A stochastic control approach to risk management under restricted information

Wolfgang J. Runggaldier
Dipartimento di Matematica Pura ed Applicata,
Universitá di Padova, Via Belzoni 7, I - 35131 - Padova, Italy
runggal@math.unipd.it

Anna Zaccaria
ARCA SGR S.p.A.
Via Mosé Bianchi 6, I - 20149 - Milano, Italy
anna.zaccaria@arcafondi.it

Abstract

We use techniques from discrete-time stochastic control under partial state information to determine a shortfall-risk minimizing investment strategy in the case when there is only restricted information on the underlying market model and transaction costs as well as shortselling constraints are present. The approach is adaptive in the sense that it takes into account all the information on the underlying model that becomes successively available to an economic agent by observing the prices in the market. As an immediate byproduct of the approach one can determine the entire shortfall distribution corresponding to the optimal strategy and to various values of the initial capital.

Keywords : Risk management, shortfall risk minimization, restricted information, transactions costs, dynamic programming.

1 Introduction

In incomplete markets the superhedging criterion allows one to eliminate completely the risk, but it requires in general too much initial capital; it corresponds in fact to a min-max-type criterion. One may then ask by how much one can lower the initial cost by accepting some risk or, dually, what is the risk corresponding to an initial capital less than what is required for superhedging. The shortfall risk minimization approach allows one to deal with these issues.

Given a market with a non-risky and a certain number of risky assets, let H_T be a liability to be hedged at some fixed future time T. Denote by $V_T(\psi)$ the value, at T, of a portfolio corresponding to a self-financing investment strategy ψ, possibly satisfying some additional constraints such as a shortselling prohibition. The problem is to find

$$J_0(S_0, V_0) := \inf_{\psi} E^P_{S_0,V_0}\left\{\ell\left([H_T - V_T(\psi)]^+\right)\right\} \tag{1}$$

for a given initial value S_0 of the asset(s) in the portfolio, for a given initial capital V_0 and where $\ell(\cdot) \geq 0$ is a suitable loss function with $\ell(0) = 0$. For $\ell(z) = 1_{\{z>0\}}$, the right hand side in (1) corresponds to the smallest shortfall probability. Problems of the type (1) have attracted recently considerable attention (see e.g. Cvitanić and Karatzas (1998), Cvitanić (1998a), Cvitanić (1998b), Föllmer and Leukert (1998), Föllmer and Leukert (1999), Pham (1998), Spivak and Cvitanić (1998)). Letting, for given S_0,

$$V_0^*(S_0) := \inf\{V_0 \mid J_0(S_0, V_0) = 0\} \tag{2}$$

it follows that $V_0 > V_0^*(S_0)$ implies $V_T(\psi^*) \geq H_T$, $P-$a.s. (*superhedging*) with ψ^* the optimal strategy in (1). For superhedging, the choice of the underlying probabilistic model for the evolution of the (risky) assets is thus irrelevant as long as it induces probability measures that are equivalent. However, for the more general problem of risk minimization in (1) the probabilistic structure of the underlying model matters, but the true model is almost never known exactly. A possibility is then to use a min-max-type criterion as in Cvitanić and Karatzas (1998) trying to find

$$\inf_{\psi} \sup_{P\in\mathcal{P}} E^P_{S_0,V_0}\left\{\ell\left([H_T - V_T(\psi)]^+\right)\right\}, \tag{3}$$

where $\mathcal{P}$ is a family of "real world probability measures". However, such a criterion does not allow to incorporate additional information on the underlying model as it becomes successively available.

An adaptive approach, corresponding to a Bayesian-type criterion, appears thus more appropriate. Such adaptive approaches have already been dealt with in the literature (see e.g. Cvitanić and Karatzas (1998), Cvitanić (1998a), Cvitanić (1998b), Spivak and Cvitanić (1998) and, in the context of portfolio optimization, in Browne and Whitt (1996), Lakner (1995), Lakner (1998), Karatzas and Zhao (1999)). In all these papers the uncertainty is only in the stock appreciation rate. The tools are mainly probabilistic in nature (involving also measure transformation) and are furthermore based on convex duality. An explicit solution is essentially possible only in simpler cases and transaction costs are not taken into account.

In this paper we follow an approach based on stochastic control under partial state observation as well as on stochastic adaptive control. This allows one to treat with essentially the same degree of difficulty : general uncertainties in the model, transactions costs, portfolio constraints, and a variety of risk criteria. An important byproduct of our approach is that, with little extra effort, one can determine the entire shortfall distribution corresponding to the optimal strategy ψ^* and to various values V_0 of the initial capital. This allows the economic agent to fully assess the impact of various choices of the initial capital.

In the next section 2 we describe our general model that takes possible uncertainties into account. In section 3 we present the solution in the case of full as well as of restricted information on the uncertain quantities in the model. In section 4 we consider a particular case (a binomial model) for which we discuss numerical results in subsection 4.1.

2 The general model

Consider a market with an underlying risky asset, in addition to a nonrisky one, where the price S_t of the risky asset evolves according to

$$dS_t = a(Z_t)\, S_t dt + \sigma(Z_t)\, S_t dw_t. \tag{4}$$

For simplicity we consider only a single risky asset (S_t is scalar), but the procedure described below can be extended without substantial difficulties to a vector S_t. The process Z_t is assumed to be a discrete-time (with step Δ), finite-state and homogeneous Markov chain, independent of the Wiener process $\{w_t\}$ and with given transition probability matrix $P = \{p_{i,j}\},\ \ (i, j = 1, \cdots, K)$. Contrary to S_t, the process $\{Z_t\}$ may not be observable. With the transition probabilities for $\{Z_t\}$ given, (4) represents a probabilistically well defined model. Notice, however, that different trajectories of Z_t generate different models for S_t that, while leading to the same solution for the superhedging problem (they all induce probability measures that are mutually equivalent), lead to different solutions when the problem is that of risk minimization as in (1). Therefore, if Z_t is not observable, it is important to use all the information as it becomes successively available from the observations of $\{S_t\}$ in order to get the best assessment concerning $\{Z_t\}$.

Suppose that the time step Δ corresponds to the smallest "trading interval". If trading takes place at discrete time steps Δ, then it suffices to have also the dynamics of S_t over the same time grid. With $S_n := S_{n\Delta}$, from (4) one has

$$S_{n+1} = S_n \exp\left\{\left[a(Z_n) - \frac{1}{2}\sigma^2(Z_n)\right]\,\Delta + \sigma(Z_n)\,\Delta w_n\right\} := S_n \cdot \omega_n \tag{5}$$

where, conditionally on Z_n, ω_n is log-normal, i.e.

$$\log \omega_n \sim \mathcal{N}\left(\left[a(Z_n) - \frac{1}{2}\sigma^2(Z_n)\right]\,\Delta\,;\ \sigma^2(Z_n)\Delta\right). \tag{6}$$

Letting N be such that $T = N\Delta$, the problem becomes that of determining

$$J_0(S_0, Z_0, V_0) = \min_{\psi \in A}\ E^P_{S_0,Z_0,V_0}\left\{\ell\left([H_N - V_N(\psi)]^+\right)\right\}, \tag{7}$$

where S_0, Z_0, V_0 are given initial values. Denoting by ψ_n^0 and ψ_n^1 the amount and the number of units of the nonrisky and risky assets respectively, that are held in the portfolio in period n (here and in the sequel we consider as "$n-$th period" the interval $[n\Delta, (n+1)\Delta)$), the admissible set A of strategies $\psi_n = (\psi_n^0, \psi_n^1)$ corresponds to requiring, first, a self financing property, i.e.

$$
\left\{\begin{array}{lcl}
\psi_0^0 + \psi_0^1 S_0 & = & V_0 \\[2ex]
\psi_n^0(1+r) + \psi_n^1 S_n & \geq & \psi_{n+1}^0 + \psi_{n+1}^1 S_n + \gamma\,|\psi_{n+1}^1 - \psi_n^1|\, S_n \\[2ex]
 & & \qquad\qquad\qquad\qquad n = 0, \cdots\cdots, N-1,
\end{array}\right. \tag{8}
$$

where r is the (constant) interest rate and γ is the transactions cost rate. In addition to this, for admissibility we shall also require the shortselling constraints

$$
\psi_n^0\,,\ \psi_n^1 \geq 0 \qquad \text{for} \quad n = 0, \cdots, N-1. \tag{9}
$$

Finally, we assume that $\psi_n = (\psi_n^0, \psi_n^1)$ is predictable with respect to the observation $\sigma-$algebra $\mathcal{F}_n^S := \sigma\{S_m,\ m \leq n\}$ (notice that ψ_n^1 is integer-valued).

Having now a discrete-time model, our solution method will be based on backwards *Dynamic Programming* (DP), analogous to one of the algorithms in Edirisinghe et al. (1993) for superhedging with transactions costs. Let

$$
V_n^+ := \psi_n^0(1+r) + \psi_n^1 S_n, \tag{10}
$$

which represents the "incoming value" of the portfolio in period $n+1$, and as "state" X_n in period n take

$$
X_n = (S_n, Z_n, V_n^+, \psi_n^1). \tag{11}
$$

The admissibility condition ($\psi \in A$) can then be expressed as

$$
\left\{\begin{array}{lclcl}
\psi_0^1 & \in & \Gamma_0(S_0, V_0) & := & \{\psi_0^1 |\, \psi_0^1 \geq 0,\ \psi_0^1 S_0 \leq V_0\} \\[2ex]
\psi_0^0 & = & V_0 - \psi_0^1 S_0 & &
\end{array}\right. \tag{12}
$$

$$
\left\{\begin{array}{lcl}
\psi_{n+1}^1 & \in & \Gamma_{n+1}(S_n, V_n^+, \psi_n^1) := \{\psi_{n+1}^1 |\, \psi_{n+1}^1 \geq 0,\ \psi_{n+1}^1 S_n \leq V_n^+ - \gamma\,|\psi_{n+1}^1 - \psi_n^1|\, S_n\} \\[2ex]
\psi_{n+1}^0 & = & V_n^+ - \gamma\,|\psi_{n+1}^1 - \psi_n^1|\, S_n - \psi_{n+1}^1 S_n \quad ; \quad n = 0, \cdots, N-1
\end{array}\right. \tag{13}
$$

showing that one can restrict oneself to just one decision variable, namely the component ψ_n^1 of ψ_n. This also implies that

$$
\left\{\begin{array}{lcl}
V_0^+ & = & [V_0 - \psi_0^1 S_0]\,(1+r) + \psi_0^1 S_0 := V_0^+(S_0, V_0, \psi_0^1) \\[2ex]
V_{n+1}^+ & = & [V_n^+ - \gamma\,|\psi_{n+1}^1 - \psi_n^1|\, S_n - \psi_{n+1}^1 S_n]\,(1+r) + \psi_{n+1}^1 S_n \cdot \omega_n \\[2ex]
 & := & V_{n+1}^+(S_n, V_n^+, \psi_n^1;\ \psi_{n+1}^1, \omega_n)
\end{array}\right. \tag{14}
$$

thereby defining implicitly the functions $V_0^+(\cdot)$, $V_{n+1}^+(\cdot)$ and where (see (6)) $\omega_n \sim p(\omega_n | Z_n)$, a lognormal distribution with mean $\left[a(Z_n) - \frac{1}{2}\sigma^2(Z_n)\right]$ and variance $\sigma^2(Z_n)\Delta$.

We shall compute the DP algorithm using discretized values of all the components of the state X_n. To this effect we have to discretize the random variables ω_n, which leads to discretized values of S_n (see (5)) and V_n^+ (see (14)). We take a finite partition $\{\Omega^i\}_{i=1,\cdots,M}$ of the positive halfline and choose representative elements $\omega^i \in \Omega^i$. The discretized random variables ω_n then take values in the set $\{\omega^i\}_{i=1,\cdots,M}$ with probabilities

$$p^i(Z_n) := P\{\omega_n = \omega^i | Z_n\} = \int_{\Omega^i} p(\omega_n | Z_n)\, d\omega_n \tag{15}$$

3 The dynamic programming algorithm (DP)

We consider two situations according to whether $\{Z_n\}$ is fully observed or not (full and restricted information respectively) and derive the DP solution for the case when H_N is given by a simple claim $H_N(S_N)$.

3.1 $\{Z_n\}$ is fully observed

The DP algorithm proceeds backwards according to the following steps :

$$\begin{aligned} &J_N(S_{N-1}, Z_{N-1}, V_{N-1}^+, \psi_{N-1}^1) \\ &\quad = \inf_{\psi_N^1 \in \Gamma_N(S_{N-1}, V_{N-1}^+, \psi_{N-1}^1)} \Big\{ \textstyle\sum_{i=1}^M p^i(Z_{N-1}) \ell \big([H_N(S_{N-1}\omega^i) \\ &\qquad\qquad - V_{N-1}^+ + \gamma\, |\psi_N^1 - \psi_{N-1}^1|\, S_{N-1} + \psi_N^1 S_{N-1}(1-\omega^i)\big]^+ \big) \Big\}, \end{aligned}$$

$$\begin{aligned} &J_n(S_{n-1}, Z_{n-1}, V_{n-1}^+, \psi_{n-1}^1) \\ &\quad = \inf_{\psi_n^1 \in \Gamma_n(S_{n-1}, V_{n-1}^+, \psi_{n-1}^1)} \Big\{ \textstyle\sum_{k=1}^K p_{Z_{n-1},k} \sum_{i=1}^M p^i(Z_{n-1}) \\ &\qquad\qquad J_{n+1}\left(S_{n-1}\omega^i, k, V_n^+(S_{n-1}, V_{n-1}^+, \psi_{n-1}^1; \psi_n^1, \omega^i), \psi_n^i\right) \Big\} \\ &\qquad\qquad\qquad\qquad \text{for} \quad n = N-1, \cdots\cdots, 1, \end{aligned} \tag{16}$$

and

$$J_0(S_0, Z_0, V_0) = \inf_{\psi_0^1 \in \Gamma_0(S_0, V_0)} J_1\left(S_0, Z_0, V_0^+(S_0, V_0, \psi_0^1), \psi_0^1\right),$$

where the sets $\Gamma_n(\cdot)$ are defined in (12) and (13), the functions $V_n^+(\cdot)$ are as defined in the right hand side of (14), $p_{h,k}$ are the elements of the transition probability matrix P of $\{Z_n\}$ and $p^i(Z_n)$ are given in (15).

It is worth mentioning that one can also determine a-priori bounds for for the possible values of V_n^+ as well as of ψ_n^1 (see (28) below in the binomial case).

3.2 $\{Z_n\}$ is not observed

If $\{Z_n\}$ is not observed, then its values have to be inferred from past and current observations of S_n, i.e. they have to be estimated on the basis of $\mathcal{F}_n^S := \sigma\{S_m,\, m \le n\}$. Since the DP algorithm depends on $\{Z_n\}$ only through the values of $p^i(Z_{n-1})$, that furthermore enter the recursions linearly, by the "smoothing property" of conditional expectations it is easily seen that, also in the present case, the DP recursions are given by the previous steps (16), except that :

1. the values of $p^i(Z_{n-1})$ have to be replaced by $E\left\{p^i(Z_{n-1}) \,|\, \mathcal{F}_{n-1}^S\right\} = \sum_{k=1}^K \pi_{n-1}^k p^i(k)$ where

   (17) $$\pi_{n-1}^k := P\left\{Z_{n-1} = k \,|\, \mathcal{F}_{n-1}^S\right\};$$

2. letting $\pi_n := (\pi_n^1, \cdots, \pi_n^K)$, the original state $(S_n, Z_n, V_n^+, \psi_n^1)$ has to be replaced by $(S_n, \pi_n, V_n^+, \psi_n^1)$. Notice also that, to specify the $K-$vector π_n, it suffices to specify $K-1$ of its components;

3. the transition probabilities $p_{Z_{n-1},k}$ for $\{Z_n\}$ are replaced by corresponding probabilities p_{π_{n-1},π_n} for $\{\pi_n\}$.

The transition probabilities p_{π_{n-1},π_n} can be obtained by means of the recursive Bayes' formula (discrete time filtering equation). To this effect let

(18) $$\pi_n^k(i) = P\{Z_n = k \,|\, \pi_{n-1}, \omega_{n-1} = \omega^i\}\,; \quad i = 1, \cdots, M$$

Proposition 3.1 (Recursive Bayes' formula) *One has, for $i = 1, \cdots, M$; $k = 1, \cdots, K$,*

(19) $$\pi_n^k(i) = \frac{\sum_{h=1}^K p_{hk}\, p^i(h)\, \pi_{n-1}^h}{\sum_{h,\ell=1}^K p_{h\ell}\, p^i(h)\, \pi_{n-1}^h}$$

where p_{hk} are the known transition probabilities for $\{Z_n\}$ and (cfr. (15)) $p^i(h) = P\{\omega_{n-1} = \omega^i |\, Z_{n-1} = h\}$.

Proof : Notice that, by $\omega_n = S_{n+1}/S_n$, one has $\mathcal{F}_n^S = \mathcal{F}_{n-1}^\omega$. In our setting, namely conditionally on $\omega_{n-1} = \omega^i$, one then obtains

$$P\left\{Z_n = k \,|\, \mathcal{F}_n^S\right\} = P\left\{Z_n = k \,|\, \mathcal{F}_{n-1}^\omega\right\}$$

(20) $$\propto \sum_{h=1}^K P\left\{Z_n = k \,|\, Z_{n-1} = h\right\} P\left\{Z_{n-1} = h \,|\, \mathcal{F}_{n-1}^\omega\right\}$$

$$\propto \sum_{h=1}^K P\left\{Z_n = k \,|\, Z_{n-1} = h\right\} P\left\{\omega_{n-1} = \omega^i |\, Z_{n-1} = h\right\} P\left\{Z_{n-1} = h \,|\, \mathcal{F}_{n-2}^\omega\right\}$$

where $\propto$ stands for "proportional to".

Notice that the only possible transitions are from π_{n-1} to $\pi_n(i) = (\pi_n^1(i), \cdots, \pi_n^K(i))$, $(i = 1, \cdots, M)$ and their probabilities are $P\{\omega_{n-1} = \omega^i | \pi_{n-1}\} = \sum_{h=1}^K \pi_{n-1}^h p^i(h)$.

The DP algorithm in the so modified form is also called "Bayesian DP" (see e.g. Martin (1967), Van Hee (1978)).

3.3 A computational remark

At each transition, both S_n and π_n have M offsprings causing the number of possible states in the DP algorithm to grow along an only partly recombining tree and this leads to computational intractability for larger horizons.

A natural approximation approach (for related approaches see e.g. Hinderer (1979), Whitt (1978)) allows however to approximate $\{S_n\}$ and $\{\pi_n\}$ by processes over a fixed finite state space, so that the number of possible values for the state in the DP algorithm remains the same throughout. This approximation approach is based on the following procedure.

Concerning S_n, consider a partition $\{\Sigma^i\}_{i=1,\cdots,L}$ of the state space of $\{S_n\}$ and pick representative elements $\sigma^i \in \Sigma^i$. Consider then, instead of $\{S_n\}$, a process with state space $\{\sigma^i\}_{i=1,\cdots,L}$ and transition probabilities as follows :

- if $\{Z_n\}$ is observed

$$p_{\sigma^i,\sigma^j}(Z_n) := P\left\{S_{n+1} \in \Sigma^j | S_n = \sigma^i, Z_n\right\} = \sum_{\{\ell:\sigma^i\omega^\ell \in \Sigma^j\}} p^\ell(Z_n) \tag{21}$$

  where the rightmost expression follows from (5) and (15), whereas

- if $\{Z_n\}$ is not observed

$$p_{\sigma^i,\sigma^j}(\pi_n) = \sum_{\{\ell:\sigma^i\omega^\ell \in \Sigma^j\}} \sum_{k=1}^K \pi_n^k p^\ell(k). \tag{22}$$

Analogously, consider a partition $\{\Pi^i\}_{i=1,\cdots,Q}$ of the state space of $\{\pi_n\}$ (the $M-$dimensional simplex) as well as representative elements $\pi^i \in \Pi^i$. Approximate then $\{\pi_n\}$ by a process with state space $\{\pi^i\}_{i=1,\cdots,Q}$ and transition probabilities

$$p_{\pi^i,\pi^j} = \sum_{\{\ell:\pi^i(\ell)\in\Pi^j\}} \sum_{k=1}^K \pi^{i,k} p^\ell(k), \tag{23}$$

where the vector $\pi^i(\ell) = (\pi^{i,1}(\ell), \cdots, \pi^{1,K}(\ell))$ is obtained from π^i by the recursive Bayes' formula, namely in the same way as the $\pi_n(i) = (\pi_n^1(i), \cdots, \pi_n^K(i))$ is obtained via (19) from the vector π_{n-1}, and where $\pi^{i,k}$ denotes the $k-th$ component of π^i.

Notice, finally, that the just described approximation approach is equivalent to approximating the optimal cost-to-go functions $J_n(S_{n-1}, Z_{n-1}, V^+_{n-1}, \psi^1_{n-1})$ in (16) by piecewise constant functions.

3.4 The case of $Z_n \equiv Z$ (unobserved)

Here we consider the special case of our model (4), when the coefficients are functions of a single random variable Z instead of being processes depending on the driving process Z_t. We may also consider this as the particular case of our model that arises when the state space of the Markov chain Z_n collapses to a single element and thus the process Z_n degenerates to a single random variable, whose distribution coincides with that of Z_0.

In this case, if Z is unobserved, it plays the role of an unknown parameter. Applying the previous procedure, some simplifications take place such as the reduction of the recursive Bayes' formula (19) to the ordinary Bayes' formula.

One may, however, also follow an alternative procedure. Instead of considering (see (15)) $p^i(Z)$ with Z unknown, one may consider as unknown directly the vector $p = (p^1, \cdots, p^M)$, where $p^i = P\{\omega = \omega^i\} \in (0,1)$ and $\sum_{i=1}^M p^i = 1$. Instead of updating the distribution of Z, one can then adopt the Bayesian point of view and use the ordinary Bayes formula to successively update the initial (prior) density $h(p)$ of p. Since $P\{\omega \,|\, p\}$ is the multinomial distribution, a conjugate family of distributions for p are the Dirichlet distributions (see e.g. Martin (1967)). With a prior

$$h_0(p) \propto (p^1)^{\alpha_0^1 - 1} \cdots (p^M)^{\alpha_0^M - 1}\,; \quad (\alpha_0^i > 0\,;\; i = 1, \cdots, M) \tag{24}$$

the posterior in period n becomes

$$h_n(p) \propto (p^1)^{\alpha_n^1 - 1} \cdots (p^M)^{\alpha_n^M - 1} \tag{25}$$

$$\text{where} \qquad \alpha_n^i = \alpha_0^i + i_n \qquad \text{with} \qquad i_n = \sum_{m=0}^{n-1} 1_{\{\omega_m = \omega^i\}} \tag{26}$$

The DP recursions are then again given by the steps (16) except that :

1. the values of $p^i(Z_{n-1})$ have to be replaced by $\frac{\alpha_{n-1}^i}{\sum_{j=1}^M \alpha_{n-1}^j}$;
2. the original state $(S_n, Z_n, V_n^+, \psi_n^1)$ is replaced by $(\alpha_n, V_n^+, \psi_n^1)$ (notice that, for given S_0 and α_0, also the value of S_n is determined by α_n);
3. the transition probabilities $p_{Z_{n-1},k}$ for $\{Z_n\}$ are replaced by $p_{\alpha_{n-1},\alpha_n}$ for $\{\alpha_n\}$.

To obtain $p_{\alpha_{n-1},\alpha_n}$ notice that if $S_n/S_{n-1} = \omega^i$ for some $i \in \{1, \cdots, M\}$, then α_{n-1} makes a transition to α_n with all components remaining the same except for the $i-$th one that is increased by 1. The corresponding probability is $P\{\omega_{n-1} = \omega^i |\, \alpha_{n-1}\} = \frac{\alpha_{n-1}^i}{\sum_{j=1}^M \alpha_{n-1}^j}$.

4 The Binomial case ($M = 2$)

For $M = 2$, the discretized ω_n become Binomial random variables (recall that $\omega_n = S_{n+1}/S_n$). A reasonable choice for the partition $\{\Omega^i\}_{i=1,2}$ is then to take $\Omega^1 = (0, 1+r]$, $\Omega^2 = (1+r, +\infty)$, while $d \in \Omega^1$, $u \in \Omega^2$ can be chosen arbitrarily. One immediately has

$$
\begin{aligned}
p^u(Z_n) &:= P\{\omega_n = u \,|\, Z_n\} = \int_{\Omega^2} p(\omega_n |\, Z_n)\, d\omega_n \\
&= P\{\log \omega_n \geq \log(1+r)\} = 1 - \Phi\left(\frac{\log(1+r) - \left[a(Z_n) - \frac{1}{2}\sigma^2(Z_n)\right]\Delta}{\sigma(Z_n)\sqrt{\Delta}}\right) \\
p^d(Z_n) &:= P\{\omega_n = d \,|\, Z_n\} = 1 - p^u(Z_n)
\end{aligned}
\tag{27}
$$

where $\Phi(\cdot)$ is the standard normal distribution function.

To perform the DP algorithm more efficiently in this case, notice that $S_n = S_0 u^{u_n} d^{n-u_n}$, where u_n denotes the number of "up" movements. For given S_0, the value of S_n can thus be characterized by the pair (n, u_n) with $u_n \leq n$. It is then rather straightforward to derive the following a-priori bounds for the values of V_n^+ and ψ_n^1 in the state vector X_n of the DP algorithm :

$$
\begin{cases}
\delta^n V_0 \leq V_n^+ \leq V_0(1+r)^{n+1}(1+u)^{u_n}(1+d)^{n-u_n} \\
\qquad\qquad \text{with} \quad \delta := (1-\gamma)\min\{1, d\} \\
\\
0 \leq \psi_n^1 \leq \frac{V_0}{S_0}(1+r)^{n+1}\left(\frac{1+u}{u}\right)^{u_n}\left(\frac{1+d}{d}\right)^{n-u_n} .
\end{cases}
\tag{28}
$$

For the case when $\{Z_n\}$ is not observed, nothing particularly changes with respect to the procedure as outlined in section 3.2. On the other hand, in the special case when Z_n degenerates to a single random variable Z (observed or not), according to the alternative approach of section 3.4, $p^u = P\{\omega = u\}$ can be considered as the only (known or unknown) parameter in the model. For the case when p^u is unknown, the (conjugate) Dirichlet distribution then particularizes to the Beta distribution :

$$
h_n(p^u) \propto (p^u)^{\alpha_n^1 - 1}(1 - p^u)^{\alpha_n^2 - 1} \;, \quad (\alpha_n^1, \alpha_n^2 > 0)
\tag{29}
$$

where, recalling that S_n can be characterized by the pair (n, u_n) with u_n denoting the total number of "up" movements,

$$
\alpha_n^1 = \alpha_0^1 + u_n \qquad ; \qquad \alpha_n^2 = \alpha_0^2 + n - u_n \,.
\tag{30}
$$

Consequently, in the DP algorithm $p^u(Z_{n-1})$ has to be replaced by $\alpha_{n-1}^1/(\alpha_{n-1}^1 + \alpha_{n-1}^2) = (\alpha_0^1 + u_{n-1})/(\alpha_0^1 + \alpha_0^2 + n - 1)$ and the state X_n simplifies considerably being given by $(n, u_n, V_n^+, \psi_n^1)$ with (integer) $u_n \leq n$. Conditionally on this state, the probability of an "up" movement is $\alpha_n^1/(\alpha_n^1 + \alpha_n^2) = (\alpha_0^1 + u_n)/(\alpha_0^1 + \alpha_0^2 + n)$.

Notice finally that when computing $J_1(\cdot)$ in (16), $p^u(Z_0)$ has to be replaced by $\alpha_0^1/(\alpha_0^1+\alpha_0^2)$. This implies that the prior distribution for p^u cannot be the noninformative prior with $\alpha_0^1=\alpha_0^2=0$.

4.1 Numerical results for the Binomial case

In our computations we have considered the binomial model in the situation when, instead of the Markov chain Z_n, we have a single random variable Z (observed and unobserved), which is equivalent to having p^u as the only parameter (known and unknown respectively). Furthermore, we have considered the following two particular cases of criterion (1), namely :

$$
\begin{aligned}
& E^P_{S_0,V_0}\left\{(H_N - V_N(\psi))^+\right\} && \textit{(shortfall risk)} \\
& P_{S_0,V_0}\left\{(H_N - V_N(\psi)) > 0\right\} && \textit{(shortfall probability)}
\end{aligned}
\tag{31}
$$

In the first case the optimal strategy turns out to be unique (in general : if $\ell(z)$ is convex, then $J_n(\cdot,\psi^1_{n-1})$ can be shown to be convex), whereas in the second case it is not uniquely determined ($\ell(z)=1_{\{Z>0\}}$ is not convex). As a consequence, it made sense to determine also the shortfall risk minimizing strategy among those that minimize the shortfall probability.

In order to give an example of the situations that may arise, in the following Figures and the Table we consider a horizon of just $N=5$. As claim we take a European call option, i.e. $H_N(S_N)=(S_N-K)^+$ and put $S_0=100, K=100$, $r=.03, \gamma=.04$, $d=0.9, u=1.3$, $p^u=p^d=0.5$. For the Bayesian case we let the prior $h_0(p)$ be a uniform distribution ($\alpha_0^1=\alpha_0^2=1$).

In Fig.1 we consider the case of a known value of $p^u=0.5$. We show two pairs of trajectories : in the first pair, the trajectory characterized by a white square plots the minimal shortfall risk against initial capital, while the one with a white disk shows the shortfall probability corresponding to the risk minimizing strategy. In the second pair, the trajectory characterized by a black disk represents the minimal shortfall probability, more precisely, the shortfall probability corresponding to the strategy that minimizes shortfall risk among those minimizing shortfall probability, while with a black square we represent the shortfall risk corresponding to this latter strategy.

Notice that, as expected, in the first pair the shortfall risk decreases with the initial capital, but the corresponding shortfall probability does not exhibit a monotonic behaviour. This can be explained by looking at Table 1, that describes the entire shortfall distribution corresponding to the risk minimizing strategy for three close values of the initial capital : for $V_0=25.4$ there are only two events leading to a shortfall with moderate shortfall values. Lowering the initial capital to $V_0=21.7$, there are many events leading to a shortfall and the shortfall values are again moderate except for the last one. Lowering further the initial capital, but only very slightly, to $V_0 = 21.6$, there are again only few events leading to shortfall, but the corresponding values start to become more significant. Since we consider a binomial model with $p^u = p^d = 0.5$, all events leading to shortfall have the same probability and this implies that with few shortfall events we have a small shortfall probability, whereas with many events this probability becomes large; since the number of shortfall events is non monotonic, so is the shortfall probability.

Fig. 2 shows the same as Fig. 1, but for the case when p^u is unknown and given a uniform prior.

We finally point out that in Table 1 we have an example of the entire shortfall distribution corresponding to the risk minimizing strategy for three possible choices of the initial capital. As mentioned in the Introduction, this distribution can be obtained for the various values of the initial capital as a rather immediate byproduct of our procedure.

References

BROWNE, S., and W.WHITT (1996) : "Portfolio choice and the Bayesian Kelly criterion," *Adv.Appl.Prob.*, 28, 1145-1176.

CVITANIĆ, J. and I. KARATZAS (1998) : "On dynamic measures of risk," Preprint, Columbia University.

CVITANIĆ, J. (1998) : "Methods of Partial Hedging," *Asia-Pacific Financial Markets*, 1-28.

CVITANIĆ, J. (1998) : "Minimizing expected loss of hedging in incomplete and constrained markets," Preprint, Department of Statistics, Columbia University.

EDIRISINGHE C., V. NAIK, and R. UPPAL (1993) : "Optimal Replication of Options with Transactions Costs and Trading Restrictions," *J. of Financial and Quantitative Analysis*, 28, 117-138.

FÖLLMER, H. and P.LEUKERT (1999) : "Quantile hedging," *Finance and Stochastics*, 3, 251-273.

FÖLLMER, H. and P.LEUKERT (1998) : "Efficient Hedging : Cost versus Shortfall Risk," Preprint Humboldt-Universität zu Berlin.

HINDERER, K. (1979) : "On approximate solutions of finite-stage dynamic programs," in *Dynamic Programming and Its Applications* ed. M.Puterman, Academic Press, NY, 289-317.

KARATZAS, I. and X. ZHAO (1999) : "Bayesian adaptive portfolio optimization," Preprint, Columbia University.

LAKNER, P. (1995) : "Utility maximization with partial information," *Stochastic Processes and their Applications*, 56, 247-273.

LAKNER, P. (1998) : "Optimal trading strategy for an investor : the case of partial information," *Stochastic Processes and their Applications*, 76, 77-97.

MARTIN, J.J. (1967) : *Bayesian Decision Problems and Markov Chains*, New York : John Wiley and Sons, Inc.

PHAM, H. (1998) : "Dynamic L^p- Hedging in Discrete Time under Cone Constraints," Preprint Université de Marne-la-Vallée and CREST.

SPIVAK, G. and J. CVITANIĆ (1998) : "Maximizing the probability of a perfect hedge," Preprint, Department of Statistics, Columbia University.

VAN HEE, K.M. (1978) : *Bayesian control of Markov chains*, Mathematical Centre Tracts 95, Amsterdam.

WHITT, W. (1978) : "Approximations of dynamic programs I," *Math. Oper. Res.*, 3, 231-243.