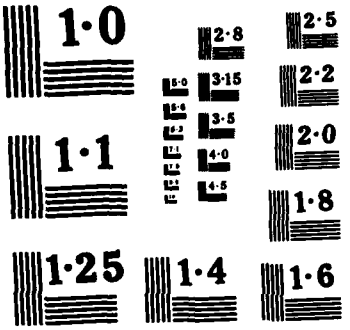
AD-A158 167 APPROXIMATION OF THE INITIAL RESERVE FOR KNOWN RUIN PROBABILITIES(U) WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER E W FREES MAY 85 MRC-TSR-2822 1/1

UNCLASSIFIED DAAG29-80-C-0041 F/G 12/1 NL

END

FILMED

DTIC



NATIONAL BUREAU OF STANDARDS
MICROCOPY RESOLUTION TEST CHART

MRC Technical Summary Report #2822

APPROXIMATION OF THE INITIAL RESERVE FOR KNOWN RUIN PROBABILITIES

Edward W. Frees

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53705

May 1985

(Received April 1, 1985)

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

APPROXIMATION OF THE INITIAL RESERVE FOR KNOWN RUIN PROBABILITIES

Edward W. Frees

Technical Summary Report #2822

May 1985

ABSTRACT

An important problem in the study of actuarial risk theory is approximating the probability of ruin within finite time based on a specified initial reserve. In this paper we address the similar, but mathematically different, problem of how to approximate a desired initial reserve given a pre-specified probability of ruin. Although the procedures have desirable asymptotic properties such as consistency and asymptotic normality, there they are computer-intensive and would not have been practicable before the wide spread availability of high-speed computers. The procedures rely on simulated realizations of a general risk process. Thus, these they can be used in many of the models of risk processes that appear in the literature such as the Compound Poisson, ARMA and Stochastic Discounting models. Examples of several models are given to demonstrate the versatility of the procedure and to demonstrate that the procedures are computationally feasible.

AMS (MOS) Subject Classifications: Primary 62P05, Secondary 62L20, 60K10, 62G30.

Key Words: stochastic approximation; quantile estimation; kernel estimation

Work Unit Number 4 - Statistics and Probability

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

SIGNIFICANCE AND EXPLANATION

The probability of ruin is the probability that claims against a risk-taking enterprise exceed its initial capital or reserve plus income at some point of time. If ruin occurs before a fixed time, such as one or ten years, this is called the finite horizon time problem and otherwise, the infinite horizon time problem. Even in the simplest models, ruin probabilities can be difficult to compute. There is a vast literature on this aspect of the problem dating back at least to the early 1900's.

An important problem in the study of actuarial risk theory is approximating the probability of ruin within finite time based on a specified initial reserve. In this paper we address the similar, but mathematically different, problem of how to approximate a desired initial reserve given a pre-specified probability of ruin. Although the procedures have desirable asymptotic properties such as consistency and asymptotic normality, they are computer-intensive and would not have been practicable before the wide spread availability of high-speed computers. The procedures rely on simulated realizations of a general risk process. Thus, they can be used in many of the models of risk processes that appear in the literature such as the Compound Poisson, ARMA and Stochastic Discounting models. Examples of several models are given to demonstrate the versatility of the procedure and to demonstrate that the procedures are computationally feasible.

APPROXIMATION OF THE INITIAL RESERVE FOR KNOWN RUIN PROBABILITIES

Edward W. Frees

§1. Introduction

Let $U(t)$ be the net liability of a risk-accepting enterprise by time t. An important parameter in risk theory is the probability that $U(t)$ exceeds an initial reserve u at some time t prior to or at T, the horizon time. This probability is denoted by

$$\psi(u) = P(\sup_{0 \le t \le T} U(t) > u) \qquad (1.1)$$

and, when T is finite, is called the finite horizon time probability of ruin. In this paper, we are only concerned with finite T and thus suppress the explicit dependence of $\psi(u)$ on T. The net liability $U(t)$ is a stochastic process indexed by the time parameter t which may be discrete or continuous. Some of the several examples which appear in the literature are given below.

Example 1.1. The Compound Poisson Process.

Let $\{N(t), 0 \le t < \infty\}$ be a Poisson process with intensity $\rho < 1$. Let $X_1, X_2, \ldots$ be i.i.d. claim random variables which are independent of $\{N(t)\}$. Then the number of claims by time t is

$$\sum_{k=1}^{N(t)} X_k \ ,$$

a compound Poisson process. If premiums are assumed to arrive at a known steady rate, say, P per unit time, then the net liability by time t is

$$U_1(t) = \sum_{k=1}^{N(t)} X_k - P\,t \ .$$

Example 1.2. ARMA model.

The introduction of the autoregressive moving average (ARMA) model in the literature on ruin probabilities is due to Gerber (1982). Here, let $\phi_1, \ldots, \phi_p$, $\theta_1, \ldots, \theta_q$ be parameters, $\{\varepsilon_i\}_{i=1}^{\infty}$ be i.i.d. r.v's, and recursively define the gain during the t^{th}

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

year by

$$G_t = \phi_1 G_{t-1} + \ldots + \phi_p G_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \ldots + \theta_q \varepsilon_{t-q} \quad .$$

To complete the specification of G_t, assume that $G_j = g_j$, $j = 0,\ldots,-p+1$, and $\varepsilon_k = e_k$, $k = 0,\ldots,-q+1$, where $\{g_j\}$ and $\{e_k\}$ are known constants. Two assumptions, somewhat unusual in the time-series ARMA literature, made by Gerber are that the noise r.v's $\{\varepsilon_i\}$ are bounded and have positive mean. In this model the net liability is

$$U_2(t) = -\sum_{j=1}^{t} G_t \quad .$$

Note that in this model, the net liability is indexed by a discrete time parameter.

Example 1.3. Stochastic Discounting model.

Let $\{C_t\}$ and $\{D_t\}$ be sequences of r.v.'s where C_t is used for the net cash flow during the t^{th} time period and D_t is the discounting factor from the t^{th} time period to an arbitrary initial time point. Under mild conditions on the random variables, we can define the prospective reserve at time t

$$V_t = D_t^{-1} E_t\left(\sum_{k=t+1}^{T} C_k D_k\right) \quad ,$$

where E_t is the conditional expectation at time t. More formally, we could define a nondecreasing sequence of sub sigma-fields $\{F_t\}_{t=1}^{T}$ such that C_t and D_t are F_t-measurable and use $E_t(\cdot) = E(\cdot \mid F_t)$. In this model, the net liability is the sum of discounted net cash flow by time t plus the prospective reserve, i.e.,

$$U_3(t) = \sum_{j=1}^{t} D_j C_j + V_t \quad .$$

Development of this discrete time parameter model and some of its properties is due to Papatriandafylou and Waters (1984). See Bühlmann (1976) and Gerber (1976) for earlier approaches.

Papers appearing in the literature give methods for calculating or approximating $\psi(u)$ for a specified initial reserve u. In this paper we give the theory for methods used to determine u based on a pre-specified ruin probability, say α. In other words, defining

$$\psi^*(u) = \psi(u) - \alpha \quad ,$$

we wish to find the root (or zero) of the monotone function $\psi^*(u)$, say θ, i.e.,

$\psi^*(\theta) = 0$. The problem of root finding is more difficult than simply evaluating the ruin probability function but is also important to the risk manager. A naive approach would be to use an iterative Newton-Raphson technique,

$$\hat{\theta}_{n+1} = \hat{\theta}_n - \psi^*(\hat{\theta}_n)/\dot{\psi}^*(\hat{\theta}_n), \quad n = 1,2,...$$

where $\hat{\theta}_n$ is the approximation of θ at the n^{th} stage and $\dot{\psi}^*$ is the derivative of ψ^*. While this procedure converges geometrically under mild assumptions on the function ψ^* and starting value $\hat{\theta}_1$, it requires evaluation of ψ^* and $\dot{\psi}^*$ at each stage, a computationally burdensome task. Instead, in §3 we introduce approximations of ψ^* and $\dot{\psi}^*$ at the n^{th} stage, ψ_n^* and $\dot{\psi}_n^*$, $n = 1,2,...$, which are stochastic in nature and very inexpensive to evaluate. This suggests the following stochastic approximation (SA) estimator,

$$\theta_{n+1} = \theta_n - a_n \psi_n^*(\theta_n)/\dot{\psi}_n^*(\theta_n) \quad n = 1,2,... \tag{1.2}$$

where $\{a_n\}$ is a decreasing sequence which is defined more explicitly in §3. The idea of modifying the Newton-Raphson procedure using these approximations is due to Robbins and Monro (1951) and has been used successfully in diverse areas such as bioassay and electrical engineering. In §3 this approximation technique is developed more fully and several desirable asymptotic properties are given. However, as with the deterministic Newton-Raphson algorithm, the performance of the recursive algorithm (1.2) is heavily dependent on the initial value θ_1. For this reason, in §2, we introduce the method of quantile estimation to approximate θ. The quantile estimation method initially converges quickly and turns out to be asymptotically equivalent to the approximation given in (1.2). However, for a large number of iterations it is computationally burdensome. Thus, in §4 we recommend a combination of these two procedures, using the quantile estimation procedure in the initial stages and the computationally simpler algorithm (1.2) in the latter stages. A criterion for stopping the algorithm (1.2) is also given which has desirable properties. Several examples are given throughout the paper. In §5 we close with some concluding remarks and further recommendations.

§2. Quantile Estimation

Define the random variable $Z = \sup_{0<t<T} U(t)$ and let F be the distribution function (d.f.) associated with Z, i.e., $F = 1-\psi$. Let $Z_1,\ldots,Z_N$ be i.i.d. random variables with d.f. F which can be thought of as N independent, computer simulated, realizations of Z. For example, in the Compund Poisson case, let $\{N_i(t),\ 0 \leq t \leq T\}$ be the i^{th} realization of a Poisson process and let $\{X_{ki}\}_{k=1}^{\infty}$ be i.i.d. claim r.v.'s, $i = 1,\ldots,N$. Then,

$$Z_i = \sup_{0<t<T} \left(\sum_{k=1}^{N_i(t)} X_{ki} - P\,t\right) . \tag{2.1}$$

We remark here that although we assume the availability of an infinite number of X_{ki}'s, only a finite number are required for computer simulation, since $N_i(T) < \infty$ almost surely (a.s.), $i = 1,\ldots,N$.

Define $Z_{(1)},\ldots,Z_{(N)}$ to be the order statistics associated with $Z_1,\ldots,Z_N$. The usual quantile estimator of θ that satisfies $F(\theta) = 1-\alpha$ is

$$\hat{\theta}_N = \begin{cases} Z_{(N(1-\alpha))} & N(1-\alpha) = \text{integer} \\ Z_{([N(1-\alpha)]+1)} & N(1-\alpha) \neq \text{integer} \end{cases} , \tag{2.2}$$

where $[\cdot]$ denotes the greatest integer part. That this approximation of θ has desirable properties is given in the following

Lemma 2.1.

(a) Suppose that θ is the unique solution x of $F(x-) \leq 1-\alpha \leq F(x)$. Then

$$\lim_{N\to\infty} \hat{\theta}_N = \theta \quad \text{a.s.} \quad . \tag{2.3}$$

(b) Suppose further that F has density f in a neighborhood of θ, that $f(\theta) > 0$, and that f is continuous at θ. Then

$$\sqrt{N}\,(\hat{\theta}_N-\theta) \to_D N(0,\ \alpha(1-\alpha)/f^2(\theta)) \tag{2.4}$$

where $\to_D N(0,c)$ indicates convergence in distribution to a mean zero normal r.v. with variance c.

The proof of this lemma and of other, deeper properties of quantile estimates can be found, for example, in Serfling (1980, chapter 2). The interpretation of equation (2.3) is that the quantile estimator $\hat{\theta}_N$ is correct in the limit, i.e., as the number of simulations tends to ∞. Equation (2.4) provides more precise information about the rate of convergence. (2.4) is also useful in providing interval estimates of θ, as follows. Define a kernel estimate of $f(\theta)$,

$$\hat{f}_N(\hat{\theta}_N) = N^{-1} \sum_{j=1}^{N} k((z_j - \hat{\theta}_N)/c_N)/c_N \tag{2.5}$$

where

$$k(u) = .75\ (1-u^2)I(-1\leq u\leq 1)\ , \tag{2.6}$$

$I(\cdot)$ is the indicator function of a set and $\{c_n\}$ is a sequence of positive constants tending to zero. The kernel $k(\cdot)$ defined in (2.6) enjoys certain optimality properties which are due to Epanechnikov (1969). With this choice of a kernel function we have the following

Lemma 2.2.

Choose $\{c_n\}$ so that $c_n \to 0$ and $(n\ c_n)^{-1} \log n \to \infty$. Assume f exists and is uniformly continuous over the real line and assume the conditions of Lemma 2.1 hold. Then,

$$\sup_x \left|\hat{f}_N(x) - f(x)\right| \to 0 \tag{2.7}$$

and

$$\begin{aligned} P(\hat{\theta}_N - z_{\gamma/2}(\alpha(1-\alpha)/N)^{1/2}/\hat{f}_N(\hat{\theta}_N) \leq \theta \\ \leq \hat{\theta}_N + z_{\gamma/2}(\alpha(1-\alpha)/N)^{1/2}/\hat{f}_N(\hat{\theta}_N)) \to 1-\gamma \end{aligned} \tag{2.8}$$

where z_γ is the $(1-\gamma)^{th}$ quantile of the standard normal distributions.

The property in (2.7) is a special case of Silverman (1978, Theorem A) who proves (2.7) for a broad class of kernel functions. (2.8) is an immediate result of (2.7) and Lemma 2.1. Other forms of interval estimates can be found, for example, in Serfling (1980, Chapter 2.6). The form in (2.8) is suitable for comparison with interval estimates introduced in §3 and §4.

Example 2.1

To illustrate the calculation of the approximation $\hat{\theta}_N$ and its properties, we assumed the Compound Poisson process of Example 1.1 for a model. We assumed the Poisson process has intensity parameter $\rho = .8$, claims are mean 1 exponential r.v.'s and premiums arrive with unit intensity ($P = 1$). In this particularly simple model, it is easy to check that the regularity conditions of Lemmas 2.1 and 2.2 hold. Further, exact values of $\psi(u)$ can be calculated, c.f., Asmussen (1984). All computations were done on a VAX 11/750 owned and operated by the Department of Statistics at the University of Wisconsin-Madison. The IMSL Fortran subroutines produced the random deviates.

Tables 1 and 2 give values of $\hat{\theta}_N$, $\hat{f}_N(\hat{\theta}_N)$, and the lower and upper 95% confidence limits for $N = 500$ and $N = 1000$, respectively. The lower 95% confidence limit is defined by

$$\hat{\theta}_L = \hat{\theta}_N - 1.96\ (\alpha(1-\alpha)/N)^{1/2}/\hat{f}_N(\hat{\theta}_N)$$

and $\hat{\theta}_U$, the upper 95% confidence limit, is defined similarly. After some trial and error, we used bandwidths $C_{500} = .40$ and $C_{1000} = .20$, respectively. In each table, values are given for ruin probabilities $\psi(\theta) = 1\%, 5\%, 10\%, 40\%$ and horizon times $T = 100, 500$. These ruin probabilities were chosen to represent a range which is typically of interest to the actuarial community. The horizon times were selected so that for the larger $T = 500$, $\psi(u)$ is close to the infinite horizon time case while for $T = 100$, $\psi(u)$ is significantly smaller. In examining Tables 1 and 2, the reader should recall that in the infinite horizon time case the ruin probability has a particularly simple form in this example, i.e., it can be shown that

$$\lim_{T\to\infty} \psi(u) = .8\ \exp\{-.2u\} \quad . \tag{2.9}$$

Thus, for example, if $T = 500$ and $\alpha = .01$, using (2.9) we have $.01 \approx .8\ \exp\{-.2u\}$ which implies that the desired initial reserved level $u \approx 21.91$. In this simple case, several other approximations and bounds are available which may be used in interpreting Tables 1 and 2.

TABLE 1

Compound Poisson Process - Quantile Estimator

N = 500

T	α	$\hat{\theta}_N$	$f_N(\hat{\theta}_N)$	$\hat{\theta}_L$	$\hat{\theta}_u$
100	.01	17.39	.00730	16.20	18.59
	.05	12.01	.01279	10.52	13.51
	.10	8.94	.01773	7.46	10.42
	.40	3.07	.10055	2.64	3.49
500	.01	23.27	.00559	21.71	24.83
	.05	14.17	.02052	13.24	15.10
	.10	10.43	.02920	9.53	11.33
	.40	3.47	.06432	2.80	4.14

TABLE 2

Compound Poisson Process - Quantile Estimator

N = 1000

T	α	$\hat{\theta}_N$	$f_N(\hat{\theta}_N)$	$\hat{\theta}_L$	$\hat{\theta}_u$
100	.01	19.57	.00375	17.93	21.22
	.05	13.16	.01726	12.38	13.94
	.10	9.99	.01742	8.93	11.06
	.40	3.36	.09621	3.04	3.68
500	.01	21.22	.00375	19.58	22.87
	.05	14.57	.01392	13.60	15.54
	.10	10.11	.02451	9.35	10.87
	.40	3.48	.05312	2.91	4.06

Advantages of the quantile estimator defined in (2.2) are that it has desirable asymptotic properties, Lemma 2.1 and 2.2, and that it is easily understood. However, without using truncation devices specific to a problem, computation of $\hat{\theta}_N$ requires storing $Z_1,\ldots,Z_N$, a vector of length N. As the risk manager requires more accuracy, N increases and this storage problem becomes computationally burdensome. To circumvent this computation problem, an asymptotically equivalent stochastic approximation estimator is introduced in the following section.

§3. SA Estimator

The stochastic approximation literature has branched into many directions since initiated by Robbins and Monro (1951). Ljung and Söderström (1983) give an elementary introduction to SA with special emphasis on applications to electrical engineering and automatic control. More mathematical treatments are given by Kushner and Clark (1978) and Chen (1985). Fabian (1971) provides a nice overview of some statistical SA estimators and a bibliography of early papers. See Nevel'son and Has'minskii (1976) for a more theoretical treatment of SA and its relationship with optimal stochastic recursive estimation. Although the mathematical theory of SA techniques has received much attention over the years, there seems to be a small supply of papers on its more applied aspects. Thus, in this paper we relegate the proofs of our results to the Appendix. In this section we give the assumptions of a general SA estimator of the initial reserve and its resulting properties. §4 deals with practical implementation issues of the general procedure.

As in §2, let $Z_1, Z_2, \ldots$ be i.i.d. r.v's with d.F. F. Let f_0 and θ_1 be arbitrary r.v's having finite second moments and let $\{\theta_n\}$ be the sequence of r.v.'s used to estimate θ, defined in (3.2) below. Let $F_n = \sigma(f_0, \theta_1, Z_j,\ j = 1,\ldots,n-1)$ be the sigma-field generated by past events at the n^{th} stage. Let $\{a_n\}$ and $\{c_n\}$ be sequences of positive r.v.'s such that a_n and c_n are F_n-measurable and

$$\lim_{n\to\infty} a_n n = 1 \text{ a.s.}, \quad \lim_{n\to\infty} c_n = 0 \text{ a.s.}$$

For positive constants B_1 and B_2 and some $\delta > 0$, define the estimate of $f(\theta)$ at the n^{th} stage by

$$f_n = \max(\min(f_n^*, B_2 n^{\delta}), B_1) \tag{3.1}$$

where

$$f_n^* = n^{-1} \sum_{j=1}^{n} k((Z_j - \theta_j)/c_j)/c_j$$

and $k(\cdot)$ is defined in (2.6). We define the general SA estimator recursively by

$$\theta_{n+1} = \theta_n + a_n(I(Z_n > \theta_n) - \alpha)/f_{n-1} \qquad n = 1,2,\ldots \ . \tag{3.2}$$

process. Thus, they can be used in many of the models of risk processes
appear in the literature such as the Compound Poisson, ARMA and Stochastic
ounting models. Examples of several models are given to demonstrate the
atility of the procedure and to demonstrate that the procedures are
utationally feasible.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER #2822	2. GOVT ACCESSION NO. AD-A158 162
	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Approximation of the Initial Reserve for Known Ruin Probabilities	5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Edward W. Frees	8. CONTRACT OR GRANT NUMBER(s) DAAG29-80-C-0041
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53705	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Work Unit Number 4 - Statistics & Probability
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office P.O. Box 12211 Research Triangle Park, North Carolina 27709	12. REPORT DATE May 1985 13. NUMBER OF PAGES 21
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)	15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

stochastic approximation, quantile estimation, kernel estimation

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

An important problem in the study of actuarial risk theory is approximating the probability of ruin within finite time based on a specified initial reserve. In this paper we address the similar, but mathematically different, problem of how to approximate a desired initial reserve given a pre-specified probability of ruin. Although the procedures have desirable asymptotic properties such as consistency and asymptotic normality, they are computer-intensive and would not have been practicable before the wide spread availability of high-speed computers. The procedures rely on simulated realizations of a general

(continued)

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

McLeish, D. (1976). Functional and random central limit theorems for the Robbins-Monro process. J. Appl. Probab. 13, 148-154.

Nevel'son, M. and Has'minskii, R. (1976). Stochastic Approximation and Recursive Estimation. American Mathematical Society, Providence, RI.

Papatriandafylou, A. and Waters, H. (1984). Martingales in Life Insurance. Paper presented at the Risk Theory meeting in Oberwolfach.

Robbins, H. and Monro, S. (1951). A stochastic approximation method. Ann. Math. Statist. 22, 400-407.

Robbins, H. and Siegmund, D. (1971). A convergence theorem for nonnegative almost supermartingales and some applications. Optimizing Methods in Statistics (J. S. Rustagi, Ed.), 233-257. Academic Press, New York.

Ruppert, D., Reish, R., Deriso, R. and Carroll, R. (1984). Optimization using stochastic approximation and Monte-Carlo simulation (with application to harvesting of Atlantic menhaden). Biometrics 40, 535-545.

Serfling, R. (1980). Approximation Theorems of Mathematical Statistics. Wiley, New York.

Sielken, R. (1973). Stopping times for stochastic approximation procedures. Z. Wahrscheinlichkeitstheorie Verw. Geb. 26, 67-75.

Silverman, B. (1978). Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. Ann. Statist. 6, 177-184.

EWF/jvs

References

Anscombe, F. (1952). Large-sample theory of sequential estimation. Proc. Cambridge Phil. Soc. 48, 600-607.

Asmussen, S. (1984). Approximations for the probability of ruin with finite time. Scand. Act. J. , 31-57.

Bühlmann, H. (1976). A probabilistic approach to long term insurance. Twentieth Intl. Congress of Act., Tokyo.

Chen, H. (1985). Recursive Estimation and Control for Stochastic Systems. Wiley, New York.

Chow, Y. and Robbins, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean. Ann. Math. Statist. 36, 457-462.

Dvoretsky, A. (1956). On stochastic approximation. Proc. Third Berkeley Symp. Math. Statist. Probab. 1 (J. Neyman, ed.), 39-55. Univ. of California Press.

Epanechnikov, V. (1969). Nonparametric estimation of multivariate probability density. Theory Probab. Appl. 14, 153-158.

Fabian, V. (1968). On asymptotic normality in stochastic approximation. Ann. Math. Statist. 39, 1327-1332.

Fabian, V. (1971). Stochastic approximation. Optimizing Methods in Statistics (J. S. Rustagi, ed.), 439-460.

Frees, E. (1985). Weak convergence of stochastic approximation processes with random indices. Sequential Ann.: Comm. Statist. C, June, 1985.

Frees, E. and Ruppert, D. (1985). Sequential nonparametric age replacement policies. Ann. Statist. , June, 1985).

Gerber, H. (1976). A probabilistic model for life contingencies and a delta-free approach to contingency reserves. Trans. Soc. Act. 28, 127-141.

Kushner, H. and Clark, D. (1978). Stochastic Approximation Methods for Constrained and Unconstrained Systems. Springer-Verlag, New York.

Ljung, L. and Söderström, T. (1983). Theory and Practice of Recursive Identification. MIT Press, Cambridge, MA.

and

$$B_n = n^{-1} \sum_{j=1}^{n} \{c_j^{-1} E(k((Z_j-\theta_j)/c_j)|F_j) - f(\theta_j)\} \quad .$$

By (A.2), sufficient for the proof of the Lemma is to show that $A_n \to 0$ and $B_n \to 0$. From a change of variable,

$$\begin{aligned} & c_j^{-1} E(k((Z_j-\theta_j)/c_j)|F_j) - f(\theta_j) \\ &= \int c_j^{-1} k((u-\theta_j)/c_j) f(u)du - f(\theta_j) \\ &= \int k(u)(f(\theta_j+uc_j) - f(\theta_j))du \\ &\to 0 \end{aligned}$$

by the Bounded Convergence Theorem. This and Kronecker's Lemma show that $B_n \to 0$. By construction, $\{n A_n\}$ is a zero-mean martingale and thus standard martingale convergence theorems can be used to show $A_n \to 0$. ∎

Proof of Theorem 3.2:

With $X_n = \theta_n - \theta$, we re-write (3.2) to get

$$X_{n+1} = X_n - a_n/f_{n-1}(I(Z_n \le \theta_n) - F(\theta)) \quad . \tag{A.3}$$

Use the martingale difference

$$V_n = n\, a_n/f_{n-1}\{I(Z_n \le \theta_n) - F(\theta_n)\} \quad ,$$

in (A.3) to get

$$X_{n+1} = X_n - a_n/f_{n-1}\{F(\theta_n) - F(\theta)\} - n^{-1}V_n \quad .$$

We can write $F(\theta_n) - F(\theta) = (\theta_n-\theta)f(\eta_n)$, where η_n satisfies $|\eta_n-\theta| \le |\theta_n-\theta|$. With $\Gamma_n = n\, a_n f(\eta_n)/f_{n-1}$, we have

$$X_{n+1} = X_n(1 - n^{-1}\Gamma_n) - n^{-1}V_n \quad . \tag{A.4}$$

By Theorem 3.1 and Lemma A.1, $\Gamma_n \to 1$. By (3.1), $E(V_n^2|F_n) = n^2a_n^2/f_{n-1}^2\, F(\theta_n)(1 - F(\theta_n))$ is bounded and tends to $\alpha(1-\alpha)/f(\theta)^2$. The result of the theorem is now an easy application of Theorem 2.2 of Fabian after satisfying

$$E(I(V_n^2 > \epsilon)V_n^2 \mid F_n) \to 0 \quad \text{for each} \quad \epsilon > 0 \quad .$$

This is easily seen by applying a conditional Hölder's inequality. ∎

Appendix

The proofs of Theorems 3.1 and 3.2 are applications of results of Robbins and Siegmund (1971) and Fabian (1968). See Frees and Ruppert (1985, Appendix) for a convenient reference of these results. Theorem 3.3 can be proved by modifying Sielken's (1973) arguments and will not be given here. All relationships between random variables are meant to hold almost surely unless otherwise specified.

Proof of Theorem 3.1:

Define $X_n = \theta_n - \theta$. From (3.2), after subtracting θ, squaring and taking conditional expectations with respect to F_n, we get

$$E(X_{n+1}^2 \mid F_n) = X_n^2 + 2a_n X_n / f_{n-1}((1-\alpha) - F(\theta_n)) + a_n^2 / f_{n-1}^2 (E((I(Z_n > \theta_n) - \alpha)^2 \mid F_n)) \quad . \tag{A.1}$$

Define $\beta_n = a_n^2 / f_{n-1}^2 (E((I(Z_n > \theta_n) - \alpha)^2 \mid F_n))$. Thus,

$$E(X_{n+1}^2 \mid F_n) = X_n^2 - 2a_n X_n / f_{n-1}(F(\theta_n) - F(\theta)) + \beta_n \quad . \quad .$$

By (3.1), $\sum \beta_n < \infty$. From Theorem 1 of Robbins and Siegmund, we have that $\lim X_n^2$ exists and is finite and that $\sum a_n X_n / f_{n-1}(F(\theta_n) - F(\theta)) < \infty$. By (3.1), $\sum a_n / f_{n-1} = \infty$ and by A.1, $(x-\theta)(F(x) - F(\theta)) > 0$ for all $x \neq \theta$. Thus, we have the result. †

Prior to the proof of Theorem 3.2, we give the following preparatory

Lemma A.1.

Assume A1 and A2. Then, for f_n defined in (3.1),

$$\lim_{n \to \infty} f_n = f(\theta) \quad .$$

Proof:

By Theorem 3.1 and Kronecker's Lemma, we have

$$g_n = n^{-1} \sum_{j=1}^{n} f(\theta_j) \to f(\theta) \quad . \tag{A.2}$$

Define $f_n^* - g_n = A_n + B_n$, where

$$A_n = n^{-1} \sum_{j=1}^{n} c_j^{-1} \{k((Z_j - \theta_j)/c_j) - E(k((Z_j - \theta_j)/c_j) \mid F_j)\}$$

§5. Summary and Conclusions

In contrast to other papers in the literature concerning ruin probability calculations, in this paper the probability of ruin is a fixed, known quantity while the initial reserve necessary to achieve this probability is the parameter of interest. Other components, both stochastic and deterministic, of the ruin probability are considered known. In this paper we have given the theory for procedures which are computer-intensive and yet, which can operate in a simple computing environment. The quantile estimation technique of §2 requires little more than a good random number generator; not even a numerical integration routine is necessary. The stochastic approximation technique of §3 also requires a large amount of computation but eliminates the problem of storing a large vector. Together these procedures are suitable for use in a micro-computing environment.

We have chosen to separate the risk reserve process $R(t)$ into two components, the initial reserve and the net liability, where

$$R(t) = u - U(t) \quad . \tag{5.1}$$

Although $R(t)$ is the traditional process used in ruin probability approximations, the separation was important because we wanted to explicitly denote the parameter of interest u. Restating (1.1), we have

$$\psi(u) = P\left(\inf_{0<t<T} R(t) < 0 \right) \quad .$$

A crucial point of our derivations is that $R(t)$ and hence $\inf_{0<t<T} R(t)$ is a monotone function of u. In less general models, $\inf_{0<t<T} R(t)$ may also be a monotone function of other parameters such as premiums, loading expenses, interest, mortality, etc. In specific models, the techniques of this paper can be extended to estimate a typical such parameter given a pre-specified ruin probability and initial reserve. For example, in the Compound Poisson process case, for a fixed α and u one might wish to estimate P. This can be considered to be a new premium principle which we intend to explore in a later paper.

Example 4.3. Stochastic Discounting model.

The performance of the two-stage estimator in Example 1.3 is illustrated through a version of Example 1 of Papatriandafylou and Waters (1984). Here, the risk is a 20 year term insurance policy to a life aged 45. The unit sum insured is to be paid at the end of the year of death and the incremental death random variable is governed by the force of mortality

$$r(x) = (.0000177)(1.11683)^x \ ,$$

i.e., a special case of Gompertz's Law. Using an interest rate of 4%, the expected annual premium with a 10% loading is $P = .0085335$ which is payable at the beginning of the year. The actual annual interest increment was assumed to be a random variable independent of mortality having a Lognormal distribution with mean 1.04 and standard deviation .02. Table 6 illustrates the performance of the two-stage estimator for $m+n = 1,000$ and 10,000 observations, respectively. The estimators show good agreement in that, even with different initial sample sizes, the confidence interval based on 10,000 observations is a strict subset of the corresponding interval based on 1,000 observations. Further, even though Papatriandafylou and Waters assumed a deterministic interest rate while we used a stochastic rate, there is a nice correspondence between our Table 6 and Table 3 of Papatriandafylou and Waters.

TABLE 6

Stochastic Discounting Model - Two-stage estimator

m	n	α	θ_n	f_n	θ_{nL}	θ_{nU}
500	500	.01	.8190	.155701	.7630	.8750
		.05	.5999	.198470	.5036	.6961
		.10	.4642	.410272	.4001	.5283
1000	9000	.01	.8364	.093327	.8144	.8584
		.05	.5809	.191120	.5573	.6045
		.10	.4198	.395416	.4041	.4358

observations are strict subsets of the corresponding intervals based on 1,000 observations. Table 5 provides a qualitative idea of how many simulations are required for a certain degree of accuracy. In Table 5 we see that the number required decreases as α increases and increases as d decreases. There does not seem to be a discernable pattern in the stopped estimates, $\theta_{N(d)}$.

TABLE 4

ARMA Model - Two-Stage Estimator

m = 500

n	α	θ_n	f_n	θ_{nL}	θ_{nu}
500	.01	1.6475	.043182	1.4455	1.8494
	.05	1.0159	.100904	.8265	1.2052
	.10	.4676	.137584	.2765	.6587
4500	.01	1.6469	.031503	1.5547	1.7392
	.05	1.0012	.082585	.9241	1.0784
	.10	.4761	.151321	.4182	.5340

TABLE 5

ARMA Model - Two-Stage Estimator
stopping rule (3.5)

m = 500 n = 4500

α		d=.5	d=.3	d=.2	d=.1
.01	N(d)	1316	1644	2311	4193
	$\theta_{N(d)}$	1.6450	1.6459	1.6451	1.6468
.05	N(d)	1004	1672	2254	3482
	$\theta_{N(d)}$	1.0084	1.0053	1.0019	1.0019
.10	N(d)	952	1350	1900	3149
	$\theta_{N(d)}$	.4688	.4652	.4708	.4694

typical since the asymptotic variance for both estimators is proportional to $\alpha(1-\alpha)/f^2(\theta)$, in general an increasing function of $\alpha < 1/2$. In the following two examples, we used larger simulation sizes to investigate these small ruin probabilities.

TABLE 3

Compound Poisson Process - Two-Stage Estimator

		m = 500		n = 500	
T	α	θ_n	f_n	θ_{nL}	θ_{nU}
100	.01	17.41	.00713	16.18	18.63
	.05	12.02	.01468	10.72	13.32
	.10	8.95	.01918	7.58	10.32
	.40	3.10	.10099	2.67	3.52
500	.01	23.27	.00279	20.15	26.39
	.05	14.17	.01472	12.87	15.49
	.10	10.42	.01844	9.00	11.85
	.40	3.50	.06500	2.84	4.16

Example 4.2 ARMA model.

The performance of the two-stage estimator in Example 1.2 is illustrated in the simple AR(1) model,

$$G_t = .2\, G_{t-1} + \varepsilon_t \ ,$$

with $G_0 = g_0 = 0$. The error random variables $\{\varepsilon_i\}$ were assumed to be standard normal but truncated at plus and minus 4 and then shifted to have mean 1. Thinking of the units of time as months, we used $T = 120$ for a ten year time horizon. In Table 4, we see a good agreement, for selected values of α, between the estimator at $n+m = 1{,}000$ observations and 5,000 observations. Further, the confidence intervals based on 5,000

approaches, we advocate the following two-stage estimator. This compromise estimator is also used in the examples which follow.

In the first stage, let m i.i.d. observations of Z be available for estimation and use the quantile estimation techniques of §2 to get estimates

$$\theta_1 = \hat{\theta}_m \quad \text{and} \quad f_0 = \hat{f}_m(\hat{\theta}_m) \quad . \tag{4.2}$$

The choice of m will depend on the application and computing power available. In the following examples we used $m = 500$ and $m = 1000$. In the second stage, use the SA estimator of §3 with the initial estimates given in (4.2). Note that in this suggested procedure θ_1 and f_0 are nondegenerate random variables with finite second moments. In this application, because f_0 is a good initial estimate of $f(\theta)$, we prefer to use f_0 in the updating of f_n. Thus, instead of using $k_A = m$ as in (4.1) to define the sequence $\{a_n\}$, use $a_1 = 1/(m f_0)$ and

$$a_{n+1} = n^{-1}(\textstyle\sum_{j=1}^{n} \xi_j)/(\sum_{j=1}^{n} \xi_j + m f_0) \ , \ n = 1,2,\ldots$$

where $\xi_j = c_j^{-1} k((Z_j - \theta_j)/c_j)$. Thus,

$$a_{n+1}/f_n = [\textstyle\sum_{j=1}^{n} \xi_j + m f_0]^{-1} \tag{4.3}$$

or a_{n+1}/f_n is $(n+m)^{-1}$ times the reciprocal of the weighted average of the estimates f_n and f_0, where the weights are the number of observations available for each estimate. To illustrate this technique we now give several examples.

Example 4.1 Compound Poisson model.

Table 3 illustrates the performance of the two-stage estimator as in Examples 2.1 and 1.1. Qualitatively, the estimators perform properly; the value of the estimator decreases as the ruin probability increases and increases as the time horizon increases. Comparing Tables 2 and 3 for large values of α, there is little difference in the estimator and resulting confidence intervals. Because of this corroboration between different approximation techniques, it appears that $N = n+m = 1000$ observations is a reasonable size for estimation with the associated confidence intervals. However, for small values of α, there are discrepancies in comparing estimators and confidence intervals. This result is

§4. A Two-Stage Procedure

The asymptotic theory of §3 shows that the SA estimator θ_n defined in (3.2) has many desirable properties. However, the asymptotic properties leave a wide range of parameters, used to define θ_n, whose choice affects the finite sample properties of θ_n. In this section we discuss the choice of these parameters and give several examples.

As with deterministic Newton-Raphson algorithms, it is well known that the choice of the initial starting value θ_1 in SA procedures is important. Although θ_1 does not appear in the limit theorems 3.1 - 3.3, it has a marked effect in finite samples. A good choice of f_0 is also important but, because of the averaging taking place in the definition of f_n (see (3.1) and below), the choice of f_0 is not as crucial.

Because the asymptotic properties of §3 are valid for any sequence $\{a_n\}$ such that $a_n n \to 1$, researchers in SA often use the simple form $a_n = 1/n$. However, in finite samples this causes the estimate θ_n to fluctuate widely in the early stages as compared to larger values of n. Applied researchers, cf., Ruppert et al. (1984), suggest an asymptotically equivalent form,

$$a_n = 1/(n + k_A) \quad , \tag{4.1}$$

where k_A is some positive number. Some justification is given for this choice of a_n in (4.1) by Dvoretsky (1956). The choice of the boundary constants B_1 and B_2 in the estimator of the density of θ, see (3.1), depends on the particular problem. In our examples we did not find the choice to be crucial and used $B_1 = .0001$ and $B_2 = 10{,}000$. The choice of $\{c_n\}$ does not appear to be crucial, either. We used a common choice, $c_n = Cn^{-1/3}$, which was also used in other studies, cf., Frees (1985).

Although asymptotically equivalent, we found that the quantile estimator of §2 initially converged more quickly than the stochastic approximation estimator of §3. This can be justified heuristically by arguing that in updating the estimator from the N^{th} to the $(N+1)^{st}$ stage, the quantile estimator uses information from all $N+1$ observations while the SA estimator is Markovian in nature, i.e., it uses only a_n, f_{n-1}, θ_n and Z_n to update the estimator. Of course, the quantile estimator is computationally more tedious than the SA estimator. In order to take advantage of the strengths of both

Let $2d > 0$ be the desired length of the confidence interval. One criterion is to halt the experiment when the stage of the experiment reaches the stopping variable

$$N(d) = \inf\{n \geq p : d > z_{\gamma/2}(\alpha(1-\alpha)/n)^{1/2} a_n/f_n\} \quad . \tag{3.5}$$

Because the justification for the use of $N(d)$ is asymptotic in nature, we usually require that the experiment run for at least p stages, where p is fixed and known. That this criterion is desirable is provided in the following

Theorem 3.3.

Assume A1 and A2. Then

$$\lim_{d \to 0} P(|\theta_{N(d)} - \theta| < d) = 1-\gamma \quad . \tag{3.6}$$

To prove this result, one would show that Theorem 3.2 is still valid when the limit goes through random indices. Anscombe (1952) has provided a sufficient condition which he termed uniform continuity in probability. In the SA literature this technique was first used by Sielken (1973) for a different problem. Other methodologies to achieve this random convergence have been used by McLeish (1976) and Frees (1985). McLeish showed convergence through random indices by using weak functional central limit theorems. Frees showed random convergence by using strong invariance principles. We remark that (3.6) is not all one might hope for. A better result is $P(|\theta_{N(d)} - \theta| < d) \geq 1-\gamma$, that is, true for all fixed d. However, this result is not available even in the simple case considered by Chow and Robbins. Numerical studies by Sielken (1973) and Frees (1985) indicate that (3.6) is adequate for practical purposes. The interpretation of (3.6) is that the confidence interval

$$(\theta_{N(d)} - d, \ \theta_{N(d)} + d) \tag{3.7}$$

covers the parameter θ with probability $1-\alpha$, at least asymptotically. Intervals of the form (3.7) are known as sequential fixed-width confidence intervals and are useful in simulation and sequential statistical experiments.

Specific choices of the parameters in (3.2), $\{a_n\}$, $\{c_n\}$, θ_1 and f_0, are discussed in §4. To ensure that the estimator defined in (3.2) has desirable asymptotic properties, we will use the following assumptions.

A1. The unique solution x of $F(x-) \leq 1-\alpha \leq F(x)$ is θ.

A2. The density of $F(x)$, $f(x)$, exists for all x, is bounded, $f(\theta) > 0$, and f is continuous at θ.

We have the following properties whose proofs appear in the Appendix.

Theorem 3.1

Assume A1. Then,

$$\lim_{n\to\infty} \theta_n = \theta \quad \text{a.s.}$$

Theorem 3.2

Assume A1 and A2. Then

$$\sqrt{n}\,(\theta_n-\theta) \to_D N(0,\ \alpha(1-\alpha)/f^2(\theta)) \tag{3.3}$$

and

$$P(\theta_n - z_{\gamma/2}(\alpha(1-\alpha)/n)^{1/2}/f_n \leq \theta \leq \theta_n + z_{\gamma/2}(\alpha(1-\alpha)/n)^{1/2}/f_n) \to 1-\gamma \ . \tag{3.4}$$

Comparing (2.4) and (3.3) we see that $\hat{\theta}_N$ and θ_n have the same asymptotic distribution. We remark that A1 is the same assumption as part (a) of Lemma 2.1. Assumption A2 is slightly stronger than the assumption of part (b) of Lemma 2.1 in that A2 requires the existence and boundedness of $f(x)$ for all x, not just in a neighborhood of θ. Because of the local nature of SA estimator, A1 and A2 are weaker than the assumptions of Lemma 2.2.

An important aspect of any simulation experiment is the development of a reasonable stopping criterion or rule for halting the recursive estimation procedure. We now give such a procedure, originally due to Chow and Robbins (1965) in the framework of sequential estimation of the mean of a standard normal distribution, which produces a sequential fixed width confidence interval.

END

FILMED

9-85

DTIC