

---

---

Décomposition  
de problèmes d'optimisation stochastique  
avec mesure de risque  
sur un ensemble fini de scénarios

---

---

RAPPORT DE STAGE DE FIN D'ÉTUDES  
MASTER PARISIEN DE RECHERCHE OPÉRATIONNELLE  
TÉLÉCOM PARISTECH

HENRI GERARD

SOUS LA DIRECTION DE MICHEL DE LARA

27 JANVIER 2016

## Résumé

Dans ce rapport, nous examinons des problèmes d'optimisation stochastique avec une contrainte de mesurabilité. Nous étudions comment des méthodes de résolution de tels problèmes s'étendent du cas risque neutre au cas sensible au risque. Nous nous intéressons successivement à des problèmes statiques puis à deux pas de temps. Les deux types de méthodes de résolution étudiées sont une méthode par imbrication du type programmation dynamique et une méthode par dualisation. Des expériences numériques sont également menées.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Exemple du vendeur de journaux</b>	<b>2</b>
2.1	Introduction . . . . .	2
2.2	Vendeur de journaux à 1 période . . . . .	2
2.3	Vendeur de journaux à 2 périodes . . . . .	12
2.4	Conclusion . . . . .	14
<b>3</b>	<b>Optimisation à 1 période avec contraintes de mesurabilité</b>	<b>18</b>
3.1	Introduction . . . . .	18
3.2	Formulation d'un problème d'optimisation stochastique sous contraintes d'information . . . . .	18
3.3	Résolution par imbrication . . . . .	20
3.4	Résolution par dualisation des contraintes . . . . .	24
3.5	Conclusion . . . . .	28
<b>4</b>	<b>Optimisation à 2 périodes avec contraintes de mesurabilité</b>	<b>29</b>
4.1	Introduction . . . . .	29
4.2	Résolution par imbrication . . . . .	29
4.3	Résolution par dualisation des contraintes . . . . .	32
4.4	Conclusion . . . . .	34
<b>5</b>	<b>Perspectives d'application à un problème multi-vendeur</b>	<b>35</b>
5.1	Introduction . . . . .	35
5.2	Modélisation . . . . .	35
5.3	Perspectives de résolution . . . . .	36
5.4	Conclusion . . . . .	37
<b>6</b>	<b>Conclusion</b>	<b>38</b>
<b>A</b>	<b>Appendices</b>	<b>39</b>
A.1	Rappels sur les mesures de risque . . . . .	39
A.2	Rappels d'analyse convexe . . . . .	45
	<b>Bibliographie</b>	<b>47</b>

# Remerciements

Je tiens à remercier toutes les personnes qui m'ont guidé et apporté leur aide pendant la durée de ce stage.

Je remercie tout d'abord le groupe de travail d'optimisation stochastique qui m'a encadré : Michel De Lara mon tuteur et futur directeur de thèse, Pierre Carpentier, Jean-Philippe Chancelier et Vincent Leclère pour leurs conseils avisés.

Je pense aussi à Isabelle Simunic pour toute l'aide qu'elle m'a fourni notamment dans les démarches administratives.

J'adresse également mes remerciements à Olivier Hudry pour m'avoir donné goût à l'optimisation et m'avoir facilité les démarches d'admission au MPRO.

Enfin je salue mon binôme de Master Stéphane Bonanno à qui je souhaite "bonne continuation" dans le milieu entrepreneurial. Je salue tout autant mes co-stagiaires Houzhi Li, Ricardo Rincón et Valentin Foucher à qui je souhaite "bon courage" pour son master de recherche.

# Symboles et abréviations

$\mathbb{N}$	Ensemble des entiers naturels
$\mathbb{R}$	Ensemble des nombres réels
$\overline{\mathbb{R}}$	$\mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$
$\mathbb{E}$	Espérance mathématique
$\mathbb{F}$	Mesure de risque
$\mathbb{P}$	Mesure de probabilité
$\langle \cdot, \cdot \rangle$	Produit scalaire
s.c.i.	Semi-continue inférieurement
s.c.s.	Semi-continue supérieurement
$\mathbb{P}$ -p.s.	Presque sûrement relativement à la mesure de probabilité $\mathbb{P}$

# 1 Introduction

Nous examinons comment des méthodes de résolution de problèmes d'optimisation stochastique s'étendent du cas risque neutre au cas sensible au risque.

Dans le Chapitre 2, nous étudions un problème jouet appelé *problème du vendeur de journaux* qui nous permet d'introduire différents concepts et formalismes mathématiques. Des expériences numériques sont également menées afin d'illustrer les résultats théoriques.

Dans le Chapitre 3, nous étudions des problèmes d'optimisation stochastique avec une contrainte de mesurabilité dans le cas statique. Dans un premier temps nous étudions une méthode de résolution par imbrication que nous généralisons du cas risque neutre au cas sensible au risque. Dans un second temps, nous dualisons la contrainte de mesurabilité et nous étendons la résolution du cas risque neutre au cas sensible au risque.

Dans le Chapitre 4, nous étudions des problèmes dynamiques et nous regardons comment les résultats obtenus changent dans le cas statique.

Dans le Chapitre 5, nous présentons une extension du problème du Chapitre 2 à un problème multi-vendeurs.

## 2 Exemple du vendeur de journaux

### 2.1 Introduction

Mettons nous à la place d'un vendeur de journaux. Chaque jour, nous devons vendre des quotidiens à une clientèle. Nous achetons donc des journaux à un certain prix et nous les revendons plus chers. Notre objectif est de maximiser notre profit. Cela nous amène à nous poser la question suivante : combien de journaux faut-il acheter en début de journée ?

En effet, plus on achète de périodiques, plus il est possible de faire du profit mais plus le risque est grand d'avoir des invendus. Cette question va nous permettre au travers de ce chapitre d'introduire les concepts mathématiques dont nous aurons besoin dans les chapitres suivants.

$$x = 2 \tag{2.1}$$

### 2.2 Vendeur de journaux à 1 période

On désire vendre des quotidiens. Le profit que l'on cherche à maximiser peut se représenter par la fonction critère

$$j(u, d) = -cu + p \min\{u, d\} , \tag{2.2}$$

où l'on distingue les éléments suivants :

- les paramètres  $c$ , coût d'achat d'un journal et  $p$ , prix de vente d'un journal avec  $c < p$ ,
- le nombre de journaux commandés  $u$  (variable de contrôle ou de décision),
- la demande  $d$ .

L'équation (2.2) définit une fonction déterministe qui, à une demande  $d$  et un nombre  $u$  de journaux commandés, associe le profit

$$j(u, d) = \begin{cases} (p - c)u , & \text{si } u \leq d \\ -cu + pd , & \text{si } u > d . \end{cases} \tag{2.3}$$

À  $d$  fixé, comme  $(p - c) > 0$  et  $-c < 0$ ,  $u \mapsto j(., d)$  est une fonction affine par morceaux dont le maximum est atteint en

$$u^\# = d . \tag{2.4}$$

## 2.2.1 Formulation du problème d'optimisation stochastique

### Aléas et variable aléatoire

En pratique, la demande  $d$  n'est pas connue et différents scénarios peuvent se réaliser. On se donne alors un ensemble  $\Omega$  fini de scénarios. Les variables aléatoires réelles  $\mathbf{X}$  sur  $\Omega$  sont ainsi les éléments de  $\mathbb{R}^\Omega$ , c'est-à-dire des fonctions qui à un scénario  $\omega$  associent une valeur réelle  $\mathbf{X}(\omega)$ .

### La demande

On représente la demande par une variable aléatoire entière  $\mathbf{D}$ . Cela signifie que  $\mathbf{D}$  est un élément de  $\mathbb{N}^\Omega$ .

### Le contrôle

Le nombre de journaux commandés est appelé contrôle car c'est une décision prise par le marchand. On le note  $u \in \mathbb{N}$ . Si ce contrôle dépend de  $\omega$ , c'est une variable aléatoire. On parle alors de contrôle stochastique et on le note  $\mathbf{U}$ .

### La fonction profit

Le profit est représenté par la fonction  $j : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}$  définie en (2.2) qui à un contrôle  $u$  et à une demande  $d$  associe un profit  $j(u, d)$ .

### Le problème d'optimisation

On s'intéresse à un problème d'optimisation stochastique où la demande  $\mathbf{D}$  est aléatoire et le contrôle  $\mathbf{U}$  est stochastique. On se donne ainsi une loi  $\mathbb{P}$  qui va nous permettre de calculer l'espérance de la fonction critère. Cette loi est caractérisée par la donnée des  $\mathbb{P}(\{\omega\})_{\omega \in \Omega}$  car  $\Omega$  est fini. Dans la suite on notera pour plus de simplicité

$$\mathbb{P}(\{\omega\}) = \mathbb{P}(\omega) . \quad (2.5)$$

Le problème d'optimisation s'énonce

$$\max_{\mathbf{U} \in \mathbb{U}^\Omega} \mathbb{E}_{\mathbb{P}}(j(\mathbf{U}, \mathbf{D})) . \quad (2.6)$$

En remarquant que

$$\max \text{ profit} = - \min -\text{profit} = - \min \text{ coût} , \quad (2.7)$$

on peut aussi énoncer le problème suivant

$$\min_{\mathbf{U} \in \mathbb{U}^\Omega} \mathbb{E}_{\mathbb{P}}(-j(\mathbf{U}, \mathbf{D})) . \quad (2.8)$$

C'est la formulation (2.8) que l'on va étudier.

Dans le problème (2.8),  $j(\mathbf{U}, \mathbf{D})$  représente la variable aléatoire

$$\begin{aligned} j(\mathbf{U}, \mathbf{D}) : \Omega &\rightarrow \mathbb{R} \\ \omega &\rightarrow j(\mathbf{U}(\omega), \mathbf{D}(\omega)) . \end{aligned} \quad (2.9)$$

On va à présent s'intéresser à 2 cas extrêmes. Dans un premier temps, on connaît la demande quand on prend la décision  $\mathbf{U}$  et  $\mathbf{U}$  dépend de  $\omega$ . Dans un second temps, on ne connaît pas la demande quand on prend la décision  $\mathbf{U}$  et  $\mathbf{U}$  ne dépend pas de  $\omega$ .

### 2.2.2 Formulation hasard-décision

Dans cette formulation, le contrôle dépend de la réalisation de l'aléa et est noté  $\mathbf{U}$ . Le problème est

$$\min_{\mathbf{U} \in \mathbb{U}^\Omega} \mathbb{E}_{\mathbb{P}}(-j(\mathbf{U}, \mathbf{D})) . \quad (2.10)$$

Ce problème est équivalent à

$$\mathbb{E}_{\mathbb{P}}\left(\min_{u \in \mathbb{U}} -j(u, \mathbf{D})\right) , \quad (2.11)$$

et peut se résoudre  $\omega$  par  $\omega$ . D'après (2.4), une solution optimale est donnée par

$$\mathbf{U}(\omega) = \mathbf{D}(\omega) , \quad \forall \omega \in \Omega . \quad (2.12)$$

### 2.2.3 Formulation décision-hasard

Ici, on prend une décision puis le hasard se produit. Le contrôle ne doit pas dépendre de  $\omega$  et on le note donc  $u \in \mathbb{U}$  (ce qui est différent de  $\mathbf{U} \in \mathbb{U}^\Omega$ ). La demande reste elle aléatoire, dépend donc de  $\omega$  et est notée  $\mathbf{D}$ . On s'intéresse alors au problème d'optimisation stochastique suivant

$$\min_{u \in \mathbb{U}} \mathbb{E}_{\mathbb{P}}\left(-j(u, \mathbf{D})\right) . \quad (2.13)$$

*Remarque 2.1.* On aurait pu écrire le problème (2.13) avec un contrôle stochastique  $\mathbf{U}$  à condition d'ajouter les contraintes

$$\mathbf{U}(\omega) = \mathbf{U}(\omega') , \quad \forall \omega \in \Omega , \quad (2.14)$$

ou les contraintes équivalentes

$$\mathbf{U}(\omega) = \mathbb{E}_{\mathbb{P}}(\mathbf{U}) , \quad \forall \omega \in \Omega . \quad (2.15)$$

Les contraintes (2.14) et (2.15) sont en nombre fini car  $\Omega$  est fini.

**Proposition 2.2.** *Le contrôle optimal  $u_{\mathbb{E}}^{\sharp} \in \mathbb{N}$  du problème (2.13) est caractérisé par*

$$\mathbb{P}\left(\mathbf{D} \leq u_{\mathbb{E}}^{\sharp} - 1\right) \leq \frac{p-c}{p} \leq \mathbb{P}\left(\mathbf{D} \leq u_{\mathbb{E}}^{\sharp}\right) . \quad (2.16)$$

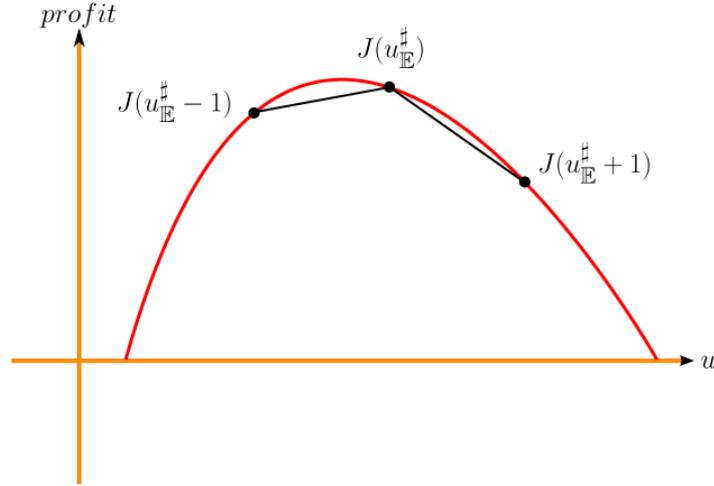


FIGURE 2.1 – Représentation des inégalités vérifiées par la décision optimale  $u^\# \in \mathbb{N}$

*Démonstration.* On définit la fonction  $J(u) = \mathbb{E}_{\mathbb{P}}(j(u, \mathbf{D}))$  et on étudie ses variations. On a :

$$J(u) = -cu + p\mathbb{E}_{\mathbb{P}}(\min\{u, \mathbf{D}\}) , \quad (2.17a)$$

$$\min\{u, \mathbf{D}\} = u\mathbf{1}_{u < \mathbf{D}} + \mathbf{D}\mathbf{1}_{u \geq \mathbf{D}} , \quad (2.17b)$$

$$\min\{u + 1, \mathbf{D}\} = (u + 1)\mathbf{1}_{u+1 < \mathbf{D}} + \mathbf{D}\mathbf{1}_{u+1 \geq \mathbf{D}} , \quad (2.17c)$$

$$\min\{u + 1, \mathbf{D}\} = (u + 1)\mathbf{1}_{u < \mathbf{D}} + \mathbf{D}\mathbf{1}_{u \geq \mathbf{D}} , \quad (2.17d)$$

$$\min\{u + 1, \mathbf{D}\} - \min\{u, \mathbf{D}\} = \mathbf{1}_{u < \mathbf{D}} . \quad (2.17e)$$

En rappelant que  $\mathbb{E}_{\mathbb{P}}(\mathbf{1}_{\mathbf{D} > u}) = \mathbb{P}(\mathbf{D} > u)$ , on obtient ainsi

$$\underbrace{J(u + 1) - J(u)}_{\text{variation marginale}} = \underbrace{-c + p\mathbb{P}(\mathbf{D} > u)}_{\text{décroissant en } u} . \quad (2.18)$$

La variation marginale du critère étant décroissante en  $u$ , un  $u$  optimal vérifiera, comme on peut le visualiser sur la figure 2.1, les inégalités

$$J(u_{\mathbb{E}}^\#) - J(u_{\mathbb{E}}^\# - 1) \geq 0 \geq J(u_{\mathbb{E}}^\# + 1) - J(u_{\mathbb{E}}^\#) . \quad (2.19)$$

Ces inégalités peuvent se reformuler grâce à l'égalité (2.18) en

$$\mathbb{P}(\mathbf{D} > u_{\mathbb{E}}^\#) \leq \frac{c}{p} \leq \mathbb{P}(\mathbf{D} > u_{\mathbb{E}}^\# - 1) , \quad (2.20)$$

ce qui, en faisant la transformation  $\mathbb{P}(\mathbf{D} > u_{\mathbb{E}}^\#) = 1 - \mathbb{P}(\mathbf{D} \leq u_{\mathbb{E}}^\#)$ , nous donne le résultat (2.16). □

## 2.2.4 Formulation d'un problème avec mesure de risque

Jusqu'ici, on s'est intéressé à un profit en moyenne c'est-à-dire que l'on accorde de l'importance aux scénarios en fonction de leurs fréquences d'apparition. On veut à présent accorder plus de poids aux scénarios qui conduisent à de faibles profits. Pour cela, on introduit une mesure de risque que l'on notera  $\mathbb{F}$ .

On formule alors, dans un cadre décision-hasard, le problème d'optimisation stochastique suivant

$$\min_{u \in \mathbb{U}} \mathbb{F}(-j(u, \mathbf{D})) . \quad (2.21)$$

On note le critère

$$J(u) = \mathbb{F}(-j(u, \mathbf{D})) . \quad (2.22)$$

### Généralités sur les mesures de risque

**Définition 2.3.** Une *mesure de risque*  $\mathbb{F}$  est une fonction  $\mathbb{R}^\Omega \rightarrow \mathbb{R}$  qui, à une variable aléatoire  $\mathbf{X}$ , associe un nombre réel. Cette mesure de risque est dite :

- *croissante* si  $\mathbf{X} \leq \mathbf{Y} \Rightarrow \mathbb{F}(\mathbf{X}) \leq \mathbb{F}(\mathbf{Y})$ ,
- *invariante par translation* si  $\forall m \in \mathbb{R}, \mathbb{F}(\mathbf{X} + m) = \mathbb{F}(\mathbf{X}) + m$ ,
- *convexe* si  $\mathbb{F}(\lambda \mathbf{X} + (1 - \lambda) \mathbf{Y}) \leq \lambda \mathbb{F}(\mathbf{X}) + (1 - \lambda) \mathbb{F}(\mathbf{Y}), \forall \lambda \in [0, 1]$ ,
- *positivement homogène* si  $\forall \lambda \geq 0, \mathbb{F}(\lambda \mathbf{X}) = \lambda \mathbb{F}(\mathbf{X})$ .

Une mesure de risque  $\mathbb{F}$  est dite

- *monétaire* si elle est croissante et invariante par translation,
- *cohérente* si elle est monétaire, convexe et positivement homogène.

*Remarque 2.4.* On peut remarquer que l'espérance classique  $\mathbb{F} = \mathbb{E}_{\mathbb{P}}$  est une mesure de risque qui vérifie les cinq propriétés énoncées dans la Définition 2.3. Elle est donc en particulier cohérente.

### Cas d'une mesure de risque monétaire homogène

On va à présent comparer le problème (2.13) dans le cas risque neutre au problème (2.21) dans le cas sensible au risque.

**Proposition 2.5.** *Pour le problème (2.21), si  $\mathbb{F}$  est une mesure de risque monétaire homogène alors le critère (2.22) s'écrit*

$$J(u) = cu + p\mathbb{F}(-\min\{u, \mathbf{D}\}) . \quad (2.23)$$

*Démonstration.* Le critère s'écrit initialement

$$J(u) = \mathbb{F}(-j(u, \mathbf{D})) = \mathbb{F}(cu - p \min\{u, \mathbf{D}\}) . \quad (2.24)$$

La mesure de risque  $\mathbb{F}$  étant invariante par translation, on a

$$J(u) = cu + \mathbb{F}(-p \min\{u, \mathbf{D}\}) . \quad (2.25)$$

Comme la mesure de risque  $\mathbb{F}$  est homogène, le critère s'écrit

$$J(u) = cu + p\mathbb{F}(-\min\{u, \mathbf{D}\}) , \quad (2.26)$$

ce qui conclut la preuve.  $\square$

On note une certaine analogie entre les formulations (2.17a) et (2.26) ce qui n'est pas étonnant puisque, comme énoncé dans la Remarque 2.4, l'espérance est une mesure de risque cohérente.

La variation marginale du critère, qui se met sous la forme

$$J(u+1) - J(u) = c + p \left( \mathbb{F}(-\min\{u+1, \mathbf{D}\}) - \mathbb{F}(-\min\{u, \mathbf{D}\}) \right), \quad (2.27)$$

ne conduit pas à des résultats analytiques intéressants pour déduire l'effet de la prise en compte du risque sur la décision optimale.

### Cas d'une mesure de risque particulière : la $\text{CVAR}_\beta$

Pour dégager une formulation analytique et ainsi comprendre l'impact du risque sur la décision optimale, on s'intéresse à une mesure de risque particulière, la Conditional Value at Risk pour laquelle nous adoptons la définition suivante. Cette définition n'est pas la définition "officielle" mais est une des formes les plus pratiques. (Voir [7]).

**Définition 2.6.** Soit un réel  $\beta \in [0, 1[$  et  $\mathbf{X} \in \mathbb{R}^\Omega$  une variable aléatoire. La *Conditional Value-at-risk* de la variable aléatoire  $\mathbf{X}$  est

$$\text{CVAR}_\beta(\mathbf{X}) = \inf_{s \in \mathbb{R}} \left\{ \frac{\mathbb{E}_{\mathbb{P}}((\mathbf{X} - s)^+)}{1 - \beta} + s \right\}. \quad (2.28)$$

Dans cette définition,  $\beta$  est un paramètre d'aversion au risque. Il est connu que (voir [7]) lorsque  $\beta$  est nul, la formule (2.28) coïncide avec l'espérance mathématique

$$\text{CVAR}_0(\mathbf{X}) = \mathbb{E}_{\mathbb{P}}(\mathbf{X}), \quad (2.29)$$

et que la limite de  $\text{CVAR}_\beta(\mathbf{X})$  lorsque  $\beta \rightarrow 1$  est la mesure de risque *Worst Case*

$$\text{CVAR}_1(\mathbf{X}) = \lim_{\beta \rightarrow 1} \text{CVAR}_\beta(\mathbf{X}) = \sup_{\omega \in \Omega} \mathbf{X}(\omega). \quad (2.30)$$

Nous allons, au travers de cette mesure de risque particulière, essayer de comprendre l'impact du risque mesuré par le paramètre  $\beta$  d'aversion au risque, sur la décision optimale, noté  $u_\beta^\#$ .

Le problème d'optimisation stochastique étudié s'écrit

$$\min_{u \in \mathbb{U}} \text{CVAR}_\beta(-j(u, \mathbf{D})). \quad (2.31)$$

Dans le cas où la demande  $d$  et le nombre de journaux commandés  $u$  sont continus, on sait d'après [4] qu'une solution optimale de (2.31) est caractérisé par

$$\mathbb{P}(\mathbf{D} \leq u_\beta^\#) = \frac{p - c}{p}(1 - \beta). \quad (2.32)$$

Le résultat (2.32) nous permet de comprendre qualitativement l'influence du paramètre d'aversion au risque  $\beta$  sur le nombre optimal de journaux commandés  $u$ .

On remarque tout d'abord que  $u_\beta^\sharp$  décroît quand  $\beta$  augmente. Ainsi, plus l'on est sensible au risque, moins l'on commande de journaux. En effet, on a

$$\frac{p-c}{p}(1-\beta) \leq \frac{p-c}{p}, \quad (2.33)$$

$$\mathbb{P}\left(\mathbf{D} \leq u_\beta^\sharp\right) \leq \mathbb{P}\left(\mathbf{D} \leq u_{\beta=0}^\sharp\right), \quad (2.34)$$

$$u_\beta^\sharp \leq u_{\beta=0}^\sharp = u_{\mathbb{E}}^\sharp. \quad (2.35)$$

### Formulation d'un programme linéaire dans le cas de la CVAR

Dans l'optique d'effectuer des applications numériques en Nsp<sup>1</sup>, on commence par reformuler le problème (2.31) comme un programme linéaire que l'on peut alors résoudre à l'aide d'un solveur. Cette reformulation a été proposée dans [8]. Le problème étudié est alors

$$\min_{\alpha \in \mathbb{R}, u \in \mathbb{U}} \alpha + \frac{1}{1-\beta} \mathbb{E}_{\mathbb{P}}\left([\ -j(u, \mathbf{D}) - \alpha]^+\right). \quad (2.36a)$$

**Proposition 2.7.** *Le problème (2.36) est équivalent au programme linéaire*

$$\min_{\alpha \in \mathbb{R}, u \in \mathbb{U}, \mathbf{X} \in \mathbb{R}^\Omega} \alpha + \frac{1}{1-\beta} \sum_{\omega \in \Omega} \mathbf{X}(\omega) \mathbb{P}(\omega), \quad (2.37a)$$

$$s.c. \quad \begin{cases} p\mathbf{D}(\omega) \geq cu - \mathbf{X}(\omega) - \alpha, & \forall \omega \in \Omega \\ 0 \geq (c-p)u - \mathbf{X}(\omega) - \alpha, & \forall \omega \in \Omega \\ \mathbf{X}(\omega) \geq 0, & \forall \omega \in \Omega \end{cases} \quad (2.37b)$$

au sens où les valeurs minimales coïncident.

Afin de démontrer la Proposition 2.7, nous avons tout d'abord besoin de la Définition 2.8 et du Lemme 2.9 ci-dessous.

**Définition 2.8.** Soit  $(\mathbf{X}, \mathbf{Y}) \in \mathbb{R}^\Omega \times \mathbb{R}^\Omega$ . On dit que  $\mathbf{X}$  est inférieur à  $\mathbf{Y}$  si

$$\forall \omega \in \Omega, \quad \mathbf{X}(\omega) \leq \mathbf{Y}(\omega) \text{ et on note } \mathbf{X} \leq \mathbf{Y}. \quad (2.38)$$

Remarquons que l'ordre défini par (2.38) n'est pas un ordre total.

**Lemme 2.9.** *Soit  $g : \mathbb{U} \times \mathbb{R} \rightarrow \mathbb{R}$  et  $h : X^\Omega \rightarrow \mathbb{R}$  une fonction croissante. Le problème d'optimisation*

$$v_o = \min_{\mathbf{X} \in \mathbb{R}^\Omega, u \in \mathbb{U}} h(\mathbf{X}), \quad (2.39a)$$

$$s.c. \quad \mathbf{X}(\omega) = \max\{0, g(u, \mathbf{D}(\omega))\}, \quad \forall \omega \in \Omega, \quad (2.39b)$$

*est équivalent au problème d'optimisation relâché*

---

1. <http://cermics.enpc.fr/jpc/nsp-tiddly/mine.html>

$$v_r = \min_{\mathbf{X} \in \mathbb{R}^\Omega, u \in \mathbb{U}} h(\mathbf{X}) , \quad (2.40a)$$

$$\text{s.c. } \mathbf{X}(\omega) \geq 0 , \quad \forall \omega \in \Omega , \quad (2.40b)$$

$$\mathbf{X}(\omega) \geq g(u, \mathbf{D}(\omega)) , \quad \forall \omega \in \Omega , \quad (2.40c)$$

au sens où les valeurs des problèmes coïncident.

*Démonstration.* On sait déjà que  $v_r \leq v_o$  car le problème relâché (2.40) minimise sur un ensemble admissible plus large que le problème (2.39). Soit  $(u^\sharp, \mathbf{X}^\sharp)$  une solution du problème relâché. On construit le couple

$$(u^*, \mathbf{X}^*) = (u^\sharp, \max\{0, g(u^\sharp, \mathbf{D})\}) , \quad (2.41)$$

qui est solution admissible du problème relâché (2.40) et du problème (2.39). Or  $\mathbf{X}^* \leq \mathbf{X}^\sharp$  et donc  $h(\mathbf{X}^*) \leq h(\mathbf{X}^\sharp)$  par croissance de  $h$ . Donc à partir de toute solution optimale de (2.40), on construit une solution admissible de (2.39) qui fait mieux que  $v_r$ . Par suite  $v_o \leq v_r$  ce qui conclut la preuve.  $\square$

Nous pouvons à présent démontrer la Proposition 2.7.

*Démonstration.* Comme  $\Omega$  est fini, le problème (2.36) se réécrit

$$\min_{\alpha \in \mathbb{R}, u \in \mathbb{U}} \alpha + \frac{1}{1 - \beta} \sum_{\omega \in \Omega} [cu - p \min\{u, \mathbf{D}(\omega)\} - \alpha]^+ \mathbb{P}(\omega) , \quad (2.42)$$

Il est évident que le problème (2.42) est équivalent au problème

$$(P_{\text{original}}) : \min_{\alpha \in \mathbb{R}, u \in \mathbb{U}, \mathbf{X} \in \mathbb{R}^\Omega} \alpha + \frac{1}{1 - \beta} \sum_{\omega \in \Omega} \mathbf{X}(\omega) \mathbb{P}(\omega) \quad (2.43a)$$

$$\text{s.c. } \mathbf{X}(\omega) = \max\{0, cu - p \min\{u, \mathbf{D}(\omega)\} - \alpha\} \quad \forall \omega \in \Omega . \quad (2.43b)$$

D'après le Lemme 2.9, on sait que le problème (2.43) est équivalent au problème relâché

$$(P_{\text{relâché}}) : \min_{\alpha \in \mathbb{R}, u \in \mathbb{U}, \mathbf{X} \in \mathbb{R}^\Omega} \alpha + \frac{1}{1 - \beta} \sum_{\omega \in \Omega} \mathbf{X}(\omega) \mathbb{P}(\omega) , \quad (2.44a)$$

$$\text{s.c. } \mathbf{X}(\omega) \geq 0 , \quad \forall \omega \in \Omega \quad (2.44b)$$

$$\mathbf{X}(\omega) \geq cu - p \min\{u, \mathbf{D}(\omega)\} - \alpha , \quad \forall \omega \in \Omega . \quad (2.44c)$$

Il est alors évident que la contrainte (2.44c) est équivalente aux contraintes (2.45c) et (2.45d) de

$$(P_{\text{final}}) : \min_{\alpha \in \mathbb{R}, u \in \mathbb{U}, \mathbf{X} \in \mathbb{R}^\Omega} \alpha + \frac{1}{1 - \beta} \sum_{\omega \in \Omega} \mathbf{X}(\omega) \mathbb{P}(\omega) , \quad (2.45a)$$

$$\text{s.c. } \mathbf{X}(\omega) \geq 0 , \quad \forall \omega \in \Omega \quad (2.45b)$$

$$\mathbf{X}(\omega) \geq cu - p\mathbf{D}(\omega) - \alpha , \quad \forall \omega \in \Omega \quad (2.45c)$$

$$\mathbf{X}(\omega) \geq cu - pu - \alpha , \quad \forall \omega \in \Omega . \quad (2.45d)$$

Cette dernière équivalence conclut preuve.  $\square$

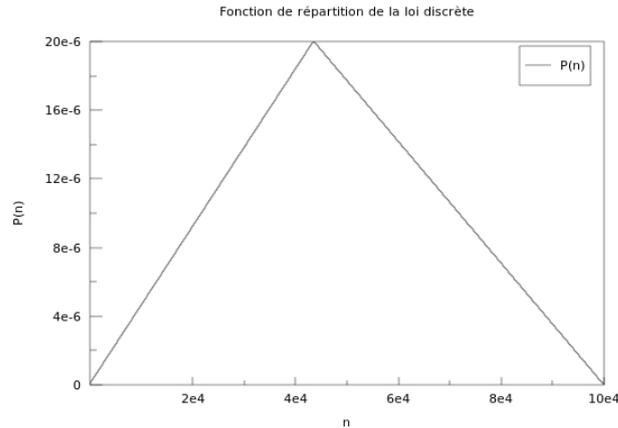


FIGURE 2.2 – Représentation de la loi de probabilité de la demande  $D$

### Applications numériques

Nous allons à présent confronter résultats numériques et théoriques. Les résultats numériques sont présentés comme des profits car plus naturels à analyser. On considère les données suivantes

- le prix d’achat initial est  $c_0 = 0,70$  €,
- le prix de vente est  $p = 1,20$  €,
- la demande maximale est  $\bar{D} = 100\,000$ ,
- les scénarios sont les demandes possibles, soit  $\Omega = \{0, \dots, \bar{D}\}$ ,
- la demande est donnée par la variable aléatoire  $D(\omega) = \omega \in \{0, \dots, \bar{D}\}$ ,
- la loi de probabilité est une loi unimodale triangulaire de paramètre  $\lambda = 43600$  dont la représentation est donnée par la Figure 2.2.

On a vu au § 2.2.4 que le cas  $\beta = 0$  correspond au cas risque neutre. On observe sur la figure 2.3 que plus le paramètre d’aversion au risque  $\beta$  augmente, plus le nombre optimal  $u_\beta^\#$  de journaux commandés diminue ce qui est cohérent avec le résultat théorique (2.32).

Connaissant le  $u_\beta^\#$ , on a effectué des simulations afin de comparer les profits dans le cas risque neutre ( $\beta = 0$ ) et un cas fortement sensible au risque ( $\beta = 0.9$ ). Les résultats sont obtenus pour 1 000 000 simulations.

La figure 2.4 nous montre l’histogramme des profits obtenus après les simulations. On observe que dans le cas sensible au risque, les profits sont plus faibles mais plus certains. L’échelle logarithmique nous permet d’observer les faibles fréquences en comparaison des pics.

Evolution du nombre optimal de journaux à commander en fonction de l'aversion au risque

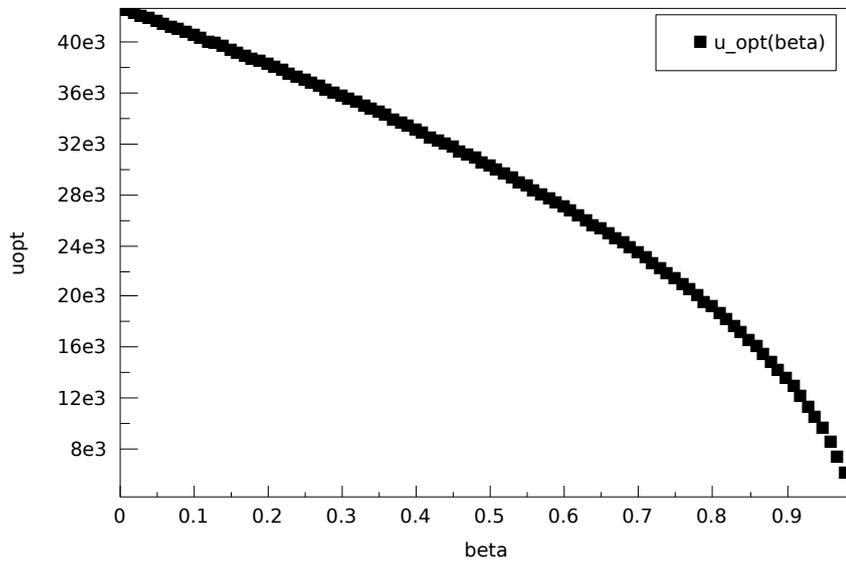


FIGURE 2.3 – Évolution du nombre optimal de journaux à commander en fonction du paramètre  $\beta$  d'aversion au risque. Les résultats correspondent à des  $\beta_i = i/100$

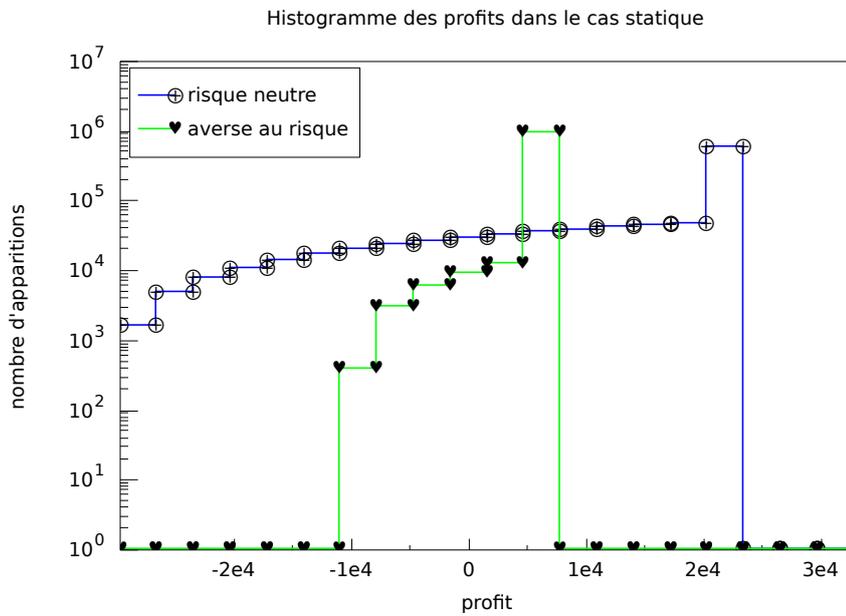


FIGURE 2.4 – Histogramme des profits obtenus lors des simulations

## 2.3 Vendeur de journaux à 2 périodes

Nous étudions à présent le cas d'un vendeur de journaux hebdomadaires qui peut faire deux commandes dans la semaine.

### Formulation du problème

On note  $u_0$  le nombre de journaux commandés au pas de temps 0 avant de connaître la demande  $\mathbf{D}$  et on note  $u_1$  le nombre de journaux commandés au pas de temps 1.

Le profit est défini par

$$\begin{aligned} j(u_0, u_1, d) = & -c_0 u_0 + p \min\{u_0, d\} \\ & -c_1 u_1 + p \min\{d - \min\{u_0, d\}, u_1\}, \end{aligned} \quad (2.46)$$

et la variante du vendeur de journaux étudiée est

$$\min_{u_0 \in \mathbb{U}, \mathbf{U}_1 \in \mathbb{U}^\Omega} \mathbb{F}(-j(u_0, \mathbf{U}_1, \mathbf{D})). \quad (2.47)$$

On note le critère

$$J(u_0, \mathbf{U}_1) = j(u_0, \mathbf{U}_1, \mathbf{D}). \quad (2.48)$$

Comme dans le § 2.2, on effectue une première commande en début de semaine de  $u_0$  journaux à un prix  $c_0$ . Puis, plus tard dans la semaine, un devin révèle quel scénario  $\omega$  se réalise. On peut alors effectuer une seconde commande de  $u_1$  journaux au prix  $c_1$  avec  $c = c_0 < c_1 < p$ .

Dans un premier temps, on ne connaît pas la demande quand on prend la décision  $\mathbf{U}_0$ . Noter  $u_0$  est une manière de représenter une variable aléatoire constante sur  $\Omega$ . Dans un second temps, on connaît la demande quand on prend la décision  $\mathbf{U}_1$ .

### Résolution de la 2ème période

**Proposition 2.10.** *Si  $u_0^\sharp$  est optimale pour le problème (2.47), alors le couple  $(u_0^\sharp, \mathbf{U}_1^\sharp)$  où  $\mathbf{U}_1^\sharp$  est défini par*

$$\mathbf{U}_1^\sharp = \mathbf{D} - \min\{u_0^\sharp, \mathbf{D}\}, \quad (2.49)$$

*est une solution optimale du problème (2.47).*

*Démonstration.* Soit  $u_0$  et  $d$  fixés. La fonction critère (2.46) ne dépend alors que de  $u_1$ . D'après (2.4), on a

$$\forall u_0, \forall d, -j(u_0, u_1, d) \geq -j(u_0, d - \min\{u_0, d\}, d). \quad (2.50)$$

Ceci étant vrai pour tout  $d$ , on a

$$-j(u_0, \mathbf{U}_1, \mathbf{D}) \geq -j(u_0, \mathbf{D} - \min\{u_0, \mathbf{D}\}, \mathbf{D}), \quad (2.51)$$

et comme la mesure de risque  $\mathbb{F}$  est croissante, on peut l'appliquer de chaque côté de l'inégalité pour obtenir

$$\mathbb{F}(-j(u_0, \mathbf{U}_1, \mathbf{D})) \geq \mathbb{F}(-j(u_0, \mathbf{D} - \min\{u_0, \mathbf{D}\}, \mathbf{D})) . \quad (2.52)$$

Soit  $u_0^\sharp$  une solution optimale du premier pas de temps du problème (2.47). On a d'après (2.52)

$$\mathbb{F}(-j(u_0^\sharp, \mathbf{U}_1, \mathbf{D})) \geq \mathbb{F}(-j(u_0^\sharp, \mathbf{D} - \min\{u_0^\sharp, \mathbf{D}\}, \mathbf{D})) . \quad (2.53)$$

Donc  $(u_0^\sharp, \mathbf{D} - \min\{u_0^\sharp, \mathbf{D}\})$  est solution optimale du problème (2.47) ce qui conclut la preuve.  $\square$

### Résolution 1ère période

La Proposition 2.10 nous permet de restreindre notre étude à celle de la fonction profit

$$\begin{aligned} j(u_0, d - \min\{u_0, d\}, d) &= -c_0 u_0^\sharp + \cancel{p \min\{u_0^\sharp, d\}} \\ &\quad - c_1 d + c_1 \min\{u_0^\sharp, d\} \\ &\quad + pd - \cancel{p \min\{u_0^\sharp, d\}} . \end{aligned} \quad (2.54)$$

Le profit (2.54) apparaît alors comme une fonction uniquement de  $(u_0, d)$ . On peut donc étudier le critère

$$\begin{aligned} J(u_0) &= \mathbb{F}\left(-j(u_0, \mathbf{D} - \min\{u_0, \mathbf{D}\}, \mathbf{D})\right) \\ &= \mathbb{F}(c_0 u_0 - c_1 \min\{u_0, \mathbf{D}\} + (c_1 - p)\mathbf{D}) , \\ &= c_0 u_0 - c_1 \mathbb{F}(\min\{u_0, \mathbf{D}\} + (c_1 - p)\mathbf{D}) \end{aligned} \quad (2.55)$$

afin de trouver un  $u_0^\sharp$  puis lui associer le  $\mathbf{U}_1^\sharp$  par (2.49).

### Influence de la dynamique dans le cas risque neutre

**Proposition 2.11.** *Une solution optimale du problème (2.47) dans le cas risque neutre est donnée par*

$$\begin{aligned} \mathbb{P}\left(\mathbf{D} \leq u_0^\sharp - 1\right) &\leq \frac{c_1 - c_0}{c_1} \leq \mathbb{P}\left(\mathbf{D} \leq u_0^\sharp\right) , \\ \mathbf{U}_1^\sharp &= \mathbf{D} - \min\{u_0^\sharp, \mathbf{D}\} . \end{aligned} \quad (2.56)$$

*Démonstration.* La démonstration reprend celle de la Proposition 2.2 dans le cas statique en utilisant (2.55) comme fonction critère.  $\square$

Comme  $(c_1 - c_0)/c_1 < (p - c_0)/p$ , on a

$$u_{0,\mathbb{E}}^\sharp \leq u_{\mathbb{E}}^\sharp . \quad (2.57)$$

On voit donc que la possibilité de faire une deuxième commande a une influence sur le contrôle optimal initial. Quand on peut faire une deuxième commande, on a intérêt à commander moins à la première période pour se prémunir des demandes faibles quitte à commander plus au deuxième pas de temps pour combler une demande éventuellement forte.

Le cas  $c_1 = p$  correspond au cas dégénéré du vendeur à un pas de temps. Le cas  $c_1 = c_0$  correspond au cas où l'on ne commande pas de journaux.

*Remarque 2.12.* On remarque que  $u_0^\sharp$  ne dépend pas du prix de vente  $p$ . En effet, si l'on ne commande rien au premier pas de temps et tout au deuxième pas de temps, on est sûr de ne pas avoir de perte. Si l'on veut augmenter son profit, il faut commander au premier pas de temps et la différence avec le profit assuré est alors  $(p - c_0)u_0 - (p - c_1)u_0 = (c_1 - c_0)u_0^\sharp$ . C'est donc cette différence de gains par rapport au profit de base qu'on cherche à maximiser.

### Formulation d'un programme linéaire

La fonction critère (2.55) ne fait intervenir que la variable de contrôle  $u_0$ . On prend pour mesure de risque la  $\text{CVAR}_\beta$  et on adapte la démonstration de la Proposition 2.7 avec la fonction critère (2.55) pour obtenir le contrôle optimal  $u_0^\sharp$  dans un cadre dynamique avec un programme linéaire.

### Applications numériques

On reprend les paramètres numériques utilisés dans le cas statique et on donne en plus  $c_1 = 1,00$  €, le prix d'achat d'un journal hebdomadaire à la deuxième période. Dans ce cas, on compare l'évolution du  $u^\sharp$  statique à  $u_0^\sharp$  dynamique. On remarque, comme le montre la figure 2.5, que le contrôle optimal initial dans le cas dynamique est toujours plus faible que dans le cas statique. Cependant, plus on est sensible au risque, plus cette différence est faible. On cherche en effet à se prémunir des pertes donc on anticipe les scénarios les pires.

On a également effectué des simulations dans le cas dynamique. Les figures 2.6 et 2.7 permettent de comparer les résultats entre le cas risque neutre et le cas sensible au risque avec un coefficient d'aversion au risque  $\beta = 0.9$ . On compare également le cas statique et le cas dynamique.

Les profits dans le cas sensible au risque sont toujours plus faibles mais plus certains que ce soit dans le cas statique ou dans le cas dynamique. Le cas dynamique a de plus tendance à augmenter les profits.

## 2.4 Conclusion

En tant que vendeur de journaux, nous avons vu qu'il est possible d'adopter différentes attitudes pour gérer le commerce : regarder un profit moyen ou se prémunir des profits faibles par exemple. L'optimisation apparaît alors comme une aide à la décision puisque la solution optimale dépend de l'attitude choisie.

On a ainsi perçu au travers de ce chapitre et de cet exemple les effets de la prise en compte du risque dans une décision ainsi que ceux de la prise en compte du caractère multi-étapes d'un système et donc de l'information.

Plusieurs termes techniques ont été introduits succinctement dans ce chapitre et seront détaillés plus tard : espace de scénario, variable aléatoire, mesurabilité, mesure de risque, information, loi de probabilité.

Nous allons à présent introduire un formalisme pour traiter de manière plus générale certains problèmes d'optimisation stochastique.

Evolution du nombre optimal de journaux à commande en fonction de l'aversion au risque dans le cas dynamique

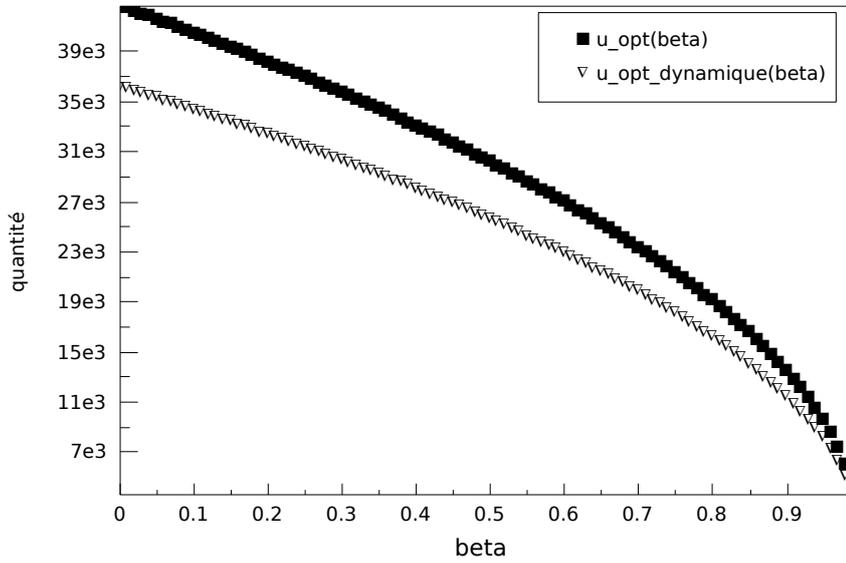


FIGURE 2.5 – Comparaison du contrôle optimal  $u_0^\#$  dans le cas statique et dans le cas dynamique en fonction du paramètre d'aversion au risque  $\beta$ .

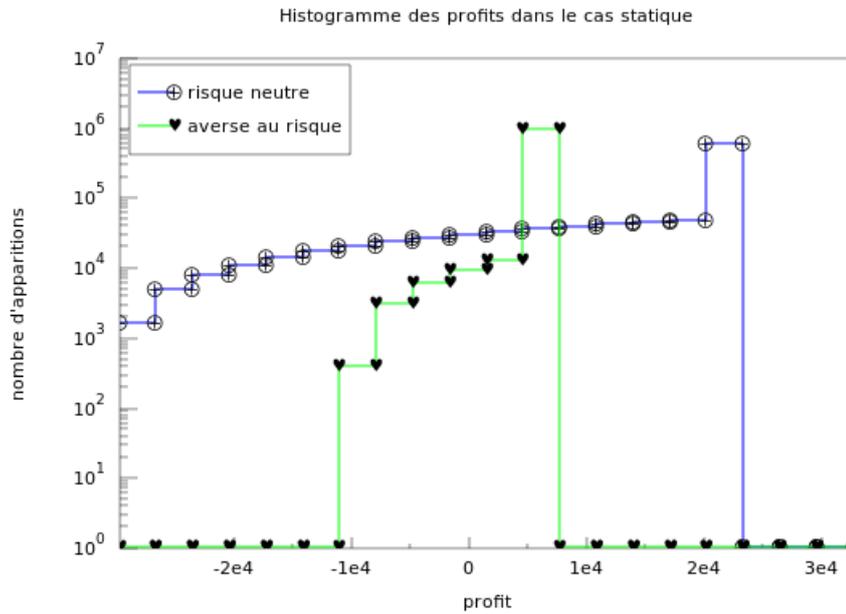


FIGURE 2.6 – Histogrammes des profits dans le cas statique. Le profit est donné pour  $c = 0.7$ ,  $p = 1.2$  et le paramètre d'aversion au risque utilisé est  $\beta = 0,9$

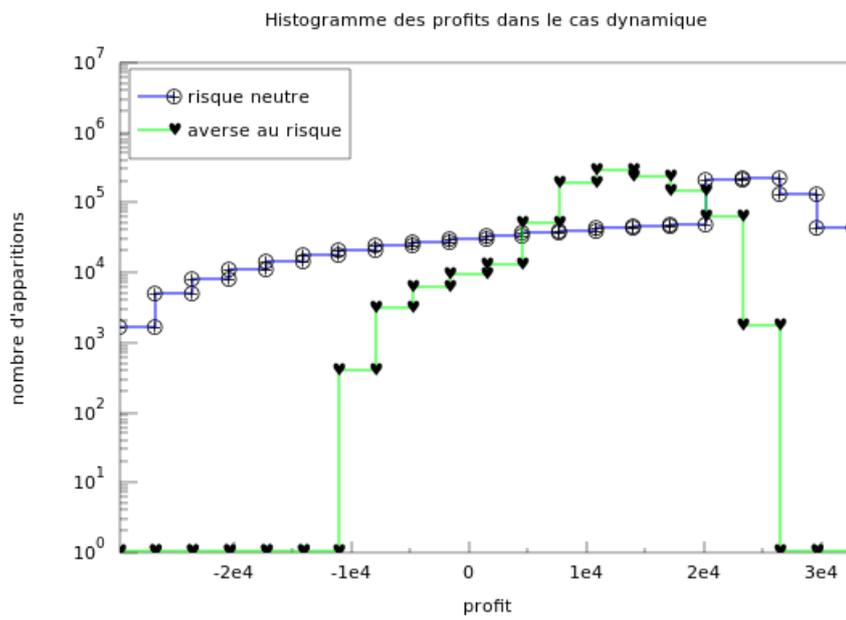


FIGURE 2.7 – Histogrammes des profits dans le cas dynamique. Le profit est donné pour  $c_0 = 0.7$ ,  $c_1 = 1.0$ ,  $p = 1.2$  et le paramètre d'aversion au risque utilisé est  $\beta = 0,9$

# 3 Optimisation stochastique à 1 période avec contraintes de mesurabilité

## 3.1 Introduction

Ce chapitre présente un formalisme mathématique et des outils techniques permettant de traiter des problèmes d'optimisation stochastique statiques avec des contraintes d'information.

Les démonstrations et notions seront présentées dans le cas fini. Néanmoins, le lecteur est invité à consulter [11] et [15] qui généralisent la plupart des résultats au cas continu.

## 3.2 Formulation d'un problème d'optimisation stochastique sous contraintes d'information

Nous étudions le problème

$$\min_{\mathbf{U} \preceq \mathcal{G}} \mathbb{F}(j(\mathbf{U}, \mathbf{W})) . \tag{3.1}$$

Les différents objets nécessaires à la compréhension de ce problème sont :

- la mesure de risque  $\mathbb{F}$ ,
- le critère  $j$ ,
- les variables aléatoire  $\mathbf{U}, \mathbf{W}$ ,
- une tribu  $\mathcal{G}$ .

**Le coût et les variables aléatoires** Pour définir nos variables aléatoires, on prend l'univers  $\Omega$  et on se donne deux espaces  $\mathbb{U}$  et  $\mathbb{W}$ . Le contrôle  $\mathbf{U}$  est alors un élément de  $\mathbb{U}^\Omega$  et le bruit  $\mathbf{W}$  est un élément de  $\mathbb{W}^\Omega$ . Le critère  $j : \mathbb{U} \times \mathbb{W} \rightarrow \mathbb{R}$  permet alors de définir la variable aléatoire  $j(\mathbf{U}, \mathbf{W}) : \omega \mapsto j(\mathbf{U}(\omega), \mathbf{W}(\omega))$  et le critère  $J(\mathbf{U}) = \mathbb{F}(j(\mathbf{U}, \mathbf{W}))$

Dans l'équation (3.5), il nous reste à préciser ce que veut dire  $\mathbf{U} \preceq \mathcal{G}$ .

## Rappels sur les tribus

Nous rappelons ici brièvement quelques notions mathématiques sur les tribus et ce qu'elles modélisent. On peut trouver plus de détails dans [3].

Une tribu  $\mathcal{G}$  sur  $\Omega$  est un façon mathématique de représenter de l'information.

**Définition 3.1.** Une tribu  $\mathcal{G}$  sur  $\Omega$  est une collection non vide de sous-ensembles de  $\Omega$  qui est stable par union finie et par passage au complémentaire.

La Définition 3.1 est la plus générale. En pratique, quand  $\Omega$  est fini, une tribu peut être caractérisée par ses atomes qui forment une partition de  $\Omega$ .

**Définition 3.2.** Soit  $\mathcal{G}$  une collection de sous-ensembles de  $\Omega$ . Un atome de  $\mathcal{G}$  est un sous-ensemble  $G$  non vide tel que si  $K \in \mathcal{G}$  et  $K \subset G$  alors on  $K = \emptyset$  ou  $K = G$ .

**Définition 3.3.** Soit  $E$  un ensemble de parties. La plus petite tribu contenant  $E$  est notée  $\sigma(E)$ .

**Proposition 3.4.** Soit une tribu  $\mathcal{G}$  et  $\{G_a, a \in A\}$  l'ensemble de ses atomes. Alors

$$\mathcal{G} = \sigma(G_a, a \in A) , \quad (3.2)$$

et  $\{G_a, a \in A\}$  forme une partition de  $\Omega$ .

À présent nous allons définir la mesurabilité de variables aléatoires.

**Définition 3.5.** Soit  $\mathbf{X}$  à valeurs dans un espace mesurable  $(\mathbb{X}, \mathcal{X})$ , la tribu notée  $\sigma(\mathbf{X}) = \{\mathbf{X}^{-1}(B) | B \in \mathcal{X}\}$  est la tribu engendrée par  $\mathbf{X}$ . On dit que  $\mathbf{X}$  est  $\mathcal{G}$ -mesurable si  $\sigma(\mathbf{X}) \subset \mathcal{G}$ , ce que l'on note  $\mathbf{X} \preceq \mathcal{G}$  :

$$\mathbf{X} \preceq \mathcal{G} \Leftrightarrow \sigma(\mathbf{X}) \subset \mathcal{G} . \quad (3.3)$$

**Proposition 3.6.** Une variable aléatoire  $\mathbf{X}$  est  $\mathcal{G}$ -mesurable si elle est constante sur chaque atome de la tribu  $\mathcal{G}$ .

Nous illustrons ces définitions au travers d'un exemple.

**Exemple 3.7.** La figure 3.1 illustre les différents éléments présentés au travers de cet exemple. Soit l'ensemble de scénario  $\Omega = \{\omega_1, \omega_2, \omega_3\}$ .

On définit la tribu  $\mathcal{G} = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3\}, \Omega\}$ . On remarque que  $\mathcal{G}$  est bien stable par union et par passage au complémentaire. Les atomes de  $\mathcal{G}$  sont les éléments  $G = \{\omega_1, \omega_2\}$  et  $\bar{G} = \{\omega_3\}$ . On remarque que les atomes de  $\mathcal{G}$  forment bien une partition. Les scénarios  $\omega_1$  et  $\omega_2$  sont dits indiscernables car dans le même atome  $G$ . Pour notre exemple,  $\mathbf{X}$  est  $\mathcal{G}$ -mesurable si  $\mathbf{X}(\omega_1) = \mathbf{X}(\omega_2)$ .

**Rappel sur l'espérance et l'espérance conditionnelle** Nous rappelons les formulation de l'espérance et de l'espérance conditionnelle d'une variable aléatoire  $\mathbf{X}$ . On se place dans le cas où  $\mathcal{G} = \sigma(G_a, a \in A)$

$$\mathbb{E}_{\mathbb{P}}(\mathbf{X}) = \sum_{\omega \in \Omega} \mathbf{X}(\omega) \mathbb{P}(\omega) , \quad (3.4a)$$

$$\mathbb{E}_{\mathbb{P}}(\mathbf{X} \mid \mathcal{G}) = \sum_{a \in A, \mathbb{P}(G_a) > 0} \frac{\mathbb{E}_{\mathbb{P}}(\mathbf{1}_{G_a} \mathbf{X})}{\mathbb{P}(G_a)} \mathbf{1}_{G_a} . \quad (3.4b)$$

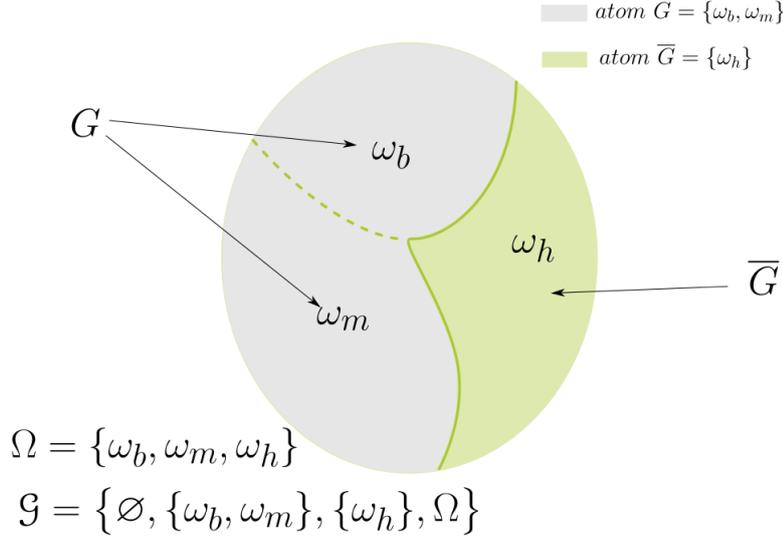


FIGURE 3.1 – Représentation graphique de l’univers  $\Omega$ , de la tribu  $\mathcal{G}$  et de ses atomes  $\{G, \bar{G}\}$ .

### 3.3 Résolution par imbrication

Ici nous cherchons à résoudre le problème (3.1) en le décomposant en sous-problèmes d’optimisation, un par atome de la tribu  $\mathcal{G}$ .

#### 3.3.1 Cas risque neutre

Dans le cas risque neutre, le problème (3.1) s’écrit

$$\min_{U \preceq \mathcal{G}} \mathbb{E}_{\mathbb{P}}(j(U, \mathbf{W})) . \quad (3.5)$$

**Présentation de l’objectif** Nous cherchons les conditions pour pouvoir écrire (3.5) avec une formulation du type

$$\min_{U \preceq \mathcal{G}} \mathbb{E}_{\mathbb{P}}(j(U, \mathbf{W})) = \mathbb{E}_{\mathbb{P}}\left(\min_{u \in U} \mathbb{E}_{\mathbb{P}}(j(u, \mathbf{W}) \mid \mathcal{G})\right) . \quad (3.6)$$

Ceci permettrait de décomposer le problème (3.5) en sous-problèmes représentés par

$$\min_{u \in U} \mathbb{E}_{\mathbb{P}}(j(u, \mathbf{W}) \mid \mathcal{G}) . \quad (3.7)$$

L’équation (3.7) désigne en fait plusieurs problèmes, un pour chaque atome  $G_a$  (voir la Proposition 3.6) de la tribu  $\mathcal{G}$ . Les solutions de ces sous-problèmes permettent de reconstituer une solution de problème initial (3.5).

Nous étudions dans un premier temps le problème (3.5) avec une tribu  $\mathcal{G}$  quelconque. Nous regarderons ensuite ce que donne le résultat énoncé en (3.8) dans les cas particuliers du Chapitre 2 ; à savoir lorsqu’on a accès aucune information avec  $\mathcal{G} = \sigma(\emptyset, \Omega)$  et lorsque l’on a accès toute l’information avec  $\mathcal{G} = \sigma(\{\omega\}_{\omega \in \Omega})$ .

**Cas général :  $\mathcal{G}$  est quelconque**

**Proposition 3.8** ([3]). *Le problème (3.5) se réécrit*

$$\min_{\mathbf{U} \preceq \mathcal{G}} \mathbb{E}_{\mathbb{P}}(j(\mathbf{U}, \mathbf{W})) = \mathbb{E}_{\mathbb{P}}\left(\min_{u \in \mathbb{U}} \mathbb{E}_{\mathbb{P}}(j(u, \mathbf{W})) \mid \mathcal{G}\right), \quad (3.8)$$

et le minimum est atteint pour la variable aléatoire définie par :

$$\mathbf{U}^{\#}(\omega) = u_a^{\#} \in \arg \min_{u \in \mathbb{U}} \frac{\mathbb{E}_{\mathbb{P}}(\mathbf{1}_{G_a} j(u, \mathbf{W}))}{\mathbb{P}(G_a)}, \quad \forall \omega \in G_a, \quad \forall a \in A. \quad (3.9)$$

**Cas sans information  $\mathcal{G} = \sigma(\{\emptyset, \Omega\})$**

Dans le cas où  $\mathcal{G}$  est la tribu triviale  $\sigma(\{\emptyset, \Omega\})$ , la variable aléatoire  $\mathbf{U}$  est constante sur  $\Omega$ . L'espérance conditionnelle par rapport à la tribu triviale est alors, comme l'indique la formule (3.4b), l'espérance classique. La Proposition 3.8 donne :

$$\min_{\mathbf{U} \preceq \sigma(\{\emptyset, \Omega\})} \mathbb{E}_{\mathbb{P}}(j(\mathbf{U}, \mathbf{W})) = \mathbb{E}_{\mathbb{P}}\left(\min_{u \in \mathbb{U}} \mathbb{E}_{\mathbb{P}}(j(u, \mathbf{W}))\right) = \min_{u \in \mathbb{U}} \mathbb{E}_{\mathbb{P}}(j(u, \mathbf{W})). \quad (3.10)$$

On voit que dans le cas où l'on ne possède aucune information, la variable aléatoire  $\mathbf{U}$  peut s'identifier à une valeur unique.

**Cas avec information complète  $\mathcal{G} = \sigma(\{\omega\}_{\omega \in \Omega})$**

Dans le cas où l'information est complète, on distingue chaque scénario  $\omega$ . Ainsi, d'après l'expression (3.4b), l'espérance conditionnelle d'une variable aléatoire par rapport à  $\mathcal{G}$  est égale à la variable aléatoire :

$$\mathbb{E}_{\mathbb{P}}(\mathbf{X} \mid \mathcal{G}) = \mathbf{X}. \quad (3.11)$$

La Proposition 3.8 et on obtient :

$$\min_{\mathbf{U} \preceq \sigma(\{\omega\}_{\omega \in \Omega})} \mathbb{E}_{\mathbb{P}}(j(\mathbf{U}, \mathbf{W})) = \mathbb{E}_{\mathbb{P}}\left(\min_{u \in \mathbb{U}} \mathbb{E}_{\mathbb{P}}(j(u, \mathbf{W}) \mid \mathcal{G})\right) = \mathbb{E}_{\mathbb{P}}\left(\min_{u \in \mathbb{U}} j(u, \mathbf{W})\right). \quad (3.12)$$

Toute solution de (3.12) est donnée par

$$\mathbf{U}(\omega) \in \arg \min_{u \in \mathbb{U}} j(u, \mathbf{W}(\omega)). \quad (3.13)$$

### 3.3.2 Extension de la résolution avec mesure de risque

Nous avons, dans le § 3.3.1, résolu le problème (3.5) par imbrication. On cherche à résoudre le problème (3.1) avec ce genre de méthode.

$$\min_{\mathbf{U} \preceq \mathcal{G}} \mathbb{F}(j(\mathbf{U}, \mathbf{W})). \quad (3.14)$$

et la formulation à laquelle on souhaite aboutir est du type suivant

$$\min_{\mathbf{U} \preceq \mathcal{G}} \mathbb{F}(j(\mathbf{U}, \mathbf{W})) = \mathbb{F}\left(\min_{u \in \mathbb{U}} \mathbb{F}(j(u, \mathbf{W}) \mid \mathcal{G})\right), \quad (3.15)$$

où l'on donne un sens à  $\mathbb{F}(\cdot \mid \mathcal{G})$  dans le paragraphe suivant.

## Mesure de risque conditionnelle et cohérence temporelle

Dans le cas risque neutre, on a naturellement une espérance conditionnelle sachant  $\mathcal{G}$  et le résultat important suivant :

$$\mathbb{E}_{\mathbb{P}}(\mathbf{X}) = \mathbb{E}_{\mathbb{P}}(\mathbb{E}_{\mathbb{P}}(\mathbf{X} \mid \mathcal{G})) \quad (3.16)$$

On aimerait, dans le cas d'une mesure de risque quelconque, définir une mesure de risque sachant  $\mathcal{G}$  et avoir un résultat du même type

$$\mathbb{F}(\mathbf{Z}) = \mathbb{F}\left(\underbrace{\mathbb{G}(\mathbf{Z})}_{\mathcal{G}\text{-mesurable}}\right). \quad (3.17)$$

On va définir ce qu'est une mesure de risque conditionnelle (Voir [2],[12],[13],[14]). Pour cela, on introduit les objets suivants :

- On note  $\mathcal{Z}$  l'ensemble des variables aléatoires sur  $\Omega$ ,
- On note  $\mathcal{Z}_{\mathcal{G}}$  l'ensemble des variables aléatoires  $\mathcal{G}$ -mesurables

Une mesure de risque conditionnellement à la tribu  $\mathcal{G}$  est une fonction  $\mathbb{G}$  à valeurs dans  $\mathcal{Z}_{\mathcal{G}}$

$$\begin{aligned} \mathbb{G} : \mathcal{Z} &\rightarrow \mathcal{Z}_{\mathcal{G}} \\ \mathbf{Z} &\mapsto \mathbb{G}(\mathbf{Z}). \end{aligned} \quad (3.18)$$

**Définition 3.9.** Une mesure de risque conditionnelle  $\mathbb{G}$  est dite :

- *monotone* si pour  $(\mathbf{X}, \mathbf{Y}) \in \mathcal{Z}^2$ ,  $\mathbf{X} \leq \mathbf{Y} \Rightarrow \mathbb{G}(\mathbf{X}) \leq \mathbb{G}(\mathbf{Y})$ .
- *invariante par translation* : si pour  $(\mathbf{M}, \mathbf{X}) \in \mathcal{Z}_{\mathcal{G}} \times \mathcal{Z}$  on a  $\mathbb{G}(\mathbf{X} + \mathbf{M}) = \mathbb{G}(\mathbf{X}) + \mathbf{M}$ .
- *convexe* si  $\forall \mathbf{X}, \mathbf{Y} \in \mathcal{Z}_{\mathcal{G}}$ ,  $\mathbb{G}(\lambda \mathbf{X} + (1 - \lambda)\mathbf{Y}) \leq \lambda \mathbb{G}(\mathbf{X}) + (1 - \lambda)\mathbb{G}(\mathbf{Y})$ .
- *positivement homogène* : si pour  $\lambda \geq 0$  et  $\mathbf{X} \in \mathcal{Z}$  on a,  $\mathbb{G}(\lambda \mathbf{X}) = \lambda \mathbb{G}(\mathbf{X})$ .

La Définition 3.9 est le pendant de la Définition 2.3 lorsque les variables aléatoires sont à valeurs dans  $\mathcal{Z}_{\mathcal{G}}$  et non pas dans  $\mathbb{R}$ .

**Définition 3.10.** Le couple de mesures de risque  $(\mathbb{F}, \mathbb{G})$ , données dans cet ordre, vérifie la propriété de cohérence temporelle si

$$\mathbb{F}\left(\underbrace{\mathbf{Z}}_{\in \mathcal{Z}}\right) = \mathbb{F}\left(\underbrace{\mathbb{G}(\mathbf{Z})}_{\in \mathcal{Z}_{\mathcal{G}}}\right), \quad \forall \mathbf{Z} \in \mathcal{Z}. \quad (3.19)$$

### Formulation imbriquée

La proposition suivante généralise la formulation (3.8).

**Proposition 3.11.** Soit  $(\mathbb{F}, \mathbb{G})$  une suite de mesures de risque vérifiant la propriété de cohérence temporelle (3.19). Supposons que  $\arg \min_{u \in \mathbb{U}} \mathbb{G}(j(u, \mathbf{W}))$  est non vide pour tout  $u \in \mathbb{U}$  et tout  $\omega \in \Omega$ . Alors le problème (3.14) s'écrit

$$\min_{\mathbf{U} \preceq \mathcal{G}} \mathbb{F}(j(\mathbf{U}, \mathbf{W})) = \mathbb{F}\left(\min_{u \in \mathbb{U}} \mathbb{G}(j(u, \mathbf{W}))\right), \quad (3.20)$$

et le minimum est atteint pour la variable aléatoire définie par :

$$\mathbf{U}^{\sharp}(\omega) = u_a^{\sharp} \in \arg \min_{u \in \mathbb{U}} \mathbb{G}(j(u, \mathbf{W}))(\omega), \quad \forall \omega \in G_a, \quad \forall a \in A. \quad (3.21)$$

*Démonstration.* Comme  $\mathbb{G}$  est une mesure de risque conditionnelle à valeurs dans  $\mathcal{Z}_{\mathcal{G}}$ , on a :

$$\forall \mathbf{U} \in \mathbb{U}^{\Omega}, \quad \mathbb{G}(j(\mathbf{U}, \mathbf{W})) \in \mathcal{Z}_{\mathcal{G}}. \quad (3.22)$$

On note  $g(u, \omega)$  la valeur prise par la variable aléatoire  $\mathbb{G}(j(u, \mathbf{W}))$  en  $\omega$  :

$$g(u, \omega) = \mathbb{G}(j(u, \mathbf{W}))(\omega). \quad (3.23)$$

La variable aléatoire  $\omega \mapsto g(u, \omega)$  est  $\mathcal{G}$ -mesurable puisque  $\mathbb{G}(j(u, \mathbf{W}))$  l'est. Donc, si  $\mathcal{G} \subset \sigma(G_a, a \in A)$ , où  $G_a$  sont les atomes de  $\mathcal{G}$ ,

$$\forall \omega, \omega' \in G_a, \quad g(u, \omega) = g(u, \omega'), \quad \forall a \in A. \quad (3.24)$$

et

$$\forall \omega, \omega' \in G_a, \quad \arg \min_{u \in \mathbb{U}} g(u, \omega) = \arg \min_{u \in \mathbb{U}} g(u, \omega'), \quad \forall a \in A. \quad (3.25)$$

Pour tout  $a$  élément de  $A$ , on sélectionne un  $u_a^{\#}$  appartenant à l'argmin ci-dessus et on note  $\mathbf{U}^{\#}$  la variable aléatoire telle que

$$\mathbf{U}^{\#}|_{G_a} = u_a^{\#}. \quad (3.26)$$

On en déduit que pour tout  $\mathbf{U}$   $\mathcal{G}$ -mesurable,

$$g(\mathbf{U}^{\#}(\omega), \omega) \leq g(\mathbf{U}(\omega), \omega), \quad \omega \in \Omega. \quad (3.27)$$

On obtient alors la suite d'inégalités suivantes

$$\begin{aligned} \min_{u \in \mathbb{U}} \mathbb{G}(j(u, \mathbf{W})) &\leq \mathbb{G}(j(\mathbf{U}, \mathbf{W})) && \text{par définition de } \mathbf{U}^{\#} \text{ en (3.26)} \\ \mathbb{F}\left(\min_{u \in \mathbb{U}} \mathbb{G}(j(u, \mathbf{W}))\right) &\leq \mathbb{F}\left(\mathbb{G}(j(\mathbf{U}, \mathbf{W}))\right) && \text{par croissance de } \mathbb{F} \\ \mathbb{F}\left(\min_{u \in \mathbb{U}} \mathbb{G}(j(u, \mathbf{W}))\right) &\leq \mathbb{F}(j(\mathbf{U}, \mathbf{W})) && \text{par cohérence temporelle (3.19)} \\ \mathbb{F}\left(\min_{u \in \mathbb{U}} \mathbb{G}(j(u, \mathbf{W}))\right) &\leq \inf_{\mathbf{U} \preceq \mathcal{G}} \mathbb{F}(j(\mathbf{U}, \mathbf{W})) && \text{par définition de l'inf.} \end{aligned}$$

Or en remplaçant  $\mathbf{U}$  par  $\mathbf{U}^{\#}$ , les inégalités précédentes deviennent des égalités et l'inf est atteint.  $\square$

### 3.3.3 Résumons nous

Nous avons résolu une classe de problèmes d'optimisation stochastique dans le cas risque neutre par décomposition en sous-problèmes, atomes par atomes, moyennant un théorème d'interversion entre un minimum et une espérance. Cette décomposition nous conduit alors à une formulation dite imbriquée.

En introduisant les mesure de risque conditionnelles et la propriété de cohérence temporelle, nous avons ensuite généralisé cette décomposition du cas risque neutre au cas sensible au risque.

Nous allons à présent traiter la contrainte de mesurabilité par dualité.

## 3.4 Résolution par dualisation des contraintes

Nous portons à présent attention aux contraintes de mesurabilité. Ces contraintes vu au §3.2 couplent les  $\omega$  entre eux. Nous allons chercher à dualiser les contraintes de mesurabilité pour décomposer un problème d'optimisation stochastique en plusieurs sous-problèmes.

### 3.4.1 Formulation du problème dans le cas risque neutre

Le problème étudié est le problème (3.5) que nous pouvons réécrire comme ceci

$$\min_{U \in \mathcal{U}^\Omega} \mathbb{E}_{\mathbb{P}}(j(U, \mathbf{W})) , \quad (3.28a)$$

$$\text{s.c. } U \preceq \mathcal{G} . \quad (3.28b)$$

Nous allons nous intéresser à une formulation fonctionnelle de la contrainte de mesurabilité (3.28b) en utilisant :

$$U \preceq \mathcal{G} \Rightarrow \mathbb{E}_{\mathbb{P}}(U \mid \mathcal{G}) = U , \quad \mathbb{P}\text{-p.s.} . \quad (3.29)$$

On remarquera que la réciproque de (3.29) n'est pas toujours vérifiée comme on peut le voir dans l'exemple suivant où la loi de probabilité  $\mathbb{P}$  ne charge pas tous les points.

**Exemple 3.12.** On se donne les objets suivants

- $\Omega = \{\omega_1, \omega_2, \omega_3\}$ ,
- $\mathbb{P} = (0, 2/3, 1/3)$  au sens où  $\mathbb{P}(\omega_1) = 0, \text{etc.}$ ,
- $\mathbf{X} = (2, 1, 3)$  au sens où  $\mathbf{X}(\omega_1) = 2, \text{etc.}$ ,
- $\mathcal{G} = \{\emptyset, \{\omega_1, \omega_2\}, \{\omega_3\}, \Omega\}$ .

On trouve alors en utilisant la formulation (3.4b) de l'espérance conditionnelle que  $\mathbb{E}_{\mathbb{P}}(\mathbf{X} \mid \mathcal{G}) = (1, 1, 3)$ . Comme  $\mathbb{P}(\omega_1) = 0$  on a

$$\underbrace{\mathbb{E}_{\mathbb{P}}(\mathbf{X} \mid \mathcal{G})}_{\mathcal{G}\text{-mesurable}} = \underbrace{\mathbf{X}}_{\text{non } \mathcal{G}\text{-mesurable}} , \quad \mathbb{P}\text{-p.s.} , \quad (3.30)$$

puisque les deux variables aléatoires sont égales sauf sur un ensemble de probabilité nulle. Cependant  $\mathbf{X}$  n'est pas  $\mathcal{G}$ -mesurable car attribue deux valeurs différentes à  $\omega_1$  et à  $\omega_2$  alors que  $\mathcal{G}$  ne fait pas la distinction entre ces deux scénarios.

Nous travaillerons a présent sur le problème relaxé suivant :

$$\min_{U \in \mathcal{U}^\Omega} \mathbb{E}_{\mathbb{P}}(j(U, \mathbf{W})) , \quad (3.31a)$$

$$\text{s.c. } \mathbb{E}_{\mathbb{P}}(U \mid \mathcal{G}) = U , \quad \mathbb{P}\text{-p.s.} . \quad (3.31b)$$

Nous allons traiter les contraintes (3.31b) par dualité.

### 3.4.2 Résolution dans le cas risque neutre

Nous prenons pour cette section  $\mathbb{U} = \mathbb{R}$  mais on peut généraliser les résultats au cas où  $\mathbb{U} = \mathbb{R}^N$ . Comme nous sommes dans le cas risque neutre, nous disposons d'un produit scalaire naturel, associé à l'espérance que l'on adopte, et qui est défini par :

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbb{E}_{\mathbb{P}}(\mathbf{X}\mathbf{Y}), \quad \mathbb{P}\text{-p.s.} . \quad (3.32)$$

S'étant doté d'un produit scalaire, on peut définir le Lagrangien du problème (3.31) qui s'écrit, sur  $\mathbb{R}^{\Omega} \times \mathbb{R}^{\Omega}$  :

$$\mathcal{L}(\mathbf{U}, \boldsymbol{\pi}) := \mathbb{E}_{\mathbb{P}}(j(\mathbf{U}, \mathbf{W})) + \langle \boldsymbol{\pi}, \mathbb{E}_{\mathbb{P}}(\mathbf{U} \mid \mathcal{G}) - \mathbf{U} \rangle, \quad (3.33)$$

$$= \mathbb{E}_{\mathbb{P}}(j(\mathbf{U}, \mathbf{W})) + \mathbb{E}_{\mathbb{P}}\left(\boldsymbol{\pi} [\mathbb{E}_{\mathbb{P}}(\mathbf{U} \mid \mathcal{G}) - \mathbf{U}]\right). \quad (3.34)$$

Le problème primal et le problème dual sont alors respectivement définis par :

$$\min_{\mathbf{U} \in \mathbb{R}^{\Omega}} \max_{\boldsymbol{\pi} \in \mathbb{R}^{\Omega}} \mathcal{L}(\mathbf{U}, \boldsymbol{\pi}), \quad (3.35a)$$

$$\max_{\boldsymbol{\pi} \in \mathbb{R}^{\Omega}} \min_{\mathbf{U} \in \mathbb{R}^{\Omega}} \mathcal{L}(\mathbf{U}, \boldsymbol{\pi}). \quad (3.35b)$$

*Remarque 3.13.* Il est connu [10] que (3.35a) est équivalent à (3.31) et que (3.35b) est une borne inférieure de (3.35a).

**Théorème 3.14.** *Sous les hypothèses*

1. *l'ensemble  $\mathbb{U}$  est convexe,*
2. *la fonction  $j$  est convexe, sci et propre en  $u$  sur  $\mathbb{U}$  pour tout  $w \in \mathbb{W}$ ,*

*le problème (3.31) est équivalent à*

$$\max_{\boldsymbol{\pi} \in \mathbb{R}^{\Omega}} \mathbb{E}_{\mathbb{P}}\left(\min_{u \in \mathbb{R}} j(u, \mathbf{W}) + (\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi} \mid \mathcal{G}) - \boldsymbol{\pi}) \cdot u\right). \quad (3.36)$$

*et si les hypothèses ne sont pas vérifiées, la formulation (3.36) fournit alors une borne inférieure.*

*Démonstration.* Les contraintes presque sûres formulées avec une espérance conditionnelle sont affines, en nombre fini car  $\Omega$  est fini donc qualifiées. Les hypothèses du Théorème A.34 nous assurent alors que le problème primal et le problème dual sont équivalents, et on étudie donc le problème dual. Dans un premier temps, on reformule le Lagrangien (3.33) afin de transférer la contrainte de mesurabilité de  $\mathbf{U}$  au multiplicateur  $\boldsymbol{\pi}$ . Pour cela, on remarque que

$$\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi} \cdot (\mathbb{E}_{\mathbb{P}}(\mathbf{U} \mid \mathcal{G}) - \mathbf{U})) = \mathbb{E}_{\mathbb{P}}((\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi} \mid \mathcal{G}) - \boldsymbol{\pi}) \cdot \mathbf{U}). \quad (3.37)$$

Le Lagrangien (3.33) se réécrit donc

$$\mathcal{L}(\mathbf{U}, \boldsymbol{\pi}) = \mathbb{E}_{\mathbb{P}}(j(\mathbf{U}, \mathbf{W})) + \mathbb{E}_{\mathbb{P}}((\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi} \mid \mathcal{G}) - \boldsymbol{\pi}) \cdot \mathbf{U}), \quad (3.38)$$

$$= \mathbb{E}_{\mathbb{P}}(j(\mathbf{U}, \mathbf{W}) + (\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi} \mid \mathcal{G}) - \boldsymbol{\pi}) \cdot \mathbf{U}), \quad (3.39)$$

et le problème dual (3.35b) s'écrit

$$\max_{\boldsymbol{\pi} \in \mathbb{R}^\Omega} \min_{U \in \mathbb{R}^\Omega} \mathbb{E}_{\mathbb{P}}(j(\boldsymbol{U}, \boldsymbol{W}) + (\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi} \mid \mathcal{G}) - \boldsymbol{\pi}) \cdot \boldsymbol{U}) . \quad (3.40)$$

On observe alors que la variable  $\boldsymbol{U}$  n'a plus de contraintes de mesurabilité. En appliquant la Proposition (3.8), on se retrouve dans le cas (3.12).

La borne inférieure est une application immédiate de la Remarque 3.13 et résultat d'interversion (3.8).  $\square$

### Cas sans information

On rappelle que le cas sans information correspond au cas où  $\mathcal{G}$  est la tribu triviale c'est-à-dire  $\mathcal{G} = \sigma(\{\emptyset, \Omega\})$ . Dans ce cas, l'espérance conditionnelle par rapport à  $\mathcal{G}$  est l'espérance mathématique. Le problème (3.36) se réécrit alors

$$\max_{\boldsymbol{\pi} \in \mathbb{R}^\Omega} \mathbb{E}_{\mathbb{P}}\left(\min_{u \in \mathbb{R}} j(u, \boldsymbol{W}) + (\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi}) - \boldsymbol{\pi}) \cdot u\right) . \quad (3.41)$$

### Cas avec information complète

Le cas avec information complète est le cas où  $\mathcal{G} = \sigma(\{\omega\}_{\omega \in \Omega})$ . L'espérance conditionnelle d'une variable aléatoire est alors égale à la variable aléatoire elle-même c'est-à-dire

$$\mathbb{E}_{\mathbb{P}}(\boldsymbol{U} \mid \mathcal{G}) = \boldsymbol{U} . \quad (3.42)$$

Le problème (3.36) se réécrit alors

$$\max_{\boldsymbol{\pi} \in \mathbb{R}^\Omega} \mathbb{E}_{\mathbb{P}}\left(\min_{u \in \mathbb{R}^\Omega} j(u, \boldsymbol{W}) + (\boldsymbol{\pi} - \boldsymbol{\pi}) \cdot u\right) . \quad (3.43)$$

Comme  $\boldsymbol{\pi} - \boldsymbol{\pi} = 0$ , le maximum en  $\boldsymbol{\pi}$  n'intervient plus et on obtient

$$\mathbb{E}_{\mathbb{P}}\left(\min_{u \in \mathbb{R}} j(u, \boldsymbol{W})\right) . \quad (3.44)$$

### 3.4.3 Extension de la résolution dans le cas avec une mesure de risque

Le but de cette section est de généraliser le résultat (3.36) en remplaçant l'espérance  $\mathbb{E}$  par une mesure de risque  $\mathbb{F}$ .

L'écriture des contraintes de mesurabilité devient plus délicate. Dans le cas risque neutre, l'espérance conditionnelle était naturellement associée à l'espérance et permettait d'écrire une contrainte de mesurabilité linéaire. Dans le cas avec risque, non seulement il faut réussir à se doter d'une mesure de risque conditionnelle  $\mathbb{G}$  cohérente temporellement avec  $\mathbb{F}$  mais celle-ci n'a pas de raison d'être linéaire.

On se donne une loi de probabilité  $\mathbb{P}$  et le produit scalaire (3.32). Le problème étudié est

$$\min_{\boldsymbol{U}} \mathbb{F}(j(\boldsymbol{U}, \boldsymbol{W})) , \quad (3.45a)$$

$$\text{s.c. } \mathbb{E}_{\mathbb{P}}(\boldsymbol{U} \mid \mathcal{G}) = \boldsymbol{U} . \quad (3.45b)$$

Nous pouvons définir le Lagrangien du problème (3.45) qui s'écrit :

$$\mathcal{L}(\mathbf{U}, \boldsymbol{\pi}) = \mathbb{F}(j(\mathbf{U}, \mathbf{W})) + \langle \boldsymbol{\pi}, \mathbb{E}_{\mathbb{P}}(\mathbf{U} \mid \mathcal{G}) - \mathbf{U} \rangle, \quad (3.46)$$

$$= \mathbb{F}(j(\mathbf{U}, \mathbf{W})) + \mathbb{E}_{\mathbb{P}}\left(\boldsymbol{\pi}(\mathbb{E}_{\mathbb{P}}(\mathbf{U} \mid \mathcal{G}) - \mathbf{U})\right), \quad (3.47)$$

$$= \mathbb{F}(j(\mathbf{U}, \mathbf{W})) + \mathbb{E}_{\mathbb{P}}\left((\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi} \mid \mathcal{G}) - \boldsymbol{\pi})\mathbf{U}\right). \quad (3.48)$$

Le problème primal et le problème dual sont alors respectivement défini par

$$\min_{\mathbf{U} \in \mathbb{R}^{\Omega}} \max_{\boldsymbol{\pi} \in \mathbb{R}^{\Omega}} \mathcal{L}(\mathbf{U}, \boldsymbol{\pi}), \quad (3.49a)$$

$$\max_{\boldsymbol{\pi} \in \mathbb{R}^{\Omega}} \min_{\mathbf{U} \in \mathbb{R}^{\Omega}} \mathcal{L}(\mathbf{U}, \boldsymbol{\pi}). \quad (3.49b)$$

**Proposition 3.15.** *Sous les hypothèses*

1. *l'ensemble  $\mathbb{U}$  est convexe,*
2. *la mesure  $\mathbb{F}$  est monétaire convexe,*
3. *le critère  $j$  est convexe en  $u$  sur  $\mathbb{U}$  pour tout  $w \in \mathbb{W}$ ,*

*alors le problème (3.45) est équivalent à*

$$\max_{\boldsymbol{\pi} \in \mathbb{R}^{\Omega}} \min_{\mathbf{U} \in \mathbb{R}^{\Omega}} \mathbb{F}\left(j(\mathbf{U}, \mathbf{W}) + \mathbb{E}_{\mathbb{P}}\left((\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi} \mid \mathcal{G}) - \boldsymbol{\pi})\mathbf{U}\right)\right). \quad (3.50)$$

*Démonstration.* La croissance de  $\mathbb{F}$  et sa convexité assurent que la fonction à minimiser dans l'équation (3.45a) est convexe en  $\mathbf{U}$ .

Le transfert de la contrainte de mesurabilité de  $\mathbf{U}$  à  $\boldsymbol{\pi}$  est détaillé dans la preuve de la Proposition 3.14.

Enfin le caractère invariant par translation de la mesure de risque  $\mathbb{F}$  permet de faire entrer la contrainte dualisée, qui se présente comme un scalaire (résultat d'un produit scalaire), à l'intérieur de la mesure de risque.  $\square$

A ce stade, on ne peut pas rentrer le minimum en  $\mathbf{U}$  de l'équation (3.50) sous la mesure de risque car, à l'intérieur de la mesure de risque, il reste l'opération espérance qui couple la variable aléatoire  $\mathbf{U}$  en  $\omega$ . L'espérance intérieure disparaît dans le cas risque neutre en raison de la série d'égalités

$$\mathbb{E}_{\mathbb{P}}(\mathbf{X} + \mathbb{E}_{\mathbb{P}}(\mathbf{Y})) = \mathbb{E}_{\mathbb{P}}(\mathbf{X}) + \mathbb{E}_{\mathbb{P}}(\mathbf{Y}) = \mathbb{E}_{\mathbb{P}}(\mathbf{X} + \mathbf{Y}). \quad (3.51)$$

Dans le cas d'une mesure de risque invariante par translation, on peut seulement dire

$$\mathbb{F}(\mathbf{X} + \mathbb{E}_{\mathbb{P}}(\mathbf{Y})) = \mathbb{F}(\mathbf{X}) + \mathbb{E}_{\mathbb{P}}(\mathbf{Y}). \quad (3.52)$$

**Proposition 3.16.** *Soit une mesure de risque  $\mathbb{F}$  qui s'écrit*

$$\mathbb{F}(\mathbf{X}) = \max_{\mathbb{Q} \in \mathcal{Q}_{\mathbb{F}}} \mathbb{E}_{\mathbb{Q}}(\mathbf{X}), \quad (3.53)$$

*où  $\mathcal{Q}_{\mathbb{F}}$  est un ensemble de lois de probabilité convexe, fermé et borné. Supposons que pour tout  $\mathbb{Q}$  de  $\mathcal{Q}_{\mathbb{F}}$ ,  $\mathbb{Q}$  est absolument continu par rapport à la loi  $\mathbb{P}$ . Si on se place sous les hypothèses de la Proposition 3.15, alors le problème (3.50) est équivalent à*

$$\max_{\boldsymbol{\pi} \in \mathbb{R}^{\Omega}} \max_{\mathbb{Q} \in \mathcal{Q}_{\mathbb{F}}} \mathbb{E}_{\mathbb{P}}\left(\min_{u \in \mathbb{R}} j(u, \mathbf{W}) \frac{d\mathbb{Q}}{d\mathbb{P}} + [\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi} \mid \mathcal{G}) - \boldsymbol{\pi}]u\right). \quad (3.54)$$

*Démonstration.* On commence par reformuler le problème (3.50) grâce à (3.53)

$$\max_{\pi \in \mathbb{R}^\Omega} \min_{U \in \mathbb{R}^\Omega} \max_{Q \in \mathcal{Q}_F} \mathbb{E}_Q \left( j(U, \mathbf{W}) + \mathbb{E}_P \left( (\mathbb{E}_P(\pi | \mathcal{G}) - \pi) U \right) \right). \quad (3.55)$$

On définit alors la fonction  $f : \mathcal{Q}_F \times \mathbf{U} \rightarrow \mathbb{R}$ , par

$$f(Q, U) = \mathbb{E}_Q \left( j(U, \mathbf{W}) + \mathbb{E}_P \left( (\mathbb{E}_P(\pi | \mathcal{G}) - \pi) U \right) \right). \quad (3.56)$$

La fonction  $f$  est convexe-concave, sci-scs sur deux ensembles convexes. Elle est de plus coercive en  $Q$  car  $\mathcal{Q}_F$  est borné. Elle est coercive en  $U$  car  $j$  est fortement convexe, que l'espérance suivant  $\mathbb{P}$  est linéaire et que la mesure  $\mathbb{F}$  est convexe croissante.

On sait alors que

$$\min_{U \in \mathbb{R}^\Omega} \max_{Q \in \mathcal{Q}} f(Q, U) = \max_{Q \in \mathcal{Q}} \min_{U \in \mathbb{R}^\Omega} f(Q, U). \quad (3.57)$$

Toutes les lois de l'ensemble  $\mathcal{Q}_F$  sont absolument continues par rapport à la loi  $\mathbb{P}$ . On peut alors utiliser le théorème de Radon-Nikodym pour réécrire l'espérance suivant  $Q$  comme une espérance suivant  $\mathbb{P}$ . Ceci nous permet d'utiliser la série d'égalités (3.51) et de faire disparaître l'espérance intérieure. On peut alors rentrer le minimum grâce à la Proposition 3.8 ce qui conclut la preuve.  $\square$

Les mesures de risque cohérentes vérifie les hypothèses de la Proposition 3.16 d'après [7]. L'ensemble  $\mathcal{Q}_F$  associé à la  $\text{CVAR}_\beta$  est polyhedral (voir [6]) ce est intéressant pour des applications numériques.

### 3.4.4 Résumons nous

Nous nous sommes intéressé à une forme relaxée des contraintes de mesurabilité ce qui nous a permis de les dualiser au sens de Lagrange.

Nous avons alors vu comment il est possible, en s'intéressant au problème dual, de s'affranchir de la contrainte de mesurabilité sur  $U$  dans le problème de minimisation. Ainsi, en utilisant un théorème d'interversion entre l'espérance et le minimum, on se ramène à des sous-problèmes que l'on traite  $\omega$  par  $\omega$ .

L'étude dans le cas d'une mesure de risque générale a été l'occasion d'introduire les mesures de risque invariantes en loi. Nous avons aussi introduit les mesures de risque cohérentes qui s'écrivent comme un maximum d'espérance sous différentes lois de probabilités.

## 3.5 Conclusion

Nous avons, au travers de ce chapitre, présenté des outils mathématiques permettant de traiter de manière générale des problèmes d'optimisation stochastique à 1 pas de temps avec contraintes de mesurabilité.

Nous avons proposé des méthodes de résolution par imbrication et par dualisation.

Nous allons à présent nous intéresser à des problèmes à deux pas de temps de décision.

# 4 Optimisation stochastique à 2 périodes avec contraintes de mesurabilité

## 4.1 Introduction

Nous étudions dans ce chapitre les problèmes d'optimisation stochastique dynamiques. On introduit la dimension temporelle et on regarde son impact du temps sur l'écriture des contraintes de mesurabilité et nous généralisons les résultats obtenus dans le cas statique.

Comme dans le Chapitre 2, nous nous intéresserons à des problèmes à deux pas de temps tout en sachant que les résultats se généralisent à des problèmes à plusieurs pas de temps.

## 4.2 Résolution par imbrication

Nous spécifions dans cette section les résultats énoncés en §3.3 au cas dynamique. Nous verrons notamment pourquoi un problème dynamique ne peut pas se traiter comme un problème statique lorsqu'on a des contraintes de mesurabilités.

### 4.2.1 Formulation du problème dans le cas risque neutre

#### Énoncé du problème

Nous énonçons le problème que nous allons étudier. On se donne deux tribus

$$\mathcal{F}_0 \subset \mathcal{F}_1, \quad \mathcal{F}_0 \neq \mathcal{F}_1, \quad (4.1)$$

et le problème s'énonce

$$\min_{U_0 \in \mathcal{U}_0^\Omega, U_1 \in \mathcal{U}_1^\Omega} \mathbb{E}_{\mathbb{P}}(j(U_0, U_1, \mathbf{W})), \quad (4.2a)$$

$$\text{s.c.} \begin{cases} U_0 \preceq \mathcal{F}_0, \\ U_1 \preceq \mathcal{F}_1. \end{cases} \quad (4.2b)$$

Pour comprendre ce problème, quatre éléments sont à définir plus précisément — les variables de décisions  $U_0$  et  $U_1$ ,

- la variable de bruit  $\mathbf{W}$ ,
- le coût  $j$ ,
- les tribus  $\mathcal{F}_0$  et  $\mathcal{F}_1$ .

**Discussion à propos des variables aléatoires et du critère.** On associe à chacune des trois variables aléatoires  $(\mathbf{U}_0, \mathbf{U}_1, \mathbf{W})$  un espace de valeurs ; respectivement  $\mathbb{U}_0$ ,  $\mathbb{U}_1$ ,  $\mathbb{W}$ . Ainsi,  $\mathbf{U}_0 \in \mathbb{U}_0^\Omega$ ,  $\mathbf{U}_1 \in \mathbb{U}_1^\Omega$ ,  $\mathbf{W} \in \mathbb{W}^\Omega$ .

Le coût est la fonction

$$j : \mathbb{U}_0 \times \mathbb{U}_1 \times \mathbb{W} \rightarrow \mathbb{R}, \quad (4.3)$$

et on rappelle à nouveau

$$j(\mathbf{U}_0, \mathbf{U}_1, \mathbf{W}) : \omega \mapsto j(\mathbf{U}_0(\omega), \mathbf{U}_1(\omega), \mathbf{W}(\omega)). \quad (4.4)$$

## 4.2.2 Discussion à propos des contraintes

Les contraintes de mesurabilité (4.2b) sont données pas de temps par pas de temps et sont appelées contraintes de non anticipativité. Elles ne peuvent s'agréger en une seule contrainte. En effet, si on pose  $\mathbf{U} = (\mathbf{U}_0, \mathbf{U}_1)$ , on peut écrire

$$\left. \begin{array}{l} \mathbf{U}_0 \preceq \mathcal{G} \\ \mathbf{U}_1 \preceq \mathcal{G} \end{array} \right\} \Leftrightarrow \mathbf{U} \preceq \mathcal{G}, \quad (4.5)$$

mais on ne peut pas écrire les contraintes

$$\mathbf{U}_0 \preceq \mathcal{F}_0, \quad (4.6a)$$

$$\mathbf{U}_1 \preceq \mathcal{F}_1, \quad (4.6b)$$

comme une seule contrainte de mesurabilité car  $\mathcal{F}_0 \subset \mathcal{F}_1$  et  $\mathcal{F}_0 \neq \mathcal{F}_1$ .

**Présentation de l'objectif du chapitre** Nous cherchons les conditions pour pouvoir écrire le problème (4.2) sous une formulation du type

$$\mathbb{E}_{\mathbb{P}} \left( \min_{u_0 \in \mathbb{U}_0} \mathbb{E}_{\mathbb{P}} \left( \min_{u_1 \in \mathbb{U}_1} \mathbb{E}_{\mathbb{P}} (j(u_0, u_1, \mathbf{W}) \mid \mathcal{F}_1) \mid \mathcal{F}_0 \right) \right). \quad (4.7)$$

Ceci permettrait de décomposer le problème (4.2) par pas de temps. On résout deux problèmes

- d'abord un problème de minimisation en  $u_1$  à  $u_0$  fixé

$$\min_{u_1 \in \mathbb{U}_1} \mathbb{E}_{\mathbb{P}} (j(u_0, u_1, \mathbf{W}) \mid \mathcal{F}_1), \quad (4.8)$$

- puis un problème de minimisation en  $u_0$

$$\min_{u_0 \in \mathbb{U}_0} \mathbb{E}_{\mathbb{P}} \left( \min_{u_1 \in \mathbb{U}_1} \mathbb{E}_{\mathbb{P}} (j(u_0, u_1, \mathbf{W}) \mid \mathcal{F}_1) \mid \mathcal{F}_0 \right) \quad (4.9)$$

On parle alors de résolution backward. C'est la méthode employée dans le Chapitre 2 lorsqu'on a résolu le problème du vendeur de journaux à deux pas de temps en donnant d'abord  $\mathbf{U}_1^\#$  avant de donner  $u_0^\#$ .

### 4.2.3 Résolution par imbrication dans le cas risque neutre

**Proposition 4.1.** *Le problème (4.2) peut s'écrire*

$$\mathbb{E}_{\mathbb{P}} \left( \min_{u_0 \in \mathbb{U}_0} \mathbb{E}_{\mathbb{P}} \left( \min_{u_1 \in \mathbb{U}_1} \mathbb{E}_{\mathbb{P}}(j(u_0, u_1, \mathbf{W}) \mid \mathcal{F}_1) \mid \mathcal{F}_0 \right) \right). \quad (4.10)$$

*Démonstration.* On commence par réécrire le problème (4.2) comme

$$\min_{\mathbf{U}_0 \preceq \mathcal{F}_0} \min_{\mathbf{U}_1 \preceq \mathcal{F}_1} \mathbb{E}_{\mathbb{P}}(j(\mathbf{U}_0, \mathbf{U}_1, \mathbf{W})). \quad (4.11)$$

D'après (3.8), on peut faire entrer le minimum en  $\mathbf{U}_1$  sous l'espérance pour obtenir

$$\min_{\mathbf{U}_0 \preceq \mathcal{F}_0} \mathbb{E}_{\mathbb{P}} \left( \min_{u_1 \in \mathbb{U}_1} \mathbb{E}_{\mathbb{P}}(j(\mathbf{U}_0, u_1, \mathbf{W}) \mid \mathcal{F}_1) \right) \quad (4.12)$$

En toute rigueur, pour effectuer l'opération ci-dessus, il faut justifier que l'on peut minimiser  $\mathbb{E}_{\mathbb{P}}(j(\mathbf{U}_0, u_1, \mathbf{W}) \mid \mathcal{F}_1)$  à  $u_0$  fixé ce que nous détaillons pas ici mais qui est détaillé dans [9]. Retenons que cette opération est possible parce que  $\mathcal{F}_0 \subset \mathcal{F}_1$ .

Toujours d'après (3.8), on peut ensuite faire entrer le minimum en  $\mathbf{U}_0$  sous l'espérance pour obtenir

$$\mathbb{E}_{\mathbb{P}} \left( \min_{u_0 \in \mathbb{U}_0} \mathbb{E}_{\mathbb{P}} \left( \min_{u_1 \in \mathbb{U}_1} \mathbb{E}_{\mathbb{P}}(j(u_0, u_1, \mathbf{W}) \mid \mathcal{F}_1) \mid \mathcal{F}_0 \right) \right). \quad (4.13)$$

□

*Remarque 4.2.* On rappelle qu'on peut écrire

$$\mathcal{F}_1 = \sigma(G_a^1, a \in A), \quad (4.14)$$

où les  $G_a^1$  sont les atomes de la tribu  $\mathcal{F}_1$ .

Le problème

$$\min_{u_1 \in \mathbb{U}_1} \mathbb{E}_{\mathbb{P}}(j(u_0, u_1, \mathbf{W}) \mid \mathcal{F}_1) \quad (4.15)$$

est alors à entendre comme la variable aléatoire issue de l'opération de minimisation sur chaque atome  $G_a^1$  de la partition engendrée par  $\mathcal{F}_1$ .

### 4.2.4 Extension de la résolution dans le cas avec mesure de risque

Le but de cette section est de généraliser l'imbrication aux problèmes d'optimisation stochastique dynamique avec risque. Le problème que l'on étudie s'écrit :

$$\min_{\mathbf{U}_0 \in \mathbb{U}_0^\Omega, \mathbf{U}_1 \in \mathbb{U}_1^\Omega} \mathbb{F}(j(\mathbf{U}_0, \mathbf{U}_1, \mathbf{W})), \quad (4.16a)$$

$$\text{s.c.} \begin{cases} \mathbf{U}_0 \preceq \mathcal{F}_0, \\ \mathbf{U}_1 \preceq \mathcal{F}_1. \end{cases} \quad (4.16b)$$

On se donne pour cette section trois mesures de risque  $(\mathbb{F}, \mathbb{F}_0, \mathbb{F}_1)$  vérifiant la condition (3.19) de cohérence temporelle c'est-à-dire telles que :

$$\begin{aligned} \mathbb{F}(\mathbb{F}_0(\mathbf{X})) &= \underbrace{\mathbb{F}(\mathbf{X})}_{\in \mathbb{R}}, \quad \forall \mathbf{X} \\ \mathbb{F}_0(\underbrace{\mathbb{F}_1(\mathbf{X})}_{\mathcal{F}_1\text{-mesurable}}) &= \underbrace{\mathbb{F}_0(\mathbf{X})}_{\mathcal{F}_0\text{-mesurable}}, \quad \forall \mathbf{X}. \end{aligned} \quad (4.17)$$

**Proposition 4.3.** *Soit  $(\mathbb{F}, \mathbb{F}_0, \mathbb{F}_1)$  une mesure de risque dynamique cohérente temporellement alors le problème (4.16) est équivalent à*

$$\mathbb{F}\left(\min_{u_0 \in \mathcal{U}_0} \mathbb{F}_0\left(\min_{u_1 \in \mathcal{U}_1} \mathbb{F}_1(j(u_0, u_1, \mathbf{W}))\right)\right). \quad (4.18)$$

*Démonstration.* La démonstration copie la démonstration de la Proposition 4.1 en utilisant les résultats généralisés de §3.3.1 au lieu de ceux du cas risque neutre de §3.3.2.  $\square$

On peut donc, sous la bonne axiomatique, utiliser une méthode par imbrication pour résoudre des problèmes d'optimisation stochastique sensibles au risque.

#### 4.2.5 Résumons nous

Nous avons vu comment utiliser les résultats obtenus dans le cadre statique pour décomposer le problème pas de temps par pas de temps. On résout alors le problème en backward que ce soit dans un cadre risque neutre ou sensible au risque.

Nous allons à présent regarder les résultats que l'on obtient lorsqu'on relaxe les contraintes de mesurabilité dans le cas de problèmes dynamiques.

### 4.3 Résolution par dualisation des contraintes

Nous allons à présent étudier dans un cadre dynamique les problèmes relaxés du §3.4.

#### 4.3.1 Formulation du problème dans le cas risque neutre

Nous définissons un problème relaxé de (4.2) par

$$\min_{\mathbf{U}_0 \in \mathcal{U}_0^\Omega, \mathbf{U}_1 \in \mathcal{U}_1^\Omega} \mathbb{E}_{\mathbb{P}}(j(\mathbf{U}_0, \mathbf{U}_1, \mathbf{W})). \quad (4.19a)$$

$$\text{s.c.} \begin{cases} \mathbb{E}_{\mathbb{P}}(\mathbf{U}_0 \mid \mathcal{F}_0) = \mathbf{U}_0 \\ \mathbb{E}_{\mathbb{P}}(\mathbf{U}_1 \mid \mathcal{F}_1) = \mathbf{U}_1 \end{cases}. \quad (4.19b)$$

Le lecteur est invité à relire le §3.4.1 pour la discussion sur la relaxation des contraintes (4.2b) et le §4.2.2 pour la discussion sur le caractère dynamique des problèmes (4.2) et (4.19).

### 4.3.2 Résolution dans le cas risque neutre

Dans le cas risque neutre, on rappelle qu'on dispose d'un produit scalaire naturel entre deux variables aléatoires, qui est défini par

$$\langle \mathbf{X}, \mathbf{Y} \rangle = \mathbb{E}_{\mathbb{P}}(\mathbf{X}\mathbf{Y}) . \quad (4.20)$$

On peut alors définir le Lagrangien du problème (4.19) par

$$\begin{aligned} \mathcal{L}(\mathbf{U}_0, \mathbf{U}_1, \boldsymbol{\pi}_0, \boldsymbol{\pi}_1) = & \mathbb{E}_{\mathbb{P}}(j(\mathbf{U}_0, \mathbf{U}_1, \mathbf{W})) \\ & + \langle \boldsymbol{\pi}_0, \mathbb{E}_{\mathbb{P}}(\mathbf{U}_0 \mid \mathcal{F}_0) - \mathbf{U}_0 \rangle \\ & + \langle \boldsymbol{\pi}_1, \mathbb{E}_{\mathbb{P}}(\mathbf{U}_1 \mid \mathcal{F}_1) - \mathbf{U}_1 \rangle \end{aligned} \quad (4.21)$$

On se place dans le cas où  $\mathbb{U}_0 = \mathbb{U}_1 = \mathbb{R}$  pour la suite. Les résultats se généralisent facilement à des dimensions supérieures.

**Proposition 4.4.** *Sous les hypothèses*

1. *l'ensemble  $\mathbb{U}$  est convexe,*
2. *la fonction  $j$  est conjointement convexe, sci et propre en  $u_0$  et  $u_0$  sur  $\mathbb{U}_0 \times \mathbb{U}_1$  pour tout  $w \in \mathbb{W}$ ,*

*alors le problème (4.19) est équivalent à*

$$\begin{aligned} \max_{\boldsymbol{\pi}_0 \in \mathbb{R}_0^{\Omega}, \boldsymbol{\pi}_1 \in \mathbb{R}_1^{\Omega}} \mathbb{E}_{\mathbb{P}} \left( \min_{u_0 \in \mathbb{R}, u_1 \in \mathbb{R}} j(u_0, u_1, \mathbf{W}) + (\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi}_0 \mid \mathcal{F}_0) - \boldsymbol{\pi}_0) \cdot u_0 \right. \\ \left. + (\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi}_1 \mid \mathcal{F}_1) - \boldsymbol{\pi}_1) \cdot u_1 \right) . \end{aligned} \quad (4.22)$$

*Si les hypothèses ne sont pas vérifiées, le problème (4.22) fournit cependant une borne inférieure du problème (4.19).*

*Démonstration.* La preuve reprend celle de la Proposition 3.14 en posant  $\mathbf{U} = (\mathbf{U}_0, \mathbf{U}_1)$  et en posant  $\boldsymbol{\pi} = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1)$ .  $\square$

### 4.3.3 Extension de la résolution dans le cas avec risque

Les discussions concernant l'absence de produit scalaire naturel, les mesures de risque invariante en loi et l'écriture des contraintes de mesurabilité ont été faites dans le §3.4.3.

Le problème étudié est

$$\min_{\mathbf{U}_0 \in \mathbb{U}_0^{\Omega}, \mathbf{U}_1 \in \mathbb{U}_1^{\Omega}} \mathbb{F}(j(\mathbf{U}_0, \mathbf{U}_1, \mathbf{W})) . \quad (4.23a)$$

$$\text{s.c.} \begin{cases} \mathbb{E}_{\mathbb{P}}(\mathbf{U}_0 \mid \mathcal{F}_0) = \mathbf{U}_0 \\ \mathbb{E}_{\mathbb{P}}(\mathbf{U}_1 \mid \mathcal{F}_1) = \mathbf{U}_1 \end{cases} . \quad (4.23b)$$

On procède alors en deux étapes comme dans §3.4.3 en énonçant d'abord un résultat général avant de s'intéresser aux mesures de risque cohérentes.

**Proposition 4.5.** *Sous les hypothèses*

1. *l'ensemble  $\mathbb{U}$  est convexe,*

2. la mesure de risque  $\mathbb{F}$  est croissante, invariante par translation et convexe,  
3. la fonction  $j$  est conjointement convexe en  $u_0$  et  $u_1$  sur  $\mathbb{U}_0 \times \mathbb{U}_1$  pour tout  $w$ ,  
le problème (4.23) est équivalent à

$$\begin{aligned} \max_{\boldsymbol{\pi}_0 \in \mathbb{R}_0^\Omega, \boldsymbol{\pi}_1 \in \mathbb{R}_1^\Omega} \min_{\mathbf{U}_0 \in \mathbb{U}_0^\Omega, \mathbf{U}_1 \in \mathbb{U}_1^\Omega} \mathbb{F}(j(\mathbf{U}_0, \mathbf{U}_1, \mathbf{W}) + \mathbb{E}_{\mathbb{P}}\left(\left(\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi}_0 \mid \mathcal{F}_0) - \boldsymbol{\pi}_0\right) \cdot \mathbf{U}_0\right) \\ + \mathbb{E}_{\mathbb{P}}\left(\left(\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi}_1 \mid \mathcal{F}_1) - \boldsymbol{\pi}_1\right) \cdot \mathbf{U}_1\right)). \end{aligned} \quad (4.24)$$

Si les hypothèses ne sont pas vérifiées, le problème (4.24) fournit cependant une borne inférieure du problème (4.23).

*Démonstration.* La preuve reprend celle de la Proposition 3.15 en posant  $\mathbf{U} = (\mathbf{U}_0, \mathbf{U}_1)$  et en posant  $\boldsymbol{\pi} = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1)$ .  $\square$

**Proposition 4.6.** Soit  $\mathbb{F}$  est une mesure de risque cohérente de la forme (3.53) associée à l'ensemble de loi de probabilité  $\mathbb{Q}_{\mathbb{F}}$ . On se place sous les hypothèses de la Proposition 4.5.

Si de plus  $\mathbb{Q}_{\mathbb{F}}$  est un convexe fermé borné,  $j$  est conjointement fortement convexe en  $u_0$  et  $u_1$  sur  $\mathbb{U}_0 \times \mathbb{U}_1$  alors le problème (3.50) est équivalent à

$$\begin{aligned} \max_{\boldsymbol{\pi}_0 \in \mathbb{R}_0^\Omega, \boldsymbol{\pi}_1 \in \mathbb{R}_1^\Omega} \max_{\mathbb{Q} \in \mathbb{Q}_{\mathbb{F}}} \mathbb{E}_{\mathbb{P}}\left(\min_{u_0 \in \mathbb{U}_0, u_1 \in \mathbb{U}_1} j(u_0, u_1, \mathbf{W}) \frac{d\mathbb{Q}}{d\mathbb{P}} + \left(\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi}_0 \mid \mathcal{F}_0) - \boldsymbol{\pi}_0\right) \cdot u_0 \right. \\ \left. + \left(\mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi}_1 \mid \mathcal{F}_1) - \boldsymbol{\pi}_1\right) \cdot u_1\right). \end{aligned} \quad (4.25)$$

Si les hypothèses ne sont pas vérifiées, le problème (4.25) fournit cependant une borne inférieure du problème (4.23) et du problème (4.24).

*Démonstration.* La preuve reprend celle de la proposition 3.16 en posant  $\mathbf{U} = (\mathbf{U}_0, \mathbf{U}_1)$  et en posant  $\boldsymbol{\pi} = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1)$ .  $\square$

#### 4.3.4 Résumons nous

Nous avons étudié des problèmes d'optimisation stochastique dynamique dans lesquels les contraintes de mesurabilité étaient relaxées.

Après avoir étudié le cas risque neutre, nous nous sommes intéressés à des mesures de risque plus générales. Ceci nous a permis d'introduire la notion de mesure de risque dynamique comme suite de mesures de risque conditionnelles cohérentes temporellement.

## 4.4 Conclusion

Nous avons tout d'abord vu au cours de ce chapitre que les contraintes de mesurabilité d'un problème dynamique ne peuvent pas s'agréger en une seule contrainte et donc qu'on ne peut appliquer d'emblée les méthodes du cas statique.

Puis nous avons utilisé les résultats du cas statique pour énoncer des méthodes de décomposition par blocs de scénarios et par dualisation des contraintes dans un cadre dynamique.

Nous avons étendu les résultats connus de dualisation des contraintes dans le cas risques neutre à des mesures de risque plus générales.

# 5 Perspectives d'application

## 5.1 Introduction

Le problème que l'on étudie est une extension du problème de vendeur de journaux exposé au Chapitre 2 à un problème multi-vendeurs. On distingue  $K$  kiosques faisant chacun face à une demande aléatoire. Ces  $K$  kiosques s'approvisionnent auprès du même imprimeur qui ne peut imprimer qu'une quantité limitée de journaux.

Ces  $K + 1$  acteurs travaillent pour la même entreprise et agissent donc en équipe en essayant de maximiser leurs revenus globaux. Le problème est illustré sur la figure 5.1.

## 5.2 Modélisation

### Paramètres du problème

- temps discret  $t \in \{0, 1\}$
- des kiosques à journaux  $k \in \{1, \dots, K\}$ ,
- une demande aléatoire  $\mathbf{D}_k$  au kiosque  $k$  révélée au temps 1,
- une tribu donnée par  $\mathcal{F}_0 = \sigma(\emptyset, \Omega) \subset \mathcal{F}_1 = \sigma(\mathbf{D}_1, \dots, \mathbf{D}_K)$ ,
- Un imprimeur  $I$  à capacité maximale d'impression  $Q^{\max}$ ,
- le coût d'impression  $c^t$  d'un journal à l'instant  $t$ ,
- le prix de vente  $p_k$  d'un journal au kiosque  $k$ ,

### Variables de décision du problème

- la quantité à commander  $\mathbf{U}_k^t$  au kiosque  $k$  à l'instant  $t$ ,
- Pour plus de confort, on note

$$\mathbf{U}^t = \begin{pmatrix} \mathbf{U}_1^t \\ \vdots \\ \mathbf{U}_3^t \end{pmatrix}, \quad u^t = \begin{pmatrix} u_1^t \\ \vdots \\ u_3^t \end{pmatrix}. \quad (5.1)$$

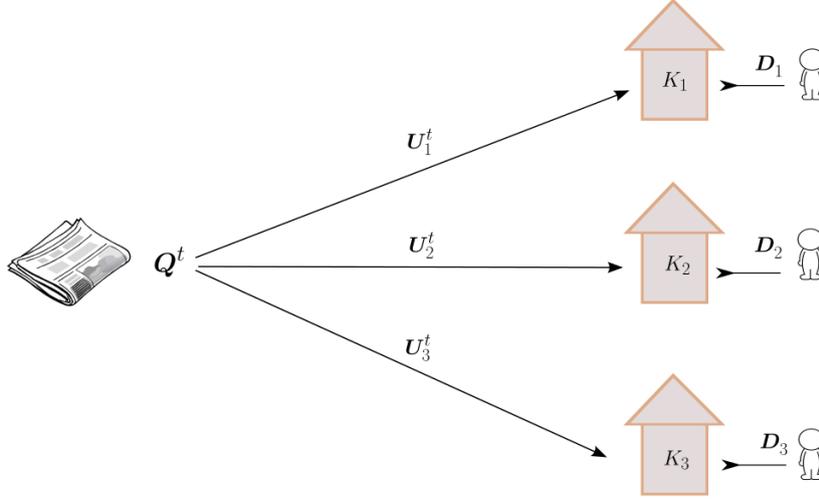


FIGURE 5.1 – Représentation schématique du problème de journaux multi-vendeurs

### Fonction objectif

$$\min_{\mathbf{U}_k^t \in (\mathbf{U}_k^t)^\omega} \mathbb{F} \left( c^0 \sum_k \mathbf{U}_k^0 - \sum_k [p_k \min\{\mathbf{U}_k^0, \mathbf{D}_k\}] \right) \quad (5.2)$$

$$+ c^1 \sum_k \mathbf{U}_k^1 - \sum_k [p_k \min\{\mathbf{U}_k^1, \mathbf{D}_k - \min\{\mathbf{D}_k, \mathbf{U}_k^0\}\}] \right) . \quad (5.3)$$

On différencie plusieurs types de contraintes

— contraintes couplantes sur la capacité de production

$$0 \leq \sum_k \mathbf{U}_k^t(\omega) \leq Q^{\max}, \quad \forall \omega \in \Omega, \quad \forall t \in \{0, 1\}, \quad (5.4)$$

— contraintes de mesurabilité

$$\mathbf{U}^t \preceq \mathcal{F}_t, \quad \forall t \in \{1, 2\}. \quad (5.5)$$

## 5.3 Perspectives de résolution

On prend comme mesure de risque  $\mathbb{F} = \text{CVAR}_\beta$ . D'après la proposition (4.6), et d'après la représentation duale (A.18) de la  $\text{CVAR}_\beta$ , on peut formuler le problème

d'optimisation

$$\begin{aligned} \max_{\boldsymbol{\pi}^0, \boldsymbol{\lambda}^0} \max_{\mathbb{Q} \in \mathcal{Q}_{\text{CVAR}_\beta}} \mathbb{E}_{\mathbb{P}} \left( \min_{u^0, u^1} \frac{d\mathbb{Q}}{d\mathbb{P}} \left( c^0 \sum_k u_k^0 - \sum_k [p_k \min\{u_k^0, \mathbf{D}_k\}] \right. \right. \\ \left. \left. + c^1 \sum_k u_k^1 - \sum_k [p_k \min\{u_k^1, \mathbf{D}_k - \min\{\mathbf{D}_k, u_k^0\}\}] \right) \right), \\ \left. + \sum_k \left[ \left( \mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi}_k^0) - \boldsymbol{\pi}_k^0 \right) u_k^0 \right] \right) \end{aligned} \quad (5.6a)$$

$$\text{s.c. } 0 \leq \sum_k \mathbf{U}_k(\omega) \leq Q^{\max}, \quad \forall \omega \in \Omega, \quad \forall t \in \{1, 2\}; \quad (5.6b)$$

$$\frac{\mathbb{Q}}{\mathbb{P}} \leq \frac{1}{1 - \beta}, \quad \forall \omega \in \Omega. \quad (5.6c)$$

À  $(\boldsymbol{\pi}^0, \boldsymbol{\lambda}^0, \mathbb{Q})$  fixés, on est amené à résoudre pour chaque  $\omega$  les sous problèmes

$$\begin{aligned} \min_{u^0, u^1} \frac{d\mathbb{Q}}{d\mathbb{P}}(\omega) \left( c^0 \sum_k u_k^0 - \sum_k [p_k \min\{u_k^0, \mathbf{D}_k(\omega)\}] \right. \\ \left. + c^1 \sum_k u_k^1 - \sum_k [p_k \min\{u_k^1, \mathbf{D}_k(\omega) - \min\{u_k^0, \mathbf{D}_k(\omega)\}\}] \right), \quad (5.7a) \\ \left. + \sum_k \left[ \left( \mathbb{E}_{\mathbb{P}}(\boldsymbol{\pi}_k^0) - \boldsymbol{\pi}_k^0(\omega) \right) u_k^0 \right] \right) \end{aligned}$$

$$\text{s.c. } 0 \leq \sum_k u_k^t \leq Q^{\max}. \quad (5.7b)$$

## 5.4 Conclusion

On a dans ce chapitre posé un problème multi-vendeurs. Dans ce genre de problèmes, une contrainte couplante lie les acteurs entre eux. On a utilisé une formulation fonctionnelle des contraintes de mesurabilité pour pouvoir les dualiser.

Notre programme de travail est d'utiliser les résultats des Chapitres 3 et la  $\text{CVAR}_\beta$  comme mesure de risque, pour décomposer le problème en plusieurs sous-problèmes que l'on peut résoudre de façon simultanée.

## 6 Conclusion

Au travers de ce rapport, nous avons étudié la décomposition des problèmes d'optimisation stochastique sous contraintes de mesurabilité ; soit en utilisant une méthode par imbrication, soit en utilisant une méthode de dualisation des contraintes de mesurabilité.

Nous avons vu comment changent les méthodes de résolution entre le cas statique et le cas dynamique. Nous avons étendus des résultats du cas risque neutre au cas sensible au risque en remplaçant l'espérance par une mesure de risque. Nous avons cherché les conditions pour reformuler des problèmes d'optimisation sensible au risque comme des problèmes d'optimisation linéaires, dans le but de faciliter leur résolution numérique.

Mes contributions personnelles dans ce rapport sont :

- des résultats numériques sur un “toy problem” permettant de saisir l'impact du risque et de l'information dans un problème d'optimisation stochastique,
- l'extension de la méthode par dualisation des contraintes en remplaçant l'espérance dans le critère par une mesure de risque cohérente, que ce soit dans un cadre statique ou dynamique,
- un résultat d'égalité entre un problème de maximum et un problème linéaire dans un cadre stochastique qui devra être étendu à des cas plus généraux que ceux présentés dans le rapport.

# A Appendices

## A.1 Rappels sur les mesures de risque

Les ouvrages de référence pour cet appendice sont [1], [7], [12]. On utilisera les notations suivantes :

1.  $(\Omega, \mathcal{F})$  est un espace mesurable,
2.  $\Delta(\Omega)$  est l'ensemble des lois de probabilités sur  $\Omega$ ,
3.  $\mathcal{X}$  est un sous-espace vectoriel de  $L^\infty$  qui contient les constantes ; c'est un ensemble de variables aléatoires  $\mathbf{X} : \Omega \rightarrow \mathbb{R}$  représentant une perte monétaire,
4.  $\mathbb{F} : \mathcal{X} \rightarrow \mathbb{R}$  est une mesure de risque.

### A.1.1 Définition et propriétés des mesures de risque monétaires

**Définition A.1.** Une mesure de risque  $\mathbb{F}$  est dite monétaire si

- $\mathbb{F}$  est *croissante* :  $\mathbf{X} \leq \mathbf{Y} \Rightarrow \mathbb{F}(\mathbf{X}) \leq \mathbb{F}(\mathbf{Y})$ ,
- $\mathbb{F}$  est *invariante par translation* :  $\forall m \in \mathbb{R}, \mathbb{F}(\mathbf{X} + m) = \mathbb{F}(\mathbf{X}) + m$ .

*Remarque A.2.* En optimisation, on cherche souvent à minimiser les coûts et la variable aléatoire  $\mathbf{X}$  représente une perte monétaire. Plus une variable aléatoire est grande, plus les pertes qu'elle représente sont grandes et plus la position est risquée à maintenir.

En économie, on cherche à maximiser les profits. Les mesures de risque sont alors décroissantes et l'invariance par translation s'écrit  $\mathbb{F}(\mathbf{X} + m) = \mathbb{F}(\mathbf{X}) - m$ . Cette définition est équivalente à celle présentée ci-avant dans le sens où un profit négatif est un coût positif et vice versa.

Si la mesure de risque  $\mathbb{F}$  est invariante par translation, quitte à translater toutes les variables aléatoires de  $m$ , on peut considérer que  $\mathbb{F}$  est normalisée c'est-à-dire  $\mathbb{F}(0) = 0$ .

La propriété d'invariance par translation est une propriété forte. En effet, cela impose des propriétés topologiques sur la mesure de risque. On a la proposition suivante.

**Proposition A.3.** Une mesure de risque monétaire  $\mathbb{F}$  est Lipschitz continue par rapport à la norme  $\|\cdot\|_\infty$

*Démonstration.*

$$\begin{aligned}
\mathbf{X} - \mathbf{Y} &\leq \|\mathbf{X} - \mathbf{Y}\|_\infty \\
\mathbf{X} &\leq \mathbf{Y} + \|\mathbf{X} - \mathbf{Y}\|_\infty \\
\mathbb{F}(\mathbf{X}) &\leq \mathbb{F}(\mathbf{Y} + \|\mathbf{X} - \mathbf{Y}\|_\infty) \quad \text{par croissance} \\
\mathbb{F}(\mathbf{X}) &\leq \mathbb{F}(\mathbf{Y}) + \|\mathbf{X} - \mathbf{Y}\|_\infty \quad \text{par invariance par translation} \\
\mathbb{F}(\mathbf{X}) - \mathbb{F}(\mathbf{Y}) &\leq \|\mathbf{X} - \mathbf{Y}\|_\infty
\end{aligned} \tag{A.1}$$

Par symétrie des rôles joués par  $\mathbf{X}$  et  $\mathbf{Y}$  on obtient :

$$|\mathbb{F}(\mathbf{X}) - \mathbb{F}(\mathbf{Y})| \leq \|\mathbf{X} - \mathbf{Y}\|_\infty . \tag{A.2}$$

ce qui conclut la preuve.  $\square$

**Définition A.4.** Une mesure de risque  $\mathbb{F}$  est dite

- *convexe* si  $\mathbb{F}(\lambda\mathbf{X} + (1 - \lambda)\mathbf{Y}) \leq \lambda\mathbb{F}(\mathbf{X}) + (1 - \lambda)\mathbb{F}(\mathbf{Y})$  pour  $\lambda \in [0; 1]$
- *positivement homogène* si  $\forall \lambda \geq 0, \mathbb{F}(\lambda\mathbf{X}) = \lambda\mathbb{F}(\mathbf{X})$

Une mesure de risque à la fois monétaire, convexe et positivement homogène, est dite *cohérente*.

**Proposition A.5.** Si  $\mathbb{F}$  est une mesure de risque monétaire positivement homogène, alors elle est convexe si et seulement si

- $\mathbb{F}$  est sous-additive c'est-à-dire  $\mathbb{F}(\mathbf{X} + \mathbf{Y}) \leq \mathbb{F}(\mathbf{X}) + \mathbb{F}(\mathbf{Y})$

## A.1.2 Exemples de mesures de risque

On introduit les notations pour la fonction de répartition

- $F_{\mathbf{X}}(x) = \mathbb{P}(\mathbf{X} \leq x)$
- $F_{\mathbf{X}}^{-1}(p) = \inf \{x | F_{\mathbf{X}}(x) \geq p\}$

Soit une variable aléatoire  $\mathbf{X}$  et  $\alpha$  un nombre entre 0 et 1. On peut définir les mesures de risque suivantes.

**Value at Risk :**

$$VaR_\alpha(\mathbf{X}) = \inf \{m \in \mathbb{R} | \mathbb{P}(\mathbf{X} \leq m) \geq \alpha\} . \tag{A.3a}$$

ce qui donne avec les notations adoptées :

$$VaR_\alpha(\mathbf{X}) = F_{\mathbf{X}}^{-1}(\alpha) . \tag{A.3b}$$

**Tail Conditional Expectation :**

$$TCE_\alpha(\mathbf{X}) = \mathbb{E}_{\mathbb{P}}(\mathbf{X} | \mathbf{X} \geq VaR_\alpha) . \tag{A.4}$$

**Mesure de risque entropique :**

$$RE_\alpha(\mathbf{X}) = \frac{1}{1 - \alpha} \log \left( \mathbb{E}_{\mathbb{P}}(e^{(1-\alpha)\mathbf{X}}) \right) . \tag{A.5}$$

**Worst Conditional Expectation :**

$$WCE_\alpha(\mathbf{X}) = \sup \{ \mathbb{E}_{\mathbb{P}}(\mathbf{X} | A); A \in \mathcal{F}, \mathbb{P}(A) \geq 1 - \alpha \} . \tag{A.6}$$

**Conditional Value at Risk :**

$$CVaR_\alpha(\mathbf{X}) = \inf s \in \mathbb{R} \left\{ \frac{\mathbb{E}[(\mathbf{X} - s)^+]}{1 - \alpha} + s \right\} . \quad (\text{A.7})$$

**Expected Shortfall :**

$$ES_\alpha(\mathbf{X}) = \frac{1}{1 - \alpha} \mathbb{E}[(\mathbf{X} - VaR_\alpha)^+] + VaR_\alpha . \quad (\text{A.8})$$

**Average Value at Risk :**

Pour tout  $\alpha$  dans  $[0,1)$

$$AVaR_\alpha(\mathbf{X}) = \frac{1}{1 - \alpha} \int_\alpha^1 VaR_p(\mathbf{X}) dp . \quad (\text{A.9})$$

et par extension, pour  $\alpha = 0$ , on utilise la mesure de risque Worst Case :

$$AVaR_1(\mathbf{X}) = \text{ess.sup}(\mathbf{X}) . \quad (\text{A.10})$$

**Théorème A.6.** *Avec les définitions ci-dessus, on a :*

$$CVaR_\alpha = ES_\alpha = AVaR_\alpha . \quad (\text{A.11})$$

**Proposition A.7.** —  $VaR_\alpha$ ,  $TCE_\alpha$  et  $RE_\alpha$  ne sont en général pas cohérentes.  
—  $WCE_\alpha$  et  $AVaR_\alpha$  sont cohérentes.

### A.1.3 Représentation duale des mesures de risque monétaires convexes

#### Cas général

On rappelle que le dual topologique de  $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$  est l'ensemble des mesures finiment additive absolument continues devant  $\mathbb{P}$  notée  $\mathcal{M}_{1,f}$ .

**Théorème A.8** (Théorème de représentation des mesures de risque monétaires).  
*Toute mesure de risque monétaire convexe  $\mathbb{F}$  sur  $\mathcal{X}$  est de la forme*

$$\mathbb{F}(\mathbf{X}) = \max_{\mathbb{Q} \in \mathcal{M}_{1,f}} (\mathbb{E}_{\mathbb{Q}}[X] - c_{\min}(\mathbb{Q})) . \quad (\text{A.12})$$

où la fonction de pénalité  $c_{\min}$  est donnée par

$$c_{\min}(\mathbb{Q}) := \sup_{\mathbf{X} \in \mathcal{A}_{\mathbb{F}}} \mathbb{E}_{\mathbb{Q}}[\mathbf{X}] . \quad (\text{A.13})$$

De plus, toute fonction de pénalité  $c$  telle que  $\mathbb{F}(\mathbf{X}) = \max_{\mathbb{Q} \in \mathcal{M}_{1,f}} (\mathbb{E}_{\mathbb{Q}}[X] - c(\mathbb{Q}))$  vérifie  $c(\mathbb{Q}) \geq c_{\min}(\mathbb{Q})$

*Remarque A.9.* La démonstration de ce théorème montre qu'on peut se restreindre aux mesures de **probabilités** absolument continues devant  $\mathbb{P}$

*Remarque A.10.* Si  $\mathbb{F}$  est convexe et semi-continue inférieurement alors on peut toujours se ramener au cas où  $c$  est convexe. En effet, d'après le théorème de Fenchel-Moreau, on a  $\mathbb{F}(\mathbf{X}) = \mathbb{F}^{**}(\mathbf{X})$  et donc

$$\mathbb{F}(\mathbf{X}) = \sup_{\mathbb{Q} \in \mathcal{M}_{1,f}} \{\mathbb{E}_{\mathbb{Q}}[\mathbf{X} - \mathbb{F}^*(\mathbb{Q})]\} . \quad (\text{A.14})$$

où  $\mathbb{F}^*$  est convexe et semi-continue inférieurement. On identifie alors  $c$  et  $\mathbb{F}^*$ .

**Théorème A.11** (Théorème de représentation des mesures de risque cohérentes.).

*La fonction de pénalité  $c_{\min}$  d'une mesure de risque cohérente  $\mathbb{F}$  prend uniquement les valeurs 0 ou  $+\infty$ . En particulier,*

$$\mathbb{F}(\mathbf{X}) = \max_{\mathbb{Q} \in \mathcal{Q}_{\max}} \mathbb{E}_{\mathbb{Q}}[\mathbf{X}] . \quad (\text{A.15})$$

avec l'ensemble convexe

$$\mathcal{Q}_{\max} := \{\mathbb{Q} \in \mathcal{M}_{1,f} \mid c_{\min} = 0\} . \quad (\text{A.16})$$

## Application aux exemples

**Proposition A.12.**

$$RE_{\alpha}(\mathbf{X}) = \sup_{\mathbb{Q} \in \mathcal{M}_1} \left\{ \mathbb{E}_{\mathbb{Q}}[\mathbf{X}] - \frac{1}{1-\alpha} H(\mathbb{Q}|\mathbb{P}) \right\} . \quad (\text{A.17})$$

où  $H(\mathbb{Q}|\mathbb{P}) := \mathbb{E}_{\mathbb{Q}}(\log \frac{d\mathbb{Q}}{d\mathbb{P}})$

**Proposition A.13.**

*L'ensemble  $\mathcal{Q}_{\max}$  associé à la mesure de risque AVaR $_{\alpha}$  est défini par :*

$$\mathcal{Q}_{\max} := \left\{ \mathbb{Q} \in \mathcal{M}_1(\mathbb{P}); \frac{d\mathbb{Q}}{d\mathbb{P}} \leq \frac{1}{1-\alpha} \right\} . \quad (\text{A.18})$$

*L'ensemble  $\mathcal{Q}_{\max}$  associé à la mesure de risque WCE $_{\alpha}$  est défini par :*

$$\mathcal{Q}_{\max} := \{ \mathbb{P}[\cdot | A]; A \in \mathcal{F}, \mathbb{P}[A] \geq 1 - \alpha \} . \quad (\text{A.19})$$

*Remarque A.14.* On remarque que, pour la AVaR, quand  $\alpha$  tend vers 1, on retrouve bien la mesure Worst Case qui est le supremum sur l'ensemble des lois de probabilités.

## A.1.4 Construction de mesures de risque

**Par sommation**

**Théorème A.15.**

*Si  $\mathbb{F}_{\alpha}$  est une famille de mesures de risque paramétrée par  $\alpha \in [a; b]$ , alors pour toute mesure  $d\mu(\alpha)$  sur  $[a; b]$  vérifiant  $\int_a^b d\mu(\alpha) = 1$ , la fonctionnelle définit par  $\mathbb{F} = \int_a^b \mathbb{F}_{\alpha} d\mu(\alpha)$  est une mesure de risque.*

**Corollaire A.16.**

$$M_\mu(\mathbf{X}) := \int_0^1 \alpha AVaR_\alpha(\mathbf{X}) d\mu(\alpha) . \quad (\text{A.20})$$

est une mesure de risque tant que  $\int_0^1 \alpha d\mu(\alpha) = 1$

**Exemple A.17.** Si on prend la mesure  $d\mu(\alpha) = -\delta'(\alpha - \beta)d\alpha$ , on a  $VaR_\beta(\mathbf{X}) = M_\mu(\mathbf{X})$ . Ici,  $\delta'$  désigne l'objet formel qu'est la dérivée première de la mesure de dirac et dont la définition formelle peut être donné à partir d'une intégration par parties.

**Par majoration et minoration**

**Théorème A.18.** Soit  $(\mathbb{F}_i)_{i \in I}$  une famille de mesures de risque convexes et monétaires de fonction de pénalité respective  $c_i$ . Si  $\sup_{i \in I} \mathbb{F}_i(\mathbf{X}) < +\infty$  alors

$$\mathbb{F}(\mathbf{X}) := \sup_{i \in I} \mathbb{F}_i(\mathbf{X}) , \quad (\text{A.21})$$

définit une mesure de risque monétaire convexe avec fonction de pénalité  $c(\mathbb{Q}) := \inf_{i \in I} c_i(\mathbb{Q})$

**Proposition A.19.**

$$VaR_\alpha(\mathbf{X}) = \min\{\mathbb{F}(\mathbf{X}); \mathbb{F} \text{ cohérente et sci telle que } \mathbb{F} \geq VaR_\alpha\} . \quad (\text{A.22})$$

## A.1.5 Mesures de risque dynamiques

On va à présent introduire de nouveaux concepts afin de pouvoir traiter des problèmes dynamiques. On se donne pour cela :

1.  $\{\emptyset, \Omega\} = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \dots \subset \mathcal{F}_T = \mathcal{F}$  une filtration de l'espace  $\Omega$
2.  $\mathcal{X}_t$  est l'ensemble des variables aléatoires  $\mathcal{F}_t$  – mesurable
3.  $\mathcal{X}_{s,t} := \mathcal{X}_s \times \mathcal{X}_{s+1} \times \dots \times \mathcal{X}_t$
4.  $\mathbf{X}_{s,t}$  désigne un élément de  $\mathcal{X}_{s,t}$
5. Le relation d'ordre  $\mathbf{X}_{s,t} \leq \mathbf{Y}_{s,t}$  désigne la relation d'ordre presque sûre composable par composante.

**Définition A.20.** Une mesure de risque  $\mathbb{F}_{s,t} : \mathcal{X}_{s,t} \rightarrow \mathcal{X}_s$  est dite conditionnelle si elle est monotone c'est-à-dire si

$$\mathbf{X}_{s,t} \leq \mathbf{Y}_{s,t} \Rightarrow \mathbb{F}_{s,t}(\mathbf{X}_{s,t}) \leq \mathbb{F}_{s,t}(\mathbf{Y}_{s,t}) . \quad (\text{A.23})$$

**Définition A.21.** Une mesure de risque dynamique  $\varrho_{\mathcal{T},T}$  est une suite de mesures de risque conditionnelles  $\mathbb{F}_{t,T}$  où  $t$  varie de manière croissante dans un ensemble  $\mathcal{T} \subset \{1, \dots, T\}$  non vide.

**Définition A.22.** La mesure de risque dynamique  $\varrho_{\{1,3,4\},8}$  désigne la séquence  $(\mathbb{F}_{1,8}, \mathbb{F}_{3,8}, \mathbb{F}_{5,8})$

*Remarque A.23.* On peut définir des mesures de risque conditionnelles à partir d'une mesure de risque dynamique par

$$\mathbb{F}_{s,t}(\mathbf{X}_s, \dots, \mathbf{X}_t) := \mathbb{F}_{s,T}(\mathbf{X}_s, \dots, \mathbf{X}_t, 0, \dots, 0) . \quad (\text{A.24})$$

**Définition A.24.** Une mesure de risque dynamique  $\rho_{\mathcal{T}}$  est dite temporellement cohérente si pour tout  $1 \leq s < t \leq T$  et tout  $\mathbf{X}_{s,T}, \mathbf{Y}_{s,T}$ , on a :

$$\begin{aligned} & \forall k \in s, \dots, t-1, X_k = Y_k \\ & \text{et} \\ & \mathbb{F}_{t,T}(\mathbf{X}_t, \dots, \mathbf{X}_T) \leq \mathbb{F}_{t,T}(\mathbf{Y}_t, \dots, \mathbf{Y}_T) \quad . \\ & \quad \downarrow \\ & \mathbb{F}_{s,T}(\mathbf{X}_s, \dots, \mathbf{X}_T) \leq \mathbb{F}_{s,T}(\mathbf{Y}_s, \dots, \mathbf{Y}_T) \end{aligned} \quad (\text{A.25})$$

**Définition A.25.** Une mesure de risque conditionnelle à un pas de temps est une mesure de la forme  $\mathbb{F}_{t,t+1}$ .

Une telle mesure peut être :

- invariante par translation : si  $\mathbf{M} \in \mathcal{X}_s$  et  $\mathbf{X} \in \mathcal{X}_t$  alors  $\mathbb{F}_{s,t}(\mathbf{X} + \mathbf{M}) = \mathbb{F}_{s,t}(\mathbf{X}) + \mathbf{M}$
- convexe :  $\forall \mathbf{X}, \mathbf{Y} \in \mathcal{X}_{t+1}$ ,  $\mathbb{F}_t(\lambda \mathbf{X} + (1-\lambda)\mathbf{Y}) \leq \lambda \mathbb{F}_t(\mathbf{X}) + (1-\lambda)\mathbb{F}_t(\mathbf{Y})$
- positivement homogène : si  $\lambda \geq 0$ ,  $\mathbb{F}_t(\lambda \mathbf{X}) = \lambda \mathbb{F}_t(\mathbf{X})$

*Remarque A.26.* De même que pour les mesures de risque classiques, à partir des définitions (A.20), on peut définir des mesures de risque conditionnelles monétaires, convexe, positivement homogènes et cohérentes

On va à présent étudier les représentations duales des mesures de risque conditionnelles cohérentes. Pour cela on associe à une mesure de risque conditionnelle une autre fonction :

$$\mathbb{F}_{s,t}^{\omega}(\mathbf{X}) := [\mathbb{F}_{s,t}(\mathbf{X})](\omega), \mathbf{X} \in \mathcal{X}_t . \quad (\text{A.26})$$

**Définition A.27.** On note  $\mathcal{Q}_t$  l'ensemble des lois de probabilités sur  $\mathcal{X}_t$ . Pour chaque  $\omega$ , on définit un ensemble de mesures de probabilités noté  $\mathcal{Q}_{s,t}(\omega)$  défini comme l'ensemble des éléments  $\mathbb{Q}'$  qui vérifient pour tout  $A \in \mathcal{F}_t$

$$\mathbb{Q}' = 1 \text{ si } \omega \in A \text{ et } 0 \text{ sinon} . \quad (\text{A.27})$$

**Théorème A.28** (Représentation des mesures de risque conditionnelles cohérentes). *Soit  $\mathbb{F}_{s,t}$  une mesure de risque conditionnelle monétaire, convexe et sci. Alors :*

$$\mathbb{F}_{s,t}^{\omega}(\mathbf{X}) = \sup_{\mathbb{Q} \in \mathcal{Q}_{s,t}^{\omega}} \mathbb{E}_{\mathbb{Q}}[\mathbf{X} - (\mathbb{F}^{\omega})^*(\mathbb{Q})] . \quad (\text{A.28})$$

*Réciproquement, une fonction  $\mathbb{F} : \mathcal{X}_t \rightarrow \mathcal{X}_s$  qui peut être écrit sous cette forme pour une fonction  $(\mathbb{F}^{\omega})^*$  propre, est une mesure de risque conditionnelle monétaire convexe sci.*

## A.2 Rappels d'analyse convexe

Les ouvrages de référence pour ce chapitre sont [5] et [10].

**Théorème A.29.** *Soit le problème :*

$$\min_{u \in \mathbb{U}^{ad}} J(u) . \quad (\text{A.29})$$

Si

1.  $J$  est une fonction convexe, s.c.i et coercive sur  $\mathbb{U}^{ad}$
2.  $\mathbb{U}^{ad}$  est convexe et fermé
3.  $\text{dom}(J) \cap \mathbb{U}^{ad} \neq \emptyset$

alors

*Il existe au moins une solution au problème (A.29).*

*L'ensemble des solutions est un sous-ensemble convexe fermé.*

*Si  $J$  est strictement convexe, la solution est alors unique.*

**Théorème A.30.** *Soit  $f : \mathcal{X} \rightarrow \bar{\mathbb{R}}$  une fonction convexe qui est finie et continue en au moins un point  $x$ . Alors le sous-différentiel de  $f$  est non vide en tout point de l'intérieur de son domaine, lui-même non vide, et en particulier en  $x$ .*

**Définition A.31.** Pour une fonction  $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R}$  et des sous-ensembles convexes fermés  $X \subset \mathcal{X}$  et  $Y \subset \mathcal{Y}$ , on dit que  $(x^\#, y^\#) \in X \times Y$  est un de point selle de  $f$  sur  $X \times Y$  si

$$\forall x \in X, \forall y \in Y, f(x^\#, y) \leq f(x^\#, y^\#) \leq f(x, y^\#) . \quad (\text{A.30})$$

**Proposition A.32.** *Le point  $(x^\#, y^\#)$  est un point selle de  $f$  sur  $X \times Y$  si et seulement si*

$$f(x^\#, y^\#) = \sup_{y \in Y} f(x^\#, y) = \min_{x \in X} \sup_{y \in Y} f(x, y) = \max_{y \in Y} \inf_{x \in X} f(x, y) = \inf_{x \in X} f(x, y^\#) . \quad (\text{A.31})$$

**Théorème A.33** (Théorème d'interversion min-max). *Soit  $f : X \times Y \rightarrow \mathbb{R}$  une fonction convexe-concave, sci-scs. Si*

- *l'ensemble  $X$  est un convexe compact*
- *l'ensemble  $Y$  est un convexe*
- *les marginales sont continues*

alors

$$\min_{x \in X} \sup_{y \in Y} f(x, y) = \sup_{y \in Y} \min_{x \in X} f(x, y) \quad (\text{A.32})$$

**Théorème A.34** (Théorème général d'existence de point selle).

*Soit  $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$  une fonction convexe-concave sci-scs. Soit  $X$  et  $Y$  deux sous-ensembles convexes fermés de  $\mathcal{X}$  et  $\mathcal{Y}$ , respectivement. On suppose aussi qu'il existe  $x^\circ \in X$  tel que*

$$\lim f(x^\circ, y) = -\infty \text{ lorsque } \|y\| \rightarrow +\infty \text{ et } y \in Y . \quad (\text{A.33a})$$

*et qu'il existe  $y^\circ \in Y$  tel que*

$$\lim f(x, y^\circ) = +\infty \text{ lorsque } \|x\| \rightarrow +\infty \text{ et } x \in X . \quad (\text{A.33b})$$

*ou bien  $Y$ , respectivement  $X$ , est borné. Alors, il existe un point selle de  $f$  sur  $X \times Y$*

On donne à présent les rappels concernant les conditions de qualification de contraintes. Pour ce rappel, on décrit l'ensemble  $\mathbb{U}^{ad}$  des solutions admissibles par un ensemble de contraintes égalités et inégalités notées :

- $\Theta_i(u) \leq 0, i = 1, \dots, m$
- $\Psi_j(u) = 0, j = 1, \dots, p$
- On peut écrire les contraintes égalités  $A(u) = b$  avec  $b \in \mathbb{R}^p$

On peut alors définir un cône analytique de la manière suivante :

$$K(u) := \sum_{i=1}^m \lambda_i \partial \Theta_i(u) + A^* \mu, \lambda \in \Lambda(u), \mu \in \mathbb{R}^p . \quad (\text{A.34a})$$

avec

$$\Lambda(u) := \{ \lambda \in \mathbb{R}^m \mid \lambda_i \Theta_i(u) = 0, i = 1, \dots, m \} . \quad (\text{A.34b})$$

**Proposition A.35.** *Avec les définitions ci-dessus, on a  $K(u) \subset (U^{ad})_u^\perp$ . L'hypothèse d'égalité est appelée hypothèse de qualification des contraintes.*

Soit le problème :

$$\min_{u \in U^{ad}} J(u) . \quad (\text{A.35a})$$

$$\Theta(u) \in -C . \quad (\text{A.35b})$$

On définit le Lagrangien élargi de ce problème par :

$$L(u, \lambda) \stackrel{\text{def}}{=} \begin{cases} J(u) + \langle \lambda, \Theta(u) \rangle & \text{si } \lambda \in C^* \\ -\infty & \text{sinon} \end{cases} . \quad (\text{A.36})$$

**Théorème A.36.** *On suppose qu'il existe une solution au problème (A.35) et on fait l'hypothèse de qualification des contraintes suivantes :*

$$0 \in \overbrace{\Theta(U^{ad})}^{\circ} + C . \quad (\text{A.37})$$

alors le Lagrangien du problème (A.35) admet un point selle et il existe des multiplicateurs optimaux associés à une solution optimale.

**Corollaire A.37.**

Si

1.  $J$  est convexe, sci
2.  $\Theta$  est  $C$ -convexe et continue
3.  $U^{ad}$  est borné ou  $J$  est coercive en  $u$  sur  $U^{ad}$
4. la condition de qualification des contraintes (A.37) est vérifiée

alors le Lagrangien  $L$  admet un point selle.

**Définition A.38.** On dit que l'application  $\Theta : \mathcal{U} \rightarrow \mathcal{V}$  est  $C$ -convexe si pour tout  $u$  et  $v$  de  $\mathcal{U}$  et tout  $\alpha$  dans  $[0;1]$  on a

$$\Theta(\alpha u + (1 - \alpha)v) - \alpha \Theta(u) - (1 - \alpha) \Theta(v) \in -C \quad (\text{A.38})$$

# Bibliographie

- [1] Carlo Acerbi. Spectral measures of risk : a coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26(7) :1505–1518, 2002.
- [2] Alejandro Balbás, José Garrido, and Silvia Mayoral. Coherent risk measures in a dynamic framework. *Preliminary version*, 2002.
- [3] Pierre Carpentier, Jean-Philippe Chancelier, Guy Cohen, and DE Michel. *Stochastic Multi-Stage Optimization : At the Crossroads between Discrete Time Stochastic Control and Stochastic Programming*, volume 75. Springer, 2015.
- [4] Xin Chen, Melvyn Sim, David Simchi-Levi, and Peng Sun. Risk aversion in inventory management. *Operations Research*, 55(5) :828–842, 2007.
- [5] Ivar Ekeland and Roger Temam. Convex analysis and 9 variational problems. 1976.
- [6] Csaba I Fábián. Handling cvar objectives and constraints in two-stage stochastic models. *European Journal of Operational Research*, 191(3) :888–911, 2008.
- [7] Hans Föllmer and Alexander Schied. *Stochastic finance : an introduction in discrete time*. Walter de Gruyter, 2011.
- [8] R Tyrrell Rockafellar and Stanislav Uryasev. Optimization of conditional value-at-risk. *Journal of risk*, 2 :21–42, 2000.
- [9] R Tyrrell Rockafellar and Roger J-B Wets. *Variational analysis*, volume 317. Springer Science & Business Media, 2009.
- [10] Ralph Tyrell Rockafellar. *Convex analysis*. Princeton university press, 2015.
- [11] RT Rockafellar and RJB Wets. Variational analysis. 1998.
- [12] Andrzej Ruszczyński. Risk-averse dynamic programming for markov decision processes. *Mathematical programming*, 125(2) :235–261, 2010.
- [13] Andrzej Ruszczyński and Alexander Shapiro. Conditional risk mappings. *Mathematics of operations research*, 31(3) :544–561, 2006.
- [14] Alexander Shapiro. Minimax and risk averse multistage stochastic programming. *European Journal of Operational Research*, 219(3) :719–726, 2012.
- [15] C Strugarek. Interchange of minimization and integration with measurability constraints. *Preprint CERMICS <http://cermics.enpc.fr/reports>*, 321, 2006.