TRANSACTIONS OF SOCIETY OF ACTUARIES
1984 VOL. 36

OPTIMAL RUIN CALCULATIONS USING PARTIAL STOCHASTIC INFORMATION

PATRICK L. BROCKETT* AND SAMUEL H. COX, JR.

ABSTRACT

In this paper we show how to obtain tight upper and lower bounds on $E[h(X)]$ for a given function h, and a random variable X with three known moments μ, σ^2, and ρ. The improvement possible when we have the additional knowledge that X is unimodal is also discussed. We show how these bounds can be used in calculating the probability of ruin and in setting initial reserve levels when we have only incomplete information concerning the statistical distribution of the loss variable.

I. INTRODUCTION

In this paper we shall present some mathematical formulas concerning bounds of fairly general function of a random variable when the only knowledge we have about the random variable is the lower order raw moments. Thus we could obtain tight upper and lower bounds on the variance σ^2 of a random variable X given only its mean μ, obtain tight bounds on the skewness (third central moment) ρ in terms of the mean and variance, obtain bounds on the expected value of $h(x)$ where $h^{(4)}(x) > 0$ in terms of the first three moments, and so on. We then show how to use the technique of Kemperman (1971) to obtain even tighter bounds when we assume additionally that the random variable in question is unimodal. As an application, we bound the ruin probability of risk theory.

II. OPTIMAL BOUNDS USING PARTIAL MOMENT KNOWLEDGE

Our goal in this section is to show how to determine tight upper and lower bounds for the expected value of a function of some random variable with given moments. The tool necessary for this development is the celebrated Markov-Krein theorem from the theory of Chebychev systems of functions.

* Mr. Brockett, not a member of the Society of Actuaries, is Associate Professor of Actuarial Science and Finance and with the Applied Research Labs, University of Texas at Austin.

Stated mathematically, the problem considered in this section is the following: Given a random variable X on $[a,b]$ with central moments $\mu = (\mu_0, \mu_1, \mu_2, \ldots, \mu_k)$, and given a function $h(x)$ on $[a,b]$ find the tightest possible bounds on $E[h(X)]$, the expectation of $h(X)$. This problem has numerous applications, among which are optimal choice of retention limits in stop loss reinsurance with partial information (compare DeVylder, 1983) optimal critical claim size in a bonus system with partial information, (compare DePril and Goovaerts, 1983) and ruin theory calculations when there is only partial stochastic information concerning the size of loss distribution. It is the latter which we shall use throughout to illustrate, although the technique presented is very general and capable of many insurance applications.

Let $\mu = (\mu_1, \mu_2, \ldots, \mu_k)$ be a point in the moment space M_k consisting of all k tuples consistent with the first k moments of some probability measure ν on the interval $[a,b]$, i.e., such that

$$\mu_i = \int_a^b x^i \, \nu \, (dx)$$

for $i = 1,2,\ldots,k$, for some measure ν. Additionally, let $h(x)$ be a function which is $(k+1)$ times differentiable with $h^{(k+1)}(x) > 0$ on $[a,b]$. As a corollary to the Markov-Krein theorem from Chebychev systems (compare Karlin and Studden, 1966) the following theorem is developed in Brockett (1984).

THEOREM

2.1a) If the mean μ is given, then for any random variable X on $[a,b]$ with mean μ we have

$$h(\mu) \le E[h(X)] \le h(a)p + h(b)(1-p),$$

$$p = \frac{(b-\mu)}{(b-a)}$$

The measure ν_1 which assigns mass p to the point b and $1-p$ to the point a is called the upper principal representation for the moment point μ.

b) If μ and σ^2 are given, then for any function h with $h^{(3)}(x) \ge 0$, and any random variable X on $[a,b]$ with mean μ, and variance σ^2, we have

$$h(a)p + h(\xi_1)(1-p) \le E[h(X)] \le h(\xi_2)q + h(b)(1-q),$$

where

$$p = \frac{\sigma^2}{\sigma^2 + (a-\mu)^2}, \quad \xi_1 = \mu - \frac{\sigma^2}{a-\mu}$$

$$\xi_2 = \mu - \frac{\sigma^2}{b-\mu} \text{ and } q = \frac{(b-\mu)^2}{\sigma^2+(b-\mu)^2}.$$

The measure v_1 which assigns probability q to the point ξ_2 and $(1-q)$ to the point b is called the upper principal representation of (μ,σ^2). The measure v_0 which assigns probability p to the point a and probability $1-p$ is the point ξ_1 is called the lower principal representation of the moment point (μ,σ^2).

c) If μ,σ^2, and ρ are given, then for any function h with $h^{(4)}(x) \ge 0$, and any random variable X on $[a,b]$ with mean μ, variance σ^2, and third moment $\rho = E(X-\mu)^3$, we have

$$h(\eta_1)q + h(\eta_2)(1-q) \le E[h(X)] \le h(a)p_1 + h(\xi)p_2 + h(b)(1-p_1-p_2),$$

where

$$\xi = \frac{\rho-(a+b-2\mu)\sigma^2}{(a-\mu)(b-\mu)+\sigma^2} + \mu,$$

$$p_1 = \frac{\sigma^2 + (\xi-\mu)(b-\mu)}{(b-a)(\xi-a)}, \tag{2.2}$$

$$p_2 = \frac{\sigma^2+(b-\mu)(a-\mu)}{(\xi-b)(\xi-a)},$$

and

$$\eta_1 = \frac{\rho-\sqrt{\rho^2+4\sigma^6}}{2\sigma^2} + \mu,$$

$$\eta_2 = \frac{\rho+\sqrt{\rho^2+4\sigma^6}}{2\sigma^2} + \mu, \tag{2.3}$$

$$q = \frac{1}{2} + \frac{\rho}{2\sqrt{\rho^2+4\sigma^6}}.$$

The measure ν_1 which assigns probabilities p_1 to a, p_2 to ξ, and $1-p_1-p_2$ to b is called the upper principal representation, and the measure ν_0 which assigns probabilites q to η_1 and $1-q$ to η_2 is called the lower principal representation of (μ,σ^2,ρ).

Notice that in all possible situations the upper and lower principal representations of the given set of moments are themselves probability measures possessing the given moments. Thus the inequalities in theorem 2.1 are actually tight, in other words, attainable bounds which cannot be improved without requiring additional information about the random variable involved. Another useful fact which should be noted is that the upper and lower principle representations do not depend in any way upon the function h which is used.

As mentioned previously, theorem 2.1 has numerous applications (for example, when $h(x)=(x-t)^2$ and μ is given we derive bounds on stop loss variance). In the next section we shall explicitly develop one such application.

III. APPLICATION: ASSESSING THE PROBABILITY OF RUIN USING INCOMPLETE LOSS DISTRIBUTION INFORMATION

Consider the collective risk model as described in the new Society of Actuaries study note on risk theory [1]. We shall only briefly sketch the model here since the development in [1] is quite complete. The cash surplus at time t is defined to be

$$U(t) = u + ct - S(t), \; t\geq 0.$$

Here $U(0) = u$ is the initial surplus, c is the rate at which premiums are credited to the fund in dollars per year, and S is the stochastic claims process:

$$S(t) = X_1 + \ldots + X_{N(t)},$$

where $N(t)$ is a Poisson process with parameter λ and the $X_i\geq 0$ are the independent and identically distributed loss variables.[1] Ruin is said to occur if $U(t)\leq 0$, that is, if the cash surplus falls below zero.

[1]As noted in Bowers et al.[1], the usual compound Poisson model for the claim process $S(t)$ can be extended to an autoregression model with dependent X_i's. Again an "adjustment coefficient" is the pertinent determinant in the formula for the probability of ruin. Our method of analysis can easily be extended to incorporate this generality. We leave it to the reader to make the obvious modifications, after consulting our section II and the development in Bowers et al.[1].

We are interested here in determining the probability that there is eventual ruin as a function of the initial reserve $U(0)=u$. Let us denote this probability by $\psi(u)$. The main theorem of chapter 12 of Bowers et al. (1982) is that the ruin probability is exactly equal to

$$\psi(u) = \frac{e^{-Ru}}{E[e^{-RU(T)} \mid T<\infty]}, \tag{3.2}$$

where $T=\inf\{t:t\geq 0 \text{ and } U(t)<0\}$ is the time of ruin, and R is the so-called adjustment coefficient which depends upon three things: the distribution of the losses, X, the frequency with which losses occur, λ, and the load factor θ, which was used for setting the premiums. By definition the adjustment coefficient is the smallest positive solution to the equation

$$1 + (1 + \theta)\mu r = M_X(r),$$

where X is a random variable having the common distribution of the losses, $\mu=E[X]$, and $\lambda\mu(1 + \theta) = c$ is the premium charged, and $M_X(r)=E[e^{rX}]$ is the moment generation function for X. As shown in [1], there is a unique positive solution R to the above equation provided only that $\theta>0$.

Assuming now that we have only partial knowledge about the loss variable X, we do not know $M_X(r)$ and hence cannot directly solve (3.2). We can, however, use the partial information about the moments of X to determine bounding curves on $M_X(r)=E[\exp(rX)]$. We have from theorem 2.1 with $h(x)=\exp(rx)$ for $r>0$,

$$M_0(r) \leq M_X(r) \leq M_1(r),$$

where $M_0(r)$ and $M_1(r)$ denote the moment generating functions of the lower and upper principal representations for the moments of the loss distribution which we have. Thus if $a \leq X \leq b$ and μ, $\sigma^2 = \text{var}(X)$ and $\rho = E[(X-\mu)^3]$ are known, then one can numerically solve, by the Newton-Raphson method or the successive bisection method, for example, for R_0 and R_1, the adjustment coefficients corresponding to the upper and lower principal representations of the given moments. For all $r \geq 0$ the curves satisfy $M_0(r) \leq M_X(r) \leq M_1(r)$ since the extremal measures v_0 and v_1 do not depend upon r in $h(x)=\exp(rx)$. At $r = 0$, the functions M_0, M_X and M_1 are equal to 1 and have slopes of μ. Since $\theta>0$, the graph of $1+(1+\theta)\mu r$ has a slope which is strictly greater than μ, and hence this straight line starts out strictly above the curves M_0, M_X and M_1. For $r > 0$, the curves M_X, M_0, and M_1 are convex since all moment-generating functions of positive variables are. Hence $1+(1+\theta)\mu r$ intersects these curves exactly twice, once at zero and once at a positive value. The intersections for the positive values are precisely in order M_1, M_X and M_0 from left to right as shown in Figure 1. Hence the corresponding adjustment coefficients must satisfy $R_1 < R < R_0$ as pictured in the following chart. (See Bowers et al., 1982, for a rigorous proof that the curves are convex and intersect the line exactly one time for $r > 0$.)

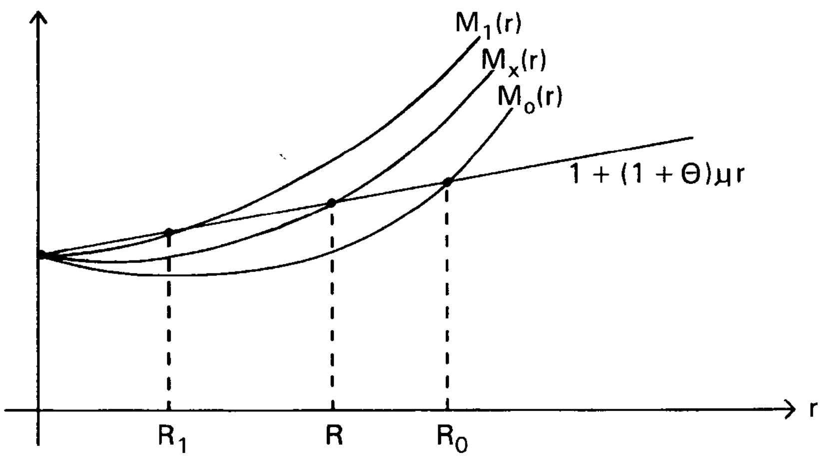

Figure 1:
Bounds on the Adjustment Coefficient using partial information

As noted in Bowers et al. (1982), if $a \leq X \leq b$ then, from the above formula for $\psi(u)$ we easily obtain the bounds on the ruin probability, namely $\exp(-R_X(u+b)) \leq \psi(u) \leq \exp(-R_X u)$. Now given the moments of the loss distribution, X, we may determine the upper and lower principal representations from the formulas of theorem 2.1, and hence we can solve numerically for R_0 and R_1 on the computer by successive bisection or Newton-Raphson techniques. These values are just the adjustment coefficients corresponding to the principal representations assuming ν_0 and ν_1 respectively are the "true" distribution for X. We find the following bounds on the ruin probability using partial information:

$$\exp(-R_0(u+b)) \leq \psi(u) \leq \exp(-R_1 u).$$

As an example, consider a group medical insurance policy which covers from the first dollar of loss up to a maximum of $5000. Assume that moments μ, σ^2, and ρ of claim size are known and the rate of frequency of claims, λ, is also known. We will approximate R, the adjustment coefficient, using the numerical values $a = 0$, $b = 5{,}000$, and the mean $\mu = 139$. Next we will include the information that $\sigma^2 = 39{,}975$ into the calculation, and finally include the skewness measure $\rho = 57{,}320{,}000$. It should be emphasized that the bounds on R are tight in the sense that both equalities are possible. These bounds cannot be improved without specifically obtaining more information about X.

The equations defining the principal representations in theorem 2.1 were solved in the two moment and three moment cases. Then, using the principal representations as the distribution for X, the adjustment coefficient R_1 and R_0 were found exactly as described earlier and as described in Bowers et al. [1]. Table 1 presents the numerical results. The values of R_1 can be used to give upper bounds on $\psi(u)$, the ruin probability, because $\psi(u) \leq e^{-Ru} \leq e^{R_1 u}$. If it is desired to know the value of u which will insure that the ruin

TABLE 1

BOUNDS UPON THE ADJUSTMENT COEFFICIENT USING ONLY MOMENT INFORMATION.*

Premium Load Factor	Number of Moments	$R_1 \times 10^4$	$R_1 \times 10^4$
$\theta = .1$			
	1	0.375	13.503
	2	3.021	4.400
	3	3.741	3.913
$\theta = .2$			
	1	0.708	25.482
	2	4.504	8.303
	3	5.958	6.753
$\theta = .3$			
	1	1.001	36.233
	2	5.522	11.806
	3	7.345	8.948
$\theta = .4$			
	1	1.278	45.973
	2	6.239	14.980
	3	8.305	10.722

*Notice particularly how quickly the width of the interval of indeterminancy decreases as more moments are included.

probability $\psi(u)$ is smaller than some permissible level, say .05, then one can solve the equation $e^{-R_1 u} = .05$ for u and obtain $u = -ln(.05)/R_1$. For load factors 0 = .1, .2, .3, .4, and for one, two or three moments known about X, this technique results in the following values of u given in table 2. One cannot improve upon these bounds unless more knowledge about X is gathered since the distribution ν_1 is entirely consistent with the given moment knowledge.

TABLE 2

THE INITIAL RESERVE u NEEDED TO INSURE THE PROBABILITY OF EVENTUAL INSOLVENCY LESS THAN .05, USING ONLY MOMENT INFORMATION CONCERNING THE LOSS VARIABLE X.

Premium Load Factor	Number of Moments Used	Intial Reserve Required $-ln.05/R_1$
$\theta = .1$		
	1	\$79,886
	2	9,916
	3	8,008
$\theta = .2$		
	1	42,313
	2	6,651
	3	5,028
$\theta = .3$		
	1	29,927
	2	5,425
	3	4,079
$\theta = .4$		
	1	23,441
	2	4,802
	3	3,607

IV. IMPROVEMENTS IN THE BOUNDS WHEN THE RANDOM VARIABLES ARE KNOWN TO BE UNIMODAL

Often more is known about the return distribution than just the first few central moments. For example, the distribution is frequently known to be unimodal. In this section we show how to use this information to improve the Chebychev system bounds. Slightly more general and advanced presentations of these results may be found in Kemperman (1971) and Karlin and Studden (1966).

The starting point for incorporating unimodality into the bounds is the result due to L. Shepp who gives the following interpretation of Khinchine's famous characterization of unimodality. He shows that a random variable Z

which is unimodal about zero has the representation $Z = UV$ where U and V are independent, and U is uniformly distributed on [0,1]. (See Feller, II, 1971, p. 158). We now use the technique of Kemperman (1971) to transfer the moment problem from the original variable Z to the auxillary variable V. We then find the appropriate bounds for V and then transform back to obtain bounds for Z. To make this more precise, we note that if X is unimodal with mode m, then $Y = X\text{-}m$ is unimodal with mode 0, and hence by Khinchine's theorem $Y = UV$ with U uniform on [0,1]. Using Kemperman's techniques we transform the moment problem on X to one on V by noting that for any function $h, E[h(Y)] = E[h^*(V)]$, where

$$h^*(x) = E[h(UV)|V=x] = \frac{1}{x}\int_0^x h(t)dt$$

In particular, the relationship between the moments of Y and the moments of V is found by taking $h(x) = x^k$. We then have

$$h^*(x) = \frac{1}{(k+1)}x^k;$$

so, $E[V^k] = (k+1)E[Y^k]$. The mean, variance, and skewness of the auxillary variable V are now calculated in terms of the mean, variance and skewness of Y, and hence of X, as follows

$$\begin{aligned}\mu_V &= E[V] = 2E[Y] = 2(\mu_X - m),\\ \sigma_V^2 &= E[V^2) - (E[V])^2 = 3E[Y^2] - 4(E[Y])^2\\ &= 3\sigma_Y^2 - (E[Y]^2 = 3\sigma_X^2 - (\mu - m)^2,\end{aligned}$$

$$\begin{aligned}\rho_V &= E[V-\mu_V]^3 = E[V^3] - 3E[V^2]E[V] + 2(E[V])^3\\ &= 4E[Y^3] - 18E[Y^2]E[Y] + 16(E[Y])^3 \qquad (4.1)\\ &= 4\rho_Y - 6E[Y]\sigma_Y^2 + 2(E[Y])^3\\ &= 4\rho_X - 6(\mu_X - m)\sigma_X^2 + 2(\mu_X - m)^3.\end{aligned}$$

Using this same integral relationship again, we note that finding bounds on $E[h(X)]$ subject to X being unimodal with mode m and given moments is equivalent to finding bounds on $E[h^*(V)]$ subject to the moment constraints (4.1) on V, where

$$h^*(x) = \frac{1}{x}\int_0^x h(t+m)dt.$$

By theorem 2.1 we already know the solution to this transformed problem. Substituting the moments from (4.1) into the formulas describing the bounds in theorem 2.1 yields the results stated below. Note that the end points for X, $a \leq X \leq b$, transform into end points $a^* \leq V \leq b^*$ where $a^* = a-m$ and $b^* = b-m$. Summarizing, we now have the following theorem.

THEOREM

4.1 Suppose X is unimodal with mode m, mean μ_X, variance $\sigma_X{}^2$, third central moment (skewness) ρ_X, and X is bounded between a and b. Let

$$h^*(x) = \frac{1}{x}\int_0^x h(t+m)dt.$$

i) If $h^{(3)}(x)>0$, then the best possible bounds on $E[h(X)]$ using only unimodality, mean and variance are

$$h^*(a-m)q + h^*(\xi_1)(1-q) \leq E[h(X)] \leq h^*(\xi_2)(1-p) + h^*(b-m)p,$$

where

$$\xi_1 = 2\mu_X - 2m - \frac{3\sigma_X^2 - (\mu_X - m)^2}{a+m-2\mu_X},$$

$$q = \frac{3\sigma_X^2 - (\mu_X - m)^2}{3\sigma_X^2 - (\mu_x - m)^2 + (a + m - 2\mu_X)^2},$$

$$\xi_2 = 2\mu_X - 2_m - \frac{3\sigma^2 - (\mu_X - m)^2}{b + m - 2\mu_X},$$

$$p = \frac{3\sigma_X^2 - (\mu_X - m)^2}{3\sigma_X^2 - (\mu_X - m)^2 + (b + m - 2\mu_X)^2}.$$

ii) If $h^{(4)}(x)>0$, then the best possible bounds on $E[h(X)]$ using only unimodality, mean, variance, and skewness are

$$h^*(\eta_1)q + h^*(\eta_2)(1-q) \leq E[h(X)] \leq h^*(a-m)p_1 + h^*(\xi)p_2 + h^*(b-m)(1-p_1-p_2),$$

where p_1, p_2, ξ, η_1, η_2, and q are obtained by substituting $a^* = a - m$ and $b^* = b - m$ together with the moments

μ_V, σ_V^2, and p_V of equations (4.1)

into the formulas of theorem 2.1(c).

We now return to our prototype example involving ruin theory.

V. THE PROBABILITY OF RUIN WHEN THE LOSS VARIABLE IS KNOWN TO BE UNIMODEL

In this section we continue the application of section III; however, now we shall include the information concerning unimodality as outlined in the previous section.

The critical impact of the loss distribution X on the probability of ruin $\psi(u)$ with initial reserve u came through the adjustment coefficient R. This was the point of intersection of the moment generating function $M_X(r)=E(e^{rX})$ with the line $1+(1+\theta)\mu r$. If we know the first few moments of X and its mode m, then we transform the moment problems involved in the calculation of $\mu_X(r)=E[\exp(rX)]$ as follows (compare section IV):

$$M_X(r)=E[e^{rX}]=e^{rm}\ E[e^{r(X-m)}]=e^{rm}Eh^*(V),$$

where

$$h^*(x)=\frac{1}{x}\int_0^x e^{ry}dy=\frac{e^{rx}-1}{rx}$$

The auxillary variable V has moments given by (4.1). We now use theorem 2.1 to bound $E[h^*(V)]$, and hence bound $M_X(r)$. The bounds for $M_X(r)$ correspond to the moment generating functions for the upper and lower principal representations for V given the moments (4.1) multiplied by e^{rm}. Once we have determined the principal representations, we find their adjustment coefficients by finding their intersection with the line $1+(1+\theta)\mu r$. To be explicit, we find the intersection of the curves

$$e^{rm}E_{\nu_0}[h^*(V)]=E_{\nu_0}\left[\frac{e^{r(m+V)}-e^{rm}}{rV}\right]$$

and

$$e^{rm}E_{\nu_1}[h^*(V)]=E_{\nu_1}\left[\frac{e^{r(m+V)}-e^{rm}}{rV}\right]$$

with the line $1+(1+\theta)\mu r$ in order to find the bounds on the adjustment coefficient. This is shown graphically in figure 2.

As a numerical illustration, we return to the example of section III. Here we shall assume additionally that the loss distribution is known to be unimodal with the most likely or modal value $m=37.5$. Our best bounds on the adjustment coefficient R are now obtained by translating the original loss variable moments from X to V, via equation (4.1), then using theorem 2.1 to find the explicit formulas for the principal representations for V, and then calculating the bounds for $M_X(r)$. The corresponding numerical values for

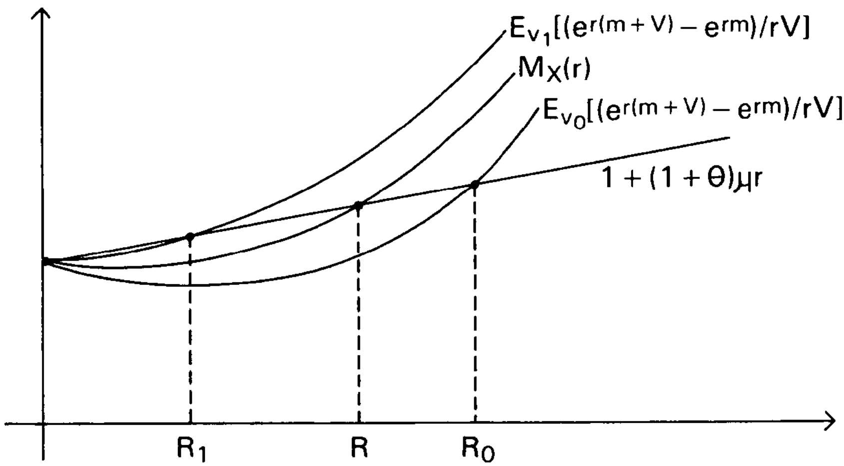

Figure 2:
The best bounding curves for $M_X(r)$ given unimodality, and their corresponding bounds upon the adjustment coefficient.

to find the explicit formulas for the principal representations for V, and then calculating the bounds for $M_X(r)$. The corresponding numerical values for the adjustment coefficient R as determined by the computer using successive bisection to solve (3.2) are given in table 3. Note that in each situation, the bounds obtained by using the unimodality assumption are strictly tighter than those obtained without unimodality. These bounds cannot be improved without further information about the loss variable X. Additionally the bounding extremal measures are of the form $UV + m$ where V has one of the principal representations for its distribution, and U is uniform. Hence, the extremal measures are continuous and unimodal and have the desired moments and mode.

The upper and lower bounds on the adjustment coefficient can be translated into estimates for the initial reserve u needed to insure a probability of eventual ruin of a prespecified size in the same way as outlined in table 2. The calculations involving unimodality will yield more accurate estimates for this initial reserve than would the calculations given in table 2.

VI. SUMMARY

It is very common in risk management and insurance to need to calculate $E[h(X)]$ for some random variable X and some function h. In most cases the statistical distribution for X is not known with certainty, and approximations

TABLE 3

BOUNDS ON THE ADJUSTMENT COEFFICIENT FOR THE LOSS DISTRIBUTION OF TABLE 1*

Premium load Factor θ	Number of moments used	Lower bound $\times 10^4$	Upper bound $\times 10^4$
θ = .1			
	2	3.32	4.35
	3	3.81	3.91
θ = .2			
	2	5.15	8.12
	3	6.21	6.72
θ = .3			
	2	6.37	11.43
	3	7.79	8.87
θ = .4			
	2	7.26	14.38
	3	8.90	10.58

*If the loss is known to be unimodal.

must be made. In this paper we have shown how to use the moments of X, and possible unimodality, to obtain bounds upon $E[h(X)]$ which are as tight as possible or, cannot be improved. The very important problem of risk of ruin calculations with incomplete information concerning the loss distribution was used as an illustration, and the rather easy numerical computations resulted in quite tight bounds on both the probability of ruin, and the required initial reserve.

REFERENCES

1. BOWERS, N.L., H. GERBER, J. HICKMAN, D. JONES AND NESBITT, C. *Risk Theory* study notes for Part 5A Society of Actuaries Examination, Itasca, IL: Society of Actuaries, 1982.
2. BROCKETT, P.L. "Chebychev Systems as a Unifying Technique in some Finance and Economic Problems." Working paper, Department of Finance, University of Texas at Austin, (in preparation 1984)
3. DEPRIL, N., AND GOOVAERTS, M. "Bounds for the Optimal Critical Claim Size of a Bonus System," *Insurance: Mathematics and Economics*, II, 1983, 27-32.
4. DEVYLDER, F. "Maximization, Under Equality Constraints, of a Functional of a Probability Distribution," *Insurance: Mathematics and Economics*, II, 1983, 1-16.
5. FELLER, W. *Introduction to Probability Theory and Its Applications*, II, New York: Wiley Press, 1971.
6. KARLIN, S., AND STUDDEN, W. *Tchebyceff Systems With Applications in Analysis and Statistics*, New York: Interscience, 1966.
7. KEMPERMAN, J.H.B. "Moment Problems with Convexity Conditions." *Optimizing Methods in Statistics*. New York and London: Academic Press, 1971, 115-178.