Sensitivity Analysis for Ruin Probabilities

Canonical Risk Model

Felisa J. Vázquez-Abad and Phong LeQuoc
Department of Computer Science and Operations Research
University of Montreal, Montréal, Quebec H3C 3J7,
e-mail : vazquez@IRO.UMontreal.ca

Revised February 13, 2000
Submitted to *Journal of the Operational Research Society*
Special Issue: Progress in Simulation Research

Abstract

The surplus process of an insurance portfolio is defined as the wealth obtained by the premium payments minus the reimboursements made at the times of claims. When this process becomes negative (if ever), we say that ruin has occurred. The general setting is the Gambler's Ruin Problem. In this paper we address the problem of estimating derivatives (sensitivities) of ruin probabilities with respect to the rate of accidents. Estimating probabilities of rare events is a challenging problem, since naïve estimation is not applicable. Solution approaches are very recent, mostly through the use of Importance Sampling techniques. Sensitivity estimation is an even harder problem for these situations. We shall study three methods for estimating ruin probabilities: one via importance sampling (IS), and two others via indirect simulation: the storage process (SP), which restates the problems in terms of a queueing system, and the convolution formula (CF). To estimate the sensitivities, we apply the RPA method to IS, the IPA method to SP and the Score Function method to CF. Simulation methods are compared in temrs of their *efficiency*, a criterion that appropriately weighs precision and CPU time. As well, we indicate how other criteria such as set-up time and prior formulæ development may actually be problem-dependent.

Keywords: Sensitivity analysis, rare event simulation, importance sampling, estimation efficiency.

1 Introduction: Ruin Probabilities

The canonical model in Risk Theory assumes that claims due to accidents arrive according to a Poisson process $N(t)$ of rate λ. The successive claim amounts, denoted $\{Y_i\}$, are i.i.d. random variables with general distribution G and premiums are received at a constant rate c. If the initial endowment is $u > 0$, the wealth of the insurance company, known as the *surplus process* is:

$$U(t) = u + ct - \sum_{i=1}^{N(t)} Y_i, \qquad t \geq 0 \tag{1}$$

The event epochs of the process $N(t)$ are denoted by $\{T_n, n \geq 0\}$, and $W_n = T_n - T_{n-1}$ are the interarrival times. The cumulative claims process $S(t) = \sum_{i=1}^{N(t)} Y_i$ is a compound Poisson process. We shall often write $U_n = U(T_n)$ to denote the embedded discrete event process and $S_n = \sum_{i=1}^{n} Y_i$, with an obvious abuse of notation. If we set $\tau = \min\{n : U(T_n) < 0\} = \min\{n : u + cT_n \leq S_n\}$ then the ruin probability is

$$\psi(u, \lambda) = P\{\tau < \infty\}$$

and it is a measure of the credit risk of the company. Call $\beta = E[Y_1]$. If $c \leq \lambda\beta$, then $\psi(u, \lambda) = 1$ for all initial endowment u, as shown in Gerber (1979). As a consequence of this result, it is common to assume that premiums satisfy $c > \lambda\beta$, which we shall do.

We briefly review some actuarial notation. The *relative security loading* θ is defined by $c = (1+\theta)\lambda\beta$. If the moment generating function of Y exists, call $\gamma > 0$ the maximal number such that $M_Y(r) = E[e^{rY}] < \infty$ for all $r < \gamma$. The *adjustment coefficient* R is defined as the (unique) non-zero solution of the implicit equation:

$$\lambda + cR = \lambda M_Y(R), \quad R < \gamma \tag{2}$$

and $\psi(u, \lambda) < e^{-Ru}$, known as *Lundberg's exponential bound*, under the assumption that $R < \infty$ exists, as stated in de Vylder (1996).

The *sensitivity* of a function $\psi(u, \lambda)$ to λ is defined as the corresponding derivative (if it exists). It can yield important information for forecasting, prevision and optimization. In Vázquez-Abad and Zubieta (2000) it is shown that for a large number of models, sensitivities to the parameters β, c and λ bear simple relationships, so it is enough to know how to estimate just one of them. Estimting ψ by simulation is an example of a problem where absorption (ruin, in this case) may or may not happen in a particular trajectory, making it difficult to establish a stopping rule for the simulated trajectories. Section 2 contains a review of three methods for simulating ruin probabilities: one via importance sampling (IS), and two others via indirect simulation: the storage process (SP), which restates the problems in terms of a queueing system, and the convolution formula (CF). Section 3 introduces the methods for estimating the sensitivities. The RPA method is used for IS, the IPA method for SP and the Score Function method for CF. We have chosen a different sensitivity method per simulation method, based on the ease of implementation and applicability of sensitivity estimation methods for each problem. Section 4 explains our simulation results, where we have chosen various scenarios, including the difficult case where claims have a Pareto distribution, which induces long-range dependencies. Our findings show that the performance of the estimators, measured in terms of precision (variance), execution time and set-up time, may greatly depend on the problem. For example, while the SP+RPA method is robust and need no set up time, it can be the worst to perform for some distributions of the claim sizes. The CF+SF, while yielding overall the best results is very hard to implement in many situations, since the so-called ladder distribution of the claim sizes has to be computed in advance and generated within the simulation. Which method to use in practice will largely depend on the needs of the estimation: if the estimation has to be performed very often with different scenarios (including use of historical data for the empirical distribution of claim sizes) then neither IS nor CF would be recommended. However, if the distribution is known and one very accurate estimation of risk must be performed to aid in critical decision making, robustness of SP may well be outperformed by CF or IS, even if a considerable amount of time and effort may have to be invested setting up the methods.

2 Estimating Ruin Probabilities

For general distributions G there is no analytical expression for the ruin probability available, and one often resorts to simulation techniques. However, estimating $\psi(u, \lambda)$ via direct simulations of the process $U(t)$ would require estimating $P(A)$, where A is a rare event. For paths where ruin is not achieved, we don't have any infallible stopping rule for the simulation. Evaluating the sensitivity of ψ to the various model parameters may be even more difficult. In this work, we shall focus on the estimation of the sensitivities to the arrival rate λ.

2.1 Importance Sampling

Let $h[(W_i, Y_i); i = 1, \dots, \tau] = \mathbf{1}_{\{\tau<\infty\}}$, then by definition, $E[h] = \psi(u, \lambda)$, where E denotes expectation with respect to the measure of the process when the accidents occur according to a Poisson process with rate λ and the claims are i.i.d.$\sim G$.

The method of importance sampling presented in Asmussen and Nielsen (1995) is based on a change of measure under which ruin is certain. Call $g(y)$ the density of Y_1 (or $P[Y_1 = y]$ in the case of discrete claim distributions) and assume that $Y_i > 0$ a.s. Suppose that the moment generating function $M_Y(\cdot)$ exists and $R < \infty$ is well defined. The approach prescribes simulating the process $U(t)$ where accidents now occur as a Poisson process at a rate

$$\lambda_0 = \lambda + cR$$

and the claims follow the exponentially tilted distribution (also known as the Esscher transform of Y or the Lundberg transformation):

$$\tilde{g}(y) = \frac{e^{Ry} g(y)}{M_Y(R)},$$

Under some assumptions on G, Asmussen (1985) shows that under the new measure, ruin is certain. Under the original measure, $P(\tau < \infty) < 1$. The new probability $\tilde{P}$ is such that $\{(W_i, Y_i), i = 1, \dots, \tau\}$ are i.i.d., and $\tilde{P}(\tau < \infty) = 1$. Consider the measure $P(\cdot)$ restricted to the event $A = \{\tau < \infty\} \subset \Omega$, where Ω is the underlying probability space. The restricted measure is a defective probability defined on the inherited σ-algebra $\sigma(A) \subset \sigma(\Omega) = \mathcal{F}$ and is absolutely continuous w.r.t. the new measure, which is a probability on A, so we can apply Radon-Nykodim's Theorem (see Chung (1974)) to obtain $E[h\mathbf{1}_{\{A\}}] = \tilde{E}[Lh]$ where L is the Radon-Nykodim derivative, also known as likelihood-ratio:

$$\psi(u,\lambda) = \tilde{E}\left[\left(\frac{\lambda M_Y(R)}{\lambda_0}\right)^{\tau} \exp \sum_{i=1}^{\tau}[(\lambda_0 - \lambda)W_i - RY_i]\right] = \tilde{E}\left[e^{R[cT_\tau - S_\tau]}\right] \tag{3}$$

Notice that by definition, $cT_\tau - S_\tau < -u < 0$ for any $u > 0$, ensuring that the estimator has a variance bounded by one (see Vázquez-Abad (1999) for the details).

2.2 Storage Process

There is a dual relationship between the surplus process (1) and the queueing, or storage process defined by:

$$X(t) = \sum_{i=1}^{N(t)} Y_i - ct,$$

for $t \le t_1 = \inf(t : X(t) < 0)$, which defines a stopping time t_1. $X(t)$ is extended by setting $X(\tau + t) = \sum_{i=N(\tau)}^{N(\tau+t)} Y_i - ct$ for $t \le t_1 + t_2 = \min(t > t_1 : X(t) < 0)$, and so on. This process describes the workload dynamics of a single server $M/G/1$ queue with a rate of service of c units per unit time, and the sequence $t_1, t_2, \dots$ are regeneration points (see Taylor and Karlin (1994)). Changing the time units to make the server rate unity, the incoming rate becomes λ/c and the mean service time β, so that if $c > \lambda\beta$ then $\lambda/c < \beta^{-1}$ and the queue is stable, so that cycles are finite a.s., which implies that the limit: $\rho(u) = \lim_{t\to\infty} P\{X(t) > u\}$ exists and is unique. Dufresne and Gerber (1989) and Asmussen (1985) show that $1 - \psi(u, \lambda)$ and $1 - \rho(u)$ satisfy the same integro-differential renewal equation for general claim amount distribution G, namely:

$$c\frac{\partial}{\partial u}\varphi(u) = \lambda\varphi(u) - \lambda \int_{-\infty}^{u} \varphi(u)\, dG(y) \tag{4}$$

with the same condition $\lim_{u\to\infty} \varphi(u) = 1$, which etablishes $\psi(u, \lambda) = \rho(u)$. Call D_n the amount of time that $X(\cdot)$ spends above the threshold u within the interval $[T_{n-1}, T_n)$. Then it follows from the general Queueing Theory (see Kleinrock (1976), Taylor and Karlin (1994)) that

$$\psi(u,\lambda) = \lim_{t\to\infty} \frac{1}{t} \sum_{n=1}^{N(t)} D_n, \quad \text{w.p.1} \tag{5}$$

2.3 Convolution Formula

We shall briefly review the arguments used in actuarial textbooks such as Bowers et al. (1986). Let $t_1 = \min(t : U(t) < U(0))$ be the first moment that there is a loss, and let L_1 be the amount of the loss. Similarly, let $t_2 = \min(t > t_1 : U(t) < U(t_1))$ and $L_2 = U(t_1) - U(t_2)$, etc. Given any surplus level $U(t) = u$, the probability that the process will ever fall bellow u is independent of u and equals $\psi(0, \lambda)$, which follows from the memoryless properties of Poisson processes and the assumption of i.i.d. claim amounts. Therefore, the total number N of consecutive losses $L_1, L_2, \dots, L_N$ has a Geometric distribution with parameter $1 - \psi(0, \lambda)$. The quantity $\psi(0, \lambda)$ can be calculated using the Initial Value Theorem for Laplace transforms: $\lim_{s \to \infty} s \int e^{-su} \varphi(u)\, du = \varphi(0)$ applied to (4), yielding $\lambda\beta/c$. Next, since claim amounts are i.i.d, and the residual times are exponential, $L_1, \dots, L_N$ have the same distribution as a sequence of independent random variables with common distribution, which can be calculated as:

$$P[L_i < x] = F(x) = \frac{1}{\beta} \int_0^x (1 - G(x))\, dx$$

Finally, the *maximal aggregate loss* is defined as $L = \max(S(t) - ct; t > 0)$. On one hand, $\psi(u, \lambda) = P(L > u)$, and on the other hand, $L = L_1 + \ldots + L_N$, where N is a Geometric random variable with parameter $1 - \psi(0, \lambda)$, and $L_1, \dots, L_N$ are i.i.d. with distribution F. Therefore:

$$\psi(u, \lambda) = E\left\{ \mathbf{1}_{\{\sum_{i=1}^{N} L_i > u\}} \right\}, \quad N \sim \text{Geom}\left(1 - \frac{\lambda\beta}{c}\right), \quad L_i \sim F \tag{6}$$

and L_i are independent, and independent of N.

3 Estimating the Sensitivities

Several approaches have been proposed to estimate gradients using only one simulated trajectory of the system, such as the methods described in Brémaud and Vázquez-Abad (1992), Glasserman (1991), Fu and Hu (1992), Glynn (1987), Pflug (1990), Reiman and Weiss (1989), among others.

In the simple context of estimating the derivative of $E[X]$ we can write:

$$\frac{\partial E[X]}{\partial \lambda} = \frac{\partial}{\partial \lambda} \int x f(x, \lambda)\, dx = \frac{\partial}{\partial \lambda} \int_0^1 X(u, \lambda)\, du$$

where the random variable X has density f depending on λ and, under an alternative representation, X is a function of a uniform random variate. This latter representation is often used in the context of simulation. Infinitesimal Perturbation Analysis (IPA) uses the stochastic derivative $X'(u, \lambda)$ as estimator and it is valid only when the interchange between the derivative and the expectation is possible. It is almost universally accepted as the best estimation whenever applicable.

An alternative estimator can be obtained under a change of measure approach, when $f'(x,\lambda) = S(x,\lambda)f(x,\lambda)$, so that the product $X\,S(X,\lambda)$ is an unbiased estimator of the derivative:

$$\frac{\partial E[X]}{\partial \lambda} = \frac{\partial}{\partial \lambda}\int x f(x,\lambda)\,dx = \int x S(x,\lambda) f(x,\lambda)\,dx.$$

The function S is called the Score Function and the method is somethimes known as the "likelihood-ratio" method, as was called by Reiman and Weiss (1989). In many problems, however, this method suffers from large variance and much of the recent research effort related to this method has focused on variance reduction.

Fu and Hu (1992) introduced a conditioning argument to smooth out discontinuities of $X(u,\lambda)$ and still apply IPA, by writing $E[X] = E[E[X|Z]]$, where Z is conveniently chosen so that $E[X|Z]$ is no longer discontinuous. This method is called Smoothed Perturbation Analysis (SPA). Pflug (1990) writes $f'(x,\lambda)$ in terms of the (unique) weighted difference of two densities, which is possible through the use of the Hahn Banach decomposition Theorem for signed measures, and the method is accordingly called the Weak Derivative (WD) method. For the particular case when λ is the rate of a Poisson process, both SPA and WD give the same estimator as the Rare Perturbation Analysis (RPA) method introduced by Brémaud and Vázquez-Abad (1992).

3.1 RPA with Importance Sampling

The main difficulty in evaluating the derivative of (3) is that the the process $N(t)$ is a.s. discontinuous in λ, and τ is piecewise constant, so the stochastic derivative of (3) is zero a.s., which prvents the application of IPA. However, the phantom RPA estimators of Brémaud and Vázquez-Abad (1992) can be used. First, for the process with rate $\lambda - \Delta\lambda$, consider the change of measure where claims arrive according to a Poisson process with rate $\lambda_0 - \Delta\lambda$ and their amount has density $\tilde{g}(y) = e^{Ry}g(y)/M_Y(R)$, with the original adjustment coefficient satisfying (2), and $\lambda_0 = \lambda + cR$. The condition for certain ruin is $(\lambda_0 - \Delta\lambda)\tilde{E}[Y] > c$, so by continuity, small enough values of $\Delta\lambda$ will also imply $\tilde{P}_{\lambda_0-\Delta\lambda}[\tau < \infty] = 1$ for this change of measure. Then we have:

$$\begin{aligned}\psi_{\Delta\lambda}(u,\lambda) &= \frac{\psi(u,\lambda) - \psi(u,\lambda-\Delta\lambda)}{\Delta\lambda}\\ &= \frac{1}{\Delta\lambda}\left(\tilde{E}_{\lambda_0}\left[e^{R[cT_\tau - S_\tau]}\right] - \tilde{E}_{\lambda_0-\Delta\lambda}\left[e^{R[cT_\tau - S_\tau]}\right]\right)\end{aligned}$$

where we have explicitly indicated the rate of the corresponding Poisson process in the subscript of the expectation. Claim amounts follow the same distribution for the *nominal* process with accident rate λ_0 as for the *perturbed* process with accident rate $\lambda_0 - \Delta\lambda$.

Consider the Poisson process $N(t)$ with rate λ_0 and define $\{\eta_i\}$ as a sequence of i.i.d. Bernoulli variables with parameter $1-\lambda_0/\Delta\lambda$. Given a trajectory $\{(W_i, Y_i), i=1,2,\ldots\}$ of the *nominal* process, a *phantom* system is associated with a particular sequence η and is defined as the surplus process obtained with interarrival times and claim amounts $\{(W_i, \eta_i Y_i), i=1,2,\ldots\}$. That is, if $\eta_i=0$, the i-th claim is disregarded, as if it had disappeared. We call such claims the *phantom* claims. The phantom surplus process has exactly the desired distribution for the perturbed process. Now $\tau(\eta)$ becomes a function of the sequence $\eta=\{\eta_i, i\leq\tau(0)\}$, and we denote by $\tau(0)$ the nominal stopping time where no claims are disregarded. Conditioning on the sets of m phantoms, we obtain for the finite difference:

$$\begin{aligned}\psi_{\Delta\lambda}(\lambda_0) &= \frac{1}{\Delta\lambda}\tilde{E}_{\lambda_0}\left\{\sum_{m=0}^{\tau(0)}\binom{\tau(0)}{m}\left(\frac{\Delta\lambda}{\lambda_0}\right)^m\left(1-\frac{\Delta\lambda}{\lambda_0}\right)^{\tau(0)-m}\times\right.\\ &\qquad\left. E_m\left\{e^{R(cT_{\tau(0)}-S_{\tau(0)})}-e^{R(cT_{\tau(\eta)}-S_{\tau(\eta)})}\right\}\right\}\end{aligned}$$

where E_m is the expectation given m phantoms, and given the nominal trajectory. For $m=0$ the two processes coincide and the corresponding term in the finite difference is zero. It can be shown (see Brémaud and Vázquez-Abad (1992), Vázquez-Abad and Davis (1995)) that

$$\sum_{m\geq 2}\binom{\tau(0)}{m}\left(\frac{\Delta\lambda}{\lambda_0}\right)^m\left(1-\frac{\Delta\lambda}{\lambda_0}\right)^{\tau(0)-m}\leq\frac{1}{2}\tau(0)^2\left(\frac{\Delta\lambda}{\lambda_0}\right)^2$$

so the contribution of the sum for $m\geq 2$ to the finite difference will be $\mathcal{O}(\Delta\lambda)$ provided that $E[\tau(0)^2]<\infty$. In this case, we obtain the limit:

$$\frac{\partial}{\partial\lambda}\psi(u,\lambda)=\frac{1}{\lambda_0}E_{\lambda_0}\left\{\sum_{j=1}^{\tau(0)}\left({K_0}^{\tau(0)}e^{\mu[cT(0)-S(0)]}-{K_0}^{\tau(j)}e^{\mu[cT(j)-S(j)]}\right)\right\}\tag{7}$$

where we use equiprobability of the $\tau(0)$ claims to be chosen as the phantom one, given that there is only one phantom claim, and $\tau(j), T(j), S(j)$ refer to the corresponding quantities when the j-th claim is disregarded.

In Vázquez-Abad (1999) the Virtual RPA formula is obtained using $\lambda+\Delta\lambda$ with $\Delta\lambda>0$, noticing that if instead of λ_0 we now use $\lambda_0+\Delta\lambda$ with $\Delta\lambda>0$, the perturbed process behaves as if an additional Poisson stream of rare claims (called the virtual claims) was added to the nominal process.

3.2 IPA for the Storage Process

Consider the embedded process evaluated at the times $T_n, n=1,2,\ldots$, that satisfies $X_{n+1}=\max(Y_{n+1}, X_n-cW_{n+1}+Y_{n+1})$, where

$$W_n=-\frac{1}{\lambda}\log(1-U_n),\quad U_n\sim U[0,1]\tag{8}$$

are i.i.d. exponential with parameter λ and $\{Y_i\}$ are i.i.d. with distribution G, independent of λ. According to (5), we are interested in the adapted process D_n that counts the time spent above level u within $[T_n, T_{n+1})$, that is:

$$D_n = \begin{cases} 0 & \text{if } X_n \leq u \\ \min(W_{n+1}, r_n) & \text{otherwise} \end{cases}$$

where $r_n = (X_n - u)/c$ is the time for the process to reach the level u if no claims were to arrive. Consider now the same trajectory $\{(U_i, Y_i); i = 1, 2, \ldots\}$ but now use $\lambda - \Delta\lambda$ instead of λ in (8). The ensuing process is called the *perturbed* process. Let $\mathcal{F}_n$ be the σ-algebra generated by $\{Y_1, (U_{i+1}, Y_i); 0 \leq i \leq n\}$. Notice first that such a change will not propagate from one regenerative cycle to the next, which follows from the observation that the ensuing embedded process is always below the nominal. When a cycle starts, both nominal and perturbed processes reach the same level Y_n and their evolution will therefore be identical–in distribution, to the first cycle (even though the time continuous process has an infinitesimal time shift). We thus concentrate on evaluating the derivatives within one cycle, using the *regenerative formula* (see Taylor and Karlin 1994) for (5):

$$\psi(u, \lambda) = \frac{E\left[\sum_{n=1}^{\tau} D_n\right]}{ET_\tau} = \lim_{t\to\infty} \frac{1}{t} \sum_{n=1}^{N(t)} D_n, \quad \text{w.p.1} \tag{9}$$

The stochastic derivative of the amount D_n is given by:

$$\frac{\partial}{\partial\lambda} D_n = \begin{cases} 0 & \text{if } X_n \leq u \\ -\dfrac{W_{n+1}}{\lambda} & \text{if } X_n > u \text{ and } W_{n+1} < r_n \\ \dfrac{\partial}{\partial\lambda} r_n = \dfrac{1}{\lambda} \displaystyle\sum_{i=1}^{n} W_i & \text{if } X_n > u \text{ and } W_{n+1} > r_n \end{cases} \tag{10}$$

where the derivative of $\frac{\partial}{\partial\lambda} r_n = \frac{1}{c} \frac{\partial}{\partial\lambda} X_n$ has been calculated recursively via the process X_n as follows (recall that we only consider claims within one cycle):

$$\begin{aligned} \frac{\partial}{\partial\lambda} X_1 &= -c\frac{\partial W_1}{\partial\lambda} = \frac{cW_1}{\lambda} \\ \frac{\partial}{\partial\lambda} X_2 &= \frac{\partial}{\partial\lambda} X_1 - c\frac{\partial W_2}{\partial\lambda} = \frac{c(W_1 + W_2)}{\lambda} \\ \frac{\partial}{\partial\lambda} X_n &= \frac{c}{\lambda} \sum_{i=1}^{n} W_i \end{aligned}$$

Interchange between the derivative and the expectation for $\sum_n D_n$ in the numerator of (9) is justified if the possible discontinuities or "critical events" have probabilities bounded by $\Delta\lambda$ and their contribution to the finite difference of the estimator $\sum_n D_n$ has uniformly bounded expectation (in $n, \Delta\lambda$). There are two types of such changes, which we shall call level changes and residual changes, explained presently. In the remainder of the discussion, call $\tilde{X}_n$ the perturbed process when $\lambda - \Delta\lambda$ is used instead of λ, and $\delta_n = \tilde{W}_n - W_n = K(\Delta\lambda) \log(U_n)$.

Level changes: Since $\tilde{X}_n \leq X_n$ a.s, if $X_n \leq u$ then $\tilde{X}_n \leq u$ and $\tilde{D}_n = D_n$. If $X_n > u$ but $\tilde{X}_n \leq u$, then their difference is bounded by $D_n - \tilde{D}_n = D_n \leq r_n < \delta_1 + \ldots + \delta_n$, since otherwise we would have had $\tilde{X}_n > u$ also (see Figure 1).

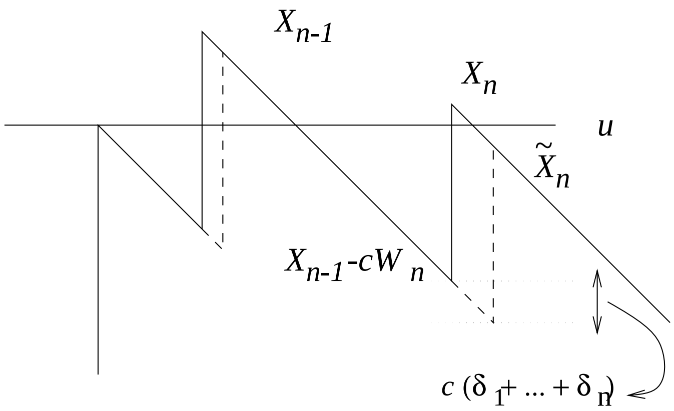

Figure 1: Critical changes in surplus level

Now for any $n \leq \tau$, $(\delta_1 + \ldots + \delta_n) \leq K(\Delta\lambda) \sum_{i=1}^{\tau} \log(U_i) \leq K(\Delta\lambda)\, T_\tau$. In addition, the probability that a level change occurs can be bounded as follows.

$$\begin{aligned} P\{X_n > u, \tilde{X}_n \leq u | \mathcal{F}_{n-1}\} &= P\{Z_{n-1} < Y_n < Z_{n-1} + c(\delta_1 + \ldots + \delta_n)\} \\ &\leq B(\delta_1 + \ldots + \delta_n) \end{aligned}$$

where $Z_{n-1} = u - (X_{n-1} - cW_n) \in \mathcal{F}_{n-1}$ and we assume that the density of Y_1 is bounded by $B < \infty$. Calling E_1 the set of trajectories for which $\{\exists n \leq \tau : X_n > u, \tilde{X}_n \leq u\}$, then by the foregoing,

$$E\left\{\sum_{n=1}^{\tau} \frac{D_n - \tilde{D}_n}{\Delta\lambda} | E_1\right\} \leq E\left\{\frac{\tau B(\delta_1 + \ldots + \delta_\tau)^2}{\Delta\lambda}\right\} \leq E\left\{\frac{\tau B K^2(\Delta\lambda) T_\tau^2}{\Delta\lambda}\right\}$$

which will converge to 0 as $\Delta\lambda \to 0$ whenever $E[\tau T_\tau^2] < \infty$.

Residual changes: Given $\tilde{X}_n > u$, the second type of critical event, denoted by $E_2(n)$, is when $W_{n+1} < r_n$ but $\tilde{W}_{n+1} > \tilde{r}_n$, as depicted in Figure 2. We have $D_n - \tilde{D}_n = W_{n+1} - \tilde{r}_n < r_n - \tilde{r}_n = \delta_1 + \ldots + \delta_n$. In this case we have:

$$\begin{aligned} P[E_2(n) | \mathcal{F}_{n-1}] &= E\left[P\left(W_{n+1} < r_n \text{ and } c\frac{\lambda}{\lambda - \Delta\lambda} W_{n+1} > \tilde{r}_n | r_n, \mathcal{F}_{n-1}\right)\right] \\ &\leq E\left[P\left(r_n\left(1 - \frac{\Delta\lambda}{\lambda}\right) \leq W_{n+1} < r_n | r_n, \mathcal{F}_{n-1}\right)\right] \\ &= E\left[e^{-\lambda r_n}(e^{\Delta\lambda r_n} - 1) | \mathcal{F}_{n-1}\right] \end{aligned}$$

where we have used $\tilde{r}_n \leq r_n, \tilde{W}_n = \lambda/(\lambda - \Delta\lambda) W_n$, and $W_{n+1} \sim \exp(\lambda)$ to evaluate the conditional expectation.

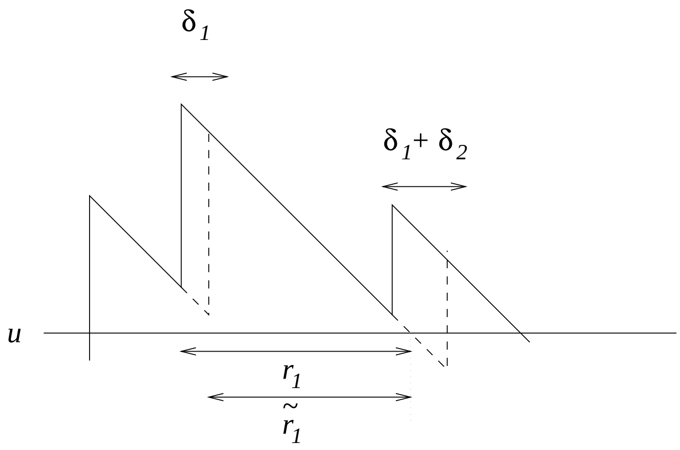

Figure 2: Critical changes in residual times

Finally, as $\Delta\lambda \to 0$, we obtain

$$\frac{1}{\Delta\lambda} P(E_2(n)|\mathcal{F}_{n-1}) \to E\left(r_n e^{-\lambda r_n}|\mathcal{F}_{n-1}\right) \leq T_\tau$$

Call E_2 the event that $E_2(n)$ occurs for at least one $n = 1, \ldots, \tau$. Then

$$E\left\{\sum_{n=1}^{\tau} \frac{D_n - \tilde{D}_n}{\Delta\lambda}|E_2\right\} \leq E\left\{\tau T_\tau(\delta_1 + \ldots + \delta_\tau)\right\} \leq E\left\{\tau K(\Delta\lambda) T_\tau^2\right\}$$

which also tends to zero if $E[\tau T_\tau^2] < \infty$. These results establish:

$$\frac{\partial}{\partial\lambda} E\left[\sum_{n=1}^{\tau} D_n\right] = E\left[\sum_{n=1}^{\tau} \frac{\partial}{\partial\lambda} D_n\right]$$

using (10) for the stochastic derivatives. From (9), it follows that:

$$\frac{\partial}{\partial\lambda}\psi(u,\lambda) = \frac{1}{ET_\tau}\left\{\frac{\partial}{\partial\lambda} E\left[\sum_{n=1}^{\tau} D_n\right] - \frac{E\left[\sum_{n=1}^{\tau} D_n\right]}{ET_\tau}\frac{\partial ET_\tau}{\partial\lambda}\right\}$$

and the IPA derivative of the cycle length (which is unbiased due to absolute continuity) can be obtained directly from the expression $T_\tau = \sum_{n=1}^{\tau} W_n$, yielding $-T_\tau/\lambda$ as the derivative estimator. For the estimation, a total of M cycles of the simulated process are used to build the sample averages (which are consistent estimators of the individual quantities), that is:

$$\begin{aligned}
\psi(u,\lambda) &= \lim_{M\to\infty} \frac{\sum_{m=1}^{M}\sum_{n=1}^{\tau_m} D_n^{(m)}}{\sum_{m=1}^{M} T_{\tau_m}} \equiv \lim_{M\to\infty} \hat{\psi}(M) \\
ET_\tau &= \lim_{M\to\infty} \frac{1}{M}\sum_{m=1}^{M} T_{\tau_m} \equiv \lim_{M\to\infty} \hat{T}_\tau(M) \\
\frac{\partial}{\partial\lambda} E\left[\sum_{n=1}^{\tau} D_n\right] &= \lim_{M\to\infty} \frac{1}{\lambda M}\sum_{m=1}^{M}\sum_{n=1}^{\tau_m} \frac{\partial}{\partial\lambda} D_n \equiv \lim_{M\to\infty} \hat{D}(M)
\end{aligned}$$

where (10) is used for estimating $\hat{D}_M$ and the limits are w.p.1. Finally, this yields:

$$\frac{\partial}{\partial\lambda}\psi(u,\lambda) = \lim_{M\to\infty}\frac{\hat{D}(M)}{\hat{T}_\tau(M)} + \frac{\hat{\psi}(M)}{\lambda} \tag{11}$$

which is estimated over M cycles. As explained in Bratley, Fox, and Schrage (1987) and in LeQuoc (1998), regeneratve simulation also allows the estimation of appropriate confidence intervals.

3.3 SF for the Convolution Formula

For (6) IPA does not work because the pathwise derivative of $N \sim \text{Geom}(1-\lambda\beta/c)$ is zero, and the path dependency on λ is present only in N. We present here the Score Function method. Let $p_\lambda(n) = P[N = n] = \left(1-\frac{\lambda\beta}{c}\right)\left(\frac{\lambda\beta}{c}\right)^n$. Then the usual change of measure argument yields

$$E_\lambda[h(N)] = \sum_{n=1}^{\infty} h(n)\frac{p_\lambda(n)}{p_{\lambda_0}(n)}p_{\lambda_0}(n) = E_{\lambda_0}\left[\frac{p_\lambda(N)}{p_{\lambda_0}(N)}h(N)\right].$$

Let $h(N) = \mathbf{1}_{\{\sum_{i=1}^N L_i > u\}}$, where $L_i \sim F$ as in (6). Taking derivatives, we have:

$$\frac{\partial}{\partial\lambda}\psi(u,\lambda) = E_{\lambda_0}\left[\boldsymbol{L}(N)S(N)\mathbf{1}_{\{\sum_{i=1}^N L_i > u\}}\right]$$

where

$$\begin{aligned}
\boldsymbol{L}(N) &= \frac{p_\lambda(N)}{p_{\lambda_0}(N)} = \frac{\left(1-\frac{\lambda\beta}{c}\right)\left(\frac{\lambda\beta}{c}\right)^N}{\left(1-\frac{\lambda_0\beta}{c}\right)\left(\frac{\lambda_0\beta}{c}\right)^N}, \\
S(N) &= \frac{\partial}{\partial\lambda}\log\boldsymbol{L}(N) = \beta\left(\frac{N}{\lambda\beta} - \frac{1}{c-\lambda\beta}\right)
\end{aligned}$$

Since $S(N)$ is independent of λ_0, we can use again the change of measure argument to obtain:

$$\frac{\partial}{\partial\lambda}\psi(u,\lambda) = E_\lambda\left[S(N)\mathbf{1}_{\{\sum_{i=1}^N L_i > u\}}\right] \tag{12}$$

4 Simulation Experiments

Our simulations were all performed in Ultra Sparc Stations using SIMOD, a simulation package by L'Ecuyer (1993) which compiles in MODULA2. We used $\lambda = 1, c = 1.5$ and experimented with three different claim amount distributions. For the surplus process ((3) and (7)) as well as for the convolution formula ((6) and (12)), we estimate the confidence intervals by the Central Limit Theorem using independent replications. For the storage process ((5) and (11)) we implement the regenerative method described in Bratley, Fox, and Schrage (1987) and Ross (1997) to estimate the confidence intervals. The efficiency of the estimators is defined as the inverse of the product of the mean CPU time and the variance of the estimators, both quantities being estimated with several replications of the simulation runs. We call $\psi'(u,\lambda) = \frac{\partial}{\partial\lambda}\psi(u,\lambda)$. Details of most of the algorithms are in LeQuoc (1998).

4.1 Exponential Claims

In this section $\beta = E[Y] = 1.0$. The model for exponential claims can be solved analytically, which gives us a basis for comparison with the true value of both $\psi(u,\lambda)$ and $\psi'(u,\lambda)$. Here $R = \beta - \lambda/c$ and

$$\psi(u,\lambda) = \frac{\lambda\beta}{c}\exp\left\{\left(\frac{\lambda\beta}{c}-1\right)\frac{u}{\beta}\right\}, \quad u \geq 0.$$

So that the derivative w.r.t. λ is:

$$\psi'(u,\lambda) = \beta\,(c+\lambda u)\exp\left\{\left(\frac{\lambda\beta}{c}-1\right)\frac{u}{\beta}\right\}$$

Ruin Probability $\psi(u,\lambda)$				
	$u=0$	$u=2$	$u=4$	$u=6$
$\psi(u,\lambda)$	0.66667	0.34228	0.17573	0.09022
IS	0.675	0.340	0.175	0.090
IC (90%)	(0.669,0.682)	(0.337,0.344)	(0.173,0.177)	(0.090,0.091)
SP	0.671	0.349	0.182	0.095
IC (90%)	(0.665,0.676)	(0.341,0.357)	(0.174,0.189)	(0.088,0.101)
CF	0.665	0.339	0.176	0.091
IC (90%)	(0.661,0.668)	(0.336,0.343)	(0.173,0.179)	(0.089,0.093)
Derivative $\psi^{'}(u,\lambda)$				
	$u=0$	$u=2$	$u=4$	$u=6$
$\psi^{'}(u,\lambda)$	0.66667	0.79864	0.64435	0.45112
IS+RPA	0.686	0.778	0.636	0.443
IC (90%)	(0.642,0.731)	(0.738,0.819)	(0.608,0.664)	(0.426,0.461)
SP+IPA	0.671	0.802	0.669	0.484
IC (90%)	(0.665,0.676)	(0.776,0.829)	(0.628,0.709)	(0.447,0.520)
CF+SF	0.668	0.802	0.651	0.460
IC (90%)	(0.653,0.683)	(0.788,0.816)	(0.638,0.665)	(0.447,0.473)

Table 1: Exponential Claim Distribution

Importance Sampling: The change of measure in (3) and (7) for this example prescribes simulating the surplus process with Poisson arrivals at rate $\lambda_0 = c/\beta$ and claim amounts $Y \sim \exp(\lambda/c)$. We performed 3,500 independent replications of a cycle.

Storage Process: There is no set-up required for the implementation of equations (5) and (11). We used 30,000 cycles for the estimation.

Convolution Formula: According to equations (6) and (12), we first generate a geometric random variable N with parameter $(1-\lambda\beta/c)$ and then generate $L_1, \ldots, L_N$ as independent random variables with exponential distribution with mean β. 50,000 independent replications were performed.

Table 1 shows the results for IS (3), SP (5) and CF (6) for the ruin probability and RPA (7), IPA (11) and SF (12) for the derivatives. We show also the confidence intervals (IC). As u increases the ruin probability and its derivative become harder to evaluate, since the event of ruin becomes rarer and rarer. All confidence intervals in Table 1 are similar, CF and CF-SF being the most accurate ones, particularly for the derivative ψ'. The computaional effort, however, is very different for every method. For IS, cycles end when ruin occurs, and the initial surplus level is u: as this increases, the cycle length (and thus the CPU time) increase. In addition, the bookkeeping required for the phantom systems (which are longer than the nominal) will also hinder performance. In contrast, both SP and CF share the property that the simulation lengths do not depend on u, however, for each regenerative cycle (or iteration) the probability of overflow being computed may yield many vanishing samples, entailing worse variance performance. To adequately reflect the compromise between precision and length of the simulations, Table 2 compares the estimated efficiencies. Clearly for this case the benefit from IS which is evident in estimating the ruin probabilities, is outweighed by the CPU time required for the derivative estimation and overall CF+SF seems to be more robust.

Ruin Probability				
	$u=0$	$u=2$	$u=4$	$u=6$
IS	8.305×10^4	18.907×10^4	52.221×10^4	167.346×10^4
SP	2.094×10^4	0.906×10^4	0.956×10^4	1.347×10^4
CF	12.059×10^4	11.802×10^4	18.253×10^4	31.693×10^4
Derivative				
	$u=0$	$u=2$	$u=4$	$u=6$
IS+RPA	0.178×10^3	1.281×10^3	1.9861×10^3	4.256×10^3
SP+IPA	20.938×10^3	0.818×10^3	0.351×10^3	0.420×10^3
CF+SF	14.336×10^3	15.946×10^3	17.116×10^3	19.999×10^3

Table 2: Efficiency for Exponential Claim Distribution

4.2 Weibull Claims

The general form of this distribution is $g(y) = ba^b y^{b-1} e^{-(ay)^b}$ for $y > 0$. With $a = 1/\sqrt{2}$ and $b = 2$, the density of the claims becomes:

$$g(y) = ye^{-\frac{y^2}{2}} \quad y > 0.$$

which implies $E[Y] = \sqrt{\pi/2}$.

Importance Sampling: The change of measure for (3) and (7) requires evaluating the adjustment coefficient with the implicit equation (2), with:

$$M_Y(r) = \int_0^\infty ye^{-\frac{y^2}{2}+yr}dy = 1 + \sqrt{2\pi}\,\Phi(-r)\,re^{\frac{r^2}{2}}$$

where $\Phi(\cdot)$ is the standard Normal distribution. Using *Mathematica*, we found $R \approx 0.19683$.

The simulation of the surplus process is performed with Poisson arrivals at rate $\lambda_0 = \lambda + cR$ and claim amounts $Y \sim \tilde{G}$, where

$$\tilde{g}(y) = \frac{1}{1.32141} ye^{-\frac{y^2}{2}+0.214272y}, \quad y > 0,$$

In order to generate these variables, we used a combination of three methods as described in Ross (1997), namely composition, inversion and acceptance/rejection. The density $\tilde{g}$ is shown in Figure 3 along with the dominating piecewise linear function:

$$t(y) = \begin{cases} 0.575486 & 0 \le y < 2 \\ -0.266460y + 0.847347 & 2 \le y < 3 \\ -0.159885y + 0.095931 & 3 \le y \le 6 \end{cases}$$

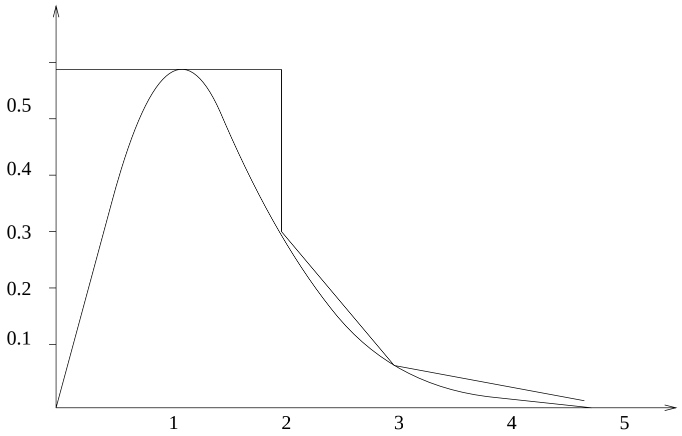

Figure 3: Density for Importance Sampling

Let $C = \int_0^6 t(y)dy = 1.40412$ and define the density $r(y) = t(y)/C = \sum_{j=1}^3 p_j r_j(y)$ which can be interpreted as the composition of three random variables with $p_j = P[Y \in I_j]$, with $I_1 = [0, 2)$, $I_2 = [2, 3)$,

$I_3 = [3,6]$. Choose first an interval according to p_j, and generate X from the conditional distribution r_j using the inversion method, then apply acceptance/rejection by generating an independent uniform variate U and accepting the sample point $Y = X$ only if $U < g(X)/t(X)$. This method accepts $C^{-1} \approx 71\%$ of all the generated values. We performed 3,500 independent replications of a cycle.

Storage Process: There is no set-up required for the implementation of equations (5) and (11). We used 30,000 cycles for the estimation.

Convolution Formula: According to equations (6) and (12), we must generate a geometric random variable N with parameter $(1-\lambda\beta/c)$ and then generate $L_1, \ldots, L_N$ as independent random variables with distribution:

$$F(y) = \frac{2}{\sqrt{2\pi}} \int_0^x e^{\frac{-y^2}{2}} dy = 2[\Phi(y) - 0.5]$$

which can be achieved most efficiently by noticing that this is the distribution of a standard normal random variable. 50,000 independent replications were performed.

Table 3 shows the results for IS (3), SP (5) and CF (6) for the ruin probability and RPA (7), IPA (11) and SF (12) for the derivatives. We show also the confidence intervals (IC).

Ruin Probability $\psi(u,\lambda)$				
	$u=0$	$u=2$	$u=4$	$u=6$
IS	0.838	0.569	0.370	0.241
IC (90%)	(0.835,0.841)	(0.568,0.571)	(0.369,0.371)	(0.240,0.242)
SP	0.836	0.567	0.368	0.240
IC (90%)	(0.832,0.839)	(0.559,0.575)	(0.358,0.378)	(0.230,0.251)
CF	0.834	0.564	0.367	0.241
IC (90%)	(0.830,0.838)	(0.559,0.569)	(0.361,0.373)	(0.236,0.246)
Derivative $\psi'(u,\lambda)$				
	$u=0$	$u=2$	$u=4$	$u=6$
IS+RPA	0.894	1.728	1.935	1.863
IC (90%)	(0.795,0.993)	(1.637,1.818)	(1.850,2.020)	(1.796,1.929)
SP+IPA	0.836	1.714	1.934	1.792
IC (90%)	(0.832,0.839)	(1.660,1.767)	(1.850,2.019)	(1.693,1.893)
CF+SF	0.799	1.669	1.930	1.798
IC (90%)	(0.740,0.857)	(1.635,1.740)	(1.880,1.980)	(1.748,1.848)

Table 3: Weibull Claim Distribution

In these experiments the theoretical value was not available, as for the exponential claim distribution.

All CI's however show very consistent results in the estimation. Once again IS seems to perform well especially for increasing u, which is when IS is better suited to yield performance improvement. Nontheless, the effort required to generate the claims according to the new change of measure outweighs the benefits for moderate values of u, as shown in Table 4. Because for this example it was straightforward to generate the ladder height distribution $F(y)$, once again consideration of the computational effort makes CF+SF a very robust estimator, whose efficiency seems constant for u.

Ruin Probability				
	$u = 0$	$u = 2$	$u = 4$	$u = 6$
IS	15.976×10^4	23.460×10^4	35.814×10^4	63.878×10^4
SP	1.883×10^4	0.379×10^4	0.248×10^4	0.239×10^4
CF	11.112×10^4	5.978×10^4	6.238×10^4	7.868×10^4
Derivative				
	$u = 0$	$u = 2$	$u = 4$	$u = 6$
IS+RPA	1.441×10^2	0.836×10^2	0.641×10^2	0.797×10^2
SP+IPA	188.296×10^2	0.857×10^2	0.348×10^2	0.247×10^2
CF+SF	6.035×10^2	7.3185×10^2	7.8615×10^2	7.847×10^2

Table 4: Efficiency for Weibull Claim Distribution

4.3 Pareto Claims

Our last experiments use Pareto distributed claim amounts. Historical data seems to agree rather well with the assumption of a Pareto distribution, which makes this model important in practice. We use:

$$g(y) = \frac{a}{(1+y)^{a+1}} \quad y > 0, a > 0.$$

with $a = 2$, so that $E[Y] = 1$ and $\mathrm{Var}[Y] = \infty$. The moment generating function does not exist. In order to generate the claims we use the inversion method, with $Y = (1-U)^{-1/a} - 1$.

Importance Sampling: We can no longer apply the Lundberg change of measure, since $M_Y(r)$ does not exist. However, we can still apply Importance Sampling using Poisson arrivals with rate $\lambda_0 > c/\beta$ and same claim distribution as for the original process. In this case, (3) has to be modified using:

$$\psi(u,\lambda) = E_{\lambda_0}\left(\frac{\lambda}{\lambda_0}\right)^{\tau} \exp\{(\lambda_0 - \lambda)T_\tau\}$$

and (7) becomes:

$$\frac{\partial}{\partial\lambda}\psi(u,\lambda) = E_{\lambda_0}\left\{\frac{\tau(0)}{\lambda}\left(\frac{\lambda}{\lambda_0}\right)^{\tau(0)} e^{(\lambda_0-\lambda)T_{\tau(0)}} - \frac{1}{\lambda_0}\sum_{j=1}^{\tau(0)}\left(\frac{\lambda}{\lambda_0}\right)^{\tau(j)} e^{(\lambda_0-\lambda)T_{\tau(j)}}\right\}$$

which is not necessarily a good estimator (as we shall verify from the experiments), since T_τ no longer has a moment generating function either.

Storage Process: There is no set-up required for the implementation of equations (5) and (11). Recall, however, that to justify interchange of derivative and expectation we used as a sufficient condition $E[\tau T_\tau^2] < \infty$, which is no longer true. We used 30,000 cycles for the estimation.

Convolution Formula: According to equations (6) and (12), we must generate a geometric random variable N with parameter $(1-\lambda\beta/c)$ and then generate $L_1, \ldots, L_N$ as independent random variables with distribution:

$$F(y) = (a-1)\int_0^x \frac{1}{(1+y)^a}dy = \int_0^x \frac{(a-1)}{(1+y)^a}dy$$

which is the Pareto distribution with parameter $(a-1)$. We used for $a = 2$, $L_i = U_i/(1-U_i)$ with U_i independent uniform variates. While $E[L_i] = \infty$, the Central Limit Theorem can still be applied because our estimators are indicator functions, which are a.s. bounded. We performed 50,000 independent replications.

Ruin Probability $\psi(u,\lambda)$				
	$u=0$	$u=2$	$u=4$	$u=6$
IS	0.634	0.339	0.223	0.217
IC (90%)	(0.582, 0.687)	(0.311, 0.367)	(0.204, 0.242)	(0.149, 0.286)
SP	0.669	0.447	0.348	0.285
IC (90%)	(0.653, 0.685)	(0.422, 0.474)	(0.318, 0.379)	(0.252, 0.318)
CF	0.665	0.446	0.345	0.283
IC (90%)	(0.661, 0.668)	(0.442, 0.449)	(0.342, 0.349)	(0.280, 0.287)
Derivative $\psi'(u,\lambda)$				
	$u=0$	$u=2$	$u=4$	$u=6$
IS+RPA	1.329	1.284	1.039	1.775
IC (90%)	(1.034, 1.625)	(1.113, 1.455)	(0.898, 1.181)	(0.738, 2.812)
SP+IPA	0.669	0.793	0.715	0.657
IC (90%)	(0.653, 0.685)	(0.734, 0.853)	(0.670, 0.759)	(0.601, 0.712)
CF+SF	0.668	0.789	0.756	0.703
IC (90%)	(0.653, 0.683)	(0.775, 0.804)	(0.742, 0.771)	(0.689, 0.717)

Table 5: Pareto Claim Distribution

Table 5 shows the results for IS (3), SP (5) and CF (6) for the ruin probability and RPA (7), IPA (11) and SF (12) for the derivatives, with the confidence intervals (IC). The theoretical values of the ruin probability and its derivative are no longer known (other than for $u = 0$, which is independent of the claim distribution, as shown in de Vylder (1996)). Nonetheless, inspection of the estimation makes it clear that IS is not a good method. We performed extensive simulation results to assess the pathological behaviour (expected from the theoretical requirements of finite moments). In order to obatain a reasonable coverage, extremely long samples would have been required, we show only those with same computational budget as the previous examples, for comparison purposes. Table 6 compares the estimated efficiencies.

Ruin Probability				
	$u = 0$	$u = 2$	$u = 4$	$u = 6$
IS	0.230×10^3	0.373×10^3	0.603×10^3	0.0358×10^3
SP	1.908×10^3	0.719×10^3	0.548×10^3	0.466×10^3
CF	143.785×10^3	123.409×10^3	137.345×10^3	152.912×10^3
Derivative				
	$u = 0$	$u = 2$	$u = 4$	$u = 6$
IS+RPA	0.0733×10^2	0.0982×10^2	0.109×10^2	0.1570×10^2
SP+IPA	19.08×10^2	1.394×10^2	2.565×10^2	1.606×10^2
CF+SF	75.969×10^2	79.124×10^2	83.18×10^2	86.025×10^2

Table 6: Efficiency for Pareto Claim Distribution

The method CF dramaticaly outperforms the other methods inthis experiment, as can be seen from their efficiencies. To understand why, notice that the particular choice of of distribution resulted in a deceivingly simple ladder height distribution $F(y)$ that can be generated at virtually no cost. The combination of these factors: IS no onger applicable, especially to estimate the derivatives, and CF being extremely easy to generate, explain why the results tilt so notably towards CF and CF+SF.

5 Conclusion

Estimating derivatives of rare events is a challenging problem and still unsolved for many models. In this work we have focused on the study of Ruin Probabilities under the canonical model of Risk Theory. While indirect estimation methods can be built, our results show that the performance and applicability of the methods is problem dependent. The Importance Sampling method is excellent whenever it is easy to implement, but it will in general require a considerable set-up time and the ensuing distributions may be hard to generate, which affects performance. For Weibull claims, we optimized the generation of the claims under the new measure by using a hybrid method, but for the Pareto claims the choice of change of measure is very delicate, especially for the derivative. The convolution formula seems to yield overall the best results, but it is also necessary to calculate in advance the distribution F and look for an efficient way of generating it. The main drawback of the method is that, as of today, it is only applicable in the linear case studied here. In Vázquez-Abad (1999) we analyze the model of surplus processes with interest, in which case the method is not applicable, while the other two are. As should be apparent, the Storage Process method is the more flexible and robust, not needing much set-up time. It is also capable of handling empirical distributions for the claim amounts.

In addition to point estimation, both the SP and CF are particularly well suited for functional estimation. The simulations follow a process dynamics which is *independent* of u, and therefore it is straightforward to calculate in parallel various statistics for different predetermined values of u (or adaptively seek new values for function evaluation). This was used in Vázquez-Abad (1999) and could also prove extremely valuable if a surface response is to be estimated: one assumes, for example, that $\psi(\cdot, \lambda)$ (or its derivative) are polynomial of certain degree, piecewise in u. By simulating at various values of u a curve is then fitted to yield the overall estimation of a function, instead of just point values. If IS were to be used for this problem, as many simulations of the cycles have to be performed as values of u are there to evaluate. In contrast, the evaluation of the parallel statistics requires marginal added cost, which could greatly compensate the benefits from IS, if the latter is applicable.

We would like to conlcude with a word of caution: which method outperforms the rest is largely dependent on the problem. Extensions of our research include multidimensional and stratified portfolios, and non-linear dynamics.

Acknowledgements

The authors wish to thank the reviewers for their valuable suggestions. This work was supported in part by NSERC-Canada grant # WFA0184198 and an FCAR team grant. The work was completed while the first author was on leave at the Department of Electrical and Electronic Engineering, Melbourne University.

References:

Asmussen, S. 1985. Conjugate processes and the simulation of ruin problems. *Stoch Proc Appl*, 20:213–229.

Asmussen, S., and H. Nielsen. 1995. Ruin probabilities via local adjustment coefficients. *J. Appl Probab*, 32:736–755.

Bowers, N., H. Gerber., J. Hickman., D. Jones., and C. Nesbitt. 1986. *Actuarial Mathematics.* New York: Society of Actuaries.

Bratley, P., B. Fox., and L. Schrage. 1987. *A Guide to Simulation.* Second ed. New York: Springer-Verlag.

Brémaud, P., and F. Vázquez-Abad. 1992. On the pathwise computation of derivatives with respect to the rate of a point process: The phantom rpa method. *Queueing Systems: Theory and Applications*, 10:249–270.

Chung, K. 1974. *A course in probability theory*. Second ed. New York: Academic Press.

de Vylder, E. 1996. *Advanced Risk Theory*. Éditions de l'Université de Bruxelles.

Dufresne, F., and H. Gerber. 1989. Three methods to calculate the probability of ruin. *ASTIN Bulletin*, 19:79–90.

Fu, M. C., and J. Hu. 1992. Extensions and generalizations of smoothed perturbation analysis in a generalized semi-markov process framework. *IEEE Trans. on Autom. Control*, 37(10):1483–1500.

Gerber, H. 1979. An introduction to risk theory. Monograph 8, University of Pennsylvania, Philadelphia. Huebner Foundation.

Glasserman, P. 1991. *Gradient Estimation via Perturbation Analysis*. Boston: Kluwer Academic.

Glynn, P. 1987. Likelihood ratio gradient estimation: An overview. *Proceedings of the 1987 Winter Simulation Conference*, 336–375.

Kleinrock, L. 1976. *Queueing systems*. New York: Wiley & Sons.

L'Ecuyer, P. 1993. Simod: Définition fonctionnelle et guide d'utilisation. Technical repport 221, Département d'iro, Université de Montréal.

LeQuoc, P. 1998. Estimation de la probabilité de ruine et sa dérivée par simulation. Master's thesis, IRO, Université de Montréal.

Pflug, G. C. 1990. On line optimization of simulated markov processes. *Math. of Operations Research*, 15:381–395.

Reiman, M., and A. Weiss. 1989. Sensitivity analysis for simulations via likelihood ratios. *Operations Research*, 37:830–844.

Ross, S. 1997. *Simulation*. Second ed. Boston: Academic Press.

Taylor, H., and S. Karlin. 1994. *Introduction to Stochastic Processes*. Revised edition ed. New York: Academic Press, Inc.

Vázquez-Abad, F. 1999. Pathwise derivative estimation of ruin probabilities. *Insurance Mathematics and Economics*. in press.

Vázquez-Abad, F., and K. Davis. 1995. Efficient implementation of the phantom rpa method, with an application to a priority queueing system. Technical Report 989, Département d'iro, Université de Montréal, University of Montreal.

Vázquez-Abad, F., and L. Zubieta. 2000. Ersätz estimation for sensitivity analysis: a methodology. To be submitted.