UNIVERSITÀ DEGLI STUDI DI PADOVA

Sede Amministrativa: Università degli Studi di Padova
Dipartimento di Matematica Pura e Applicata

DOTTORATO DI RICERCA IN MATEMATICA
CICLO XIII

Robust and adaptive strategies for shortfall risk minimization under model uncertainty

Coordinatore: Ch.mo Prof. Alberto Facchini
Supervisore: Ch.mo Prof. Wolfgang Runggaldier

Dottorando: Gino Favero

To all people
who love me

A tutti quelli
che mi vogliono bene

Contents

Introduzione – Introduction **7**

1 Notations and generalities **13**
1.1 The stochastic optimal control problem 13
1.2 Model uncertainty 18

2 An overview of literature **23**
2.1 The minimax approach 23
2.2 The bayesian approach for the binomial market model 25
2.3 Robustness for diffusion processes 26
2.4 Robustness for Markov control processes 28

3 Our results **31**
3.1 General robustness results for stochastic control 31
3.2 Robust and adaptive approaches for the binomial model . . . 36
3.3 Extensions of the previous results 44

Bibliography **51**

Appendix: enclosed papers

G. Favero, W. J. Runggaldier, *A Robustness Result for Stochastic Control*, Università di Padova, Dipartimento di Matematica Pura e Applicata, Preprint n. 21 (2000) . . . . (10)

G. Favero, *Shortfall risk minimization under model uncertainty in the binomial case: adaptive and robust approaches*, Math. Meth. Oper. Res. 53, 493-503 (2001) . . . . (10)

G. Favero, T. Vargiolu, *Shortfall risk minimization in the binomial model* (temporary title), work in progress . . . . (14)

Introduzione

Molte applicazioni della matematica possono essere viste come situazioni in cui un utente cerca di raggiungere una situazione “desiderata” intervenendo su un sistema che evolve disturbato da eventi incontrollabili e imprevedibili. I modelli di questo tipo sono detti *problemi stocastici controllati* e sono studiati da molti anni (si veda, per esempio, [9]).

In generale, un problema stocastico controllato consiste di una data dinamica di stato e di una funzione obiettivo. La *dinamica* descrive l'evoluzione del sistema considerato in termini del *controllo* scelto dall'utente e dei *disturbi*, che sono simulati da processi stocastici su qualche spazio di probabilità soggiacente. La *funzione costo* (o *funzione ricavo*) è qualche funzionale (dipendente, in generale, dallo stato del sistema e dalla scelta dei controlli) che può rappresentare, per esempio, una perdita (o un introito) che l'utente cerca di minimizzare (o massimizzare); la *funzione valore* è quindi definita tramite un opportuno funzionale della funzione costo (o ricavo) in modo da rispecchiare le priorità dell'utente. L'utente è a questo punto interessato a determinare la *politica ottima*, cioè a scegliere i controlli, basandosi sulla conoscenza dell'evoluzione passata del sistema, in modo da massimizzare (o minimizzare) la funzione valore rispetto a tutti i controlli possibili.

Tuttavia, è pressoché inevitabile che nella costruzione di un modello matematico non possano essere presi in considerazione tutte le componenti del problema. Per esempio, alcuni termini dinamici potrebbero non essere compresi nel modello, o la distribuzione di probabilità dei disturbi potrebbe essere conosciuta solo fino a una certa approssimazione, e così via. In questi casi si dice comunemente che si ha un'“informazione incompleta” sul modello.

In letteratura, sono stati proposti molti modi per affrontare la situazione di informazione incompleta. Per esempio, si potrebbe decidere di scegliere la politica ottima in modo da ottenere la migliore prestazione possibile contro quello che si crede possa essere il peggior caso possibile: questo è il cosiddetto approccio *minimax* (studiato, per esempio, in [1], [3], [18] o, in forma più generale, in [4], [16]). In alternativa, si potrebbe cercare di formulare il modello in modo da dare agli elementi incerti una struttura di elementi aleatori su un opportuno spazio di probabilità, aggiornandone la distribuzione a mano a mano che con il passare del tempo aumenta la quantità di informazioni sul loro conto: questo è il cosiddetto approccio *bayesiano* (si veda, per esempio, [17]).

Un altro possibile modo per prendere in considerazione l'incertezza è il cosiddetto approccio *robusto* (si vedano, per esempio, [2], [10], [11], [12]), che è quello esaminato nella presente tesi. L'idea è quella di includere il problema stocastico controllato (P) con cui si ha a che fare in una classe $\mathcal{Q}$ di problemi e di dimostrare che la politica ottima π^Q per un particolare problema (Q) in questa classe dà, in qualche senso, risultati accettabili quando viene usata per controllare qualsiasi altro sistema in $\mathcal{Q}$ e, in particolare, anche il "vero" sistema (P) (il problema (Q), in letteratura, è chiamato "problema nominale" o, talvolta, "approssimante"). In sostanza, si stima per ogni problema $(R) \in \mathcal{Q}$ la differenza tra il valore ottimo del problema (R) e il valore che si ottiene quando (R) è controllato con la strategia π^Q; si definisce poi l'*indice di robustezza* come l'estremo superiore di tutte queste differenze. Lo scopo dell'approccio robusto è allora quello di ottenere una *disuguaglianza di robustezza*, cioè una quantità superiore all'indice di robustezza calcolabile a partire dalla proprietà note della classe $\mathcal{Q}$: questa quantità è allora anche un limite superiore della prestazione che si ottiene usando il controllo π^Q nel problema (P). Si noti che l'approccio robusto può essere utilizzato anche quando una formulazione "completa" del modello, cioè una formulazione che prenda in considerazione tutti gli aspetti del modello stesso, risultasse intrattabile dal punto di vista teorico o computazionale. In generale, allora (come si osserva, per esempio, in [2]), questo approccio si può anche usare per approssimare il controllo ottimo del problema (P) quando il controllo ottimo stesso non è esplicitamente calcolabile. Da questo punto di vista, l'approccio robusto ammette quindi una formulazione "duale": ci si può, cioè, chiedere quale sia il livello di precisione che si deve avere nella formulazione del modello per riuscire a ottenere un "controllo ottimo approssimato" che fornisca la desiderata precisione nella prestazione.

Il capitolo 1 stabilisce le notazioni che saranno usate nel corso dell'intero testo. Nella sezione 1.1 si definisce il problema stocastico controllato in entrambe le situazioni considerate classicamente, cioè sia quando si considera il tempo come un parametro continuo sia quando si considera l'evoluzione del sistema in un insieme discreto di istanti prefissati. Si dà particolare enfasi al caso in tempo discreto con orizzonte finito, che è quello studiato nelle nostre ricerche, e si introduce brevemente l'algoritmo della programmazione dinamica, uno strumento classico che si può impiegare per ottenere soluzioni esplicite. La sezione 1.2 richiama gli approcci al caso di informazione incompleta di cui abbiamo parlato sopra. Si espone anche brevemente il problema di robustezza in un'accezione generale, e si evidenzia come l'approccio robusto può esserne dedotto.

Il capitolo 2 raccoglie alcuni risultati riguardanti gli approcci suddetti, che abbiamo tratto dalla letteratura disponibile all'inizio della nostra ricerca. La sezione 2.1 è una panoramica dell'approccio minimax come è formulato in [3] e ripreso in [4] e [16]; sono anche brevemente esposti i risultati di [1] e [18], esempi di applicazione dell'approccio minimax a un particolare modello. Nella sezione 2.2 si riprende l'articolo [17] che, oltre a essere il punto di partenza dei nostri articoli [5] e [7], contiene anche il risultato dell'applicazione a un particolare modello dell'approccio bayesiano. Le sezioni 2.3 e 2.4 sono dedicate a alcuni risultati ottenuti attraverso l'approccio robusto: si riportano, rispettivamente, i risultati di [10] e [2] nel caso a tempo continuo e quelli di [11] e [12] nel caso a tempo discreto. Queste due sezioni contengono anche alcune osservazioni che possono essere applicate alle conclusioni ottenute da noi in [6].

I risultati della nostra ricerca sono raccolti nel capitolo 3. La sezione 3.1 riporta i teoremi di robustezza contenuti in [6]. Supponiamo qui di avere un'informazione parziale sulla legge di probabilità P dei disturbi del nostro problema, cioè, in dettaglio, di sapere solamente che P è in qualche modo una "perturbazione" di una legge Q nota. Suggeriamo quindi due possibili definizioni per dei problemi "approssimanti" e deduciamo una disuguaglianza di robustezza per ciascuno di essi. Le due disuguaglianze ottenute sono poi confrontate tra loro e con i risultati di [11] e [12]. La sezione 3.2 nasce dall'applicazione dell'approccio robusto di [6] a un caso particolare preso dalla finanza matematica, cioè alla minimizzazione dello scoperto medio nel celebre "modello binomiale di mercato" proposto da Cox, Ross e Rubinstein. L'esame di questo modello ha portato alla scoperta di alcune sue notevoli proprietà che hanno giustificato un articolo autonomo ([5]) e permesso di elaborare una strategia "adattativa" nel caso di informazione incompleta. La sezione 3.3 e [7] raccolgono la ricerca attualmente in evoluzione, mirata a estendere i risultati di [5] al problema più generale di minimizzazione del rischio di scoperto nel mercato binomiale.

Gli articoli [6, 5, 7] sono allegati per comodità del lettore.

Introduction

Many applications of mathematics can be formulated in terms of a user trying to reach a "desirable" situation by controlling an evolving system which, in turn, is disturbed by some uncontrollable and unforeseeable events. Models of this kind are called *stochastic control problems* and have been widely studied for many years (see [9] as an example).

In general, a stochastic control problem consists of a given state dynamics and of a value function. The *dynamics* describes the evolution of the system we are dealing with in terms of the *control* chosen by the user and of the *disturbances* which are supposed to be stochastic processes on some underlying probability space. The *cost* (or *gain*) *function* is some functional (generally depending on the state of the system and of the controls) that may represent, for instance, a loss (or a gain) that the user wants to minimize (or maximize). A *value function* is then defined as some functional of the cost (gain) function which translates the user preferences. The user is interested in determining an *optimal policy*, that is, a choice of the controls based on the past evolution of the system so as to minimize (or maximize) the value function with respect to all the possible controls.

When constructing a model it is nevertheless almost inevitable that some of its aspects cannot be taken into consideration. For example, some dynamics might not be included in the model, or the distribution of the disturbances might be known only up to a certain degree of accuracy, and so on. This is what is commonly referred to as an "incomplete information" about the system.

In the literature, many ways have been studied to approach the situation when there is incomplete information. For example, one might decide to choose the optimal policy in such a way as to get the best possible performance against what is believed to be the worst possible case: this is the so-called *minimax* approach (applied, e.g., in [1], [3], [18] or, in a more general form, in [4], [16]). As an alternative, one might want to formulate the model in such a way to give the sources of uncertainty a structure of random elements on a suitable probability space and to update their distribution as more information becomes available with time: this is the so-called *bayesian* approach (see, e.g., [17]).

Another possible way to take into account the uncertainty is the so-called *robust* approach (see, e.g., [2], [10], [11], [12]), which is the setting of our work. The idea is to include the stochastic control problem (P) we are dealing with (and that we call the "real" system) in a whole class $\mathcal{Q}$ of problems, and show that the optimal policy π^Q for one chosen problem (Q) in this class (that in the literature is called the "nominal" or sometimes the "approximating" problem) gives, in some sense, acceptable results when used to control every other system in $\mathcal{Q}$ and, in particular, also the real system. For every system $(R) \in \mathcal{Q}$ one looks at the gap between the optimal value of the problem (R) and the value obtained when (R) is controlled by means of π^Q, and defines the *robustness index* as the supremum of all these gaps. The aim of the robustness approach is then to give a *robustness inequality*, i.e., an upper bound for the robustness index based on the (known) properties of the class $\mathcal{Q}$, which also measures the performance of π^Q when it is applied to the problem (P). Note that the robust approach can also be applied to the case in which a "complete" formulation of the problem, taking into account all the information, would result intractable either from a computational or a theoretical point of view. In general, then, (as observed, e.g., in [2]) this approach can also be seen as a way to approximate the optimal policy of the problem (P) when it is not explicitly computable. From this point of view, then, the robust approach also admits a "dual" formulation, that is, it allows to know the degree of precision needed in the formulation of the "approximating" model (Q) in order to obtain an "approximating optimal control" performing with the desired degree of precision.

Chapter 1 sets the notations that will be used throughout the entire text. In Section 1.1 the stochastic optimal control problem is defined in both the continuous time and discrete time settings, with particular emphasis on the discrete time, finite horizon case which is the setting of our research. In this situation, the main tool used to obtain explicit solutions is the dynamic programming algorithm, which is also introduced. In section 1.2 the mentioned approaches to the case when there is incomplete information about the model are recalled. It is also emphasized how the robust approach can be deduced from the more general robustness problem, which is briefly explained.

Chapter 2 gathers some results, which were available in literature when our research began, about all the approaches mentioned above to the incomplete information case. Section 2.1 is an overwiew of the minimax approach as formulated in [3] and extended in [4] and [16]. As an application of the minimax approach to a particular model, [1] and [18] are also briefly summarized here. Section 2.2 introduces the setting of [17] which, besides being a starting point for our papers [5] and [7], features also the application of the bayesian approach to a particular model. Sections 2.3 and 2.4 are dedicated to some results obtained by means of the robust approach reporting, respectively, results of [10] and [2] for the continuous time setting and results of [11] and [12] for the discrete time setting. These two sections also feature some remarks that can be applied to the conclusions we obtain in [6].

Our results are reported in chapter 3. In section 3.1 the robustness theorems gathered in [6] are reported. We suppose to have partial information about the probability law P of the disturbances of our problem, namely, that we know P to be in some sense a "perturbation" of a known probability law Q. We then suggest two possible definitions for an "approximating" problem, and deduce a robustness inequality for each of them. The inequalities obtained are then compared to each other and to the results of [11] and [12]. Section 3.2 arises from the application of the robust approach of [6] to a particular case taken from mathematical finance, namely, the mean shortfall minimization problem in the well-known Cox, Ross and Rubinstein "binomial market model". The investigations about this model led to the discovery of some noteworthy properties of the model itself that deserved a paper on its own (namely, [5]), and also allowed to formulate an "adaptive" strategy for the incomplete information case. Section 3.3 and [7] gather the research currently in progress, aimed at the extension of the results of [5] to the more general setting of the shortfall risk minimization problem.

Papers [6, 5, 7] are also included for the convenience of the reader.

Chapter 1

Notations and generalities

This chapter is intended to be a reference for all the sequel and is divided into two parts. In 1.1 the standard notations for the stochastic control problem are described, with particular emphasis on the discrete time, finite horizon case which is the setting of our work. The so-called “robust approach” is specified in 1.2, where the key concepts of robustness index and robustness inequality are also defined.

1.1 The stochastic optimal control problem

Suppose that we are dealing with a system whose state can be described as an element of a suitable space X (typically, X is some subset of the standard d-dimensional real space $\mathbb{R}^d$). We shall write the state of the system as $x_t \in X$, $t \in \mathbb{R}^+$ to stress the dependence on time, and we want to imagine that the evolution of the state of this system, defined by a suitable *dynamics*, can be affected by two components: a *control* π_t chosen by the user and some *disturbance* ξ_t with stochastic nature. The control and the disturbances are supposed to take values in sets A and Ξ respectively, which can also be supposed to be subsets of some standard multi-dimension real spaces with suitable dimensions d' and d'' (possibly different from each other and from d). In general, the sets A and Ξ can be supposed to be dependent on time or on the current state but, for simplicity, we shall suppose them to be given and the same for all states and times.

Two different settings are now possible, namely, in continuous time or in discrete time, depending on whether one supposes the evolution of the involved processes to be observed continuously or only at particular moments. Moreover, both of these families are divided into the two subclasses of “infinite horizon” and “finite horizon” problems, depending on whether one is interested in monitoring the evolution of the system forever or just until a final time T, which is indeed called the *horizon* of the problem.

Continuous time problems

In the setting of *infinite horizon, continuous time control problems*, the underlying probability space is a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, P)$ whose filtration $(\mathcal{F}_t)_t$ is supposed to satisfy some standard conditions (typically, it is increasing and right continuous). A d-dimensional function $f\colon \mathbb{R}_+ \times X \times A \to X$ and a matrix function $\sigma\colon \mathbb{R}_+ \times X \times A \to \mathcal{M}_{d\times d''}$, both measurable, are given, called the *drift* and the *volatility* function respectively. The dynamics is written, in its general form, as

$$\mathrm{d}x_t = f_t(x_t, \pi_t)\mathrm{d}t + \sigma_t(x_t, \pi_t)\mathrm{d}\xi_t, \tag{1.1}$$

where the differential notation is to be suitably interpreted depending on the context (e.g., as an integral with respect to a jump process, or as an Itô stochastic integral if ξ_t is a martingale with continuous paths).

Typically, the user is supposed to decide the control π_t at time t before knowing the value x_t taken by the state at time t. To reflect this situation, the disturbance process is supposed to be adapted to the filtration $(\mathcal{F}_t)_t$, and the control process is supposed to be *predictable*, i.e., measurable with respect to the σ-algebra generated on $\Omega \times \mathbb{R}_+$ by all the left-continuous adapted processes. Since every predictable process is adapted, then, the state process x also turns out to be adapted. In the case when x_t has continuous paths, indeed, the distinction between adapted and predictable controls is not strictly necessary, and one can suppose all processes to be simply adapted to the filtration. We suppose the initial state x_0 to be given and well determined, but it could also be considered a random variable with respect to $\mathcal{F}_0$ with known distribution.

A control π will be called *admissible* if it induces a solution of (1.1) unique in distribution, and Π will indicate the class of all admissible controls.

A measurable function $c\colon \mathbb{R}_+ \times X \times A \to \mathbb{R}$ bounded from below, called the *instantaneous cost function* (here in general form), is given. The aim of the problem is then to minimize, with respect to all the admissible controls $\pi \in \Pi$, the *value function*, that is defined starting from the instantaneous cost function according to the user's preferences.

As an example of a choice for the value function, one may define a *cost function* $C(x, \pi)$ (depending on the path $x = (x_t)_t$ and on the choice of the control $\pi = (\pi_t)_t$ as

$$C(x, \pi) := \int_0^{+\infty} \alpha^t c_t(x_t, \pi_t) \mathrm{d}t$$

where $\alpha \in (0, 1]$ is a constant, called the *discount factor*, possibly chosen to ensure integral convergence and/or take into account, e.g., price inflation. The aim of the problem is then to minimize the value function $\Phi_{x_0}(\pi) := \mathrm{E}\{C(x, \pi) \mid x_0\}$. In this situation, we shall write the expected value conditioned to the initial state x_0 with the notation $\mathrm{E}_{x_0}\{C(x, \pi)\}$. We shall also sometimes write $\Phi^P_{x_0}(\pi) = \mathrm{E}^P_{x_0}\{C(x, \pi)\}$ to stress the fact that the expected value is computed according to the probability law P.

This is just one of the several possible ways to define the value function for an infinite horizon, continuous time stochastic control problem. Another criterion widely studied in literature is, e.g., the so-called *long-run expected average cost* (see [12] for a definition in the discrete time setting). Moreover, a variation of the given example can be defined, e.g., by considering the expected value of some convex function of the cost function to take into consideration also higher order moments: one might thus aim at minimizing $\mathrm{E}_{x_0}\{C(x, \pi) + \lambda C^2(x, \pi)\}$ for some $\lambda > 0$ (in literature, this is sometimes called the "mean-variance approach") or $\mathrm{E}_{x_0}\{\exp C(x, \pi)\}$. In some other cases, as we shall see again in the following section and chapter, instead of considering the expected value with respect to the probability law of the disturbance, one may want to be prepared for the "worst" possible case and define the value function $\sup\{C(x, \pi)\}$, where the supremum is taken over all possible disturbance processes.

Since we are just interested in giving some basic notations, in the following part of this chapter we shall always refer to the case in which the value function is defined as the expected value of the (discounted) cost function. The reader should nevertheless keep in mind that everything written below can be adapted to the various choices of the value function.

When the system has to be monitored until a given *finite horizon* $T > 0$, only a few changes to the previous notations have to be made. In this case, the underlying filtered space will be $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ and similarly the functions f, σ and c are defined on the time interval $[0, T]$ instead of the whole half-line $\mathbb{R}_+$. A major difference is that here also a so-called *terminal cost* can be taken into account under the form of a measurable function on the final state $b \colon X \to \mathbb{R}$, bounded from below. Thus, the cost function in this situation is defined as

$$C(x, \pi) := \int_0^T c_t(x_t, \pi_t) \mathrm{d}t + b(x_T)$$

(and there's no longer the need to take into account any discount factor).

The *solution* of a continuous time stochastic control problem is a pair $(\pi^*_{x_0}, \Phi^*_{x_0})$ called, respectively, *optimal control* and *optimal value*, such that $\Phi^*_{x_0} = \Phi^P_{x_0}(\pi^*_{x_0}) = \inf_\pi \Phi^P_{x_0}(\pi)$. The notation $\pi^*_{x_0}$ is used to stress the fact that the optimal control depends on the initial state x_0, since it is reasonable to expect that different initial states induce, in general, different optimal controls. For the purposes of this text, no hypotheses about the dependence of the optimal control/value on the initial state are needed.

Discrete time problems

An *infinite horizon, discrete time stochastic control problem* describes the situation when the user is interested in the state process x_t only at the discrete times $t = 0, 1, \dots$. The underlying probability space is a filtered space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0,1,\dots}, P)$, and usually (unlike the continuous time case) no particular conditions have to be satisfied by the increasing filtration $(\mathcal{F}_t)_t$. The control disturbance process ξ, just as in the continuous time case, is supposed to be adapted to the filtration.

The control at time t is supposed to be chosen on the basis of the information gathered up to and including time t, but in literature one can find two possible notations for this situation. One possibility is indeed to consider π_t as the choice made by the user at time t ($t = 0, 1, \dots$), and thus imagine the control π_t to "hold" in the time interval $[t, t+1)$. In this setting, the control process is naturally required to be adapted to the filtration. On the other hand, one might interpret π_t as the control applied in the t'th time period ($t = 1, 2, \dots$), that is, $[t-1, t)$. In this case π_t is chosen at time $t-1$, and thus the control process must be predictable (which, in the discrete time case, has the meaning of "adapted to the shifted filtration $(\mathcal{F}_{t-1})_t$").

Given a d-dimensional measurable function $F\colon \mathbb{N} \times X \times A \times \Xi \to X$, the dynamics of a discrete time stochastic control problem is defined recursively as

$$x_{t+1} = F_t(x_t, \pi_t, \xi_{t+1}), \qquad t = 0, 1, \dots, \tag{1.2}$$

so that also x is an adapted process. We suppose the initial state x_0 to be given, but (as we observed when dealing with the continuous case) it might be a random variable $\in \mathcal{F}_0$ with known distribution, for instance $x_0 = \xi_0$.

We want to remark that it is possible to formulate the discrete time problems without using filtered probability spaces and stochastic processes. Indeed, one might simply consider a given sequence ξ of random variables and define the same dynamics (1.2) as before, with the condition that the control π is chosen in such a way that $\pi_t = \pi_t(x_0, \ldots, x_t)$ be a function of the states up to and including time t. This formulation can be actually verified to be equivalent to the previous one, and in the sequel we shall often refer to the one of the two that will allow for simpler notations.

If the ξ_t are independent random variables, it can be verified that (1.2) allows to consider x_n as a controlled Markov process with transition kernel $p(x_{n+1} \mid x_n, \pi_n)$. Usually, this formulation is quite common to Operations Research and Management settings.

The *cost function* is (like in the continuous time case) a measurable function of the state and of the control processes. In its general form, given a measurable function $c\colon \mathbb{N} \times X \times A \to \mathbb{R}$ bounded from below, and a discount factor $\alpha \in (0, 1]$, it is defined as

$$C(x, \pi) := \sum_{t=0}^{\infty} \alpha^t c_t(x_t, \pi_t).$$

The aim of a discrete time control problem is the same of the continuous time one, that is, to minimize with respect to all possible choices of the control π some *value function* defined on the basis of the cost function. For our purposes, as explained above, we shall consider the value function to be defined as $\Phi_{x_0}(\pi) := \mathrm{E}_{x_0}\{C(x, \pi)\}$.

The differences between an infinite horizon and a *finite horizon* discrete time problem are quite similar to the differences between the corresponding problems in continuous time. Namely, if the system is monitored until the given horizon $T > 0$, the dynamics is defined only until time T, no discount factor need to be considered and a measurable "terminal cost" function $b\colon X \to \mathbb{R}$ (bounded from below) can be taken into account. Since this is the problem which our work is concerned with, we want to summarize here all its features for future reference.

$$(P) \qquad \begin{cases} x_{t+1} = F_t(x_t, \pi_t, \xi_{t+1}), & t = 0, 1, \ldots, T-1 \\ C(x, \pi) := \sum_{t=0}^{T-1} c_t(x_t, \pi_t) + b(x_T) \\ \Phi(x_0, \pi) := \mathrm{E}_{x_0}\{C(x, \pi)\} \\ \inf_\pi \Phi(x_0, \pi) \end{cases} . \tag{1.3}$$

The *solution* of a discrete time stochastic control problem is, just like in the continuous time case, a pair $(\pi^*_{x_0}, \Phi^*_{x_0})$ called optimal control and optimal value, such that $\Phi^*_{x_0} = \Phi^P_{x_0}(\pi^*_{x_0}) = \inf_\pi \Phi^P_{x_0}(\pi)$. We want to point out that sometimes, in the literature, a control is defined as a measurable function

$\pi\colon \mathbb{N} \times X \to A$ of the time and of the state with the meaning that, at each time $t = 0, \dots, T-1$, the system is controlled by $\pi_t := \pi_t(x_t)$ (in the "adapted" notation, or $\pi_t = \pi_t(x_{t-1})$, for $t = 1, \dots, T$ in the "predictable" notation). In this setting, e.g., the existence of $\pi^*_{x_0} = \operatorname{argmin} \Phi^P_{x_0}(\pi)$ for every initial state x_0 might not be sufficient for the existence of an optimal control, unless the set of all the possible initial states is countable or $\pi^*_{x_0}$ satisfies some "uniformity" condition with respect to x_0 (see, e.g., [13]). The proposed notation has been chosen because, as already mentioned, for our purposes the dependence of the optimal control/value on the initial state need not be considered.

In the finite horizon case, the discrete version of the *dynamic programming principle* can be applied, which is defined as follows. Suppose that $(J_t(x))_{t=0}^T$ is a sequence of functions which satisfy the following backwards recursions:

$$\begin{cases} J_T(x) = b(x) \\ J_t(x) = \inf\limits_a \left\{ c_t(x,a) + \mathrm{E}\{J_{t+1}(x_{t+1}) \mid x_t = x, \pi_t = a\} \right\} \\ \qquad\;\; = \inf\limits_a \left\{ c_t(x,a) + \mathrm{E}\{J_{t+1}(f(x,a,\xi_{t+1}))\} \right\} \quad \text{for } t < T. \end{cases}, \tag{1.4}$$

Then, for every $t = 0, \dots, T-1$, J_t is the *optimal cost-to-go* defined as

$$J_t(x) := \inf_{\pi_t, \dots, \pi_{T-1}} \mathrm{E}\left\{ \sum_{s=t}^{T-1} c_s(x_s, \pi_s) + b(x_T) \;\Big|\; x_t = x \right\}$$

and, in particular, $\Phi^*_{x_0} = J_0(x_0)$. Moreover, if there exists a $\pi^* = \pi^*_t(x)$ that realizes the inf in (1.4), then π^* is an optimal control.

Note that from (1.4) it follows that the optimal control π^*, when it exists, only depends on the current value $x = x_t$ of the state.

We want to remark that, in the finite horizon case, the assumption that the functions c and b are bounded from below also implies that both the cost function and its expected value are bounded from below. As a consequence, it is not restrictive to suppose that the value function always assumes positive values, and we shall always use this assumption in the sequel.

1.2 Model uncertainty

There are many reasons to suppose that a stochastic optimal control problem cannot be solved explicitly. When this happens, one is naturally led to simplify or to approximate or to change the model in such a way as to avoid the problems that may be caused from a rigorous formulation.

For instance, it could happen that our problem does not admit a solution, or maybe that its solution is too burdensome to obtain from a computational point of view. In other cases, which are the ones we are interested in, one could have only an approximate idea of the true dynamics of the system or of the actual costs that will be faced or of the probability law of the disturbances. This situation is referred to as "model uncertainty" or "incomplete information", and causes what by some authors is referred to as "model risk" to stress the fact that an uncertain model necessarily generates some errors in the computation of the optimal control and value.

There are many ways to face the problem of incomplete information. For example, one might consider the unknown elements of the problem as random variables on a suitable probability space. Then, the observation of the system evolution will allow the user to get more and more information as time increases, and this will be taken into account by updating the distribution of these random variables. This is the so-called Bayesian approach, which is used, e.g., in [17] (see also 2.2).

Another possible way is to consider the whole set of values that the unknown quantities can assume and try to optimize the worst possible case. In such a way one will have, for instance, a set of possible cost functions and will try to minimize over all controls the supremum of all the possible costs. This kind of approaches will lead to minimax problems similar to the one considered in [3] (also summarized in 2.1).

A third way, which is the one that we want to study, is called the *robust* approach and is exposed below for the discrete time, finite horizon case.

The robust approach

We call "robust approach" to the incomplete information case an interpretation of the so-called *robustness problem*, which can be summarized as follows.

A family $\mathcal{R}$ of stochastic optimal control problems is given, and each problem $(R) \in \mathcal{R}$ admits an optimal control $\pi^R_{x_0}$ and an optimal value $\Phi^{R,*}_{x_0}$. Having chosen a particular problem $(Q) \in \mathcal{R}$, called the *nominal* problem, one wants to investigate how the optimal control $\pi^Q_{x_0}$ behaves when used to control systems evolving according to the dynamics of other problems $(R) \in \mathcal{R}$. Of course $\Phi^R_{x_0}(\pi^Q) - \Phi^{R,*}_{x_0} \geqslant 0$, because $\pi^Q_{x_0}$ cannot induce in the problem (R) a better performance than the optimal control $\pi^R_{x_0}$ itself. Moreover, in general, $\pi^Q_{x_0}$ is not optimal for the problem (R), so that the inequality above is verified in the strict sense. The quantity

$$\Delta^{\mathcal{R},Q}_{x_0} := \sup_{(R)\in\mathcal{R}} \Phi^R_{x_0}(\pi^Q) - \Phi^{R,*}_{x_0}$$

measures this "sub-optimality" of the control $\pi^Q_{x_0}$, and is called the *robustness index* of $\pi^Q_{x_0}$ with respect to the problems of the class $\mathcal{R}$. Sometimes the robustness index is not easily determinable, and in this case one may aim at giving a *robustness inequality*, i.e., an upper bound for the robustness index that can be calculated from known properties of $\mathcal{R}$.

This setting can be adapted to the incomplete information case as follows. Suppose that we are dealing with the problem (P) described in (1.3), but that we cannot determine exactly the dynamics F or the cost function C or the probability law P. In this case, we may formulate a new problem (Q) using a dynamics F^Q, a cost function C^Q and a probability measure Q, and interpret the problem

$$(Q) \qquad \begin{cases} x_{t+1} = F^Q_t(x_t, \pi_t, \xi^Q_{t+1}), \qquad t = 0, 1, \ldots, T-1 \\ \inf_\pi \Phi^Q_{x_0}(\pi) \\ \Phi^Q_{x_0}(\pi) := \mathrm{E}^Q_{x_0}\{C^Q(x, \pi)\} \\ C^Q(x, \pi) = \sum_{t=0}^{T-1} c^Q_t(x_t, \pi_t) + b^Q(x_T) \end{cases}, \tag{1.5}$$

which we suppose to be able to solve, as the nominal problem of a suitable class $\mathcal{R}$ containing also the problem (P). We shall also say that (Q) is an *approximating problem* of (P), as its solution allows to explicitly compute a control for the problem (P), which we suppose not to be solvable. For what we have seen above, the performance of the optimal control $\pi^Q_{x_0}$ in the problem (P) is measured by the *robustness index*

$$\Delta_{x_0}(P, Q) := \Phi^P(x_0, \pi^Q_{x_0}) - \Phi^*_{x_0} \geqslant 0, \tag{1.6}$$

which, in turn, is bounded from above by any robustness inequality we can give for any class $\mathcal{R}$ containing both (P) and (Q).

Of course the robust approach, here defined for finite horizon, discrete time problems, can be applied to any other stochastic control problem as well. Indeed, once the approximating problem is suitably constructed, the robustness index can be defined exactly as in (1.6) also for continuous time and/or infinite horizon problems.

Note also that the robustness problem can be approached from different viewpoints. As an example, we may want to consider controls $\pi^{\mathcal{R}}$ that need not be optimal for any of the problems $(R) \in \mathcal{R}$ but, in exchange, give "equally good" performance in all problems of the class. This will lead us to solve a problem such as $\inf_\pi \sup_{(R)\in\mathcal{R}}\{\Phi^R_{x_0}(\pi)\}$, or, in other words, to study the robustness of a minimax control. In other cases, the robustness problem can be "mixed" with a bayesian approach, in the sense that, e.g., we could look for a way to progressively "shrink", on the basis of the successively accumulated information, the class containing the nominal and the real problems (possibly choosing different nominal problems as time goes by) and thus hopefully improve the robustness inequality we can obtain.

Problems analogous to this one have been recently investigated, for example, by Dupuis, James and Petersen ([2]), by Gordienko and Lemus-Rodríguez ([10]) and by Gordienko and Salem ([11, 12]). In [10] (see also 2.3) the authors deal with the case of a continuous time control problem “approximated” by another problem on the same underlying probability space (i.e., only dynamics and target function are different in the two problems). In [11] and [12] (see 2.4) the authors investigate the case when the state process is a Markov control process and the control problem is approximated by another problem with different transition kernels and target function. In [2] the authors use the robust approach for a particular class of continuous time control problems coming from an engineering setting. They consider the “real” problem as a parametric “perturbation” of the nominal problem, and the behaviour of the robustness inequality with respect to the perturbating parameter is investigated.

The aim of our work is to investigate the case when the dynamics F and the target function C are supposed to be completely determined, and the only source of uncertainty in the model is the underlying probability measure P. In other words, we want to determine $\Delta_{x_0}(P, Q)$ as defined above when $F^Q = F$, possibly defining in a suitable form the cost functional C^Q. It turns out (see [6] and 3.1) that a “good” robustness inequality is very hard to obtain, as the model is in general very sensitive to even small changes in the underlying probability measure.

Chapter 2

An overview of literature

In this chapter some results are gathered which were available in literature when our research began.

Papers [10] and [11, 12] (summarized in 2.3 and 2.4 respectively) are examples of applications of the robust approach, and allow to understand the typical results and problems that this approach implies. Indeed, examining the results of these papers will generate some remarks that can be applied also to the general results of our paper [6].

On the other hand, papers [3, 16] and [17] (see 2.1 and 2.2) are examples of other approaches to incompleteness of information, i.e., the minimax and the Bayesian approach respectively. Moreover, [17] gave us the idea of a particular model to apply the robust approach to, and has thus been the starting point for papers [5] and [7].

2.1 The minimax approach

In [3], Dai Pra et al. use the following approach to uncertainty in stochastic control problems. Given a discrete time, finite horizon problem whose dynamics is like in problem (P), the authors suppose that at time t our only knowledge of the disturbance ξ_t is its support $\Xi_t(x_t, \pi_t)$ (possibly dependent on the state and the control). In particular, no hypothesis is made on the probability law of the disturbances, so that probability theory plays no role in this approach. Instead of the expectation of the target function C, the following value function is considered:

$$\Phi(x_0, \pi) := \sup_{\xi_t \in \Xi_t(x_t, \pi_t)} \{C(x, \pi)\}.$$

In other words, instead of optimizing the "average" of all the possible outcomes of the system, the user wants to prepare for the worst of them.

The key result of the paper is that this problem can be solved via a modified version of the dynamic programming algorithm, which in this case takes the following form. If the target function is given in the general form $C(x,\pi) := \sum_{t=0}^{T-1} c_t(x_t,\pi_t) + b(x_T)$ as in (1.3) and one defines inductively

$$\begin{aligned} J_T(x_T) &:= b(x_T), \\ J_t(x_t) &:= \inf_{\pi_t} \sup_{\xi_t \in \Xi_t(x_t,\pi_t)} \{c_t(x_t,\pi_t,\xi_t) + J_{t+1}(F_t(x_t,\pi_t,\xi_t))\}, \end{aligned}$$

then (if $\Phi^*_{x_0}$ is the optimal value of the problem as in chapter 1) $\Phi^*_{x_0} = J_0(x_0)$. This way, the optimization problem becomes a "minimax" problem and explicit solutions can be achieved.

Another more general way to take into account the uncertainty of the disturbances ξ_t is described by Runggaldier in [16, Section 6] (see also [4]). Here, one is supposed to know that at time t the disturbances belong to random sets with known distributions, i.e., $\xi_t \in \Xi_t(\omega)$ where ω is the generic element of a suitable underlying probability space.

In this setting, there are many possible ways to define the value function. Two of the most natural ones, for instance, are

$$\begin{aligned} \Phi^\dagger(x_0,\pi) &:= \mathrm{E}\sup_{\xi_0} \ldots \mathrm{E}\sup_{\xi_{T-1}} \{C(x,\pi)\}, \\ \Phi^\ddagger(x_0,\pi) &:= \mathrm{E}\sup_{\{\xi_t\}} \{C(x,\pi)\}. \end{aligned}$$

It turns out (see [4]) that the dynamic programming algorithm can be applied to the optimization of $\Phi^\dagger(x_0,\pi)$, but that in general it is not possible to approach in the same way the optimization of $\Phi^\ddagger(x_0,\pi)$.

It is noteworthy that this model is in some sense a "combination" of the stochastic approach (P) and the pure minimax approach of [3]: indeed, the problem (P) is recovered in the particular case when the random sets collapse to a single point, and the situation of [3] is recovered when the sets $\Xi_t(\omega)$ are constant with respect to ω.

A similar approach has been followed also, e.g., by Cvitanić and Karatzas in [1], and by Talay and Zheng in [18]. In [1], the authors consider the incomplete information case for a particular continuous time stochastic problem arising from an economic application, namely the analogous of the "mean shortfall minimization problem" defined in 2.2 and 3.2. They approach the problem with a minimax technique, as they aim at finding an optimal control that minimizes the supremum of the value function over a family of probability measures. Moreover, Talay and Zheng show in [18] that it is possible to apply in a suitable way the continuous version of the dynamic programming principle to the minimax approach of [1].

2.2 The bayesian approach for the binomial market model

In [17], Runggaldier et al. study a well-known market model, that is, the binomial one. The model is quickly summarized here, but for a more precise definition of the model, see 3.2.

Suppose that two assets are tradable on a discrete time market, namely a *bond* B (that without loss of generality – see also 3.2 – can be supposed to be constant and equal to one and represents, e.g., some cash account) and a *stock* S whose starting value S_0 is given. The stock is supposed to evolve stochastically, at each time t, either going "up" or "down" by fixed and known percentages u and d (with $0 < d < 1 < u$). The probability p of the stock going up at time t is also supposed to be fixed and the same for all $t = 0, \dots, T-1$ (with T a known finite horizon), so that the dynamics of the stock can be written as

$$S_{t+1} = S_t \xi_{t+1}, \quad t = 0, \dots, T-1, \qquad P\{\xi_t = u\} = 1 - P\{\xi_t = d\} = p.$$

A function H of the final state S_T is given, and the user (called, to reflect the economic setting, the *investor*) wants to determine a *portfolio* (that is, a control $\pi = (\alpha, \beta)$ corresponding to the decision to hold, at each time t, α_t units of the stock and β_t units of the bond) so as to possibly possess, at time T, exactly $H(S_T)$ units of money whatever the evolution of the stock is. If this happens, one says that the user has succeeded in *hedging*, or *replicating*, the function H by means of the portfolio π. The portfolio is moreover supposed (as it is customary to do in economic models) to be self-financing, i.e., once the initial capital V_0 is decided, no money can be added or withdrawn until the final time T. This way, once the control is chosen, the only gains or losses come from the evolution of the stock, and indeed if V_t^π is the value of our portfolio at time t when the control π is used, then V_t evolves according to the following dynamics:

$$V_{t+1}^\pi(V_t, S_t, \alpha_t, \xi_{t+1}) = V_t^\pi + \alpha_t S_t (\xi_{t+1} - 1), \qquad V_0^\pi := V_0.$$

It is a classical result that there exists a "critical" initial capital V_0^* such that the function $H(S_T)$ can be replicated by a suitable portfolio if and only if $V_0 \geqslant V_0^*$. As a consequence, if $V_0 < V_0^*$ some loss will have to be taken into account. This way the problem can be seen as a dynamic stochastic control problem whose aim is to minimize the expected *shortfall*, that is, the positive part of the difference between $H(S_T)$ and the final value V_T of our portfolio:

$$\inf_\pi \mathrm{E}_{S_0, V_0} \left\{ (H(S_T) - V_T^\pi)^+ \right\}.$$

The optimal control and the optimal value for this problem can be obtained explicitly by using the dynamic programming algorithm, and turn out to depend on the probability p.

The authors face then the problem of incomplete information in the model, that is, the case in which p is not known by the investor. The Bayesian approach, that they follow, can be explained as follows.

Consider p as a random variable on a suitable probability space, and let $h_0(p) \propto p^{\lambda_0}(1-p)^{\mu_0}$ be an "a-priori" Beta distribution obtained by fixing arbitrarily two parameters $\lambda_0, \mu_0 > -1$. Let then, for each time t, $u_t := \#\{s < t \mid \xi_s = u\}$ be the total number of "up-movements" of the stock cumulated up to time t and define $\lambda_t := \lambda_0 + u_t$, $\mu_t := \mu_0 + t - u_t$. The "posterior" distribution of p at time t thus becomes $h_t(p) \propto p^{\lambda_t}(1-p)^{\mu_t}$. This way, at each time t it will be possible to calculate, according to the distribution h_t, $p_t := \mathrm{E}\{p\} = \frac{\lambda_t+1}{\lambda_t+\mu_t+2}$.

Now, the dynamic programming technique can still be applied, and it turns out at each time t the probability p_t replaces the (unknown) probability p. In general, this approach would be very heavy from a computational point of view (as the dynamic programming requires to solve $T-t$ recursive steps to determine the optimal control at time t). In this particular case, however, it turns out that one can use the explicit form of the control for the complete information case (see 3.2 for the details). Thus, the bayesian way proves to be a quite efficient approach to the incomplete information in the binomial case.

2.3 Robustness for diffusion processes

As an example of the robust approach applied to a continuous time problem, one can consider [10]. Here, Gordienko and Lemus-Rodríguez suppose to deal with a continuous time, finite horizon stochastic control problem

$$\begin{cases} \mathrm{d}x_t = f_t(x_t, \pi_t)\mathrm{d}t + \sigma_t(x_t, \pi_t)\mathrm{d}\xi_t \\ \inf_\pi \mathrm{E}_{x_0}\{\int_0^T c_t(x_t, \pi_t)\mathrm{d}t\} \end{cases}$$

and to build an approximating problem choosing a different drift f^Q, a volatility σ^Q and an istantaneous cost c^Q. The disturbance process ξ_t is supposed to be a standard brownian motion, and the same for both the "real" and the approximating problems.

As explained in 1.2, the optimal control π^Q of the approximating problem is applied to the original problem thus yelding a value $\Phi_{x_0}(\pi^Q)$ greater than the optimal $\Phi^*_{x_0}$. The robustness index is now defined in the same way as in (1.6), namely $\Delta_{x_0} := \Phi_{x_0}(\pi^Q) - \Phi^*_{x_0}$, and some upper bound on its value is investigated. The results of Gordienko and Lemus-Rodríguez can be gathered as follows.

2.3.1 Theorem. *Suppose that f and σ are continuous in the state and in the control and both lipschitz and sub-linear in the state, in the sense that there exist constants $K_1, K_2, K' > 0$ such that*

$$\begin{aligned} |f_t(x,\pi) - f_t(y,\pi)| &\leqslant K_1|x-y|,\\ \|\sigma_t(x,\pi) - \sigma_t(y,\pi)\| &\leqslant K_2|x-y|,\\ |f_t(x,\pi)| + \|\sigma(t,x,\pi)\| &\leqslant K'(1+|x|) \end{aligned}$$

uniformly for all times t and controls π. Suppose also c to be lipschitz in the state, i.e., there exists $C > 0$ such that

$$|c_t(x,\pi) - c_t(y,\pi)| \leqslant C|x-y|$$

uniformly for all times and controls.

Define $\vartheta := 2(K_1^2T + K_2^2)$ and

$$\begin{aligned} \Delta_f &:= \sup_{t\in[0,T]} \Big\{ \mathrm{e}^{-\frac{\vartheta t}{2}} \int_0^t \sup_{x,a} |f_s(x,a) - f_s^Q(x,a)| \mathrm{d}s \Big\},\\ \Delta_\sigma &:= \sup_{t\in[0,T]} \Big\{ \mathrm{e}^{-\vartheta t} \int_0^t \sup_{x,a} \|\sigma_s(x,a) - \sigma_s^Q(x,a)\| \mathrm{d}s \Big\}, \\ \Delta_c &:= \int_0^T \sup_{x,a} |c_s(x,a) - c_s^Q(x,a)| \mathrm{d}s. \end{aligned} \tag{2.1}$$

Define also

$$R := \frac{C\sqrt{2T(\mathrm{e}^{\vartheta T} - 1)}}{\sqrt{\vartheta}[1 - \sqrt{1 - \mathrm{e}^{-\vartheta T}}]}.$$

Then

$$\Delta_{x_0} \leqslant 2\big[\Delta_c + R\sqrt{\Delta_f^2 + \Delta_\sigma}\big]. \tag{2.2}$$

Even if these results themselves are not particularly interesting for our work, they are cited here because they lead to some important considerations. It is noteworthy, for example, how strong the conditions on the coefficients f, σ and c have to be in order to get some robustness result.

To get an idea of what the given upper bound can be, consider the (rather natural) case in which, to calculate (2.1), the functions $|f_s(x,a) - f_s^Q(x,a)|$ and $||\sigma_s(x,a) - \sigma_s^Q(x,a)||$ have to be replaced by some uniform majorants.

Then Δ_f and Δ_σ are the supremum of functions proportional to $t\,e^{-\frac{\vartheta t}{2}}$, so that it can be verified by a straightforward calculation that they are proportional to $\frac{1}{\vartheta}$. On the other hand, R can be checked to be asymptotically equivalent to $e^{\frac{3}{2}\vartheta T}\sqrt{T/\vartheta}$ and ϑ is linearly increasing with respect to the horizon T. So, the upper bound given in (2.2) is exponentially increasing in the horizon.

These considerations lead to the conclusion that, even under these quite strong assumptions, the given upper bound for the robustness index turns out to be very sensitive to the distance between the "real" and the approximating problems. This is not limited to the continuous time setting, as we are going to see below.

2.4 Robustness for Markov control processes

In [11], Gordienko and Salem deal with a particular case of robustness problem, very close to the one we study in [6]. Namely, they deduce a robustness inequality in the same sense of (1.6) in the setting of a Markov control process with infinite horizon and "one-stage" cost function as explained below.

Roughly speaking, as already mentioned in the definition of the discrete time problems in section 1.1, the state x_t of the problem (P) is a Markov process if the disturbances ξ_t are independent. Actually, Gordienko and Salem suppose a stronger condition, namely that the transition kernel does not depend on time t: this is the situation, for instance, of a dynamics F independent of t and ξ_t (independent and) identically distributed. Moreover, the cost function is said to be "one-stage" if the target function C is given in the particular form

$$C(x,\pi) := \sum_{t=0}^{\infty} \alpha^t c(x_t,\pi_t),$$

where $c(x,a)$ is a given function depending only on the state and the control (and in particular independent of t) and α is a given discount factor. Also, the set of feasible actions is supposed to depend only on the state of the system and not on the time, i.e., to every state x one associates a set $A(x)$ such that at time t the control π_t has to be chosen in the set $A(x_t)$.

The problem (P) is approximated by a problem (Q) which differs in the cost function and in the transition probability, i.e., in the underlying probability law. Using our notations, the authors construct an optimal control $\pi^Q_{x_0}$ for the approximating problem (Q) and derive an upper bound for the robustness index $\Delta_{x_0}(P,Q)$ defined in (1.6). The results obtained by Gordienko and Salem can be summarized as follows.

2.4.1 Theorem. *Let (P) be a discrete time, infinite horizon Markov control problem with one-stage cost function. Consider the approximating problem (Q) obtained from the problem (P) by taking another probability law Q and another cost function $c^Q(x,a)$.*

If there exist a measurable function $w\colon X \to \mathbb{R}_+$ and a constant $0 \leqslant \beta < 1$ such that for all $x \in X$, $a \in A(x)$ one has

$$c(x,a) \leqslant w(x), \qquad \mathrm{E}^P\{w(x_{t+1}) \mid x_t = x\} \leqslant \frac{\beta}{\alpha} w(x), \tag{2.3}$$

then

$$\Delta_{x_0}(P,Q) \leqslant 2w(x_0)(1-\beta)^{-1}[\delta_1 + \delta_2 \alpha (1-\beta)^{-1}] \tag{2.4}$$

where

$$\begin{aligned} \delta_1 &:= \sup_{x,a} [w(x)]^{-1} |c(x,a) - c^Q(x,a)| \\ \delta_2 &:= \sup_{x,a} [w(x)]^{-1} \mathrm{E}^Q \Big\{ w(x_{t+1}) \cdot \big| 1 - \tfrac{\mathrm{d}P}{\mathrm{d}Q}(x_{t+1}) \big| \;\Big|\; x_t = x, \pi_t = a \Big\}. \end{aligned} \tag{2.5}$$

Note that in the case when $c^Q(x,a) = c(x,a)$, namely, when the only difference between the real and the approximating problems is the underlying probability law, the robustness inequality of the above theorem reduces to

$$\Delta_{x_0}(P,Q) \leqslant 2\alpha(1-\beta)^{-2} w(x_0) \delta_2.$$

In [12], the same authors consider this case together with the "long-run expected average cost" and show that, under strong recurrence and ergodicity assumptions, this $O((1-\beta)^{-2}$ bound can be improved to $O((1-\beta)^{-1})$. While for more details and the proof of the theorems we refer to [11] and [12], we want now to discuss some of the major problems that arise when dealing with robust approximations as those that can be deduced from the cited results.

The conditions (2.3) can be interpreted as follows. The first one, as it is straightforward to see, means that the one-stage cost function has to be bounded by some positive measurable function $w(x)$ uniformly in the control. The second one implies that the expected discounted cost has to be exponentially decreasing in time, because the existence of a β behaving like in (2.3) implies that

$$\mathrm{E}^P\{w(x_t)\} = \mathrm{E}^P\left\{ \mathrm{E}^P\{w(x_t) \mid x_{t-1}\}\right\} = \ldots \leqslant \frac{\beta^t}{\alpha^t} w(x_0),$$

i.e., that the average discounted cost of one step, dominated by $\alpha^t \, \mathrm{E}^P\{w(x_t)\}$, tends exponentially to 0 uniformly in the control as $t \to +\infty$. It is clear, then, that these conditions are actually very strong, even if they seem quite reasonable.

In spite of this strong condition, however, the given upper bound of the robustness index is strongly sensitive to the rate β of convergence to zero of the discounted cost — indeed, the right hand side of (2.4) goes to infinity as $\beta \to 1$.

Another problem, that is quite evident, arises when considering the expressions involved in determining the upper bound. Consider, e.g., the computation of δ_2 in (2.5). The more $\frac{\mathrm{d}P}{\mathrm{d}Q}$ is close to 1 in correspondence to the x's where $w(x)$ is "big", the smaller δ_2 will be. This is quite reasonable, because the more precisely we shall estimate the probability of "high" costs, the more accurate our approximation will be. Similar considerations can be made about the role that β plays, because of course less accuracy is needed in the approximation when the discounted cost quickly decreases to zero. This stresses the importance to have some insight into the probability P to get the best possible values for β and δ_2, and makes clear that the robustness index is in general very sensitive to the degree of knowledge of the real world probability measure P. On the other hand, there seems to be no general method to be able either to estimate the P-expected value of (2.3) or to know the distribution of high values of $w(x)$ better than the distribution of its low values.

We shall see in 3.1 that these considerations can be applied unchanged to the general result of [6].

Chapter 3

Our results

This chapter synthesizes our research.

Our starting point was to determine a robustness inequality for discrete time, finite horizon stochastic control problem under assumptions as general as possible. This investigation has led to [6], summarized in 3.1, whose results in some sense reflect those of [10] and [11] cited in 2.3 and 2.4.

Inspired by [17], then, we decided to choose the binomial market model as a particular case for an application of the robust approach. This study quickly led to the discovery of some interesting properties of the binomial model that deserved a paper on its own, namely [5] (summarized in 3.2), and also lead us to look for going deeper into this investigation. The work currently in progress is gathered in [7] and summarized in 3.3.

3.1 General robustness results for stochastic control

We suppose a discrete time, finite horizon stochastic control problem to be given in the form (P) defined in (1.3). We want to investigate the case in which the probability P is unknown, and so we have to build an approximating problem (Q) based on an hypothetical measure Q. Our aim, in particular, is to give a robustness inequality (see 1.2 for the definition) based only upon the Radon-Nikodym derivative $\frac{\mathrm{d}P}{\mathrm{d}Q}$, which is supposed to exist (i.e., P is supposed to be absolutely continuous with respect to Q).

If the problem (Q) is defined simply by substituting the measure Q for the measure P, the following problem is obtained:

$$(Q') \qquad \begin{cases} x_{t+1} = F_t(x_t, \pi_t, \xi^Q_{t+1}), & t = 0, 1, \ldots, T-1 \\ \inf_\pi \Phi^Q(x_0, \pi) & \\ \Phi^Q(x_0, \pi) := \mathrm{E}^Q_{x_0}\{C(x, \pi)\} & \end{cases} . \tag{3.1}$$

Suppose that that Q is such that $\left|1 - \frac{\mathrm{d}P}{\mathrm{d}Q}\right| \leqslant \widetilde{\gamma}$. Since, as already said at the end of chapter 1, we can always suppose $C(x, \pi) > 0$, from a straightforward calculation one has

$$\left| \mathrm{E}^Q_{x_0}\{C(x, \pi)\} - \mathrm{E}^P_{x_0}\{C(x, \pi)\} \right| = \left| \mathrm{E}^Q_{x_0}\left\{C(x, \pi)\big(1 - \tfrac{\mathrm{d}P}{\mathrm{d}Q}\big)\right\} \right| \leqslant \widetilde{\gamma} M_{x_0} \tag{3.2}$$

with $M_{x_0} := \sup_\pi \Phi^Q(x_0, \pi)$. As a consequence, one gets the robustness inequality

$$\Delta_{x_0}(P, Q) \leqslant 2\widetilde{\gamma} M_{x_0}. \tag{3.3}$$

It is clear that this approximation is unsatisfactory from many points of view. First of all, since both $\widetilde{\gamma}$ and M_{x_0} are upper bounds for the Radon-Nikodym derivative and the value function respectively, the right hand side of (3.2) is in general much greater than both the optimal values $\Phi^*_{x_0}$ and $\Phi^Q_{x_0}$ themselves, and thus uninformative. There is not even any information on whether the optimal value $\Phi^Q_{x_0}$ is actually greater or smaller than the real optimal value $\Phi^*_{x_0}$.

A way to improve the bounds seems then to redefine suitably the cost function C in such a way to obtain what we call *lower* and *upper* approximating problems. A problem (Q_l) is said to be a *lower approximating problem* of (P) if for every initial state x_0 and control π one has $\Phi^{Q_l}(x_0, \pi) \leqslant \Phi(x_0, \pi)$; analogously, a problem (Q_u) is called an upper approximating problem if $\Phi^{Q_u}(x_0, \pi) \geqslant \Phi(x_0, \pi)$.

By (3.2) one can define the following cost functions for two approximating problems, respectively, lower and upper:

$$\Phi^{Q_l}(x_0, \pi) := (1 - \widetilde{\gamma})\Phi^Q(x_0, \pi) \leqslant \Phi(x_0, \pi)$$
$$\Phi^{Q_u}(x_0, \pi) := (1 + \widetilde{\gamma})\Phi^Q(x_0, \pi) \geqslant \Phi(x_0, \pi).$$

Note that Φ^{Q_u} is always positive, while for Φ^{Q_l} there are two possibilities. If $\widetilde{\gamma} \geqslant 1$, then $\Phi^{Q_l} \leqslant 0$ and it may even be unbounded from below, so that the lower approximating problem that has Φ^{Q_l} as value function does not give any information. On the other hand, in the interesting case $\widetilde{\gamma} < 1$ (i.e., when P and Q are next to each other) one has $\Phi^{Q_l} > 0$ and, if $\pi^*_{x_0}$ is the P-optimal control, then

$$(1 - \widetilde{\gamma})\Phi^Q(x_0, \pi^*_{x_0}) \leqslant \Phi(x_0, \pi^*_{x_0}) \leqslant \Phi(x_0, \pi^Q) \leqslant (1 + \widetilde{\gamma})\Phi^Q(x_0, \pi^Q).$$

So, in this situation, $\pi^*_{x_0}$ must belong to the class

$$\Pi^* := \{\pi \in \Pi \mid (1-\widetilde{\gamma})\Phi^Q(x_0,\pi) \leqslant (1+\widetilde{\gamma})\Phi^Q(x_0,\pi^Q_{x_0})\} \subseteq \Pi$$

and we can restrict the computation of $M_{x_0} := \sup_\pi \Phi^Q(x_0,\pi)$ to the subclass Π^*, where we get $M_{x_0} = \frac{1+\widetilde{\gamma}}{1-\widetilde{\gamma}}\Phi^Q(x_0,\pi^Q_{x_0})$. This way, in the case $\widetilde{\gamma} < 1$, one can improve the robustness inequality obtained in (3.3) to

$$\Delta_{x_0}(P,Q) \leqslant 2\frac{\widetilde{\gamma}(1+\widetilde{\gamma})}{1-\widetilde{\gamma}}\Phi^Q(x_0,\pi^Q_{x_0}), \tag{3.4}$$

whose right hand side, being based upon the Q-optimal value $\Phi^Q(x_0,\pi^Q_{x_0})$ and no longer upon a sup, is in general much smaller than the right hand side of (3.3).

As an alternative approach, one can try to treat the Radon-Nikodym derivative $\frac{\mathrm{d}Q}{\mathrm{d}P}$ by using exponentiation together with the linearity of the expected value. In this situation, we suppose the knowledge of the measure P to be limited to an upper bound $\overline{\gamma}$ for the random variable $\log\left(\frac{\mathrm{d}P}{\mathrm{d}Q}\right)$. Note that if such a $\overline{\gamma}$ exists, necessarily the support sets for the measures P and Q coincide, i.e., P and Q are equivalent.

Define $\Gamma(x,\pi) := \log\big(C(x,\pi)\big) + \log\left(\frac{\mathrm{d}P}{\mathrm{d}Q}\right)$ and write

$$\Phi(x_0,\pi) = \mathrm{E}^P_{x_0}\{C(x,\pi)\} = \mathrm{E}^Q_{x_0}\{\mathrm{e}^{\Gamma(x,\pi)}\}. \tag{3.5}$$

To switch the exponential function and the expected value, one may want to use Jensen's inequality $\mathrm{E}\{\mathrm{e}^X\} \geqslant \mathrm{e}^{\mathrm{E}\{X\}}$. Unfortunately, this might lead to estimates too "loose", because in general such an inequality is quite far from being an equality.

The idea is then to "partition" the variability domain of the r.v. $\Gamma(x,\pi)$ in order to get the sum of (conditional) expected values of r.v.'s with "small" variation, which can be estimated more precisely. To do this, we need the further hypotheses that $C(x,\pi)$ be bounded, namely, that there are $\overline{M} > \overline{m}$ such that

$$\mathrm{e}^{\overline{m}} \leqslant C(x,\pi) \leqslant \mathrm{e}^{\overline{M}} \qquad \text{for every } \pi.$$

Define now $m := \overline{m} - \overline{\gamma}$, $M := \overline{M} + \overline{\gamma}$ and consider a partition $m = m_1 < M_1 = m_2 < \cdots < M_{n-1} = m_n < M_n = M$ of $[m,M]$. Writing $I_i(x_0,\pi) := \{m_i \leqslant \Gamma(x,\pi) \leqslant M_i \mid x_0\}$ and $Q'_i(x_0,\pi) := Q(I_i(x_0,\pi))$ $(i = 1,\dots,n)$, equation (3.5) becomes

$$\Phi(x_0,\pi) = \sum_{i=1}^{N} Q'_i(x_0,\pi)\,\mathrm{E}^Q\{\mathrm{e}^{\Gamma(x,\pi)}\,|I_i(x_0,\pi)\}.$$

It is now clear that the problems with cost functionals

$$L^Q(x_0,\pi) := \textstyle\sum_i Q'_i(x_0,\pi)\,\mathrm{e}^{m_i}$$
$$U^Q(x_0,\pi) := \textstyle\sum_i Q'_i(x_0,\pi)\,\mathrm{e}^{M_i}$$

are, respectively, a lower and an upper approximating problem for (P). They, however, still depend on the probability P which enters $\Gamma(x,\pi)$ and thus also the probabilities $Q'_i(x_0,\pi)$.

To eliminate this dependence on P, define

$$Q''_i(x_0,\pi) := Q\{m_i + \overline{\gamma} \leqslant \log(C(x,\pi)) \leqslant M_i - \overline{\gamma} \mid x_0\}$$
$$Q_i(x_0,\pi) := Q\{m_i - \overline{\gamma} \leqslant \log(C(x,\pi)) \leqslant M_i + \overline{\gamma} \mid x_0\}$$

for every i. Clearly, $Q''_i(x_0,\pi) \leqslant Q'_i(x_0,\pi) \leqslant Q_i(x_0,\pi)$ and both Q and Q'' only depend on P via $\overline{\gamma}$, which is supposed to be known. A lower and an upper approximating problem "independent" of P can then be defined as

$$\Phi^{Q_l}(x_0,\pi) := \textstyle\sum_i Q''_i(x_0,\pi)\,\mathrm{e}^{m_i}$$
$$\Phi^{Q_u}(x_0,\pi) := \textstyle\sum_i Q_i(x_0,\pi)\,\mathrm{e}^{M_i}\,.$$

To estimate the distance between the "real" problem (P) and the upper approximating problem just defined, calculate

$$\begin{aligned}
&\Phi^{Q_u}(x_0,\pi) - \Phi(x_0,\pi) \leqslant \\
&\quad\leqslant \Phi^{Q_u}(x_0,\pi) - \Phi^{Q_l}(x_0,\pi) \\
&\quad= \textstyle\sum_i Q_i(x_0,\pi)(\mathrm{e}^{M_i} - \mathrm{e}^{m_i}) + \sum_i (Q_i(x_0,\pi) - Q''_i(x_0,\pi))\,\mathrm{e}^{m_i}\,.
\end{aligned}$$

Choose now a $\delta > 0$ and fix the partition $m = m_1 < M_1 = m_2 < \cdots < M_{n-1} = m_n < M_n = M$ in such a way that $\mathrm{e}^{M_i} - \mathrm{e}^{m_i} = \delta$ for all $i = 1,\ldots,n$. Suppose furthermore that the r.v. $\log C(x,\pi)$ has a probability density function $p^Q_{\log C(x,\pi)}$ under the measure Q such that, given x_0, $p^Q_{\log C(x,\pi)} \leqslant K_{x_0}$ for all π. Then, from the above expression, one gets

$$\begin{aligned}
\Phi^{Q_u}(x_0,\pi) - \Phi(x_0,\pi) &\leqslant \delta + 4K_{x_0}\overline{\gamma}\textstyle\sum_i (\mathrm{e}^m + i\delta) \\
&= \delta + 4K_{x_0}\overline{\gamma}n\Big(\,\mathrm{e}^m + \frac{n-1}{2}\delta\Big) \\
&< \delta + 2(\mathrm{e}^M + \mathrm{e}^m)K_{x_0}n\overline{\gamma} \\
&= \delta + 2(\mathrm{e}^{2M} - \mathrm{e}^{2m})K_{x_0}\frac{\overline{\gamma}}{\delta}.
\end{aligned}$$

Now, we want to choose δ in such a way as to minimize the expression just obtained. It is immediate to verify that $\delta + 2(\mathrm{e}^{2M} - \mathrm{e}^{2m})K_{x_0}\frac{\overline{\gamma}}{\delta}$ is a convex function which for $\delta(\overline{\gamma}) := \sqrt{2(\mathrm{e}^{2M} - \mathrm{e}^{2m})K_{x_0}\overline{\gamma}}$ attains its absolute minimum of $2\delta(\overline{\gamma})$. In the end, then, this "exponential" approach leads to the robustness inequality

$$\Delta_{x_0}(P,Q) \leqslant 4\sqrt{2(\mathrm{e}^{2M} - \mathrm{e}^{2m})K_{x_0}\overline{\gamma}}. \tag{3.6}$$

Note that, in general, there is no exact partition $m = m_1 < M_1 = m_2 < \cdots < M_{n-1} = m_n < M_n = M$ such that $\mathrm{e}^{M_i} - \mathrm{e}^{m_i} = \delta(\overline{\gamma})$, because $\frac{\mathrm{e}^M - \mathrm{e}^m}{\delta(\overline{\gamma})}$ might not be an integer. The above argument makes nevertheless clear that the optimal partition is the one with $\overline{n} := \lceil \frac{\mathrm{e}^M - \mathrm{e}^m}{\delta(\overline{\gamma})} \rceil$ points, obtained by setting $m_1 := m$, $M_i = m_{i+1} := \log(\mathrm{e}^m + i\delta(\overline{\gamma}))$ $(i = 1, \ldots, \overline{n} - 1)$ and $M_{\overline{n}} := M$.

The two approaches considered, called respectively the "direct" and the "exponential" approach, are quite hard to compare with each other.

Carrying out some calculations, it is possible to bound from both sides the ratio $\frac{\overline{\varepsilon}}{\widetilde{\varepsilon}}$, where $\widetilde{\varepsilon}$ and $\overline{\varepsilon}$ indicate, respectively, the right hand sides of (3.3) and (3.6). The bounds obtained are, roughly speaking, the product of three terms depending, respectively, on the Radon-Nykodim derivative $\frac{\mathrm{d}P}{\mathrm{d}Q}$, on the upper bound K_{x_0} for the Q-density of the r.v. $\log(C(x,\pi))$ and on the support $[\mathrm{e}^m, \mathrm{e}^M]$ of the target function. A joint minimization (see [6] for the details) shows then that the minimum value for $\frac{\overline{\varepsilon}}{\widetilde{\varepsilon}}$ is approximately 8.14.

This allows to conclude that the direct approach (even in its coarser form (3.3)) gives much more precise results than the exponential one. Note also that, as already mentioned, $\log\left(\frac{\mathrm{d}P}{\mathrm{d}Q}\right)$ can be bounded only if P and Q are equivalent, while this condition is not necessary for the existence of an upper bound for $\left|1 - \frac{\mathrm{d}P}{\mathrm{d}Q}\right|$. As a consequence, the direct approach can be applied to a wider class of approximations. Moreover, the direct approach does not need any hypotheses either on the distribution or on the variability domain of the r.v. $\log(C(x,\pi))$, so that it can be applied to atomic or discrete models or even to unbounded value functions, unlike the exponential one.

On the other hand, the exponential approach features a value function whose form is different from the original one, and this might be an advantage from a computational point of view when $C(x,\pi)$ has a complicated structure. Indeed, the exponential approach gives both quite good results and very simple value functions in the case when P and Q are very near, and the idea of partitioning the variability domain of C gives nontrivial simplifications even in the case $Q = P$.

Note that the robustness inequality (3.3) is from many points of view similar to the one obtained by Gordienko and Salem in [11] (see 2.4). In particular, if there is a random variable $w(x)$ (depending on all the path $x =$

$(x_t)_t$) such that $C(x,\pi) \leqslant w(x)$ for all π and we can calculate (or estimate) $\delta_{x_0} := \mathrm{E}^Q_{x_0}\left\{w(x)\left|1-\frac{\mathrm{d}P}{\mathrm{d}Q}\right|\right\}$, then (3.3) can be improved to $\Delta_{x_0}(P,Q) \leqslant 2\delta_{x_0}$.

In general, however, these results resemble those of [11] and [10] in showing that the approximation of a problem by means of one based upon a different probability measure is, in general, very sensitive to the distance between the "real" and the approximating measure.

3.2 Robust and adaptive approaches for the binomial model

The *binomial market model*, first introduced by Cox, Ross and Rubinstein, is one of the most popular discrete-time models in mathematical finance. The market under consideration is monitored at the discrete times $t = 0, \dots, T$ and is composed of two assets: a riskless *bond* $B = (B_t)_{t=0}^T$ (that represents, e.g., a bank account) and a risky *stock* $S = (S_t)_{t=0}^T$.

The bond is a real valued, positive deterministic process. As it is customary to do, the bond will be chosen as a "numéraire", that is, all processes at time t are discounted with respect to the value B_t of the process B at time t. In other words, this means that it is not restrictive to suppose the process B to be constant, and indeed in the sequel we shall suppose $B_t \equiv 1$.

The stock is a real valued, positive stochastic process on a suitable probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t=0}^T, P)$ that evolves as follows. Fix two positive real numbers $u > d$ and let $(\xi_t)_{t=1}^T$ be a sequence of independent and identically distributed random variables that take values in the set $\{d,u\}$ with law

$$p := P\{\xi_1 = u\} = 1 - P\{\xi_1 = d\}.$$

Then, given the initial value S_0 for the stock, the binomial model supposes its dynamics to be written as

$$S_{t+1} = S_t \xi_{t+1} \qquad t = 0, \dots, T-1.$$

This way, the evolution of the stock can be represented as an element of the set $\{d,u\}^T$, and its final state assumes the values $S_0 u^k d^{T-k}$ with law $P\{S_T = S_0 u^k d^{T-k}\} = \binom{T}{k} p^k (1-p)^{T-k}$. This is where the name of the model comes from. Note that a natural definition for the underlying probability space is

$$\begin{aligned} \Omega &:= \{d,u\}^T \\ \mathcal{F} &:= \mathcal{P}(\Omega), \qquad \mathcal{F}_t := \left\{\mathcal{P}(\{d,u\}^t) \times \{d,u\}^{T-t}\right\} \\ P(\omega) &:= p^\lambda (1-p)^{T-\lambda} \qquad \text{where } \lambda := \#\{t \mid \omega_t = u\}. \end{aligned}$$

An investor is supposed to trade in this market according to the following conditions. The investor can either deposit or borrow any amount of money from the bank account, that is, there is no limit on the number of units of bond B that can be traded at each step t. Note that, in the "non discounted" setting, this also means that we want the interest rate paid for a bank account to be the same as asked for a bank loan, without limits on their entity. Analogously, the investor can buy or sell any amount (positive or negative) of stock S at each time t. In particular, the so-called "short selling" is allowed, that is, it is possible to own a negative amount of stock. It is useful to mention the fact that we allow any *real* value for the amounts of both stock and bond owned by the investor. Finally, no costs are charged to the investor for the transactions, that is, any amount of money can be transferred to the bond from the stock or viceversa without any bank commissions. This is what is commonly referred to as a "frictionless" market.

We call *investment strategy* a sequence $\pi = (\pi_t)_{t=0}^{T-1} = \big((\alpha_t, \beta_t)_{t=0}^{T-1}\big)$ which represents the decision to hold, in each time interval $[t, t+1)$ up to the last $[T-1, T)$, α_t units of the stock S and β_t units of the bond B. (In particular, by analogy with [17], we choose the first of the two possible notations described in section 1, namely, the control is adapted to the filtration). The *portfolio* is defined as the set of the assets owned by the investor, and its *value* at time t is thus naturally defined by

$$V_t := \alpha_t S_t + \beta_t B_t = \alpha_t S_t + \beta_t$$

(recall that we consider α to be an adapted process). We want the investment strategies to be *self-financing*, that is, once the *initial capital* V_0 is fixed, the investor is not allowed to add nor to withdraw any amount of capital from the portfolio. This implies that any loss or gain is determined by the evolution of the assets (i.e., of the stock), and that at each time t the investor can only decide to redistribute the wealth he/she owns in such a way that

$$\alpha_t S_t + \beta_t = \alpha_{t-1} S_t + \beta_{t-1}.$$

As a consequence, the value of the portfolio can be expressed as a stochastic process that follows the dynamics

$$V_{t+1}^{\alpha} = V_t + \alpha_t S_t(\xi_t - 1), \tag{3.7}$$

where we use the notation V_{t+1}^{α} to stress the fact that under the self-financing conditions only the first first component of the strategy π affects explicitly the evolution of the portfolio.

A financial market is said to feature an *arbitrage opportunity* if there exists a self-financing investment strategy starting from $V_0^{\alpha} = 0$ such that

$V_T^\alpha \geqslant 0$ almost surely and $P\{V_T^\alpha > 0\} > 0$. Roughly speaking, this means that an investor can invest in such a way as to be sure not to "lose" anything and to have a positive probability to "gain" something. We want the market to be *arbitrage-free*, that is, there cannot be such an opportunity or, in other words, any investor willing to gain must take the risk of losing something. It is a classical result that the absence of arbitrage is equivalent to the existence of an *equivalent martingale measure*, that is, a measure P^* equivalent to P such that S is a martingale under the law P^*. In the binomial case, the market is arbitrage-free if and only if $0 < p < 1$, $0 < d < 1 < u$, and the unique equivalent martingale measure is defined by

$$p^* := P^*\{\xi_1 = u\} = \frac{1-d}{u-d}, \qquad 1 - p^* = P^*\{\xi_1 = d\} = \frac{u-1}{u-d}. \tag{3.8}$$

A *simple european contingent claim* is a real valued function of the final state $H\colon \mathbb{R}_+ \to \mathbb{R}$. As an example, the function $H(S_T) := (S_T - K)^+$ represents the payoff of an *european call option with strike price* K, that is, a contract between an investor and (say) a bank that gives the bearer the right (but not the obligation) to buy, at time T, one unit of stock at the fixed price K. The problem studied and solved by Cox, Ross and Rubinstein is that of *pricing* and *hedging* these claims in the binomial market, and can be explained as follows.

Consider an investor buying an european call option with strike price K. If the investor is paying nothing for this contract, then at time T he can realize an arbitrage by buying one unit of stock at price $\min\{S_T, K\}$ and immediately selling it at the price S_T. The user has then to pay a *fair price* $C_t^*(S_t)$ for buying at time t the contract from the bank in such a way that neither the investor nor the bank can realize an arbitrage by means of the contingent claim. To *price* an option means exactly to determine the fair price (or the set of fair prices) of that option.

The bank selling, say, an european call option with strike price K is in a risky position, in the sense that the contract is going to cost the bank $H(S_T) = (S_T - K)^+$ at time T. To avoid this risk, the bank wants to find a self-financing investment strategy $\alpha^H = (\alpha_s^H)_{s=t}^{T-1}$ that, starting from a value $V_t^*(S_t)$ (as low as possible) at time t, yields $V_T^{\alpha^H} = H(S_T)$ almost surely. If such a strategy exists, it is called *hedging* or *replicating strategy.*

A classical result is that the arbitrage-free price of an european contingent claim coincides with the optimal initial capital for the hedging strategy and is given by the *Cox, Ross and Rubinstein valuation formula*

$$C_t^*(S_t) := \mathrm{E}^*\{H(S_T) \mid S_t\}, \tag{3.9}$$

where $\mathrm{E}^*\{\cdot\}$ denotes the expectation with respect to the equivalent martingale measure P^* defined in (3.8). The hedging strategy is given as a function of the current time and the current value of the underlying stock as

$$\alpha_t^H(S_t) := \frac{C^*_{t+1}(S_t u) - C^*_{t+1}(S_t d)}{S_t(u-d)}, \tag{3.10}$$

and indeed, using the formula $\mathrm{E}^*\left\{\mathrm{E}^*\{H(S_T) \mid S_{t+1}\} \,\middle|\, S_t\right\} = \mathrm{E}^*\{H(S_T) \mid S_t\}$ it is straightforward to check that, if $V_t = C^*_t(S_t)$, then $V^{\alpha^H}_{t+1} = C^*_{t+1}(S_{t+1})$ with probability 1.

Sometimes it may happen that the arbitrage-free price of a claim is believed to be too high to make investors willing to pay it. In other situations, it may happen that the financial institution which sells the claim does not want to endow the whole sell price of the claim for the hedging strategy. In both cases, one is led to start an investment strategy with an initial capital $V_0 < V_0^*(S_0)$.

The results of Cox, Ross and Rubinstein together with the absence of arbitrage in the binomial market model ensure that no hedging strategy starting with capital V_0 can exist. More in detail, any self-financing strategy α starting with capital V_0 leads to a final value V_T^α that is smaller that $H(S_T)$ with positive probability. In financial terms, one says that the *shortfall* $(H(S_T) - V_T^\alpha)^+$ is greater than zero with positive probability.

A quite natural approach to this situation is the following. The user who aims at hedging the claim H will define a *loss function* $\ell\colon \mathbb{R}_+ \to \mathbb{R}_+$ to express his/her preferences about "negative" positions. In detail, the loss function will be increasing, such that $\ell(0) = 0$, and respectively concave or convex depending on the propension or aversion of the user to get into "risky" positions (i.e., situations such that $H(S_T) - V_T^\alpha \gg 0$). Once the function ℓ is fixed, the user will look for the strategy that minimizes the *shortfall risk*

$$J_0(S_0, V_0) := \mathrm{E}_{S_0,V_0}\{\ell\big([H(S_T) - V_T^\alpha]^+\big)\}. \tag{3.11}$$

The notion of shortfall risk has been known for quite some time. In the more economically oriented literature, it was also known under the name of "lower partial moments". It is present in the financial mathematics literature mainly since the fundamental paper [8] by Föllmer and Leukert has appeared.

More specifically, we aim at solving the discrete time stochastic optimal control problem

$$\begin{cases} \begin{pmatrix} S_{t+1} \\ V^\alpha_{t+1} \end{pmatrix} = \begin{pmatrix} S_t \xi_{t+1} \\ V_t + \alpha_t S_t(\xi_t - 1) \end{pmatrix} \\ \inf_\alpha \mathrm{E}_{S_0,V_0}\left\{\ell\big([H(S_T) - V_T^\alpha]^+\big)\right\} \end{cases}.$$

As a particular case, e.g., one might define the function $\ell(x) := 1_{(0,+\infty)}(x)$: in this case, the problem of minimizing the shortfall risk is the problem of minimizing the probability of positive shortfall.

As already mentioned in 2.2, Runggaldier et al. have solved the shortfall risk minimization problem for the binomial model in the case $\ell(x) = x$ (that we shall sometimes call the "mean shortfall minimization problem"). By explicitly solving the dynamic programming algorithm (see (1.4) for the definition), they have shown that the optimal control is computed according to the formula

$$\alpha_t^* = \begin{cases} \dfrac{C_{t+1}^*(S_t d) - V_t}{S_t(d-1)} & \text{if } p < p^* \\ \dfrac{C_{t+1}^*(S_t u) - V_t}{S_t(u-1)} & \text{if } p > p^* \end{cases} \tag{3.12}$$

and that the optimal value is

$$J_0^*(S_0, V_0) = \min\left\{\tfrac{p}{p^*}, \tfrac{1-p}{1-p^*}\right\}^T (C_0^*(S_0) - V_0)^+$$

(the case $p = p^*$ is not considered, as both the stock and the portfolio are martingales under the probability P, and so the optimal control is no longer unique).

This result has two noteworthy features. First of all, it is quite remarkable that the mean shortfall is exponentially decreasing to 0 in the horizon. Indeed, this apparently implies that an "almost perfect" hedging could be achieved with any initial capital, provided that the expiration date of the contract is "far" enough. The second feature is that, unlike the replicating strategy of Cox, Ross and Rubinstein seen in (3.10), the optimal strategy and the optimal value depend on the probability p.

As already mentioned in 2.2, the authors also explore the bayesian approach to the case of incomplete information, that is, the case in which p is unknown to the investor. Their result is that, if the prior distribution for p is proportional to $p^{\lambda_0}(1-p)^{\mu_0}$ and one sets $\lambda_t := \lambda_0 + \#\{s < t \mid \omega_s = u\}$, $\mu_t := \mu_0 + \#\{s < t \mid \omega_s = d\}$, then the bayesian optimal control is

$$\alpha_t^b = \begin{cases} \dfrac{C_{t+1}^*(S_t d) - V_t}{S_t(d-1)} & \text{if } \frac{\lambda_t+T-t+1}{\lambda_t+\mu_t+T-t+2} < p^* \\ \left[\frac{C_{t+1}^*(S_t d)-V_t}{S_t(d-1)}, \frac{C_{t+1}^*(S_t u)-V_t}{S_t(u-1)}\right] & \text{if } \frac{\lambda_t+1}{\lambda_t+\mu_t+T-t+2} < p^* < \frac{\lambda_t+T-t+1}{\lambda_t+\mu_t+T-t+2} \\ \dfrac{C_{t+1}^*(S_t u) - V_t}{S_t(u-1)} & \text{if } \frac{\lambda_t+1}{\lambda_t+\mu_t+T-t+2} > p^* \end{cases}$$

(see [17] for the details of the proof), where the "interval" notation means that, in the "undecided" case $\frac{\lambda_t+1}{\lambda_t+\mu_t+T-t+2} < p^* < \frac{\lambda_t+T-t+1}{\lambda_t+\mu_t+T-t+2}$, any value in the interval is an optimal choice for α_t. Note that, even if the bayesian approach would theoretically require to recalculate the dynamic programming algorithm at each t, the formula obtained allows the investor to calculate the optimal control on the basis of known values at each time t. This way, the bayesian approach is very efficient from a computational point of view. Indeed, its only flaw is the fact that, for a typical evolution of the portfolio, the choice of the optimal policy is "undecided" (i.e., $\frac{\lambda_t+1}{\lambda_t+\mu_t+T-t+2} < p^* < \frac{\lambda_t+T-t+1}{\lambda_t+\mu_t+T-t+2}$) until a few steps from the final date T.

The dependence upon p of the optimal control (3.12) led us to try to apply the robust approach studied in [6] (see 3.1) to the mean shortfall minimization problem in the binomial case. To do this, we suppose not to know the probability p and we solve the problem for an hypothetical measure Q with law $q := Q\{\xi_1 = u\}$. We want then to use the results of [6] (see 3.1) to give a robustness inequality (as defined in 1.2) based only on the Radon-Nikodym derivative of Q with respect to P.

Since the final state in this model is a discrete random variable, the "exponential" approach of [6] is not applicable. To estimate the robustness inequality (3.3) given by the "direct" approach, then, we have to calculate

$$\widetilde{\gamma} := \sup \left|1 - \frac{\mathrm{d}P}{\mathrm{d}Q}\right|, \qquad M_{S_0,V_0} := \sup_\alpha \mathrm{E}_{S_0,V_0} \left\{[H(S_T) - V_T^\alpha]^+\right\}.$$

The sup of $\left|1 - \frac{\mathrm{d}P}{\mathrm{d}Q}\right|$ can be immediately calculated by considering that the Radon-Nikodym derivative is concentrated at the points $S_0u^kd^{T-k}$ and that

$$\frac{\mathrm{d}P}{\mathrm{d}Q}\left(S_0u^kd^{T-k}\right) = \left(\frac{p}{q}\right)^k \left(\frac{1-p}{1-q}\right)^{T-k} = \left(\frac{1-p}{1-q}\right)^T \left(\frac{p(1-q)}{q(1-p)}\right)^k$$

is either increasing or decreasing in k depending on whether $\frac{p(1-q)}{q(1-p)}$ is greater or less than 1. As a consequence, $\left|1 - \frac{\mathrm{d}P}{\mathrm{d}Q}\right|$ is bounded from above by

$$\widetilde{\gamma} := \max\left\{\left|1 - \left(\frac{p}{q}\right)^T\right|, \left|1 - \left(\frac{1-p}{1-q}\right)^T\right|\right\}.$$

To determine M_{S_0,V_0} is a more complicated matter, as in general the expression $\mathrm{E}_{S_0,V_0}\{[H(S_T) - V_T^\alpha]_+\}$ is not bounded from above if we consider all the admissible choices for the control α. To get an upper bound on this quantity, then, we need to redefine the class of "admissible" controls, and this can be achieved by examining the proof of [17, Theorem 4.1]. Indeed, this proof makes clear that at each step t the argument of the inf in the dynamic programming algorithm is a piecewise linear function of α_t

which is strictly decreasing for $\alpha_t \leqslant \frac{C^*_{t+1}(S_t d)-V_t}{S_t(d-1)}$ and strictly increasing for $\alpha_t \geqslant \frac{C^*_{t+1}(S_t u)-V_t}{S_t(u-1)}$. So, the optimal control necessarily has to assume, at each time t, either the value $\frac{C^*_{t+1}(S_t d)-V_t}{S_t(d-1)}$ or $\frac{C^*_{t+1}(S_t u)-V_t}{S_t(u-1)}$ depending on whether, respectively, the function to be minimized is increasing or decreasing in the interval $\left[\frac{C^*_{t+1}(S_t d)-V_t}{S_t(d-1)}, \frac{C^*_{t+1}(S_t u)-V_t}{S_t(u-1)}\right]$. This way, we can define

$$\alpha_t^{(1)}(S_t, V_t) := \frac{C^*_{t+1}(S_t d) - V_t}{S_t(d-1)}, \qquad \alpha_t^{(2)}(S_t, V_t) := \frac{C^*_{t+1}(S_t u) - V_t}{S_t(u-1)}, \tag{3.13}$$

and consider only the investment strategies belonging to the class

$$\Pi := \big\{(\alpha_t)_t \mid \alpha_t \in \{\alpha_t^{(1)}(S_t, V_t), \alpha_t^{(2)}(S_t, V_t)\} \text{ for every } t\big\}. \tag{3.14}$$

A slight modification of the cited proof allows indeed to see that if a control $\alpha \in \Pi$ is chosen, one has

$$\mathrm{E}_{S_t,V_t}\left\{[H(S_T) - V_T^\alpha]^+\right\} = \Big(\frac{p}{p^*}\Big)^{\lambda_t}\Big(\frac{1-p}{1-p^*}\Big)^{T-t-\lambda_t}(C_t^*(S_t) - V_t)^+ \tag{3.15}$$

where $\lambda_t := \#\{s > t \mid \alpha_t = \alpha_t^{(1)}(S_t, V_t)\}$. (From now on, when there will be no ambiguity, we shall not make explicit the dependence of $\alpha_t^{(i)}$ on the current values of the stock and the portfolio, and write $\alpha_t^{(i)}$ instead of $\alpha_t^{(i)}(S_t, V_t)$.)

In particular, then,

$$M_{S_0,V_0} := \sup_\alpha \mathrm{E}_{S_0,V_0}\left\{[H(S_T) - V_T^\alpha]^+\right\} = \max\left\{\frac{p}{p^*}, \frac{1-p}{1-p^*}\right\}^T (C_0^*(S_0) - V_0)^+,$$

and the right hand side of the robustness inequality (3.3) can be calculated.

Note that the upper bound obtained this way can be dramatically big. Indeed, both $\widetilde{\gamma}$ and M_{S_0,V_0} are exponentially increasing in the horizon T, and this seems to suggest that the optimal strategy for the nominal/approximating problem (Q) could lead to great errors when used to control (P). This impression is supported by observing that, by (3.15) and the expression for the optimal control found by Runggaldier et al. as reported in (3.12), if p and q lie on opposite sides of p^* then the Q-optimal control is indeed the "P-worst" control, i.e., the one that realizes the sup in the calculation of M_{S_0,V_0}.

This allows also to better understand the results of the bayesian approach reported above. It is indeed clear that, when the investor is not sure on whether p is greater or less than p^*, choosing either $\alpha = \alpha^{(1)}$ or $\alpha = \alpha^{(2)}$ could mean choosing the "worst" control in the sense explained above. This explains the "undecidedness" of the bayesian control in the case $\frac{\lambda_t+1}{\lambda_t+\mu_t+T-t+2} < p^* < \frac{\lambda_t+T-t+1}{\lambda_t+\mu_t+T-t+2}$. On the other hand, it seems that the bayesian approach, being based on the dynamic programming algorithm with an unknown value for p, could lead to big errors in the average.

These considerations seem to suggest that one should look for an adaptive control that minimizes the shortfall risk on a "step-by-step" basis, instead of relying on the DP algorithm. A way to formulate such a control comes from (3.14) and (3.15). From (3.15) comes indeed that, at each time t, the effect of choosing $\alpha_t = \alpha_t^{(1)}$ (respectively, $\alpha_t = \alpha_t^{(2)}$) is to multiply the shortfall risk by $\frac{p}{p^*}$ (respectively, $\frac{1-p}{1-p^*}$). As a consequence, one is led to choose, respectively, $\alpha_t = \alpha_t^{(1)}$ or $\alpha_t = \alpha_t^{(2)}$ depending on whether, given the information gathered up to time t, he/she believes "more likely" to be $\frac{p}{p^*} < \frac{1-p}{1-p^*}$ (i.e., $p < p^*$) or $\frac{p}{p^*} > \frac{1-p}{1-p^*}$ (i.e., $p > p^*$). So, defining $\mathcal{F}_t^S := \sigma\{S_s \mid s \leqslant t\}$ the σ-algebra that represents the information gathered up to time t and supposing that at each time t we can estimate $r_n := P\{p \leqslant p^* \mid \mathcal{F}_t^S\}$, we have proved that the best adaptive choice in the class Π to minimize the shortfall risk is the control

$$\alpha_t^\dagger = \begin{cases} \alpha^{(1)} & \text{if } r_n \geqslant 1/2 \\ \alpha^{(2)} & \text{if } r_n > 1/2 \end{cases}.$$

To investigate the behaviour of the strategies seen so far (namely: the "deterministic" strategy α^* defined in (3.12) for the complete information case, the "robust" strategy α^Q defined as α^* for an hypothetical value q used instead of p, the bayesian strategy α^b of [17] and the adaptive strategy $\alpha^\dagger$ just defined), a computer program written by the author has been used. In the case of an european call option, the program randomly determines a possible evolution of the stock and computes the evolution of the portfolio according to the four investment strategies taken into consideration. When the bayesian strategy is "undecided", the program uses as a value for the optimal control the middle point of the interval of possible choices.

The typical output of the program is shown in [5, Figure 1], and the outcomes were, at first, hard to interpret. In most of the simulation runs, indeed, all controls gave coincident portfolio evolutions and yielded perfect hedging of the claim, even in the case $(q-p^*)(p-p^*) < 0$ where π^Q is the P-worst control. The only exception to this situation is the bayesian strategy, which also leads to perfect hedging but coincides with the others only from some t on.

Trying to understand how the "best" and the "worst" controls could yield the same results led to a property of the strategy α^* defined in (3.12) that was not stressed in [17]. One starts by showing, by using straightforward calculation, that $\alpha_t^{(1)}(s,v) = \alpha_t^{(2)}(s,v)$ if and only if $v = C_t^*(s)$ (see (3.9)), and that in this case they both coincide with the replicating control $\alpha_t^H(s) = \frac{C_{t+1}^*(su)-C_{t+1}^*(sd)}{s(u-d)}$ of Cox, Ross and Rubinstein (see (3.10)). Hence it follows that if $V_t = C_t^*(S_t)$ for some t, then choosing either $\alpha_s = \alpha_s^{(1)}$ or $\alpha_s = \alpha_s^{(2)}$ for $t \leqslant s \leqslant T$ is the same as following the CRR hedging strategy, and thus leads to perfect hedging. Another straightforward calculation shows that if $\alpha_t = \alpha_t^{(1)}$ and $\xi_{t+1} = d$ then $V_{t+1} = C_{t+1}^*(S_{t+1})$, and analogously if $\alpha_t = \alpha_t^{(2)}$ and $\xi_{t+1} = u$ then $V_{t+1} = C_{t+1}^*(S_{t+1})$.

These results show that the controls belonging to the class Π of (3.14) are "quasi-replicating" in the following sense. Given a control $\alpha \in \Pi$, define $\omega(\alpha)$ by

$$\big(\omega(\alpha)\big)_n := \begin{cases} u & \text{if } \alpha_n = \alpha_n^{(1)} \\ d & \text{if } \alpha_n = \alpha_n^{(2)} \end{cases}. \tag{3.16}$$

Then *any strategy* $\alpha \in \Pi$ *gives perfect hedging on all events* $\omega \neq \omega(\alpha)$. In particular, if one defines $\lambda_T^\alpha := \#\{t \mid \alpha_t = \alpha_t^{(1)}\}$, any strategy $\alpha \in \Pi$ leads to perfect hedging with probability $1 - p^{\lambda_T^\alpha}(1-p)^{T-\lambda_T^\alpha}$, which is next to 1 when T is big. Moreover, it is straightforward to derive from this result and (3.15) that in the "critical" event $\omega(\alpha)$ one has

$$H(S_T(\omega(\alpha))) - V_T^\alpha(\omega(\alpha)) = \Big(\frac{1}{p^*}\Big)^{\lambda_T}\Big(\frac{1}{1-p^*}\Big)^{T-\lambda_T}[C_0^*(S_0) - V_0]^+.$$

This explains why the deterministic strategy α^* and the robust strategy α^Q behave the same way in most events (namely, on all $\omega \notin \{\omega(\alpha^*), \omega(\alpha^Q)\}$), and supports the decision to restrict the class of admissible controls to Π. On the other hand, it is noteworthy that in the "critical" event, the final shortfall is exponentially increasing in the horizon. As we shall see in the next section, this is typical of the so-called "risk prone" situation, which classically refers to the case in which the function ℓ in (3.11) is concave. In other words, the results of [5] together with those of [7] also show that the case $\ell(x) = x$ can be considered a particular case of the case in which ℓ is concave.

3.3 Extensions of the previous results

As seen in the previous section, the results of [5] (and those of [17], upon which [5] is based) refer to the "mean shortfall minimization problem", i.e., the particular case $\ell(x) = x$ for the shortfall risk minimization problem as defined in 3.11. A natural question is whether some similar properties hold for a more general choice of the function ℓ, and the work currently in progress (see [7]) is an investigation into this question. This section is written in the same setting and with the same notations of the previous one.

The dynamic programming algorithm proved to be a master tool in solving the mean shortfall minimization problem. Also, many of the results of [5] (e.g., the "quasi-replicating" property of the controls belonging to the class Π defined in (3.14)) come in a more or less direct way from the DP algorithm itself. It is then reasonable that the investigation about the general shortfall risk minimization problem starts from understanding which properties of the optimal cost-to-go can be deduced from the DP approach. The research is then conducted separately in the "concave" (i.e., risk prone investor) and in the "convex" (i.e., risk averse investor) case, as it is reasonable to suppose that investors with opposite attitudes towards the possibility of getting into risky positions will choose strategies with different behaviours.

The final goal of our study is to show that, as it is reasonable to expect, some continuity properties of optimal values and optimal strategies with respect to the form of the loss function hold. Moreover, we also want to investigate the robustness (in the classical sense of the robustness problem defined in section 1.2) of the optimal strategies with respect to different loss functions. This part of the work, together with the investigation about the properties of the optimal strategies/values in the convex case, is still in progress.

The dynamic programming algorithm

The dynamic programming algorithm in the shortfall risk minimization problem takes the form

$$\begin{aligned} J_T(s,v) &:= \ell\big([H(s)-v]^+\big) \\ J_t(s,v) &:= \inf_\alpha \mathrm{E}\{J_{t+1}(S_{t+1},V^\alpha_{t+1}) \mid S_t = s, V^\alpha_t = v\},\ n = 0,\dots,N-1, \end{aligned}$$

and the recursive step can be written more explicitly as

$$J_t(s,v) = \inf_\alpha \{pJ_{t+1}(su, v+\alpha s(u-1)) + (1-p)J_{t+1}(sd, v+\alpha s(d-1))\}.$$

For simplicity of notation, we use the intermediate recursive definition

$$\begin{aligned} j^u_t(s,v,\alpha) &:= pJ_t(su, v+\alpha s(u-1)) \\ j^d_t(s,v,\alpha) &:= (1-p)J_t(sd, v+\alpha s(d-1)), \end{aligned} \tag{3.17}$$

so that $J_t(s,v) := \inf_\alpha\{j^u_{t+1}(s,v,\alpha) + j^d_{t+1}(s,v,\alpha)\}$.

From the assumption that ℓ is increasing, it is straightforward to prove by backwards induction that $j_t^u(s,v,\alpha)$ and $j_t^d(s,v,\alpha)$ are decreasing in v and, respectively, increasing and decreasing in α for every $t = 1,\ldots,T$. Hence one obtains the immediate conclusion that also $J_t(s,v)$ is decreasing in α for every t. Moreover, since $J_T(s,v) = 0$ for $v \geqslant H(s) = C_T^*(s)$ (where $C_t^*(\cdot)$ is the CRR valuation formula as defined in (3.9)), carrying out some calculations it can be shown that for every $t = 0,\ldots,T$ it is

$$\begin{aligned} j_{t+1}^u(s,v,\alpha) = 0 \qquad & \text{for } \alpha \leqslant \alpha_t^{(2)}(s,v), \\ j_{t+1}^d(s,v,\alpha) = 0 \qquad & \text{for } \alpha \geqslant \alpha_t^{(1)}(s,v) \end{aligned} \tag{3.18}$$

where

$$\alpha_t^{(1)}(s,v) := \frac{C_{t+1}^*(sd) - v}{s(d-1)}, \qquad \alpha_t^{(2)}(s,v) := \frac{C_{t+1}^*(su) - v}{s(u-1)}$$

are defined analogously to (3.13). Since a straightforward calculation shows that $\alpha_t^{(1)}(s,v) < \alpha_t^{(2)}(s,v)$ if and only if $v < C_t^*(s)$, from this result it follows that $J_t(s,v) = 0$ for $v \geqslant C_t^*(s)$, because in this case it can be checked that $j_{t+1}^u(s,v,\alpha) = j_{t+1}^d(s,v,\alpha) = 0$ for any $\alpha \in [\alpha_t^{(2)}(s,v), \alpha_t^{(1)}(s,v)]$.

Equation (3.18), together with the monotonicity of j_t^u and j_t^d with respect to α, also shows that $j_t^u(s,v,\alpha)+j_t^d(s,v,\alpha)$ is increasing in α for $\alpha \geqslant \alpha_t^{(1)}(s,v)$ and decreasing in α for $\alpha \leqslant \alpha_t^{(2)}(s,v)$. This implies that, when $v < C_t^*(s)$, $J_t(s,v)$ can be calculated according to the formula

$$J_t(s,v) = \inf_{\alpha \in [\alpha_t^{(1)}(s,v), \alpha_t^{(2)}(s,v)]} \{ j_{t+1}^u(s,v,\alpha) + j_{t+1}^d(s,v,\alpha) \}. \tag{3.19}$$

In particular, since we shall prove that j_t^u and j_t^d are continuous in α for every t both in the concave and in the convex cases, the inf is realized as a min for some $\overline{\alpha} \in [\alpha_t^{(1)}(s,v), \alpha_t^{(2)}(s,v)]$, so that this result has as an immediate consequence the existence of an optimal strategy.

Risk prone investor, ℓ concave

Since the expected value is a linear functional and the minimum of linear functions is a concave function, it is straightforward to prove that when $\ell(x)$ is a concave function on $\mathbb{R}_+$, then $J_t(s,v)$ is also concave in v on the half-line $(-\infty, C_t^*(s)]$ for every $t = 0,\cdot,N$. Hence it follows that when $v < C_t^*(s)$, then $j_{t+1}^u(s,v,\alpha) + j_{t+1}^d(s,v,\alpha)$ is concave with respect to α on the interval $[\alpha_t^{(1)}(s,v), \alpha_t^{(2)}(s,v)]$.

It is well known that a concave function on an interval can attain its minimum value only at the extremal points, so that formula (3.19) allows us to conclude that we can restrict the class of admissible controls to the class Π defined in (3.14). In particular, analogously to the case $\ell(x) = x$, the admissible controls for this problem are "quasi-replicating" as shown at the end of the previous section.

To determine the optimal strategy in the concave case is a more complicated matter than in the case of the mean shortfall minimization. Indeed, the condition to choose between $\alpha_t^* = \alpha_t^{(1)}$ and $\alpha_t^* = \alpha_t^{(2)}$ is no longer as simple as it is in the case $\ell(x) = x$. (see (3.12)). By backwards induction based on the DP algorithm, it is still possible to determine the shortfall risk associated to every admissible strategy $\alpha \in \Pi$ as

$$\mathrm{E}_{S_t,V_t}\left\{\ell\big([H(S_T) - V_T^{\alpha}]^+\big)\right\} = p^{\lambda_t}(1-p)^{T-t-\lambda_t}\,\ell\left(\frac{C_t^*(S_t) - V_t}{(p^*)^{\lambda_t}(1-p^*)^{T-t-\lambda_t}}\right)$$

where, as in previous section, $\lambda_t := \#\{s > t \mid \alpha_s = \alpha_s^{(1)}\}$. (As in the previous section, we still use the convention to write for short $\alpha_t^{(i)}$ instead of $\alpha_t^{(i)}(S_t, V_t)$.)

A straightforward consequence is that the optimal value for every $t = 0, \ldots, T$ is

$$J_t(S_t, V_t) = \min_{k=t,\ldots,T} p^{k-t}(1-p)^{T-k}\,\ell\left(\frac{C_t^*(S_t) - V_t}{(p^*)^{k-t}(1-p^*)^{T-k}}\right).$$

Hence it might seem that to determine the optimal strategy it is necessary to find, at each time t, the minimum of all the possible $T-t+1$ "outcomes" and to choose a strategy accordingly. Unfortunately, if $\overline{k}_t$ realizes the min in the above formula, any strategy $\alpha \in \Pi$ such that $\#\{s > t \mid \alpha_s = \alpha_s^{(1)}\} = \overline{k}_t - t$ is optimal, and then, since there are $\binom{T-t}{\overline{k}_t - t}$ such strategies, this result is still insufficient to give an algorithm to determine the choice of the optimal strategy. (Indeed, as it is shown with an example, it might happen that the minimizer $\overline{k}_t$ is different both from t and T. This was not the case in the mean shortfall minimization, because – as it can be deduced from (3.12) – the optimal strategy is to choose either $\alpha_t^* = \alpha_t^{(1)}$ for every t or $\alpha_t^* = \alpha_t^{(2)}$ for every t, which respectively correspond to the cases in which, for every t, either $\overline{k}_t = t$ or $\overline{k}_t = T$.)

As a possible solution to determine an optimal control, we propose an algorithm based on a "branch and bound" like procedure, and show that the optimal control can be computed by the formula

$$\alpha_t^* = \begin{cases} \alpha_t^{(1)}(S_t, V_t) & \text{if } p^{T-t}\ell\left(\frac{C_t^*(S_t)-V_t}{(p^*)^{T-t}}\right) \leqslant (1-p)^{T-t}\ell\left(\frac{C_t^*(S_t)-V_t}{(1-p^*)^{T-t}}\right) \\ \alpha_t^{(2)}(S_t, V_t) & \text{if } p^{T-t}\ell\left(\frac{C_t^*(S_t)-V_t}{(p^*)^{T-t}}\right) > (1-p)^{T-t}\ell\left(\frac{C_t^*(S_t)-V_t}{(1-p^*)^{T-t}}\right) \end{cases}.$$

Since this algorithm yields the optimal control in linear time with respect to the horizon, it is also optimal from a computational point of view. For the exact definition of the algorithm and details about the proof that it actually leads to the optimal value, we refer to [7].

Risk averse investor, ℓ convex

It is reasonable to expect that, by analogy with the concave case, $J_t(s, v)$ is convex in v for every t when ℓ is convex. This property is nevertheless not as straightforward to obtain as its concave counterpart, because (as already observed) the minimum is a strictly concave operator.

To prove the convexity of the optimal cost-to-go, we start by considering that, since ℓ is convex, it is possible to show that the functions j_T^u and j_T^d defined as in (3.17) are convex in v and α in the sense that for every s and for every convex combination $(\overline{v}, \overline{\alpha}) = \lambda_1(v_1, \alpha_1) + \lambda_2(v_2, \alpha_2)$, it is

$$j_T^{\dagger}(s, \overline{v}, \overline{\alpha}) \leqslant \lambda_1 j_T^{\dagger}(s, v_1, \alpha_1) + \lambda_2 j_T^{\dagger}(s, v_2, \alpha_2)$$

(with $\dagger = u, d$). In particular, then, this ensures that the inf in the recursion to determine J_{T-1} according to the dynamic programming is realized as a min for some α. Take now every convex combination $v = \lambda' v' + \lambda'' v''$, and choose α, α' and α'' so as to minimize, respectively, $[j_T^u + j_T^d](s, v, \cdot)$, $[j_T^u + j_T^d](s, v', \cdot)$ and $[j_T^u + j_T^d](s, v'', \cdot)$. By exploiting the proved convexity of j_T^u and j_T^d and the minimality of α, one then has

$$\begin{aligned} \lambda' J_{T-1}(s, v') + \lambda'' J_{T-1}(s, v'') &= \lambda'[j_T^u + j_T^d](s, v', \alpha') + \lambda'[j_T^u + j_T^d](s, v'', \alpha'') \\ &\geqslant [j_T^u + j_T^d](s, v, \lambda'\alpha' + \lambda''\alpha'') \\ &\geqslant [j_T^u + j_T^d](s, v, \alpha) = J_{T-1}(s, v), \end{aligned}$$

which proves the convexity of J_{T-1} as desired. The convexity of $J_t(s, v)$ with respect to v for every t is proved by backwards induction in a similar way.

As already noticed, this has as an immediate consequence the existence of an optimal solution for the shortfall risk minimization problem in the convex case. Moreover, if ℓ is strictly convex, J_t is also strictly convex for every t, and this allows to conclude that the optimal strategy is unique. The investigation about the properties of these optimal strategies is, as already mentioned at the beginning of this section, still in progress.

Acknowledgements

Tiziano Vargiolu, researcher in Padua, has managed to employ for years a huge patience to bear my tales about astounding progresses and insurmountable obstacles in my research. It is in my nature to need constant dialogue and encouragement to move on, and this is why I am quite sure that without Tiziano's help this thesis would never have been possible. For the same reason, even if it is not customary to mention the thesis supervisor in the acknowledgements, professor **Wolfgang Runggaldier** deserves some special thanks. More than that, it is right and proper to mention my gratitude for him having stood all the times that some unlucky event delayed or damaged my progresses. Indeed, I cannot be sure that he did not lose the trust in my possibilities in spite of all these events; anyway, even if he did, he surely managed not to let me understand – and that is worth quite a lot more than two pence to me. I am also in debt with **Evgueni Gordienko** and **Matthew James**: the ideas that led to [6] arose mainly in discussions with them.

While Runggaldier and Vargiolu just had to bear me during office hours in the last four years, my parents **Adriana** and **Giancarlo** were forced to deal with me each and every day of my life, i.e, for almost twenty-eight years. I am sure that this cannot have been easy, and that it can have been possible just thanks to a whole lot of love that I cannot but hope to deserve. I want also to thank my grandmother **Adele** and my uncles and aunts **Antonella and Paolo** and **Gastone and Marilena**, who never saved themselves any trouble to help and support me. Indeed, my whole family has always been very united and we never ceased to love one another; I think that I will never be able to find the words to say how much this is important to me. I also remember my grandmother **Ottavia** and my grandfather **Giovanni**, now disappeared but always part of the wonderful family of which I have the chance to be a member.

Anyway, the one that maybe has stood me the most in the last two years is my beloved girlfriend **Fabrizia**. Even if it must have been harder for her than for anyone else, she always managed to make me feel all her love and dedication; once again, I cannot but hope to deserve her being near for a very long time to come, and to be able to love her back as deep as she will ever need. I am also in big debt with her parents **Maurizio** and **Milena**, who often make me feel more like a godson than a future son-in-law. I think it is enough to mention that their hospitality came to the point of lending me a room of their house, where the greatest part of this thesis has been written.

It is worth to remember the years I spent in the paduan “Goliardia”. From what I have already said, it might be already clear that my self-esteem is not as big as one might want; anyway, it would have been even worse without my goliardic past. Moreover, in those years I had the chance to practice a hard training to the ancient art of finding the way home while completely drunk at late night: without that surviving skills, the all-nighters spent at research and at writing the thesis would certainly have been harder than they have been. In the last years, baseball (first with Limena team, now with Padua amateurs) and role-playing game (Dungeons & Dragons with friends from Thiene and Padua) have also generated great moments of psycho-physic outlets which have always been quite valuable.

I want to thank also the guys at **Associazione Dottorandi e Dottori di Ricerca Italiani** (ADI – Italian Ph. D. Association) who have carried out an impressive amount of good work since their association was born. Moreover, many of them have become nice e-mail friends: with them we decided to found the “ADI-sinibiti” mailing list, which is now a daily source of good humour. The **Università degli Studi di Padova**, Padua University, and its **Dipartimento di Matematica Pura e Applicata**, Pure and Applied Mathematics Department, also deserve to be mentioned for the rooms and materials they made available to me as a Ph. D. student. In particular, I want to thank **Elena**, **Marinella** and, above all, **Pierina**, for their constant helpfulness even to my most incredible requests. A special mention is deserved by my fellows doctoral students **Clara**, **Mattia** and **Gherardo**, who have beared my excesses with surprising patience. Gherardo, in particular, has become a great friend of mine also outside the university environment, and his presence and advice are almost irreplaceable.

Finally, I thank everyone who has believed in my possibilities and supported me. In spite of my pessimism, I have to admit that they are quite a lot, surely too numerous to mention individually. I also want to thank the – not too few – people that wish me anything evil: surely it is their right to do so, and they make me remember that there are people in the world whose absence can have no effect on one’s happiness.

Bibliography

[1] J. Cvitanić, I. Karatzas, *On dynamic measures of risk*, Finance & Stochastics 3(4), 451-482 (1999).

[2] P. Dupuis, M. R. James, I. Petersen, *Robust properties of risk-sensitive control*, Math. Control, Systems & Signals 13, 318-332 (2000).

[3] P. Dai Pra, L. Meneghini, W. J. Runggaldier, *Explicit Solutions for multivariate, discrete-time control problems under model uncertainty*, Systems & Control Letters 34, 169-176 (1998).

[4] P. Dai Pra, C. Rudari, W. J. Runggaldier, *On dynamic programming for sequential decision problems under a general form of uncertainty*, Math. Meth. Oper. Res. 45, 81-107 (1997).

[5] G. Favero, *Shortfall risk minimization under model uncertainty in the binomial case: adaptive and robust approaches*, Math. Meth. Oper. Res. 53, 493-503 (2001).

[6] G. Favero, W. J. Runggaldier, *A Robustness Result for Stochastic Control*, 39th IEEE Conference on Decision and Control proceedings, "short paper" n. 1024 (vol. 1, 3349-3350), Sydney (2000). Also Università di Padova, Dipartimento di Matematica Pura e Applicata, Preprint n. 21 (2000).

[7] G. Favero, T. Vargiolu, *Shortfall risk minimization in the binomial model* (temporary title), work in progress.

[8] H. Föllmer, P. Leukert, *Efficient hedging: cost versus shortfall risk*, Finance and Stoch. 4, 117-146 (2000).

[9] W. H. Fleming, R. W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York (1975).

[10] E. Gordienko, E. Lemus-Rodríguez, *Estimation of robustness for controlled diffusion processes*, Stoch. Anal. and Appl. 17 (3), 421-441 (1999).

[11] E. Gordienko, F. S. Salem, *Robustness inequality for Markov control processes with unbounded cost*, Systems & Control Letters 33, 125-130 (1998).

[12] E. Gordienko, F. S. Salem, *Estimates of stability of Markov control processes with unbounded cost*, Kybernetica 36 (2), 195-210 (2000).

[13] K. Hinderer, *Foundations of Non-stationary Dynamic Programming with Discrete Time Parameter*, Springer-Verlag, Berlin, Heidelberg (1970).

[14] L. Lazrieva, T. Toronjadze, *Stochastic volatility model with small randomness. Robust parameter estimate and hedging*, A. Razmadze Mathematical Institute of Georgian Academy of Sciences and Caucasus School of business, report LT-001 (2001).

[15] M. Musiela, M. Rutkowski, *Martingale Methods in Financial Modelling*, Springer-Verlag, Berlin (1997).

[16] W. J. Runggaldier, *Concepts and methods for discrete and continuous time control under uncertainty*, Insurance: Math. and Econ. 22, 25-39 (1998).

[17] W. J. Runggaldier, B. Trivellato, T. Vargiolu, *A Bayesian Adaptive Control Approach to Risk Management in a Binomial Model*, to appear in the Proceedings of the Ascona '99 Seminar on Stochastic Analysis, Random Fields and Applications, Birkhauser (2001).

[18] D. Talay, Z. Zheng, *Worst case model risk management*, INRIA internal report (2001).