Stockholms universitet

Cost-of-Capital Estimation using Least-Squares Monte Carlo

Adam Pettersson

Masteruppsats i försäkringsmatematik
Master Thesis in Actuarial Mathematics

Masteruppsats 2022:8
Försäkringsmatematik
Juni 2022

www.math.su.se

Matematisk statistik
Matematiska institutionen
Stockholms universitet
106 91 Stockholm

Mathematical Statistics
Stockholm University
Master Thesis **2022:8**
http://www.math.su.se

Cost-of-Capital Estimation using Least-Squares Monte Carlo

Adam Pettersson*

June 2022

Abstract

This thesis evaluates how well LSM (Least-Squares Monte Carlo) succeeds in estimating the cost-of-capital margin derived in Engsner, Lindholm, and Lindskog [1]. Using Gaussian processes to model the cash flows arising from an insurance portfolio, we can explicitly calculate the value of the margin and thus evaluate how well LSM works. While Gaussian assumptions are not entirely realistic in the insurance context we consider, the idea is that LSM might work well also for non-Gaussian assumptions if it works well under Gaussian assumptions.

We consider first an insurance portfolio where the individual risks are independent and identically distributed. It turns out that LSM can estimate the margin very well in this one-cohort setting. The results are less satisfactory in the two-cohort case examined, and more work would be needed to improve the LSM algorithm used. In particular, modeling correlation between different cohorts in a heterogeneous insurance portfolio proved to be complicated

*Postal address: Mathematical Statistics, Stockholm University, SE-106 91, Sweden. E-mail: adam.gideon.pettersson@gmail.com. Supervisor: Filip Lindskog.

Contents

Acknowledgements
I am grateful to Filip Lindskog for his guidance and sound advice. I also owe thanks to my colleagues at EY, who have been flexible and helped me find time for the thesis. Lastly, I want to thank my family and my girlfriend for their support.

1 Introduction

This introductory chapter will briefly look at the Solvency 2 Directive. We then target a specific object, the risk margin defined in the directive. The design of the risk margin has received some criticism. We look at why and what alternatives to the risk margin are available.

1.1 Solvency 2

The European Commission transposed the Solvency 2 directive into national law in January 2016. It includes three pillars and aims to unite the European insurance market and improve customer protection.

Pillar 1 contains quantitative requirements for technical provisions and solvency capital. The purpose of technical provisions is to ensure insurers have enough assets to cover their insurance undertakings. Technical provisions are definied as a sum of the discounted best estimate of the insurance liabilities and a risk margin. Even if the best estimates are unbiased estimates of the liability cash flow, the actual costs will deviate from the estimates due to the stochastic nature of future insurance liability cash flows. Therefore, the insurer needs extra capital to be prepared for years where actual costs exceed expected costs. Ensuring the availability of this extra capital is not free, and insurers are obliged to set aside capital for a risk marginal, defined conceptually as the amount that would have to be paid to another insurance company in order for them to take on the best estimate liability, and mathematically as:

$$\mathrm{RM(D)} := \mathrm{CoC}\sum_{t=1}^{T} \frac{SCR_t(D)}{(1+r_{0,i})^t}. \tag{1}$$

CoC is a constant representing the cost-of-capital rate, set to 0.06, and $r_{i,j}$ is the yearly compounded forward rate for period i to j. The values $SCR_t(D)$ are stochastic and represent the Solvency Capital Requirement as calculated at time t. The Solvency Capital Requirement needs to be calculated yearly and should cover basically all other risk than those directly connected to insurance liability costs. It is not necessary to define the SCR's further here.The key point is that they are impossible to calculate at time 0. Hence, it is not clear from equation 1 how to calculate the risk margin.

Insurers can calculate the risk margin either using an internal formula approved by the national legislator, or a standard formula provided by the European Insurance and Occupational Pensions Authority, EIOPA (2015, Paragraph 1.114, Method 2). The standard formula is based on (1) but replaces the stochastic SCR with expected values,

$$\mathrm{RM(D)} := \mathrm{CoC}\frac{SCR_0(D)}{BE_{\mu,1}(D)} \sum_{i=1}^{T} \frac{BE_{\mu,i}(D)}{(1+r_{0,i})^i}.$$

Here $BE_{\mu,i}(D)$ is the best estimate of remaining liabilities at time i, which in a life insurance context could take the form

$$\mathrm{BE}_{\mu,i}(D) := \sum_{j=1}^{T} \frac{\mathrm{E}[D_j;\mu]}{(1+r_{i-1,j})^{j-i+1}}, \quad i = 1, ..., T,$$

where $\mathrm{E}[D_j;\mu]$ denotes expected value of D_j using mortality rate μ.

The definition of the risk margin has rendered criticism. For example, the Society of Actuaries has in an article, Pelkiewicz et al. [3], expressed concerns that the risk margin is too volatile and too conservative,. In response to the insurance market's criticism, the European Commission ordered a draft of reworked Solvency 2 guidelines from EIOPA. It is still unknown what shape these will take and to what extent they will be able to address problems with the risk margin. In the draft, [4], it is suggested that the risk margin is defined as

$$RM = \mathrm{CoC} \cdot \sum_{t \geq 0} \frac{\max(\lambda^t;\ \text{floor}\ ) \cdot SCR_t}{(1+r_{t+1})^{t+1}}$$

$$\text{where}\ \ \mathrm{CoC} = 6\%, \lambda = 0.975,\ \text{floor}\ = 50\%$$

The idea is that the exponential lambda approach, if adopted, will lead to a risk margin better mimicking the actual cost-of-capital. We will not look further into this topic, only note that discussions regarding the risk margin are still ongoing.

1.2 Objective

Alternatives to the solvency 2:s standard formula for the risk margin have been proposed, as in Möhr [5] and Engsner, Lindholm, and Lindskog [1]. These articles use dynamic risk measures and dynamic utility functions to define a cost-of-capital margin. In [1], an expression for a cost-of-capital margin is described. However, a difficulty is that the cost-of-capital margin is hard or impossible to calculate analytically without making strong assumptions about the underlying cash flows, such as Gaussian cash flows. Another possibility is to use simulation, as in Engsner [6].

In this thesis, we examine a method known as LSM (Least Squares Monte Carlo) for calculating the cost of capital margin from [1]. We restrict ourselves to Gaussian cash flows. That way, we can compute the cost-of-capital margin exactly and thus evaluate the LSM algorithm's performance. If the LSM algorithm's margin succeeds in estimating the exact margin value, this may indicate that the LSM is suitable also for situations where Gaussian cash flows can not be assumed.

2 Theoretical background

This section begins with a derivation of the cost-of-capital margin using economic arguments. We present the definitions, assumptions, and results needed to compute the cost-of-capital margin explicitly. The presentation and terminology follows that of [1].

Next comes an introduction to the least-squares Monte Carlo algorithms. We look first at its original area of use and how to interpret it for our purposes. The section ends with a description of the life insurance assumptions and models used throughout this thesis.

2.1 The cost-of-capital margin

In this section, the goal is to provide a more accurate interpretation of the market-consistent value of a liabilitys's cash flow. We introduce the margin for the cost of capital as it is definied in [1]. From that article we also lend terminology and definitions.

We consider time periods (usually years or quarters) and associated times 0,1,..., T, where times 0 and T correspond to the start and end point of the insurance liability, respectively. For this we have a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F})$. $\mathbb{F} = (\mathcal{F})_{t=0}^{T}$ with $\{\emptyset, \Omega\} = \mathcal{F}_0 \subseteq ... \subseteq \mathcal{F}_T = \mathcal{F}$. Based on the filtered probability space, an insurance undertaking's cost cash flow can be described by means of a $\mathbb{F}$-adapted stochastic process $X^0 = (X_t^0)_{t=1}^{T}$.

If a cash flow comprises capital costs from capital requirements on both assets and liabilities, the value of cash flow is affected by the value of the assets. This dependence on assets is not optimal from a regulatory perspective, as the cash flows of two identical commitments are then valued differently depending on the company's assets. For a more uniform valuation, European Commision (2015, Article 38) advocates valuing liabilities through a reference undertaking. The reference undertaking means a hypothetical transfer to a separate entity whose access only serves to match the liability value. Based on such a reference undertaking, we will gradually define the value of the liability's cost of capital margin.

The valuation takes place in two steps. At time 0, a portfolio is purchased with the sole purpose of replicating the liability cash flow. The replicating portfolio has a market price of π and generates a cash flow $X^s = (X_t^s)_{t=1}^{T}$. The value of the original liability can now be expressed as the sum of π and the value of the residual cash flow $X := X^0 - X^s$, calculated through repeated one-period replication. The replication uses capital from a capital provider with limited liability and requires compensation for capital costs. The value $V_0(X)$ of the residual cash flow is what we define as the cost of capital margin.

In defining $V_0(X)$ mathematically, we need first to clarify what is meant by capital requirements. We define capital requirements as the insurance company needing to have a certain amount of capital available, depending on the cost from the liability cash flow. At time t, the value of the resiudual liability including costs during the current year is modeled as

$$-X_{t+1} - V_{t+1}(X).$$

Note that this is a random variable, as we are looking at costs from year $t+1$ at time t. Let R_t be a risk measure that indicates what capital must be available at a time t linked to next year's costs. At time t, the insurance company then needs to hold capital of $R_t(-X_{t+1} - V_{t+1}(X))$. At the same time, the capital $V_t(X)$ is available, the current value of the residual cash flow. If $V_t(X)$ is not sufficient to meet the capital requirement, the missing capital will be

$$C_t := R_t(-X_{t+1} - V_{t+1}(X)) - V_t(X). \tag{2}$$

We assume that a capital provider agrees to provide C_t at time t. The balance at time t can now be written as

$$R_t(-X_{t+1} - V_{t+1}(X)) - X_{t+1}.$$

Any surplus goes to the capital provider as compensation for assisting with capital. Moreover, the capital provider has no obligation to provide more than

$$R_t(-X_{t+1} - V_{t+1}(X)) - X_{t+1} - V_{t+1}(X).$$

For a capital provider to accept providing capital, the deal must be sufficently good. Using the conditional expectation $\mathrm{E}_t(X) = \mathrm{E}[X|\mathcal{F}_t]$ to measure utility, we express the capital provider's acceptance conditions at time t as

$$\mathrm{E}_t\left((R_t(-X_{t+1} - V_{t+1}(X)) - X_{T+1} - V_{t+1}(X))_+\right) \geq (1+\eta_t)C_t. \tag{3}$$

Here η_t is a value that quantifies how much utility the capital provider requires to accept providing capital. A capital provider would accept providing capital only if the expected return is larger than η_t. All numerical examples and implementations in this thesis use the conditional expected value as utility function.

Setting $Y_{t+1} := X_{t+1} + V_{t+1}(X)$, and combining equations (2) and (3), we get $Y_{t+1} := X_{t+1} + V_{t+1}(X)$,

$$V_t(X) \geq R_t(-Y_{t+1}) - \frac{1}{1+\eta_t}\mathrm{E}_t\left((R_t(-Y_{t+1}) - Y_{t+1})_+\right).$$

We replace the inequality with an equality. Otherwise, the interpretation would be that the capital provider gets a larger compensation than it would need to accept the deal. Hence we get

$$V_t(X) = R_t(-Y_{t+1}) - \frac{1}{1+\eta_t}\mathrm{E}_t\left((R_t(-Y_{t+1}) - Y_{t+1})_+\right). \tag{4}$$

2.1.1 Mathematical details

Here the content of the previous section is supplemented with mathematical definitions and results cited from [1]. We first look into the risk measures used to express capital requirements.

Definition 1. For $p \in [0, \infty]$, a dynamic monetary risk measure $(R_t)_{t=0}^{T-1}$ is a sequence of mappings $R_t : L^p(\mathcal{F}_{t+1}) \to L^p(\mathcal{F}_t)$

$$\text{if } \lambda \in L^p(\mathcal{F}_t) \text{ and } Y \in L^p(\mathcal{F}_{t+1}), \text{ then } R_t(Y + \lambda) = R_t(Y) - \lambda, \tag{5}$$

$$\text{if Y, } \tilde{Y} \in L^p(\mathcal{F}_{t+1}) \text{ and } Y \leq \tilde{Y}, \text{ then } R_t(Y) \geq R_t(\tilde{Y}), \tag{6}$$

$$R_t(0) = 0 \tag{7}$$

With the above definition at hand, we can express the cost of capital margin.

Theorem 1 (Proposition 1 in [1])**.** *Fix $p \in [0, \infty]$. Let $(R_t)_0^{T-1}$ be given by Definition 1, let $\mathrm{E}_t : L^p(\mathcal{F}_{t+1}) \to L^p(\mathcal{F}_t)$ be the conditional expectation, and let $\eta_t \in L^0_+(\mathcal{F}_t)$.*

(i)

$$W_t(Y) := R_t(-Y) - \frac{1}{1 + \eta_t} \mathrm{E}_t \left((R_t(Y) - Y)_+ \right) \tag{8}$$

is a mapping from $L^p(\mathcal{F})_{t+1})$ to $L^p(\mathcal{F})_t)$ having the properties

$$if \lambda \in L^p(\mathcal{F}_t) \textit{ and } Y \in L^p(\mathcal{F}_{t+1}) \textit{ then } W_t(Y + \lambda) = W_t(Y) + \lambda, \tag{9}$$

$$\textit{if } Y, \tilde{Y} \in L^p(\mathcal{F}) \textit{ and } Y \leq \tilde{Y}, \textit{ then } W_t(Y) \leq W_t(\tilde{Y}) \tag{10}$$

$$W_t(0) = 0. \tag{11}$$

(ii)

Let $(X_t)_{t=1}^T$ be a $\mathbb{F}$-adapted cash flow with $X_t \in L^p(\mathcal{F}_t)$ for every t. The cost-of-capital margin $V_t(X)$ in (3) satisfies

$$V_t(X) = W_t \circ \cdots \circ W_{T-1}(X_{t+1} + \cdots X_T), \tag{12}$$

where $W_t \circ \cdots \circ W_{T-1}$ denotes the composition of mappings $W_t, ..., W_{T-1}$

The Value-at-risk is the risk measure used when defining the solvency capital requirements in the Solvency 2 guidelines and will be our risk measure of choice. Conceptually, the value at risk is the most sizeable loss that can occur if we exclude the 0.5% most sizeable losses, or more formally as

Definition 2. Let X be a real random variable. The value-at-risk (VaR) at the quantile level α is defined as

$$\mathrm{VaR}_\alpha(X) = \inf\{L \in \mathbb{R} : P(L + X < 0) \leq \alpha\}.$$

In a dynamic risk measure context, as defined in (2.1), the Value-at-Risk is slightly different to inclose the cash flow's time-dependent nature. Based on example 1 in [1], we make the definition as follows

Definition 3. Let $X \in L^0(\mathcal{F}_{t+1})$. The value at risk at time t and level α is

$$\mathrm{VaR}_{t,\alpha}(X) = \text{ess inf}\{L \in L^0(\mathcal{F}_t) : P(L + X < 0|\mathcal{F}_t) \leq \alpha\}.$$

2.1.2 Cash Flow Models

For an explicit calculation of the risk margin, we need stronger assumptions about the underlying cash flow. We first assume that cash flows are auto-regressive time series, leading to theorem 2. Auto-regressive time series is a sufficiently general class of models to describe many types of cash flows. However, this generality is at the expense of the feasibility of calculating the cost of capital margin. Assuming instead that the cash flow follows a Gaussian process allows for explicit calculation as in theorem 3.

Theorem 2 (Proposition 5 in [1])**.** *Fix $p \in [1, \infty]$ and let W_t be given by (8), with R_t and E_t satisfying the condition in Proposition 4, described in appendix A2. Let $(Z_t)_{t=1}^T$ be an $\mathbb{F}$-adapted sequence of random variables such that, for each t, $Z_{t+1} \in L^p(\mathcal{F}_{t+1})$ is independent of $\mathcal{F}_t$. Let $(\alpha_t)_{t=1}^T$ be a nonrandom sequence of real numbers. Let $X_0 := 0, \quad X_{t+1} := \alpha_{t+1} X_t + Z_{t+1}, \quad t = 0, \ldots, T-1$, and set*

$$\begin{aligned} &\beta_T := 1, \quad \beta_t := 1 + \beta_{t+1}\alpha_{t+1}, \quad t \in \{1, \ldots, T-1\}, \\ &\delta_T := 0, \quad \delta_t := \delta_{t+1} + |\beta_{t+1}| W_t (\operatorname{sign}(\beta_{t+1}) Z_{t+1}) \\ &\quad t \in \{0, \ldots, T-1\} \end{aligned}$$

Then, for $t = 0, \ldots, T-1, \delta_t \in L^0(\mathcal{F}_0)$ and

$$V_t(X) = \delta_t + \beta_{t+1}\alpha_{t+1} X_t \in L^p(\mathcal{F}_t).$$

In particular,

$$V_0(X) = \sum_{t=0}^{T-1} |\beta_{t+1}| W_t (\operatorname{sign}(\beta_{t+1}) Z_{t+1})$$

The result above allows us to calculate the cost of capital margin with a backward recursion. However, it generally becomes difficult to calculate the cost-of-capital margin using theorem 2 unless the underlying distribution has quantiles easy to calculate. We therefore move on to theorem 3, which uses Gaussian assumption and leads to an explicit formula for the cost-of-capital margin.

Theorem 3 (Proposition 6 in [1]). *Let $(X, \mathbb{G})$ be a zero mean Gaussian model, ϵ_1 standard normal and suppose that $\eta_t = \eta_0 \ \forall \ t \ \in \{0, ..., T\}$. Let W_t be given by (8), with R_t and E_t satisfying the condition in Proposition 4. Then, for $t \in \{0, ..., T-1\}$,*

$$\begin{aligned} V_{t,\mathbb{G}}(X) = &\mathrm{E}\left[\sum_{s=t+1}^{T} X_s \mid \mathcal{G}_t\right] \\ &+ \sum_{s=t+1}^{T} \mathrm{Var}\left(\mathrm{E}\left[\sum_{u=s}^{T} X_u \mid \mathcal{G}_s\right] \mid \mathcal{G}_{s-1}\right)^{1/2} W_0\left(\epsilon_1\right). \end{aligned}$$

Moreover,

$$\begin{aligned} V_{0,\mathbb{G}}(X) = &\sum_{s=1}^{T}\Bigg(\mathrm{Var}\left(\sum_{u=s}^{T} X_u \mid \mathcal{G}_{s-1}\right) \\ &- \mathrm{Var}\left(\sum_{u=s}^{T} X_u \mid \mathcal{G}_s\right)\Bigg)^{1/2} W_0\left(\epsilon_1\right). \end{aligned}$$

We may decompose and rewrite the latest expression as,

$$\sum_{s=1}^{T} \mathrm{Var}\left(X_s + \mathrm{E}\left[\sum_{u=s+1}^{T} X_u \mid \mathcal{G}_s\right] - \mathrm{E}\left[\sum_{u=s}^{T} X_u \mid \mathcal{G}_{s-1}\right] \mid \mathcal{G}_{s-1}\right)^{1/2} W_0\left(\epsilon_1\right),$$

meaning that the cost of capital margin is proportional to the sum of conditional standard deviations of the prediction error of the remaining residual cash flow. Proofs and further details on topics in this subsection can be found in [1].

2.1.3 Example calculation

To provide some intutition and clarity regarding the cost-of-capital margin, we here consider an example small enough to allow for calculations by hand while still being non-trivial. R_t will be as in definition 4 and E_t the conditional expectation. Let $T = 3$ and the cash flow $(X_t)_{t=0}^{T}$ follow a Gaussian process with mean zero and covariance matrix given by

$$\Sigma = \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & \rho \\ \rho^2 & \rho & 1 \end{pmatrix}, \ \rho = 0.5.$$

Our aim is to calculate the cost-of-capital margin using the formula from theorem 3:

$$\sum_{s=1}^{T}\left(\mathrm{Var}\left(\sum_{u=s}^{T} X_u \mid g_{s-1}\right) - \mathrm{Var}\left(\sum_{u=s}^{T} X_u \mid g_s\right)\right)^{1/2} W_0\left(\epsilon_1\right).$$

We calculate (numerically) beforehand

$$\begin{aligned}W_0(\epsilon_0) &= R(\epsilon_0) - \frac{1}{1+\eta}E\Big[\big(R(\epsilon_0)-\epsilon_0\big)_+\Big]\\ &= 2.575829 - \frac{1}{1.06}2.577202 = 0.1442583.\end{aligned}$$

If we define $Y = (X_t)_{t=s}^{T}$ and $Z = (X_t)_{t=0}^{s-1}$, the conditional covariance matrix for $(X_t)_{t=s}^{T}|(X_t)_{t=0}^{s-1} \Leftrightarrow Y|Z$ is given by

$$\Sigma_{Y|Z} = \Sigma_{YY} - \Sigma_{YZ}\Sigma_{ZZ}^{-1}\Sigma_{ZY}, \text{ where } \Sigma = \begin{pmatrix}\Sigma_{ZZ} & \Sigma_{ZY}\\ \Sigma_{YZ} & \Sigma_{YY}\end{pmatrix}.$$

This allows us to, as an intermediate step, calculate conditional covariance matrices. Below $\underset{\mathcal{G}_i}{\Sigma}$ denotes the conditional covariance matrix given $\mathcal{G}_i$.

$$\underset{\mathcal{G}_0}{\Sigma} = \Sigma, \qquad \underset{\mathcal{G}_3}{\Sigma} = 0$$

$$\underset{\mathcal{G}_1}{\Sigma} = \begin{pmatrix}1 & 0.5\\ 0.5 & 1\end{pmatrix} - \begin{pmatrix}0.5\\ 0.25\end{pmatrix}1^{-1}\begin{pmatrix}0.5 & 0.25\end{pmatrix} = \begin{pmatrix}0.75 & 0.375\\ 0.375 & 0.9375\end{pmatrix},.$$

$$\underset{\mathcal{G}_2}{\Sigma} = 1 - \begin{pmatrix}0.5 & 0.25\end{pmatrix}\begin{pmatrix}1 & 0.5\\ 0.5 & 1\end{pmatrix}^{-1}\begin{pmatrix}0.5\\ 0.25\end{pmatrix} = 0.75.$$

We now get

$$\begin{aligned}&\left(\mathrm{Var}\left(\sum_{u=1}^{T} X_u \mid \mathcal{G}_0\right) - \mathrm{Var}\left(\sum_{u=1}^{T} X_u \mid \mathcal{G}_1\right)\right)^{1/2} W_0(\epsilon_1)\\ +&\left(\mathrm{Var}\left(\sum_{u=2}^{T} X_u \mid \mathcal{G}_1\right) - \mathrm{Var}\left(\sum_{u=2}^{T} X_u \mid \mathcal{G}_2\right)\right)^{1/2} W_0(\epsilon_1)\\ +&\left(\mathrm{Var}\left(\sum_{u=3}^{T} X_u \mid \mathcal{G}_2\right) - \mathrm{Var}\left(\sum_{u=3}^{T} X_u \mid \mathcal{G}_3\right)\right)^{1/2} W_0(\epsilon_1)\end{aligned}$$

$$\Leftrightarrow$$

$$\begin{aligned}&\Big((3\cdot 1 + 4\cdot 0.5 + 2\cdot 0.25) - (0.75 + 0.9375 + 2\cdot 0.375)\Big)^{1/2} W_0(\epsilon_1)\\ &\qquad + \Big(0.75 + 0.9375 + 2\cdot 0.375) - (0.75)\Big)^{1/2} W_0(\epsilon_1)\\ &\qquad + \Big(0.75) - (0)\Big)^{1/2} W_0(\epsilon_1)\\ &= (3.0625^{1/2} + 1.6875^{1/2} + 0.75^{1/2})\cdot 0.14425 = 0.565.\end{aligned}$$

Deploying the backwards recursion from theorem 4, with implementation described in appendix A1, returns the exact same value.

2.2 Monte Carlo methods

Monte Carlo simulation is a technique used to calculate a value by random sampling. The goal is to calculate a deterministic value that is too difficult to calculate with analytical or numerical methods. In mathematical statistics, a common area of application of Monte Carlo simulation is calculating some expected value of a function of a random variable. For a random vector X with density function $f(x)$, and a function $g(\cdot)$, we might want to determine the value of

$$\mathrm{E}[g(X)] = \int g(x)f(x)dx. \tag{13}$$

Using Monte Carlo simulation, this is estimated as

$$\frac{\sum_{i=1}^{n} g(x_i)}{n}, \ x_i\text{:s sampled from } X. \tag{14}$$

The strong law of large numbers would now guarantee that the expression (14) converges to the value in (13) for large enough values on the sample n.

The cost-of-capital margin does typically not allow for explicit computation of the margin unless we restrict our attention to special cases, such as Gaussian cash flows. However, a straightforward Monte Carlo approximation would not be a realistic option for calculating the cost-of-capital margin. To see why, we recognize from theorem 1 that

$$V_t(X) = W_t \circ \cdots \circ W_{T-1}(X_{t+1} + \cdots X_T) = W_t\big(X_{t+1} + V_{t+1}(X)\big).$$

If we were to draw samples $X_{t+1}^{(i)}$ of X_{t+1}, the value V_{t+1} conditional on a draw $X_{t+1}^{(i)}$ is deterministic but incalculable in most cases. To find this unknown value, we would need to estimate V_{t+1} using new simulations for each draw $X_{t+1}^{(i)}$. I.e., we need to use simulations within a simulation, so-called nested simulations. The nested simulation approach is illustrated in figure 1. Since the liabilities we consider in this thesis might span over decades, a nested simulation could require too many simulations to generate stable estimates of the cost-of-capital margin.

Instead of a full-scale nested Monte Carlo approach, we consider a method called least-squares Monte Carlo. The following section describes the method's original area of use. After that, we will move on to a more general description, and lastly, we will consider an insurance setting.

2.2.1 The Longstaff-Schwartz Algorithm

In Longstaff and Schwartz [8], the Longstaff-Schwartz algorithm for pricing options was introduced. The key idea is to replace the inner simulations in a nestled simulation with a regression model. This elimination of nested simulations can

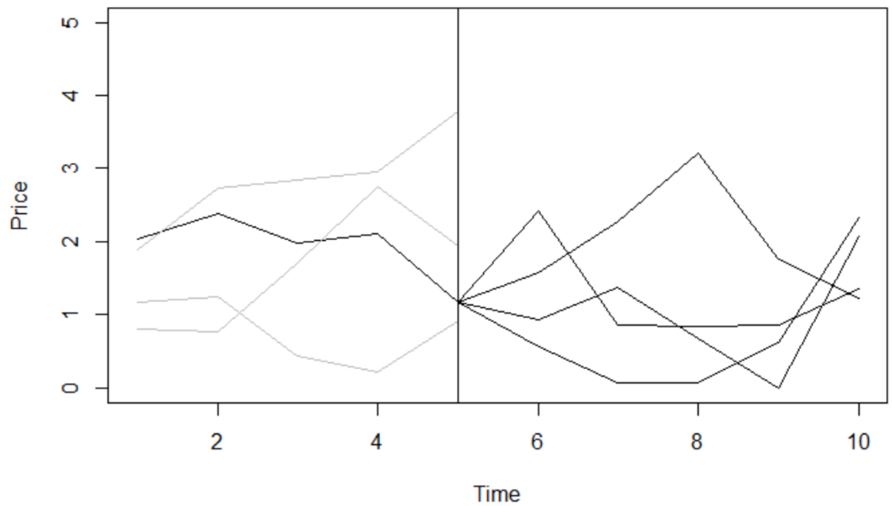


Figure 1: A figure of a nested simulation, illustrating how it for each path is necessary to simulate new paths at each time step. In particular, we see for the black path how new paths are simulated to calculate the $t = 6$ value from the $t = 5$ value.

significantly reduce the size of the computational task while maintaining a low error in the right circumstances.

For the algorithm, a complete probability space $(\Omega, \mathcal{F}, P)$ and a finite number of times $t \in 0, ..., T$ are assumed, where T is the option's expiration date. The option has a payout-function $g(S_t)$, where S_t is the price vector of the option's underlying random assets at time t. The value V_t of an American option equals the earnings from at all times $t \in 0, ...T$ deploying an optimal strategy. This means that for any t, the alternative that provides the greatest expected discounted earning is selected. At time T, the expiration date, the optimal strategy is clear. The holder triggers the option if it is in the money, otherwise not. To find the most fruitful alternative at any other time $t_i \in 0, ..., T$, we need to compare the earnings from immediate triggering of the option, with that of holding on to the option. I.e, we consider

$$max\left(g(S_{t_i}), \mathrm{E}[V_{t_{i+1}}|S_{t_i}]\right). \tag{15}$$

The conditional expecation in (15) is represented using basis functions $\Phi(\cdot)$ and constants β,

$$\mathrm{E}[V_{t_{i+1}}|S_{t_i}] = \sum_{j=0}^{m} \beta_j \Phi_j(S_{t_i}). \tag{16}$$

The parameters β from (16) are estimated using an ordinary least-squares method.

Algorithm 1 LSM for a simple put option

1 Simulate m paths $S_t^{(i)}$, $i \in 1, ..., m$ and $t \in 1, ...T$.
2 Set $\hat{V}_T^{(i)} = (K - S_T^{(i)})_+$ for all $i \in 1, ..., m$.
3 Fit coefficients β_t with $\hat{V}_T^{(i)}$ as response variables and $S_{T-1}^{(i)}$ as explanatory variables, using some form of regression.
4 Estimate the conditional expectation $\mathrm{E}[V_T^{(i)}|S_{T-1}^{(i)}]$ from the regression model in step 3 and denote this value by $\hat{C}_T^{(i)}$.
5 Set $\hat{V}_{T-1}^{(i)} = max\left(\hat{C}_T^{(i)}, (K - S_{T-1}^{(i)})_+\right)$.
6 Proceed with steps (3 to 5) until $\hat{V}_0^{(i)}$ is found for all i:s.
7 The LSM estimate is given by $\hat{V}_0 = \mathrm{mean}(\hat{V}_0^{(i)})$.

To examplify the algorithm we consider an American put option with value depending on a single risky asset S_t. For this example, we assume that the risky asset $S_t = S(t)$ behaves according to assumptions from a Black - Scholes market.

At times $t \in 0, 1, 2, ..., T$, the owner has the opportunity to sell the underlying stock for a settled amount K. Say $K = 100$. If the current stock price is 85 dollars, triggering the option would mean a 15-dollar surplus for the put options owner. On the other hand, the surplus might increase further if the owner waits longer to trigger the option. Hence, the optimal strategy for the owner is to compare the surplus gained from triggering at time t with the surplus expected from holding on to the option. Mathematically, we say that at time t,

$$V_t = max\left((K - S_t)_+, \mathrm{E}[V_{t+1}|\mathcal{F}_t]\right), \tag{17}$$

where V_t is the option's value at time t. Algorithm 1 shows how the option's value at time 0 is calculated using the Longstaff-Schwartz algorithm.

2.2.2 Least Square Monte Carlo

In the previous section, an American option was valued using a least-squares Monte Carlo approach. In that context, the problem consists of an optimal stopping part and a part where we estimate a conditional expectation. Our interest lies in the latter part, where least-squares is used to find the conditional expectation without additional Monte Carlo simulations. In the previous section's example, S_t was the price of a risky asset. In our upcoming explorations, S_t will be the number of survivors at time t Formally, for t in $0, ..., T-1$ and $V_T = 0$, we want to solve a recursion on the form

$$V_t = W_t\big(g(S_{t+1}) + V_{t+1}\big). \tag{18}$$

with W_t defined as in theorem 1, $g(S_{t+1}) = S_{t+1} - S_t$ is the number of deaths and R_t the Value-at-risk. To find an approximate solution for (18) we require

M simulated paths of S_t, i.e. $S_t^{(i)}$ for i in 1,...,M and t in 1,...,T. We also need a set of basis functions

$$\Phi(x) = \{1, \Phi_1(x), ..., \Phi_N(x)\},$$

required to be linearly independent. One simple example is the set of basis functions $\Phi(x) = \{1, x, x^2, ..., x^N\}$. In relation to the basis functions, we have for each t a vector $\beta_t = (\beta_0, ..., \beta_N)$. Next we define

$$\mathbf{\Phi}_t = \begin{pmatrix} 1 & \Phi_1(S_t^{(1)}) & ... & \Phi_N(S_t^{(1)}) \\ \vdots & \vdots & \vdots & \\ 1 & \Phi_1(S_t^{(M)}) & ... & \Phi_N(S_t^{(M)}) \end{pmatrix}, \ Y_t = \begin{pmatrix} Y_t^{(1)} \\ \vdots \\ Y_t^{(M)} \end{pmatrix},$$

and that for such $\mathbf{\Phi}_t$ and Y_t,

$$\widehat{\beta}_t := \left(\mathbf{\Phi}_t^{\mathrm{T}} \mathbf{\Phi}_t\right)^{-1} \mathbf{\Phi}_t^{\mathrm{T}} Y_t.$$

The LSM version of (18) is now, when $Y_t^{(i)} = W_t\big(g(S_t^{(i)}) + V_{t+1}^{(i)}\big)$,

$$\hat{V}_t = \widehat{\beta}_t \ \Phi_t(S_t). \tag{19}$$

2.2.3 Life Insurance setting

In our upcoming calculations, we will consider life insurance portfolios consisting at time $t = 0$ of individuals in some age cohorts. In general, each age cohort will be individuals of a certain age. So one cohort for people who are 50, one cohort for 51, for all cohorts present in the portfolio. The insurance contracts considered in this thesis all concern mortality. No other elements of risk are present or examined. Therefore, some care will be given to model the mortalities properly. We will first have some assumptions outlined, followed by our method of choice and some examples illustrating the behaviors of mortality.

We assume first that fatalities occur independent of each other. The probability that an x years old individual starting at time t dies in the next year is denoted by

$$q_x(t) = 1 - \frac{S_x(t+1)}{S_x(t)}. \tag{20}$$

Note that the individuals of an age cohort will not of the exact same age. Some will be a few months older than the others. Still, they are all assumed to have the same probability of dying in the next 1-year period. This is in generally assumed to even out.

$S_x(t)$ in (20) is here the survival function for an x years old individual. The survival function can be calculated as

$$S_x(t) = exp\left\{-\int_x^{x+t} \mu_x(s)ds\right\},$$

where $\mu_x(s)$ is the hazard rate at time s for an x years old individual. Details about fitting a hazard rate model are found in appendix 2, together with some graphs.

3 Computational approach

This section looks at implementing the explicit formula for the cost-of-capital margin given Gaussian cash flows. After that, we look at implementing an LSM algorithm for a homogenous underlying insurance portfolio. Finally, we look at a more advanced approach with a heterogeneous portfolio consisting of two cohorts.

Throughout this section, we ignore time effects on the value of money; 1 SEK today is assumed to be worth as much as 1 SEK is in 30 years. Of course, such an assumption is not realistic but allows us to focus on how the LSM algorithm performs. Discounting cash flows with some deterministic interest rate would not be difficult, but would require focusing on something that does not affect the LSM algorithm.

3.1 Gaussian Cost-of-Capital margin

In the previous sections, we estimated the mean and covariance-variances of a stocastic cash flow describing a life insurance portfolio,

$$\mu = (\mu_0, ..., \mu_T)^T,\ \Sigma = \begin{pmatrix} \sigma_{11} & ... & \sigma_{1T} \\ \vdots & \vdots & \vdots \\ \sigma_{T1} & ... & \sigma_{TT} \end{pmatrix}.$$

We assumed that these estimates are the actual characteristics of the cash flow. However, we exchanged the underlying distribution with a normal distribution to allow for explicit computation of the cost-of-capital with theorem 3. Theorem 4 provides an alternative that is less cumbersome to implement. For all numerical results in this thesis, theorem 4 was used. The implementation is described in algorithm 4 in appendix A1.

Theorem 4 (Proposition 10 in [1])**.** *Let $X = (X_t)_{t=1}^T \sim N_T(0, \Sigma)$ where Σ is invertible, with $\mathbb{G}$ as its natural filtration. Let W_t be given by (8), with R_t and E_t satisfying the condition in Proposition 4. Then*

$$V_t(X) = \begin{cases} 0, & t = T \\ \left(v^{(t)}\right)^T X_{1:t} + k_t, & t \in \{1, \dots, T-1\} \\ k_0, & t = 0, \end{cases}$$

with $k_t \in \mathbb{R}$ and $v_t \in \mathbb{R}^t$ to be calculated recursively as

$$k_t := k_{t+1} + W_t \left(\left(1 + v_{t+1}^{(t+1)}\right) \Sigma_{t+1|1:t}^{1/2} \epsilon_{t+1} \right)$$
$$\left(v^{(t)}\right)^T := \left(v_{1:t}^{(t+1)}\right)^T + \left(1 + v_{t+1}^{(t+1)}\right) \Sigma_{t+1:1:t} \Sigma_{1:t,1:t}^{-1}$$

and initial conditions $k_T = 0, v^{(T)} = 0, v^{(0)} = 0$. Here $(\epsilon_t)_{t=1}^T$ is a sequence of independent standard normally distributed random variables such that ϵ_{t+1} is $\mathcal{G}_{t+1}$-measurable and independent of $\mathcal{G}_t$.

3.2 Set-up and assumptions

3.2.1 One-cohort case

We consider first a cohort of 1000 individuals signing life insurance paying out one SEK at the time of death, with a duration of 30 years. We assume that the entire portfolio consists of 50 years old men with identical mortality law. The yearly number of deaths is assumed to follow a binomial process. Under these binomial assumptions, we calculate the yearly number of expected mortalities and the covariance matrix. We fix these values as the true characteristics of the stochastic process. However, we exchange the binomial assumption with a normal assumption at this point. In other words, we consider a Gaussian process with the binomial process' characteristics so that we can compute the cost-of-capital margin using theorem 3. Note that despite this change of distribution, the process is still Markovian. While there is a correlation between the number of deaths at different time steps, we can calculate the upcoming year's death toll's expectancy and variance if we know the number of survivors at time t. Since the process is Markovian, equally well as sampling the entire vector of deaths, we may start at time 0 and simulate the number of deaths year by year.

The Markov property enables us to use an LSM algorithm that does not need empirically estimated variances and covariances. Algorithm 2 describes this algorithm from the point where sample paths have been simulated using μ and Σ computed using standard results regarding binomial distributions. The design of the algorithm is similar to the design used in Engsner [6]. In figure 2, the results are illustrated and compared to the true values.

3.2.2 Two-cohort case

In a real-life insurance portfolio, there are usually many different cohorts. We consider such a two-cohort portfolio to evaluate the LSM algorithm's ability to estimate the cost-of-capital margin of aggregated cash flows arising from correlated cohorts. The starting point will be a Gaussian process where we pre-condition on a covariance structure. Since the overall purpose is to evaluate the LSM algorithm, we will prioritize computability over realism.

In the one-cohort case, we pre-conditioned on a covariance matrix. This was straightforward, since we were able to use the moments from a corresponding

Algorithm 2 LSM algorithm for the one-cohort life insurance portfolio

$\hat{\beta}_T = (0, 0)$
D_t^{static} = static repliacating portfolio, $t \in 1, ..., T$
$S_t^{(i)},\ i \in 1, ..., M$ and $t \in 0, ..., T$ simulated paths

for t in T-1:0 **do**
  for i in 1:M **do**
    $\mu_{t+1} = S_t^{(i)} q_{50}(t)$
    $\sigma_{t+1}^2 = S_t^{(i)} q_{50}(t) p_{50}(t)$
    Sample $S_{t+1}^{(i,j)}$ from $N(\mu_{t+1}, \sigma_{t+1}),\ j \in 1, .., n$
    $D_{t+1}^{(i,j)} = S_{t+1}^{(i,j)} - S_t^{(i)}$
    $Y_{t+1}^{(i,j)} = D_{t+1}^{(i,j)} - D_{t+1}^{static} \ + \ \hat{\beta}_{t+1} \cdot (1,\ S_{t+1}^{(i,j)})^T$
    $R_t^{(i)} = \text{sort}\big(Y_{t+1}^{(i,j)}\big)[\text{ceil}(0.995n))]$ ▷ Empirical Value-at-Risk
    $E_t^{(i)} = \frac{1}{n} \sum_j \big(R_t^{(i)} - Y_{t+1}^{(i,j)}\big)_+$
    $W_t^{(i)} = R_t^{(i)} - \frac{1}{1+\eta} E_t^{(i)}$
  end for
  α_t, β_t = OLS parameters from redressing W_t on S_t
  $\hat{\beta}_t = (\alpha_t, \beta_t)$
end for

$V_0 = \frac{1}{M} \sum W_0^{(i)}$ ▷ The cost-of-capital margin binomial process. When we now consider a 2-dimensional stochastic process, it is still possible to use the binomial moments, but we also need to handle the cross-covariances. The sum of two Gaussians is not necessarily Gaussian. To ensure normality in the aggregated cash flows, we need to make stronger assumptions. The idea is to simulate the two-dimensional length T stochastic process as a one-dimensional length $2T$ Gaussian process. Such a setup lets us control the cross-covariances, and the Gaussian aggregated cash flow is a linear transformation away.

Let N_t^A, N_s^B be the number of survivors for cohort A and B, at times t and s, and D_t^A, D_s^B the corresponding number of year-wise fatalities. Let $t \leq s$ and $c \in [0,1]$. For the purpose of testing the LSM algorithm's ability to estimate the cost-of-capital margin when the cash flow comes from correlated cohorts, we use a structure where

$$\operatorname{corr}(D_t^A, D_s^B) = \frac{c}{s-t+1} \Leftrightarrow \operatorname{cov}(D_t^A, D_s^B) = \frac{c\sqrt{\mathrm{V}(D_t^A)\,\mathrm{V}(D_s^B)}}{s-t+1}.$$

The in-cohort expectancies and variances are modelled as in the one-cohort case,

$$\operatorname{cov}(D_t, D_{s+1}) = n_0 S(t) q_t S(s+1) q_{s+1},$$

$$\mathrm{V}(D_{t+1}) = N_t q_t p_t,\ \mathrm{E}[D_{t+1}] = N_t q_t.$$

We consider for ease of notation only the case where $T = 2$. Then we define $X_{AB} = (D_1^A, D_2^A, D_1^B, D_2^B)$ and

$$\Sigma_0^{AB} = \begin{pmatrix} \sigma^2_{A,1} & \sigma^2_{A,12} & \sigma^2_{AB,11} & \sigma^2_{AB,12} \\ \sigma^2_{A,12} & \sigma^2_{A,2} & \sigma^2_{AB,21} & \sigma^2_{AB,22} \\ \sigma^2_{AB,11} & \sigma^2_{AB,21} & \sigma^2_{B,1} & \sigma^2_{B,12} \\ \sigma^2_{AB,12} & \sigma^2_{AB,22} & \sigma^2_{B,12} & \sigma^2_{B,2} \end{pmatrix},$$

where the components of the covariance matrix has been estimated using the covariance structure defined above. We simulate from this Gaussian process and create an aggregate cash flow as

$$KX_{AB},\ K = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

It follows that the aggregate cash flow is Gaussian too, with

$$\mu = K\mu_{AB},\ \Sigma = K\Sigma_{AB}K^T.$$

We may now compute the exact CoC margin using the explicit formula from theorem 4. Figure 2 illustrates the impact of the implictly definied c parameter when the 1000 individuals from the one-cohort case are divided into two equally sized cohorts. As expected, the margin stays the same when c = 0, increases when the correlation is positive and decreases otherwise.

We will proceed similarly to the one-cohort case but with some modifications to

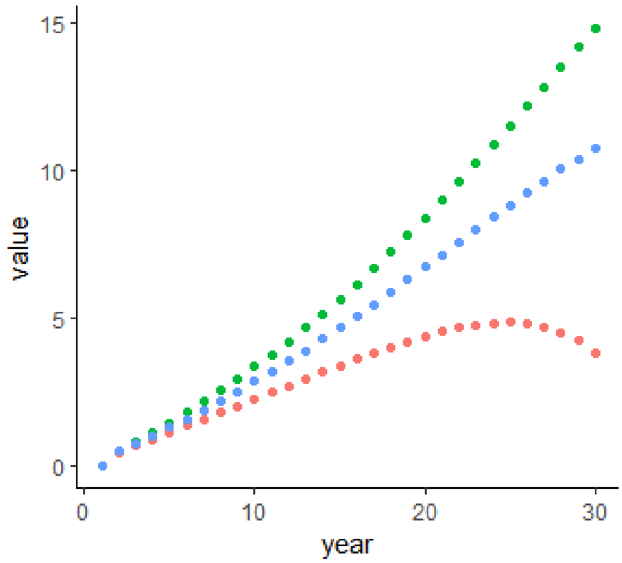


Figure 2: The cost-of-capital margin for increasingly long durations, with c = 0.1 (green), c = 0 (blue) and c = -0.1 (red)

handle the added dimension. Most importantly, the model's inner simulations will be drawn from a 2-dimensional normal distribution with a covariance that has to be estimated, and the regression part will include at least one additional parameter. With these modifications at hand, the only pre-requirement for the LSM algorithm will be M simulated paths of the two-dimensional vector of survivors and estimated cohort-correlations at times 1,..,T. Algorithm 3 in the appendix describes the modified LSM algorithm used for the two-cohort case.

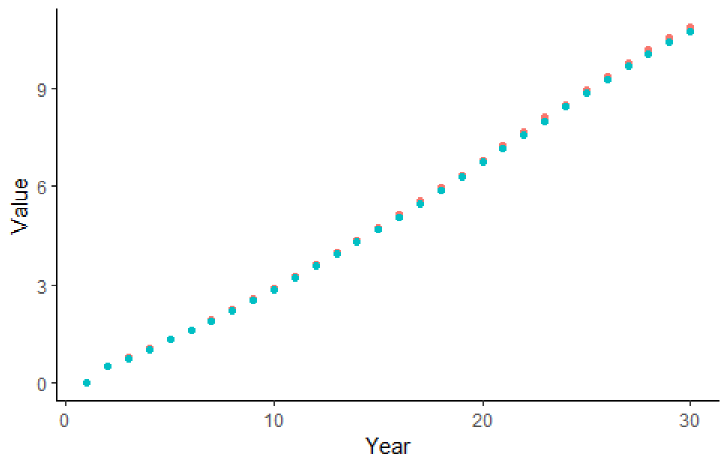
4 Results

This section contains the results from the one-cohort and two-cohort cases, carrying out calculations in the programming language R. We use the built-in function LM for regression and the function corr to estimate correlation in the two-cohort case.

4.1 One-cohort case

For the one-cohort case, we test two different models. In the first model, the regression model consists of only an intercept term, without regard to the number of survivors. The second model uses an intercept parameter and a slope parameter. All runs use 1000 external simulations. The number of internal simulations is 1000, 5000, and 10000. Table 1 shows how the respective models

Figure 3: One-cohort CoC margins V_0 for duration ranging from 1 to 30



LSM (red) compared to Gaussian (blue)

Table 1: Results from the one-cohort case

T	M	n	intercept	intercept + slope	Gaussian
15	1000	1000	4.95	4.74	4.67
15	1000	5000	4.92	4.68	4.67
15	1000	10000	4.89	4.68	4.67
30	1000	1000	13.68	10.82	10.73
30	1000	5000	13.39	10.72	10.73
30	1000	10000	13.55	10.71	10.73

perform. The model with only an intercept term estimates the margin reasonably well but with a significant deviation. The more advanced model estimates the margin almost entirely without deviation for sufficiently large values of n.

Table 2: Results from the two-cohort case, with models 1-6 definied as in table 2

c	M	n	M1	M2	M3	M4	M5	Gaussian
-0.1	1000	1000	12.55	10.31	10.32	10.34	10.22	3.86
-0.1	1000	5000	12.32	10.17	10.16	10.13	10.14	3.86
-0.1	1000	25000	12.26	10.11	10.13	10.13	10.18	3.86
0.1	1000	1000	12.10	13.36	11.36	11.39	11.36	14.83
0.1	1000	5000	11.95	11.21	11.26	11.24	11.20	14.83
0.1	1000	25000	12.14	11.29	11.19	11.24	11.25	14.83

Table 3: Description of models in table 2

Model	Expression
M1	$\beta_0 + \beta_1 A$
M2	$\beta_0 + \beta_1 A + \beta_2 B$
M3	$\beta_0 + \beta_1 A + \beta_2 B + \beta_3 AB$
M4	$\beta_0 + \beta_1 A + \beta_2 B + \beta_3 AB + \beta_4 A^2 + \beta_5 B^2$
M5	$\beta_0 + \beta_1 A + \beta_2 B + \beta_3 AB + \beta_4 A^2 + \beta_5 B^2 + \beta_6 A^3 + \beta_7 B^3$

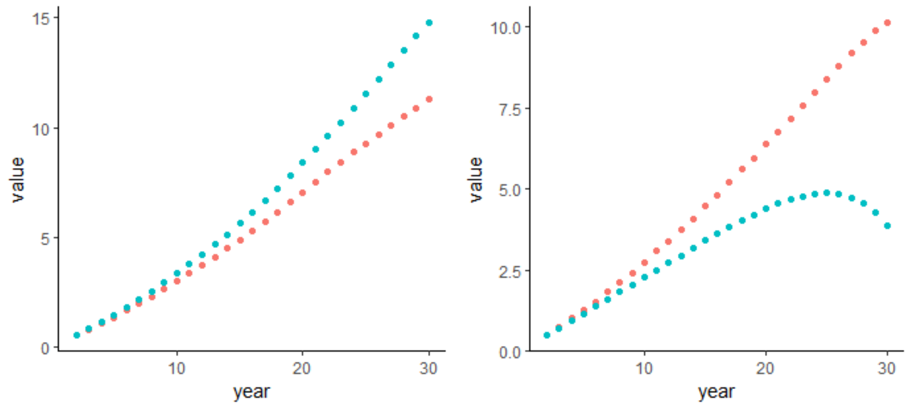
4.2 Two-cohort case

In the case of two cohorts, we examine the performance of the LSM algorithm for three different values of parameter c. We use five different basis functions. When c = 0, the same layout is obtained as in the one-cohort case. Quite rightly, we also get the same value on the cost-of-capital margin. In this case, the LSM algorithm has no problem estimating the margin as long as the base function considers the number of survivors of both cohorts. When c deviates from 0, The Gaussian margin becomes as in Figure 4, and the LSM algorithm's estimates for $T = 30$ become as in Table 2. Although gradually more advanced sets of basis functions are used, the algorithm does not succeed in estimating the margin. The algorithm understands that the margin becomes larger when c is positive and smaller otherwise, but without converging towards the correct value.

5 Discussion

We have seen how to value a residual cash flow due to imperfect replication based on the cost of capital. Under assumptions about Gaussian cash flows, we calculated this cost-of-capital margin exactly for an insurance portfolio and compared this margin with that of a valuation based on a least-squares Monte

Figure 4: Cost-of-capital margins from the two-cohort case. The red dots are estimates from the LSM algorithm run with $M = 1000$, $n = 5000$ and model M3 for the basis functions.



(a) $c = 0.1$ **(b)** $c = -0.1$

Carlo estimation. It was possible to estimate the exact value without deviation for a homogeneous insurance portfolio.

We split the insurance portfolio into two correlated cohorts and compared two different scenarios - one with a positive correlation between the cohorts and one with a negative correlation. A positive correlation led to an increased value of the cost-of-capital margin and a negative correlation to a decreased value. In this more complex case, the LSM algorithm did not succeed in estimating the cost-of-capital margin. Different basic functions of varying valence were tested but did not significantly affect the estimate. The LSM algorithm captured the direction the cost-of-capital margin moved due to the correlation but failed to estimate the magnitude.

Future work that implements more well-thought-out approaches to the LSM algorithm in the two-cohort case would be interesting. The thesis used a perhaps naive direct generalization of the one-cohort modeling the cash flow as if it were Markovian. More consideration in estimating the actual conditional covariance matrix at each time step of the iteration could help improve the LSM algorithm.

The LSM algorithm used could have been better evaluated if an underlying covariance structure that preserves the Markov property of the cash flow had been used. As table 2 shows, increasing the degree of the basis functions did not improve performance, indicating that the estimated covariance matrix was biased.

A next step given a better LSM algorithm would be to replace the covariance structure with estimates from actual data. The theory presented in appendix A3 could be a good breeding ground for such estimation. It is also possible to increase the number of cohorts by generalizing the simulation approach used in the two-covariance case.

References

[1] H. Engsner, M. Lindholm, and F. Lindskog. "Insurance valuation: A computable multi-period cost-of-capital approach". In: *Insurance Mathematics and Economics 72, 250-264.* (2017).

[2] EIOPA. "Commission Delegated Regulation (EU) 2015/35 of 10 October 2014". In: *Official Journal of the European Union* (2015).

[3] A. Pelkiewicz et al. "A review of the risk margin – Solvency II and beyond". In: *British Actuarial Journal, 25, E1* (2020).

[4] EIOPA. "Opinion on the 2020 review of Solvency II". In: (2020).

[5] C. Möhr. "Market-consistent valuation of insurance liabilities by cost of capital." In: *ASTIN Bull. 41, 315-341* (2011).

[6] H. Engsner. "Least Squares Monte Carlo applied to Dynamic Monetary Utility Functions". In: *Papers 2101.10947, arXiv.org, revised Apr 2021* (2021).

[7] European Commision. "Commission Delegated Regulation (EU) 2015/35 of 10 October 2014." In: *Official Journal of the European Union.* (2015).

[8] F. Longstaff and E. Schwartz. "Valuing American options by simulation: A simple least-squares approach". In: *The Review of Financial Studies, 14 (1),113-147.* (2001).

[9] N. Brouhns, M. Denuit, and J. Vermunt. "A Poisson log-bilinear regression approach to the construction of projected lifetables". In: *Insurance: Mathematics and Economics, 31, (3), 373-3930* (2002).

[10] A. M. Villegas, V. K. Kaishev, and P. Millossovich. "StMoMo: an R package for stochastic mortality modeling." In: *Journal of Statistical Software 84(3), 1–38* (2018).

[11] Human Mortality Database HMD. "University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany)." In: (2014).

Appendix

A Implementation details

In section 3, the Gaussian cost-of-capital margin is calculated using theorem 6 through the implementation in algorithm 4 below.

Algorithm 3 Computing the Gaussian CoC margin

```
Import(mean, cov) ▷ Importing Gaussian characteristics for yearly number
of fatalities
function R(mean, sd)                          ▷ Value at risk for normal variable
    return quantile(0.995, mean, sd)
end function
function U(mean, sd)                          ▷ Selected utility function
    R ← R(mean, sd)
    integrand(x) ← (x + R)· density(x, mean, sd)
    U ← integrate(integrand(x), lower = -R, upper = ∞)
    return U
end function
function W(mean, sd)                          ▷ W used in the recursion
    eta ← 0.06
    R ← R(mean, sd)
    U ← U(mean, sd)
    W ← R - U/(1 + eta)
    return W
end function
T ← 30
v, k ← list(T), list(T)
v[T] ← rep 0 T+1 times
k[T] ← 0
for t in (T-1):2 do
    cVar ← C[t+1, t+1] - C[t+1, 1:t] * inv(C[1:t, 1:t]) * C[1:t, t+1]
    sd ← (1 + v[t+1][t+1]) * sqrt(cVar)
    Wt ← W(0, sd)
    k[t] ← k[t+1] + Wt
    v[t] ← v[t+1][1:t] + (1+v[t+1][t+1]) * C[t+1, 1:t] * inv(C[1:t, 1:t])
end for
```

Algorithm 4 LSM algorithm for a two-cohort life insurance portfolio

$\hat{\beta}_T$ = repeat(0, N) ▷ length N vector of 0's
D_t^{static} = static repliacating portfolio, $t \in 1, ..., T$
$S_t^{(i)}$, simulated 2-dimensional paths with $i \in 1, ..., M$ and $t \in 0, ..., T$
corr = length T vector of estimated correlations

for t in T-1:0 **do**
  for i in 1:M **do**
    $\mu_{t+1} = \left(S_t^{(i,A)} q_{50}^A(t),\ S_t^{(i,B)} q_{50}^B(t) \right)$
    $\left(\sigma_{A,t+1}^2,\ \sigma_{B,t+1}^2\right) = \left(S_t^{(A,i)} q_{50}^A(t) p_{50}^A(t),\ S_t^{(A,i)} q_{50}^A(t) p_{50}^A(t) \right)$
    $\Sigma_{t+1} = \begin{pmatrix} \sigma_{A,t+1}^2 & \mathrm{corr}_{t+1}\sigma_{A,t+1}\sigma_{B,t+1} \\ \mathrm{corr}_{t+1}\sigma_{A,t+1}\sigma_{B,t+1} & \sigma_{B,t+1}^2 \end{pmatrix}$
    Sample $S_{t+1}^{(i,j)}$ from $N(\mu_{t+1}, \Sigma_{t+1}),\ j \in 1, .., n$
    $D_{t+1}^{(i,j)} = S_{t+1}^{(i,j)} - S_t^{(i)}$
    $Y_{t+1}^{(i,j)} = D_{t+1}^{(i,j)} \cdot (1,1) - D_{t+1}^{static} \ + \ \hat{\beta}_{t+1} \cdot \Phi(S_{t+1}^{(i,j)}),\ j \in 1, .., n$
    $R_t^{(i)} = \mathrm{sort}\big(Y_{t+1}^{(i,j)}\big)[\mathrm{ceil}(0.995n))]$ ▷ Empirical Value-at-Risk
    $E_t^{(i)} = \frac{1}{n} \sum_j \big(R_t^{(i)} - Y_{t+1}^{(i,j)}\big)_+$
    $W_t^{(i)} = R_t^{(i)} - \frac{1}{1+\eta} E_t^{(i)}$
  end for
  $\hat{\beta}_t$ = OLS parameters from redressing W_t on S_t
end for

$V_0 = \frac{1}{M} \sum W_0^{(i)}$ ▷ The cost-of-capital margin

A2 Additional mathematical details

In section 2 some details about risk measures were omitted. These details are necessary to derive theorems 2, 3 and 4. Although the conditional expectation E_t has been sufficient for our purposes, we here include a more general definition of a dynamic utility function to bring justice to theorem 5. We cite straight off [1].

Definition 4. For $p \in [0, \infty]$, a dynamic monetary utility function $(U_t)_{t=0}^{T-1}$ is a sequence of mappings $U_t : L^p(\mathcal{F}_{t+1}) \to L^p(\mathcal{F}_t)$ satisfying

$$\text{if } \lambda \in L^p(\mathcal{F}_t) \text{ and } Y \in L^p(\mathcal{F}_{t+1}), \text{ then } U_t(Y+\lambda) = U_t(Y) + \lambda, \tag{21}$$

$$\text{if Y, } \tilde{Y} \in L^p(\mathcal{F}_{t+1}) \text{ and } Y \leq \tilde{Y}, \text{ then } U_t(Y) \leq U_t(\tilde{Y}), \tag{22}$$

$$U_t(0) = 0 \tag{23}$$

We also define, for probability measures M^R and M^U on $(0,1)$, for $Y \in L^p\left(\mathscr{F}_{t+1}\right)$,

$$R_t(Y) := \int_0^1 F_{t,-Y}^{-1}(u) dM^R(u), \tag{24}$$

$$U_t(Y) := \int_0^1 F_{t,Y}^{-1}(u) dM^U(u). \tag{25}$$

Theorem 5 (Proposition 4)**.** *Suppose there exist $\mu_0 \in (0, 1/2)$ and $\bar{m} \in (0, \infty)$, such that, for $k = R,\ U$,*

$$\begin{aligned} &\max\left(M^k((u,v)), M^k((1-v, 1-u))\right) \\ &\quad \leq \bar{m}(v-u) \text{ for all } 0 < u < v < u_0. \end{aligned}$$

Fix $p \in [1, \infty]$.
(i) $\mathrm{R_t}$ in (24) and U_t in (25) are well-defined as mappings from $L^p\left(\widetilde{F}_{t+1}\right)$ to $L^p\left(\mathscr{F}_t\right)$ and satisfy $(5)-(7)$ and $(21)-(23)$, respectively.
(ii) If $Y \in L^p\left(\mathscr{F}_{t+1}\right)$ and, for any Borel set $A \subset \mathbb{R}$

$$\mathbb{P}\left(Y \in A \mid \mathscr{F}_t\right) = \mathbb{P}\left(Y^{(1)} + Y^{(2)} Y^{(3)} \in A \mid \mathscr{F}_t\right),$$

where $Y^{(1)} \in L^p\left(\mathscr{F}_t\right), 0 < Y^{(2)} \in L^0\left(\mathscr{F}_0\right)$, and $Y^{(3)} \in L^p\left(\mathscr{F}_{t+1}\right)$ is independent of $\mathscr{F}_t$, then

$$R_t(Y) = -Y^{(1)} + Y^{(2)} R_t\left(Y^{(3)}\right),$$

$$U_t(Y) = Y^{(1)} + Y^{(2)} U_t\left(Y^{(3)}\right),$$

where $R_t\left(Y^{(3)}\right), U_t\left(Y^{(3)}\right) \in \sim^0 \left(\mathscr{F}_0\right)$. Moreover, $R_t\left(Y^{(3)}\right) = R_0\left(\widetilde{Y}_{(3)}^{(3)}\right)$ and $U_t\left(Y^{(3)}\right) = U_0\left(\tilde{Y}^{(3)}\right)$ for $\widetilde{Y}^{(3)} \in L^p\left(\widetilde{F}_1\right)$ with $Y^{(3)}$ and $\widetilde{Y}^{(3)}$ equally distributed. For W_t in (13), $W_t(Y) = Y^{(1)} + Y^{(2)} W_t\left(Y^{(3)}\right)$, and if $\eta_t \in L_+^0\left(\mathscr{F}_0\right)$, then $W_t\left(Y^{(3)}\right) \in L^0\left(\mathscr{F}_0\right)$. Further, if $\eta_t = \eta_0$, then $W_t\left(Y^{(3)}\right) = W_0\left(\tilde{Y}^{(3)}\right)$ for $\tilde{Y}^{(3)} \in L^P\left(\mathscr{F}_1\right)$ with $Y^{(3)}$ and $\tilde{Y}^{(3)}$ equally distributed.

A3 Life insurance details

For modelling the hazard rates, we used a Makeham distribution with parameters $\alpha = 0.001$, $\beta = 0.000012$ and $\gamma = 0.101314$ throughout the thesis.

Definition 5. Makeham law of mortality states that the hazard rate for an x years old individual is given by

$$\mu_x = \alpha + \beta e^{\gamma x}, \text{ for } \alpha,\ \beta,\ \gamma > 0.$$

We could equally well have used the Lee-Carter inspired model described in Brouhns, Denuit, and Vermunt [9]. The later approach offers a statistically sound way of forecasting and simulating future years. The thesis never got to a stage where this was relevant, but there are several ways the Brouhn model could have enrichened the thesis.

In implementing the model from Brouhn, the number of fatalities $D_x(t)$ in the period starting at t are assumed to be Poisson distributed. With exposure $w_x(t)$, we assume that

$$D_x(t) \sim Po\left(w_x(t)\mu_x(t)\right).$$

The hazard rate itself is modelled as

$$log(\mu_x(t)) = \alpha_x + \beta_x \kappa_t, \tag{26}$$

where α_x is the age dependent effect, β_x the trend sensitivity at age x and κ_t the general mortality trend. Fitting the model to human mortality data, we can use the hazard rate to simulate a life insurance portfolio.

Let $N_x(t)$ be the number of x years old alive at time t. Assuming a portfolio initially consisting of $k+1$ age groups, we start off with $\left(N_{x_0}(0), \dots, N_{x_k}(0)\right)$, This is modelled as a discrete multivariate stochastic process where $t \in 0, 1, 2, ..., T$. In particular, using the Brouhn assumptions from this section, $t+1$ given t is assumed to be distributed as

$$N_{x+1}(t+1) \sim \text{Binom}(N_x(t), p_x(t)), \text{ with } p_x(t) = 1 - q_x(t).$$

Note that even if the different age groups are be modelled independently, conditional on the values on $N_x(t)$ and $p_x(t)$, correlation between the age groups will still be present, since we use (21) to model the hazard rate. The second term, the time-effect, is what gives rise to age group correlation.

Simulating data

In this section, we implement the mortality model from Brouhn. Using the R package StMoMo from Villegas, Kaishev, and Millossovich [10], we generate a Gaussian cash flow model corresponding to a specific life insurance portfolio.

As a starting point, we consider 1000 males in each of the ages from 50 to

60. Hence we start with 11000 individuals all signed up for the same type of insurance product. These numbers are chosen to give a somewhat realistic representation of an insurance portfolio without requiring too much extra work, but the starting position can easily be simplified or widened. We could equally well restrict our attention to one age group of 50 years old males, widen it to a larger span of ages and not only men. The insurance product analyzed is a traditional life insurance product paying one SEK if a fatality occurs and otherwise nothing, over the years $0, ..., T$. The LSM method and the normal approximation allow for more complex insurance products, but this set-up is sufficient since it possesses interest's underlying cash flow structure.

Fitting the mortality model

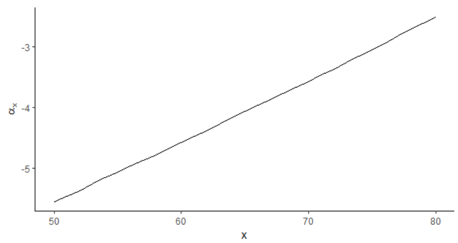
Based on the HMD [11], we fit a Lee Carter model. We use ages 30-90 and years 1960-2017. This gives a poisson model with predictor

$$log\left(\mu_x(t)\right) = \alpha_x + \beta_x \kappa_t, \tag{27}$$

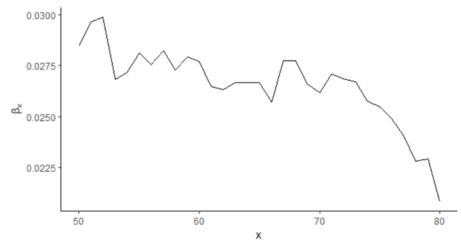
where α, β_x and κ_t all are fitted using the *fit* function provided in StMoMo. Figure 2 shows the estimated parameters. From the same model we simulate future mortality rates for years beyond 2017. The simulation uses techqnique from time-series analysis to simulate the κ_t. We use the StMoMo function *simulate* for this purpose.

The StMoMo package simulates a realistic life insurance portfolio development using a Lee Carter model. This corresponds to fitting parameters $\alpha_x(t), \beta_x(t)$ and $\kappa_x(t)$ for x and t values such that the outcome is observable. For future time values, $\alpha_x(t), \beta_x(t)$ are extrapolated and $\kappa_x(t)$ simulated as a time series. This means that each round of simulation renders a set of hazard rates and corresponding portfolio set-up.

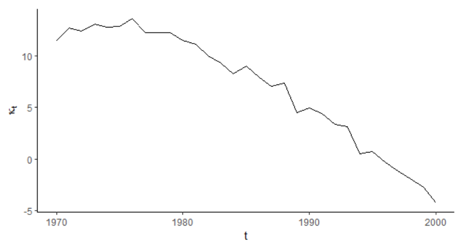
Based on these simulated trajectories, we simulate year-by-year fatalities resulting in a cash flow. We fit a multivariate normal variable on these cash flows. From the fitted normal cash flow, we can simulate new cash flows. While a Gaussian distribution is not the go-to choice for modelling mortalities, approximating the aggregated cash flow with a normal distribution is not so far fetched. More importantly, simulating from a Gaussian distrubtion allows for explicit computation of the cost-of-capital margin, using theorem 5. This is necessary, since this allows us to evaluate the performance of the least squares Monte carlo algorithm. This fatality development is illustrated in figure 3. Clearly, since the analyzed insurane product only pays out at deaths, the quantities illustrated in figure 3 correspond to accumulated costs and yearly costs, respectively. We perform a large number of simulations, $N = 5 \cdot 10^5$ providing us with the same of number of aggregated cash flows. Each cash flow is now a vector X of length 10, such that element 1 is the number of deaths in year 1, and so on. Examples of such simulations are shown in figure 3. We estimate the cash flows mean



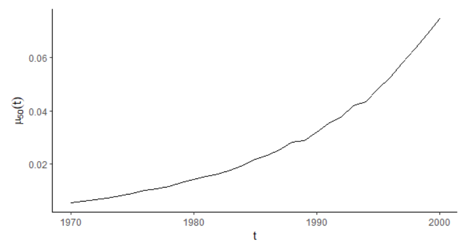
(a) α_x as a function of age



(b) β_x as a function of age



(c) κ_t as a function of time



(d) α_{50} as a function of time

Figure 5: Above graphs of first the three fitted parameters, and in the bottom right we see resulting mortality rates for 50 years old men.

vector and covariance matrix using empiricial estimators

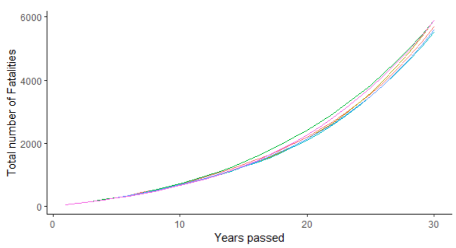
$$\hat{x} = \frac{1}{N}\sum_{i=1}^{N} x^{(i)}, \text{ where } x^{(i)} = (x_1^{(i)}, ..., x_T^{(i)})$$

$$\hat{\Sigma} = \frac{1}{N-1}(X - 1_N\hat{x})^T(X - 1_N\hat{x}),$$

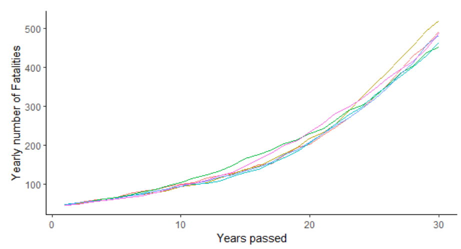
$$X = \begin{pmatrix} x_1^{(1)} & ... & x_t^{(1)} \\ \vdots & \vdots & \vdots \\ x_1^{(N)} & ... & x_t^{(N)} \end{pmatrix}.$$

Since the number of simulations N is large, the empirical covariance matrix converges element-wise to the fourth decimal value. We may simulate similar cash flows from a multivariate normal distribution using these estimates. Figures 3 to 4 illustrate simulated cash flows from the Lee-Carter model and Gaussian flows' corresponding approximations.

Figure 6: Below the development for six simulated sets of parameters.

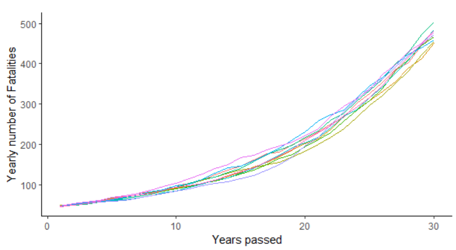


(a) Fatalities counted cumulative

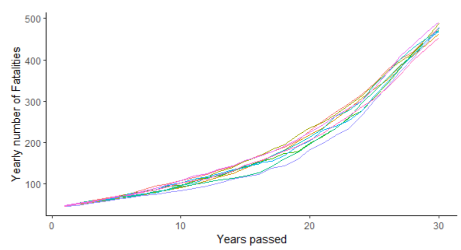


(b) Fatalities counted incrementally

Figure 7: Incremental cash flows from simulated parameter sets



(a) Lee Carter simulation



(b) Gaussian approximation simulation