l'évidence finale correspondant à la fonction de crédibilité Bel issue de cette opération. Bien entendu, il est nécessaire que la somme orthogonale obtenue à chaque niveau, donc à chaque nouvelle entrée de Bel_i , existe. L'ordre dans lequel les différentes Bel_i rentrent dans la somme orthogonale n'a pas d'importance. Cela relève de la propriété d'associativité de la règle de Dempster.

Le théorème suivant expose les cas dans lesquels une collection de fonctions de crédibilité ne peut être combinée sous forme de somme orthogonale.

Théorème 3.2. *Supposons que Bel_1, \dots, Bel_n soient des fonctions de crédibilité portant sur le même cadre de discernement Θ . Notons par $\mathcal{C}_1, \dots, \mathcal{C}_n$ leurs coeurs. Alors, les conditions suivantes sont équivalentes :*

1. $\mathcal{C}_1 \cap \cdots \cap \mathcal{C}_n = \emptyset$.
2. Il existe D_1, \dots, D_n de Θ , tels que $D_1 \cap \cdots \cap D_n = \emptyset$, mais $Bel_i(D_i) = 1$ pour $i = 1, \dots, n$.
3. Il existe des sous-ensembles D et E de Θ tels que $D \cap E = \emptyset$, de plus

$$D = F_1 \cap \cdots \cap F_r$$

et

$$E = G_1 \cap \cdots \cap G_s$$

pour certains sous-ensembles $F_1, \dots, F_r, G_1, \dots, G_s$ dont chacun bénéficie d'un degré de crédibilité de 1 pour au moins une Bel_i .

4. La tentative de construire les sommes orthogonales

$$Bel_1 \oplus Bel_2,$$

$$(Bel_1 \oplus Bel_2) \oplus Bel_3,$$

$$((Bel_1 \oplus Bel_2) \oplus Bel_3) \oplus Bel_4,$$

⋮

$$(((Bel_1 \oplus Bel_2) \cdots) \oplus Bel_{n-1}) \oplus Bel_n$$

échouera ; en d'autres termes, avant que la liste soit complétée, on sera confronté à une somme orthogonale qui n'existe pas.

5. Une tentative de combiner les Bel_i répétitivement dans n'importe quel autre ordre échouera également.

La version complète de ce théorème et sa preuve ont été présentées par Shafer [95].

Une collection de fonctions de crédibilité Bel_1, \dots, Bel_n sur un cadre de discernement Θ est appelée **combinable** si elle ne satisfait pas aux conditions du théorème 3.2. La fonction de crédibilité qui résulte de la combinaison d'une telle collection s'appelle sa **somme orthogonale** et est notée $Bel_1 \oplus \dots \oplus Bel_n$. Il découle du théorème 3.1 que le coeur de $Bel_1 \oplus \dots \oplus Bel_n$ est égal à l'intersection des coeurs des Bel_i avec $i = 1, \dots, n$.

De même qu'on avait imaginé un tableau à deux dimensions pour combiner deux fonctions de crédibilité, on peut également représenter le cas présent à l'aide d'un tableau à non plus deux, mais à n dimensions. La démarche qui consiste à éliminer les produits de masses non nuls correspondant à l'ensemble vide, de même que celle du réajustement, sont identiques dans les deux cas.

3.3 Poids du conflit

Le fait de devoir opérer un réajustement à travers le facteur K peut paraître contraignant, cependant, celui-ci s'avère être un outil précieux pour décrire le niveau de conflit ou de contradiction entre deux ou plusieurs fonctions de crédibilité.

Rappelons qu'en cas de combinaison de deux fonctions de crédibilité, Bel_1 et Bel_2 , nous nous servions d'un tableau à deux dimensions dont chaque cellule correspondait à la masse attribuée à l'intersection de deux éléments focaux. Cette masse, étant le produit de deux masses d'éléments focaux, est obligatoirement supérieure à 0 et cela même lorsque l'intersection des éléments focaux correspond à l'ensemble vide. Pour remédier à ce problème, il fallait écarter tous les produits de masses de sous-ensembles disjoints. Il est évident que plus il y aura de sous-ensembles disjoints entre les deux fonctions de crédibilité, plus la quantité de masse à écarter sera grande. Notons cette quantité

$$k = \sum_{\substack{i,j \\ A_i \cap B_j = \emptyset}} m_1(A_i)m_2(B_j). \quad (3.7)$$

Le facteur d'ajustement se présente alors comme

$$K = \frac{1}{1 - k}. \quad (3.8)$$

On peut ainsi aisément constater que K peut servir de mesure du conflit ou de la contradiction entre les deux fonctions de crédibilité. Afin de donner une échelle à ce conflit allant de 0 à $+\infty$, Shafer [95] préfère considérer le logarithme naturel de K

$$\ln(K) = \ln\left(\frac{1}{1 - k}\right) = -\ln(1 - k) \quad (3.9)$$

qu'il appelle **poids du conflit** entre Bel_1 et Bel_2 et qu'on notera $\mathcal{K}(Bel_1, Bel_2)$.

Ainsi, lorsque Bel_1 et Bel_2 ne se contredisent absolument pas ou, autrement dit, n'ont aucuns éléments focaux disjoints, $k = 0$ et par conséquent $\mathcal{K}(Bel_1, Bel_2) = 0$. De même, lorsque Bel_1 et Bel_2 se contredisent clairement, au point que tous leurs éléments focaux sont disjoints et $Bel_1 \oplus Bel_2$ n'existe pas, $k = 1$ et $\mathcal{K}(Bel_1, Bel_2) = +\infty$.

La notion de conflit peut bien entendu être généralisée à une collection Bel_1, \dots, Bel_n de fonctions de crédibilité. Dans ce cas, $ln(K)$ représente le poids du conflit entre Bel_1, \dots, Bel_n et est noté $\mathcal{K}(Bel_1, \dots, Bel_n)$.

Les poids de conflit peuvent être combinés additionnellement.

Théorème 3.3. *Supposons que Bel_1, \dots, Bel_n soient des fonctions de crédibilité sur un cadre de discernement Θ et supposons également que $Bel_1 \oplus \dots \oplus Bel_n$ existe. Alors*

$$\mathcal{K}(Bel_1, \dots, Bel_n, Bel_{n+1}) = \mathcal{K}(Bel_1, \dots, Bel_n) + \mathcal{K}(Bel_1 \oplus \dots \oplus Bel_n, Bel_{n+1}). \quad (3.10)$$

Ce théorème et sa preuve ont été présentés par Shafer [95].

Chapitre 4

Conditionnement des fonctions de crédibilité

La règle de combinaison de Dempster a montré comment assimiler la crédibilité actuelle, exprimée par Bel_1 , avec la crédibilité relevant d'une nouvelle source d'évidence, exprimée par Bel_2 , pour obtenir la crédibilité finale résultant de cette combinaison, exprimé par $Bel_1 \oplus Bel_2$.

Supposons à présent que l'effet de la nouvelle évidence sur le cadre de discernement Θ est d'établir un certain sous-ensemble $B \subset \Theta$ avec certitude. Alors, Bel_2 bénéficiera d'un degré de crédibilité total égal à 1 pour la proposition correspondante à B ainsi que pour toutes les propositions impliquées par celle-ci :

$$Bel_2(A) = \begin{cases} 1 & \text{si } B \subset A \\ 0 & \text{si } B \not\subset A. \end{cases} \quad (4.1)$$

Notons que dans ce cas, le sous-ensemble B est le seul élément focal de Bel_2 et que sa masse est également égale à 1. Une telle fonction Bel_2 est combinable avec Bel_1 aussi longtemps que $Bel_1(\bar{B}) < 1$.

Théorème 4.1. *Supposons que Bel_2 soit définie par l'équation (4.1) et que Bel_1 soit une autre fonction de crédibilité sur Θ . Alors, Bel_1 et Bel_2 sont combinables si et seulement si $Bel_1(\bar{B}) < 1$. Si Bel_1 et Bel_2 sont combinables, notons par $Bel_1(\cdot|B)$ la somme orthogonale $Bel_1 \oplus Bel_2$ et par Pl_1^* et $Pl_1^*(\cdot|B)$ les fonctions de plausibilité pour Bel_1 et pour $Bel_1 \oplus Bel_2$ respectivement. Alors,*

$$Bel_1(A|B) = \frac{Bel_1(A \cup \bar{B}) - Bel_1(\bar{B})}{1 - Bel_1(\bar{B})} \quad (4.2)$$

et

$$Pl_1^*(A|B) = \frac{Pl(A \cap B)}{Pl_1^*(B)} \quad (4.3)$$

pour tout $A \subset \Theta$.

Ce théorème et sa preuve ont été présentés par Shafer [95].

L'équation (4.3) porte le nom de **règle de Dempster de conditionnement**.

Chapitre 5

Critiques de la théorie mathématique de l'évidence

On trouve dans la littérature de nombreuses critiques de la théorie mathématique de l'évidence en général et de la règle de combinaison de Dempster en particulier. Il est vrai que celle-ci prend une place centrale dans la théorie de Dempster et Shafer.

Les critiques formulées se concentrent essentiellement sur trois aspects :

1. l'aspect théorique,
2. l'aspect mathématique et
3. l'aspect empirique.

1. Aspect théorique

Les critiques de la théorie mathématique de l'évidence d'un point de vue théorique lui reprochent le manque de clarté de son cadre formel. Par exemple, Voorbraak [119] critique la formulation initiale de Shafer quant aux conditions d'application de la règle de combinaison de Dempster. Il les estime beaucoup trop vagues. Selon Voorbraak, l'indépendance des cadres de discernement devrait être imposée comme condition nécessaire à l'application de la règle de combinaison de Dempster. Cette indépendance n'est pas imposée dans la théorie mathématique de l'évidence. Par ailleurs, la plupart des exemples de Shafer [95] ne la respectent pas. Une alternative tenant compte des dépendances a été développée par Dubois et Prade [24], présentant ainsi une généralisation de la règle de combinaison de Dempster.

2. Aspect mathématique

Kyburg [65] a montré que les fonctions de crédibilité de la théorie mathématique de l'évidence sont des cas particuliers d'enveloppes inférieures d'un ensemble convexe de mesures de probabilité admissibles. Ainsi, le résultat de l'application de la règle de combinaison de Dempster est comparable à un ensemble de mesures de probabilités admissibles vérifiant des contraintes ou appartenant à un ensemble donné.

Fabiani [28] soutient que dans le cas de deux fonctions de crédibilité bayésiennes, Bel_1 et Bel_2 , la règle de combinaison de Dempster est identique à la combinaison de mesures de probabilités en une mesure de probabilité produit sur l'espace produit des deux espaces de probabilité qui sont respectivement associés à Bel_1 et Bel_2 . Cependant, pour les situations en dehors de ce cas précis, il cite les travaux de Lemmer [48], [69] qui, considérant les fonctions de crédibilité en termes d'enveloppes inférieures d'ensembles convexes de mesures de probabilité, montre que la règle de combinaison de Dempster ne peut s'appliquer à des données de nature statistiques. D'après Lemmer, cette règle conduit à des résultats incohérents si elle est appliquée à des taux d'erreurs ou à des fréquences d'arrivée.

3. Aspect empirique

Les critiques sur le plan empirique proviennent toutes d'auteurs ayant appliqué la règle de combinaison de Dempster dans leurs travaux. Ainsi, Dezert, Perneel, Van De Velde et Acheroy [22], [89] ont révélé une faiblesse importante de la règle de Dempster. Celle-ci réside dans la sensibilité de la règle aux petites valeurs de masse au voisinage de 0. Perneel, Van De Velde et Acheroy [89] se plaignent également du grand temps de calcul que la règle de Dempster implique, contrairement à d'autres méthodes d'inférence. En outre, Appriou [6], [7] ajoute que l'absence de sémantique claire ne permet pas de se guider vers une méthode systématique souhaitable d'un point de vue empirique. Enfin, plusieurs auteurs ont demandé (voir [23], [122] et [123]) à ce que le cadre d'application de règle de Dempster soit amendé. Concernant l'étude de stabilité menée par Wu, Apostakis et Okrent [122], Dubois et Prade [24] soutiennent que la règle de Dempster est discontinue au voisinage de conflit total et, dans ce cas, elle rejette presque toute l'information disponible.

Malgré ces critiques, la théorie mathématique de l'évidence est toutefois plébiscitée par de nombreux auteurs. Dans le Chapitre 1, ont été mentionnées de nombreuses applications de cette théorie à des problèmes économiques ainsi que les avantages qui en résultent. Par ailleurs, notre recherche s'est basée essentiellement sur la théorie mathématique des indications, présentée dans la Partie 2, qui offre une solution à la plupart de ces critiques. De ce fait, celles-ci ne se sont pas révélées pertinentes pour ce travail.

Chapitre 6

Problème du casino

Le problème du casino est un exemple à la fois simple et intéressant. Il sera repris tout au long de ce travail, afin d'illustrer certaines lacunes tant de la théorie mathématique de l'évidence, que de la théorie mathématique des indications.

Exemple 6.1. Imaginons le client d'un casino qui désire jouer à la roulette. Sa question porte évidemment sur le nombre qui sortira gagnant à la fin d'un tour de roulette. Rappelons que dans ce jeu un seul nombre est gagnant à la fin d'un tour. L'ensemble des réponses Θ est défini dans ce cas par

$$\Theta = \{0, 1, 2, \dots, 36\}$$

et représente tous les nombres qui peuvent sortir gagnants à la fin d'un tour dans un jeu de roulette. Il s'agit pour l'instant de toute l'information dont on dispose. Nous sommes en présence d'une fonction de crédibilité Bel_1 , qui représente le degré de crédibilité que nous accordons à chacun des nombres de sortir gagnant. Comme la table de roulette est supposée équilibrée, alors les numéros sortants suivent une distribution uniforme. Cela implique que Bel_1 est une fonction de crédibilité bayésienne et que

$$Bel_1(i) = \frac{1}{37},$$

avec $i = 0, \dots, 36$.

Il s'agit d'un cas tout à fait classique, avec une seule fonction de crédibilité Bel_1 . Que ce passerait-il, si nous obtenions une information supplémentaire concernant le problème qui nous intéresse ? Dans un cas standard, nous la modéliserons par Bel_2 et la combinerions avec Bel_1 , selon la règle de combinaison de Dempster. Toutefois, comment procéder si on ne peut modéliser cette information supplémentaire par une fonction de crédibilité, car celle-ci influence celle véhiculée par Bel_1 et la remet même en question ? Par exemple, si quelqu'un vous informe qu'une table de roulette sur dix est truquée ?

La théorie mathématique de l'évidence ne permet pas de donner une solution à ce problème. Elle ne permet pas de modéliser des situations de remise en question de l'information par des fonctions de crédibilité. C'est pourquoi, on développera dans la Partie 3 une solution à ce problème qu'on illustrera par plusieurs exemples.

Deuxième partie

Théorie mathématique des indications

La théorie mathématique des indications a été développée par Kohlas et Monney [54] [59] ; elle a vu le jour relativement récemment, vers 1990. Ces deux chercheurs proposent une approche alternative de l'étude de l'évidence. Ils introduisent le concept d'**indication** (en anglais «hint») en tant que modèle explicite d'une masse d'information incertaine et imprécise. Une indication inclut un ensemble de différentes interprétations possibles de l'information disponible. Elle est utilisée pour juger à quel niveau elle supporte une hypothèse donnée, ainsi qu'à quel niveau l'hypothèse serait plausible à la lumière de cette indication. La mesure de ces deux niveaux est donnée par le **degré de support** et par le **degré de plausibilité** respectivement. En effet, dans leur théorie des indications, Kohlas et Monney préfèrent le terme «support» au terme «crédibilité». Ils justifient ce choix par le fait que les différentes interprétations possibles d'une indication nous procurent des arguments en faveur ou contre les hypothèses. Ainsi, les indications peuvent plus ou moins supporter les hypothèses et les rendre plus ou moins plausibles. D'ailleurs, on trouve un point de vue similaire dans un ouvrage de Pearl [85].

Dans l'optique du modèle des indications, la théorie de l'évidence est considérée non pas comme une théorie de la crédibilité, mais comme une théorie du support. En effet, dans la théorie proposée par Shafer [95], la notion de fonction de crédibilité définie axiomatiquement est assez distante de l'information originelle, de la source de l'évidence, ou autrement dit, de l'indication. De plus, la théorie mathématique de l'évidence de Shafer, telle qu'elle a été présentée en 1976, se limite à des cadres de discernement finis. Dans un article, Shafer [97] a également étudié le cas des cadres de discernement infinis, où les fonctions de crédibilité sont axiomatiquement définies comme des capacités de Choquet monotones et d'ordre infini. En utilisant le théorème de la représentation intégrale de Choquet [17], [18], il dérive une allocation de la probabilité pour les fonctions de crédibilité ; cependant, celle-ci pose quelques problèmes quant à l'application de la règle de combinaison de Dempster.

Pour cette raison, Kohlas et Monney se tournent vers le concept originel de Dempster des applications à valeurs multiples et l'utilisent comme point de départ. En effet, les indications permettent une définition logique et claire de la règle de combinaison de Dempster. A travers cette règle, que Kohlas et Monney définissent de manière plus générale, on peut combiner des indications **stochastiquement dépendantes ou indépendantes**, privant ainsi d'objet l'une des critiques majeures de la théorie mathématique de l'évidence.

Une indication représente une part d'information, relative à une certaine question. Elle inclut un ensemble de différentes interprétations possibles de cette information qui ne sont pas nécessairement également probables. Chacune de ces interprétations détermine des propositions, impliquées et possibles dans l'ensemble des réponses à la question, appelé **cadre de discernement** ; cela nous amène à un concept plus général que l'application à valeur multiple. Le concept de l'indication correspond également dans son essence mathé-

matique au modèle de message codé, proposé par Shafer [99] comme modèle de mesure des fonctions de crédibilité.

A noter enfin qu'il est possible de faire le parallèle entre la théorie mathématique des indications et celle de la fiabilité. On y trouve une analogie importante. En effet, la théorie mathématique des indication peut être conçue, ainsi que l'a montré Kohlas [53], comme une théorie de la fiabilité du raisonnement avec des arguments qui ne sont pas totalement fiables. Un autre lien peut être établi entre la théorie mathématique des indications et l'ATMS (Assumption Based Truth Maintenance). Dans cette conception, les indications sont utilisées pour différentes hypothèses ainsi que pour leur vraisemblance (voir l'ouvrage de Laskey et Lehner [66]).

Chapitre 7

Concepts de base

7.1 Imprécision, incertitude et contradiction

Une indication n'est rien d'autre qu'une part d'information relative à une certaine question. Le problème est de représenter de manière quantifiée cette part d'information, précise ou **imprécise**, claire ou **ambiguë**, afin de s'approcher le plus possible de la réponse correcte à la question posée. C'est justement l'objet de la théorie mathématique des indications.

Notons que les indications sont la plupart du temps imprécises, en ce sens qu'elles ne peuvent pas conduire à une réponse unique à la question posée. De même, elles sont souvent ambiguës, en ce sens qu'il n'est pas toujours clair de quelle manière faut-il interpréter l'information qu'elles contiennent. De plus, une indication peut être **contradictoire** au niveau **interne** aussi bien qu'au niveau **externe**. Une contradiction interne s'exprime par le fait que les différentes interprétations possibles d'une indication peuvent aboutir à des réponses partiellement ou totalement contradictoires. Par opposition, il y a contradiction externe lorsque les interprétations de plusieurs indications mènent à des réponses mutuellement exclusives ou contradictoires.

Bien que les indications, comme point de départ d'une théorie mathématique, puissent sembler mal appropriées, vu qu'elles ne supposent pas les propriétés de consistance souhaitées pour une analyse simplifiée, leur atout majeur réside dans le fait qu'elles représentent la réalité **telle qu'elle est** et non pas telle qu'on suppose qu'elle est dans un objectif de faciliter son étude. Le modèle mathématique des indications développé par Kohlas et Monney est justement conçu de manière à représenter fidèlement la réalité et les informations qui en découlent tout en faisant face aux difficultés inhérentes à la complexité de celle-ci.

Comme nous l'avons déjà souligné, la théorie mathématique des indications ne se limite pas, contrairement à la théorie mathématique de l'évidence de Shafer [95], à des cadres

finis. Soit Θ un ensemble quelconque (cadre de discernement), fini ou non, dont les éléments θ représentent les réponses possibles à une certaine question sous considération. On suppose qu'un des éléments de Θ , θ^* représente la vraie réponse à cette question ; θ^* est cependant inconnue. Notons au passage qu'il s'agit là d'une **hypothèse d'un monde fermé**. L'hypothèse d'un monde ouvert, qui admet la possibilité que la vraie réponse θ^* se trouve en dehors de Θ , peut cependant également être envisagée (voir Smets [106]). Evidemment, l'ensemble vide représente la **proposition impossible** et Θ , la proposition qui est **toujours vraie**.

7.2 Interprétations et mesure de probabilité

Soit Ω l'ensemble des interprétations possibles de l'information véhiculée par une indication. On suppose cet ensemble fini, c'est-à-dire

$$\Omega = \{\omega_1, \dots, \omega_m\}. \quad (7.1)$$

Il existe une seule et unique interprétation qui soit correcte, ω_j ; il est cependant impossible de l'identifier avec certitude. On suppose par contre que la probabilité $p(\omega_j)$ que ω_j soit correcte est connue pour chaque interprétation ω_j avec $j = 1, \dots, m$. Il n'est pas nécessaire de considérer les interprétations de probabilité nulle. De ce fait, elles seront éliminées. Soit $p(\omega_j) > 0$ avec $j = 1, \dots, m$ la probabilité que l'interprétation ω_j soit l'interprétation correcte. On suppose que

$$\sum_{j=1}^m p(\omega_j) = 1, \quad (7.2)$$

étant donné que parmi toutes les interprétations possibles, une doit être correcte. Il a été ainsi introduit un espace de probabilité (Ω, P) qui représente l'**incertitude** rattachée à l'indication. Les interprétations ne sont pas nécessairement toutes équiprobables.

Il faut faire à ce stade une remarque importante. La théorie mathématique des indications n'aborde absolument pas le problème de l'estimation de la probabilité que ω_j soit l'interprétation correcte. Celle-ci est simplement supposée connue, comme il vient d'être mentionné.

7.3 Ensembles focaux

Soit ω_j l'interprétation correcte de l'indication. Supposons qu'il soit possible de prouver que θ^* se trouve dans un certain sous-ensemble $\Gamma(\omega_j)$ de l'ensemble des réponses, c'est-à-dire du cadre de discernement, Θ . Γ est une application multivaluée de Ω dans Θ ou plus

précisément une application de Ω dans l'ensemble de tous les sous-ensembles de Θ , noté $\mathcal{P}(\Theta)$, définie par :

$$\Gamma : \begin{cases} \Omega \rightarrow \mathcal{P}(\Theta) \\ \omega \rightarrow \Gamma(\omega) \subset \Theta. \end{cases} \quad (7.3)$$

Notons par contre que même si nous savons que ω_j est l'interprétation correcte de l'indication, cela ne nous permet pas pour autant de fixer la réponse correcte de manière précise, mais uniquement de restreindre l'ensemble des possibilités à un sous-ensemble de Θ . C'est justement ce que l'on appelle l'**imprécision** de l'indication. Par contre, si ω_j n'était pas l'interprétation correcte, on ne pourrait que conclure que θ^* se trouve dans l'union de tous les autres sous-ensembles de $\Gamma(\omega_i)$ différents de $\Gamma(\omega_j)$. Autrement dit, dans ce cas,

$$\theta^* \in \bigcup_{i \neq j} \Gamma(\omega_i). \quad (7.4)$$

Comme dans la théorie mathématique de l'évidence, les sous-ensembles $\Gamma(\omega_j)$ de Θ portent le nom d'**ensemble focaux** de l'indication et ne sont pas nécessairement tous différents. De plus, la vraie réponse θ^* se trouve avec certitude dans la réunion des ensembles focaux, appelé **coeur** de l'indication et noté par \mathcal{C} , c'est-à-dire

$$\theta^* \in \mathcal{C} = \bigcup_j \Gamma(\omega_j). \quad (7.5)$$

Supposant ω_j comme l'interprétation correcte, on pourrait se demander quels sont les sous-ensembles de Θ , autrement dit les propositions, impliquées par cette interprétation. Soit \mathcal{S} la famille de ces sous-ensembles. L'objectif étant de trouver la réponse correcte, donc de localiser θ^* dans Θ , il est nécessaire de déterminer $\Gamma(\omega_j)$ comme le plus petit sous-ensemble de Θ pour lequel il est possible de prouver qu'il contient la réponse correcte. Nous n'aurions, en effet, aucun intérêt à choisir par exemple $\Gamma(\omega_j) = \Theta$, car, tout en étant sûrs que cet ensemble contient θ^* , ce choix ne fournirait aucune information. En déterminant $\Gamma(\omega_j)$ le plus précisément possible, nous restreignons les possibilités à un sous-ensemble de Θ . L'information en résultant est caractérisée bien sûr par un certain niveau d'incertitude puisque nous n'avons pas localisé l'emplacement précis de la réponse, mais juste un sous-ensemble parmi les emplacements possibles. Ce niveau d'incertitude correspond à l'imprécision rattachée à l'indication. Dans le cas d'une indication tout-à-fait précise, nous serions capables de réduire $\Gamma(\omega_j)$ à un singleton de Θ . Inversement, dans le cas d'une indication tout-à-fait imprécise, nous ne pourrions réduire $\Gamma(\omega_j)$ à un sous-ensemble de Θ et l'ensemble focal de ω_j serait donné par l'ensemble Θ lui-même. Enfin, une telle interprétation ω_j dont l'ensemble focal ne réduit nullement Θ à un sous-ensemble, c'est-à-dire une interprétation qui ne permet pas de restreindre les réponses possibles, est appelée **interprétation vide**.

7.4 Indications

Une **indication** sur un ensemble de réponses Θ est représentée par un quadruplet $\mathcal{H} = (\Omega, P, \Gamma, \Theta)$ où :

- Ω est l'ensemble des interprétations et est fini,
- P est une mesure de probabilité sur cet ensemble, définie par sa fonction de probabilité p ,
- Γ est une application multivaluée de Ω dans Θ
- et enfin, Θ est le cadre de discernement, ou autrement dit, l'ensemble des réponses.

Exemple 7.1. Prenons comme exemple l'évolution du cours d'une action cotée en bourse. On est intéressé à prévoir si le lendemain le cours va

- augmenter (situation symbolisée par a),
- baisser (situation symbolisée par b),
- ou finalement rester stable (situation symbolisée par c).

Autrement dit, l'ensemble des réponses à notre question est donné par $\Theta = \{a, b, c\}$. Un ami analyste financier nous dit que, d'après son analyse, le cours va sûrement augmenter le lendemain, indication représentée par \mathcal{H} . Les interprétations possibles de cette indication sont :

- ω_1 : l'ami a fait une analyse correcte,
- ω_2 : l'ami s'est trompé dans son analyse.

Si l'ami a fait une analyse correcte, alors le cours de l'action augmentera effectivement et $\Gamma(\omega_1) = \{a\}$. Si par contre son analyse comporte des erreurs, ses prévisions ne sont plus fiables et dans ces conditions nous n'avons plus aucune idée sur l'évolution du cours de l'action. Dans ce dernier cas, il n'est pas possible de prouver que la réponse correcte θ^* se trouve dans un ensemble plus petit que le cadre de discernement lui-même et par conséquent $\Gamma(\omega_2) = \Theta = \{a, b, c\}$. La probabilité $p(\omega_1)$ de la première interprétation est égale à la fiabilité de l'analyse de l'ami financier, quant à la probabilité de la deuxième interprétation $p(\omega_2)$, elle est égale à $1 - p(\omega_1)$.

On peut distinguer un certain nombre de classes d'indications particulières (voir [54], [78]) :

1. Une indication est appelée **vide** lorsque toutes ses interprétations ω_j sont vides, c'est-à-dire lorsqu'elle ne fournit aucune information en rapport avec la réponse à la question posée. Dans ce cas, $\Gamma(\omega_j) = \Theta$ pour tout $\omega_j \in \Omega$.
2. Une indication est appelée **simple** s'il existe un $F \subset \Theta$ tel que soit $\Gamma(\omega_j) = F$, où $F \subset \Theta$, soit $\Gamma(\omega_j) = \Theta$ pour chaque interprétation ω_j .

3. Une indication est appelée **sûre** lorsque $\Gamma(\omega_j) = F$, où $F \subset \Theta$, pour toutes les indications ω_j , étant donné que dans ce cas θ^* se trouve avec certitude dans F .
4. Une indication est appelée **consonnante** lorsque les ensembles focaux sont emboîtés, par exemple $\Gamma(\omega_1) \subset \dots \subset \Gamma(\omega_m)$. Dans ce cas, les interprétations vont dans la même direction, mais selon des degrés de précision divers. Dans cet exemple, ω_1 sera la plus précise, $\Gamma(\omega_1)$ étant le plus petit.
5. Une indication est appelée **précise** lorsque tous les ensembles focaux sont des singletons. Dans ce cas, Γ n'est plus une application multivaluée et peut être entendue comme une variable aléatoire. De ce fait, la théorie des variables aléatoires peut être considérée comme un cas particulier de celle des indications.
6. Une indication est appelée **bayésienne** lorsque tous les ensembles focaux sont disjoints. Ainsi, chaque indication précise est également bayésienne du fait que deux singletons sont forcément disjoints.

Notons encore que

- si une indication est vide, alors elle est
 - simple,
 - sûre et
 - consonnante ;
- si une indication est sûre, alors elle est
 - simple et
 - consonnante.

Exemple 7.2. Dans l'exemple 7.1, l'indication \mathcal{H} donnée par l'ami financier est simple et consonnante, mais n'est ni vide, ni sûre, ni précise, ni bayésienne.

Chapitre 8

Support et plausibilité

8.1 Famille des propositions impliquées

La plupart du temps, on est d'avantage intéressé par une certaine proposition discernée par le cadre Θ plutôt que par la réponse précise à la question considérée. Soit \mathcal{S} la famille des sous-ensembles de Θ impliqués par l'interprétation correcte d'une indication et soit H un sous-ensemble de Θ représentant cette proposition. A présent, on désire savoir jusqu'à quel niveau une indication \mathcal{H} relative à Θ prouve la proposition H ; autrement dit, jusqu'à quel niveau l'indication supporte l'hypothèse selon laquelle H contient la réponse correcte.

Soit l'indication $\mathcal{H} = (\Omega, P, \Gamma, \Theta)$ l'indication considérée. Si ω est l'interprétation correcte de \mathcal{H} , alors la réponse correcte se trouve obligatoirement dans $\Gamma(\omega)$. Une proposition $H \subseteq \Theta$ est dite **impliquée** par l'interprétation ω de l'indication \mathcal{H} si H contient $\Gamma(\omega)$. Cette implication résulte du fait que si $\theta^* \in \Gamma(\omega)$, alors $\theta^* \in H$. On dit dans ce cas que l'interprétation ω **supporte** ou prouve l'hypothèse selon laquelle $\theta^* \in H$. Mathématiquement, la **famille des propositions impliquées** par ω se présente comme

$$\mathcal{S}(\omega) = \{H \subseteq \Theta : \Gamma(\omega) \subseteq H\}. \quad (8.1)$$

De même, du point de vue des interprétations cette fois, on peut écrire que H est impliquée si et seulement si l'interprétation correcte se trouve dans l'ensemble

$$u(H) = \{\omega \subseteq \Omega : \Gamma(\omega) \subseteq H\}. \quad (8.2)$$

Pourtant, tant que l'interprétation correcte est inconnue, il est incertain si la proposition H peut ou non être inférée depuis l'indication. Toutefois, la probabilité de l'événement (8.2), exprimée par $P(u(H))$, mesure la vraisemblance que H puisse effectivement être inférée.

Exemple 8.1. Par rapport à l'exemple 7.1 de l'évolution du cours de l'action, la famille des propositions (sous-ensembles) impliquées par ω_1 est donnée par tous les sous-ensembles qui contiennent $\Gamma(\omega_1)$, donc $\{a\}$. Par conséquent,

$$\mathcal{S}(\omega_1) = \{\{a\}, \{a, b\}, \{a, c\}, \{a, b, c\}\}.$$

De même, la famille des propositions impliquées par ω_2 est donnée par

$$\mathcal{S}(\omega_2) = \{a, b, c\}.$$

En effet, le cadre de discernement Θ est le seul ensemble qui contienne $\Gamma(\omega_2)$.

Du point de vue des interprétations cette fois-ci, on obtient en appliquant l'équation (8.2) :

H	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$u(H)$	\emptyset	ω_1	\emptyset	\emptyset	ω_1	ω_1	\emptyset	Ω

TAB. 8.1 – Exemple : interprétations impliquant H

Voici quelques propriétés de la famille $\mathcal{S}(\omega)$:

- 1 Si $H \in \mathcal{S}(\omega)$ et $H \subseteq H'$, alors $H' \in \mathcal{S}(\omega)$.
- 2 Si $H_i \in \mathcal{S}(\omega)$ pour $i = 1, \dots, n$, alors $\bigcap_{i=1}^n H_i \in \mathcal{S}(\omega)$.
- 3 $\Theta \in \mathcal{S}(\omega)$ et $\emptyset \notin \mathcal{S}(\omega)$.

Notons que si \mathcal{S} représente la famille de toutes les propositions (sous-ensembles) impliquées par l'interprétation correcte, alors elle peut être considérée comme un **filtre**. En effet, chaque ensemble contenant un sous-ensemble impliqué le sera également à son tour, de même que si deux sous-ensembles sont impliqués, leur intersection l'est aussi.

8.2 Famille des propositions possibles

Il existe une autre famille intéressante de sous-ensembles de Θ . Il s'agit de la famille \mathcal{P} de tous les sous-ensembles possibles sous l'interprétation ω . Une proposition $H \subseteq \Theta$ est appelée **possible** ou **plausible** sous l'interprétation ω de \mathcal{H} si son complément \bar{H} n'est pas impliqué par ω . Mathématiquement, la **famille des propositions possibles** sous ω se présente comme

$$\mathcal{P}(\omega) = \{H \subseteq \Theta : \Gamma(\omega) \cap H \neq \emptyset\}. \quad (8.3)$$

De même, du point de vue des interprétations cette fois, nous pouvons écrire

$$v(H) = \{\omega \subseteq \Omega : \Gamma(\omega) \cap H \neq \emptyset\}. \quad (8.4)$$

Enfin, la famille \mathcal{P} possède les propriétés suivantes :

1. Si $H \in \mathcal{P}(\omega)$ et $H \subseteq H'$, alors $H' \in \mathcal{P}(\omega)$.
2. Si $H_i \in \mathcal{S}(\omega)$ pour $i = 1, \dots, n$, alors $\bigcup_{i=1}^n H_i \in \mathcal{P}(\omega)$.
3. $\Theta \in \mathcal{P}(\omega)$ et $\emptyset \notin \mathcal{P}(\omega)$.

Exemple 8.2. Par rapport à l'exemple 7.1 de l'évolution du cours de l'action, la famille des propositions (sous-ensembles) possibles sous ω_1 est donnée par tous les sous-ensembles de Θ dont les compléments ne sont pas impliqués par ω_1 , c'est-à-dire dont les compléments ne contiennent pas $\Gamma(\omega_1)$, donc $\{a\}$. Par conséquent,

$$\mathcal{P}(\omega_1) = \{\{a\}, \{a, b\}, \{a, c\}, \{a, b, c\}\}.$$

De même, la famille des propositions dont les compléments ne sont pas impliqués par ω_2 est donnée par :

$$\mathcal{P}(\omega_2) = \{\{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}\}.$$

En effet, de tous les sous-ensembles possibles de Θ , le seul dont le complément contient $\Gamma(\omega_2)$, donc le seul qui sera exclu de $\mathcal{P}(\omega_2)$, est \emptyset , vu que $\Gamma(\omega_2)$ est égal au cadre de discernement Θ .

Du point de vue des interprétations cette fois-ci, nous pouvons écrire en nous basant sur l'équation (8.4) :

H	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$v(H)$	\emptyset	Ω	ω_2	ω_2	Ω	Ω	ω_2	Ω

TAB. 8.2 – Exemple : interprétations rendant H possible

Il existe dans la théorie mathématique des indication la relation suivante entre les familles des propositions impliquées et des propositions possibles sous ω :

$$\mathcal{S}(\omega) \subseteq \mathcal{P}(\omega), \tag{8.5}$$

étant donné que si $H \in \mathcal{S}(\omega)$ alors $\bar{H} \notin \mathcal{S}(\omega)$. Cette relation s'explique par le fait que le sous-ensemble des interprétations pour lesquelles il est possible de prouver que θ^* se trouve dans H est non seulement plus petit, mais également inclus dans le sous-ensemble des interprétations qui ne contredisent pas cette hypothèse. En effet, une proposition doit d'abord être possible avant d'être nécessaire. Par exemple, si ω est une interprétation vide, alors aucune proposition n'est impliquée par elle ; par contre, tout sous-ensemble non vide de Θ est possible sous elle.

8.3 Fonctions de support et de plausibilité

Le **degré de support** d'une proposition H selon l'indication \mathcal{H} est donné par la vraisemblance que H soit inférée (voir équation (8.2)) :

$$sp(H) = P(u(H)). \quad (8.6)$$

Il peut également être défini à partir de la famille \mathcal{S} des sous-ensembles de Θ impliqués par l'interprétation correcte :

$$sp(H) = \sum_{j=1}^m \{p(\omega_j) : H \in \mathcal{S}(\omega_j)\}. \quad (8.7)$$

Enfin, en revenant à l'équation (8.1), on peut aussi l'exprimer par :

$$sp(H) = \sum_{j=1}^m \{p(\omega_j) : \Gamma(\omega_j) \subseteq H\}. \quad (8.8)$$

Le degré de support est défini pour tout sous-ensemble H de Θ . Ainsi apparaît la fonction

$$sp : \mathcal{P}(\Theta) \rightarrow [0, 1], \quad (8.9)$$

appelée **fonction de support** associée à l'indication.

De même, le **degré de plausibilité** d'une proposition H selon l'indication \mathcal{H} est donné par :

$$pl(H) = P(v(H)). \quad (8.10)$$

Il peut également être défini à partir de la famille \mathcal{P} des sous-ensembles de Θ possibles sous l'interprétation correcte :

$$pl(H) = \sum_{j=1}^m \{p(\omega_j) : H \in \mathcal{P}(\omega_j)\}. \quad (8.11)$$

En revenant à l'équation (8.3), on peut aussi l'exprimer par :

$$pl(H) = \sum_{j=1}^m \{p(\omega_j) : \Gamma(\omega_j) \cap H \neq \emptyset\}. \quad (8.12)$$

Enfin, le degré de plausibilité est également défini pour tout sous-ensemble H de Θ . Ainsi apparaît la fonction

$$pl : \mathcal{P}(\Theta) \rightarrow [0, 1], \quad (8.13)$$

appelée **fonction de plausibilité** associée à l'indication. Notons que pour ne pas confondre la notion de **fonction de plausibilité** de la théorie de l'évidence (plus restrictive) avec celle de la théorie des indications, elles seront notées Pl et pl respectivement.

Exemple 8.3. Revenons à l'exemple 7.1 de l'évolution du cours de l'action. Supposons que la probabilité que l'ami fasse une analyse correcte, vu ses capacités, est de 0.7, autrement dit $p(\omega_1) = 0.7$. Par conséquent, la probabilité qu'il se trompe est donnée par $p(\omega_2) = 1 - 0.7 = 0.3$. On est en mesure à présent de définir la fonction de support liée à l'indication \mathcal{H} de l'ami. Il s'agit de définir sp pour tous les sous-ensembles (propositions) de Θ . Voici leur liste :

$$\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \{a, b, c\}.$$

Chacun de ces sous-ensembles est également appelé **proposition** ou encore **hypothèse**, car c'est justement dans cette optique que nous les considérons. Nous essayerons ainsi de mesurer le degré de support de l'hypothèse selon laquelle le cours de l'action restera stable $sp(c)$ ou encore celui de l'hypothèse inverse, c'est-à-dire qu'il ne reste pas stable, mais sans spécifier s'il augmente ou baisse $sp(\{a, b\})$.

À présent, déterminons la fonction de support liée à l'indication \mathcal{H} en calculant le degré de support pour tous les sous-ensembles de Θ . Pour ce faire, il suffit d'appliquer au choix une des deux équations, (8.7) ou (8.8). On obtient les valeurs suivantes :

H	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$sp(H)$	0	0.7	0	0	0.7	0.7	0	1

TAB. 8.3 – Exemple : degrés de support

Déterminons également la fonction de plausibilité liée à l'indication \mathcal{H} . Pour calculer les degrés de plausibilité de tous les sous-ensembles de Θ , nous appliquons soit l'équation (8.11), soit l'équation (8.12). Voici les résultats obtenus :

H	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$pl(H)$	0	1	0.3	0.3	1	1	0.3	1

TAB. 8.4 – Exemple : degrés de plausibilité

8.4 Propriétés élémentaires des fonctions de support et de plausibilité

Voici à présent quelques propriétés importantes des fonctions de support et de plausibilité, présentées par Kohlas et Monney [62] :

Théorème 8.1.

1. $sp(\emptyset) = pl(\emptyset) = 0$ et $sp(\Theta) = pl(\Theta) = 1$.
2. $sp(\cdot) \leq pl(\cdot)$.
3. $sp(H) = 1 - pl(\bar{H})$ et $pl(H) = 1 - sp(\bar{H})$ pour tout $H \subseteq \Theta$.
4. Si $A \subseteq B$, alors $sp(A) \leq sp(B)$ et $pl(A) \leq pl(B)$.
5. Soit $H_1 \supseteq H_2 \supseteq \dots$ une suite d'ensembles dans Θ . Alors

$$sp\left(\bigcap_{i=1}^{\infty} H_i\right) = \lim_{i \rightarrow \infty} sp(H_i). \quad (8.14)$$

6. Soit $H_1 \subseteq H_2 \subseteq \dots$ une suite d'ensembles dans Θ . Alors

$$pl\left(\bigcup_{i=1}^{\infty} H_i\right) = \lim_{i \rightarrow \infty} pl(H_i). \quad (8.15)$$

7. Pour tout sous-ensemble $H \subseteq \Theta$

$$pl(H) = \max\{pl(I) : I \text{ fini}, I \subseteq H\}, \quad (8.16)$$

$$sp(C) = \min\{sp(C) : C \text{ cofini}, I \supseteq H\}. \quad (8.17)$$

(C est cofini si \bar{C} est fini.)

8. Pour tous sous-ensembles disjoints H_1 et H_2 dans Θ ,

$$sp(H_1 \cup H_2) \leq sp(H_1) + pl(H_2) \leq pl(H_1 \cup H_2). \quad (8.18)$$

Les assertions 5 et 6 de ce théorème impliquent que les fonctions de support et de plausibilité sont **continues**. Les preuves des assertions 1 à 7 ont été présentées par Kohlas et Monney [59] et celle de l'assertion 8 par Monney [78].

Théorème 8.2.

1. La fonction de support d'une indication est **monotone d'ordre infini**, c'est-à-dire

$$sp(H) \geq \sum \left\{ (-1)^{|I|+1} sp\left(\bigcap_{i \in I} H_i\right) : \emptyset \neq I \subseteq \{1, \dots, n\} \right\} \quad (8.19)$$

pour tout $n \geq 1$ et pour tout $H \subseteq \Theta$, où les $H_i \subseteq \Theta$ sont tels que $H \supseteq H_i$ pour tout $i = 1, \dots, n$.

2. La fonction de support d'une indication est **alternée d'ordre infini**, c'est-à-dire

$$pl(H) \leq \sum \left\{ (-1)^{|I|+1} pl\left(\bigcup_{i \in I} H_i\right) : \emptyset \neq I \subseteq \{1, \dots, n\} \right\} \quad (8.20)$$

pour tout $n \geq 1$ et pour tout $H \subseteq \Theta$, où les $H_i \subseteq \Theta$ sont tels que $H \subseteq H_i$ pour tout $i = 1, \dots, n$.

Ce théorème a été présenté par Kohlas et Monney [62]; Shafer [95] en a présenté la preuve pour des cadres de discernement finis. En effet, plusieurs des théorèmes proposés par Kohlas et Monney ont été inspirés des travaux de Shafer.

D'après l'équation (8.14), toute fonction de support est une capacité de Choquet d'ordre infini. En effet, dans les années 50, Choquet [17] a développé une théorie mathématique des capacités. Une capacité a été définie axiomatiquement comme une fonction C non négative, monotone, continue et définie sur le système de tous les ensembles compacts satisfaisant l'axiome suivant :

Axiome 8.1. $C(A \cup B) + C(A \cap B) \leq C(A) + C(B)$.

De nos jours, toute fonction portant sur des ensembles et satisfaisant cet axiome est appelée une **capacité de Choquet**. Le théorème suivant montre non seulement que toute fonction de support est une capacité de Choquet d'ordre infini, mais qu'en plus, toute capacité de Choquet d'ordre infini est la fonction de support d'une certaine indication, à condition que Θ soit un ensemble fini.

Théorème 8.3. *Soit Θ un ensemble fini et $sp : \mathcal{P}(\Theta) \rightarrow [0, 1]$ une capacité de Choquet d'ordre infini telle que $sp(\emptyset) = 0$ et $sp(\Theta) = 1$. Alors il existe une indication dont la fonction de support coïncide avec sp .*

Ce théorème a été présenté par Kohlas et Monney [62]. On en trouve la preuve dans la théorie mathématique de l'évidence de Shafer [95]. Celle-ci résulte en effet d'un résultat de théorie combinatoire sur les inverses de Möbius.

8.5 Fonction de masse

Le concept de **masse** apparaît également dans la théorie mathématique des indications. Il y est définie comme le regroupement de toutes les interprétations qui fournissent le même ensemble focal. Ainsi, pour tout sous-ensemble $A \subseteq \Theta$, nous avons la **masse** :

$$m(A) = \sum \{p(\omega_j) : \Gamma(\omega_j) = A\}. \quad (8.21)$$

Nous pouvons également déterminer la masse à partir de l'application Γ dans $\mathcal{P}(\Theta)$. En effet, pour une indication $\mathcal{H} = (\Omega, P, \Gamma, \Theta)$ portant sur le cadre Θ , nous pouvons définir

$$\Gamma^{-1}(A) = \{\omega \in \Omega : \Gamma(\omega) = A\} \quad (8.22)$$

pour tout $A \subseteq \Theta$. La fonction

$$m : \mathcal{P}(\Theta) \rightarrow [0, 1], \quad (8.23)$$

donnée par

$$m(A) = P(\Gamma^{-1}(A)) = P(\{\omega \in \Omega : \Gamma(\omega) = A\}) \quad (8.24)$$

s'appelle **fonction de masse** associée à l'indication.

Il est évident que $m(A)$ est nulle si A n'est pas un ensemble focal. Il s'en suit également que $0 \leq m(A) \leq 1$ et que

$$\sum \{m(A) : A \subseteq \Theta\} = 1. \quad (8.25)$$

Cette somme est bien définie, étant donné que $m(A) = 0$ sauf pour un ensemble fini.

Le support d'une proposition implique toute proposition qui la contient. En effet, du degré de support attribué à une certaine proposition, une partie supporte une ou plusieurs propositions que celle-ci contient, tandis que la partie restante supporte exclusivement cette proposition. C'est justement cette partie restante que représente la notion de masse.

Soit $\mathcal{F} \subseteq \mathcal{P}(\Theta)$ la famille finie des ensembles focaux de \mathcal{H} , alors

$$u(H) = \sum \{\Gamma^{-1}(F) : F \in \mathcal{F}, F \subseteq H\}. \quad (8.26)$$

Similairement, nous avons

$$v(H) = \sum \{\Gamma^{-1}(F) : F \in \mathcal{F}, F \cap H \neq \emptyset\}. \quad (8.27)$$

Par ailleurs, à partir des équations (8.8) et (8.12), nous pouvons définir les fonctions de support et de plausibilité en termes de masses. Tant que \mathcal{F} est une famille finie et que $m(F) = 0$ pour chaque $F \notin \mathcal{F}$, il s'en suit que

$$sp(H) = \sum \{m(A) : A \subseteq H\} \quad (8.28)$$

et

$$pl(H) = \sum \{m(A) : A \cap H \neq \emptyset\}. \quad (8.29)$$

Deux indications, par exemple \mathcal{H}_1 et \mathcal{H}_2 , ayant les mêmes fonctions de support et de plausibilité sont dites **équivalentes**. Notons également qu'il suffit que deux indications aient les mêmes fonctions de masse pour être considérées comme équivalentes. Nous écrivons alors $\mathcal{H}_1 \simeq \mathcal{H}_2$. Il s'agit-là d'une relation **réflexive**, **symétrique** et **transitive**. Enfin, deux indications sont également équivalentes si les quatre éléments qui composent chacune d'entre elles sont identiques.

Le théorème suivant nous montre comment calculer la masse à partir de la fonction de support.

Théorème 8.4. *Soit m la fonction de masse de l'indication $\mathcal{H} = (\Omega, P, \Gamma, \Theta)$ portant sur un cadre fini Θ . Si sp est la fonction de support de \mathcal{H} , alors*

$$m(A) = \sum_{B \subseteq A} (-1)^{|A-B|} sp(B) \quad (8.30)$$

pour tout $A \subseteq \Theta$.

La démonstration de ce théorème a été présentée par Shafer [95].

Exemple 8.4. Revenons à l'exemple 7.1 de l'évolution du cours de l'action. En appliquant l'équation (8.21), nous sommes en mesure de calculer les masses pour tous les sous-ensembles de l'ensemble des réponses, c'est-à-dire les masses de toutes les propositions. Il s'agit de la partie de support soutenant exclusivement chacune de ses propositions et non leurs sous-ensembles. Nous obtenons ainsi les résultats suivants :

A	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$m(A)$	0	0.7	0	0	0	0	0	0.3

TAB. 8.5 – Exemple : masses

Il est aisé de constater que la somme des masses est effectivement égale à 1. Si les sous-ensembles $\{a, b\}$ et $\{a, c\}$ ont des masses nulles, c'est parce que justement ils ne possèdent aucune partie de support qui leur est propre. Malgré un degré de support total de 0.7 pour chacun d'entre eux, celui-ci n'est autre que le degré de support de $\{a\}$ qu'ils contiennent les deux.

8.6 Indications et fonctions de crédibilité

Après avoir introduit les concepts d'indication, de fonctions de support, de plausibilité et de masse de la théorie mathématique des indications, il serait intéressant de comparer les différences entre cette théorie avec la théorie mathématique de l'évidence. Brachinger et Monney [12] proposent dans un article récent une distinction très claire entre ces deux théories :

- **Théorie mathématique de l'évidence**

Pour un cadre de discernement Θ , la fonction de masse est définie par $m : 2^\Theta \rightarrow [0, 1]$ si et seulement si

1. $m(\emptyset) = 0$,
2. $\sum_{A \subset \Theta} m(A) = 1$.

En outre,

$$Bel : 2^\Theta \rightarrow [0, 1]$$

est une fonction de crédibilité s'il existe une fonction de masse m telle que

$$Bel(A) = \sum_{B \subset A} m(B), \quad (8.31)$$

pour tout $A \subset \Theta$.

- **Théorie mathématique des indications**

Chaque indication

$$\mathcal{H} = (\Omega, P, \Gamma, \Theta)$$

génère une fonction de crédibilité, car l'application

$$m : 2^\Theta \rightarrow [0, 1]$$

définie par

$$m(A) = P(\{\omega \in \Omega : \Gamma(\omega) = A\})$$

est en effet une fonction de masse dans le sens de la théorie mathématique de l'évidence. La fonction de crédibilité associée à l'indication \mathcal{H} est obtenue évidemment à partir de l'équation (8.31). Il est important de souligner que cette fonction de crédibilité coïncidera avec la fonction de support associée à cette indication, donnée par l'équation (8.6).

Une fonction de crédibilité est complètement spécifiée par le concept primaire de fonction de masse, telle qu'on l'entend dans la théorie mathématique de l'évidence. A l'opposé, la fonction de masse telle qu'on l'entend dans la théorie mathématique des indications n'est pas un concept primaire, mais le résultat de la composition d'un espace de probabilité (Ω, P) et de l'application multivaluée Γ associés à une indication $\mathcal{H} = (\Omega, P, \Gamma, \Theta)$. Le degré de support d'une proposition H est déterminé à partir des interprétations qui supportent cette proposition, ou autrement dit, à partir des interprétations dont l'ensemble focal est inclus dans H .

Chapitre 9

Combinaison d'indications selon la règle de Dempster

9.1 Règle de combinaison de Dempster selon la théorie mathématique des indications

La règle de Dempster présentée dans la Partie 1 est également utilisée dans la théorie mathématique des indications. De même que nous combinions plusieurs fonctions de crédibilité pour en obtenir une nouvelle, nous le ferons avec des indications. Supposons en effet que nous disposons de plusieurs indications concernant la réponse à une même question, donc de plusieurs indications relatives à un même cadre de discernement Θ . Il est nécessaire de tenir compte de toute l'information véhiculée par ces indications pour évaluer le support et la plausibilité de l'hypothèse qui nous intéresse. Pour ce faire, il suffit de déterminer une nouvelle indication en combinant celles dont on dispose, autrement dit il s'agit de combiner l'information des différentes indications pour obtenir une image globale de l'évidence disponible. L'information finale, résultant de l'indication obtenue par cette méthode, reprend tous les éléments d'information véhiculés par les indications que l'on combine.

Rappelons que pour chaque indication, il existe une interprétation correcte ω et que la réponse à la question considérée se trouve dans l'ensemble focal $\Gamma(\omega)$ correspondant à cette interprétation. En combinant des indications, on obtient plusieurs interprétations correctes, une pour chaque indication. Le vecteur ainsi formé s'appelle **interprétation correcte combinée**. Parmi tous les vecteurs que l'on peut former en choisissant une interprétation par indication, il n'en existe qu'un seul qui soit l'interprétation correcte combinée.

Logiquement, la réponse à la question considérée se trouvera dans l'intersection des en-

sembles focaux correspondant aux interprétations formant ce vecteur. Toutefois, il arrive que la combinaison de certaines interprétations mène à une intersection d'ensembles focaux qui soit vide. Cette interprétation combinée ne peut être à l'évidence correcte ; elle porte le nom d'**interprétation combinée contradictoire**. En éliminant de notre analyse toutes les interprétations combinées contradictoires, nous réduisons le champ de travail et la recherche de l'interprétation combinée correcte en devient plus simple. En ce sens, les interprétations contradictoires peuvent être interprétées comme une nouvelle source d'information qui doit être prise en compte dans l'information globale relative à la question considérée. Au départ, chaque interprétation combinée, contradictoire ou pas, bénéficie d'une certaine probabilité d'être correcte. La prise en compte de la nouvelle source d'information se fait en éliminant les interprétations combinées contradictoires de l'analyse et en **conditionnant** cette probabilité aux interprétations combinées non contradictoires.

9.2 Combinaison de deux indications

Commençons par le cas le plus simple : la combinaison de deux indications. Soient $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$ et $\mathcal{H}_2 = (\Omega_2, P_2, \Gamma_2, \Theta)$ les deux indications que nous désirons combiner. Notons que Θ reste invariable pour les deux.

Il est possible de combiner deux ou plusieurs indications dont les ensembles de réponses diffèrent, à condition qu'ils soient compatibles. On ne traitera pas ce cas de figure, étudié par Kohlas et Monney [62], auxquels on renvoie le lecteur.

Bien que les indications \mathcal{H}_1 et \mathcal{H}_2 donnent lieu à plusieurs interprétations possibles, il n'existe qu'une seule interprétation correcte pour chacune d'entre elles ; appelons-les ω_1 et ω_2 pour \mathcal{H}_1 et \mathcal{H}_2 respectivement. De ce fait, il n'existera qu'une seule interprétation combinée $(\omega_1, \omega_2) \in \Omega_1 \times \Omega_2$ qui soit correcte, avec l'ensemble focal $\Gamma(\omega_1, \omega_2)$, où $\Omega_1 \times \Omega_2$ représente évidemment le nouvel ensemble des interprétations et $\Gamma(\omega_1, \omega_2)$ représente l'**intersection** des ensembles focaux $\Gamma(\omega_1)$ et $\Gamma(\omega_2)$. En effet, $\Gamma(\omega_1) \cap \Gamma(\omega_2)$ est le plus petit sous-ensemble de Θ pour lequel il est possible de prouver qu'il contient la réponse correcte θ^* . Enfin, la probabilité p^* que (ω_1, ω_2) soit la réponse correcte n'est autre que la probabilité jointe de p_1 et p_2 , donc la probabilité que ω_1 et ω_2 soient toutes deux correctes. Si les deux indications sont **indépendantes**, c'est-à-dire si les variables aléatoires qui représentent les interprétations correctes sont stochastiquement indépendantes, la probabilité p^* est donnée par le produit $p_1 p_2$. La combinaison d'indications dépendantes ne sera pas traitée dans ce travail, pour laquelle on renvoie également le lecteur aux travaux de Kohlas et Monney [62].

Comme dans la théorie de Shafer [95] présentée dans la Partie 1, il se peut que la combinaison de certaines interprétations ne puisse se réaliser. C'est notamment le cas lorsque l'intersection de leurs ensembles focaux respectifs est donnée par l'ensemble vide :

$\Gamma(\omega_1, \omega_2) = \Gamma(\omega_1) \cap \Gamma(\omega_2) = \emptyset$. Lorsque l'intersection de $\Gamma(\omega_1)$ et $\Gamma(\omega_2)$ est **vide**, il est évident que (ω_1, ω_2) ne peut être correcte et nous en déduisons que ω_1 et ω_2 sont des interprétations **contradictoires**. Comme il s'agit d'une contradiction de deux ensembles focaux provenant de deux indications différentes, nous sommes en présence du cas de contradiction externe. Le fait de savoir que deux interprétations ω_1 et ω_2 sont contradictoires implique que la probabilité $p^*(\omega_1, \omega_2)$ est nulle et de ce fait ne doit pas être prise en considération. Il s'agit là d'une **nouvelle information**, obtenue grâce à la combinaison, qu'il est nécessaire d'utiliser pour conditionner les probabilités des combinaisons d'interprétations non contradictoires au fait que la combinaison correcte ne doit pas être contradictoire.

La démarche est identique à celle de la combinaison de deux fonctions de crédibilité. Tout d'abord, on élimine toutes les combinaisons d'interprétations contradictoires ou, autrement dit impossibles, puis on multiplie les combinaisons restantes, soit celles pour lesquelles $\Gamma(\omega_1, \omega_2)$ n'est pas vide, par un **facteur d'ajustement** K pour adapter leur poids face à cette nouvelle information. En effet, le fait d'éliminer une combinaison implique que les combinaisons d'interprétations restantes ont une probabilité plus grande de contenir la réponse correcte. Pour éviter toute confusion, les probabilités des combinaisons d'interprétations avant ajustement sont exprimées par p^* , tandis que la nouvelle fonction de probabilité après ajustement est donnée par p .

Pour tout $\omega = (\omega_1, \omega_2) \in \Omega_1 \times \Omega_2$, le facteur d'ajustement se présente de la manière suivante :

$$K = \frac{1}{1 - k} = \frac{1}{\sum \{p^*(\omega) : \omega \notin C\}} \quad (9.1)$$

où

$$C = \{\omega \in \Omega_1 \times \Omega_2 : \Gamma(\omega) = \emptyset\} \quad (9.2)$$

et

$$k = \sum \{p^*(\omega) : \omega \in C\}. \quad (9.3)$$

En comparant avec l'équation (3.3), on constate que la démarche reste identique, on l'a simplement adaptée au modèle mathématique des indications. Nous n'additionnons que les probabilités (il s'agissait des masses dans la Partie 1) correspondant aux intersections non vides, puis nous prenons l'inverse de cette somme pour obtenir le facteur d'ajustement. L'indication combinée est par conséquent donnée par

$$\mathcal{H}_1 \oplus \mathcal{H}_2 = (\Omega, P, \Gamma, \Theta) \quad (9.4)$$

où

$$\Omega = \Omega_1 \times \Omega_2 \setminus C, \quad (9.5)$$

P est une mesure de probabilité définie par

$$p(\omega) = Kp^*(\omega), \quad (9.6)$$

pour tout $\omega = (\omega_1, \omega_2) \in \Omega$, et

$$\Gamma(\omega) = \Gamma(\omega_1) \cap \Gamma(\omega_2), \quad (9.7)$$

pour tout $\omega = (\omega_1, \omega_2) \in \Omega$.

Enfin, il peut arriver que toutes les interprétations soient contradictoires lors de la combinaison de deux indications \mathcal{H}_1 et \mathcal{H}_2 , autrement dit que toutes les combinaisons d'interprétations aboutissent à l'ensemble vide. On parle dans ce cas d'indications **totale**ment **contradictaires** et on dit que $\mathcal{H}_1 \oplus \mathcal{H}_2$ n'existe pas.

La procédure qui vient d'être présentée porte également le nom de **règle de combinaison de Dempster** dans les travaux de Kohlas et de Monney, bien que sa forme ait été adaptée à la théorie des indications.

Exemple 9.1. Reprenons l'exemple 7.1 sur l'évolution de cours de l'action. Nous avons l'indication de l'ami financier que nous appellerons cette fois-ci $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$. Notons ses interprétations comme suit :

- $\omega_{1,1}$: l'ami a fait une analyse correcte,
- $\omega_{1,2}$: l'ami s'est trompé dans son analyse.

Rappelons encore que $\Gamma_1(\omega_{1,1}) = \{a\}$, $\Gamma_1(\omega_{1,2}) = \Theta$, $p_1(\omega_{1,1}) = 0.7$ et $p_1(\omega_{1,2}) = 0.3$. On suppose à présent que l'on dispose d'une deuxième indication $\mathcal{H}_2 = (\Omega_2, P_2, \Gamma_2, \Theta)$ donnée par un autre financier selon lequel le cours restera stable. Cette nouvelle indication porte sur le même ensemble de réponses $\Theta = \{a, b, c\}$. Les interprétations possibles de cette indication sont :

- $\omega_{2,1}$: il a fait une analyse correcte,
- $\omega_{2,2}$: il s'est trompé dans son analyse.

On suppose également que les probabilités de ces deux interprétations soient connues, $p(\omega_{2,1}) = 0.1$ et $p(\omega_{2,2}) = 0.9$ respectivement. Ce deuxième analyste financier est en effet notoirement incompetent. Les ensembles focaux sont les suivants : $\Gamma(\omega_{2,1}) = \{c\}$ et $\Gamma(\omega_{2,2}) = \Theta$. Il s'agit à présent de combiner ces deux indications pour obtenir une nouvelle indication portant sur l'évolution du cours de l'action et calculer les nouveaux degrés de support et de plausibilité résultant de cette double information. Déterminons tout d'abord le nouvel ensemble d'interprétations Ω . Pour cela il faut combiner chaque interprétation de l'une des indications avec chacune de l'autre. Il suffit de s'imaginer un tableau dont les interprétations de \mathcal{H}_1 représentent les lignes, celles de \mathcal{H}_2 les colonnes et celles de $\mathcal{H}_1 \oplus \mathcal{H}_2$ les cellules du tableau. Nous obtenons ainsi quatre interprétations présentées à l'intérieur du tableau suivant :

Comme il s'agit d'indications indépendantes, les probabilités jointes sont données par le produit des probabilités correspondant à chaque couple d'interprétations. Ainsi, nous obtenons :

	$\omega_{2,1}$	$\omega_{2,2}$
$\omega_{1,1}$	$(\omega_{1,1}, \omega_{2,1})$	$(\omega_{1,1}, \omega_{2,2})$
$\omega_{1,2}$	$(\omega_{1,2}, \omega_{2,1})$	$(\omega_{1,2}, \omega_{2,2})$

TAB. 9.1 – Exemple : interprétations combinées

	$p(\omega_{2,1}) = 0.1$	$p(\omega_{2,2}) = 0.9$
$p(\omega_{1,1}) = 0.7$	$p^*(\omega_{1,1}, \omega_{2,1}) = 0.7 \cdot 0.1 = 0.07$	$p^*(\omega_{1,1}, \omega_{2,2}) = 0.7 \cdot 0.9 = 0.63$
$p(\omega_{1,2}) = 0.3$	$p^*(\omega_{1,2}, \omega_{2,1}) = 0.3 \cdot 0.1 = 0.03$	$p^*(\omega_{1,2}, \omega_{2,2}) = 0.3 \cdot 0.9 = 0.27$

TAB. 9.2 – Exemple : probabilités jointes

A présent, déterminons les ensembles focaux correspondant à ces quatre interprétations. Rappelons que $\Gamma(\omega_1, \omega_2) = \Gamma(\omega_1) \cap \Gamma(\omega_2)$. Nous obtenons :

	$\Gamma(\omega_{2,1}) = \{c\}$	$\Gamma(\omega_{2,2}) = \Theta$
$\Gamma(\omega_{1,1}) = \{a\}$	$\Gamma(\omega_{1,1}, \omega_{2,1}) = \emptyset$	$\Gamma(\omega_{1,1}, \omega_{2,2}) = \{a\}$
$\Gamma(\omega_{1,2}) = \Theta$	$\Gamma(\omega_{1,2}, \omega_{2,1}) = \{c\}$	$\Gamma(\omega_{1,2}, \omega_{2,2}) = \Theta$

TAB. 9.3 – Exemple : ensembles focaux joints

Nous pouvons tout de suite remarquer que les interprétations $\omega_{1,1}$ et $\omega_{2,1}$ sont contradictoires, en effet le cours de l'action ne peut à fois augmenter et rester stable, par conséquent l'intersection de ces deux propositions est donnée par l'ensemble vide. A partir du moment où la combinaison de ces deux interprétations est impossible, elle devra bien évidemment avoir une probabilité nulle et être éliminée de l'ensemble des interprétations. Calculons à présent le facteur d'ajustement K donné par l'équation (9.1). Dans notre exemple, $C = \{(\omega_{1,1}, \omega_{2,1})\}$; il s'agit de la seule combinaison d'interprétations à éliminer. Sa probabilité est de 0.07. Vu que la somme des probabilités de toutes les combinaisons est égale à 1, la somme de toutes celles qui ne sont pas contradictoires, ou autrement dit toutes celles qui sont possibles et de $1 - 0.07 = 0.93$. Par conséquent, $K = 1/0.93$. Pour obtenir la nouvelle fonction de probabilité, il faut multiplier les probabilités des combinaisons d'interprétations que nous avons gardées par ce facteur. En résumé, l'ensemble des interprétations est donné par $\Omega = \{(\omega_{1,1}, \omega_{2,2}), (\omega_{1,2}, \omega_{2,1}), (\omega_{1,2}, \omega_{2,2})\}$ et les probabilités de ses différents éléments sont :

- $p(\omega_{1,1}, \omega_{2,2}) = 0.63/0.93$,
- $p(\omega_{1,2}, \omega_{2,1}) = 0.03/0.93$,
- $p(\omega_{1,2}, \omega_{2,2}) = 0.27/0.93$.

Sur la base de ces informations, on peut à présent calculer les fonctions de support, de plausibilité et de masse pour cette nouvelle indication combinée ;

appelons-la \mathcal{H} comme nous l'avons fait dans la section précédente. Nous obtenons ainsi, en appliquant les équations (8.7) ou (8.8), les degrés de support suivants :

H	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$sp(H)$	$\frac{0}{0.93}$	$\frac{0.63}{0.93}$	$\frac{0}{0.93}$	$\frac{0.03}{0.93}$	$\frac{0.63}{0.93}$	$\frac{0.66}{0.93}$	$\frac{0.03}{0.93}$	1

TAB. 9.4 – Exemple : degrés de support d'une indication combinée

De même, grâce aux équations (8.11) ou (8.12), nous arrivons aux résultats suivants relatifs à la plausibilité :

H	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$pl(H)$	$\frac{0}{0.93}$	$\frac{0.90}{0.93}$	$\frac{0.27}{0.93}$	$\frac{0.30}{0.93}$	$\frac{0.90}{0.93}$	1	$\frac{30}{0.93}$	1

TAB. 9.5 – Exemple : degrés de plausibilité d'une indication combinée

Déterminons enfin les masses pour cette nouvelle indication combinée. On les présente dans la table suivante :

A	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$m(A)$	$\frac{0}{0.93}$	$\frac{0.63}{0.93}$	$\frac{0}{0.93}$	$\frac{0.03}{0.93}$	$\frac{0}{0.93}$	$\frac{0}{0.93}$	$\frac{0}{0.93}$	$\frac{0.27}{0.93}$

TAB. 9.6 – Exemple : masses d'une indication combinée

On constate que les valeurs pour les masses sont identiques aux probabilités. Cela n'est pas une règle générale. Pour éviter toute confusion, rappelons que la fonction de masse donne la masse d'un sous-ensemble donné, ou autrement dit la masse d'une proposition ou hypothèse, tandis que la fonction de probabilité concerne les interprétations. En résumé, la fonction de masse porte sur Θ , alors que la fonction de probabilité porte sur Ω . De plus, la masse d'un certain sous-ensemble est donnée par la somme des probabilités de toutes les interprétations dont les ensembles focaux sont donnés par ce sous-ensemble. En l'occurrence, dans notre exemple, chaque masse est donnée par une seule probabilité et non par une somme de celles-ci, car l'ensemble focal correspondant à chaque interprétation est égal à une proposition différente. Cependant, dans certains cas, il se pourrait que plusieurs interprétations aient toutes le même ensemble focal ; la masse de celui-ci sera alors calculée en additionnant les probabilités de toutes ces interprétations.

9.3 Combinaison de plusieurs indications

Du fait que la règle de combinaison de Dempster est commutative et associative, la combinaison de plusieurs indications n'est que la combinaison de deux indications répétée

plusieurs fois. Autrement dit, si nous devons combiner les indications $\mathcal{H}_1, \dots, \mathcal{H}_n$, il suffit de combiner les indications deux par deux, puis de combiner chaque nouvelle indication ainsi obtenue avec une autre indication jusqu'à ce que toutes les indications soient incluses :

$$\begin{aligned} & \mathcal{H}_1 \oplus \mathcal{H}_2, \\ & (\mathcal{H}_1 \oplus \mathcal{H}_2) \oplus \mathcal{H}_3, \\ & ((\mathcal{H}_1 \oplus \mathcal{H}_2) \oplus \mathcal{H}_3) \oplus \mathcal{H}_4, \\ & \quad \vdots \\ & ((\dots(\mathcal{H}_1 \oplus \mathcal{H}_2) \oplus \dots) \oplus \mathcal{H}_{n-1}) \oplus \mathcal{H}_n. \end{aligned}$$

A noter que la combinaison ne fonctionne que si, lors de chaque nouvelle entrée d'une indication dans la combinaison, la nouvelle indication ainsi obtenue existe. Cependant, l'ordre dans lequel les différentes \mathcal{H}_i rentrent dans la combinaison n'a aucune importance.

9.4 Cas spéciaux

Parmi les indications qui ne sont pas totalement contradictoires, on peut en trouver certaines dont aucuns des ensembles focaux ne se contredisent, ou encore d'autres dont seulement certains des ensembles focaux se contredisent.

Indications homogènes

Supposons que l'on ait r indications simples $\mathcal{H}_1, \dots, \mathcal{H}_r$, toutes avec le même ensemble focal $F \subset \Theta$ (en dehors de Θ lui-même bien sûr, qui est un ensemble focal pour les indications simples). Nous les appelons alors indications **homogènes**. Soient $p_1 = m(F), \dots, p_r = m(F)$ les masses de l'ensemble focal F pour les indications \mathcal{H}_1 à \mathcal{H}_r . Du fait que toutes les indications ont exactement les mêmes ensembles focaux F et Θ , il est évident que nous n'aurons aucune contradiction entre leurs interprétations. De plus, $\mathcal{H}_1 \oplus \dots \oplus \mathcal{H}_r$ sera elle aussi une indication simple avec les ensembles focaux F et Θ . En effet, dans la combinaison de \mathcal{H}_1 et \mathcal{H}_2 , il y aura quatre interprétations combinées, dont les ensembles focaux seront donnés par F pour trois d'entre elles et par Θ pour la dernière. De ce fait, l'indication $(\mathcal{H}_1 \oplus \mathcal{H}_2)$ est elle-même une indication simple avec F et Θ comme ensembles focaux. En la combinant avec \mathcal{H}_3 , on obtient de nouveau une indication simple avec F et Θ comme ensembles focaux, et ainsi de suite. Par conséquent, l'indication $\mathcal{H}_1 \oplus \dots \oplus \mathcal{H}_r$ aura F et Θ comme seuls ensembles focaux et sera de ce fait elle-même une indication simple. Si les indications initiales \mathcal{H}_1 à \mathcal{H}_r sont indépendantes, alors la masse de Θ est donnée par

$$m(\Theta) = (1 - p_1)(1 - p_2) \cdots (1 - p_r) \quad (9.8)$$

et celle de F par

$$m(F) = 1 - m(\Theta), \quad (9.9)$$

car la somme des masses est égale à 1. De ce fait, chaque proposition (hypothèse) H qui contient F bénéficie d'un support convergeant vers 1 et cela d'autant plus qu'augmentera le nombre d'indications simples et indépendantes à combiner.

Indications hétérogènes

Soient \mathcal{H}_1 et \mathcal{H}_2 deux indications simples et indépendantes avec deux ensembles focaux différents F_1 et F_2 respectivement, tels que $F_1 \cap F_2 \neq \emptyset$. Nous les appelons alors indications **hétérogènes**. L'indication résultant de leur combinaison $\mathcal{H}_1 \oplus \mathcal{H}_2$ aura les ensembles focaux F_1 , F_2 , $F_1 \cap F_2$ et Θ . Leurs probabilités seront $p_1(1 - p_2)$, $(1 - p_1)p_2$, p_1p_2 et $(1 - p_1)(1 - p_2)$ respectivement. Il est aisé de constater qu'il n'y a pas de contradiction entre ces deux indications. Ainsi, l'indication $\mathcal{H}_1 \oplus \mathcal{H}_2$ bénéficie de la fonction de support suivante sur Θ :

$$sp(H) = \begin{cases} 0 & \text{si } H \not\subseteq F_1 \cap F_2 \\ p_1p_2 & \text{si } F_1 \cap F_2 \subseteq H, \text{ mais } H \not\subseteq F_1, H \not\subseteq F_2, \\ p_1 & \text{si } F_1 \subseteq H, H \not\subseteq F_2 \\ p_2 & \text{si } F_2 \subseteq H, H \not\subseteq F_1 \\ 1 - (1 - p_1)(1 - p_2) & \text{si } F_1 \cup F_2 \subseteq H, H \neq \Theta \\ 1 & \text{si } H = \Theta. \end{cases} \quad (9.10)$$

Pour l'indication combinée $\mathcal{H}_1 \oplus \mathcal{H}_2$, le support de chaque proposition H , telle que $F_1 \cup F_2 \subseteq H$, est augmenté car H est supportée par les deux indications initiales \mathcal{H}_1 et \mathcal{H}_2 . Cependant, l'indication combinée n'est plus une indication simple.

Indications conflictuelles

Considérons encore une fois deux indications \mathcal{H}_1 et \mathcal{H}_2 simples et indépendantes avec deux ensembles focaux différents F_1 et F_2 respectivement, mais cette fois-ci tels que $F_1 \cap F_2 = \emptyset$. Nous les appelons alors indications **conflictuelles**. La probabilité de $F_1 \cap F_2$, donc de la contradiction, est p_1p_2 . Ainsi, nous sommes amenés à ajuster les probabilités pour les combinaisons d'interprétations non contradictoires en les multipliant par le facteur d'ajustement $K = (1 - p_1p_2)^{-1}$. Par conséquent, l'indication $\mathcal{H}_1 \oplus \mathcal{H}_2$ aura les ensembles focaux F_1 , F_2 et Θ avec les probabilités $(p_1 - p_1p_2)/(1 - p_1p_2)$, $(p_2 - p_1p_2)/(1 - p_1p_2)$ et $(1 - p_1)(1 - p_2)/(1 - p_1p_2)$ respectivement. Ainsi, nous obtenons la fonction de support suivante pour l'indication $\mathcal{H}_1 \oplus \mathcal{H}_2$:

$$sp(H) = \begin{cases} 0 & \text{si } H \not\subseteq F_1, H \not\subseteq F_2 \\ \frac{(p_1 - p_1p_2)}{(1 - p_1p_2)} & \text{si } F_1 \subseteq H, H \not\subseteq F_2 \\ \frac{(p_2 - p_1p_2)}{(1 - p_1p_2)} & \text{si } F_2 \subseteq H, H \not\subseteq F_1 \\ \frac{(1 - (1 - p_1)(1 - p_2) - p_1p_2)}{(1 - p_1p_2)} & \text{si } F_1 \cup F_2 \subseteq H, H \neq \Theta \\ 1 & \text{si } H = \Theta. \end{cases} \quad (9.11)$$

Remarquons que les degrés de support de H seront toujours plus petits dans le cas des indications conflictuelles que dans celui des indications hétérogènes. Par exemple, le degré de support combiné de la proposition H qui n'est pas supportée par \mathcal{H}_2 , c'est-à-dire pour H qui ne contient pas F_2 , $(p_1 - p_1p_2)/(1 - p_1p_2)$ est plus petit que le degré de support de H supporté par la seule indication \mathcal{H}_1 , p_1 . C'est également le cas pour la situation inverse, c'est-à-dire pour les hypothèses qui ne sont pas supportées par \mathcal{H}_1 . Cet aspect de **non-monotonie** représente une caractéristique importante du raisonnement avec de l'évidence conflictuelle.

Exemple 9.2. Pour revenir à notre exemple 9.1 de l'évolution du cours de l'action avec deux indications, nous pouvons facilement constater qu'il s'agit d'un cas de combinaison de deux indications simples et conflictuelles.

Chapitre 10

Conditionnement des indications

Imaginons une indication simple \mathcal{B} , soutenant que la réponse correcte se trouve dans un certain sous-ensemble B de Θ , ou, en d'autres termes, toutes les interprétations de \mathcal{B} impliquent B .

Soit \mathcal{H} une indication quelconque avec le même cadre de discernement que \mathcal{B} . Supposons que nous voulons considérer $\mathcal{H} \oplus \mathcal{B}$. Cette combinaison porte le nom de **conditionnement** de \mathcal{H} à B , du fait qu'elle revient à ajuster \mathcal{H} à la connaissance fournie par \mathcal{B} que la vraie réponse se trouve dans B . Si ω_j représente n'importe quelle interprétation de \mathcal{H} , alors l'interprétation combinée affectant ω_j implique forcément $\Gamma(\omega_j) \cap B$. La probabilité de la contradiction sera donnée par :

$$k = \sum \{p(\omega_j) : \Gamma(\omega_j) \cap B = \emptyset\}. \quad (10.1)$$

Si $k = 1$, alors \mathcal{H} est en totale contradiction avec \mathcal{B} et leur combinaison $\mathcal{H} \oplus \mathcal{B}$ n'existe pas. D'un autre côté, si $k < 1$, $\mathcal{H} \oplus \mathcal{B}$ est une indication ayant les mêmes interprétations que \mathcal{H} , excepté celles pour lesquelles $\Gamma(\omega_j) \cap B = \emptyset$, avec les probabilités

$$p(\omega_j|B) = \frac{p(\omega_j)}{1 - k}. \quad (10.2)$$

Exemple 10.1. Reprenons l'exemple 9.1 de l'évolution du cours de l'action. On dispose de l'indication \mathcal{H}_1 de l'ami financier selon lequel le cours de l'action augmentera. Supposons à présent que l'entreprise émettrice de l'action publie ses résultats aujourd'hui, résultats qui s'avèrent très négatifs. Cette nouvelle information peut être entendue comme une indication supplémentaire \mathcal{H}_3 . Toutes ses interprétations nous amènent à penser que le prix de l'action diminuera. Par conséquent, son seul et unique ensemble focal sera représenté par $\{b\}$, qui sera également le coeur de l'indication \mathcal{H}_3 . En d'autres termes, selon \mathcal{H}_3 , la réponse correcte θ^* est donnée par $\{b\}$ avec certitude. Rappelons que les ensembles focaux de \mathcal{H}_1 étaient :

- $\Gamma(\omega_{1,1}) = \{a\}$,
- $\Gamma(\omega_{1,2}) = \Theta$.

Nous constatons que l'ensemble focal de $\omega_{1,1}$ a une intersection vide avec l'ensemble focal de \mathcal{H}_3 . Par conséquent, $k = 0.7$. Comme $k < 1$, il s'en suit que $\mathcal{H}_1 \oplus \mathcal{H}_3$ existe et possède comme seule interprétation $\omega_{1,2}$. La probabilité de cette interprétation est égale à

$$p(\omega_{1,2}|B) = \frac{p(\omega_{1,2})}{1 - k} = \frac{0.3}{1 - 0.7} = 1.$$

Il est également possible d'exprimer le conditionnement en termes de fonction de masse, de support et de plausibilité. C'est précisément ce que présente le théorème suivant.

Théorème 10.1. *Soit \mathcal{H} une indication portant sur Θ avec la fonction de masse m , la fonction de support sp et la fonction de plausibilité pl . Soit \mathcal{B} une indication simple, telle que toutes ses interprétations aient le même ensemble focal $B \subseteq \Theta$. Alors l'indication $\mathcal{H} \oplus \mathcal{B}$ existe et si $pl(B) > 0$ et si $m(\cdot|B)$, $sp(\cdot|B)$ et $pl(\cdot|B)$ représentent les fonctions de masse, support et plausibilité respectivement de $\mathcal{H} \oplus \mathcal{B}$ – si elle existe –, alors pour tous les $A \subseteq \Theta$ et $H \subseteq \Theta$,*

$$\begin{aligned} m(A|B) &= \frac{\sum\{m(A') : A' \cap B = A\}}{\sum\{m(A') : A' \cap B \neq \emptyset\}} \\ sp(H|B) &= \frac{(sp(H \cup \bar{B})) - sp(\bar{B})}{1 - sp(\bar{B})} \\ pl(H|B) &= \frac{pl(H \cap B)}{pl(B)}. \end{aligned} \tag{10.3}$$

Ce théorème a été présenté par Kohlas et Monney [59] et sa preuve résulte des travaux de Shafer [95].

Exemple 10.2. Reprenons l'exemple 10.1. Comme il a déjà été mentionné, parmi les deux ensembles focaux de \mathcal{H}_1 , le premier, représenté par $\{a\}$, sera en contradiction avec l'ensemble focal de \mathcal{H}_3 , $\{b\}$. Par conséquent, la seule interprétation de $\mathcal{H}_1 \oplus \mathcal{H}_3$ est $\omega_{1,2}$ et sa probabilité est de 1. En appliquant le théorème 10.1, nous avons obtenu les résultats suivants pour les fonctions de masse, de support et de plausibilité de l'indication combinée $\mathcal{H}_1 \oplus \mathcal{H}_3$:

A	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$m(A \{b\})$	0	0	0	0	0	0	0	1

TAB. 10.1 – Exemple : masses conditionnelles avec contradiction

Nous constatons que par rapport à la fonction de masse de H_1 , toutes les masses sont restées identiques à l'exception de la masse de $\{a\}$ qui est passée

de $\{a\}$ à Θ . Ainsi, la masse de Θ passe de 0.3 à 1. Le sous-ensemble $\{a\}$ a perdu en effet toute sa masse après le conditionnement.

H	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$sp(H \{b\})$	0	0	0	0	0	0	0	1

TAB. 10.2 – Exemple : degrés de support conditionnels avec contradiction

Nous constatons que tous les degrés de support de $\mathcal{H}_1 \oplus \mathcal{H}_3$ sont nuls à l'exception de celui de Θ . En effet, tous les sous-ensembles de \mathcal{H}_1 qui contenaient $\{a\}$ avaient des degrés de support positifs, mais à la lumière de la nouvelle information qui nous dit que la réponse se trouve avec certitude dans $\{b\}$ selon \mathcal{H}_3 , ils ont tous perdu la totalité de leur support, à l'exception de Θ .

H	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$pl(H \{b\})$	0	1	1	1	1	1	1	1

TAB. 10.3 – Exemple : degrés de plausibilité conditionnels

Enfin, les degrés de plausibilité de $\mathcal{H}_1 \oplus \mathcal{H}_3$ sont tous passés à 1 à l'exception de l'ensemble vide, qui bénéficie d'un degré nul par définition. En d'autres termes, à la lumière de ces deux informations, nous sommes arrivés à une indication combinée ayant pour seul ensemble focal l'ensemble des réponses Θ lui-même. Autrement dit, tout est possible quant à l'évolution de l'action.

Exemple 10.3. Reprenons l'exemple 10.1 du conditionnement de \mathcal{H}_1 à \mathcal{H}_3 . Supposons la même situation, sauf une différence concernant \mathcal{H}_3 . En effet, supposons que l'entreprise ait publié des résultats très positifs. En nous basant sur le théorème 10.1, nous sommes en mesure de calculer les fonctions de masse, de support et de plausibilité de l'indication combinée $\mathcal{H}_1 \oplus \mathcal{H}_3$. Dans cet exemple, le dénominateur représenté par la somme des masses de tous les sous-ensembles non disjoints de $\{a\}$ est égal à 1. En effet, cette somme est donnée par

$$\begin{aligned} \sum \{m(A') : A' \cap \{a\} \neq \emptyset\} &= m(a) + m(\{a, b\}) + m(\{a, c\}) + m(\{a, b, c\}) \\ &= 0.7 + 0 + 0 + 0.3 \\ &= 1 \end{aligned}$$

La fonction de masse de $\mathcal{H}_1 \oplus \mathcal{H}_3$ est présentée dans la table suivante :

A	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$m(A \{a\})$	0	1	0	0	0	0	0	0

TAB. 10.4 – Exemple : masses conditionnelles sans contradiction

Nous constatons que par rapport à la fonction de masse de H_1 , toutes les masses sont restées identiques à l'exception de la masse de Θ qui est passée de Θ à $\{a\}$. Ainsi, la masse de $\{a\}$ passe de 0.7 à 1 ; autrement dit, ce sous-ensemble récupère la totalité de la masse suite au conditionnement. A présent, pour déterminer les fonctions de support et de plausibilité de $\mathcal{H}_1 \oplus \mathcal{H}_3$, nous pouvons soit nous baser sur cette fonction de masse, soit appliquer les équations du théorème 10.1. Nous opterons pour la deuxième méthode afin d'illustrer au mieux ce théorème.

Commençons par la fonction de support de $\mathcal{H}_1 \oplus \mathcal{H}_3$. Tous ses degrés de support doivent être divisés par $(1 - sp(\{b, c\})) = 1 - 0 = 1$. En effet, comme le degré de support du complément est nul, nous n'avons pas besoin de considérer le dénominateur. Les degrés de support de l'indication combinée se réduisent finalement à $sp(H|\{a\}) = sp(H \cup \{b, c\})$. Ainsi nous obtenons les résultats suivants :

H	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$sp(H \{a\})$	0	1	0	0	1	1	0	1

TAB. 10.5 – Exemple : degrés de support conditionnels sans contradiction

Nous constatons que la fonction de support de $\mathcal{H}_1 \oplus \mathcal{H}_3$ est tout à fait identique à celle de \mathcal{H}_1 , mis à part le fait que tous les degrés de support, qui étaient de 0.7 avant le conditionnement, sont passés à 1 après celui-ci.

Déterminons enfin la fonction de plausibilité de $\mathcal{H}_1 \oplus \mathcal{H}_3$. A nouveau, nous pouvons nous passer du dénominateur dans la formule que nous appliquons, vu que $pl(a) = 1$. Ainsi, les degrés de plausibilité de cette indication combinée sont donnés par $pl(H|\{a\}) = pl(H \cap \{a\})$. Voici les résultats obtenus :

H	\emptyset	$\{a\}$	$\{b\}$	$\{c\}$	$\{a, b\}$	$\{a, c\}$	$\{b, c\}$	Θ
$pl(H \{a\})$	0	1	0	0	1	1	0	1

TAB. 10.6 – Exemple : degrés de plausibilité conditionnels sans contradiction

Nous remarquons que tous les degrés de plausibilité de $\mathcal{H}_1 \oplus \mathcal{H}_3$ sont identiques à ceux de \mathcal{H}_1 à l'exception des degrés de plausibilité de $\{b\}$, $\{c\}$ et $\{b, c\}$ qui sont devenus nuls. En effet, à la lumière de la nouvelle information apportée par \mathcal{H}_3 , nous savons que la réponse correcte se trouve avec certitude dans $\{a\}$ et, de ce fait, les propositions $\{b\}$, $\{c\}$ et $\{b, c\}$, c'est-à-dire celles qui ne contiennent pas $\{a\}$, ont perdu toute leur plausibilité.

Notons que si \mathcal{H}_3 n'était pas une indication simple, par exemple si elle avait deux ensembles focaux $\{a\}$ et $\{c\}$, bien que leur union représente un sous-ensemble de Θ , nous n'aurions

pas pu appliquer les règles de calcul du théorème 10.1. En effet, ces règles ne se rapportent qu'aux cas où toutes les interprétations de l'indication supplémentaire aboutissent au même ensemble focal, donc lorsque nous avons la certitude que celui-ci contient la réponse correcte. Pour tous les autres cas de figure, il suffit d'appliquer la règle de combinaison de Dempster, telle qu'elle a été présentée dans la section 9.2.

Troisième partie

Combinaison d'indications en chaîne

Chapitre 11

Problème du casino dans la théorie mathématique des indications

*La connaissance progresse en
intégrant en elle l'incertitude,
non en l'exorcisant.*

EDGAR MORIN

Extrait de *La méthode*

Supposons que nous disposions d'une série d'indications \mathcal{H}_i avec $i = 1, \dots, n$, relatives à une même question et donc portant sur le même ensemble de réponses Θ . On sait que pour combiner ces indications, on doit utiliser la règle de combinaison de Dempster. Il s'agit alors de créer, à partir des indications \mathcal{H}_i , une nouvelle indication \mathcal{H}' qui contiendrait toute l'information véhiculée par les \mathcal{H}_i . Il est évident dans ce cas que l'ensemble de réponses de \mathcal{H}' reste Θ .

A présent imaginons que nous disposions également d'une deuxième série d'indications, quelque peu particulières, relatives à la même question que les \mathcal{H}_i , mais, à la différence du cas précédent, il existe des relations d'influence entre les indications. Autrement dit, les sources à disposition s'influencent entre elles. Par «s'influencer» nous entendons que certaines indications de la série **remettent en question** d'autres indications de cette même série. En ce sens, on dit que certaines indications sont étroitement liées où enchaînées entre elles.

Afin de mieux cerner le problème, commençons par un exemple simple.

Exemple 11.1. Imaginons le client d'un casino qui désire jouer à la roulette. Sa question porte évidemment sur le nombre qui sortira gagnant à la fin d'un

tour de roulette. Rappelons que dans ce jeu un seul nombre est gagnant à la fin d'un tour. L'ensemble des réponses Θ est défini dans ce cas par

$$\Theta = \{0, 1, 2, \dots, 36\}$$

et représente tous les nombres qui peuvent sortir gagnants à la fin d'un tour dans un jeu de roulette. On admet que la table de roulette est équilibrée, c'est-à-dire que chaque nombre entre 0 et 36 a la même probabilité de sortir gagnant. On considère que c'est toute l'information dont on dispose pour l'instant. On est en présence d'une indication précise qu'on appellera $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$. Son ensemble des interprétations Ω_1 est représenté par

$$\Omega_1 = \{\omega_{1,1}, \omega_{1,2}, \dots, \omega_{1,37}\}$$

et l'ensemble focal de chacune de ces interprétations est donné par un des 37 résultats possibles du jeu de roulette. On a en effet,

$$\Gamma_1(\omega_{1,1}) = \{0\},$$

$$\Gamma_1(\omega_{1,2}) = \{1\},$$

$$\vdots$$

$$\Gamma_1(\omega_{1,37}) = \{36\}.$$

Considérons enfin l'espace de probabilité de cette indication donné par (Ω_1, P_1) . En admettant que la table de roulette est équilibrée, c'est-à-dire que chaque résultat possible a la même probabilité de sortir à la fin d'un tour, il est clair que la fonction de probabilité sur Ω_1 est donnée par

$$p_1(\omega_{1,j}) = \frac{1}{37},$$

$j = 1, \dots, 37$. Graphiquement, l'application Γ_1 est représentée à la Figure 11.1.

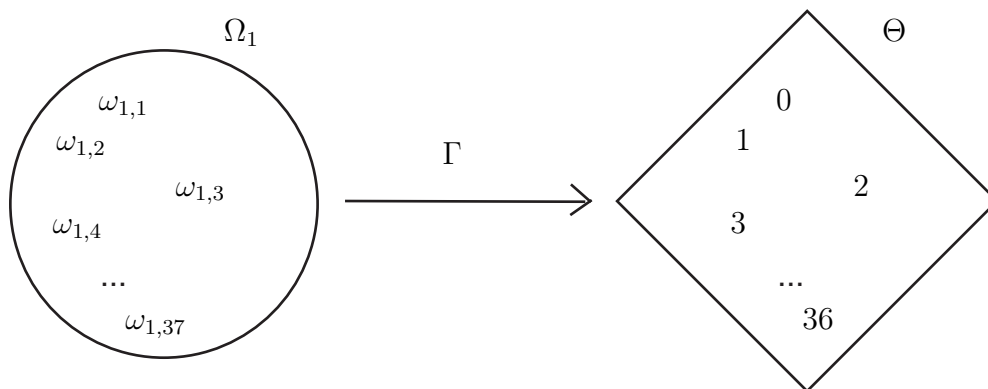
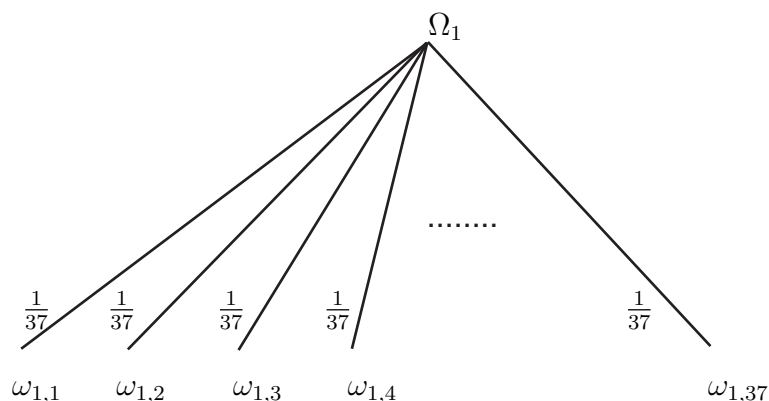


FIG. 11.1 – Exemple de la roulette avec \mathcal{H}_1

FIG. 11.2 – Arbre de probabilité de \mathcal{H}_1

L'espace de probabilité (Ω_1, P_1) de l'indication \mathcal{H}_1 peut, lui aussi, être représenté de manière graphique à travers un arbre de probabilité, comme le montre la Figure 11.1.

Enfin, si l'on s'intéresse aux fonctions de support et de plausibilité de l'indication \mathcal{H}_1 , celles-ci sont très simples à établir du fait qu'il s'agit d'une indication précise. Le degré de support de chaque nombre à la fin d'un tour est égal à son degré de plausibilité ; en effet, les deux degrés en question sont de $\frac{1}{37}$. De plus, les masses de chaque nombre possible à la fin d'un tour sont également de $\frac{1}{37}$. En résumé,

$$sp(i) = pl(i) = m(i) = \frac{1}{37},$$

avec $i = 0, \dots, 36$.

Le cas considéré jusque là est tout à fait classique, avec une seule indication. Cependant, que se passerait-il, si une deuxième indication \mathcal{H}_2 venait représenter une information supplémentaire concernant la question posée, à savoir quel nombre sortira gagnant ? Dans un cas standard, nous la combinerions avec \mathcal{H}_1 selon la règle de combinaison de Dempster. Toutefois, comment procéder si l'information véhiculée par \mathcal{H}_2 influence celle véhiculée par \mathcal{H}_1 ?

Afin d'éclaircir la compréhension de ce phénomène, revenons à l'exemple 11.1.

Exemple 11.2. Nous disposons de l'information représentée par $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$ à partir de laquelle nous avons défini les masses, les degrés de support et de plausibilité de chaque nombre possible à la fin d'un tour de la roulette.

Supposons à présent que le casino dispose de dix tables de roulette et qu'un des croupiers, que le joueur connaît bien, lui donne une information supplémentaire. En effet, il lui apprend qu'une des dix tables est truquée, mais en

précisant qu'il ignore laquelle. Cette information diffère de celle représentée par \mathcal{H}_1 , néanmoins, elle est relative à la question posée, à savoir, quel nombre sortira gagnant. Elle peut ainsi être considérée comme une indication particulière $\mathcal{H}_2 = (\Omega_2, P_2, \Gamma_2, \Theta)$. Son ensemble d'interprétations Ω_2 est donné par

$$\Omega_2 = \{\omega_{2,1}, \omega_{2,2}\},$$

où

- $\omega_{2,1}$: la table que le joueur a choisie n'est pas truquée,
- $\omega_{2,2}$: la table que le joueur a choisie est truquée.

Etant donné que neuf tables sur dix sont équilibrées et une seule est truquée, nous avons la fonction de probabilité suivante sur Ω_2 :

- $p_2(\omega_{2,1}) = 0.9$,
- $p_2(\omega_{2,2}) = 0.1$.

Quel est à présent l'ensemble des réponses de cette indication ? Rappelons que la question considérée porte sur le nombre qui sortira gagnant à la fin d'un tour de roulette. Vu que l'information véhiculée par \mathcal{H}_2 est en rapport avec la question qui nous intéresse, son ensemble de réponses est donné, comme pour \mathcal{H}_1 , par

$$\Theta = \{0, 1, 2, \dots, 36\}.$$

Il s'agit donc de combiner \mathcal{H}_1 et \mathcal{H}_2 afin d'obtenir une nouvelle indication, qu'on appellera \mathcal{M}' . On aimerait pouvoir ainsi calculer les degrés de supports et de plausibilité de chaque nombre possible à la fin d'un tour de roulette, tout en tenant compte de toute l'information disponible, soit de \mathcal{H}_1 et de \mathcal{H}_2 . Cette fois-ci, la situation est cependant particulière : l'information de \mathcal{H}_2 est certes relative à la question considérée, mais elle n'a de signification qu'en étant **liée** ou **enchaînée** à l'information de \mathcal{H}_1 . Elle n'a pas lieu d'exister sans la présence de \mathcal{H}_1 . En effet, son unique rôle est de **remettre en question** l'information qui nous est transmise par \mathcal{H}_1 . L'information de l'indication \mathcal{H}_2 seule, sans \mathcal{H}_1 , ne peut être considérée ou interprétée. Il nous est inutile de savoir qu'une table est truquée si nous ne connaissons pas le comportement des tables non truquées. Pour que \mathcal{H}_2 puisse différencier les tables truquées des tables non truquées, l'information de \mathcal{H}_1 est indispensable. Ainsi, \mathcal{H}_2 n'est pas une «vraie» indication dans le sens de la théorie mathématique des indications. Comme on le verra dans le chapitre suivant, elle est définie différemment.

Au vu de ce qui précède, on ne peut combiner \mathcal{H}_1 et \mathcal{H}_2 par la règle de combinaison de Dempster, celle-ci n'étant pas définie pour des indications du type de \mathcal{H}_2 .

On touche, à travers cet exemple, un problème particulier qui ne semble pas avoir été abordé jusqu'à présent par la théorie mathématique des indications. Il s'agit d'un cas de

figure qui sort du schéma classique de combinaison des indications. Il ne semble pas non plus que le problème ait été traité, d'un point de vue plus large, par Dempster, Shafer ou Smets.

Dans les chapitres suivants, on présentera une solution à ce problème à l'aide de la théorie mathématique des indications. Cette solution passe par la définition d'une nouvelle opération, qu'on appellera **règle de combinaison en chaîne**, en raison de l'enchaînement de l'information qu'elle exprime et de la chaîne d'indications qui en résulte. La notion de **meta-indication** sera également introduite pour exprimer le résultat de cette combinaison. Enfin, les indications du type de \mathcal{H}_2 de l'exemple qui précède, qui ne peuvent exister indépendamment et ne trouvent de sens qu'en étant combinées à d'autres indications, seront appelées **indications de confirmation**.

Chapitre 12

Chaînes d'indications, indications de confirmation et meta-indications

12.1 Construction d'une chaîne d'indications

Avant de présenter une solution générale au problème posé au chapitre précédent, revenons à l'exemple 11.2 du joueur de roulette.

Exemple 12.1. Le joueur se demande quel numéro sortira gagnant à la fin d'un tour et il dispose de deux indications, \mathcal{H}_1 et \mathcal{H}_2 , où

- \mathcal{H}_1 l'informe que chaque nombre a la même chance de sortir gagnant,
- \mathcal{H}_2 l'informe que une des dix tables de roulette est truquée.

On aimerait considérer ces deux indications simultanément afin de pouvoir calculer les degrés de support et de plausibilité des nombres qui peuvent sortir gagnants.

On connaît les interprétations possibles de ces deux indications, ainsi que la probabilité de chaque interprétation d'être l'interprétation correcte.

Commençons par considérer les interprétations de \mathcal{H}_2 :

- $\omega_{2,1}$: la table n'est pas truquée, avec $p_2(\omega_{2,1}) = 0.9$,
- $\omega_{2,2}$: la table est truquée, avec $p_2(\omega_{2,2}) = 0.1$.

Si $\omega_{2,1}$ est l'interprétation correcte de l'indication \mathcal{H}_2 , alors chaque nombre aura la même chance d'être gagnant, ou autrement dit, l'indication \mathcal{H}_1 est vraie et doit être considérée. Dans ce cas, on dit que l'indication \mathcal{H}_2 **confirme** \mathcal{H}_1 . Si, au contraire, $\omega_{2,2}$ est l'interprétation correcte, alors la table est truquée et nous ne pouvons rien dire concernant la chance d'un nombre de sortir gagnant, \mathcal{H}_1 est fausse et ne doit plus être considérée. Dans ce cas, on dit que l'indication

\mathcal{H}_2 **infirm** \mathcal{H}_1 . Le premier cas de figure réduit la situation à une indication précise, tandis que le second mène à une indication vide.

Le schéma graphique des applications Γ_1 et Γ_2 est présenté à la Figure 12.1.

Rappelons encore que par définition, un ensemble focal associé à une interprétation est le plus petit sous-ensemble du cadre de discernement pour lequel on peut prouver qu'il contient la réponse correcte. Evidemment, Ω_1 ne peut être un ensemble focal dans le cadre de la théorie mathématique des indications, car un ensemble focal doit par définition appartenir à l'ensemble des réponses, donc à Θ . On sait cependant que la fonction Γ_1 sur Ω_1 mène à Θ ; dans ce sens, je propose de considérer Ω_1 comme un **ensemble focal de second ordre**, vu que le chemin jusqu'à l'ensemble des réponses Θ passe par lui. Donc, voici les ensembles focaux de \mathcal{H}_2 :

- $\Gamma_2(\omega_{2,1}) = \Omega_1$,
- $\Gamma_2(\omega_{2,2}) = \Theta$.

En résumé,

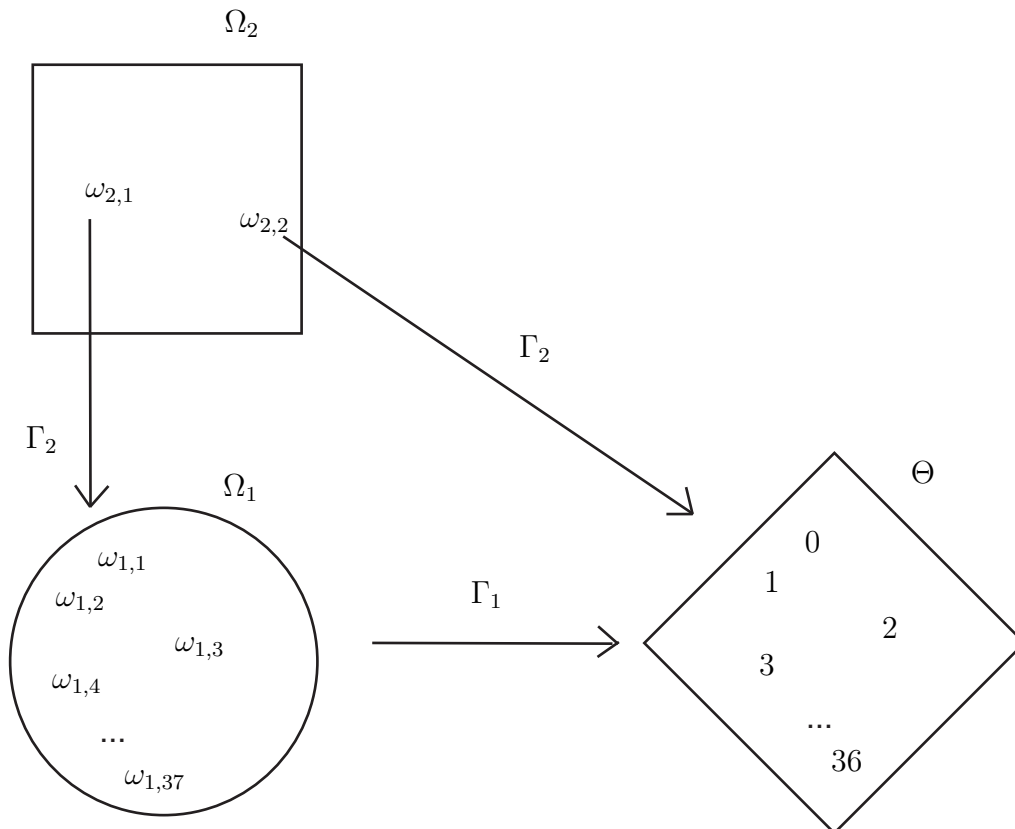


FIG. 12.1 – Exemple de la roulette avec \mathcal{H}_1 et \mathcal{H}_2

- si $\omega_{2,1}$ est l'interprétation correcte de l'indication \mathcal{H}_2 , alors une des interprétations de \mathcal{H}_1 sera correcte également et son ensemble focal est donné par un des 37 nombres qui peuvent sortir gagnants,
- si $\omega_{2,2}$ est l'interprétation correcte de l'indication \mathcal{H}_2 , alors on ne peut rien dire et son ensemble focal est donnée par Θ .

Voici à présent l'espace de probabilité de la combinaison des deux indications \mathcal{H}_1 et \mathcal{H}_2 . Il est représenté à travers l'arbre de probabilité de la Figure 12.2. Nous avons ainsi constitué un nouvel ensemble d'interprétations Ω' . L'ensemble Ω' est constitué en effet de tous les chemins qui mènent de la racine aux sommets terminaux de l'arbre. Rappelons que dans un ensemble d'interprétations, il n'y a qu'une seule et unique interprétation qui puisse être l'interprétation correcte. Dans le cas de \mathcal{M}' , soit la table est truquée et l'interprétation correcte est $\omega_{2,2}$, soit elle n'est pas truquée et l'interprétation correcte est un couple d'interprétations, formé de $\omega_{2,1}$ et d'une des interprétations de Ω_1 . On obtient ainsi :

$$\Omega' = \{(\omega_{2,1}, \omega_{1,1}), (\omega_{2,1}, \omega_{1,2}), \dots, (\omega_{2,1}, \omega_{1,37}), \omega_{2,2}\}.$$

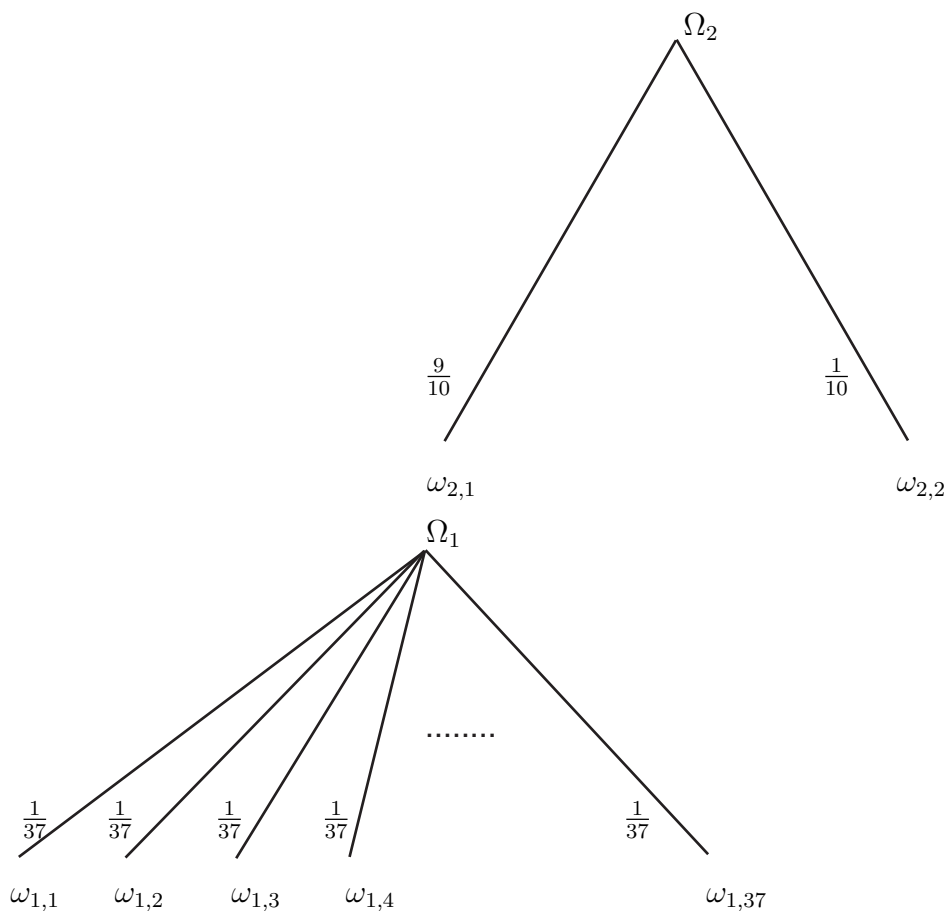


FIG. 12.2 – Arbre de probabilité de \mathcal{H}_1 et \mathcal{H}_2

Il s'agit de l'ensemble d'interprétations de \mathcal{M}' , l'indication combinant l'information de \mathcal{H}_1 et \mathcal{H}_2 .

Rappelons que la fonction de probabilité p' sur Ω' doit définir la probabilité de chaque interprétation d'être l'interprétation correcte. De même, les probabilités de toutes les interprétations de Ω' doivent sommer à 1. Cette fonction de probabilité est définie par l'arbre que nous venons de considérer. Ainsi, chaque interprétation de Ω' est représentée par un chemin de la racine de l'arbre jusqu'à un sommet terminal et sa probabilité est donnée par le produit des probabilités de tous les branchements qui lient ce sommet terminal à la racine de l'arbre.

Soit en effet A , l'événement «une certaine interprétation de Ω_1 est correcte». Soit aussi B , l'événement «une certaine interprétation de Ω_2 est correcte». La probabilité que les deux événements A et B se produisent simultanément est donnée par

$$p(B \cap A) = p(B) \cdot p(A|B),$$

où $p(B)$ donne la probabilité qu'«une certaine interprétation de Ω_2 est correcte», tandis que $p(A|B)$ donne la probabilité qu'«une certaine interprétation de Ω_1 est correcte étant donné qu'une certaine interprétation de Ω_2 est correcte».

Nous obtenons ainsi

$$\begin{aligned} - p'(\omega_{2,1}, \omega_{1,j}) &= p_2(\omega_{2,1}) \cdot p_1(\omega_{1,j}) = \frac{9}{10} \cdot \frac{1}{37} \text{ pour } j = 1, \dots, 37, \\ - p'(\omega_{2,2}) &= p_2(\omega_{2,2}) = \frac{1}{10}. \end{aligned}$$

Examinons à présent si l'on peut définir une application Γ' de Ω' dans Θ . Il faut définir les ensembles focaux correspondant aux interprétations de \mathcal{M}' . En ce qui concerne les interprétations $(\omega_{2,1}, \omega_{1,j})$, avec $j = 1, \dots, 37$, elles gardent les mêmes ensembles focaux que $\omega_{1,j}$ sous \mathcal{H}_1 , c'est-à-dire

$$\begin{aligned} \Gamma'(\omega_{2,1}, \omega_{1,1}) &= \Gamma_1(\omega_{1,1}) = \{0\}, \\ \Gamma'(\omega_{2,1}, \omega_{1,2}) &= \Gamma_1(\omega_{1,2}) = \{1\}, \\ &\vdots \\ \Gamma'(\omega_{2,1}, \omega_{1,37}) &= \Gamma_1(\omega_{1,37}) = \{36\}. \end{aligned}$$

Quant à l'ensemble focal de $\omega_{2,2}$, vu que cette interprétation ne nous fournit aucune information, nous ne pouvons aucunement réduire l'ensemble des réponses, par conséquent

$$\Gamma'(\omega_{2,2}) = \Theta.$$

Ainsi, nous avons défini une nouvelle indication $\mathcal{M}' = (\Omega', P', \Gamma', \Theta)$, en combinant l'information de \mathcal{H}_1 et \mathcal{H}_2 , où

- $\Omega' = \{(\omega_{2,1}, \omega_{1,1}), (\omega_{2,1}, \omega_{1,2}), \dots, (\omega_{2,1}, \omega_{1,37}), \omega_{2,2}\}$,
- $p'(\omega_{2,1}, \omega_{1,j}) = \frac{9}{10} \cdot \frac{1}{37}$ pour $j = 1, \dots, 37$ et $p'(\omega_{2,2}) = \frac{1}{10}$,
- $\Gamma'(\omega_{2,1}, \omega_{1,j}) = \Gamma(\omega_{1,j})$ pour $j = 1, \dots, 37$ et $\Gamma'(\omega_{2,2}) = \Theta$,
- $\Theta = \{0, 1, 2, \dots, 36\}$.

Une fois \mathcal{M}' définie, nous pouvons calculer les masses, les degrés de support et de crédibilité, comme nous le ferions pour n'importe quelle indication. Dans le cas présent, nous obtenons les résultats suivants :

- $m(i) = \frac{9}{370}$ pour $i = 0, \dots, 36$ et $m(\Theta) = 0.1$,
- $sp(i) = \frac{9}{370}$ pour $i = 0, \dots, 36$ et $sp(\Theta) = 1$,
- $pl(i) = \frac{46}{370}$ pour $i = 0, \dots, 36$ et $pl(\Theta) = 1$.

La relation en chaîne dans ce genre de problème s'exprime par une **confirmation** de l'ensemble des interprétations d'une des deux indications sur l'ensemble des interprétations de l'autre. Il est important de souligner que les deux indications portent sur le même ensemble de réponses Θ , mais qu'une partie de l'information de l'**indication qui confirme** passe par l'**indication confirmée**. De plus, l'indication qui confirme n'est pas une vraie indication. Rappelons qu'elle ne peut exister à elle seule et ne trouve de sens qu'en étant combinée à l'indication confirmée. Pour cette raison, afin d'améliorer la distinction, on appellera ce type d'indications **indications de confirmation**, notées *Con*, étant donné que leur rôle est de confirmer ou d'infirmer des indications déjà existantes. Quant aux indications qui peuvent exister indépendamment, c'est-à-dire celles définies dans la Partie 2, nous les appellerons **indications classiques** et continuerons de les noter \mathcal{H} . Enfin, nous appellerons l'indication obtenue par la combinaison en chaîne d'indications classiques avec des indications de confirmation **meta-indication** et la noterons \mathcal{M} .

12.2 Indications de confirmation

Il est nécessaire à ce stade de présenter une distinction claire entre les différents types d'indications.

Désormais, une indication \mathcal{H} , telle que l'on l'entend dans la théorie mathématique des indications, est appelée **indication classique**. En d'autres termes, une indication classique est définie par un quadruplet $(\Omega, P, \Gamma, \Theta)$, où

- Ω est l'ensemble de ses interprétations,
- P est une mesure de probabilité sur cet ensemble (fini) et est définie par sa fonction de probabilité p ,
- Γ est une application multivaluée de Ω dans Θ , c'est-à-dire

$$\Gamma : \Omega \rightarrow \Theta,$$

- et enfin, Θ est l'ensemble des réponses.

Une structure $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$ est dite **indication de confirmation**, s'il existe une indication classique $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$, telle que :

- Ω_2 est un ensemble d'interprétations et ne possède que deux éléments, $\omega_{2,c}$ et $\omega_{2,f}$, où le premier confirme \mathcal{H}_1 et le second l'infirmé,
- P_2 est une mesure de probabilité sur cet ensemble (fini) et est définie par sa fonction de probabilité p_2 ,
- Δ_2 est une application de Ω_2 dans Ω_1 et Θ , c'est-à-dire

$$\Delta_2 : \{\omega_{2,c}, \omega_{2,f}\} \rightarrow \{\Omega_1, \Theta\},$$

avec

$$\Delta_2(\omega_{2,c}) = \Omega_1$$

et

$$\Delta_2(\omega_{2,f}) = \Theta,$$

- et enfin, Θ est l'ensemble des réponses.

Il s'agit d'une définition encore très partielle des indications de confirmation. En effet, il existe d'autres types d'indications de confirmation, comme le montre l'exemple suivant.

Exemple 12.2. Soit la même situation que précédemment avec comme information disponible \mathcal{H}_1 et \mathcal{H}_2 . Rappelons que \mathcal{H}_2 , étant une indication de confirmation, sera dorénavant notée $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$. Imaginons que le joueur surprenne une discussion entre le directeur du casino et le chef du personnel. En effet, le chef se plaint que l'un de ces croupiers aurait menti à des clients. Le joueur sait en outre qu'il y a vingt croupiers employés par le casino. Cette nouvelle information peut être modélisée par un quadruplet $Con_3 = (\Omega_3, P_3, \Delta_3, \Theta)$. L'information véhiculée par Con_3 est pertinente pour la question considérée, par conséquent son ensemble de réponses est Θ . Toutefois, quel est son ensemble d'interprétations Ω_3 ? Par rapport à la question qui nous intéresse, c'est-à-dire quel nombre sortira gagnant à la fin d'un tour, nous pouvons interpréter cette information de deux manières :

- $\omega_{3,1}$: le croupier qui a dit qu'une des tables est truquée ment, avec $p_3(\omega_{3,1}) = \frac{1}{20}$,
- $\omega_{3,2}$: le croupier qui a dit qu'une des tables est truquée ne ment pas, avec $p_3(\omega_{3,2}) = \frac{19}{20}$.

Si $\omega_{3,1}$ est l'interprétation correcte de l'information contenue dans Con_3 , alors l'information contenue dans Con_2 est fautive et nous retrouvons le cas où seule \mathcal{H}_1 est pertinente. Nous devons passer par l'ensemble des interprétations Ω_1 qui nous mènera à Θ . Si au contraire $\omega_{3,2}$ est l'interprétation correcte, alors

l'information de Con_2 et vraie et nous retrouvons le cas de \mathcal{M}' de l'exemple 12.1. Dans ce cas, nous passons par Ω_2 qui nous mène soit à Θ en passant par Ω_1 , soit directement à Θ . Voici donc les ensembles focaux de Con_3 :

- $\Delta_3(\omega_{3,1}) = \Omega_1$,
- $\Delta_3(\omega_{3,2}) = \Omega_2$.

De même que lorsque nous avons combiné \mathcal{H}_1 et Con_2 , Ω_1 ne peut être l'ensemble focal de $\omega_{3,1}$ dans le cadre de la théorie mathématique des indications, car un ensemble focal doit par définition appartenir à l'ensemble des réponses, donc à Θ . On sait cependant que la fonction Γ_1 sur Ω_1 mène à Θ et, en ce sens, on considérera Ω_1 comme un ensemble focal de second ordre, vu que le chemin jusqu'à l'ensemble des réponses Θ passe par lui.

Quant à Ω_2 comme ensemble focal de $\omega_{3,2}$, il est clair qu'il ne peut contenir la réponse correcte étant un ensemble d'interprétations et non de réponses. Pourtant, on peut également le considérer en partie comme tel, puisque le chemin qui mène à Θ passe par lui. En effet, la fonction Δ_2 sur Ω_2 mène à Θ soit directement, soit en passant par Ω_1 et Γ_1 et, en ce sens, on considérera Ω_2 comme un **ensemble focal de troisième ordre**.

Remarquons enfin que la fonction Δ_3 de l'indication de confirmation Con_3 ne mène pas directement à Θ . En effet, elle est définie comme

$$\Delta_3 : \Omega_3 \rightarrow \{\Omega_1, \Omega_2\},$$

avec

$$\Delta_3(\omega_{3,1}) = \Omega_1,$$

et

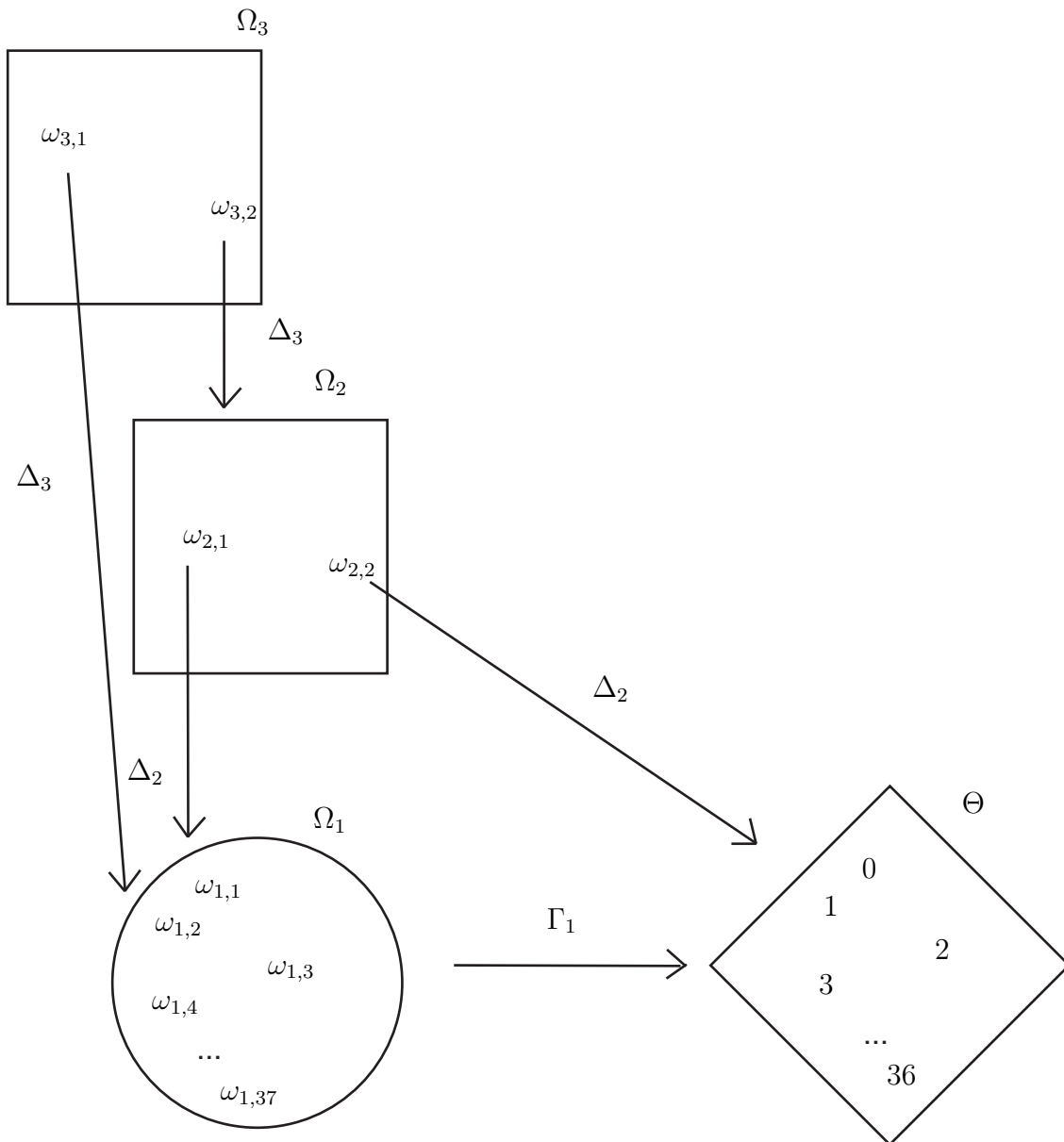
$$\Delta_3(\omega_{3,2}) = \Omega_2.$$

Le schéma graphique des applications Γ_1 , Δ_2 et Δ_3 est présenté à la Figure 12.3.

En résumé,

- si $\omega_{3,1}$ est l'interprétation correcte de l'indication Con_3 , alors une des interprétations de \mathcal{H}_1 sera correcte également,
- si $\omega_{3,2}$ est l'interprétation correcte de l'indication Con_3 , alors une des interprétations de Con_2 sera correcte également et on retrouve le cas de \mathcal{M}' que nous avons déjà considéré.

Pour revenir à Con_3 , il s'agit à nouveau d'une indication qui ne trouve de sens qu'à travers l'existence d'une autre indication, en l'occurrence Con_2 . En effet, si aucun croupier n'avait parlé au client qui désire jouer à la roulette, l'information de Con_3 lui aurait été tout à fait inutile. Cependant, comme ce n'est

FIG. 12.3 – Exemple de la roulette avec \mathcal{H}_1 , Con_2 et Con_3

pas le cas, Con_3 , ou plutôt son ensemble des interprétations Ω_3 , influencera notre modèle et en combinant les trois indications, nous obtiendrons donc une nouvelle meta-indication, différente de \mathcal{M}' , qu'on appellera \mathcal{M}'' .

La Figure 12.4 présente l'arbre de probabilité associé à ce problème.

Nous avons ainsi constitué un nouvel ensemble d'interprétations Ω'' . Il s'agit de l'ensemble d'interprétations de \mathcal{M}'' , la meta-indication combinant l'information de \mathcal{H}_1 , Con_2 et Con_3 . L'ensemble Ω'' est constitué de tous les chemins entre la racine et les sommets terminaux de l'arbre. Cependant, par souci de simplification, on notera ces chemins par l'interprétation située à leur sommet terminal, c'est-à-dire

$$\Omega'' = \{\omega_{1,1}, \omega_{1,2}, \dots, \omega_{1,37}, \omega_{2,2}\},$$

où

$$\omega_{1,j} := \{(\omega_{3,1}, \omega_{1,j})(\omega_{3,2}, \omega_{2,1}, \omega_{1,j})\} \quad (12.1)$$

et

$$\omega_{2,2} := \{(\omega_{3,2}, \omega_{2,2})\}. \quad (12.2)$$

Considérons la fonction de probabilité p'' sur Ω'' . Rappelons que chaque interprétation de Ω'' est représentée par un sommet terminal de l'arbre de la Figure 12.4 et sa probabilité est donnée par le produit des probabilités de tous les branchements qui lient ce sommet terminal à la racine de l'arbre. Contrairement au cas précédent (où l'on combinait uniquement \mathcal{H}_1 et Con_2), certaines interprétations de Ω'' apparaissent sur plusieurs sommets terminaux. Autrement dit, il y a plusieurs chemins depuis la racine de l'arbre qui mènent à ces interprétations. Pour calculer la probabilité d'une telle interprétation, il faut faire la somme des probabilités de tous les chemins qui mènent à cette interprétation. Nous obtenons ainsi :

$$\begin{aligned} - p''(\omega_{1,j}) &= \frac{1}{20} \cdot \frac{1}{37} + \frac{19}{20} \cdot \frac{9}{10} \cdot \frac{1}{37} \text{ pour } j = 1, \dots, 37, \\ - p''(\omega_{2,2}) &= \frac{19}{20} \cdot \frac{1}{10}. \end{aligned}$$

Considérons à présent l'application Γ'' de Ω'' dans Θ et définissons les ensembles focaux. Aucune des interprétations de Ω_3 ne mène directement à Θ ou à un sous-ensemble de Θ . En effet, chacune des deux flèches partant de Ω_3 passe soit par Ω_1 , soit par Ω_2 . Donc, les interprétations de Ω'' gardent les mêmes ensembles focaux que sous Ω' et Γ'' est ainsi identique à Γ' .

En résumé, nous avons défini une nouvelle meta-indication $\mathcal{M}'' = (\Omega'', P'', \Gamma'', \Theta)$ en combinant l'information de \mathcal{H}_1 , Con_2 et Con_3 , où

$$\begin{aligned} - \Omega'' &= \{\omega_{1,1}, \omega_{1,2}, \dots, \omega_{1,37}, \omega_{2,2}\}, \\ - p''(\omega_{1,j}) &= \frac{1}{20} \cdot \frac{1}{37} + \frac{19}{20} \cdot \frac{9}{10} \cdot \frac{1}{37} \text{ pour } j = 1, \dots, 37 \text{ et } p''(\omega_{2,2}) = \frac{19}{20} \cdot \frac{1}{10}, \end{aligned}$$

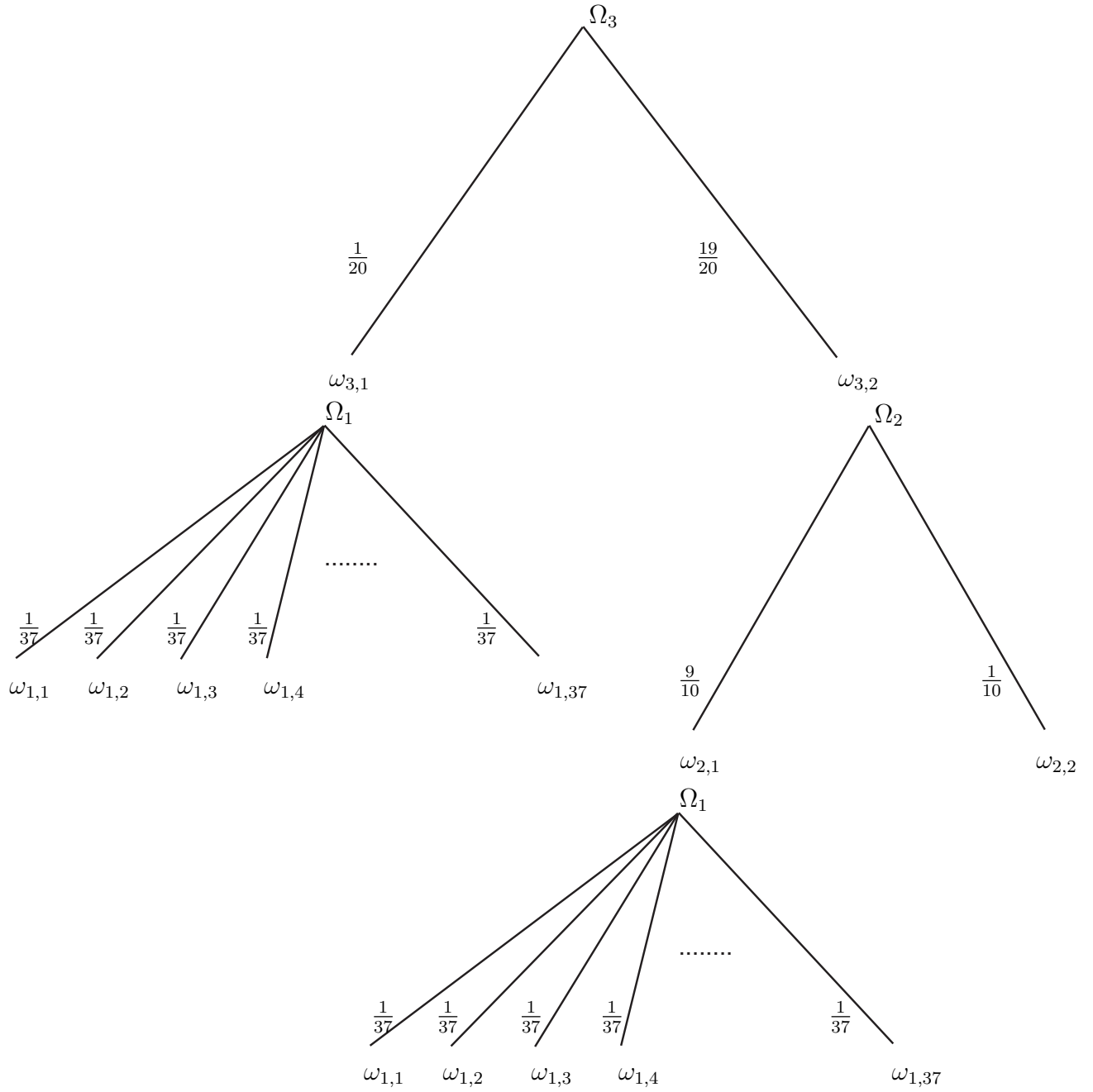


FIG. 12.4 – Arbre de probabilité de \mathcal{H}_1 , Con_2 et Con_3

- $\Gamma''(\omega_{1,j}) = \Gamma'(\omega_{1,j}) = \Gamma(\omega_{1,j})$ pour $j = 1, \dots, 37$ et $\Gamma''(\omega_{2,2}) = \Gamma'(\omega_{2,2}) = \Theta$,
- $\Theta = \{0, 1, 2, \dots, 36\}$.

Une fois \mathcal{M}'' définie, nous obtenons les masses, les degrés de support et de crédibilité suivants :

- $m(i) = \frac{181}{200} \cdot \frac{1}{37}$ pour $i = 0, \dots, 36$ et $m(\Theta) = \frac{19}{200}$,
- $sp(i) = \frac{181}{200} \cdot \frac{1}{37}$ pour $i = 0, \dots, 36$ et $sp(\Theta) = 1$,
- $pl(i) = \frac{181}{200} \cdot \frac{1}{37} + \frac{19}{200}$ pour $i = 0, \dots, 36$ et $pl(\Theta) = 1$.

Nous venons d'exposer à travers ces quelques exemples la démarche à suivre lorsqu'on veut combiner plusieurs indications en chaîne, c'est-à-dire combiner une indication classique avec des indications de confirmation. Le chapitre suivant présentera cette même démarche de manière générale.

Revenons à présent à la définition des indications de confirmation. Une indication de confirmation peut confirmer une indication classique, mais elle peut aussi confirmer une autre indication de confirmation. Une structure $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$ est donc également dite **indication de confirmation**, s'il existe une indication de confirmation $Con_1 = (\Omega_1, P_1, \Delta_1, \Theta)$, telle que :

- Ω_2 est un ensemble d'interprétations et ne possède que deux éléments, $\omega_{2,c}$ et $\omega_{2,f}$, où le premier confirme Con_1 et le second l'infirmé,
- P_2 est une mesure de probabilité sur cet ensemble (fini) et est définie par sa fonction de probabilité p_2 ,
- Δ_2 est une application de Ω_2 dans Ω_1 et Θ , c'est-à-dire

$$\Delta_2 : \{\omega_{2,c}, \omega_{2,f}\} \rightarrow \{\Omega_1, \Theta\},$$

avec

$$\Delta_2(\omega_{2,c}) = \Omega_1$$

et

$$\Delta_2(\omega_{2,f}) = \Theta,$$

- et enfin, Θ est l'ensemble des réponses.

Enfin, une indication de confirmation peut confirmer deux indications, qu'elles soient classiques ou de confirmation, si le fait de confirmer l'une correspond au fait d'infirmé l'autre. Un tel cas est présenté dans l'exemple 12.2.

Ainsi, une structure $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$ est également dite **indication de confirmation**, s'il existe deux indications (classiques ou de confirmation) $\mathcal{E}_1 = (\Omega_1, P_1, \Lambda_1, \Theta)$, et $\mathcal{E}_3 = (\Omega_3, P_3, \Lambda_3, \Theta)$, où Λ_1 et Λ_3 sont soit des applications Γ , soit des applications Δ , telles que :

- Ω_2 est un ensemble d'interprétations et ne possède que deux éléments, $\omega_{2,c}$ et $\omega_{2,f}$, où le premier confirme \mathcal{E}_1 et infirme \mathcal{E}_3 et le second infirme \mathcal{E}_1 et confirme \mathcal{E}_3 ,
- P_2 est une mesure de probabilité sur cet ensemble (fini) et est définie par sa fonction de probabilité p_2 ,
- Δ_2 est une application de Ω_2 dans Ω_1 et Ω_3 , c'est-à-dire

$$\Delta_2 : \{\omega_{2,c}, \omega_{2,f}\} \rightarrow \{\Omega_1, \Omega_3\},$$

avec

$$\Delta_2(\omega_{2,c}) = \Omega_1$$

et

$$\Delta_2(\omega_{2,f}) = \Omega_3,$$

- et enfin, Θ est l'ensemble des réponses.

On peut à présent formuler une définition générale des indications de confirmation :

Définition 12.1. Une structure $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$ est dite **indication de confirmation**,

1. s'il existe une indication (classique ou de confirmation) $\mathcal{E}_1 = (\Omega_1, P_1, \Lambda_1, \Theta)$, où Λ_1 est soit une application Γ , soit une application Δ , telle que :
 - Ω_2 est un ensemble d'interprétations et ne possède que deux éléments, $\omega_{2,c}$ et $\omega_{2,f}$, où le premier confirme \mathcal{E}_1 et le second l'infirme,
 - P_2 est une mesure de probabilité sur cet ensemble (fini) et est définie par sa fonction de probabilité p_2 ,
 - Δ_2 est une application de Ω_2 dans Ω_1 et Θ , c'est-à-dire

$$\Delta_2 : \{\omega_{2,c}, \omega_{2,f}\} \rightarrow \{\Omega_1, \Theta\}, \quad (12.3)$$

avec

$$\Delta_2(\omega_{2,c}) = \Omega_1 \quad (12.4)$$

et

$$\Delta_2(\omega_{2,f}) = \Theta, \quad (12.5)$$

- et enfin, Θ est l'ensemble des réponses,
2. ou s'il existe deux indications (classiques ou de confirmation) $\mathcal{E}_1 = (\Omega_1, P_1, \Lambda_1, \Theta)$, et $\mathcal{E}_3 = (\Omega_3, P_3, \Lambda_3, \Theta)$, où Λ_1 et Λ_3 sont soit des applications Γ , soit des applications Δ , telles que :
 - Ω_2 est un ensemble d'interprétations et ne possède que deux éléments, $\omega_{2,c}$ et $\omega_{2,f}$, où le premier confirme \mathcal{E}_1 et infirme \mathcal{E}_3 et le second infirme \mathcal{E}_1 et confirme \mathcal{E}_3 ,

- P_2 est une mesure de probabilité sur cet ensemble (fini) et est définie par sa fonction de probabilité p_2 ,
- Δ_2 est une application de Ω_2 dans Ω_1 et Ω_3 , c'est-à-dire

$$\Delta_2 : \{\omega_{2,c}, \omega_{2,f}\} \rightarrow \{\Omega_1, \Omega_3\}, \quad (12.6)$$

avec

$$\Delta_2(\omega_{2,c}) = \Omega_1 \quad (12.7)$$

et

$$\Delta_2(\omega_{2,f}) = \Omega_3, \quad (12.8)$$

- et enfin, Θ est l'ensemble des réponses.

L'ensemble focal d'une interprétation qui confirme une indication \mathcal{E} (classique ou de confirmation) est appelé **ensemble focal de minimum second ordre**, en ce sens qu'il ne mène pas directement à Θ et est toujours donné par $\Omega_{\mathcal{E}}$. Si l'ensemble focal d'une interprétation qui infirme \mathcal{E} est donné par Θ , alors il est appelé **ensemble focal de premier ordre**, en ce sens qu'il mène directement à Θ .

Notons que la différence fondamentale entre les indications de confirmation et les indications classiques réside dans l'application de l'ensemble des interprétations sur l'ensemble des réponses. Dans le cadre des indications classiques, Γ est une application multivaluée de l'ensemble des interprétations vers l'ensemble des réponses de l'indication, tandis que dans le cadre des indications de confirmation, Δ est une application de l'ensemble des interprétations vers l'ensemble des réponses de l'indication et vers l'ensemble des interprétations de l'indication confirmée, ou vers les ensembles des interprétations des deux indications confirmées.

Une seconde différence peut être notée quant aux ensembles d'interprétations. Ceux-ci ne contiennent **que deux éléments** dans le cas des indications de confirmations. On peut supposer, à tort, que dans certains cas, une indication de confirmation pourrait être interprétée de plus que deux manières. Reprenons l'exemple 11.2. L'ensemble focal de l'interprétation qui infirmait \mathcal{H}_1 était donné par Θ . En effet, si on suppose que la table est truquée, le plus petit ensemble de réponses pour lequel on peut prouver qu'il contient la réponse correcte n'est autre que Θ . Imaginons à présent que Con_2 soit légèrement différente : au lieu de «une des dix tables est truquée», supposons que le croupier ait dit «une des dix tables est truquée de manière à ce que ce soit toujours un nombre impair qui sorte». On constate que si la table est truquée effectivement de cette manière, alors Θ n'est plus le plus petit ensemble pour lequel on peut prouver qu'il contient la réponse correcte. Pour modéliser cette situation, il n'est cependant pas nécessaire de modifier la définition des indications de confirmation. Il suffit de modéliser l'information «une des dix tables est truquée de manière à ce que ce soit toujours un nombre impair qui sorte» par Con_2 et une indication classique supplémentaire, appelons-la \mathcal{H}_4 , confirmée également par Con_2

et qui représenterait la **précision** de l'information de Con_2 . Ainsi, $\omega_{2,2}$ de Con_2 dirait que la table est truquée, alors que \mathcal{H}_4 dirait comment elle est truquée. Les interprétations de Con_2 confirmeraient toutes deux une indication classique (\mathcal{H}_1 et \mathcal{H}_4) et aucune d'entre elles ne mènerait directement à Θ . Les interprétations de \mathcal{H}_4 seraient au nombre de 18 et leurs ensembles focaux respectifs seront donnés par les 18 nombres impaires d'un jeu de roulette.

Enfin, une indication confirmée n'est pas obligatoirement une indication classique. Une indication peut tout aussi bien être à la fois une indication de confirmation et une indication confirmée, lorsque par exemple elle se trouve au milieu de la chaîne d'indications. On a considéré un tel cas dans l'exemple 12.2.

12.3 Meta-indications

Commençons par définir le cas le plus simple, à savoir le cas d'une meta-indication résultant de la combinaison d'une indication classique avec une indication de confirmation. On donnera une définition générale de la notion de meta-indication au chapitre suivant.

Une structure $\mathcal{M} = (\Omega, P, \Gamma, \Theta)$ est appelée **meta-indication**, s'il existe une indication classique $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$, avec $\Omega_1 = \{\omega_{1,1}, \omega_{1,2}, \dots, \omega_{1,K}\}$ et une indication de confirmation de \mathcal{H}_1 , $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$, avec $\Omega_2 = \{\omega_{2,c}, \omega_{2,f}\}$, telles que :

- $\Omega = \{(\omega_{2,c}, \omega_{1,1}), (\omega_{2,c}, \omega_{1,2}), \dots, (\omega_{2,c}, \omega_{1,K}), \omega_{2,f}\}$ est l'ensemble de ses interprétations,
- P est une mesure de probabilité sur cet ensemble (fini) et est définie par sa fonction de probabilité

$$p(\omega_{2,c}, \omega_{1,k}) = p_2(\omega_{2,c}) \cdot p_1(\omega_{1,k}), \quad k = 1, \dots, K \quad (12.9)$$

$$p(\omega_{2,f}) = p_2(\omega_{2,f}) = 1 - \sum_{k=1}^K p(\omega_{2,c}, \omega_{1,k}) \quad (12.10)$$

- Γ est une application multivaluée de Ω dans Θ , c'est-à-dire

$$\Gamma : \Omega \rightarrow \Theta, \quad (12.11)$$

avec

$$\Gamma(\omega_{2,c}, \omega_{1,k}) = \Gamma_1(\omega_{1,k}) \subset \Theta, \quad k = 1, \dots, K \quad (12.12)$$

$$\Gamma(\omega_{2,f}) = \Theta \quad (12.13)$$

- et enfin, Θ est l'ensemble des réponses de \mathcal{H}_1 .

Comme \mathcal{M} est avant tout fonction de \mathcal{H}_1 , en ce sens que sans celle-ci la combinaison serait impossible sans celle-ci, on la notera $\mathcal{M}_{\mathcal{H}_1}$. Pour simplifier la notation, les éléments $(\omega_{2,c}, \omega_{1,j})$, $j = 1, \dots, m$ de son ensemble des interprétations Ω sont dorénavant notés $\omega_{1,j}$, $j = 1, \dots, m$.

Chapitre 13

Règle de combinaison en chaîne

13.1 Combinaison en chaîne d'une indication classique \mathcal{H}_1 avec une indication de confirmation Con_2

Commençons tout d'abord par définir la notion d'**enchaînement de l'information**.

Considérons un cas simple avec deux indications. Soient $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$ et $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$, deux indications portant sur le même ensemble de réponses Θ , telles que Con_2 confirme l'information véhiculée par \mathcal{H}_1 . L'indication Con_2 est donc une indication de confirmation et \mathcal{H}_1 une indication classique. Soient $\omega_{1,k}$, $k = 1, \dots, K$, les interprétations de Ω_1 . Soient également $\omega_{2,c}$, l'interprétation de Ω_2 qui confirme l'information véhiculée par \mathcal{H}_1 et $\omega_{2,f}$, celle qui l'infirme. Comme $\omega_{2,c}$ confirme \mathcal{H}_1 , on dit que l'information de Con_2 est liée à celle de \mathcal{H}_1 et ne peut être considérée séparément. En effet, l'ensemble focal de $\omega_{2,c}$, c'est-à-dire $\Delta_2(\omega_{2,c})$, est un ensemble focal de second ordre et est donné par Ω_1 . La fonction Δ_2 est une application

$$\Delta_2 : \Omega_2 \rightarrow \{\Omega_1, \Theta\},$$

avec

$$\Delta_2(\omega_{2,c}) = \Omega_1$$

et

$$\Delta_2(\omega_{2,f}) = \Theta,$$

et c'est ainsi que les deux indications se voient **enchaînées** l'une à l'autre. Pour cette raison justement, on dit que Con_2 n'a de sens qu'à travers l'existence de \mathcal{H}_1 .

L'interprétation $\omega_{2,f}$ de Ω_2 pourrait en théorie confirmer d'autres indications (comme ce fut le cas dans l'exemple 12.2 avec trois indications), mais puisque pour l'instant nous ne considérons que le cas avec deux indications, on admet qu'elle mène à Θ .

Par l'expression «l'interprétation $\omega_{2,c}$ de Ω_2 confirme \mathcal{H}_1 » nous entendons que l'interprétation en question **confirme** que l'information contenue dans l'indication H_1 est vraie. De plus, cela signifie que son ensemble focal est un ensemble focal de second ordre et donné par Ω_1 . Si $\omega_{2,c}$ confirme \mathcal{H}_1 , alors forcément $\omega_{2,f}$ l'infirmes, car il s'agit de deux événements complémentaires. Ainsi, si $\omega_{2,c}$ est l'interprétation correcte, cela implique que $\omega_{2,f}$ est fautive et vice-versa. Enfin, si $\omega_{2,c}$ est l'interprétation correcte de Ω_2 , alors une des interprétations de Ω_1 sera correcte également. Si, au contraire, $\omega_{2,f}$ est l'interprétation correcte, alors on ne considère plus \mathcal{H}_1 .

Graphiquement, la situation que nous venons de considérer, c'est-à-dire l'enchaînement de \mathcal{H}_1 et Con_2 , est présentée à la Figure 13.1.

Définition 13.1. Soient $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$ et $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$ deux indications sur le même ensemble de réponses Θ , telles que \mathcal{H}_1 est une indication classique et Con_2 une indication de confirmation. Soient $\omega_{2,c}$ et $\omega_{2,f}$ les deux éléments de Ω_2 . On dit que $\omega_{2,c}$ **confirme** \mathcal{H}_1 si son ensemble focal est un ensemble focal de second ordre et est donné par Ω_1 . Si tel est le cas, alors on dit que le second élément de Ω_2 , $\omega_{2,f}$, **infirmes** \mathcal{H}_1 . On dit aussi que Con_2 **confirme** \mathcal{H}_1 et que \mathcal{H}_1 est **confirmée** par Con_2 . On écrit alors

$$Con_2 \underset{\omega_{2,c}}{\varphi} \mathcal{H}_1.$$

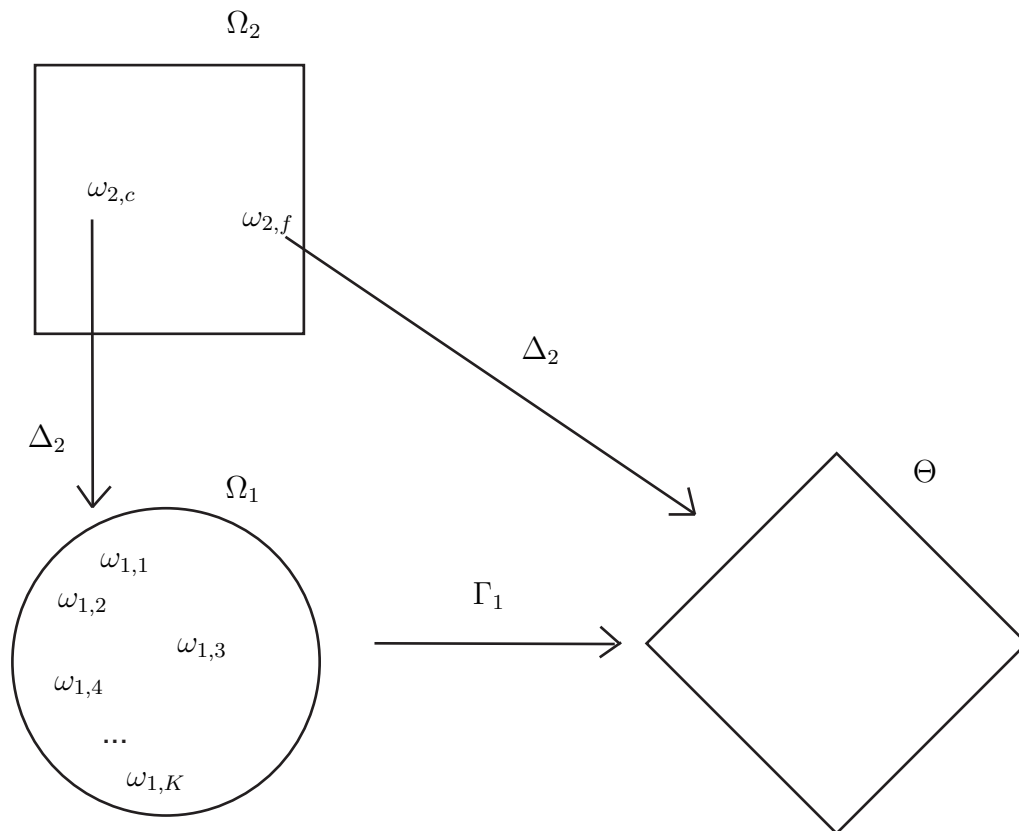


FIG. 13.1 – Combinaison en chaîne de Con_2 confirmant \mathcal{H}_1

Définissons à présent la procédure de combinaison de deux indications enchaînées. Cette procédure représente une nouvelle opération que nous appellerons **combinaison en chaîne**.

On peut se poser la question si en combinant en chaîne \mathcal{H}_1 et Con_2 , on obtient une nouvelle indication. C'est en effet le cas, car nous pouvons considérer $Con_2 \xrightarrow[\omega_{2,c}]{\mathcal{H}_1}$ comme une nouvelle indication $\mathcal{M}'_{\mathcal{H}_1}$ qui contient toute l'information de \mathcal{H}_1 et Con_2 .

L'ensemble des interprétations Ω' de cette indication $\mathcal{M}'_{\mathcal{H}_1}$ est donné par les couples $(\omega_{2,c}, \omega_{1,k})$, $k = 1, \dots, K$, et par $\omega_{2,f}$. Cependant, comme la notation a été simplifiée, on écrit $\Omega' = \{\omega_{1,1}, \omega_{1,2}, \dots, \omega_{1,K}, \omega_{2,f}\}$. Comme il a déjà été mentionné, cette notation remplace chaque couple d'interprétations (ou dans un cas de plusieurs indications enchaînées chaque n -tuplet d'interprétations) par sa seule interprétation qui possède un ensemble focal de premier ordre ou, autrement dit, qui mène directement à Θ . Pour définir cette notation de Ω' directement, on dit que Ω' est donné par l'ensemble des interprétations Ω_1 et l'ensemble des interprétations Ω_2 moins $\omega_{2,c}$. On peut l'exprimer par

$$\Omega' = \Omega_1 \cup (\Omega_2 \setminus \{\omega_{2,c}\}) = \Omega_1 \cup \{\omega_{2,f}\}. \quad (13.1)$$

Définissons à présent une fonction de probabilité p' sur Ω' . Nous connaissons les fonctions de probabilité p_1 et p_2 , c'est-à-dire la probabilité $p_1(\omega_{1,k})$, avec $k = 1, \dots, K$, de chaque interprétation de Ω_1 et la probabilité $p_2(\omega_{2,l})$, avec $l = c, f$, de chaque interprétation de Ω_2 d'être l'interprétation correcte. Etant donné que les interprétations de Ω_1 ne peuvent être correctes qu'à la condition que $\omega_{2,c}$ soit correcte, on obtient l'arbre de probabilité présenté à la Figure 13.2, où p indique la probabilité de chaque interprétation.

La fonction de probabilité p' sur Ω' est ainsi donnée par l'arbre de probabilité. Nous obtenons :

$$\begin{aligned} p'(\omega_{1,k}) &= p_2(\omega_{2,c}) \cdot p_1(\omega_{1,k}) \\ p'(\omega_{2,f}) &= p_2(\omega_{2,f}) \end{aligned} \quad (13.2)$$

avec $k = 1, \dots, K$.

Considérons enfin l'application Γ' de Ω' dans Θ . Rappelons que Ω' est composé des interprétations de Ω_1 et de celles de Ω_2 sauf $\omega_{2,c}$. Les interprétations de Ω_1 gardent les mêmes ensembles focaux, c'est-à-dire

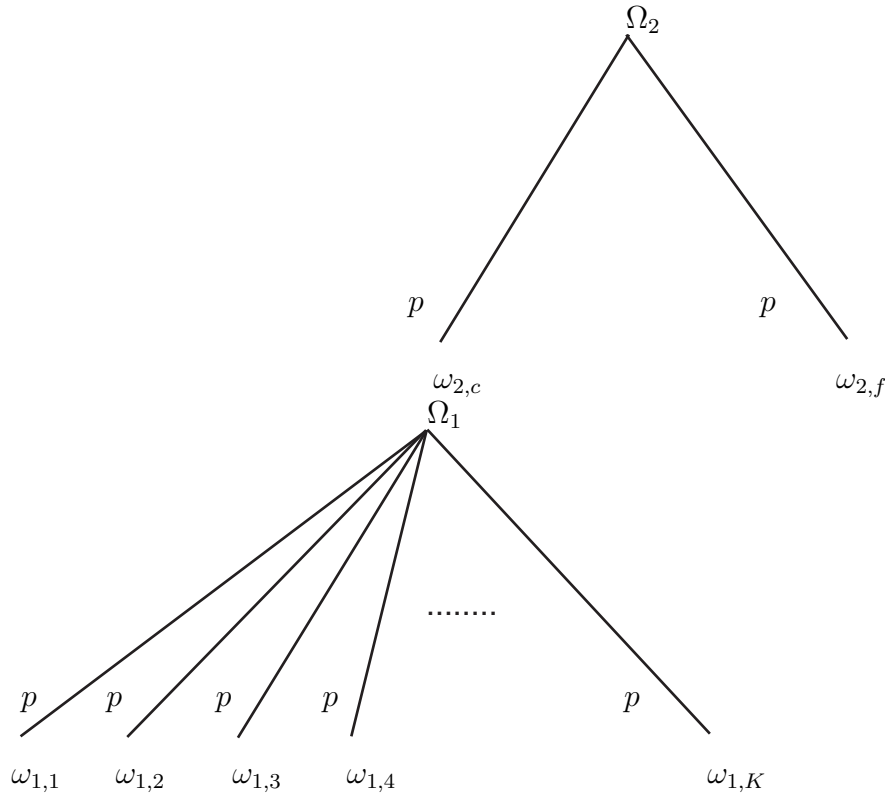
$$\Gamma'(\omega_{1,k}) = \Gamma_1(\omega_{1,k}) \quad (13.3)$$

avec $k = 1, \dots, K$. Il en est de même pour l'interprétation $\omega_{2,f}$ de Ω_2 , donc

$$\Gamma'(\omega_{2,f}) = \Delta_2(\omega_{2,f}). \quad (13.4)$$

Nous obtenons ainsi, en définissant tous les éléments qui la composent, la meta-indication $\mathcal{M}'_{\mathcal{H}_1} = (\Omega', P', \Gamma', \Theta)$, qui combine toute l'information de \mathcal{H}_1 et Con_2 . Nous écrivons alors

$$Con_2 \xrightarrow[\omega_{2,c}]{\mathcal{H}_1} \mathcal{M}'_{\mathcal{H}_1}. \quad (13.5)$$

FIG. 13.2 – Arbre de probabilité de Con_2 confirmant \mathcal{H}_1

Examinons à présent les conditions nécessaires pour que deux indications puissent être combinées en chaîne.

1. Pour combiner en chaîne deux indications, en premier lieu, il faut être en présence d'une indication de confirmation et d'une indication classique, confirmée par l'indication de confirmation. Soient $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$ l'indication classique et $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$ l'indication de confirmation.
2. Comme Con_2 est une indication de confirmation, son ensemble des interprétations Ω_2 ne possède que deux éléments $\omega_{2,c}$ et $\omega_{2,f}$. Le premier doit confirmer \mathcal{H}_1 et le second doit l'infirmier.
3. L'application Δ_2 doit être définie comme suit :

$$\Delta_2 : \Omega_2 \rightarrow \{\Omega_1, \Theta\}.$$

Plus précisément, l'ensemble focal de $\omega_{2,c}$ doit être un ensemble focal de second ordre et donné par Ω_1 , tandis que l'ensemble focal de $\omega_{2,f}$ doit être un ensemble focal de premier ordre et donné par Θ . En résumé,

- $\Delta_2(\omega_{2,c}) = \Omega_1$,
- $\Delta_2(\omega_{2,f}) = \Theta$.

Lorsque ces conditions sont respectées, alors on peut combiner \mathcal{H}_1 et Con_2 par la règle de combinaison en chaîne et ce faisant, on obtient une meta-indication notée $\mathcal{M}_{\mathcal{H}_1}$.

13.2 Combinaison en chaîne d'une indication classique \mathcal{H}_1 avec plusieurs indications de confirmation $Con_i, i = 2, \dots, n$

Nous venons de considérer la combinaison de deux indications dont l'une confirme l'autre. Que se passe-t-il lorsque nous voulons combiner en chaîne plusieurs indications dont certaines en confirment d'autres ? Dans l'exemple 12.2 du jeu de roulette avec trois indications \mathcal{H}_1, Con_2 et Con_3 , qui a été présenté au chapitre précédent, nous avons la situation suivante :

$$\begin{array}{c} Con_3 \xrightarrow[\omega_{3,2}]{} Con_2 \xrightarrow[\omega_{2,1}]{} \mathcal{H}_1. \\ \xrightarrow[\omega_{3,1}]{} \mathcal{H}_1 \end{array}$$

On peut se poser la question si cette combinaison de \mathcal{H}_1, Con_2 et Con_3 peut être considérée comme une nouvelle indication, $\mathcal{M}_{\mathcal{H}_1} = (\Omega, P, \Gamma, \Theta)$, qui contient toute l'information de \mathcal{H}_1, Con_2 et Con_3 . Comme on le verra, la démarche reste identique à la combinaison d'une indication classique avec une indication de confirmation présentée à la section précédente.

Peu importe la chaîne d'indications considérée et le nombre d'indications qui la composent, son dernier maillon sera toujours une indication classique alors que le ou les maillons qui le précèdent, des indications de confirmation. Il y aura toujours au moins une interprétation pour chaque indication de confirmation qui mènera à une autre indication (celle qu'elle confirme) au lieu de mener à l'ensemble des réponses. Quant à l'indication classique, elle représentera le maillon relié directement à l'ensemble des réponses en ce sens que toutes ses interprétations mèneront directement à cet ensemble.

Voici les conditions nécessaires pour combiner en chaîne une indication classique \mathcal{H}_1 avec plusieurs indications de confirmation $Con_i, i = 2, \dots, n$.

1. Pour combiner en chaîne une indication classique avec $(n - 1)$ indications de confirmation, on doit en premier lieu être en présence de liens de confirmation entre elles. Soient $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$ l'indication classique et $Con_i = (\Omega_i, P_i, \Delta_i, \Theta)$, $i = 2, \dots, n$, les indications de confirmation.
2. Comme Con_i , avec $i = 2, \dots, n$, sont des indications de confirmation, leurs ensembles des interprétations Ω_i ne possèdent que deux éléments $\omega_{i,c}$ et $\omega_{i,f}$. Le premier doit confirmer une indication, soit \mathcal{H}_1 , soit Con_j , avec $j \neq i$, et le second doit infirmer cette même indication. Rappelons qu'une interprétation infirmant une indication peut en confirmer une autre. De toutes les indications de confirmation Con_i , au moins une doit confirmer \mathcal{H}_1 , car celle-ci est le dernier maillon de la chaîne.
3. Pour tout $i = 2, \dots, n$, l'application Δ_i doit être définie comme suit :

$$\Delta_i : \Omega_i \rightarrow \{\Omega_1, \Omega_2, \dots, \Omega_{i-1}, \Omega_{i+1}, \dots, \Omega_n, \Theta\}. \quad (13.6)$$

Plus précisément, les ensembles focaux des $\omega_{i,c}$, $i = 2, \dots, n$ doivent être des ensembles focaux de minimum second ordre et donnés par Ω_{j_1} , avec $j_1 = 1, \dots, n$ et $j \neq i$, tandis que les ensembles focaux des $\omega_{i,f}$, $i = 2, \dots, n$, doivent être donnés soit par Θ , soit par des ensembles focaux de minimum second ordre Ω_{j_2} , avec $j_2 = 2, \dots, n$, $j \neq i$ et $j_1 \neq j_2$. En résumé,

- $\Delta_i(\omega_{i,c}) = \Omega_{j_1}$, avec $i = 2, \dots, n$, $j_1 = 1, \dots, n$ et $j_1 \neq i$,
- $\Delta_i(\omega_{i,f}) = \Theta$ ou $\Delta_i(\omega_{i,f}) = \Omega_{j_2}$, avec $i = 2, \dots, n$, $j_2 = 2, \dots, n$, $j \neq i$ et $j_1 \neq j_2$.

Lorsque toutes ces conditions sont respectées, on peut procéder à la combinaison en chaîne de $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$ avec $Con_i = (\Omega_i, P_i, \Delta_i, \Theta)$, $i = 2, \dots, n$, en suivant les étapes ci-dessous :

1. Il est nécessaire tout d'abord de définir le schéma des applications Γ_1 et Δ_i , avec $i = 2, \dots, n$, de la situation, c'est-à-dire déterminer quelle interprétation confirme quelle indication.
2. Puis, sur cette base, on doit construire l'arbre de probabilité correspondant à ce schéma, où chaque sommet terminal est donné par une interprétation, dont l'ensemble focal est un sous-ensemble de Θ , et chaque noeud correspond à une interprétation, dont l'ensemble focal est un ensemble d'interprétations d'une indication que celle-ci confirme.
3. L'étape suivante consiste à définir l'ensemble des interprétations Ω de la meta-indication. Cet ensemble est donné par toutes les interprétations représentées par les feuilles (sommets terminaux) de l'arbre. Ainsi, Ω est défini par l'union de tous les ensembles d'interprétations des indications que nous voulons combiner, moins toutes les interprétations qui confirment une indication quelconque. Autrement dit

$$\Omega = \bigcup_{i=1}^n \Omega_i \setminus \{\omega_{i,\bullet} \mid \exists_j \text{ tel que } \Delta_i(\omega_{i,\bullet}) = \Omega_j; j = 1, \dots, n; i = I + 1, \dots, n; i \neq j\}. \quad (13.7)$$

On peut également écrire

$$\Omega = \bigcup_{i=1}^n \Omega_i \setminus \Omega^*, \quad (13.8)$$

où Ω^* représente toutes les interprétations $\omega_{i,c}$ et $\omega_{i,f}$ qui confirment une indication. Il est important de noter qu'une interprétation qui infirme une indication peut très bien en confirmer une autre. Cela a d'ailleurs été le cas dans notre exemple 12.2. Ainsi, pour certaines indications de confirmation, il sera nécessaire de déduire de Ω leurs deux interprétations.

4. Puis, nous devons calculer, à partir de l'arbre de probabilité, la nouvelle fonction de probabilité p sur Ω . Pour obtenir la probabilité d'une interprétation de Ω , il suffit de multiplier les probabilités des branchements qui lient la feuille qui lui correspond à la racine de l'arbre. En d'autres termes, il faut calculer la probabilité du chemin qui

mène à cette interprétation. Si plusieurs chemins mènent à la même interprétation, il faut additionner leurs probabilités pour obtenir celle de l'interprétation.

5. Enfin, il est nécessaire de définir l'application Γ de Ω sur Θ , c'est-à-dire les ensembles focaux des interprétations de Ω . Il s'agit d'une opération très simple, car chaque interprétation de Ω garde le même ensemble focal que celui qu'elle avait dans son indication d'origine. Les interprétations qui confirment des indications n'ont pas d'ensemble focal associé de premier ordre et ne se trouvent pas dans Ω .

Ainsi, meta-indication $\mathcal{M}_{\mathcal{H}_1} = (\Omega, P, \Gamma, \Theta)$ a été définie.

13.3 Combinaison en chaîne de plusieurs indications classiques \mathcal{H}_i , $i = 1, \dots, I$ avec plusieurs indications de confirmation Con_j , $i = I + 1, \dots, n$

Nous venons de considérer la combinaison d'une indication classique avec plusieurs indications de confirmation. Que se passe-t-il lorsque nous voulons combiner en chaîne plusieurs indications classiques avec plusieurs indications de confirmation? Dans la section 12.2, nous avons exposé un cas simple où deux indications classiques étaient confirmées par la même indication de confirmation. Appelons-les \mathcal{H}_1 , \mathcal{H}_2 et Con_3 . L'interprétation de Con_3 qui infirmait \mathcal{H}_1 , confirmait \mathcal{H}_2 , et vice versa. Ainsi, nous avons la situation suivante :

$$\begin{array}{c} Con_3 \xrightarrow[\omega_{3,2}]{} \mathcal{H}_2. \\ \xrightarrow[\omega_{3,1}]{} \mathcal{H}_1 \end{array}$$

On peut aussi imaginer des situations plus complexes où plusieurs couples d'indications classiques sont confirmés chacun par une indication de confirmation. Puis, imaginons que parmi ces indications de confirmation, il existe des couples confirmés eux aussi par une indication de confirmation et ainsi de suite. Il y aurait enfin des indications de confirmation qui confirment une seule indication classique ou une seule indication de confirmation. Ainsi, il s'agirait d'une situation de combinaison en chaîne de plusieurs indications \mathcal{H}_i , $i = 1, \dots, I$ avec Con_i , $i = I + 1, \dots, n$.

On peut se poser la question si cette combinaison de \mathcal{H}_i , $i = 1, \dots, I$ avec Con_i , $i = I + 1, \dots, n$, peut être considérée comme une nouvelle indication, $\mathcal{M}_{\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_I} = (\Omega, P, \Gamma, \Theta)$, qui contient toute l'information de \mathcal{H}_i , $i = 1, \dots, I$ et Con_i , $i = I + 1, \dots, n$.

Tout comme pour les conditions de la section précédente, peu importe la chaîne d'indications que nous considérons et le nombre d'indications qui la composent, ses derniers maillons seront toujours des indications classiques alors que les maillons qui les précèdent, des indications de confirmation. Ainsi, les indications classiques représenteront les maillons

reliés directement à l'ensemble des réponses en ce sens que toutes leurs interprétations mèneront directement à cet ensemble.

Voici à présent les conditions nécessaires pour combiner plusieurs indications classiques \mathcal{H}_i , $i = 1, \dots, I$ avec plusieurs indications de confirmation Con_i , $i = I + 1, \dots, n$.

1. Pour combiner en chaîne n indications, dont I classiques et $n - I$ de confirmation en une seule meta-indication, nous devons en premier lieu être en présence de liens de confirmation entre elles. Soient $\mathcal{H}_i = (\Omega_i, P_i, \Gamma_i, \Theta)$, avec $i = 1, \dots, I$ les indications classiques et soient $Con_i = (\Omega_i, P_i, \Delta_i, \Theta)$, avec $i = I + 1, \dots, n$, les indications de confirmation.
2. Comme Con_i , avec $i = I + 1, \dots, n$, sont des indications de confirmation, leurs ensembles des interprétations Ω_i ne possèdent que deux éléments $\omega_{i,c}$ et $\omega_{i,f}$. Le premier doit confirmer soit une indication classique, soit une indication de confirmation, et le second doit l'infirmier. Rappelons qu'une interprétation infirmant une indication peut en confirmer une autre. De toutes les indications de confirmation Con_i , $i = I + 1, \dots, n$ au moins $(I/2)$ doivent confirmer les \mathcal{H}_i , $i = 1, \dots, I$. En effet, chaque \mathcal{H}_i doit être confirmée par au moins une indication de confirmation, car c'est le seul moyen pour elle d'être enchaînée aux autres indications. Comme chaque indication de confirmation peut confirmer au maximum deux indications classiques et qu'il existe I indications classiques, on aura au moins $(I/2)$ Con_i , $i = I + 1, \dots, n$ confirmant des \mathcal{H}_i , $i = 1, \dots, I$.
3. Pour tout i tel que $i = I + 1, \dots, n$, l'application Δ_i doit être définie comme suit :

$$\Delta_i : \Omega_i \rightarrow \{\Omega_1, \Omega_2, \dots, \Omega_{i-1}, \Omega_{i+1}, \dots, \Omega_n, \Theta\}. \quad (13.9)$$

Plus précisément, les ensembles focaux des $\omega_{i,c}$, $i = I + 1, \dots, n$ doivent être des ensembles focaux de minimum second ordre et donnés par Ω_{j_1} , avec $j_1 = 1, \dots, n$ et $j \neq i$, tandis que les ensembles focaux des $\omega_{i,f}$, $i = I + 1, \dots, n$, doivent être soit donnés par Θ , soit par des ensembles focaux de minimum second ordre Ω_{j_2} , avec $j_2 = 2, \dots, n$, $j \neq i$ et $j_1 \neq j_2$. En résumé,

- $\Delta_i(\omega_{i,c}) = \Omega_{j_1}$, avec $i = I + 1, \dots, n$, $j_1 = 1, \dots, n$ et $j_1 \neq i$,
- $\Delta_i(\omega_{i,f}) = \Theta$ ou $\Delta_i(\omega_{i,f}) = \Omega_{j_2}$, avec $i = I + 1, \dots, n$, $j_2 = 2, \dots, n$, $j \neq i$ et $j_1 \neq j_2$.

Lorsque toutes ces conditions sont respectées, on peut procéder à la combinaison en chaîne de n indications \mathcal{H}_i , $i = 1, \dots, I$, avec Con_i , $i = I + 1, \dots, n$, en suivant les étapes ci-dessous :

1. Il faut tout d'abord de définir le schéma des applications Γ_i , avec $i = 1, \dots, I$, et Δ_i , avec $i = I + 1, \dots, n$, de la situation, c'est-à-dire déterminer quelle interprétation confirme quelle indication.

2. Puis, sur cette base, on doit construire l'arbre de probabilité correspondant à ce schéma, où chaque sommet terminal est donné par une interprétation, dont l'ensemble focal est un sous-ensemble de Θ , et chaque noeud correspond à une interprétation, dont l'ensemble focal est un ensemble d'interprétations d'une indication qu'elle confirme.
3. L'étape suivante consiste à définir l'ensemble des interprétations Ω de la meta-indication. Cet ensemble est donné par toutes les interprétations représentées par les feuilles (sommets terminaux) de l'arbre. Ainsi, Ω est défini par l'union de tous les ensembles d'interprétations des indications que nous voulons combiner moins toutes les interprétations qui confirment une indication quelconque. Autrement dit

$$\Omega = \bigcup_{i=1}^n \Omega_i \setminus \{\omega_{i,\bullet} \mid \exists_j \text{ tel que } \Delta_i(\omega_{i,\bullet}) = \Omega_j; j = 1, \dots, n; i = I+1, \dots, n; i \neq j\}. \quad (13.10)$$

On peut également écrire

$$\Omega = \bigcup_{i=1}^n \Omega_i \setminus \Omega^*, \quad (13.11)$$

où Ω^* représente toutes les interprétations $\omega_{i,c}$ et $\omega_{i,f}$ qui confirment une indication. Comme dans le cas avec une seule indication classique, pour certaines indications de confirmation, il sera nécessaire de déduire de Ω leurs deux interprétations.

4. Puis, on doit calculer, à partir de l'arbre de probabilité, la nouvelle fonction de probabilité p sur Ω . Pour obtenir la probabilité d'une interprétation de Ω , il suffit de multiplier les probabilités des branchements qui lient la feuille qui lui correspond à la racine de l'arbre. En d'autres termes, il faut calculer la probabilité du chemin qui mène à cette interprétation. Si plusieurs chemins mènent à la même interprétation, il faut additionner leurs probabilités pour obtenir celle de l'interprétation. Autrement dit, nous devons calculer pour chaque interprétation ω_k de Ω , la somme des probabilités de tous les chemins qui mènent, depuis la racine de l'arbre, à ω_k . On appellera ces chemins **suites confirmatives**.

Une telle suite est définie par une suite

$$\mathbf{s}_c := (\omega_{1,\bullet}, \omega_{2,\bullet}, \dots, \omega_{L,\bullet}), \quad (13.12)$$

$1 \leq L \leq n - I$, d'interprétations d'indications de confirmation $\omega_{l,\bullet}$, $l = 1, \dots, L$, où :

- l'élément $\omega_{l,\bullet}$ confirme un Ω_{l+1} , c'est-à-dire $\Delta_l(\omega_{l,\bullet}) = \Omega_{l+1}$,
- pour l'ensemble Ω_1 , qui contient $\omega_{1,\bullet}$, il n'existe pas d'interprétation qui le confirme et
- le dernier élément de la suite $\omega_{L,\bullet}$ confirme l'ensemble d'interprétation de l'indication (classique ou de confirmation) d'origine de ω_k .

L'ensemble des suites confirmatives de ω_k est noté $S_c(\omega_k)$.

La probabilité d'une suite confirmative \mathbf{s}_c est donnée par

$$p(\mathbf{s}_c) = \prod_{l=1}^L p_l(\omega_l, \bullet) \quad (13.13)$$

et celle de ω_k de Ω est donnée par

$$p(\omega_k) = \sum_{\mathbf{s}_c \in S_c(\omega_k)} p(\mathbf{s}_c) p_i(\omega_k), \quad (13.14)$$

où $p_i(\omega_k)$ est la probabilité de ω_k dans son indication d'origine.

5. Enfin, il est nécessaire de définir l'application Γ de Ω sur Θ , c'est-à-dire les ensembles focaux des interprétations de Ω . Chaque interprétation de Ω garde le même ensemble focal que celui qu'elle avait dans son indication d'origine. Les interprétations qui confirment des indications n'ont pas d'ensemble focal associé de premier ordre et ne se trouvent pas dans Ω .

Ainsi, une meta-indication $\mathcal{M}_{\mathcal{H}_1, \dots, \mathcal{H}_I} = (\Omega, P, \Gamma, \Theta)$ a été définie.

On peut à présent proposer une définition générale d'une meta-indication :

Définition 13.2. Une structure $\mathcal{M} = (\Omega, P, \Gamma, \Theta)$ est appelée $(I, n - I)$ -**meta-indication**, s'il existe I indications classiques $\mathcal{H}_i = (\Omega_i, P_i, \Gamma_i, \Theta)$, $i = 1, \dots, I$, et $n - I$ indications de confirmation, $Con_j = (\Omega_j, P_j, \Delta_j, \Theta)$, $j = I + 1, \dots, n$, avec $\Omega_j = \{\omega_{j,c}, \omega_{j,f}\}$, telles que :

- pour chaque Ω_i , $i = 1, \dots, I$, il existe au moins un $\omega_{j,\bullet}$, $j = I + 1, \dots, n$, tel que $\Delta_j(\omega_{j,\bullet}) = \Omega_i$,
- $\Omega = \bigcup_{i=1}^n \Omega_i \setminus \{\omega_{i,\bullet} \mid \exists_j \text{ tel que } \Delta_i(\omega_{i,\bullet}) = \Omega_j; j = 1, \dots, n; i = I + 1, \dots, n; i \neq j\}$ et ses éléments sont notés ω_k ,
- P est une mesure de probabilité sur cet ensemble (fini) et est définie par sa fonction de probabilité

$$p(\omega_k) = \sum_{\mathbf{s}_c \in S_c(\omega_k)} p(\mathbf{s}_c) p_i(\omega_k),$$

où $p_i(\omega_k)$ est la probabilité de ω_k dans son indication d'origine et $p(\mathbf{s}_c)$ est la probabilité d'une suite confirmative de ω_k , donnée par

$$p(\mathbf{s}_c) = \prod_{l=1}^L p_l(\omega_l, \bullet),$$

- Γ est une application multivaluée de Ω dans Θ , c'est-à-dire

$$\Gamma : \Omega \rightarrow \Theta, \quad (13.15)$$

avec $\Gamma(\omega_k) = \Gamma_i(\omega_k)$, où $\Gamma_i(\omega_k)$ est l'ensemble focal de ω_k dans son indication d'origine,

- et enfin, Θ est l'ensemble commun des réponses des I indications classiques \mathcal{H}_i .

Evidemment, chaque meta-indication dans le sens de la section 12.3 est une $(I, n-I)$ -meta-indication, avec $I = n - I = 1$. Dans la suite de ce travail, chaque $(1, 1)$ -meta-indication sera, par soucis de simplification, appelée seulement «meta-indication», comme elle l'a été jusqu'ici.

On constate qu'une meta-indication a la même structure qu'une indication classique, ce qui nous mène au théorème suivant :

Théorème 13.1. *Toute meta-indication est une indication classique.*

En outre, une indication classique peut être considérée comme une meta-indication. Elle résulterait de la combinaison d'elle-même et d'une indication de confirmation, où l'interprétation qui la confirme a une probabilité de 1 et celle qui l'infirme a une probabilité de 0. Normalement, suite à une telle combinaison, à son ensemble d'interprétations s'additionnerait l'interprétation qui l'infirme, cependant, rappelons qu'il n'est pas nécessaire de considérer les interprétations de probabilité nulle et que de ce fait, elles sont éliminées. Ainsi, nous retrouvons exactement l'indication classique de départ. On peut ainsi proposer le théorème suivant :

Théorème 13.2. *Toute indication classique est une meta-indication.*

Chapitre 14

Propriétés de la règle de combinaison en chaîne

On examinera dans ce chapitre les propriétés de commutativité, d'associativité et de distributivité de l'opération de combinaison en chaîne. Disons-le d'emblée, cette règle ne possède aucune de ces propriétés. C'est ce que démontrent les développements qui suivent, qui permettent également de mettre en évidence d'autres aspects intéressants de la règle.

14.1 Commutativité

Contrairement à la règle de combinaison de Dempster, la règle de combinaison en chaîne **n'est pas commutative**. D'ailleurs, la question de la commutativité ne se pose pas du fait de l'asymétrie de la combinaison en chaîne. Illustrons cela par un exemple très simple. Soient deux indications $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$ et $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$ portant sur le même ensemble de réponses Θ . En effet

$$Con_2 \underset{\omega_{2,c}}{\rightsquigarrow} \mathcal{H}_1 \neq \mathcal{H}_1 \underset{\omega_{1,i}}{\rightsquigarrow} Con_2 \quad (14.1)$$

où $\omega_{2,c}$ est l'interprétation de \mathcal{H}_2 qui confirme \mathcal{H}_1 et $\omega_{1,i}$ est une interprétation de \mathcal{H}_1 .

La seconde partie de cette équation n'est tout simplement **pas définie**!

Il n'est pas possible de rencontrer un cas de figure où une indication classique \mathcal{H}_1 confirme une indication de confirmation Con_2 . Toutes les interprétations de Ω_1 doivent mener directement à Θ et aucune d'entre elles ne peut nous conduire à Ω_2 . Les définitions d'une indication classique et d'une indication de confirmation ne permettent pas ce type d'enchaînement.

A présent, considérons deux indications de confirmation enchaînées entre elles. Soient $Con_1 = (\Omega_1, P_1, \Delta_1, \Theta)$ et $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$ les deux indications en question. De nouveau,

$$Con_2 \underset{\omega_{2,c}}{\rightsquigarrow} Con_1 \neq Con_1 \underset{\omega_{1,c}}{\rightsquigarrow} Con_2 \quad (14.2)$$

où $\omega_{2,c}$ est l'interprétation de Con_2 qui confirme Con_1 et $\omega_{1,c}$ est l'interprétation de Con_1 qui confirme Con_2 .

Deux indications de confirmation ne peuvent se confirmer mutuellement. Logiquement, cela n'a pas de sens. Si le rôle de Con_2 est de confirmer ou infirmer l'infirmerie véhiculée par Con_1 , alors celle-ci ne peut à son tour faire de même pour Con_2 . Si tel était le cas, nous serions en présence d'un enchaînement cyclique qui non seulement serait exempt de toute logique, mais ne pourrait jamais aboutir à l'ensemble des réponses Θ . En effet, toutes les interprétations mènent directement ou indirectement à Θ . Dans le cas considéré ici, si les deux parties de l'équation étaient égales, $\omega_{2,c}$ et $\omega_{1,c}$ ne pourraient remplir cette fonction essentielle des interprétations.

En résumé, deux indications de confirmation **ne peuvent se confirmer mutuellement**. Nous aurons toujours soit Con_2 qui confirme Con_1 à travers son interprétation $\omega_{2,c}$, soit Con_1 qui confirme Con_2 à travers son interprétation $\omega_{1,c}$, mais ces deux cas ne sont pas identiques.

La question de la commutativité ne se pose donc pas pour la règle de combinaison en chaîne.

14.2 Associativité

Pas plus que pour la commutativité, la question de l'associativité ne se pose pas pour la règle de combinaison en chaîne. Cette règle **n'est donc pas associative**, comme le montrera cette section. Considérons une situation avec trois indications, $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$, $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$ et $Con_3 = (\Omega_3, P_3, \Delta_3, \Theta)$, enchaînées selon le schéma suivant :

$$Con_3 \underset{\omega_{3,c}}{\rightsquigarrow} Con_2 \underset{\omega_{2,c}}{\rightsquigarrow} \mathcal{H}_1.$$

Soit le résultat de cette combinaison en chaîne la meta-indication $\mathcal{M}''_{\mathcal{H}_1} := (\Omega'', P'', \Gamma'', \Theta)$, c'est-à-dire

$$\mathcal{M}''_{\mathcal{H}_1} = Con_3 \underset{\omega_{3,c}}{\rightsquigarrow} Con_2 \underset{\omega_{2,c}}{\rightsquigarrow} \mathcal{H}_1.$$

Notons qu'il n'y a pas d'interprétation de Ω_3 qui confirme \mathcal{H}_1 (contrairement à l'exemple 12.2 avec trois indications du chapitre précédent).

Pour que la règle de combinaison en chaîne soit une opération associative, il faudrait que nous obtenions le même résultat, peu importe dans quel ordre nous commençons à enchaîner ces trois indications. Est-ce que c'est le cas, c'est-à-dire est-ce que

$$(Con_3 \underset{\omega_{3,c}}{\wp} Con_2) \underset{\omega_{2,c}}{\wp} \mathcal{H}_1 = Con_3 \underset{\omega_{3,c}}{\wp} (Con_2 \underset{\omega_{2,c}}{\wp} \mathcal{H}_1)? \quad (14.3)$$

La partie de gauche de l'équation n'est tout simplement **pas définie**!

Si nous commençons par enchaîner Con_3 à Con_2 , nous nous retrouverons rapidement dans une impasse. En effet, si nous voulons respecter les étapes de la combinaison en chaîne définies dans la section précédente, nous devons en premier lieu construire le schéma de la situation. Etant donné que Con_3 et Con_2 sont toutes deux des indications de confirmation, il nous manque dans cette combinaison une indication classique qui, rappelons-le, doit toujours représenter le dernier maillon de la chaîne. L'indication Con_2 étant une indication de confirmation, elle ne peut occuper la place de ce maillon. De plus, lorsqu'une indication est une indication de confirmation, elle ne peut exister indépendamment de l'indication qu'elle confirme, vu qu'elle n'acquiert de sens qu'à travers celle-ci. En commençant la combinaison en chaîne par Con_3 et Con_2 , l'interprétation de Ω_2 qui confirme \mathcal{H}_1 ne pourrait être intégrée dans le schéma de la situation sans \mathcal{H}_1 elle-même. De ce fait, la combinaison en chaîne

$$(Con_3 \underset{\omega_{3,c}}{\wp} Con_2)$$

ne peut ni exister, ni aboutir à un résultat indépendamment de \mathcal{H}_1 . Les indications Con_2 et Con_3 ne peuvent à elles seules aboutir à une meta-indication.

14.3 Distributivité de la règle de combinaison en chaîne par rapport à la règle de combinaison de Dempster

Finalement, la question de la distributivité, par rapport à la règle de combinaison de Dempster, ne se pose pas non plus pour la règle de combinaison en chaîne. Cette règle **n'est pas distributive** par rapport à la règle de combinaison de Dempster, comme le montrera cette section.

Soient, en effet, $Con_1 = (\Omega_1, P_1, \Delta_1, \Theta)$ une indication de confirmation et $\mathcal{H}_2 = (\Omega_2, P_2, \Gamma_2, \Theta)$ et $\mathcal{H}_3 = (\Omega_3, P_3, \Gamma_3, \Theta)$ deux indications classiques. De même, soient $\mathcal{H}'_1 = (\Omega'_1, P'_1, \Gamma'_1, \Theta)$ une indication classique et $Con'_2 = (\Omega'_2, P'_2, \Delta'_2, \Theta)$ et $Con'_3 = (\Omega'_3, P'_3, \Delta'_3, \Theta)$ deux indications de confirmation. Si

$$Con_1 \underset{\omega_{1,c}}{\wp} (\mathcal{H}_2 \oplus \mathcal{H}_3) = (Con_1 \underset{\omega_{1,c}}{\wp} \mathcal{H}_2) \oplus (Con_1 \underset{\omega_{1,c}}{\wp} \mathcal{H}_3), \quad (14.4)$$

où $\omega_{1,c}$ est l'interprétation de Ω_1 qui confirme \mathcal{H}_2 et \mathcal{H}_3 , et si

$$(Con'_2 \oplus Con'_3) \underset{\omega_{4,c}}{\wp} \mathcal{H}'_1 = (Con'_2 \underset{\omega_{2,c}}{\wp} \mathcal{H}'_1) \oplus (Con'_3 \underset{\omega_{3,c}}{\wp} \mathcal{H}'_1), \quad (14.5)$$

où $\omega_{2,c}$ et $\omega_{3,c}$ sont les interprétations de Ω'_2 et de Ω'_3 respectivement qui confirment \mathcal{H}'_1 , et $\omega_{4,c}$ est l'interprétation de l'ensemble des interprétations Ω'_4 résultant de la combinaison de Ω'_2 avec Ω'_3 qui confirme \mathcal{H}'_1 , alors φ serait distributive par rapport à \oplus .

On peut se poser la question si une interprétation $\omega_{4,c}$ qui confirme \mathcal{H}'_1 existera après la combinaison de Con'_2 et Con'_3 . Notons également que dans ce cas, nous devons combiner par la règle de combinaison de Dempster deux indications de confirmation, opération qui pour l'instant n'est pas définie. Rappelons que la règle de combinaison de Dempster ne concerne que les indications classiques! Une adaptation de cette règle aux indications de confirmation sera présentée au Chapitre 16. En attendant, nous n'analyserons pas ce cas.

Revenons à l'équation (14.4), c'est-à-dire est-ce que

$$Con_1 \underset{\omega_{1,c}}{\varphi} (\mathcal{H}_2 \oplus \mathcal{H}_3) = (Con_1 \underset{\omega_{1,c}}{\varphi} \mathcal{H}_2) \oplus (Con_1 \underset{\omega_{1,c}}{\varphi} \mathcal{H}_3)?$$

Analysons ce que l'on obtient en combinant Con_1 , \mathcal{H}_2 et \mathcal{H}_3 de la manière suivante :

$$(Con_1 \underset{\omega_{1,c}}{\varphi} \mathcal{H}_2) \oplus (Con_1 \underset{\omega_{1,c}}{\varphi} \mathcal{H}_3).$$

Un problème surgit, car Con_1 ne peut pas confirmer deux indications classiques en même temps, à moins que l'une soit confirmée et l'autre infirmée, donc que les deux indications classiques se contredisent. Par exemple, si \mathcal{H}_2 nous dit qu'il pleuvra demain et que \mathcal{H}_3 nous dit le contraire, alors nous pouvons imaginer la situation suivante : $(Con_1 \underset{\omega_{1,c}}{\varphi} \mathcal{H}_2)$ et $(Con_1 \underset{\omega_{1,f}}{\varphi} \mathcal{H}_3)$, où Con_1 confirme et infirme chacune d'entre elles avec une certaine probabilité. La seule manière d'imaginer Con_1 qui confirme deux indications classiques en même temps, sans que confirmer l'une revienne à infirmer l'autre est de la représenter deux fois, à condition que les deux indications classiques proviennent de la même source d'information et que Con_1 confirme ou infirme cette même source. Cela dit, si c'était le cas, on devrait modéliser Con_1 par deux indications de confirmation, Con'_1 qui confirme \mathcal{H}_2 et Con''_1 qui confirme \mathcal{H}_3 . Ce serait la seule manière possible de représenter $(Con'_1 \underset{\omega'_{1,c}}{\varphi} \mathcal{H}_2)$ et $(Con''_1 \underset{\omega''_{1,c}}{\varphi} \mathcal{H}_3)$. Prenons un exemple :

Exemple 14.1. Soient \mathcal{H}_2 et \mathcal{H}_3 deux indications classiques sont fournies par Paul et soit Con_1 une indication de confirmation de émanant de Jean, qui prétend que Paul ment. Soient enfin les deux interprétations de Con_1 , $\omega_{1,c}$ et $\omega_{1,f}$, telles que :

- $\omega_{1,c}$: Jean se trompe,
- $\omega_{1,f}$: Jean a raison.

On peut considérer cette situation de deux manières. En effet, soit nous combinerons \mathcal{H}_2 et \mathcal{H}_3 par la règle de combinaison de Dempster, puis combinerons ce résultat avec Con_1 à travers la règle de combinaison en chaîne, soit nous

combinerons tout d'abord en chaîne Con_1 avec \mathcal{H}_2 et Con_1 avec \mathcal{H}_3 , puis nous combinerons par la règle de Dempster ces deux résultats. Pourtant, ces deux manières de modéliser ne représentent pas la même situation !

Dans le premier cas de figure,

$$Con_1 \underset{\omega_{1,c}}{\vartriangleright} (\mathcal{H}_2 \oplus \mathcal{H}_3),$$

nous modélisons une situation dans laquelle, si Paul ment, il le fait à la fois pour \mathcal{H}_2 et \mathcal{H}_3 . De même, s'il dit la vérité, c'est-à-dire si Jean se trompe, \mathcal{H}_2 et \mathcal{H}_3 sont vraies.

Dans le deuxième cas de figure,

$$(Con'_1 \underset{\omega'_{1,c}}{\vartriangleright} \mathcal{H}_2) \oplus (Con''_1 \underset{\omega''_{1,c}}{\vartriangleright} \mathcal{H}_3),$$

la situation qui y est représentée est différente. Cette fois-ci la situation décrite comporte certes les mêmes indications que précédemment, cependant, \mathcal{H}_2 et \mathcal{H}_3 ne sont plus considérées simultanément. On admet que Paul pourrait mentir pour l'une des deux indications qu'il donne et dire la vérité pour la seconde. En d'autres termes, Jean pourrait avoir raison pour \mathcal{H}_2 par exemple et se tromper pour \mathcal{H}_3 .

Cet exemple simple souligne toute l'importance de la manière de modéliser un problème. En partant d'une même source d'information, nous pouvons aboutir à plusieurs résultats différents si nous ne respectons pas le schéma logique du problème. Donc, l'analyse de la distributivité de la règle de combinaison en chaîne par rapport à la règle de combinaison de Dempster ne fait pas beaucoup de sens. Par contre, la question la plus importante que l'on doit se poser est : quelle est **la modélisation correcte** par rapport à une situation donnée.

14.4 Distributivité de la règle de combinaison de Dempster par rapport à la règle de combinaison en chaîne

Examinons à présent si la règle de combinaison de Dempster est distributive par rapport à la règle de combinaison en chaîne. Soient $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$, $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$ et $\mathcal{H}_3 = (\Omega_3, P_3, \Gamma_3, \Theta)$ trois indications.

Si

$$\mathcal{H}_1 \oplus (Con_2 \underset{\omega_{2,c}}{\vartriangleright} \mathcal{H}_3) = (\mathcal{H}_1 \oplus Con_2) \underset{\omega_{\bullet,c}}{\vartriangleright} (\mathcal{H}_1 \oplus \mathcal{H}_3), \quad (14.6)$$

où $\omega_{2,c}$ est l'interprétation de Ω_2 qui confirme \mathcal{H}_3 , et $\omega_{\bullet,c}$ est l'interprétation de l'ensemble des interprétations de $(\mathcal{H}_1 \oplus Con_2)$ qui confirme \mathcal{H}_3 , et si

$$(Con_2 \underset{\omega_{2,c}}{\rightsquigarrow} \mathcal{H}_3) \oplus \mathcal{H}_1 = (Con_2 \oplus \mathcal{H}_1) \underset{\omega_{\bullet,c}}{\rightsquigarrow} (\mathcal{H}_3 \oplus \mathcal{H}_1), \quad (14.7)$$

où $\omega_{2,c}$ est l'interprétation de Ω_2 qui confirme \mathcal{H}_3 , et $\omega_{\bullet,c}$ est l'interprétation de l'ensemble des interprétations de $(Con_2 \oplus \mathcal{H}_1)$ qui confirme \mathcal{H}_3 , alors \oplus serait distributive par rapport à \rightsquigarrow .

Comme précédemment, une analyse de la distributivité de la règle de combinaison de Dempster par rapport à la règle de combinaison en chaîne ne fait aucun sens. La seule manière de combiner \mathcal{H}_1 , Con_2 et \mathcal{H}_3 , serait de procéder selon un des deux schémas suivants : $\mathcal{H}_1 \oplus (Con_2 \underset{\omega_{2,c}}{\rightsquigarrow} \mathcal{H}_3)$ ou $(Con_2 \underset{\omega_{2,c}}{\rightsquigarrow} \mathcal{H}_3) \oplus \mathcal{H}_1$. Encore faudrait-il que l'un de ces deux schémas corresponde à ce que l'on aimerait modéliser.

Pour montrer qu'une analyse de la distributivité ne présente aucun intérêt, commençons par l'équation (14.6), c'est-à-dire est-ce que

$$\mathcal{H}_1 \oplus (Con_2 \underset{\omega_{2,c}}{\rightsquigarrow} \mathcal{H}_3) = (\mathcal{H}_1 \oplus Con_2) \underset{\omega_{2,\bullet}}{\rightsquigarrow} (\mathcal{H}_1 \oplus \mathcal{H}_3)?$$

Analysons à présent ce que l'on obtient en combinant \mathcal{H}_1 , Con_2 et \mathcal{H}_3 de la manière suivante :

$$(\mathcal{H}_1 \oplus Con_2) \underset{\omega_{\bullet,c}}{\rightsquigarrow} (\mathcal{H}_1 \oplus \mathcal{H}_3).$$

Le problème qui se pose à ce niveau-là est le manque de définition de l'opération $\mathcal{H}_1 \oplus Con_2$. En effet, la règle de combinaison de Dempster n'est définie que pour la combinaison d'indications classiques et non pour la combinaison d'indications de confirmation. Pour celles-ci, une adaptation de la règle de Dempster est présentée au Chapitre 16. Cependant, cette règle adaptée ne permet que la combinaison d'indications de confirmation entre elles. Qu'en est-il du cas où nous voulons combiner par la règle de Dempster des indications de différents types : classiques et de confirmation? Développer une nouvelle opération de combinaison est peut-être possible ; cela dit, est-ce nécessaire ou même utile? Combiner des indications par la règle de Dempster ou une adaptation de celle-ci revient à les placer sur un pied d'égalité, à un même niveau. Nous avons déjà considéré de nombreux exemples impliquant des indications classiques. Pour les indications de confirmation, une telle adaptation de la règle de Dempster est utilisée lorsque toute une série d'indications de confirmation confirment une même indication classique. Cette nouvelle règle nous permet de combiner toute l'information contenue dans la série pour en créer une seule et unique indication de confirmation.

Pour revenir à la combinaison d'indications de confirmation avec des indications classiques par une adaptation de la règle de Dempster, il est facile de se rendre compte qu'un tel cas de figure ne peut se produire dans la réalité. L'objectif de la règle de combinaison de Dempster et d'intégrer toute nouvelle information au moment de son apparition. Dans

notre cas, Con_2 confirme \mathcal{H}_3 . Lorsque dans l'équation (14.6) \mathcal{H}_1 apparaît, celle-ci doit être intégrée à $\mathcal{M}' = Con_2 \underset{\omega_{2,c}}{\varphi} \mathcal{H}_3$ et non à Con_2 toute seule. Autrement, nous ne disposerons pas de l'information nécessaire pour savoir par quel schéma les combiner. L'indication Con_2 est **indissociable** de \mathcal{H}_3 et de ce fait ne peut être combinée sans elle.

On peut ainsi conclure que la règle de combinaison de Dempster, **n'est pas distributive** par rapport à la règle de combinaison en chaîne.

Chapitre 15

Combinaison d'indications de confirmation entre elles

15.1 Contexte

On a déjà mentionné la nécessité d'adapter la règle de combinaison de Dempster aux indications de confirmation. Dans ce chapitre, une telle adaptation sera développée. Pour mieux comprendre la démarche, on commencera par un exemple simple. Soient deux indications de confirmation qu'il serait nécessaire de combiner par une règle, dont on voudrait qu'elle reste dans le même esprit que celle de Dempster. Il s'agit en effet de deux indications de confirmation qui confirment la même indication classique, mais qui n'ont pas de lien de confirmation entre elles. On ne peut donc pas combiner ces trois indications en chaîne. Il faudrait dans un premier temps combiner les deux indications de confirmation en une seule et unique indication de confirmation, puis combiner celle-ci à l'indication classique.

Exemple 15.1. Le client d'un casino décide de jouer au jeu de roulette. Il a le choix entre 10 tables différentes. Il en choisit une et veut maintenant savoir quel nombre sortira gagnant.

Soient les trois indications suivantes :

1. \mathcal{H}_1 : La table de roulette choisie est équilibrée (c'est-à-dire avec loi de probabilité uniforme sur les 37 nombres, allant de 0 à 36).
2. Con_2 : Un croupier et ami du joueur lui dit qu'une des dix tables est truquée, mais il ignore laquelle.
3. Con_3 : Un autre client informe le joueur que trois tables parmi les dix sont cassées, en ce sens qu'elles ne sont plus tout à fait équilibrées. Cependant, il ignore lesquelles.

Voici comment se présentent ces indications en détail :

1. $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$, avec
 - $\Omega_1 = \{\omega_{1,1}, \omega_{1,2}, \dots, \omega_{1,37}\}$
 - $\Gamma_1(\omega_{1,j}) = \{(j-1)\}$ où $j = 1, \dots, 37$
 - $p_1(\omega_{1,j}) = \frac{1}{37}$ où $j = 1, \dots, 37$
 - $\Theta = \{0, 1, 2, \dots, 36\}$
2. $Con_2 = (\Omega_2, P_2, \Delta_2, \Theta)$, avec
 - $\Omega_2 = \{\omega_{2,1}, \omega_{2,2}\}$, où
 - $\omega_{2,1}$: le joueur se trouve sur une table non truquée
 - $\omega_{2,2}$: le joueur se trouve sur une table truquée
 - $\Delta_2 : \Omega_2 \rightarrow \{\Omega_1, \Theta\}$, où
 - $\Delta_2(\omega_{2,1}) = \Omega_1$
 - $\Delta_2(\omega_{2,2}) = \Theta$
 - $p_2(\omega_{2,1}) = \frac{9}{10}$ et $p_2(\omega_{2,2}) = \frac{1}{10}$
 - $\Theta = \{0, 1, 2, \dots, 36\}$
3. $Con_3 = (\Omega_3, P_3, \Delta_3, \Theta)$, avec
 - $\Omega_3 = \{\omega_{3,1}, \omega_{3,2}\}$, où
 - $\omega_{3,1}$: le joueur se trouve sur une table non cassée
 - $\omega_{3,2}$: le joueur se trouve sur une table cassée
 - $\Delta_3 : \Omega_3 \rightarrow \{\Omega_1, \Theta\}$, où
 - $\Delta_3(\omega_{3,1}) = \Omega_1$
 - $\Delta_3(\omega_{3,2}) = \Theta$
 - $p_3(\omega_{3,1}) = \frac{7}{10}$ et $p_3(\omega_{3,2}) = \frac{3}{10}$
 - $\Theta = \{0, 1, 2, \dots, 36\}$

Le schéma graphique des applications Γ_1 , Δ_2 et Δ_3 est présenté à la Figure 15.1.

Comment combiner les trois indications de cet exemple ?

15.2 Stratégie 1

$$(Con_2 \underset{\omega_{2,c}}{\rightsquigarrow} \mathcal{H}_1) \oplus (Con_3 \underset{\omega_{3,c}}{\rightsquigarrow} \mathcal{H}_1)$$

Soit $\mathcal{M}_4 = (Con_2 \underset{\omega_{2,c}}{\rightsquigarrow} \mathcal{H}_1)$. Elle est définie par $(\Omega_4, P_4, \Gamma_4, \Theta)$, avec

$$- \Omega_4 = \{\omega_{1,1}, \omega_{1,2}, \dots, \omega_{1,37}, \omega_{2,2}\}$$

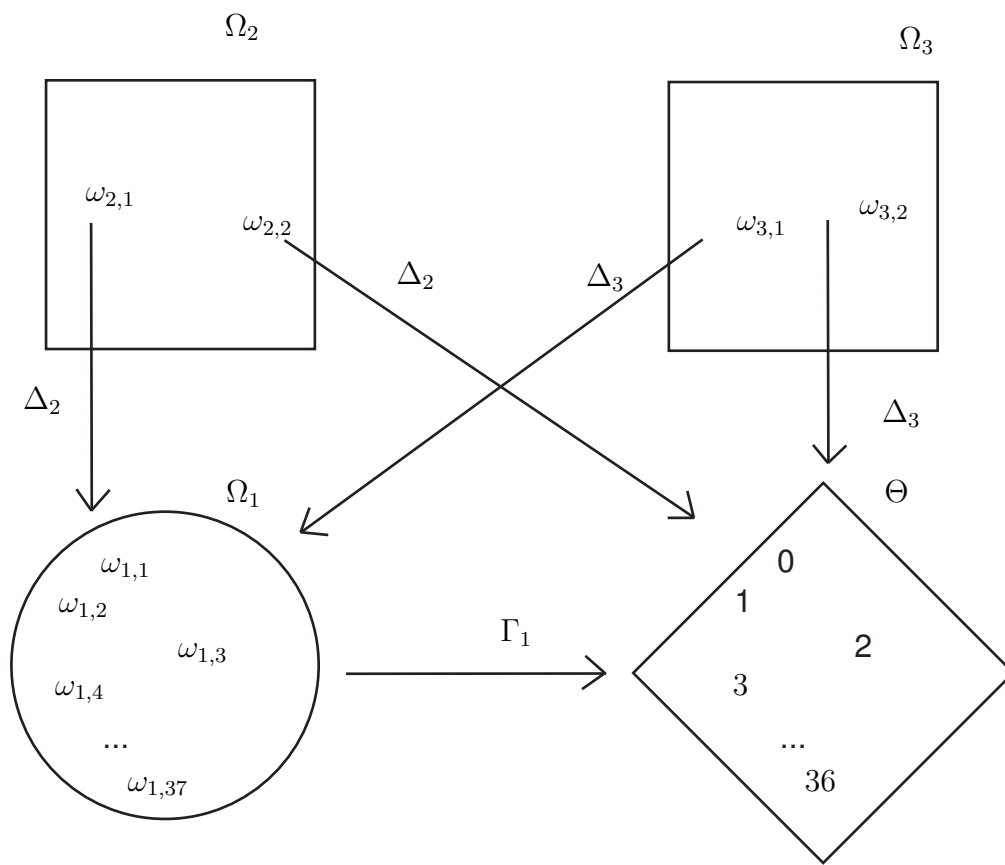


FIG. 15.1 – Exemple de la roulette avec deux indications de confirmation

- $\Gamma_4(\omega_{1,j}) = \{(j-1)\}$ où $j = 1, \dots, 37$ et $\Gamma_4(\omega_{2,2}) = \Theta$
- $p_4(\omega_{1,j}) = \frac{9}{10} \frac{1}{37}$ où $j = 1, \dots, 37$ et $p_4(\omega_{2,2}) = \frac{1}{10}$
- $\Theta = \{0, 1, 2, \dots, 36\}$

Soit $\mathcal{M}_5 = (Con_3 \underset{\omega_{2,c}}{\varphi} \mathcal{H}_1)$. Elle est définie par $(\Omega_5, P_5, \Gamma_5, \Theta)$, avec

- $\Omega_5 = \{\omega_{1,1}, \omega_{1,2}, \dots, \omega_{1,37}, \omega_{3,2}\}$
- $\Gamma_5(\omega_{1,j}) = \{(j-1)\}$ où $j = 1, \dots, 37$ et $\Gamma_5(\omega_{3,2}) = \Theta$
- $p_5(\omega_{1,j}) = \frac{7}{10} \frac{1}{37}$ où $j = 1, \dots, 37$ et $p_5(\omega_{3,2}) = \frac{3}{10}$
- $\Theta = \{0, 1, 2, \dots, 36\}$

Combinons à présent \mathcal{M}_4 et \mathcal{M}_5 par la règle de combinaison de Dempster. Rappelons que cette règle s'applique sans problème aux meta-indications \mathcal{M} , car celles-ci ont la même structure que les indications classiques \mathcal{H} .

Si l'on applique la règle de combinaison de Dempster, l'indication $\mathcal{M}_6 = \mathcal{M}_4 \oplus \mathcal{M}_5$ est définie par $(\Omega_6, P_6, \Gamma_6, \Theta)$, avec

- $$\Omega_6 = \left\{ \begin{array}{cccccc} (\omega_{1,1}, \omega_{1,1}), & (\omega_{1,1}, \omega_{1,2}), & \dots & (\omega_{1,1}, \omega_{1,37}), & (\omega_{1,1}, \omega_{2,2}) \\ (\omega_{1,2}, \omega_{1,1}), & (\omega_{1,2}, \omega_{1,2}), & \dots & (\omega_{1,2}, \omega_{1,37}), & (\omega_{1,2}, \omega_{2,2}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ (\omega_{1,37}, \omega_{1,1}), & (\omega_{1,37}, \omega_{1,2}), & \dots & (\omega_{1,37}, \omega_{1,37}), & (\omega_{1,37}, \omega_{2,2}) \\ (\omega_{3,2}, \omega_{1,1}), & (\omega_{3,2}, \omega_{1,2}), & \dots & (\omega_{3,2}, \omega_{1,37}), & (\omega_{3,2}, \omega_{2,2}) \end{array} \right\}$$
- $\Gamma_6(\omega_{1,j}, \omega_{1,j}) = \{(j-1)\}$, où $j = 1, \dots, 37$
- $\Gamma_6(\omega_{1,j}, \omega_{1,i}) = \emptyset$, où $j = 1, \dots, 37$, $i = 1, \dots, 37$ et $j \neq i$,
- $\Gamma_6(\omega_{1,j}, \omega_{2,2}) = \{(j-1)\}$, où $j = 1, \dots, 37$
- $\Gamma_6(\omega_{3,2}, \omega_{1,j}) = \{(j-1)\}$, où $j = 1, \dots, 37$
- $\Gamma_6(\omega_{3,2}, \omega_{2,2}) = \Theta$
- Fonction de probabilité P'_6 avant normalisation
 - $p'_6(\omega_{1,j}, \omega_{1,j}) = \frac{9}{10} \frac{1}{37} \frac{7}{10} \frac{1}{37}$, où $j = 1, \dots, 37$
 - $p'_6(\omega_{1,j}, \omega_{1,i}) = \frac{9}{10} \frac{1}{37} \frac{7}{10} \frac{1}{37}$, où $j = 1, \dots, 37$, $i = 1, \dots, 37$ et $j \neq i$,
 - $p'_6(\omega_{1,j}, \omega_{2,2}) = \frac{7}{10} \frac{1}{37} \frac{1}{10}$, où $j = 1, \dots, 37$
 - $p'_6(\omega_{3,2}, \omega_{1,j}) = \frac{9}{10} \frac{1}{37} \frac{3}{10}$, où $j = 1, \dots, 37$
 - $p'_6(\omega_{3,2}, \omega_{2,2}) = \frac{3}{10} \frac{1}{10}$
- Fonction de probabilité P_6 après normalisation
 - $p_6(\omega_{1,j}, \omega_{1,j}) = (\frac{9}{10} \frac{1}{37} \frac{7}{10} \frac{1}{37}) / (1 - 37 \cdot 36 \frac{9}{10} \frac{1}{37} \frac{7}{10} \frac{1}{37})$, où $j = 1, \dots, 37$
 - $p_6(\omega_{1,j}, \omega_{2,2}) = (\frac{7}{10} \frac{1}{37} \frac{1}{10}) / (1 - 37 \cdot 36 \frac{9}{10} \frac{1}{37} \frac{7}{10} \frac{1}{37})$, où $j = 1, \dots, 37$
 - $p_6(\omega_{3,2}, \omega_{1,j}) = (\frac{9}{10} \frac{1}{37} \frac{3}{10}) / (1 - 37 \cdot 36 \frac{9}{10} \frac{1}{37} \frac{7}{10} \frac{1}{37})$, où $j = 1, \dots, 37$
 - $p_6(\omega_{3,2}, \omega_{2,2}) = (\frac{3}{10} \frac{1}{10}) / (1 - 37 \cdot 36 \frac{9}{10} \frac{1}{37} \frac{7}{10} \frac{1}{37})$
- $\Theta = \{0, 1, 2, \dots, 36\}$

Supposons que nous considérons $\omega_{1,9}$. Dans \mathcal{H}_1 , cette interprétation signifie : la table est équilibrée et c'est le 8 qui sortira. Son ensemble focal est donc donné par $\{8\}$. Nous avons vu que l'ensemble focal de $\omega_{2,2}$ (la table est truquée) est donné par Θ , étant donné que nous sommes incapables de réduire l'ensemble des réponses à un sous-ensemble plus petit. A présent, si l'on combinait ces deux interprétations par la règle de combinaison de Dempster dans une nouvelle indication, leur ensemble focal serait $\{8\}$, et cela bien que la table soit truquée. Or, la seule chose que l'on sait est que la table est truquée, mais on ignore comment. Rien ne nous garantit de ce fait que le nombre 8 sortira gagnant. Le problème provient du fait que $\omega_{1,9}$ et $\omega_{2,2}$ ne sont pas contradictoires (l'intersection de leurs ensembles focaux n'est pas vide) et on peut très bien imaginer que le nombre 8 sorte gagnant d'une table truquée. Une contradiction existe malgré tout. La table ne peut à la fois être équilibrée et être truquée. Pourtant, $\omega_{1,9}$ ne dit pas uniquement que le nombre 8 sortira gagnant, mais qu'il sortira gagnant d'une table équilibrée ! Ainsi, ces deux interprétations ne peuvent être vraies en même temps, même si l'intersection de leurs ensembles focaux n'est pas vide. Nous sommes confrontés à un problème de contradiction qui dépasse la définition de la contradiction donnée par Dempster et cela malgré le fait que nous combinons des meta-indications pour lesquelles cette règle s'applique. On constate ainsi que l'application de la règle de combinaison de Dempster à deux meta-indications basées sur une même indication classique **pose problème**. En considérant \mathcal{H}_1 «doublement», on obtient des résultats problématiques.

15.3 Stratégie 2

Rappelons que dans le cadre de la règle de combinaison de Dempster, lors de la combinaison de deux ou plusieurs indications, les interprétations combinées contenant des interprétations contradictoires sont éliminées, bien que celles-ci puissent être de probabilité non nulle. La fonction de probabilité sur les interprétations combinées restantes est ensuite normalisée. Ainsi, dans l'esprit de cette règle, une combinaison d'informations contradictoires doit être éliminée du modèle, étant donné qu'elle ne pourra se réaliser.

Dans la théorie mathématique des indications de Kohlas et Monney, la définition de la contradiction passe par la fonction Γ . En effet, deux ou plusieurs interprétations sont dites «contradictoires» si l'intersection de leurs ensembles focaux aboutit à l'ensemble vide.

A la place, on propose de considérer comme contradictoires toutes les interprétations qui ne peuvent se réaliser simultanément.

Considérons en effet les ensembles des interprétations des deux meta-indications \mathcal{M}_4 et \mathcal{M}_5 .

Commençons par $\Omega_4 = \{\omega_{1,1}, \omega_{1,2}, \dots, \omega_{1,37}, \omega_{2,2}\}$. Voici la signification des 38 interprétations de cet ensemble :

1. les 37 premières $\omega_{1,j}$ signifient : *la table est équilibrée et un nombre précis entre 0 et 36 sortira gagnant,*
2. la dernière interprétation $\omega_{2,2}$ signifie : *la table est truquée (donc pas équilibrée) et on ne peut rien dire quant au nombre qui sortira gagnant.*

A présent, considérons $\Omega_5 = \{\omega_{1,1}, \omega_{1,2}, \dots, \omega_{1,37}, \omega_{3,2}\}$. Voici la signification des 38 interprétations de cet ensemble :

1. les 37 premières $\omega_{1,j}$ signifient : *la table est équilibrée et un nombre précis entre 0 et 36 sortira gagnant,*
2. la dernière interprétation $\omega_{3,2}$ signifie : *la table est cassée (donc pas équilibrée) et on ne peut rien dire quant au nombre qui sortira gagnant.*

Si l'on revient à la combinaison $\mathcal{M}_6 = \mathcal{M}_4 \oplus \mathcal{M}_5$ de la section 15.2, nous avons obtenu des interprétations combinées dans Ω_6 dont l'intersection des ensembles focaux n'est pas vide, mais lesquels, sont malgré tout contradictoires : $(\omega_{1,j}, \omega_{2,2})$ et $(\omega_{3,2}, \omega_{1,j})$, où $j = 1, \dots, 37$ et leurs ensembles focaux sont donnés par $\{(j-1)\}$. On peut expliquer ce phénomène de deux manières :

1. Il est clair que $\omega_{1,j}$ et $\omega_{2,2}$, avec $j = 1, \dots, 37$, ne peuvent être correctes simultanément ; les $\omega_{1,j}$ signifient que la table est équilibrée, tandis que d'après $\omega_{2,2}$ elle est truquée. Bien sûr, une table de casino peut être truquée sans être cassée, cependant, dans notre cas, la question porte avant tout sur le nombre qui sortira gagnant ou, autrement dit, est-ce que la table est équilibrée ou non. En résumé, qu'elle soit non équilibrée, parce qu'étant cassée, ou parce qu'étant truquée, nous importe peu ; nous désirons uniquement savoir si Ω_1 est confirmée ou non. Si une des $\omega_{1,j}$ est l'interprétation correcte, alors Ω_1 a été confirmé (la table est équilibrée) et par conséquent elle ne peut être truquée. Le problème se résume ainsi à savoir si \mathcal{H}_1 est confirmée ou non, c'est-à-dire si elle est vraie ou fausse. Lorsqu'elle est n'est pas confirmée, et peu en importe la raison, nous en déduisons simplement qu'elle est fausse.
2. Une deuxième raison de la contradiction entre $\omega_{1,j}$ et $\omega_{2,2}$, avec $j = 1, \dots, 37$ relève de la composition de Ω_4 . En effet, dans les couples $(\omega_{1,j}, \omega_{2,2})$, les $\omega_{1,j}$ proviennent de Ω_5 ; cependant, ils sont tout à fait identiques aux $\omega_{1,j}$ appartenant à Ω_4 . Rappelons que $\Omega_4 = \{\omega_{1,1}, \omega_{1,2}, \dots, \omega_{1,37}, \omega_{2,2}\}$. De même, pour tout Ω possédant m éléments,

$$\sum_{i=1}^m p(\omega_i) = 1.$$

Ainsi, un ensemble d'interprétations est défini comme un ensemble d'interprétations fini, dont une et une seule des interprétations peut être l'interprétation correcte. De ce fait, si l'une des 37 $\omega_{1,j}$ se révèle être l'interprétation correcte, l'interprétation $\omega_{2,2}$ se révèle automatiquement fausse.

Tous les arguments pour la contradiction entre $\omega_{1,j}$ et $\omega_{2,2}$ sont également valables pour les couples $(\omega_{3,2}, \omega_{1,j})$, avec $j = 1, \dots, 37$.

Ainsi, nous avons montré que des interprétations peuvent se contredire, tout en ayant une intersection non vide de leurs ensembles focaux.

Revenons à présent à la règle de combinaison de Dempster. D'après celle-ci, lorsqu'on combine des indications, il faut éliminer toutes les interprétations combinées qui contiennent des interprétations se contredisant entre elles. Le but de cette démarche est avant tout de normaliser la fonction de probabilité par rapport aux interprétations possibles uniquement, toutes celles qui ne pourraient se réaliser étant déduites de l'ensemble d'interprétations de l'indication combinée. Nous avons vu que des interprétations pouvaient se contredire, en ce sens qu'elles ne peuvent se réaliser simultanément, tout en sortant de la définition formelle de la contradiction de la règle de Dempster.

Essayons à présent de combiner \mathcal{M}_4 et \mathcal{M}_5 par la règle de combinaison de Dempster. De même, considérons que nous devons éliminer toutes les interprétations combinées contradictoires, soit toutes celles qui ne peuvent se réaliser simultanément, indépendamment du fait que leurs ensembles focaux sont disjoints ou non.

Notons cette opération par le symbole $\dot{\oplus}$, afin d'éviter la confusion avec la règle de combinaison de Dempster.

L'indication $\mathcal{M}_7 = \mathcal{M}_4 \dot{\oplus} \mathcal{M}_5$ est définie par $(\Omega_7, P_7, \Gamma_7, \Theta)$, avec

-
- $$\Omega_7 = \left\{ \begin{array}{cccccc} (\omega_{1,1}, \omega_{1,1}), & (\omega_{1,1}, \omega_{1,2}), & \dots & (\omega_{1,1}, \omega_{1,37}), & (\omega_{1,1}, \omega_{2,2}) \\ (\omega_{1,2}, \omega_{1,1}), & (\omega_{1,2}, \omega_{1,2}), & \dots & (\omega_{1,2}, \omega_{1,37}), & (\omega_{1,2}, \omega_{2,2}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ (\omega_{1,37}, \omega_{1,1}), & (\omega_{1,37}, \omega_{1,2}), & \dots & (\omega_{1,37}, \omega_{1,37}), & (\omega_{1,37}, \omega_{2,2}) \\ (\omega_{3,2}, \omega_{1,1}), & (\omega_{3,2}, \omega_{1,2}), & \dots & (\omega_{3,2}, \omega_{1,37}), & (\omega_{3,2}, \omega_{2,2}) \end{array} \right\}$$
- $\Gamma_7(\omega_{1,j}, \omega_{1,j}) = \{(j-1)\}$, où $j = 1, \dots, 37$,
- $\Gamma_7(\omega_{1,j}, \omega_{1,i}) = \emptyset$, où $j = 1, \dots, 37$, $i = 1, \dots, 37$ et $j \neq i$, contradiction !
- $\Gamma_7(\omega_{1,j}, \omega_{2,2}) = \{(j-1)\}$, où $j = 1, \dots, 37$, contradiction !
- $\Gamma_7(\omega_{3,2}, \omega_{1,j}) = \{(j-1)\}$, où $j = 1, \dots, 37$, contradiction !
- $\Gamma_7(\omega_{3,2}, \omega_{2,2}) = \Theta$
- Fonction de probabilité p'_7 avant normalisation
 - $p'_7(\omega_{1,j}, \omega_{1,j}) = \frac{9}{10} \frac{1}{37} \frac{7}{10} \frac{1}{37}$, où $j = 1, \dots, 37$
 - $p'_7(\omega_{1,j}, \omega_{1,i}) = \frac{9}{10} \frac{1}{37} \frac{7}{10} \frac{1}{37}$, où $j = 1, \dots, 37$, $i = 1, \dots, 37$ et $j \neq i$,
 - $p'_7(\omega_{1,j}, \omega_{2,2}) = \frac{7}{10} \frac{1}{37} \frac{1}{10}$, où $j = 1, \dots, 37$
 - $p'_7(\omega_{3,2}, \omega_{1,j}) = \frac{9}{10} \frac{1}{37} \frac{3}{10}$, où $j = 1, \dots, 37$
 - $p'_7(\omega_{3,2}, \omega_{2,2}) = \frac{3}{10} \frac{1}{10}$
- Fonction de probabilité P_7 après normalisation
 - $p_7(\omega_{1,j}, \omega_{1,j}) = \left(\frac{9}{10} \frac{1}{37} \frac{7}{10} \frac{1}{37} \right) / \left(1 - \left(37 \cdot 36 \frac{9}{10} \frac{1}{37} \frac{7}{10} \frac{1}{37} + 37 \frac{7}{10} \frac{1}{37} \frac{1}{10} + 37 \frac{9}{10} \frac{1}{37} \frac{3}{10} \right) \right) = \frac{63}{6438}$, où $j = 1, \dots, 37$
 - $p_7(\omega_{3,2}, \omega_{2,2}) = \left(\frac{3}{10} \frac{1}{10} \right) / \left(1 - \left(37 \cdot 36 \frac{9}{10} \frac{1}{37} \frac{7}{10} \frac{1}{37} + 37 \frac{7}{10} \frac{1}{37} \frac{1}{10} + 37 \frac{9}{10} \frac{1}{37} \frac{3}{10} \right) \right) = \frac{4107}{6438}$

$$- \Theta = \{0, 1, 2, \dots, 36\}$$

On pourrait penser avoir résolu le problème de la combinaison de \mathcal{M}_4 et \mathcal{M}_5 ; mais il n'en est rien pour l'instant. Si l'on considère le problème à sa base, on constate qu'il y a quatre situations possibles :

1. La table n'est ni truquée, ni cassée.
2. La table est truquée, mais pas cassée.
3. La table n'est pas truquée, mais elle est cassée.
4. La table est truquée et cassée.

Leurs probabilités respectives sont :

1. $\frac{9}{10} \frac{7}{10}$,
2. $\frac{1}{10} \frac{7}{10}$,
3. $\frac{9}{10} \frac{3}{10}$,
4. $\frac{1}{10} \frac{3}{10}$.

A présent, si l'on considère ces quatre situations en termes d'équilibre de la table, c'est-à-dire est-ce que \mathcal{H}_1 est confirmée ou non, nous n'avons plus que deux situations :

- a) La table est équilibrée.
- b) La table n'est pas équilibrée.

La première correspond à la situation où la table n'est ni truquée, ni cassée (seule manière de garantir son équilibre), et la seconde englobe les situations 2 à 4, c'est-à-dire toutes celles qui affectent l'équilibre de la table. Leurs probabilités respectives sont :

- a) $\frac{9}{10} \frac{7}{10}$,
- b) $1 - \frac{9}{10} \frac{7}{10}$.

Ainsi, si la table est équilibrée, chacune des interprétations $\omega_{1,j}$, avec $j = 1, \dots, 37$, de l'indication combinant toutes les informations devraient avoir une probabilité de $\frac{1}{37} \frac{9}{10} \frac{7}{10}$. Quant à la situation b), regroupant les interprétations $(\omega_{1,j}, \omega_{2,2})$, $(\omega_{3,2}, \omega_{1,j})$ et $(\omega_{3,2}, \omega_{2,2})$, où $j = 1, \dots, 37$, elle devraient être représentée par une seule interprétation disant que la table n'est pas équilibrée et de probabilité $1 - \frac{9}{10} \frac{7}{10}$.

Nous constatons que les résultats de la stratégie 2 ne correspondent pas à ceux qu'on devrait obtenir. Non seulement les interprétations correspondant à une table équilibrée voient leur probabilité baisser, par rapport à la probabilité qu'elles devraient avoir, mais en plus, la situation de table non équilibrée se résume à une seule des trois situations possibles, celle où la table est à la fois truquée et cassée, $(\omega_{3,2}, \omega_{2,2})$.

On constate ainsi que l'application de la règle de combinaison de Dempster, **dans une version modifiée**, à deux meta-indications basées sur une même indication classique **pose problème**. La stratégie 2 ne fonctionne donc pas.

15.4 Stratégie 3

Sur la base des problèmes mis en évidence dans les sections 15.2 et 15.3, il serait raisonnable de conclure que la meilleure manière de combiner \mathcal{H}_1 , Con_2 et Con_3 serait de combiner en premier lieu Con_2 avec Con_3 , étant donné que toutes deux confirment \mathcal{H}_1 , de manière à obtenir une seule et nouvelle indication de confirmation, puis de combiner celle-ci avec \mathcal{H}_1 par la règle de combinaison en chaîne. En effet, les problèmes rencontrés jusque là résultaient du fait que nous essayions de combiner deux meta-indications **basées sur la même indication classique**. Ce faisant, l'information contenue dans cette indication classique était considérée **doublement**, d'où les incohérences dans cette démarche.

Le problème majeur d'une telle démarche réside dans le fait que nous ne disposons pas d'opération permettant de combiner deux (ou plusieurs) indications de confirmation entre elles.

Une telle opération est proposée au Chapitre 16.

15.5 Règle de combinaison de Dempster et meta-indications

On a vu que la règle de combinaison en chaîne ne traite que de la combinaison entre des indications classiques et celles qui les confirment, tandis que la règle de combinaison de Dempster ne concerne que la combinaison d'indications classiques entre elles. Combiner \mathcal{H}_1 avec Con_2 , \mathcal{H}_1 avec Con_3 , puis les deux meta-indications ainsi obtenues par la règle de Dempster, n'est pas une solution non plus, comme nous l'avons constaté dans la stratégie 1. En réalité, le problème ne réside pas dans la combinaison de meta-indications par la règle de Dempster ; celle-ci leur convient parfaitement du fait que toute meta-indication est une indication classique également. Le problème vient du fait que les deux meta-indications ainsi obtenues sont toutes deux fonction de \mathcal{H}_1 . En opérant de cette manière, l'information véhiculée par \mathcal{H}_1 se retrouve doublée (car considérée deux fois) dans le résultat final. S'il s'agissait de combiner deux meta-indications fonctions de deux indications classiques distinctes, il n'y aurait aucun problème d'application de la règle de Dempster. Des expressions du type

$$(Con_a \underset{\omega_{a,c}}{\rightsquigarrow} \mathcal{H}_b) \oplus (Con_d \underset{\omega_{d,c}}{\rightsquigarrow} \mathcal{H}_e),$$

où $a \neq d$ et $b \neq e$, sont en effet parfaitement correctes et calculables, car dans ce cas, la règle de Dempster combine deux meta-indications véhiculant des informations de sources différentes. De plus, les indications de confirmation qui confirment les indications classiques doivent également être différentes. Si une même indication de confirmation Con confirmait simultanément \mathcal{H}_b et \mathcal{H}_e (configuration possible si l'une des interprétations de

Con confirme \mathcal{H}_b en infirmant \mathcal{H}_e et si la seconde confirme \mathcal{H}_e en infirmant \mathcal{H}_b , alors il suffirait tout simplement d'appliquer la règle de combinaison en chaîne.

En résumé, la règle de combinaison de Dempster, lorsqu'elle est utilisée pour combiner des meta-indications, doit être appliquée à des meta-indications fonctions d'indications classiques différentes et confirmées par des indications de confirmations différentes. Si tel est le cas, c'est-à-dire si nous sommes en présence de n indications classiques différentes confirmées par n indications de confirmations différentes, donnant lieu ainsi à n meta-indications différentes, alors pour combiner ces meta-indications entre elles nous appliquons la règle de Dempster et le résultat ainsi obtenu sera une indication classique :

$$\bigoplus_{i=1}^n (Con_i \underset{\omega_{i,c}}{\rightsquigarrow} \mathcal{H}_j), \quad (15.1)$$

avec $i = 1, \dots, n$ et $j = 1, \dots, n$.

Nous pouvons également appliquer la règle de Dempster à une combinaison de meta-indications et indications classiques. La seule condition à cette application reste évidemment qu'aucune des meta-indications ne soit construite sur la base d'une des indications classiques, situation difficilement envisageable d'ailleurs. En effet, si une indication classique venait à être confirmée se transformant ainsi en meta-indication, elle serait désormais considérée comme telle et n'interviendrait plus dans la combinaison en tant qu'indication classique. Il faut cependant rester vigilant aux indications classiques résultant de la combinaison par la règle de Dempster de plusieurs indications classiques. Celles-ci pourraient contenir des indications classiques sur lesquelles certaines meta-indications sont construites.

Chapitre 16

Adaptation de la règle de Dempster à la combinaison d'indications de confirmation

Revenons à présent au cas qui nous intéresse, c'est-à-dire comment combiner deux ou plusieurs indications de confirmation qui confirment toutes une même indication classique.

Soit $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$ une indication classique et soient Con_2, \dots, Con_n des indications de confirmation qui confirment toutes \mathcal{H}_1 . Chaque indication Con_i , avec $i = 2, \dots, n$ est définie par $(\Omega_i, P_i, \Delta_i, \Theta)$, où les deux éléments de Ω_i sont $\omega_{i,c}$ et $\omega_{i,f}$, le premier confirmant \mathcal{H}_1 et le deuxième l'infirmant. L'ensemble focal de chaque $\omega_{i,c}$ est donné par Ω_1 et de chaque $\omega_{i,f}$ par Θ respectivement.

Pour toutes les raisons susmentionnées, nous désirons en premier lieu combiner toutes les indications de confirmation Con_2, \dots, Con_n entre elles, afin d'obtenir une seule et unique indication de confirmation, puis la combiner avec \mathcal{H}_1 à travers la règle de combinaison en chaîne.

16.1 Combinaison de deux indications de confirmation

Reprenons pour cela l'exemple 15.1 avec \mathcal{H}_1 , Con_2 et Con_3 .

Exemple 16.1. Nous voulons combiner Con_2 et Con_3 de manière à obtenir une seule et unique indication de confirmation, appelons-la $Con_4 = (\Omega_4, P_4, \Delta_4, \Theta)$, qui contiendrait toute l'information véhiculée par Con_2 et Con_3 . L'indication de confirmation Con_4 , comme toute autre indication de confirmation, devrait avoir deux interprétations, $\omega_{4,c}$ et $\omega_{4,f}$,

dont la première confirmerait \mathcal{H}_1 et la seconde l'infirmerait. Leurs ensembles focaux respectifs seraient bien entendu donnés par Ω_1 et Θ .

Les indications de confirmation Con_2 et Con_3 , comme toutes les indications, ont toutes deux une seule de leurs interprétations qui est l'interprétation correcte. En les observant simultanément, nous sommes en présence des quatre cas du figure suivants :

1. $(\omega_{2,c}, \omega_{3,c})$: la table n'est ni truquée, ni cassée,
2. $(\omega_{2,f}, \omega_{3,c})$: la table est truquée, mais pas cassée,
3. $(\omega_{2,c}, \omega_{3,f})$: la table n'est pas truquée, mais elle est cassée,
4. $(\omega_{2,f}, \omega_{3,f})$: la table est truquée et cassée.

Leurs probabilités respectives sont :

1. $\frac{9}{10} \frac{7}{10}$,
2. $\frac{1}{10} \frac{7}{10}$,
3. $\frac{9}{10} \frac{3}{10}$,
4. $\frac{1}{10} \frac{3}{10}$.

Un seul de ces quatre couples correspond à la vérité, c'est-à-dire à un couple d'interprétations dont les deux composantes sont correctes.

A présent, du point de vue de Con_4 , la seule chose qui nous intéresse est de savoir si \mathcal{H}_1 est confirmée ou non ; autrement dit, si la table est équilibrée ou non. Des quatre couples d'interprétations susmentionnés, un seul correspond à une table équilibrée : $(\omega_{2,c}, \omega_{3,c})$. En effet, si le couple d'interprétations correctes est donné par $(\omega_{2,f}, \omega_{3,c})$, $(\omega_{2,c}, \omega_{3,f})$ ou $(\omega_{2,f}, \omega_{3,f})$, alors la table n'est pas équilibrée.

Ainsi, nous constatons qu'en considérant simultanément les indications de confirmation Con_2 et Con_3 , la seule situation où \mathcal{H}_1 serait confirmée est celle dans laquelle les deux interprétations qui la confirment sont toutes deux correctes. Dans tous les autres cas, elle est infirmée.

Nous obtenons pour Con_4 les deux interprétations suivantes :

1. $\omega_{4,c} = (\omega_{2,c}, \omega_{3,c})$: la table est équilibrée.
2. $\omega_{4,f} = (\omega_{2,f}, \omega_{3,c}) \cup (\omega_{2,c}, \omega_{3,f}) \cup (\omega_{2,f}, \omega_{3,f})$: la table n'est pas équilibrée.

Leurs probabilités respectives sont :

1. $p_4(\omega_{4,c}) = \frac{9}{10} \frac{7}{10}$,
2. $p_4(\omega_{4,f}) = 1 - \frac{9}{10} \frac{7}{10}$.

Voici enfin leurs ensembles focaux :

1. $\Delta_4(\omega_{4,c}) = \Omega_1$,
2. $\Delta_4(\omega_{4,f}) = \Theta$.

Nous avons réussi à définir Con_4 ainsi que toutes ses composantes. Il s'agit d'une indication de confirmation qui contient toute l'information de Con_2 et de Con_3 . On pourrait penser qu'une partie de l'information s'est perdue lors de cette combinaison, car on ne sait plus si les tables sont cassées ou truquées, mais uniquement si elles sont équilibrées. Pourtant, n'oublions pas que le rôle d'une indication de confirmation est uniquement de confirmer avec une certaine probabilité une autre indication. Ainsi, Con_2 et Con_3 nous ont été fournies afin de nous informer **dans quelle mesure nous pouvons nous fier à \mathcal{H}_1** ; c'est tout ce qui nous importe. Les raisons sous-jacentes à la confirmation ou non de \mathcal{H}_1 ne nous intéressent pas car la question posée ne porte que sur le nombre qui sortira gagnant. Dans ce sens, Con_4 réunit en elle toute l'information de Con_2 et de Con_3 en rapport avec la question posée.

Essayons à présent de généraliser cette démarche à la combinaison de plusieurs indications de confirmation.

16.2 Combinaison de plusieurs indications de confirmation

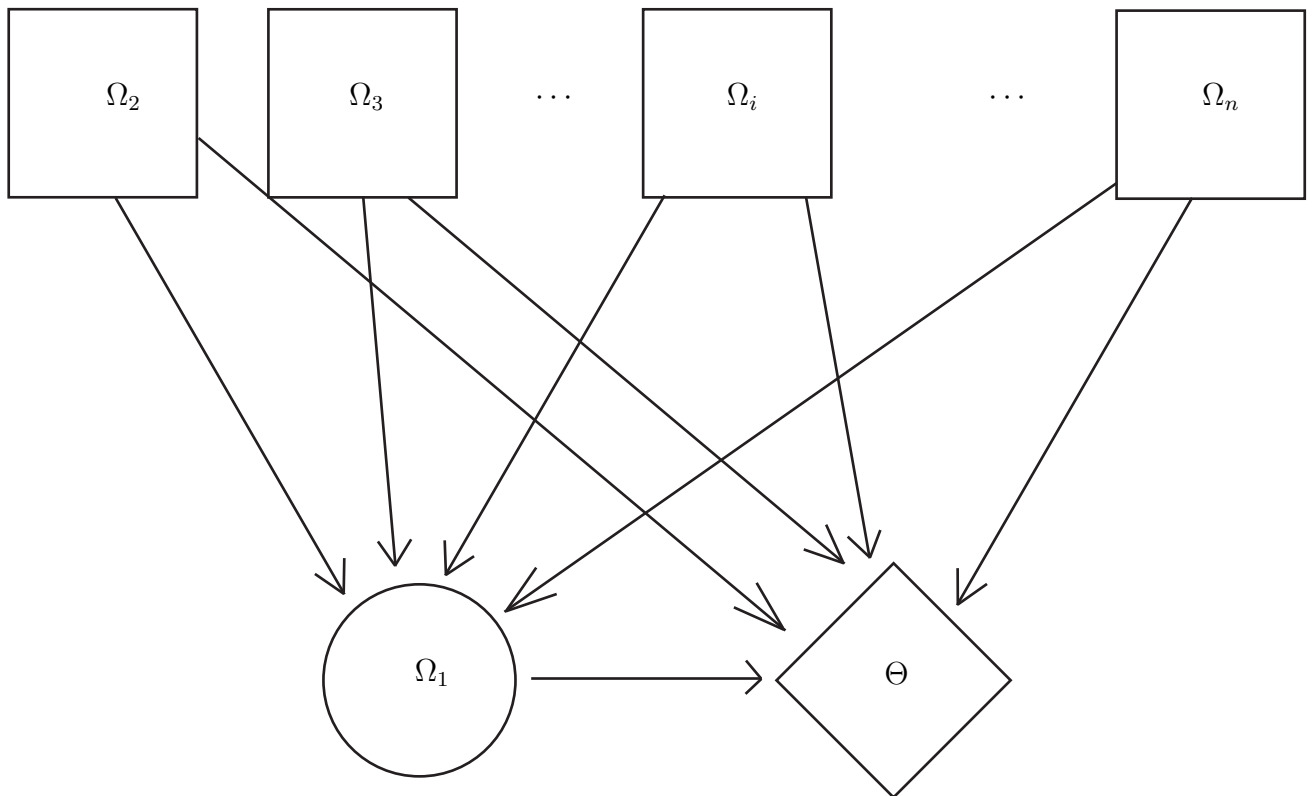
Soient Con_i , avec $i = 2, \dots, n$, plusieurs indications de confirmation, confirmant toutes la même indication $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$. Rappelons que chaque indication Con_i , avec $i = 2, \dots, n$ est définie par $(\Omega_i, P_i, \Delta_i, \Theta)$, où les deux éléments de Ω_i sont $\omega_{i,c}$ et $\omega_{i,f}$, le premier confirmant \mathcal{H}_1 et le second l'infirmant. De même, l'ensemble focal de chaque $\omega_{i,c}$ est donné par Ω_1 et de chaque $\omega_{i,f}$ par Θ respectivement. Comme aucune indication de confirmation Con_i ne confirme une autre indication à part \mathcal{H}_1 , toutes les Con_i se situent «au même niveau», c'est-à-dire qu'il est nécessaire de les combiner entre elles par une adaptation de la règle de Dempster, de manière à obtenir une seule indication de confirmation Con_N .

La Figure 16.1 illustre cette situation.

Soit $Con_N = (\Omega_N, P_N, \Delta_N, \Theta)$ l'indication de confirmation obtenue en combinant toutes les Con_i , avec $i = 2, \dots, n$. Elle doit contenir toute l'information véhiculée par les indications de confirmation qui la composent. Etant elle-même une indication de confirmation, elle n'a que deux interprétations possibles, $\omega_{N,c}$ et $\omega_{N,f}$, la première confirmant \mathcal{H}_1 et la seconde l'infirmant.

En considérant simultanément les $n - 1$ indications de confirmation Con_i , nous sommes confrontés à 2^{n-1} cas de figure différents, représentés par des $(n - 1)$ -tuplets d'interprétations :

$$\begin{aligned} & (\omega_{2,c}, \omega_{3,c}, \dots, \omega_{n,c}), \dots, (\omega_{2,c}, \omega_{3,c}, \dots, \omega_{n,f}), \dots \\ & \dots, (\omega_{2,f}, \omega_{3,f}, \dots, \omega_{n,c}), \dots, (\omega_{2,f}, \omega_{3,f}, \dots, \omega_{n,f}) \end{aligned} \quad (16.1)$$

FIG. 16.1 – Combinaison de $(n - 1)$ Con_i confirmant toutes \mathcal{H}_1

dont un seul est le $(n - 1)$ -tuplet correct, c'est-à-dire composé de $n - 1$ interprétations toutes correctes.

Notons l'ensemble de ces $(n - 1)$ -tuplets d'interprétations $\dot{\Omega}_N$, car c'est lui qui nous permettra de construire Ω_N .

Les $n - 1$ Con_i confirment ou infirment \mathcal{H}_1 pour $n - 1$ raisons différentes. De ce fait, aucun des 2^{n-1} $(n - 1)$ -tuplets différents n'est composé d'interprétations contradictoires. Une indication de confirmation Con_i peut très bien confirmer \mathcal{H}_1 pendant qu'une autre, Con_j , avec $i \neq j$, l'infirme, étant donné qu'elles le font pour des raisons différentes. Pour reprendre l'exemple de la table de roulette, son équilibre peut être mis en doute pour $n - 1$ raisons différentes : la table peut être truquée, cassée, installée en biais, mal nettoyée, etc. Cependant, la table peut être mal nettoyée sans être installée en biais ou cassée. Chacun de ces événements peut se produire indépendamment des autres et sans les contredire.

Revenons à $\omega_{N,c}$ et $\omega_{N,f}$. Quels sont les $(n - 1)$ -tuplets qui confirment \mathcal{H}_1 et par conséquent qui vont composer $\omega_{N,c}$? Le seul cas de figure où \mathcal{H}_1 est confirmée tout en considérant les $n - 1$ Con_i est celui représenté par le $n - 1$ -tuplet $(\omega_{2,c}, \omega_{3,c}, \dots, \omega_{n,c})$. En effet, pour que \mathcal{H}_1 soit confirmée, il est nécessaire que les $n - 1$ interprétations $\omega_{i,c}$ soient correctes simultanément. Nous obtenons ainsi

$$\omega_{N,c} = (\omega_{2,c}, \omega_{3,c}, \dots, \omega_{n,c}). \quad (16.2)$$

Dans tous les autres $(2^{n-1} - 1)$ cas de figure, les $n - 1$ -tuplet sont composés d'au moins une interprétation $\omega_{i,f}$ ce qui suffit pour infirmer \mathcal{H}_1 . Ainsi, $\omega_{N,f}$ est donné par l'union de tous les $(n - 1)$ -tuplets sauf $(\omega_{2,c}, \omega_{3,c}, \dots, \omega_{n,c})$, c'est-à-dire

$$\omega_{N,f} = \dot{\Omega}_N \setminus \{(\omega_{2,c}, \omega_{3,c}, \dots, \omega_{n,c})\}. \quad (16.3)$$

En ce qui concerne la fonction de probabilité de Con_N , elle est simple à définir. La probabilité que $\omega_{N,c}$ soit l'interprétation correcte est donnée par la probabilité que toutes les $\omega_{2,c}, \omega_{3,c}, \dots, \omega_{n,c}$ soient correctes simultanément. Ainsi on obtient :

$$p_N(\omega_{N,c}) = \prod_{i=2}^n p_i(\omega_{i,c}). \quad (16.4)$$

En ce qui concerne la probabilité de $\omega_{N,f}$, elle est tout simplement donnée par

$$p_N(\omega_{N,f}) = 1 - \prod_{i=2}^n p_i(\omega_{i,c}). \quad (16.5)$$

Enfin, comme toute indication de confirmation, ses ensembles focaux sont les suivants :

$$\Delta(\omega_{N,c}) = \Omega_1 \quad (16.6)$$

et

$$\Delta(\omega_{N,f}) = \Theta. \quad (16.7)$$

Ainsi, nous avons défini Con_N ainsi que toutes ses composantes.

Afin de distinguer l'opération de combinaison d'indications de confirmation, que nous venons de présenter, de l'opération de combinaison de Dempster et de celle de combinaison en chaîne, nous allons la noter par le symbole \oplus et l'appeler **règle de combinaison de confirmations**. On choisit précisément ce symbole, très proche de \oplus , car l'opération en question rappelle à maints égards l'opération de Dempster. Cependant, celle-là se distingue de celle-ci, en ce que l'opération de Dempster ne concerne que des indications classiques, tandis l'opération de combinaison de confirmations ne concerne que des indications de confirmation.

Quatrième partie

Nouvelle mesure de performance

La mesure de la performance d'un titre, ou d'un portefeuille de titres, a toujours été au coeur des préoccupations de l'analyse financière. Comment déterminer en effet, quel titre sera le plus performant demain ou à un horizon déterminé ? Voilà une question qui trouve des réponses aussi nombreuses que variées dans la littérature économique. La plupart d'entre elles sont construites sur des fondements combinant la théorie financière aux théories statistiques et économétriques. En effet, la théorie financière est indispensable pour comprendre les mécanismes régissant les places financières, pour modéliser les rapports entre les différents facteurs influençant la performance. Une fois les modèles créés, on fait appel aux théories statistiques et économétriques pour les estimer, les tester, effectuer des prévisions, etc.

Il existe de nos jours plusieurs mesures de performance, ce qui ne facilite pas la prise de décision par l'investisseur. Comment en effet s'y retrouver parmi tous les indices et ratios qui remplissent la littérature financière ? Nous présenterons dans les deux chapitres suivants les six mesures les plus connues et utilisées à ce jour. La question n'en demeure pas moins sans réponse : comment et laquelle choisir parmi les six ? De nombreuses études en finance ont essayé de proposer une solution à ce problème, soit en comparant les théories financières à l'origine des mesures de performance, soit en comparant directement les différentes mesures de performance entre elles. Notre objectif n'est pas de formuler une critique des théories financières, ni de trouver la meilleure d'entre elles. Peu importe qu'une mesure de performance se révèle meilleure qu'une autre, tant que celle qui est jugée de moins bonne qualité continue d'être utilisée par la grande majorité des spécialistes. En effet, ce qui nous intéresse avant tout, c'est la représentation que font, à tort ou à raison, les financiers de la réalité. Les six mesures présentées ultérieurement étant les plus employées, nous ne chercherons pas à comparer leurs qualités et défauts, mais à résumer toute l'information qu'elles véhiculent en **une mesure unique**. Celle-ci permettrait à l'investisseur de connaître l'idée générale que les spécialistes se font du marché, la manière dont ils le perçoivent.

Comment rassembler toute l'information en un seul indice ? Le Chapitre 19 présentera une solution à ce problème sur la base de la théorie mathématique des indications. Mais avant cela, on s'intéressera aux modèles de mesure de la performance permettant un classement des titres ou des portefeuilles. Ce classement est primordial dans le choix des investisseurs, en ce sens qu'il leur fournit un outil de comparaison.

Il existe en effet deux grands courants classiques dans la théorie financière à ce propos : la **Théorie Moderne du Portefeuille** et la **Théorie Post-Moderne du Portefeuille** ; la seconde, comme son nom l'indique, est inspirée de la première, mais elle remet en question quelques uns des fondements qui y sont à la base. Sans entrer dans les détails, puisque l'objectif de ce travail n'est ni de comparer ces deux courants, ni de les critiquer, nous nous concentrerons sur deux aspects de leurs divergences : la conception du risque et la distribution des rendements des actifs financiers.

Chapitre 17

Théorie Moderne du Portefeuille

Ce chapitre exposera brièvement les hypothèses de base de la Théorie Moderne du Portefeuille (TMP), ainsi que trois mesures de performance qui lui sont associées. Pour une présentation détaillée des fondements de la TMP ou de ces mesures de performance, on renvoie le lecteur aux Chapitres 4 à 8 du livre de Elton et Gruber [27], de même qu'au Chapitre 29 du livre de Charest, Lusztiq et Schwab [15].

17.1 Introduction

Selon la Théorie Moderne du Portefeuille, les décisions d'investissement ne sont fondées que sur deux paramètres. Il s'agit des deux premiers moments : l'espérance mathématique et la variance de la distribution des rendements. Ce modèle est connu sous le nom de **Moyenne-Variance** et est attribué à Markowitz [75]. Il découle d'hypothèses sur la fonction d'utilité, supposée quadratique, et sur la distribution des rendements, supposée normale. En effet, on fait l'hypothèse que le rendement d'un actif peut être modélisé par une variable aléatoire, \tilde{r} . Ainsi, dans toute décision d'investissement, on visera simultanément deux objectifs : augmenter le taux de rendement espéré et réduire le risque de son investissement. Le rendement espéré et le risque d'un actif P sont représentés respectivement par l'espérance mathématique et la variance de son rendement :

$$\mu_P = E(\tilde{r}_P) \quad (17.1)$$

et

$$\sigma_P^2 = E(\tilde{r}_P - \mu_P), \quad (17.2)$$

où \tilde{r}_P est le rendement de l'actif P .

Selon Markowitz, les investisseurs doivent ainsi considérer simultanément le rendement espéré et la variance du titre. Le critère Moyenne-Variance permet de générer l'ensemble

des décisions efficaces en faisant varier l'un des deux paramètres : soit l'espérance de rendement, soit le risque de l'actif. L'investisseur peut ainsi étudier l'ensemble des décisions efficaces pour sélectionner la solution satisfaisant son utilité.

Par ailleurs, Markowitz [73] montre qu'un investisseur peut réduire le risque de son portefeuille simplement en détenant des actifs peu ou pas corrélés, donc en diversifiant ses placements. Cela permet d'obtenir la même espérance de rendement en diminuant la variance du portefeuille. Cette manière de sélectionner les actifs d'un portefeuille est appelée **diversification efficace**.

Bien que le modèle de Markowitz présente de nombreuses limites et difficultés, dont nous ne traiterons pas ici, il connaît toujours des applications quotidiennes dans les milieux financiers. Il existe en effet de nombreuses critiques de la TMP. On lui reproche notamment la non prise en compte des frais de transaction et de la fiscalité dans son analyse. Cependant, notre objectif n'est ni de juger cette théorie, ni de souligner ses limites. Nous l'avons choisi pour son universalité et son application très répandue. Malgré ses défauts, la théorie de Markowitz, ainsi que les mesures de performance présentées dans ce chapitre, restent de loin les plus fréquemment utilisées par les professionnels de la finance.

17.2 Relation rendement-risque

L'observation des marchés financiers révèle que lorsque le marché est à la hausse (l'indice du marché augmente), les prix de la plupart des actifs tendent à augmenter également. De même, si le marché est à la baisse, les prix de la plupart des actifs tendent à baisser également. Cela suggère qu'il existe une relation entre les variations des prix des actifs et celles du marché, qui peut être modélisée par :

$$\tilde{r}_P = \alpha_P + \beta_P r_M + \tilde{\varepsilon}_P, \quad (17.3)$$

- \tilde{r}_P est le rendement aléatoire de l'actif P ,
- r_M est le rendement du marché,
- α_P est une composante du rendement de l'actif P indépendante du rendement du marché,
- β_P est une constante qui mesure le changement espéré du rendement de l'actif P , étant donné un certain changement du rendement du marché,
- $\tilde{\varepsilon}_P$ est une variable aléatoire d'espérance mathématique nulle et de variance $\sigma_{\tilde{\varepsilon}_P}^2$.

Rappelons enfin que le rendement discret d'un actif au temps t est défini par :

$$\tilde{r}_t = \frac{\tilde{P}_t - \tilde{P}_{t-1}}{\tilde{P}_{t-1}}, \quad (17.4)$$

où \tilde{P}_t est le prix de l'actif au temps t .

Enfin, selon la TMP, \tilde{r} est supposé être une variable aléatoire distribuée selon une loi normale :

$$\tilde{r} \sim N(\mu, \sigma^2). \quad (17.5)$$

On peut également s'intéresser à la relation entre les rendements excessifs d'un actif P et les rendements excessifs du marché. Par rendements excessifs, on entend les rendements supérieurs aux rendements sûrs, tels que les rendements des obligations de l'Etat, par exemple. Cette relation est exprimée par l'équation suivante :

$$\tilde{r}_P - r_F = \alpha_P + (r_M - r_F)\beta_P + \tilde{\varepsilon}_P, \quad (17.6)$$

où

- r_F est le taux sans risque, c'est-à-dire le taux de rendement sûr, généralement associé aux obligations de l'Etat,
- \tilde{r}_P est le rendement aléatoire de l'actif P ,
- \tilde{r}_M est le rendement du marché,
- α_P est une composante du rendement de l'actif P indépendante du rendement du marché,
- β_P est une constante qui mesure le changement espéré du rendement excessif de l'actif P étant donné un certain changement du rendement excessif du marché,
- $\tilde{\varepsilon}_P$ est une variable aléatoire d'espérance mathématique nulle et de variance $\sigma_{\tilde{\varepsilon}_P}^2$.

17.3 Mesures de performance dans la Théorie Moderne du Portefeuille

Différentes mesures de performance, combinant rendement et risque, ont été inspirées de la théorie de Markowitz. Les plus connues et utilisées sont les suivantes :

- le ratio de Sharpe,
- l'indice de Treynor,
- l'alpha de Jensen.

Le ratio de Sharpe

Le ratio de Sharpe a été développé par William F. Sharpe en 1966 (voir [101]). Celui-ci fut l'un des trois économistes à recevoir le prix Nobel pour «la Théorie Moderne du Portefeuille». Ce ratio mesure le rendement excessif espéré d'un actif P , pondéré par rapport à son risque. Il est donné par :

$$S_P = \frac{E(\tilde{r}_P) - r_F}{\sigma_P}, \quad (17.7)$$

où

- $E(\tilde{r}_P)$ est le rendement espéré de l'actif risqué P ,
- σ_P est l'écart-type du rendement de l'actif P , utilisé comme mesure du risque de l'actif,
- r_F est le taux sans risque.

Evidemment, les investisseurs ne retiennent que les actifs bénéficiant d'un ratio de Sharpe positif. Un ratio négatif ne peut correspondre qu'aux actifs dont le rendement est inférieur au taux sans risque, donc sans aucun intérêt pour les investisseurs. En ce qui concerne les ratios positifs, plus les actifs qui leurs sont associés sont performants, plus ils seront élevés. Cela s'explique aisément, car pour deux actifs de volatilité identique, nous préférons toujours celui qui a le rendement le plus élevé, ou encore, pour deux actifs de rendement identique, nous préférons toujours celui qui a l'écart-type le plus faible.

L'indice de Treynor

L'indice de Treynor a été développé en 1965 [117], par J. Treynor. Il est noté T_P et est défini de la manière suivante :

$$T_P = \frac{E(\tilde{r}_P) - r_F}{\beta_P}, \quad (17.8)$$

où β_P est utilisé comme mesure du risque de l'actif P et β_P , \tilde{r}_P et r_F sont définis dans l'équation (17.6). Cet indice utilise la même mesure du rendement que Sharpe, mais définit le risque différemment, en termes de changement espéré du rendement excessif de l'actif P étant donné un certain changement du rendement excessif du marché. Comme pour le ratio de Sharpe, les investisseurs ne s'intéresseront qu'aux portefeuilles ayant un indice de Treynor positif. Plus le portefeuille est performant, plus cet indice sera élevé.

L'alpha de Jensen

L'alpha de Jensen a été développé en 1968 par Michael J. Jensen, [44], [45]. Cet indice est défini par la différence de performance d'un actif P par rapport au marché :

$$\alpha_P = E(\tilde{r}_P - r_F - \beta_P(r_M - r_F)), \quad (17.9)$$

où α_P , β_P , \tilde{r}_P , r_M et r_F sont définis dans l'équation (17.6).

Si α est positif, alors l'actif réalise une meilleure performance que le marché. Plus la performance de l'actif sera bonne, plus la valeur de l'indice sera élevée. Une valeur négative indique une performance en-dessous du marché et, enfin, une valeur nulle indique la performance que laissait espérer l'efficacité du marché.

Chapitre 18

Théorie Post-Moderne du Portefeuille

18.1 Origine

La Théorie Post-Moderne du Portefeuille (TPMP) vit le jour suite aux critiques de la Théorie Moderne du Portefeuille. Sharpe et Markowitz reconnaissaient déjà les limites de leur modèle initial. Markowitz avait même proposé [75], d'utiliser le risque baissier, qu'il appelait «semi-variance», comme mesure adéquate du risque. Cela dit, en raison de la complexité des calculs et des moyens logistiques insuffisants dont il disposait à son époque, il a dû renoncer à cette idée, impossible à appliquer, et se contenter de la variance.

Les deux reproches de base formulés contre la TMP sont :

1. la représentation du risque, par la variance ou l'écart-type,
2. la représentation des rendements par la distribution normale.

Les progrès tant au niveau de la théorie financière que des moyens informatiques ont permis de surmonter ces limites et de poser les bases de la TPMP. Celle-ci ne réfute pas le principe du rapport entre le rendement et le risque d'un actif; elle ne fait que le présenter différemment. D'après la TPMP, la variance ou l'écart-type ne représente pas correctement le risque d'un actif. Etant une mesure symétrique, elle pénalise de manière identique les gains et les pertes. En effet, Brachinger et Weber [13] expliquent que le risque représente une caractéristique négative et relative à une certaine décision. Intuitivement, on le perçoit comme la chance de réalisation d'un préjudice, d'une nuisance ou d'une perte, relatifs à une action donnée. Il est évident que tout financier cherchera à se préserver de grandes pertes potentielles, mais s'agissant des gains, une grande volatilité semble même attractive! La TPMP résout ce problème en introduisant le concept de **risque baissier**, représenté par la volatilité à la baisse, car seule la volatilité à la baisse, en-dessous d'un certain seuil que l'investisseur s'est fixé, représente un véritable risque pour celui-ci. Tous les rendements potentiels, en-dessus de ce seuil, aussi volatils qu'ils soient, sont des rendements sans risque. Le seuil, propre à chaque investisseur, est appelé

rendement minimal acceptable (RMA). Bawa [8] et Fishburn [30] ont proposé d'utiliser les moments partiels inférieurs (MPI) comme mesure appropriée du risque baissier. Par la suite, de nouveaux modèles Moyenne-Moment Partiel Inférieur (MMPI) ont vu le jour, avec l'objectif de remplacer le modèle Moyenne-Variance de Markowitz. On les doit principalement à Bawa et Lindberg [10], qui ont développé le «Lower Partial Moment Capital Asset Pricing Model» (LMP-CAPM), à Bawa, Brown et Klein [9], dont le modèle s'appelle Asymmetric Response Model (ARM), à Lee et Rao [67] et à Harlow et Rao [38].

S'agissant du second problème évoqué, la représentation de la distribution des rendements, la TPMP se base sur l'hypothèse que les logrendements suivent une distribution normale, ce qui implique que les rendements suivent une distribution lognormale.

Le **logrendement** d'un actif au temps t , est défini par

$$\tilde{R}_t^* = \ln \left(\frac{\tilde{P}_t}{\tilde{P}_{t-1}} \right), \quad (18.1)$$

où \tilde{P}_t est le prix de l'actif au temps t .

Il est notoire qu'une variable aléatoire \tilde{x} qui n'admet que des valeurs positives suit exactement une distribution lognormale de paramètres μ_L et σ_L^2 , si la variable aléatoire transformée $\tilde{y} := \ln(\tilde{x})$ suit une distribution normale de paramètres μ_L et σ_L^2 .

On voit ainsi apparaître une nouvelle définition des rendements, différente de celle proposée dans l'équation (17.4). Selon la TPMP, un logrendement \tilde{R}^* suit une loi normale de paramètres μ_L et σ_L^2

$$\tilde{R}^* \sim N(\mu_L, \sigma_L^2). \quad (18.2)$$

Il s'en suit que le rendement, défini par

$$\tilde{R}_t = \left(\frac{\tilde{P}_t}{\tilde{P}_{t-1}} \right), \quad (18.3)$$

suit une distribution lognormale de paramètres μ_L et σ_L^2 .

Le choix de cette nouvelle définition des rendements a déjà été justifié par de nombreuses études ; parmi elles, pour ne citer que les plus connues, celles de Fama [29] et de Singleton et Wingender [105].

18.2 Moments partiels inférieurs comme mesure du risque

Bawa [8] et Bawa et Lindberg [10] sont les premiers auteurs à modéliser le risque baissier par les moments partiels inférieurs MPI d'ordre n avec un RMA de valeur t défini par :

$$MPI_P(n, t) = \int_{-\infty}^t (t - R_P^*)^n dF(R_P^*), \quad (18.4)$$

où

- R_P^* est le logrendement de l'actif P ,
- F est la distribution de probabilité de ce rendement,
- t est le RMA,
- n est le degré du moment.

L'équation (18.4) peut également être exprimée par :

$$MPI_P(n, t) = \left[E(t - \tilde{R}_P^*)^n | t > R_P^* \right]. \quad (18.5)$$

De plus, Stone [115], [116] (suivi par Bawa [8] et Fishburn [30]), a montré que la semi-variance et la variance constituent des cas particuliers des moments partiels inférieurs. Ainsi, pour $n = 2$ et $t = R_F$, on retrouve la semi-variance traditionnelle de Markowitz. Le semi-écart-type (SET) est alors défini par

$$SET = MPI_P(2, R_F) = E \left[(R_F - \tilde{R}_P^*)^2 | R_F > R_P^* \right]^{\frac{1}{2}}. \quad (18.6)$$

Il s'agit bien sûr du second moment partiel inférieur dont le RMA est égal au taux de rendement sûr R_F .

18.3 Mesures de performance dans la Théorie Post-Moderne du Portefeuille

Dans le modèle MMPI, les mesures de performance de Sharpe, Treynor et Jensen ont été remplacées par leur équivalent basé sur le risque défini en termes de SET et non plus d'écart-type. C'est ainsi que nous disposons de trois mesures nouvelles :

$$S_P^{SET} = \frac{E(\tilde{R}_P^*) - R_F}{SET}, \quad (18.7)$$

appelé ratio de Sortino,

$$T_P^{SET} = \frac{E(\tilde{R}_P^*) - R_F}{\beta_P^{SET}}, \quad (18.8)$$

appelé indice de Treynor (SET), et enfin

$$\alpha_P^{SET} = E(\tilde{R}_P^* - R_F - \beta_P^{SET}(R_M^* - R_F)), \quad (18.9)$$

appelé alpha de Jensen (SET), où

- \tilde{R}_P^* est le logrendement de l'actif P ,
- \tilde{R}_M^* est le logrendement du marché,
- R_F est le taux sans risque,

et où

$$\beta_P^{SET} = \frac{E \left[(\tilde{R}_P^* - R_F) \min(0, \tilde{R}_M^* - R_F) \right]}{E \left[\min(0, \tilde{R}_M^* - R_F) \right]^2}. \quad (18.10)$$

On constate que β_P^{SET} est défini en termes de semi-covariance avec le marché. Ainsi, seuls les comouvements entre les rendements du marché et les rendements de l'actif P , lorsque la performance du marché descend en-dessous du niveau sans risque R_F , génèrent le risque systématique.

Le ratio de Sortino a été développé par Sortino et Price [109], tandis que des variantes de l'indice de Treynor (SET) et de l'alpha de Jensen (SET) ont été utilisées par Henriksson et Merton [42], puis par Henriksson [41].

Chapitre 19

Nouvelle approche de la mesure de performance

Notre objectif premier est de rassembler toute l'information en une seule et unique mesure de performance des actifs, aussi bien celle basée sur la TMP que celle basée sur la TPMP.

19.1 Modèle de performance et incertitude

La question posée est la suivante : «*parmi n actifs considérés, quel est le classement de l'actif j , avec $i = 1, \dots, n$?*». En d'autres termes, de même que le permettent les six mesures de performance présentées dans les deux chapitres précédents, on a construit un classement de n actifs, du plus performant au moins performant. Le modèle à la base de ce classement a été appelé **Modèle de performance et incertitude** et est présenté dans ce chapitre.

Soient deux indications classiques relatives à cette question, donc relatives à un actif j :

- $\mathcal{H}_1 = (\Omega_1, P_1, \Gamma_1, \Theta)$,
- $\mathcal{H}_2 = (\Omega_2, P_2, \Gamma_2, \Theta)$.

L'indication \mathcal{H}_1 nous dit que nous devons modéliser la performance de l'actif j avec la Théorie Moderne du Portefeuille, tandis que l'indication \mathcal{H}_2 nous dit que nous devons modéliser la performance de l'actif j avec la Théorie Post-Moderne du Portefeuille. On peut imaginer que ces indications nous sont données par des experts en finance, chacun se fiant à une des deux théories en question.

L'ensemble des réponses Θ , qui est commun aux deux indications \mathcal{H}_1 et \mathcal{H}_2 , est donné par tous les rangs possibles de l'actif j dans un classement de n actifs :

$$\Theta = \{1, 2, \dots, n\}.$$

Supposons qu'il existe trois manières d'interpréter \mathcal{H}_1 ou, autrement dit, la performance dans le cadre de la TMP. En effet, on estime que la performance peut être révélée soit par le ratio de Sharpe, soit par l'indice de Treynor, soit enfin par l'alpha de Jensen. Chacune de ces trois manières de concevoir la performance peut mener avec une certaine probabilité à la réponse correcte. Ainsi Ω_1 est composé des trois éléments suivants :

- $\omega_{1,1}$: la performance d'un actif est donnée par le ratio de Sharpe,
- $\omega_{1,2}$: la performance d'un actif est donnée par l'indice de Treynor,
- $\omega_{1,3}$: la performance d'un actif est donnée par l'alpha de Jensen.

Supposons qu'il en soit de même pour \mathcal{H}_2 , dont Ω_2 est composé de :

- $\omega_{2,1}$: la performance d'un actif est donnée par le ratio de Sortino,
- $\omega_{2,2}$: la performance d'un actif est donnée par l'indice de Treynor (SET),
- $\omega_{2,3}$: la performance d'un actif est donnée par l'alpha de Jensen (SET).

Rappelons que pour une indication, une seule de ses interprétations est l'interprétation correcte. Autrement dit, une seule des trois manières d'interpréter la TMP et la TPMP nous fournira une représentation correcte de la performance d'un actif. Jusque là, la littérature financière n'a pas réussi à départager ces trois façons d'interpréter la même information. Il existe certes des études comparatives, mais aucune de ces mesures de performance n'a été abandonnée depuis. Ainsi, chaque expert a ses propres préférences et choisi la mesure qui lui convient le mieux.

Les fonctions de probabilités de \mathcal{H}_1 et \mathcal{H}_2 , p_1 et p_2 nous sont fournies par les experts. Ainsi, chaque expert défenseur de la TMP accordera une certaine probabilité aux trois mesures, donc aux trois interprétations de \mathcal{H}_1 d'être l'interprétation correcte et de le mener ainsi à la réponse correcte. Il en est de même pour les défenseurs de la TPMP et pour les trois mesures qui y sont associées. Ces probabilités peuvent évidemment varier d'une personne à l'autre, tout dépend de leurs connaissances et expériences personnelles.

Les applications Γ_1 et Γ_2 sont données par les classements obtenus par les six mesures de performance considérées. Par exemple, si le ratio de Sharpe met un certain actif en 3^{ème} position parmi n actifs considérés, alors l'ensemble focal de $\omega_{1,1}$ sera donné par le sous-ensemble $\{3\}$ de Θ .

Nous pouvons constater que \mathcal{H}_1 et \mathcal{H}_2 sont des indications précises (tous les ensembles focaux sont des singletons) et bayésiennes (tous les ensembles focaux sont disjoints). De plus, si les trois mesures de performance (de \mathcal{H}_1 ou de \mathcal{H}_2) classent un certain actif à une même position, alors l'indication en question associée à cet actif sera sûre.

A chaque actif considéré sont associées une indication \mathcal{H}_1 et une indication \mathcal{H}_2 qui lui sont propres. En effet, parmi les n actifs considérés, les \mathcal{H}_1 et \mathcal{H}_2 qui leur sont associées seront identiques en ce qui concerne Θ , les fonctions de probabilité p_1 , les fonctions de probabilité p_2 , les ensembles d'interprétations Ω_1 et les ensembles d'interprétations Ω_2 . Par contre, les applications Γ_1 et Γ_2 seront différentes pour chaque actif.

Rappelons enfin que la TMP se base sur l'hypothèse que les rendements suivent une distribution normale, c'est-à-dire

$$\tilde{r}_t \sim N(\mu, \sigma^2). \quad (19.1)$$

De même, la TPMP suppose que les rendements suivent une distribution lognormale, c'est-à-dire

$$\tilde{R}_t \sim LN(\mu_L, \sigma_L^2). \quad (19.2)$$

Que se passerait-il si pour les rendements d'un actif considéré, une de ces hypothèses, ou les deux hypothèses, ne sont pas respectées? Si tel est le cas, toutes les mesures de performance basées sur l'hypothèse qui n'est pas respectée perdent leur sens. Par exemple, il sera toujours possible de calculer le ratio de Sharpe pour un actif, mais si les rendements de cet actif ne suivent pas une distribution normale, on ne peut pas interpréter la valeur de ce ratio. On ne pourra pas non plus interpréter les valeurs du l'indice de Treynor ou de l'alpha de Jensen. Ceux-ci n'ont de sens que si l'hypothèse de normalité des rendements est respectée. Ainsi, toute la modélisation par la TMP se voit remise en question. Il ne suffit donc pas de calculer les six mesures de performance de la TMP et de la TPMP, il est également nécessaire de s'interroger si les hypothèses sur la distribution des rendements sont respectées.

On peut aisément modéliser cette situation à l'aide des indications de confirmation. En effet, nous sommes en présence d'informations (\mathcal{H}_1 et \mathcal{H}_2) qui sont remises en question : confirmées ou infirmées si les hypothèses de distribution des rendements sont respectées ou violées respectivement.

Soient deux indications de confirmation :

- $Con_3 = (\Omega_3, P_3, \Delta_3, \Theta)$,
- $Con_4 = (\Omega_4, P_4, \Delta_4, \Theta)$.

L'indication Con_3 est une indication de confirmation sur \mathcal{H}_1 . Elle confirme que la distribution des rendements est normale. De même, Con_4 est une indication de confirmation sur \mathcal{H}_2 . Elle confirme, quant à elle, que la distribution des rendements est lognormale.

Les éléments de Ω_3 , comme pour toute indication de confirmation, sont au nombre de deux :

- $\omega_{3,c}$: la distribution des rendements est normale,
- $\omega_{3,f}$: la distribution des rendements n'est pas normale.

Les éléments de Ω_4 sont bien sûr

- $\omega_{4,c}$: la distribution des rendements est lognormale,
- $\omega_{4,f}$: la distribution des rendements n'est pas lognormale.

Les probabilités $p_3(\omega_{3,c})$, $p_3(\omega_{3,f})$, $p_4(\omega_{4,c})$ et $p_4(\omega_{4,f})$ de Con_3 et Con_4 peuvent être estimées sur la base d'un échantillon d'observation des rendements de l'actif qui nous intéresse.

Pour cela, on peut partager l'échantillon en plusieurs sous-échantillons donnés par des séquences d'observations, comme il est d'usage de le faire en finance. Par exemple, si nous disposons de 100 observations et si nous désirons considérer les sous-échantillons formés de 60 observations, nous obtiendrons les 41 sous-échantillons, dont le premier sera composé de la 1^{ère} à la 60^{ème} observation, le second sera composé de la 2^{ème} à la 61^{ème} observation, etc. jusqu'au dernier, le 41^{ème}, qui sera composé de la 41^{ème} à la 100^{ème} observation. Cette manière de procéder n'est certainement pas la meilleure. En effet, les sous-échantillons ne seront pas indépendants étant donné que d'un sous-échantillon au sous-échantillon suivant, une seule observation variera. Cependant, en finance, on dispose rarement d'un échantillon d'observations assez grand pour pouvoir le partager en un nombre suffisant de sous-échantillons indépendants. Puis, à l'aide du test de Kolmogorov-Smirnov, nous pouvons, pour chaque sous-échantillon, tester l'hypothèse que ses rendements suivent une distribution normale de paramètres μ est σ^2 , ou qu'ils suivent une distribution lognormale de paramètres μ_L est σ_L^2 . L'hypothèse H_0 nous dit que la distribution des rendements est normale de paramètres μ est σ^2 (ou lognormale de paramètres μ_L est σ_L^2) et H_1 que la distribution des rendements n'est pas normale de paramètres μ est σ^2 (ou n'est pas lognormale de paramètres μ_L est σ_L^2).

Les paramètres sont estimés sur la base de l'échantillon entier, car pour chaque sous-échantillon, ce qui nous intéresse n'est pas uniquement de tester si ses rendements suivent une distribution normale ou lognormale, mais s'ils suivent une distribution normale ou lognormale **précise**. En effet, aussi bien dans la TMP que dans la TPMP, lorsqu'on suppose la distribution des rendements d'un actif, on suppose une distribution précise et non une famille de distributions. Les paramètres de cette distribution précise sont estimés sur la base de tout l'échantillon, puis sont utilisés dans le calcul des différentes mesures de performance. En ce sens, pour un actif considéré, ce qui nous intéresse n'est pas de savoir si les rendements d'un sous-échantillon suivent une distribution normale ou lognormale quelconque, mais s'ils suivent la distribution précise. En effet, ce sont les paramètres de cette distribution précise qui définissent les mesures de performance de l'actif considéré.

Une fois que le test a été effectué sur tous les échantillons, on calcule la proportion des sous-échantillons pour lesquels, avec un certain seuil de signification, par exemple de 5%, on accepte \mathcal{H}_0 . Cette proportion peut être utilisée pour estimer la probabilité $p_3(\omega_{3,c})$ ou $p_4(\omega_{4,c})$ que les rendements suivent une distribution normale précise, ou lognormale précise, selon le test.

Il existe certainement de meilleures méthodes pour estimer les fonctions de probabilité de Con_3 et Con_4 , mais on n'analysera pas en détail cette question qui relève de la théorie des tests et sort des objectifs de ce travail. On pourrait cependant estimer de plusieurs manières les probabilités que les rendements sont distribués normalement ou lognormalement, où chaque méthode aboutit à une estimation différente de ces probabilités. Puis, on pourrait considérer que ces estimations correspondent aux fonctions de probabilité

de plusieurs indications de confirmation $Con_{3,1}, Con_{3,2}, Con_{3,3}, \dots$ de \mathcal{H}_1 et de plusieurs indications de confirmation $Con_{4,1}, Con_{4,2}, Con_{4,3}, \dots$ de \mathcal{H}_4 . Enfin, on pourrait combiner $Con_{3,1}, Con_{3,2}, Con_{3,3}, \dots$ en une seule indication de confirmation Con_3 et combiner $Con_{4,1}, Con_{4,2}, Con_{4,3}, \dots$ en une seule indication de confirmation Con_4 .

Revenons au modèle présenté dans ce chapitre. Les ensembles focaux de Con_3 et de Con_4 sont bien évidemment donnés par

- $\Delta_3(\omega_{3,c}) = \Omega_1,$
- $\Delta_3(\omega_{3,f}) = \Theta,$
- $\Delta_4(\omega_{4,c}) = \Omega_2,$
- $\Delta_4(\omega_{4,f}) = \Theta.$

De même que pour \mathcal{H}_1 et \mathcal{H}_2 , il existe une indication Con_3 et une indication Con_4 spécifique pour chaque actif. Leurs ensembles focaux, leurs ensembles d'interprétations et leur ensemble des réponses sont identiques pour chaque actif, tandis que les fonctions de probabilité p_3 et p_4 ne le sont pas, comme elles sont estimées sur la base des rendements.

Nous obtenons ainsi les deux meta-indications suivantes pour chaque actif :

- $\mathcal{M}'_{H_1} = Con_3 \underset{\omega_{3,c}}{\rightsquigarrow} \mathcal{H}_1,$
- $\mathcal{M}''_{H_2} = Con_4 \underset{\omega_{3,c}}{\rightsquigarrow} \mathcal{H}_2.$

La meta-indication $\mathcal{M}'_{\mathcal{H}_1}$ est définie par le quadruplet $(\Omega', P', \Gamma', \Theta)$, où

- $\Omega' = (\omega_{1,1}, \omega_{1,2}, \omega_{1,3}, \omega_{3,f})$
- $P' :$

$$p'(\omega_{1,1}) = p_3(\omega_{3,c}) \cdot p_1(\omega_{1,1})$$

$$p'(\omega_{1,2}) = p_3(\omega_{3,c}) \cdot p_1(\omega_{1,2})$$

$$p'(\omega_{1,3}) = p_3(\omega_{3,c}) \cdot p_1(\omega_{1,3})$$

$$p'(\omega_{3,f}) = p_3(\omega_{3,f})$$
- $\Gamma' :$

$$\Gamma'(\omega_{1,1}) = \Gamma_1(\omega_{1,1})$$

$$\Gamma'(\omega_{1,2}) = \Gamma_1(\omega_{1,2})$$

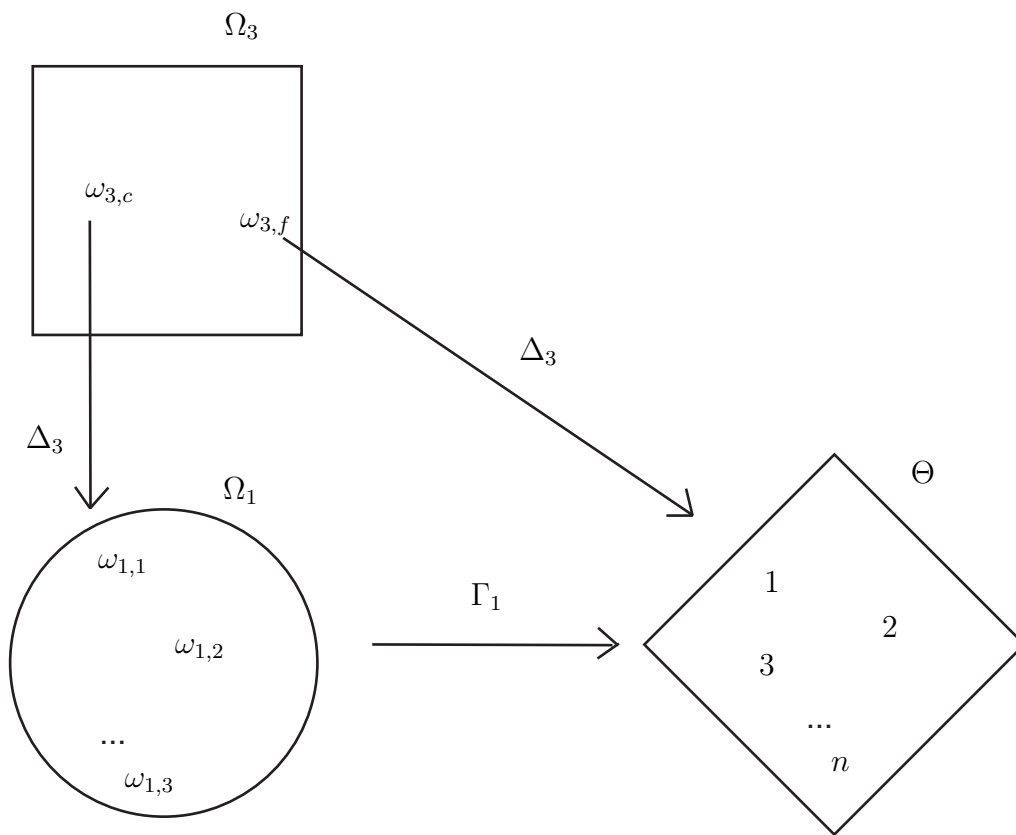
$$\Gamma'(\omega_{1,3}) = \Gamma_1(\omega_{1,3})$$

$$\Gamma'(\omega_{3,f}) = \Delta_3(\omega_{3,f}) = \Theta$$
- $\Theta = \{1, 2, \dots, n\}$

Son schéma graphique est présenté à la Figure 19.1.

De même, $\mathcal{M}''_{\mathcal{H}_2}$ est définie par le quadruplet $(\Omega'', P'', \Gamma'', \Theta)$, où

- $\Omega'' = (\omega_{2,1}, \omega_{2,2}, \omega_{2,3}, \omega_{4,f})$
- $P'' :$

FIG. 19.1 – Schéma graphique de la meta-indication \mathcal{M}'_{H_1}

$$p''(\omega_{2,1}) = p_4(\omega_{4,c}) \cdot p_2(\omega_{2,1})$$

$$p''(\omega_{2,2}) = p_4(\omega_{4,c}) \cdot p_2(\omega_{2,2})$$

$$p''(\omega_{2,3}) = p_4(\omega_{4,c}) \cdot p_2(\omega_{2,3})$$

$$p''(\omega_{4,f}) = p_4(\omega_{4,f})$$

- Γ'' :

$$\Gamma''(\omega_{2,1}) = \Gamma_2(\omega_{2,1})$$

$$\Gamma''(\omega_{2,2}) = \Gamma_2(\omega_{2,2})$$

$$\Gamma''(\omega_{2,3}) = \Gamma_2(\omega_{2,3})$$

$$\Gamma''(\omega_{4,f}) = \Delta_4(\omega_{4,f}) = \Theta$$

- $\Theta = \{1, 2, \dots, n\}$

Son schéma graphique est présenté à la Figure 19.2.

Il ne reste plus qu'à combiner $\mathcal{M}'_{\mathcal{H}_1}$ et \mathcal{M}''_{H_2} par la règle de combinaison de Dempster :

$$\mathcal{M}'_{H_1} \oplus \mathcal{M}''_{H_2} = \mathcal{M}.$$

La meta-indication \mathcal{M} permet de considérer toute l'information véhiculée par $\mathcal{M}'_{\mathcal{H}_1}$ et \mathcal{M}''_{H_2} simultanément, tout en tenant compte des informations contradictoires. Grâce à cette combinaison, on aura une fonction de support et une fonction de plausibilité uniques, relatives à \mathcal{M} . Ainsi, pour un actif j considéré, on connaîtra les degrés de support et de plausibilité de chaque rang qu'il pourrait avoir dans un classement de n actifs.

Combiner par la règle de Dempster $\mathcal{M}'_{\mathcal{H}_1}$ et \mathcal{M}''_{H_2} ne pose aucun problème, car, rappelons-le, toute meta-indication est également une indication classique.

En combinant ces deux meta-indications, nous obtenons \mathcal{M} , définie par le quadruplet $(\Omega, P, \Gamma, \Theta)$, où

-

$$\Omega = \{(\omega_{1,1}, \omega_{2,1}), (\omega_{1,1}, \omega_{2,2}), (\omega_{1,1}, \omega_{2,3}), (\omega_{1,1}, \omega_{4,f}),$$

$$(\omega_{1,2}, \omega_{2,1}), (\omega_{1,2}, \omega_{2,2}), (\omega_{1,2}, \omega_{2,3}), (\omega_{1,2}, \omega_{4,f}),$$

$$(\omega_{1,3}, \omega_{2,1}), (\omega_{1,3}, \omega_{2,2}), (\omega_{1,3}, \omega_{2,3}), (\omega_{1,3}, \omega_{4,f}),$$

$$(\omega_{3,f}, \omega_{2,1}), (\omega_{3,f}, \omega_{2,2}), (\omega_{3,f}, \omega_{2,3}), (\omega_{3,f}, \omega_{4,f})\}$$

- P^* (fonction de probabilité avant ajustement) :

$$p^*(\omega_{1,1}, \omega_{2,i}) = p'(\omega_{1,1}) \cdot p''(\omega_{2,i}), \text{ avec } i = 1, 2, 3,,$$

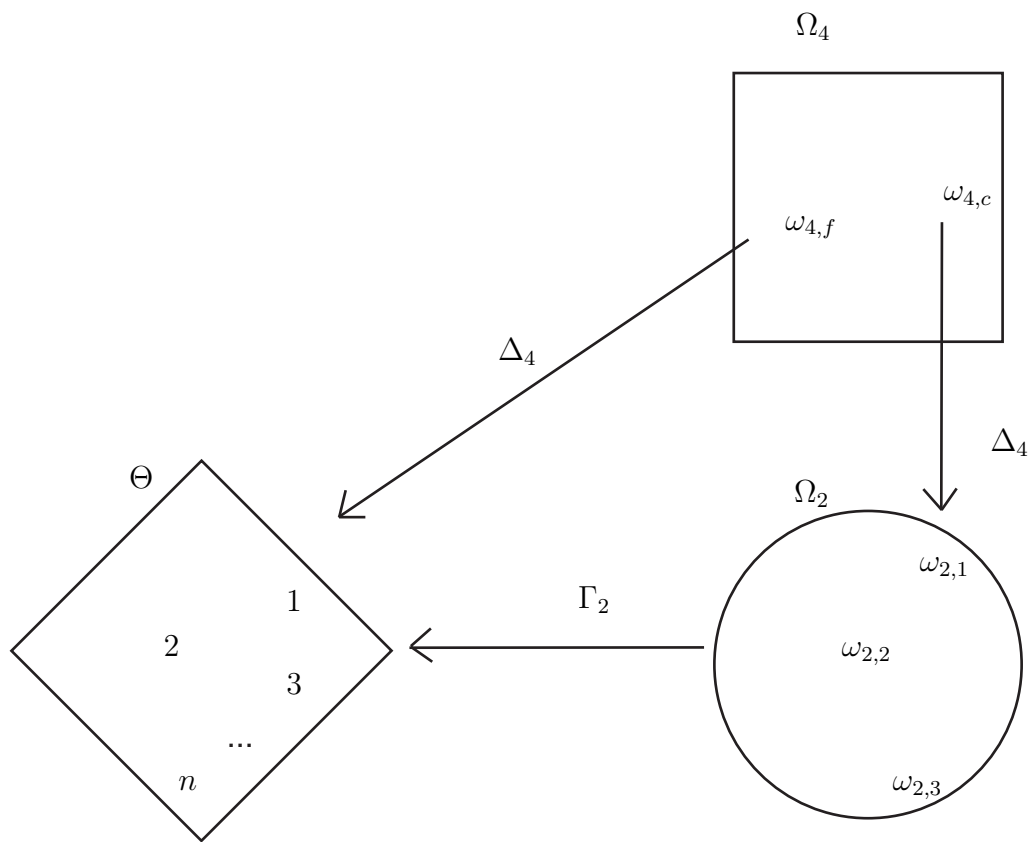
$$p^*(\omega_{1,1}, \omega_{4,f}) = p'(\omega_{1,1}) \cdot p''(\omega_{4,f})$$

$$p^*(\omega_{1,2}, \omega_{2,i}) = p'(\omega_{1,2}) \cdot p''(\omega_{2,i}) \text{ avec } i = 1, 2, 3,,$$

$$p^*(\omega_{1,2}, \omega_{4,f}) = p'(\omega_{1,2}) \cdot p''(\omega_{4,f})$$

$$p^*(\omega_{1,3}, \omega_{2,i}) = p'(\omega_{1,3}) \cdot p''(\omega_{2,i}) \text{ avec } i = 1, 2, 3,,$$

$$p^*(\omega_{1,3}, \omega_{4,f}) = p'(\omega_{1,3}) \cdot p''(\omega_{4,f})$$

FIG. 19.2 – Schéma graphique de la meta-indication \mathcal{M}''_{H_2}

$$p^*(\omega_{3,f}, \omega_{2,i}) = p'(\omega_{3,f}) \cdot p''(\omega_{2,i}) \text{ avec } i = 1, 2, 3,,$$

$$p^*(\omega_{3,f}, \omega_{4,f}) = p'(\omega_{3,f}) \cdot p''(\omega_{4,f})$$

- Γ :

$$\Gamma(\omega_{1,1}, \omega_{2,i}) = \Gamma'(\omega_{1,1}) \cap \Gamma''(\omega_{2,i}) \text{ avec } i = 1, 2, 3,,$$

$$\Gamma(\omega_{1,1}, \omega_{4,f}) = \Gamma'(\omega_{1,1}) \cap \Gamma''(\omega_{4,f})$$

$$\Gamma(\omega_{1,2}, \omega_{2,i}) = \Gamma'(\omega_{1,2}) \cap \Gamma''(\omega_{2,i}) \text{ avec } i = 1, 2, 3,,$$

$$\Gamma(\omega_{1,2}, \omega_{4,f}) = \Gamma'(\omega_{1,2}) \cap \Gamma''(\omega_{4,f})$$

$$\Gamma(\omega_{1,3}, \omega_{2,i}) = \Gamma'(\omega_{1,3}) \cap \Gamma''(\omega_{2,i}) \text{ avec } i = 1, 2, 3,,$$

$$\Gamma(\omega_{1,3}, \omega_{4,f}) = \Gamma'(\omega_{1,3}) \cap \Gamma''(\omega_{4,f})$$

$$\Gamma(\omega_{3,f}, \omega_{2,i}) = \Gamma'(\omega_{3,f}) \cap \Gamma''(\omega_{2,i}) \text{ avec } i = 1, 2, 3,,$$

$$\Gamma(\omega_{3,f}, \omega_{4,f}) = \Gamma'(\omega_{3,f}) \cap \Gamma''(\omega_{4,f})$$

Evidemment, si l'intersection des ensembles focaux de $\mathcal{M}'_{\mathcal{H}_1}$ et $\mathcal{M}''_{\mathcal{H}_2}$ aboutissent à l'ensemble vide, il sera nécessaire d'ajuster la fonction de probabilité p^* , pour obtenir la fonction de probabilité ajustée p .

- $\Theta = \{1, 2, \dots, n\}$

19.2 Nouvelle source de risque

Une fois \mathcal{M} définie pour un actif j considéré, nous obtenons une fonction de support et de plausibilité pour tous les sous-ensembles de Θ . Cela dit, ceux dont les degrés de support et de plausibilité nous intéressent particulièrement sont les éléments de Θ , c'est-à-dire $\{1\}, \{2\}, \dots, \{n\}$. Ces éléments représentent la place que l'actif considéré occupe dans le classement, parmi les n actifs classés. A chaque place se voit ainsi associé un degré de support et de plausibilité.

Contrairement à \mathcal{H}_1 et à \mathcal{H}_2 , la meta-indication \mathcal{M} n'est pas une indication simple, car ses ensembles focaux ne sont pas tous des singletons et, par conséquent, ses degrés de support ne sont pas identiques à ses degrés de plausibilité. En fait, son seul ensemble focal qui ne correspond pas à un singleton est $\Gamma(\omega_{3,f}, \omega_{4,f}) = \Gamma'(\omega_{3,f}) \cap \Gamma''(\omega_{4,f}) = \Theta$. Il représente la situation où les rendements de l'actif considéré ne suivent ni une distribution normale, ni une distribution lognormale. Cela implique que nous ne pouvons pas utiliser les six mesures classiques de performance, leur hypothèse de base n'étant pas remplie. Ainsi, le plus petit sous-ensemble de l'ensemble des réponses pour lequel nous pouvons prouver qu'il contient la réponse correcte n'est autre que Θ , l'ensemble des réponses lui-même. Dans ce cas, ne disposant d'aucune information, on ne peut se prononcer quant au classement de cet actif. La masse de Θ est égale à la probabilité $p(\omega_{3,f}, \omega_{4,f})$. Il s'agit, rappelons-le, de la partie du degré de support total accordé à Θ (qui est de 1 par définition), qui ne

supporte aucun autre sous-ensemble de Θ , mais Θ lui-même de manière exclusive. Cette masse représente l'incertitude liée au classement de l'actif considéré.

Jusque là, aussi bien dans la TMP que dans la TPMP, le risque était défini en termes de variabilité. Dans la TMP, on considère la variabilité des rendements de l'actif (ratio de Sharpe), ou la variabilité des rendements de l'actif par rapport à celle des rendements du marché (indice de Treynor et alpha de Jensen). Quant à la TPMP, elle représente le risque par une variabilité à la baisse, à nouveau soit des rendements de l'actif (ratio de Sortino) ou des rendements de l'actif par rapport à ceux des du marché (indice de Treynor (SET) et alpha de Jensen (SET)).

On propose de joindre à ces concepts une nouvelle mesure de risque, non pas substitutive, mais plutôt complémentaire, qui le définirait en termes d'incertitude, telle qu'elle est perçue dans la théorie mathématique des indications. En effet, le manque total ou partiel d'information sur la distribution des rendements implique l'impossibilité de se fier entièrement aux mesures de performances de la TMP et de la TPMP et, par conséquent, de classer l'actif considéré. Que ses rendements varient peu ou beaucoup, entre eux ou par rapport à ceux du marché, à la baisse ou à la hausse, n'y change rien. Pour que les mesures de performances de la TMP et de la TPMP aient un sens, il est nécessaire que leur hypothèse sur la distribution des rendements soit remplie. Or, ce n'est la plupart du temps pas le cas en pratique, ou alors de manière partielle. Par conséquent, une certaine incertitude se rattache à l'actif, différente du concept de variabilité. Cette incertitude, nous la considérons comme une source supplémentaire de risque, étant donné qu'elle conduit à une absence totale ou partielle de l'exactitude des classements obtenus par les six mesures de performance que nous traitons ici. La théorie classique des probabilités ignorant ce concept, nous la modélisons avec la théorie mathématique des indications qui, une fois de plus, révèle ainsi tout son intérêt.

Ainsi, cette source supplémentaire de risque est modélisée par la mesure définie comme la masse attribuée à Θ qui, bien sûr, est comprise entre 0 et 1.

Si cette masse est nulle, alors la totalité de la masse est répartie entre les singletons et, de ce fait, il n'y a aucune incertitude. Le risque qui y est lié est nul. Tel sera le cas si la probabilité que les rendements de l'actif suivent une loi normale **ou** si la probabilité que les rendements de l'actif suivent une loi lognormale valent 1. Il n'est pas nécessaire que ces deux probabilités valent 1, une seule suffit, car la masse de Θ est donnée par le produit de la probabilité que les rendements de l'actif ne suivent pas une loi normale et de celle que les rendements de l'actif ne suivent pas une loi lognormale. Il suffit que l'une des deux soit égale à 0.

A l'inverse, si cette masse vaut 1, alors la totalité de la masse est accordée à Θ , ce qui implique que les singletons sont tous de masse nulle. Alors, le risque lié à l'incertitude est total. Tel sera le cas si la probabilité que les rendements de l'actif suivent une loi normale

et si la probabilité que les rendements de l'actif suivent une loi lognormale valent toutes deux 0. Cette fois-ci, il est nécessaire que ces deux probabilités valent 0, car pour que la masse de Θ soit égale à 1 ; la probabilité que les rendements de l'actif ne suivent pas une loi normale et de celle que les rendements de l'actif ne suivent pas une loi lognormale doivent toutes les deux valoir 1.

Cette mesure de risque sera appelée **risque d'incertitude** et notée par ψ .

Définition 19.1. *Dans le cadre d'un classement de n actifs, basé sur le Modèle de performance et incertitude, le risque d'un actif j , avec $j = 1, \dots, n$, lié à l'incertitude est donné par*

$$\psi_j := m_j(\Theta), \quad (19.3)$$

où $m_j(\cdot)$ est la fonction de masse de la meta-indication \mathcal{M}_j relative à l'actif j , et est appelé **risque d'incertitude**.

Le risque d'incertitude est ainsi défini dans le cadre du Modèle de performance et incertitude. Ce modèle est lui-même basé sur les TMP et TPMP. On peut cependant aisément étendre ce même concept de risque à d'autres théories. Imaginons que, au lieu de se baser uniquement sur la TMP et la TPMP, le Modèle de performance et incertitude soit basé sur n théories financières. Imaginons également que k de ces théories soient fondées sur des hypothèses sur la distribution des rendements et que $n - k$ ne soient pas fondées sur ce type d'hypothèses. Chaque théorie financière pourra être représentée par une indication classique, puis, chaque hypothèse sur la distribution des rendements pourra être représentée par une indication de confirmation, comme cela a été fait pour la TMP et la TPMP. Il sera nécessaire ensuite de combiner chaque indication de confirmation avec l'indication classique qu'elle confirme. Nous obtiendrons ainsi k meta-indications et $n - k$ indications classiques. Enfin, en combinant toutes ces meta-indications et indications classiques en une seule et unique meta-indication \mathcal{M} par la règle de Dempster, nous retrouverons le risque d'incertitude donné par la masse du cadre de discernement de cette meta-indication \mathcal{M} .

19.3 Classement final

Concrètement, dans le cadre de ce modèle, le passage des degrés de support et de plausibilité des meta-indications \mathcal{M} des actifs considérés au classement final ces mêmes actifs s'opère de manière relativement simple. Soient $sp(1), sp(2), \dots, sp(n)$ et $pl(1), pl(2), \dots, pl(n)$ respectivement les degrés de support et de plausibilité des éléments de Θ de la meta-indication \mathcal{M}_j relative à un actif j considéré. Chacun d'entre eux indique quel est le degré de support ou de plausibilité que l'actif considéré occupe une place concrète dans

le classement. Par exemple, $sp(1)$ indique le degré de support que l'actif considéré soit classé premier.

Notre objectif est de construire un classement de n actifs, basé sur sept mesures : les six mesures de performance classiques et le risque d'incertitude. Comme les six mesures de performance sont fondées sur les critères de rendement espéré et de risque défini en termes de variabilité, notre classement sera fondé sur les trois critères suivants :

1. le rendement espéré,
2. le risque défini en termes de variabilité,
3. le risque d'incertitude.

Le premier critère doit avoir une influence positive sur le classement, tandis que les deux autres doivent l'influencer de manière négative.

Les deux premiers critères sont déjà intégrés dans les six mesures de performance traitées. Ils apparaissent dans notre modèle à travers le degré de support ou de plausibilité pour un actif accordé à chaque rang dans le classement. Quant au troisième, nous avons proposé de le définir comme la masse attribuée à Θ .

De même que pour le risque défini en termes de variabilité, un niveau élevé du risque d'incertitude doit pénaliser le classement d'un actif, tandis qu'un niveau faible devrait le favoriser.

Nous proposons ici deux manières de procéder au classement définitif, l'une basée sur les degrés de support, l'autre sur les degrés de plausibilité. Leurs résultats respectifs sont relativement proches.

19.4 Classement définitif basé sur les degrés de support

Nous disposons de n actifs que nous désirons classer du plus performant au moins performant. Nous avons calculé des degrés de support de la meta-indication \mathcal{M}_j pour chaque actif j , avec $j = 1, \dots, n$ et donc à chaque rang i , avec $i = 1, \dots, n$ se trouvant dans Θ est associé un certain degré de support, $sp(i)$. Pour déterminer les places des actifs dans le classement définitif, nous devons en premier lieu calculer un indice pour chaque actif séparément. Nous appelons cet indice SP_j pour l'actif j , car il est basé sur les degrés de support.

L'indice SP_j est construit de manière à prendre des valeurs élevées pour les actifs performants et des valeurs faibles pour les actifs peu performants. Pour obtenir SP_j , nous multiplions chaque degré de support $sp(i)$, avec $i = 1, \dots, n$, par $(n + 1 - i)$. En effet, si dans un classement de n actifs, un actif considéré est classé au rang i , la valeur $(n + 1 - i)$ correspond au rang inverse de cet actif. Ainsi, $sp(1)$ est multiplié par n et $sp(n)$ par 1.

Inverser les rangs de cette manière est nécessaire pour obtenir des valeurs élevées de l'indice pour les actifs performants et vice-versa. Puis, nous additionnons tous les produits $sp(i)(n + 1 - i)$ et obtenons ainsi l'indice SP_j . En procédant ainsi, des actifs dont les degrés de support non nuls sont associés aux places les plus performantes, se retrouvent avec des indices SP_j élevés, tandis que les actifs dont les degrés de support non nuls sont associés aux places les moins performantes, se retrouvent avec des indices SP_j faibles.

On a ainsi, pour l'actif j , l'indice suivant :

$$SP_j := \sum_{i=1}^n sp(i)(n + 1 - i), \quad (19.4)$$

qu'on appelle **indice de support**.

La pénalisation dans le classement par le risque d'incertitude est **déjà intégrée** dans l'indice SP_j ! En effet, un actif j , dont le risque d'incertitude est maximal, aurait $\psi_j = 1$. Cela est dû au fait que dans la meta-indication \mathcal{M}_j qui lui est associée, $m(\Theta) = 1$. Etant donné que la somme des masses est égale à 1 par définition, il s'en suit que tous les autres sous-ensembles de Θ ont des masses nulles, et par conséquent, des degrés de support nuls. Ainsi, l'indice de performance associé à cet actif est

$$SP_j = 0.$$

Il s'agit de la **valeur minimale** que l'indice peut prendre.

Deux actifs, par exemple k et l , peuvent avoir des classements tout à fait différents, donnés par les six mesures de performance de la TMP et la TPMP et aboutir, pourtant, à des indices nuls, $SP_k = SP_l = 0$ et se retrouver ainsi à la fin du classement. Tel sera le cas si $\psi_k = \psi_l = 1$. En effet, peu importe leurs classements respectifs donnés par les six mesures de performance, si leur risque d'incertitude est maximal, alors ces mesures perdent toute leur pertinence. Un risque d'incertitude maximal et dû à la masse de Θ qui vaut 1. Pour que cela arrive, il est nécessaire que la probabilité que les rendements de l'actif ne suivent pas une loi normale et celle que les rendements de l'actif ne suivent pas une loi lognormale valent toutes deux 1. Autrement dit, ni l'hypothèse de la TMP, ni celle de la TPMP ne sont respectées. Cela explique pourquoi, dans ce cas précis, les mesures de performance données par ces deux théories ne se reflètent plus du tout dans notre classement.

Inversement, si plusieurs actifs j ont un risque d'incertitude nul, $\psi_j = 0$, cela serait dû au fait que leurs rendements suivent, avec probabilité 1, une loi normale ou lognormale. La totalité de la masse sera répartie entre les singletons, car la masse de Θ sera nulle. Dans ce cas, les six mesures de performance vont se refléter entièrement dans notre classement. C'est en effet les rangs accordés par chacune d'elles, ainsi que les degrés de support qui permettront de départager les actifs entre eux, le risque d'incertitude n'intervenant pas dans le classement. Par exemple, si toutes les six mesures de performance classent un actif j à la première place, avec un risque d'incertitude nul, alors, $m(1) = 1$ et les masses de tous les autres singletons, de même que celle de Θ seront nulles.

On obtient ainsi

$$SP_j = \sum_{i=1}^n sp(i)(n+1-i) = n.$$

Il s'agit de la **valeur maximale** que l'indice peut prendre.

Un actif peut également obtenir l'indice de support maximal n dans d'autres circonstances. En effet, imaginons que ses rendements suivent une loi normale avec probabilité 1 et une loi lognormale avec probabilité 0. Imaginons également que les trois mesures de performance de la TMP le classent en première position. Comme notre modèle éliminera les classements des trois mesures de performance de la TPMP, car ne répondant pas à leur hypothèse de base, nous nous retrouverons dans le même cas que celui cité précédemment.

En résumé,

$$0 \leq SP_j \leq n, \quad (19.5)$$

pour tout actif j dans un classement de n actifs basé sur les degrés de support.

Le fait que l'indice de support soit borné de deux côtés permet de créer un indice de support normé, qui donne la possibilité de positionner un actif j par rapport à sa performance indépendamment des autres actifs. Ce n'est pas le cas des six mesures de performance, dont aucune n'a de borne inférieure ou supérieure.

On définit ainsi l'**indice de support normé**, donné par

$$SP_j^* := \frac{SP_j}{n}, \quad (19.6)$$

où

$$0 \leq SP_j^* \leq 1, \quad (19.7)$$

pour tout actif j dans un classement de n actifs basé sur les degrés de support.

Nous avons malgré tout besoin de n actifs, avec $n \geq 2$, pour calculer SP_j^* , car nous avons besoin des classements obtenus par les six mesures de performance pour déterminer la fonction Γ de la meta-indication relative à l'actif j .

19.5 Classement définitif basé sur les degrés de plausibilité

A nouveau, nous disposons de n actifs que nous désirons classer du plus performant au moins performant. Nous avons calculé des degrés de plausibilité de la meta-indication \mathcal{M}_j pour chaque actif j , avec $j = 1, \dots, n$ et donc à chaque rang i , avec $i = 1, \dots, n$ se trouvant dans Θ est associé un certain degré de support, $pl(i)$. Pour déterminer les places des actifs dans le classement définitif, nous devons en premier lieu calculer un indice pour chaque

actif séparément. Nous appelons cet indice PL_j pour l'actif j , car il est basé sur les degrés de plausibilité.

L'indice PL_j est construit de manière à prendre des valeurs faibles pour les actifs performants et des valeurs élevées pour les actifs peu performants. Pour obtenir PL_j , nous multiplions chaque degré de plausibilité $pl(i)$, avec $i = 1, \dots, n$, par (i) , qui représente un rang dans le classement. Puis, nous additionnons tous les produits $pl(i)i$ et obtenons ainsi l'indice PL_j . En procédant ainsi, des actifs dont les degrés de plausibilité non nuls sont associés aux places les plus performantes, se retrouvent avec des indices PL_j faibles, tandis que les actifs dont les degrés de plausibilité non nuls sont associés aux places les moins performantes, se retrouvent avec des indices SP_j élevés.

On a ainsi, pour l'actif j , l'indice suivant :

$$PL_j := \sum_{i=1}^n pl(i)i, \quad (19.8)$$

qu'on appelle **indice de plausibilité**.

De même que dans le classement basé sur les degrés de support, la pénalisation dans le classement par le risque d'incertitude est **déjà intégrée** dans l'indice PL_j . En effet, un actif j , dont le risque d'incertitude est maximal, c'est-à-dire avec $\psi_j = 1$, aurait $m(\Theta) = 1$, dans la meta-indication \mathcal{M}_j qui lui est associée. Etant donné que la somme des masses est égale à 1 par définition, il s'en suit que tous les autres sous-ensembles de Θ ont des masses nulles, et par conséquent, des degrés de plausibilité de 1. Ainsi, l'indice de performance associé à cet actif est

$$PL_j = \sum_{i=1}^n pl(i)i = \sum_{i=1}^n i = \frac{n(n+1)}{2}.$$

Il s'agit de la **valeur maximale** que l'indice peut prendre.

Deux actifs, par exemple k et l , peuvent avoir des classements tout à fait différents, donnés par les six mesures de performance de la TMP et la TPMP et aboutir, pourtant, à des indices maximaux, $PL_k = PL_l = \frac{n(n+1)}{2}$. Ils se retrouvent ainsi à la fin du classement. Comme dans le classement basé sur les indices de support, tel sera le cas si $\psi_k = \psi_l = 1$.

Inversement, si plusieurs actifs j ont un risque d'incertitude nul, $\psi_j = 0$, cela serait dû au fait que leurs rendements suivent parfaitement une loi normale ou lognormale. A nouveau, les six mesures de performance vont se refléter entièrement dans notre classement. La somme des degrés de plausibilité accordés aux singletons sera égale à 1, aucune masse n'étant accordée à Θ . Par exemple, si toutes les six mesures de performance classent un actif j à la première place, avec un risque d'incertitude nul, alors, $m(1) = 1$ et les masses de tous les autres singletons, de même que celle de Θ seront nulles.

On obtient ainsi

$$PL_j = \sum_{i=1}^n pl(i)i = 1.$$

Il s'agit de la **valeur minimale** que l'indice peut prendre.

Une actif peut également obtenir un indice de plausibilité minimal de 1 dans d'autres circonstances. Comme dans la section précédente, si ses rendements suivent une loi normale avec probabilité 1 et une loi lognormale avec probabilité 0 et que les trois mesures de performance de la TMP la classent en première position, son indice PL_j sera minimal.

En conclusion,

$$1 \leq PL_j \leq \frac{n(n+1)}{2}, \quad (19.9)$$

pour toute actif j dans un classement de n actifs basé sur les degrés de plausibilité.

Le fait que l'indice de plausibilité, de même que l'indice de support, soit borné de deux côtés, permet de créer un indice de plausibilité normé, qui donne à nouveau la possibilité de positionner un actif j par rapport à sa performance indépendamment des autres actifs.

On définit ainsi l'**indice de plausibilité normé**, donné par

$$PL_j^* = \frac{2PL_j}{n(n+1)}, \quad (19.10)$$

où

$$\frac{2}{n(n+1)} \leq PL_j^* \leq 1, \quad (19.11)$$

pour tout actif j dans un classement de n actifs basé sur les degrés de plausibilité.

On sait que

$$\lim_{n \rightarrow \infty} \frac{2}{n(n+1)} = 0,$$

par conséquent

$$0 < PL_j^* \leq 1, \quad (19.12)$$

pour tout actif j dans un classement de n actifs basé sur les degrés de plausibilité.

Enfin, pour les mêmes raisons que celles citées pour l'indice de support normé, nous avons besoin de n actifs, avec $n \geq 2$, pour calculer PL_j .

En conclusion, on notera que le classement peut varier selon le choix des indices utilisés, de support ou de plausibilité. Cela relève du fait que le support et la plausibilité expriment des concepts différents. Cependant, les actifs qui obtiennent un indice de support maximal, obtiennent toujours un indice de plausibilité minimal et sont classés premiers. De même, ceux qui obtiennent un indice de support minimal, obtiennent toujours un indice de plausibilité maximal et sont classés derniers.

Chapitre 20

Exemple d'application

Nous avons appliqué le Modèle de performance et incertitude, exposé au chapitre précédent, à 10 actions suisses cotées en bourse. Pour cela, nous avons observé leurs rendements journaliers R_t et r_t du 2.01.2001 au 31.12.2001, ainsi que ceux du Swiss Market Index (SMI).

Voici la liste des 10 actions considérées :

- **ADEN** : action de l'entreprise Adecco SA
- **BAER** : action de l'entreprise Julius Baer Holding
- **BALN** : action de l'entreprise Bâloise-Holding
- **CIBN** : action de l'entreprise Ciba Specialty Chemicals
- **CLN** : action de l'entreprise Clariant SA
- **CSGN** : action de l'entreprise Credit Suisse Group
- **KUD** : action de l'entreprise Kudelski SA
- **UHR** : action de l'entreprise Swatch Group SA
- **SCMN** : action de l'entreprise Swisscom SA
- **UBSN** : action de l'entreprise UBS SA

En tout premier lieu, on a classé les 10 actions susmentionnées d'après les six mesures de performance de la TMP et TPMP. Pour cela, on a dû estimer des paramètres pour chaque action ainsi que pour le SMI.

20.1 Classements d'après les mesures de performance de la TMP

Dans cette section, on s'est basé sur un taux sûr r_F annuel de 3%. Le taux de rendement annuel des obligations de la Confédération en 2001 a été de 3,56%. Cependant, comme la bourse était en crise durant cette année (voir la moyenne des rendements des Tables 20.1 et 20.4), on a choisi le taux de 3% pour s'approcher d'avantage de la réalité. Nous avons trouvé le taux sûr journalier en calculant

$$\sqrt[260]{(1 + 0.03)} - 1 = 0.0113694\%,$$

car les rendements du 2.01.2001 au 31.12.2001 correspondent à 260 jours ouvrables ou 260 observations.

Pour calculer les trois mesures de performance de la TMP, nous avons dû estimer l'espérance mathématique et l'écart-type des rendements journaliers \tilde{r}_t des dix actions susmentionnées et du SMI. Nous avons choisi comme estimateurs la moyenne arithmétique et l'écart-type empirique respectivement. Pour calculer l'indice de Treynor et l'alpha de Jensen, nous avons également estimé pour chaque action P , avec $P = 1, \dots, 10$, α_P et β_P , tels qu'il sont défini dans l'équation (17.6). Les résultats obtenus se trouvent dans la Table 20.1.

Sur la base des données de la Table 20.1 et du taux sûr R_F journalier, nous avons calculé le Ratio de Sharpe S_P , l'indice de Treynor T_P et l'alpha de Jensen α_P , pour chaque action P . Les valeurs de ces indices sont présentées dans la Table 20.2.

Sur la base des données de la Table 20.2, nous avons établi le classement d'après les trois

	Moyenne	Ecart-type	α_P	β_P
ADEN	0.00004	0.03225	0.00111	1.27895
BAER	-0.00143	0.02473	-0.00064	0.98076
BALN	-0.00037	0.01972	0.00036	0.91107
CIBN	-0.00008	0.01624	0.00042	0.66479
CLN	-0.00193	0.03008	-0.00093	1.20236
CSGN	0.00009	0.02484	0.00129	1.42051
KUD	-0.00127	0.04732	0.00041	1.94412
UHR	-0.00071	0.02830	0.00003	0.92615
SCMN	0.00054	0.01632	0.00091	0.53025
UBSN	0.00005	0.02234	0.00111	1.26953
SMI	-0.00081	0.01436		

TAB. 20.1 – Paramètres estimés pour les 10 actions et le SMI

	S_P	T_P	α_P
ADEN	-0.00219	-0.00006	0.00111
BAER	-0.06245	-0.00157	-0.00064
BALN	-0.02430	-0.00053	0.00036
CIBN	-0.01208	-0.00030	0.00042
CLN	-0.06789	-0.00170	-0.00093
CSGN	-0.00091	-0.00002	0.00129
KUD	-0.02914	-0.00071	0.00041
UHR	-0.02908	-0.00089	0.00003
SCMN	0.02592	0.00080	0.00091
UBSN	-0.00273	-0.00005	0.00111

TAB. 20.2 – Indices de la TMP

indices de la TMP, de l'action la plus performante à l'action la moins performante. Ces classements se trouvent dans la Table 20.3.

	S_P	T_P	α_P
ADEN	3	4	3
BAER	9	9	9
BALN	6	6	7
CIBN	5	5	5
CLN	10	10	10
CSGN	2	2	1
KUD	8	7	6
UHR	7	8	8
SCMN	1	1	4
UBSN	4	3	2

TAB. 20.3 – Classement d'après les trois indices de la TMP

20.2 Classements d'après les mesures de performance de la TPMP

Dans cette section, on s'est basé également sur un taux sûr R_F annuel de 3%.

Pour calculer les trois mesures de performance de la TPMP, nous avons dû estimer l'espérance mathématique et le semi-écart-type (SET), défini dans l'équation (18.6), des rendements journaliers \tilde{R}_t des dix actions susmentionnées et du SMI. Nous avons choisi comme estimateurs la moyenne arithmétique et le semi-écart-type (SET) empirique res-

pectivement. Pour calculer l'indice de Treynor et l'alpha de Jensen, nous avons également estimé pour chaque action P , avec $P = 1, \dots, 10$, α_P^{SET} et β_P^{SET} , tels qu'ils sont définis dans les équations (18.9) et (18.10) respectivement. Les résultats obtenus se trouvent dans la Table 20.4.

	Moyenne	SET	α_P^{SET}	β_P^{SET}
ADEN	-0.00047	0.02008	0.00912	1.74017
BAER	-0.00174	0.01611	0.00528	1.27845
BALN	-0.00056	0.01652	0.00669	1.32007
CIBN	-0.00022	0.01431	0.00499	0.95436
CLN	-0.00239	0.02378	0.00624	1.56610
CSGN	-0.00022	0.01820	0.00994	1.84056
KUD	-0.00238	0.03166	0.01093	2.40510
UHR	-0.00111	0.01950	0.00490	1.09713
SCMN	0.00040	0.01104	0.00389	0.64573
UBSN	-0.00020	0.01511	0.00847	1.57391
SMI	-0.00091	0.00954		

TAB. 20.4 – Paramètres estimés pour les 10 actions et le SMI

Sur la base des données de la Table 20.4 et du taux sûr R_F journalier, nous avons calculé le Ratio de Sortino S_P^{SET} , l'indice de Treynor (SET) T_P^{SET} et l'alpha de Jensen (SET) α_P^{SET} , pour chaque action P de notre liste. Les valeurs de ces indices sont présentées dans la Table 20.5.

Sur la base des données de la Table 20.5, nous avons établi le classement d'après les trois indices de la TPMP, de l'action la plus performante à l'action la moins performante. Ces classements se trouvent dans la Table 20.6.

	S_P^{SET}	T_P^{SET}	α_P^{SET}
ADEN	-0.02911	-0.00034	0.00912
BAER	-0.11483	-0.00145	0.00528
BALN	-0.04086	-0.00051	0.00669
CIBN	-0.02303	-0.00035	0.00499
CLN	-0.10506	-0.00160	0.00624
CSGN	-0.01811	-0.00018	0.00994
KUD	-0.07870	-0.00104	0.01093
UHR	-0.06274	-0.00112	0.00490
SCMN	0.02630	0.00045	0.00389
UBSN	-0.02045	-0.00020	0.00847

TAB. 20.5 – Indices de la TPMP

	S_P^{SET}	T_P^{SET}	α_P^{SET}
ADEN	5	4	3
BAER	10	9	7
BALN	6	6	5
CIBN	4	5	8
CLN	9	10	6
CSGN	2	2	2
KUD	8	7	1
UHR	7	8	9
SCMN	1	1	10
UBSN	3	3	4

TAB. 20.6 – Classement d’après les trois indices de la TPMP

20.3 Construction des meta-indications

Nous devons à présent construire la meta-indication définie dans la section 19.1 pour chacune de 10 actions. Nous allons appeler ces 10 meta-indications \mathcal{M}_P avec $P = 1, \dots, 10$.

Pour chaque \mathcal{M}_P , définie par le quadruplet $(\Omega_P, P_P, \Gamma_P, \Theta)$, nous connaissons déjà Ω_P , qui est identique pour les 10 actions. Nous connaissons également Γ_P , qui diffère d’une action à l’autre et qui est déterminé sur la base des classements obtenus par les six mesures de performance, comme expliqué dans la section 19.1. Nous connaissons enfin Θ qui est composé des 10 singletons suivants : $\{1\}, \{2\}, \dots, \{10\}$.

Pour déterminer chaque \mathcal{M}_P , il ne manque plus que les fonctions de probabilité p_P qui, comme nous allons le voir, vont différer pour chaque action. Pour les estimer, nous avons besoin de connaître les fonctions de probabilité de \mathcal{H}_1 et de \mathcal{H}_2 , ainsi que celles de Con_3 et Con_4 , indications définies à la section 19.1.

Commençons par Con_3 et Con_4 . Rappelons les éléments de Ω_3 ,

- $\omega_{3,c}$: la distribution des rendements est normale,
- $\omega_{3,f}$: la distribution des rendements n’est pas normale,

et ceux de Ω_4

- $\omega_{4,c}$: la distribution des rendements est lognormale,
- $\omega_{4,f}$: la distribution des rendements n’est pas lognormale.

Les fonctions de probabilité que nous cherchons sont des fonctions de probabilité sur les ensembles Ω_3 et Ω_4 . Chaque fonction représente en effet la probabilité de chacun des deux éléments d’être l’interprétation correcte.

Pour chaque action, nous disposons d’un échantillon de 260 rendements journaliers discrets r_t et d’un échantillon de 260 rendements journaliers continus R_t , $t = 1, \dots, 260$. Nous

sommes intéressés à savoir quelle est la probabilité que ces rendements suivent une loi normale ou respectivement une loi lognormale. Pour cela, nous avons composé, sur la base de chaque échantillon, 161 sous-échantillons de 100 observations comme expliqué à la section 19.1. Puis, pour chaque sous-échantillon, nous avons testé l'hypothèse H_0 , si ses rendement suivent une loi normale de paramètres précis, lorsqu'il s'agissait d'un sous-échantillon de rendements r_t , ou une loi lognormale de paramètres précis, lorsqu'il s'agissait d'un sous-échantillon de rendements R_t . Ces hypothèses ont été testées avec le test de Kolmogorov-Smirnov. Pour chaque action P , nous avons choisi une loi normale et une loi lognormale, dont les paramètres $(\mu_P, \sigma_P^2$ et μ_{LP}, σ_{LP}^2 respectivement) ont été estimés sur la base de l'échantillon entier, car nous sommes intéressés à savoir si tous les rendements suivent la même loi précise et non juste une famille de lois normale ou lognormale. Les estimations de ces paramètres se trouvent dans la Table 20.7.

Pour ces tests, nous avons enfin choisi un seuil de signification de 5%, puis nous avons observé pour chaque action, pour combien de sous-échantillons sur 161 l'hypothèse H_0 a été refusée. Nous avons choisi cette proportion comme estimateur de la probabilité de $\omega_{3,f}$ pour les sous-échantillons de rendements r_t et de $\omega_{4,f}$ pour les sous-échantillons de rendements R_t .

Nous avons obtenu ainsi les probabilités présentées dans la Table 20.8.

Les fonctions de probabilités de \mathcal{H}_1 et de \mathcal{H}_2 sont à déterminer par chaque expert en fonctions de ses connaissances et de son expérience, comme expliqué au Chapitre 19. Elles sont identiques pour chaque action. Nous avons choisi pour illustrer notre modèle les fonctions de probabilités p_1 et p_2 présentées dans la Table 20.9.

Nous pouvons à présent trouver la fonction de probabilité de chaque M_P . Nous obtenons les 10 fonctions présentées dans la Table 20.10.

	μ_P	σ_P^2	μ_{LP}	σ_{LP}^2
ADEN	0.00004	0.00104	-0.00047	0.00103
BAER	-0.00143	0.00061	-0.00174	0.00061
BALN	-0.00037	0.00039	-0.00056	0.00040
CIBN	-0.00008	0.00026	-0.00022	0.00027
CLN	-0.00193	0.00090	-0.00239	0.00092
CSGN	0.00009	0.00062	-0.00022	0.00062
KUD	-0.00127	0.00224	-0.00238	0.00223
UHR	-0.00071	0.00080	-0.00111	0.00080
SCMN	0.00054	0.00027	0.00040	0.00027
UBSN	0.00005	0.00050	-0.00020	0.00050

TAB. 20.7 – Paramètres estimés pour chaque action

	$p_3(\omega_{3,f})$	$p_4(\omega_{4,f})$
ADEN	0.14286	0.13043
BAER	0.17391	0.17391
BALN	0.62112	0.60870
CIBN	0.93789	0.94410
CLN	0.60248	0.62112
CSGN	0.46584	0.46584
KUD	0.00000	0.00000
UHR	0.11180	0.11180
SCMN	0.19255	0.19255
UBSN	0.01242	0.00000

TAB. 20.8 – Probabilités estimées de $p_3(\omega_{3,f})$ et $p_4(\omega_{4,f})$

P_1		P_2	
$p_1(\omega_{1,1})$	0.40	$p_2(\omega_{2,1})$	0.30
$p_1(\omega_{1,2})$	0.35	$p_2(\omega_{2,2})$	0.50
$p_1(\omega_{1,3})$	0.25	$p_2(\omega_{2,3})$	0.20

TAB. 20.9 – Fonctions de probabilités de H_1 et H_2

P_P^*	ADEN	BAER	BALN	CIBN	CLN	CSGN	KUD	UHR	SCMN	UBSN
$p^*(\omega_{1,1}, \omega_{2,1})$	0.089	0.082	0.018	0.000	0.018	0.034	0.120	0.095	0.078	0.119
$p^*(\omega_{1,2}, \omega_{2,1})$	0.078	0.072	0.016	0.000	0.016	0.030	0.105	0.083	0.068	0.104
$p^*(\omega_{1,3}, \omega_{2,1})$	0.056	0.051	0.011	0.000	0.011	0.021	0.075	0.059	0.049	0.074
$p^*(\omega_{3,f}, \omega_{2,1})$	0.037	0.043	0.073	0.016	0.068	0.075	0.000	0.030	0.047	0.004
$p^*(\omega_{1,1}, \omega_{2,2})$	0.149	0.136	0.030	0.001	0.030	0.057	0.200	0.158	0.130	0.198
$p^*(\omega_{1,2}, \omega_{2,2})$	0.130	0.119	0.026	0.001	0.026	0.050	0.175	0.138	0.114	0.173
$p^*(\omega_{1,3}, \omega_{2,2})$	0.093	0.085	0.019	0.000	0.019	0.036	0.125	0.099	0.081	0.123
$p^*(\omega_{3,f}, \omega_{2,2})$	0.062	0.072	0.122	0.026	0.114	0.124	0.000	0.050	0.078	0.006
$p^*(\omega_{1,1}, \omega_{2,3})$	0.060	0.055	0.012	0.000	0.012	0.023	0.080	0.063	0.052	0.079
$p^*(\omega_{1,2}, \omega_{2,3})$	0.052	0.048	0.010	0.000	0.011	0.020	0.070	0.055	0.046	0.069
$p^*(\omega_{1,3}, \omega_{2,3})$	0.037	0.034	0.007	0.000	0.008	0.014	0.050	0.039	0.033	0.049
$p^*(\omega_{3,f}, \omega_{2,3})$	0.025	0.029	0.049	0.010	0.046	0.050	0.000	0.020	0.031	0.002
$p^*(\omega_{1,1}, \omega_{4,f})$	0.045	0.057	0.092	0.023	0.099	0.100	0.000	0.040	0.062	0.000
$p^*(\omega_{1,2}, \omega_{4,f})$	0.039	0.050	0.081	0.021	0.086	0.087	0.000	0.035	0.054	0.000
$p^*(\omega_{1,3}, \omega_{4,f})$	0.028	0.036	0.058	0.015	0.062	0.062	0.000	0.025	0.039	0.000
$p^*(\omega_{3,f}, \omega_{4,f})$	0.019	0.030	0.378	0.885	0.374	0.217	0.000	0.012	0.037	0.000

TAB. 20.10 – Fonctions de probabilités (non ajustées) des 10 \mathcal{M}_P

Il s'agit évidemment des fonctions de probabilités non ajustées. Pour obtenir les fonctions de probabilités ajustées, nous devons encore éliminer tous les événements impossibles, c'est-à-dire dont l'ensemble focal est donné par l'ensemble vide. La Table 20.11 nous donne l'ensemble focal de chaque interprétation des 10 \mathcal{M}_P .

En ajustant les fonctions de probabilités de la Table 20.10, nous obtenons les fonctions de probabilité ajustées, présentées dans la Table 20.12.

Sur la base des valeurs de la Table 20.12, nous sommes à présent en mesure de calculer les masses, les degrés de support et de plausibilité des différents éléments de Θ pour chaque \mathcal{M}_P .

La Table résume les masses des 10 singletons de Θ , ainsi que de Θ lui-même pour la meta-indication \mathcal{M}_P de chaque action P .

Voici enfin dans les Tables 20.14 et 20.15 les degrés de support et de plausibilité respectivement des 10 singletons de Θ pour la meta-indication \mathcal{M}_P de chaque action P .

Disposant de toutes les informations qui nous sont nécessaires, nous sommes à présent en mesure de déterminer les indices de support SP_P , de support normé SP_P^* , de plausibilité PL_P et enfin de plausibilité normé PL_P^* pour chaque action P .

Ces indices se trouvent dans la Table 20.16.

Γ_P	ADEN	BAER	BALN	CIBN	CLN	CSGN	KUD	UHR	SCMN	UBSN
$\Gamma(\omega_{1,1}, \omega_{2,1})$	\emptyset	\emptyset	{6}	\emptyset	\emptyset	{2}	{8}	{7}	{1}	\emptyset
$\Gamma(\omega_{1,2}, \omega_{2,1})$	\emptyset	\emptyset	{6}	\emptyset	\emptyset	{2}	\emptyset	\emptyset	{1}	{3}
$\Gamma(\omega_{1,3}, \omega_{2,1})$	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset
$\Gamma(\omega_{3,f}, \omega_{2,1})$	{5}	{10}	{6}	{4}	{9}	{2}	{8}	{7}	{1}	{3}
$\Gamma(\omega_{1,1}, \omega_{2,2})$	\emptyset	{9}	{6}	{5}	{10}	{2}	\emptyset	\emptyset	{1}	\emptyset
$\Gamma(\omega_{1,2}, \omega_{2,2})$	{4}	{9}	{6}	{5}	{10}	{2}	{7}	{8}	{1}	{3}
$\Gamma(\omega_{1,3}, \omega_{2,2})$	\emptyset	{9}	\emptyset	{5}	{10}	\emptyset	\emptyset	{8}	\emptyset	\emptyset
$\Gamma(\omega_{3,f}, \omega_{2,2})$	{4}	9	{6}	{5}	{10}	{2}	{7}	{8}	{1}	{3}
$\Gamma(\omega_{1,1}, \omega_{2,3})$	{3}	\emptyset	\emptyset	\emptyset	\emptyset	{2}	\emptyset	\emptyset	\emptyset	{4}
$\Gamma(\omega_{1,2}, \omega_{2,3})$	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	{2}	\emptyset	{8}	\emptyset	\emptyset
$\Gamma(\omega_{1,3}, \omega_{2,3})$	{3}	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	\emptyset	{8}	\emptyset	\emptyset
$\Gamma(\omega_{3,f}, \omega_{2,3})$	{3}	{7}	{5}	{8}	{6}	{2}	{1}	{8}	{10}	{4}
$\Gamma(\omega_{1,1}, \omega_{4,f})$	{3}	{9}	{6}	{5}	{10}	{2}	{8}	{7}	{1}	{4}
$\Gamma(\omega_{1,2}, \omega_{4,f})$	{4}	{9}	{6}	{5}	{10}	{2}	{7}	{8}	{1}	{3}
$\Gamma(\omega_{1,3}, \omega_{4,f})$	{3}	{9}	{7}	{5}	{10}	{1}	{6}	{8}	{4}	{2}
$\Gamma(\omega_{3,f}, \omega_{4,f})$	Θ	Θ	Θ	Θ	Θ	Θ	Θ	Θ	Θ	Θ

TAB. 20.11 – Ensembles focaux des 10 \mathcal{M}_P

P_P	ADEN	BAER	BALN	CIBN	CLN	CSGN	KUD	UHR	SCMN	UBSN
$p(\omega_{1,1}, \omega_{2,1})$	0.000	0.000	0.019	0.000	0.000	0.037	0.407	0.175	0.106	0.000
$p(\omega_{1,2}, \omega_{2,1})$	0.000	0.000	0.017	0.000	0.000	0.032	0.000	0.000	0.093	0.282
$p(\omega_{1,3}, \omega_{2,1})$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$p(\omega_{3,f}, \omega_{2,1})$	0.077	0.065	0.078	0.016	0.074	0.080	0.000	0.055	0.063	0.010
$p(\omega_{1,1}, \omega_{2,2})$	0.000	0.207	0.032	0.001	0.033	0.061	0.000	0.000	0.176	0.000
$p(\omega_{1,2}, \omega_{2,2})$	0.271	0.181	0.028	0.001	0.029	0.054	0.593	0.255	0.154	0.470
$p(\omega_{1,3}, \omega_{2,2})$	0.000	0.129	0.000	0.000	0.020	0.000	0.000	0.182	0.000	0.000
$p(\omega_{3,f}, \omega_{2,2})$	0.129	0.109	0.129	0.026	0.123	0.134	0.000	0.092	0.105	0.017
$p(\omega_{1,1}, \omega_{2,3})$	0.124	0.000	0.000	0.000	0.000	0.025	0.000	0.000	0.000	0.215
$p(\omega_{1,2}, \omega_{2,3})$	0.000	0.000	0.000	0.000	0.000	0.022	0.000	0.000	0.000	0.000
$p(\omega_{1,3}, \omega_{2,3})$	0.077	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$p(\omega_{3,f}, \omega_{2,3})$	0.052	0.044	0.052	0.011	0.049	0.054	0.000	0.037	0.042	0.007
$p(\omega_{1,1}, \omega_{4,f})$	0.093	0.087	0.098	0.023	0.107	0.107	0.000	0.073	0.084	0.000
$p(\omega_{1,2}, \omega_{4,f})$	0.081	0.076	0.086	0.021	0.093	0.094	0.000	0.064	0.074	0.000
$p(\omega_{1,3}, \omega_{4,f})$	0.058	0.055	0.061	0.015	0.067	0.067	0.000	0.046	0.053	0.000
$p(\omega_{3,f}, \omega_{4,f})$	0.039	0.046	0.402	0.887	0.405	0.234	0.000	0.023	0.050	0.000

TAB. 20.12 – Fonctions de probabilités ajustées des 10 \mathcal{M}_P

	ADEN	BAER	BALN	CIBN	CLN	CSGN	KUD	UHR	SCMN	UBSN
$m(1)$	0.000	0.000	0.000	0.000	0.000	0.067	0.000	0.000	0.855	0.000
$m(2)$	0.000	0.000	0.000	0.000	0.000	0.699	0.000	0.000	0.000	0.000
$m(3)$	0.403	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.779
$m(4)$	0.481	0.000	0.000	0.016	0.000	0.000	0.000	0.000	0.053	0.221
$m(5)$	0.077	0.000	0.052	0.087	0.000	0.000	0.000	0.000	0.000	0.000
$m(6)$	0.000	0.000	0.485	0.000	0.049	0.000	0.000	0.000	0.000	0.000
$m(7)$	0.000	0.044	0.061	0.000	0.000	0.000	0.593	0.303	0.000	0.000
$m(8)$	0.000	0.000	0.000	0.011	0.000	0.000	0.407	0.638	0.000	0.000
$m(9)$	0.000	0.845	0.000	0.000	0.074	0.000	0.000	0.037	0.000	0.000
$m(10)$	0.000	0.065	0.000	0.000	0.472	0.000	0.000	0.000	0.042	0.000
$m(\Theta)$	0.039	0.046	0.402	0.887	0.405	0.234	0.000	0.023	0.050	0.000

TAB. 20.13 – Fonctions de masse des 10 \mathcal{M}_P

	ADEN	BAER	BALN	CIBN	CLN	CSGN	KUD	UHR	SCMN	UBSN
$sp(1)$	0.000	0.000	0.000	0.000	0.000	0.067	0.000	0.000	0.855	0.000
$sp(2)$	0.000	0.000	0.000	0.000	0.000	0.699	0.000	0.000	0.000	0.000
$sp(3)$	0.403	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.779
$sp(4)$	0.481	0.000	0.000	0.016	0.000	0.000	0.000	0.000	0.053	0.221
$sp(5)$	0.077	0.000	0.052	0.087	0.000	0.000	0.000	0.000	0.000	0.000
$sp(6)$	0.000	0.000	0.485	0.000	0.049	0.000	0.000	0.000	0.000	0.000
$sp(7)$	0.000	0.044	0.061	0.000	0.000	0.000	0.593	0.303	0.000	0.000
$sp(8)$	0.000	0.000	0.000	0.011	0.000	0.000	0.407	0.638	0.000	0.000
$sp(9)$	0.000	0.845	0.000	0.000	0.074	0.000	0.000	0.037	0.000	0.000
$sp(10)$	0.000	0.065	0.000	0.000	0.472	0.000	0.000	0.000	0.042	0.000

TAB. 20.14 – Degrés de support des 10 \mathcal{M}_P

	ADEN	BAER	BALN	CIBN	CLN	CSGN	KUD	UHR	SCMN	UBSN
$pl(1)$	0.039	0.046	0.402	0.405	0.405	0.301	0.000	0.023	0.905	0.000
$pl(2)$	0.039	0.046	0.402	0.405	0.405	0.933	0.000	0.023	0.050	0.000
$pl(3)$	0.442	0.046	0.402	0.405	0.405	0.234	0.000	0.023	0.050	0.779
$pl(4)$	0.519	0.046	0.402	0.420	0.405	0.234	0.000	0.023	0.103	0.221
$pl(5)$	0.116	0.046	0.454	0.491	0.405	0.234	0.000	0.023	0.050	0.000
$pl(6)$	0.039	0.046	0.887	0.405	0.454	0.234	0.000	0.023	0.050	0.000
$pl(7)$	0.039	0.090	0.463	0.405	0.405	0.234	0.593	0.326	0.050	0.000
$pl(8)$	0.039	0.046	0.402	0.415	0.405	0.234	0.407	0.661	0.050	0.000
$pl(9)$	0.039	0.891	0.402	0.405	0.479	0.234	0.000	0.060	0.050	0.000
$pl(10)$	0.039	0.111	0.402	0.405	0.877	0.234	0.000	0.023	0.092	0.000

TAB. 20.15 – Degrés de plausibilité des 10 \mathcal{M}_P

	SP_P	SP_P^*	PL_P	PL_P^*
ADEN	7.05541	0.70554	5.64562	0.06273
BAER	1.92999	0.19300	11.09010	0.12322
BALN	2.98085	0.29809	25.70302	0.28559
CIBN	0.66225	0.06622	49.36573	0.54851
CLN	0.86686	0.08669	27.93948	0.31044
CSGN	6.96392	0.69639	14.31773	0.15909
KUD	3.59322	0.35932	7.40678	0.08230
UHR	3.19693	0.31969	8.81697	0.09797
SCMN	8.96214	0.89621	4.24464	0.04716
UBSN	7.77853	0.77785	3.22147	0.03579

TAB. 20.16 – Indices de support et de plausibilité des 10 \mathcal{M}_P

Sur la base des indices de la Table 20.16, nous pouvons enfin classer les actions considérées de la plus performante à la moins performante. Leur classement d'après les indices de support est présenté dans la Table 20.17.

Quant à leur classement d'après les indices de plausibilité, il est présenté dans la Table 20.18.

Ces deux classements permettent d'avoir une vision globale de la performance des 10 actifs considérés, en intégrant à la fois les six mesures de performance classiques et le respect ou non des hypothèses sur les distributions des rendements. Ils nous permettent de trouver le «juste milieu» entre les six mesures de performance, tout en tenant compte de leur pertinence, jugée d'après les distributions des rendements. Si l'on compare avec les classements basés sur les six mesures de performance classiques, on constate que l'action de l'UBS a été favorisée par le fait que ces rendements suivent une loi normale avec une probabilité de 0.99 et une loi lognormale avec une probabilité de 1. De même, l'action de Ciba s'est retrouvée à la dernière position aussi bien dans le classement basé sur les degrés de support que celui basé sur les degrés de plausibilité. En effet, bien que classée à la 5^{ème} position par les mesures S_P , T_P , α_P , T_P^{SET} , et à la 4^{ème} et 8^{ème} position par S_P^{SET} et α_P^{SET} respectivement, elle se retrouve à présent en 10^{ème} position pour une raison toute simple : ses rendements suivent aussi bien une loi normale qu'une loi lognormale avec une probabilité de 0.06 seulement. Il s'agit évidemment d'un critère que les six mesures de performance ignorent, ce qui explique leur classement de cette action. Pourtant, comme il a déjà été mentionné à de nombreuses reprises, il est important de tenir compte des hypothèses sur les distributions des rendements. Si ces hypothèses ne sont pas respectées, il devient inutile de calculer les 6 mesures de performance classiques, celles-ci perdant tout leur sens. Ainsi, un financier qui ne se base que sur les 6 mesures de performance classiques pour prendre une décision, peut être induit en erreur. En effet, ce serait le cas

Action	Rang
SCMN	1
UBSN	2
ADEN	3
CSGN	4
KUD	5
UHR	6
BALN	7
BAER	8
CLN	9
CIBN	10

TAB. 20.17 – Classement des 10 actions P d'après les indices de support

Action	Rang
UBSN	1
SCMN	2
ADEN	3
KUD	4
UHR	5
BAER	6
CSGN	7
BALN	8
CLN	9
CIBN	10

TAB. 20.18 – Classement des 10 actions P d'après les indices de plausibilité

lors d'un classement favorable par celles-ci d'une action dont les rendements ne respectent pas les hypothèses de distributions.

En outre, même en ignorant les distributions des rendements, un financier aurait des difficultés à prendre une décision rationnelle, à moins que, par pur hasard, les six mesures classent l'action considérée au même rang. Les classements basés sur les degrés de support et sur les degrés de plausibilité permettent non seulement d'avoir une vision nettement plus globale que les six mesures de performance classiques, mais surtout une vision nettement moins dépendante d'une seule mesure de performance. Ainsi, en intégrant le plus d'information possible, les classements proposés dans notre modèle sont plus robustes que les mesures classiques de la TMP et de la TPMP.

Enfin, les classements basés sur les degrés de support et sur les degrés de plausibilité ont l'avantage d'être facilement compréhensibles et ne posent aucun problème quant à leur calcul. Ils peuvent ainsi être facilement intégrés dans la pratique, tout en permettant aux spécialistes, par le biais des fonctions de probabilité de \mathcal{H}_1 et \mathcal{H}_2 , d'y introduire une pondération entre les mesures de performance classiques, basée sur leurs propres connaissances et expériences.

Conclusion

Les débuts du raisonnement sous incertitude

Ce travail a présenté, dans un premier temps, une introduction à la théorie mathématique de l'évidence de Dempster et Shafer. Ainsi, avons-nous découvert les concepts de probabilité fréquentiste et de probabilité subjective, les notions de masse, de crédibilité, de plausibilité, et bien d'autres, permettant la représentation de situations d'information incomplète, voire inexistante, floue ou ambiguë. La théorie mathématique de l'évidence est une théorie alternative à la théorie classique des probabilités. Elle la complète ; plus encore, elle la relègue au rang de cas spécial de la théorie mathématique de l'évidence. L'apport de Dempster et Shafer est remarquable. En effet, grâce à leurs travaux, il est désormais possible de distinguer des situations jusque là confondues, malgré leurs différences évidentes. La théorie classique des probabilités ne traite de manière précise que les situations d'information complète. Quant à celles dont l'information est partielle, soit elle les ignore, soit elle émet des suppositions sur l'information manquante. En ce sens, Dempster et Shafer ont réussi à en dépasser les limites.

Dans la théorie mathématique de l'évidence, on trouve également une méthode de combinaison de plusieurs sources d'information, opération que la théorie classique des probabilités ne permet pas. Par exemple, si l'on demande à plusieurs spécialistes de donner la probabilité, selon eux, de réalisation d'un événement, grâce à l'opération de combinaison de Dempster, il devient possible de combiner leurs avis et d'évaluer cette probabilité en fonction de toutes les informations reçues. En résumé, on peut dire sans exagérer que Dempster et Shafer ont révolutionné le traitement de l'information, en développant une méthode de raisonnement sous incertitude. Leur théorie a été reprise par la suite par de nombreux chercheurs qui ont développé à leur tour des versions alternatives. Parmi ces alternatives, nous nous sommes intéressés en particulier à la théorie mathématique des indications de Kohlas et Monney.

Le concept d'indication

Commencer par une brève introduction à la théorie mathématique de l'évidence nous a semblé essentiel pour la compréhension de la théorie mathématique des indications de Kohlas et Monney, présentée en second lieu, ainsi que pour la connaissance de ses

origines. Nous avons vu ses différences essentielles par rapport à la théorie mathématique de l'évidence, notamment l'extension des cadres de discernement à des cadres infinis, le remplacement du concept de crédibilité par celui du support et, finalement, l'élimination du jugement personnel qui en découle. Nous avons ainsi découvert une théorie basée sur la notion d'indication comme modèle explicite d'une part d'information incertaine et imprécise. Nous y avons également trouvé plusieurs types d'indications, ainsi que leurs propriétés spécifiques. Enfin, nous avons présenté la règle de combinaison de Dempster dans le contexte des indications.

Le problème du casino

La Partie 3 débute par l'exposition d'un problème que nous avons appelé «le problème de casino», qui représente une situation où l'information disponible est remise en question. Ce problème est utilisé pour mettre en évidence certaines limites de la théorie mathématique des indications, limites d'ailleurs également présentes dans la théorie mathématique de l'évidence. Ce problème ayant facilement pu être généralisé à toute une série de situations, il nous a semblé essentiel d'y trouver une solution. Ainsi, en nous inspirant de la théorie mathématique des indications, nous avons développé quelques concepts nouveaux, inexistants jusque là même dans les théories parallèles à celle des indications. Il s'agit notamment des concepts d'indication de confirmation, de meta-indication, ainsi que des règles de combinaison en chaîne et de combinaison de confirmations. Ces développements nous ont permis de créer une nouvelle méthode pour modéliser non seulement des informations incertaines et imprécises, mais également leur remise en question. En effet, cette nouvelle approche rend possible la représentation fidèle et précise, ainsi que l'évaluation de schémas plus complexes, mais néanmoins inhérents au monde qui nous entoure. Ces schémas sont rencontrés dans toutes les disciplines scientifiques, car ils sont tout à fait appropriés pour la modélisation de la remise en question des hypothèses de base d'une théorie. Nous avons pu constater cela dans la Partie 4, lorsque les hypothèses sur la distribution des rendements des actifs, aussi bien dans le cadre de la Théorie Moderne du Portefeuille que dans celui de la Théorie Post-Moderne du Portefeuille ont été remises en question et évaluées. D'autres applications des concepts développés dans la Partie 3 restent à être explorés. Par exemple, leur emploi serait parfaitement adapté à des modèles de fiabilité ou à n'importe quel modèle basé sur l'hypothèse d'une certaine distribution.

Vers une nouvelle mesure de performance

La Partie 4 expose brièvement la Théorie Moderne du Portefeuille (TMP), la Théorie Post-Moderne du Portefeuille (TPMP) et les mesures de performance qui leur sont associées. Sur cette base, ainsi que sur celle des résultats de notre recherche exposés dans la Partie 3, nous avons présenté un modèle d'évaluation et de classement d'actifs financiers du plus performant au moins performant, appelé Modèle de performance et incertitude. Contrairement aux mesures de performance classiques, présentées dans la Partie 4, le modèle que nous avons développé propose deux indices de performance indépendants pour chaque

actif. La performance de chaque actif peut ainsi être évalué indépendamment de celle des autres actifs. Par ailleurs, ce modèle introduit un nouveau concept de risque, le risque d'incertitude, complémentaire et non pas substitutif aux concepts de risque déjà présents dans ces deux théories financières. Ce risque évalue dans quelle mesure les rendements d'un actif ne remplissent pas les hypothèses sur leur distribution à la base de la TMP et la TPMP. Ainsi, selon cette évaluation, les mesures de performance proposées par ces deux théories financières seront plus ou moins fiables. Par exemple, pour un actif dont le risque d'incertitude est nul, nous pouvons nous y fier complètement. A l'inverse, pour un actif dont le risque d'incertitude est maximal, c'est-à-dire de 1, ces mesures perdent toute leur pertinence. Ainsi, pour cet actif, nous nous retrouvons dans une situation d'incertitude totale et pouvons uniquement affirmer que sa place dans un classement de n actifs sera comprise entre la première et la n -ème. Autrement dit, le seul sous-ensemble de l'ensemble de réponses pour lequel nous pouvons prouver qu'il contient la réponse correcte n'est autre que l'ensemble de réponses lui-même. Peu importe que les mesures de performance de la TMP ou de la TPMP le classent comme performant ou non ; s'y fier serait incorrect tant que leur hypothèse de base n'est pas respectée. Dans notre conception, parallèlement aux risques définis en termes de variabilité, cette incertitude représente également une source de risque importante, c'est pourquoi elle contribue à pénaliser les actifs chez lesquels elle est présente. Le modèle de performance et incertitude a été finalement illustré par une application à 10 actions suisses cotée en bourse.

Recherches futures

De nombreuses pistes de recherche restent encore à être explorées. Les développements proposés dans ce travail peuvent être appliqués à de nombreux problèmes économiques ou encore à des problèmes relatifs à d'autres disciplines. Par exemple, on pourrait modéliser des problèmes typiquement microéconomiques de grande incertitude, tels que la fonction d'utilité et la remise en question de ses hypothèses de base. En finance, il serait possible d'introduire le risque d'incertitude dans plusieurs modèles déjà existants, par exemple les modèles de risque de crédit. Par ailleurs, d'un point de vue purement théorique, un approfondissement des concepts de combinaison en chaîne, d'indication de confirmation et de meta-indication serait souhaitable dans le cadre d'une recherche future. En effet, comme disait EDGAR MORIN :

«La conscience n'est jamais assurée de surmonter l'ambiguïté et l'incertitude.»

Extrait de *Le paradigme perdu*

Bibliographie

- [1] B. Anrig, R. Bissig, R. Haenni, J. Kohlas, and N. Lehmann. Probabilistic argumentation systems : Introduction to assumption-based modeling with abel. Technical Report 99-1, Institute of Informatics, University of Fribourg, 1999.
- [2] B. Anrig, R. Haenni, J. Kohlas, and N. Lehmann. Assumption-based modeling using abel. In D. Gabbay, R. Kruse, A. Nonnengart, and H.J. Ohlbach, editors, *First International Joint Conference on Qualitative and Quantitative Practical Reasoning; ECSQARU-FAPR'97*. Springer, 1997.
- [3] B. Anrig, R. Haenni, J. Kohlas, and P.-A. Monney. Probabilistic analysis of model-based diagnosis. In *IPMU'96, Proceedings of the sixth international conference, Granada, Spain*, pages 123–128, 1996.
- [4] B. Anrig, R. Haenni, and N. Lehmann. Abel - a new language for assumption-based evidential reasoning under uncertainty. Technical Report 97-01, Institute of Informatics, University of Fribourg, 1997.
- [5] B. Anrig, N. Lehmann, and R. Haenni. Reasoning with finite set constraints. Technical Report 97-11, Institute of Informatics, University of Fribourg, 1997.
- [6] A. Appriou. Probabilités et incertitude en fusion de données multi-senseurs. *Revue Scientifique et Technique de la Défense*, 11 :27–40, 1991.
- [7] A. Appriou. Formulation et traitement de l'incertain en analyse multi-senseurs. In *14ème colloque GRETSI*, Septembre 1993.
- [8] V. Bawa. Optimal rules for ordering uncertain prospects. *Journal of Financial Economics*, 2 :1169–1183, 1975.
- [9] V. Bawa, S. Brown, and R. Klein. *Asymmetric Response Asset Pricing Models : Testable Alternatives to Mean-Variance*. Mimeo, 1981.
- [10] V. Bawa and E. B. Lindenberg. Capital market equilibrium in a mean, lower partial moment framework. *Journal of Financial Economics*, 5 :189–200, 1977.
- [11] P. Besnard and J. Kohlas. Evidence theory based on general consequence relations. *Int. J. of Foundations of Computer Science*, 6(2) :119–135, 1995.
- [12] H. W. Brachinger and P.-A. Monney. The conjunction fallacy : explanations of the linda problem by the theory of hints. *International Journal of Intelligent Systems*, 18 :75–91, 2003.

- [13] H. W. Brachinger and M. Weber. Risk as a primitive : a survey of measures of perceived risk. *Research - Spektrum*, 19/3 :235–250, 1997.
- [14] M. Le Brenton and P. Ghirardato. Choquet rationality. Technical Report 1000, Division of Humanities and social sciences, California institute of technology, 1999.
- [15] Charest, Lusztig, and Schwab. *Gestion financière*. Editions du Renouveau Pédagogique Inc., 1990.
- [16] V. Chernozhukov and S. Morozov. Econometric evaluation of rational belief model. forthcoming, 2000.
- [17] G. Choquet. Theory of capacities. *Ann. Inst. Fourier (Grenoble)*, 5 :131–295, 1953.
- [18] G. Choquet. *Lectures on Analysis*. Benjamin, New York, 1969.
- [19] A. David and P. Veronesi. Option prices with uncertain fundamentals : Theory and evidence on the dynamics of implied volatilities. Technical Report 485, Graduate School of Business, University of Chicago, 2000.
- [20] J. de Kleer. An assumption-based tms. *Journal of Artificial Intelligence*, 28 :127–162, 1986.
- [21] A. P. Dempster. Upper and lower probabilities induced by a multivaluated mapping. *Annals of Mathematical Statistics*, 38 :325–339, 1967.
- [22] J. Dezert. *Vers un nouveau concept de navigation autonome d’engin. Un lien entre le filtrage à association probabiliste de données et la théorie de l’évidence*. PhD thesis, Université de Paris XI, 1990.
- [23] D. Dubois and H. Prade. Combination and propagation of uncertainty with belief functions - a reexamination. In *9th International Joint Conference on Artificial Intelligence*, pages 111–113. IJCAI, 1985.
- [24] D. Dubois and H. Prade. On the relevance of non-standard theories of uncertainty in modeling and pooling expert opinions. *Reliability Engineering and System Safety*, 36, 1992.
- [25] H. J. Einhorn and R. M. Hogarth. Decision making under ambiguity. *The Journal of Business*, 59(4) :225–250, 1986.
- [26] D. Ellsberg. Risk, ambiguity, and the savage axioms. *Quarterly Journal of Economics*, 75 :643–669, 1961.
- [27] E. Elton and M. Gruber. *Modern Portfolio Theory and Investment Analysis*. John Wiley and Sons, Inc, fifth edition, 1995.
- [28] P. Fabiabi. *Représentation dynamique de l’incertain et stratégie de prise d’information pour un système autonome en environnement évolutif*. PhD thesis, Ecole nationale supérieure de l’aéronautique et de l’espace, France, 1996.
- [29] E. Fama. Portfolio analysis in a stable paretian market. *Management Science*, 11 :404–419, 1965.

- [30] P. C. Fishburn. Mean-risk analysis with risk associated with below target returns. *American Economic Review*, 67 :61–69, 1977.
- [31] P. Ghirardato. Coping with ignorance : unforeseen contingencies and non-additive uncertainty. *Economic Theory*, 17 :247–276, 2001.
- [32] O. Guedhami, J. Roy, and M.-Ch. To. La répartition de l'actif dans l'approche du risque baissier. Technical report, Groupe de recherche en finance, Ecole des Hautes Etudes Commerciales, Montréal, Québec, 1999.
- [33] R. Haenni. A valuation-based architecture for assumption-based reasoning. In G. Coletti, D. Dubois, and R. Scozzafava, editors, *Mathematical Models for Handling Partial Knowledge in Artificial Intelligence*, pages 255–258. Plenum Press, 1995.
- [34] R. Haenni. *Propositional Argumentation Systems and Symbolic Evidence Theory*. PhD thesis, Institut für Informatik, Universität Freiburg, 1996.
- [35] R. Haenni. Assumption-based reasoning with algebraic clauses. In H.J. Zimmermann, editor, *EUFIT'97, 5th European Congress on Intelligent Techniques and Soft Computing*, pages 958–961. Verlag Mainz, 1997.
- [36] R. Haenni. Modeling uncertainty with propositional assumption-based systems. In S. Parson and A. Hunter, editors, *Applications of Uncertainty Formalisms*, Lecture Notes in Artificial Intelligence 1455, pages 446–470. Springer-Verlag, 1998.
- [37] R. Haenni and N. Lehmann. Assumption-based reasoning with finite set constraints. In *IPMU'98, Proceedings of the seventh international conference, Paris, France*, pages 1289–1295, 1998.
- [38] W. V. Harlow and R.K.S. Rao. Asset pricing in a generalized mean-lower partial moment framework. *Journal of Financial and Quantitative Analysis*, 24 :285–311, 1989.
- [39] D. Harmanec. Measures of uncertainty and information. In Fabio G. Cozman Gert de Cooman, Peter Walley and Marco Zaffalon, editors, *The Imprecise Probabilities Project*. <http://ippserv.rug.ac.be/home/ipp.html>, 1996.
- [40] K. Harrison, R. P. Srivastava, and R. D. Plumlee. Auditors' evaluations of uncertain audit evidence : Belief functions versus probabilities. In Ted Mock and Rajendra Srivastava, editors, *Belief Functions in Business Decisions*. Physica-Verlag, A Springer-Verlag Company, Heidelberg : New York, 2002.
- [41] R. Henriksson. Market timing and mutual fund performance : An empirical investigation. *Journal of Business*, 57 :73–96, 1984.
- [42] R. Henriksson and R. Merton. On market timing and investment performance. ii statistical procedures for evaluating forecasting skills. *Journal of Business*, 54 :513–533, 1981.
- [43] K. Inoue. An abductive procedure for the cms/atms. In J.P.Martins and M. Reinfrank, editors, *Truth Maintenance Systems, Lectures Notes in A.I.*, pages 34–53. Springer, 1991.

- [44] M. Jensen. The performance of mutual funds in the period 1945-1964. *Journal of Finance*, 23 :389–416, 1968.
- [45] M. Jensen. Risk, the pricing of capital assets, and the evaluation of investment portfolios. *Journal of Business*, pages 167–247, 1969.
- [46] R. Jirousek, J. Vejnarova, and J. Gemela. Possibilistic belief network constructed by operators of composition and its application to financial analysis. In Ted Mock and Rajendra Srivastava, editors, *Belief Functions in Business Decisions*. Physica-Verlag, A Springer-Verlag Company, Heidelberg : New York, 2002.
- [47] M. K. P. Jung and H. E. Fink. The effectiveness and efficiency of belief based audit procedures. In Ted Mock and Rajendra Srivastava, editors, *Belief Functions in Business Decisions*. Physica-Verlag, A Springer-Verlag Company, Heidelberg : New York, 2002.
- [48] L. N. Kanal and J. F. Lemmer. Uncertainty in artificial intelligence. In *Machine Intelligence and Pattern Recognition*. Elsevier/North-Holland, 1986.
- [49] B. Kemmerer, S. Mishra, and P. P. Shenoy. Managing venture capital investment decisions : A knowledge-based approach. Technical report, School of Business, University of Kansas, 2001.
- [50] F. Klawonn and P. Smets. The dynamic of belief in the transferable belief model and specialization-generalization matrices. In D’Ambrosio B. Dubois D., Wellman M.P. and Smets Ph., editors, *Proceedings of the 8th Conference on Uncertainty in Artificial Intelligence*, pages 130–137. Morgan Kaufmann, 1992.
- [51] F. H. Knight. *Risk, Uncertainty, and Profit*. Boston, MA : Hart, Schaffner and Marx; Houghton Mifflin Company, 1921.
- [52] J. Kohlas. *Zuverlässigkeit und Verfügbarkeit*. Teubner, Stuttgart, 1987.
- [53] J. Kohlas. The reliability of reasoning with unreliable arguments. Technical Report 168, Institute for Automation and Operational Research, University of Fribourg, 1989.
- [54] J. Kohlas. A mathematical theory of hints. Technical Report 173, Institute for Automation and Operations Research, University of Fribourg, 1990.
- [55] J. Kohlas. Mathematical foundations of evidence theory. In G. Coletti, D. Dubois, and R. Scozzafava, editors, *Mathematical Models for Handling Partial Knowledge in Artificial Intelligence*, pages 31–64. Plenum Press, 1995.
- [56] J. Kohlas. Allocation of arguments and evidence theory. *Theoretical Computer Science*, 171 :221–246, 1997.
- [57] J. Kohlas and P. Besnard. An algebraic study of argumentation systems and evidence theory. Technical Report 95–13, Institute of Informatics, University of Fribourg, 1995.

- [58] J. Kohlas and R. Haenni. Assumption-based reasoning and probabilistic argumentation systems. Technical Report 96-07, Institute of Informatics, University of Fribourg, 1996. also published in "Defeasible Reasoning and Uncertainty Management Systems : Algorithms".
- [59] J. Kohlas and P.-A. Monney. Modeling and reasoning with hints. Technical Report 174, Institute for Automation and Operations Research, University of Fribourg, 1990.
- [60] J. Kohlas and P.-A. Monney. Probabilistic assumption-based reasoning. In D. Heckermann and A. Mamdani, editors, *Proc. 9th Conf. on Uncertainty in Artificial Intelligence, Washington*, pages 485-491. Kaufmann, Morgan Publ., 1993. Short Version of Tech. Rep. 94-22.
- [61] J. Kohlas and P.-A. Monney. Representation of evidence by hints. In R.R. Yager, J. Kacprzyk, and M. Fedrizzi, editors, *Advances in the Dempster-Shafer Theory of Evidence*, pages 473-492. John Wiley, 1994.
- [62] J. Kohlas and P.-A. Monney. *A Mathematical Theory of Hints. An Approach to Dempster-Shafer Theory of Evidence*, volume 425 of *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag, 1995.
- [63] J. Kohlas, P.-A. Monney, B. Anrig, and R. Haenni. Model-based diagnostics and probabilistic assumption-based reasoning. Technical Report 96-09, Institute of Informatics, University of Fribourg, 1996.
- [64] A. Kolmogoroff. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Springer, Berlin., 1933.
- [65] H. E. Kyburg. Bayesian and non-bayesian evidential updating. *Artificial Intelligence*, 31 :271-293, 1987.
- [66] K. B. Laskey and P. E. Lehner. Assumptions, beliefs and probabilities. *Artificial Intelligence*, 41 :65-77, 1989.
- [67] W. Y Lee and R. K. S. Rao. Mean lower partial moment valuation and lognormally distributed returns. *Management Science*, 34 :446-453, 1988.
- [68] N. Lehmann. *Argumentation Systems and Belief Functions*. PhD thesis, University of Fribourg, 2001.
- [69] J. F. Lemmer. Confidence factor, empiricism and the dempster-shafer theory of evidence. In *Machine Intelligence and Pattern Recognition*, chapter II, pages 117-125. Elsevier/North-Holland, 1986.
- [70] C. Lesage. Evaluation du risque d'audit : proposition d'un modèle linguistique. Technical Report 9713, CEREG, Université de Rennes 1, 1997.
- [71] D. Lien. Futures hedging under prospect utility and knightian uncertainty. In Ted Mock and Rajendra Srivastava, editors, *Belief Functions in Business Decisions*. Physica-Verlag, A Springer-Verlag Company, Heidelberg : New York, 2002.

- [72] J. Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics*, 47 :13–37, 1965.
- [73] H. M. Markowitz. Portfolio selection. *Journal of Finance*, 7 :77–91, 1952.
- [74] H. M. Markowitz. The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly*, pages 111–133, 1956.
- [75] H. M. Markowitz. *Portfolio selection : Efficient diversification of investments*. Wiley, Yale University Press, 1959.
- [76] H. M. Markowitz. *Mean Variance Analysis in Portfolio Choice and Capital Markets*. Blackwell, 1987.
- [77] P. McBurney and S. Parsons. Using belief functions to forecast demand for mobile satellite services. In Ted Mock and Rajendra Srivastava, editors, *Belief Functions in Business Decisions*. Physica-Verlag, A Springer-Verlag Company, Heidelberg : New York, 2002.
- [78] P.-A. Monney. *Le raisonnement temporel et géométrique sous incertitude*. PhD thesis, Université de Fribourg, 1991.
- [79] G. Monroe and J. Ng. The descriptive ability of models of audit risk. In Ted Mock and Rajendra Srivastava, editors, *Belief Functions in Business Decisions*. Physica-Verlag, A Springer-Verlag Company, Heidelberg : New York, 2002.
- [80] J. Mossin. Security pricing and investment criteria in competitive markets. *American Economic Review*, 59 :749–756, 1969.
- [81] D. Nawrocki. A brief history of downside risk measures. *Journal of Investing*, 8 :9–25, 1999.
- [82] D. Nawrocki. The case for the relevancy of downside risk measures. Technical report, College of Commerce and Finance, Villanova University, 1999.
- [83] K. G. Nishimura and H. Ozaki. A note on learning under the knightian uncertainty. Technical Report CIRJE-F-113, Faculty of Economics, University of Tokyo, 2001.
- [84] A. Patel. *Auditors' Belief Revision : Recency Effects of Contrary and Supporting Audit Evidence and Source Reliability*. PhD thesis, University of Queensland, Brisbane, Australia, 2001.
- [85] J. Pearl. *Probabilistic Reasoning in Intelligent Systems : Networks of Plausible Inference*. Morgan Kaufman, San Mateo, Cal., 1988.
- [86] C. Pedersen and S. Hwang. Does downside beta matter in asset pricing? Technical report, Cass business school, 2003.
- [87] C. S. Pedersen and T. Rudholm-Alvin. Selecting a risk-adjusted shareholder performance measure. Technical report, Judge Institute Seminar, Cambridge, 2003.
- [88] C. S. Pedersen and S. E. Satchell. On the foundation of performance measures under asymmetric returns. *Journal of Quantitative Finance*, 2 :217–223, June 2002.

- [89] C. Perneel, H. Van De Velde, and M. Acheroy. A heuristic search algorithm based on belief functions. In *14th International Avignon Conference*, pages 99–107. AFIA - AI'94, EC2, 1994.
- [90] D. M. Peterson. *Financial Ratios abs Investment Results*. Lexington Books, 1974.
- [91] A. Plantinga and S. de Groot. Risk-adjusted performance measures and implied risk attitudes. *The Journal of Performance Measurement*, 6 :9–22, 2002.
- [92] R. Reiter and J. de Kleer. Foundations of assumption-based truth maintenance systems. In *Proceedings of the American Association in AI*, pages 183–188, 1987.
- [93] B. M. Rom and K. Ferguson. Post-modern portfolio theory comes of age. *The Journal of Investing*, pages 11–17, 1993.
- [94] L. J. Savage. *The Foudations of Statistics*. John Wiley, New York, 1954.
- [95] G. Shafer. *A Mathematical Theory of Evidence*. Prinston University Press, 1976.
- [96] G. Shafer. Dempster's rule of combination. Unpublished Manuscript, 1978.
- [97] G. Shafer. Allocations of probability. *The Annals of Probability*, 7 :827–839, 1979.
- [98] G. Shafer. Non-additive probabilities in the work of bernoulli and lambert. *Archive for Hitory of Exact Science*, 19 :309–370, 1979.
- [99] G. Shafer. Constructive probability. *Synthese*, 48 :827–839, 1981.
- [100] W. F. Sharpe. Capital asset prices : A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19 :425–442, September 1964.
- [101] W. F. Sharpe. Mutual fund performance. *Journal of Business*, pages 119–138, 1966.
- [102] P. P. Shenoy. Valuation network representation and solution of asymmetric decision problems. *European Journal of Operational Research*, 121(3) :579–608, 2000.
- [103] P. P. Shenoy and C. Shenoy. Bayesian network models of portfolio risk and return. In A. W. Lo Y. S. Abu-Mostafa, B. LeBaron and A. S. Weigend, editors, *Computational Finance 1999*. MIT Press, 2000.
- [104] P. P. Shenoy and C. Shenoy. Modeling financial portfolios using belief functions. In R. P. Srivastava and T. J. Mock, editors, *Belief Functions in Business Decisions*, pages 316–332. Physica-Verlag, 2002.
- [105] J. Singleton and J. Wingender. Skewness persistence in common stock returns. *Journal of Financial and Quantitative Analysis*, 13 :335–341, 1986.
- [106] P. Smets. *Belief Functions*. Academic Press, London, 1988.
- [107] P. Smets. The combination of evidence in the transferable belief model. *IEEE Pattern Analysis and Machine Intelligence*, 12 :447–458, 1990.
- [108] P. Smets. The transferable belief model for quantified belief representation. In Gabbay D. and Smets Ph., editors, *Handbook of Defeasible Reasoning and Uncertainty Management Systems*, pages 267–301. Kluwer, Doordrecht, 1998.

- [109] F. Sortino and L. Price. Performance measurement in a downside risk framework. *The Journal of Investing*, 3 :59–64, 1994.
- [110] F. A. Sortino and R. Van Der Meer. Downside risk. *Journal of Portfolio Management*, 17 :27–32, 1991.
- [111] R. P. Srivastava and D. Datta. Evaluating mergers and acquisitions : A belief function approach. In Ted Mock and Rajendra Srivastava, editors, *Belief Functions in Business Decisions*. Physica-Verlag, A Springer-Verlag Company, Heidelberg : New York, 2002.
- [112] R. P. Srivastava and L. Liu. Applications of belief functions in business decisions : A review. *Information Systems Frontiers*, 5 :359–378, 2003.
- [113] R. P. Srivastava and H. Lu. Structural analysis of audit evidence using belief functions. *Fuzzy Set and Systems*, 131 :107–120, 2002.
- [114] R. P. Srivastava and T. J. Mock. Belief functions in accounting behavioral research. *Advances in Accounting Behavioral Research*, 3 :225–242, 2000.
- [115] B. K. Stone. *Risk, Return and Equilibrium*. MIT Press, 1970.
- [116] B. K. Stone. A general class of three-parameter risk measures. *Journal of Finance*, 23 :675–685, 1973.
- [117] J. Treynor. How to rate management of investment funds. *Harvard Business Review*, 43(1) :63–75, 1965.
- [118] Y.-P. Tseng and M. Wooden. Enterprise bargaining and productivity : Evidence from the business longitudinal survey. Technical Report 8/01, Melbourne Institute of Applied Economic and Social Research, The University of Melbourne, 2001.
- [119] F. Voorbraak. On the justification of dempster’s rule of combination. *Artificial Intelligence*, 48 :171–197, 1991.
- [120] T. Wang. Updating rules for non-bayesian preferences. Econometric Society World Congress 2000 Contributed Papers, 2000.
- [121] S. T. Wierzchon and M. A. Klotek. Empirical models for the dempster-shafer theory. In Ted Mock and Rajendra Srivastava, editors, *Belief Functions in Business Decisions*. Physica-Verlag, A Springer-Verlag Company, Heidelberg : New York, 2002.
- [122] J. S. Wu, G.E. Apostakis, and D. Okrent. Uncertainties in system analysis : probabilistic versus nonprobabilistic theories. *Reliability Engineering and Systems Safety*, 30 :163–181, 1990.
- [123] L. A. Zadeh. On the validity of demspster’s rule of combination of evidence. In *Dept. of Electrical Engineering and Computer Sciences, University of California, Berkeley*, 1979.
- [124] S. Zancanella. Les fonds de placement : Définitions et mesures de la performance des fonds de placement sur le marché suisse. Technical Report 200.09, Université de Genève, 2000.