ACTUARIAL RESEARCH CLEARING HOUSE
1990 VOL. 1

A CLOSER LOOK AT THE ADJUSTMENT COEFFICIENT

Esther Portnoy
University of Illinois
Department of Mathematics
1409 W. Green Street
Urbana IL 61801

ABSTRACT

The adjustment coefficient R plays an important role in mathematical ruin theory, figuring in many elegant probabilistic theorems. Noticeably lacking has been attention to the statistical problems arising when R is estimated from data. One must then be concerned with the variability of the estimator $\hat{R}$ and its implications, for example for estimates of ruin probability. This paper presents several different methods for estimating the variance of $\hat{R}$. Examples are given in which the variance is so high that the estimate of R must be considered essentially worthless, and other means sought for estimating ruin probabilities.

0. Introduction

A standard, simple model in ruin theory assumes a portfolio producing claims of random positive sizes at random times, and with premium flowing in at a constant rate (presumably greater than the rate of aggregate claims). A common objective is to estimate the probability that the maximum aggregate loss, over a finite or infinite period, will exceed a preassigned level. This estimation is facilitated by the use of the adjustment coefficient R.

Under one particularly simple model, we assume that the claims process is compound Poisson. Let $M_X(r)$ be the moment–generating function of the distribution of single claim amount (assumed to exist for at least some positive values of r) and λ the Poisson parameter of the number process. If the rate c of premium flow exceeds $\lambda \cdot E[X]$, then the equation

$$M_X(r) = 1 + cr/\lambda \tag{1}$$

will have a unique nonzero solution R, which is positive; this is the adjustment coefficient. The ruin probability as a function of initial surplus, or equivalently the distribution of the random variable L, maximum aggregate loss, can then be expressed in terms of R. Analogous results hold under broader conditions, but this simple case will serve well to illustrate the issues raised here.

1. The statistical estimator

In practice we do not know the distribution of the claim amount X and the waiting time W between claims, but can only estimate. Suppose that we have, for a particular portfolio, some records of amounts x_i of claims and waiting times w_i between claims, which seem consistent with the assumption that X and W are independent and that W is exponentially distributed. We can use the sample to estimate $M_X(r)$ and λ, and then estimate R by the nonzero[1] solution $\hat{R}$ of the empirical equation

$$\hat{M}_X(r) = 1 + cr/\hat{\lambda} . \tag{2}$$

Reasonable choices would be $\hat{M}_X(r) = \frac{1}{n}\sum_1^n e^{rx_i}$ and $\hat{\lambda} = n/(\sum_1^n w_i)$. Then we can rewrite equation (2) as

$$\sum_1^n (\psi(r;x_i,w_i)) = 0$$

where $\psi(r;x,w) = e^{rx} - 1 - crw$. (3)

Equation (3) makes it clear that $\hat{R}$ is an M–estimator for R. M–estimators are consistent (asymptotically unbiased) and have other attractive properties, to which we will return later.

Table 1 gives a pseudo–random sample of 50 pairs (claim size, waiting time). The mean claim size is 0.9898, the mean waiting time 1.0700, and the sample correlation coefficient 0.01456. These samples were produced by a computer simulations of two simple distributions, but for the moment let us proceed as if this were empirical data.

For any $c > 0.925$, the positive solution of (3) can be found numerically; the values for some selected premium rates are given at the bottom of the table. Since these estimates are based not on definite knowledge of the joint distribution of X and W, but on a sample of data, we need to ask how good they are. One way of answering this question is to estimate $\mathrm{Var}(\hat{R})$.

2. The bootstrap method

The bootstrap [Efron, 1979, 1982] provides an attractive method of estimating $\mathrm{Var}(\hat{R})$. Let (X^*,W^*) be resampled from the empirical distribution; that is, only the pairs appearing in Table 1 can occur, each with probability 0.02. The bootstrap approximation R^* is the (nonzero) solution of equation (3), with X_i^* and W_i^* in the place of X_i and W_i. The central idea of the bootstrap method is that the distribution of R^* provides an approximation to the distribution of $\hat{R}$. Although we know exactly the distribution of (X^*,W^*), it is not feasible to calculate the distribution of R^*, so instead we approximate it using a Monte Carlo simulation.

A few remarks are in order. By resampling from the pairs listed in Table 1, rather than independently from the listed claim amounts and waiting times, we appear to avoid the assumption that X and W are independent; but in fact if they are dependent it is not clear that the solution of equation (3) has any meaning. In the present example, the pseudo–random variables were in fact generated independently, and their correlation is quite low. There would be little change in the results if we resampled independently for X^* and W^*. The decision to sample pairs was made with an eye to later comparisons with cases when X and W might be dependent.

Another matter has to do with the selection of c, the rate of premium flow. In each of the examples of this paper, c has been held constant. It may happen that the resampled data include so many pairs with larger claim sizes and/or smaller waiting times that the implied aggregate claims rate exceeds the premium. Clearly this happens more frequently when the "loading factor" $\theta = (c/\text{expected claims}) - 1$ is relatively small. Just as in real life, we set the premium rate in advance of experience, but consider how to respond when this premium turns out to be inadequate.

Table 1. A pseudo-random sample of (claim amount, waiting time):

(1.967,0.695)	(1.976,2.034)	(0.441,0.127)	(1.140,0.271)	(0.221,3.124)
(0.844,0.233)	(1.480,0.090)	(1.324,0.597)	(0.731,0.094)	(1.141,0.121)
(1.416,2.911)	(1.579,2.164)	(1.915,1.168)	(1.006,1.106)	(0.586,1.996)
(1.571,0.086)	(0.588,1.222)	(1.010,0.259)	(0.483,0.821)	(0.964,0.071)
(1.929,0.330)	(0.009,1.796)	(0.569,2.214)	(1.454,0.107)	(0.095,0.204)
(1.268,0.062)	(1.824,0.602)	(1.341,0.441)	(0.116,0.784)	(0.462,4.328)
(1.714,2.179)	(0.022,0.110)	(1.148,6.015)	(1.590,0.291)	(0.201,0.088)
(0.466,1.078)	(1.288,0.315)	(0.267,0.291)	(1.482,1.082)	(1.769,1.176)
(0.877,3.877)	(0.541,0.969)	(1.932,2.405)	(0.251,0.899)	(1.227,0.507)
(0.929,0.285)	(0.370,1.108)	(1.327,0.026)	(0.340,0.007)	(0.299,0.735)

Premium rate:	1.110	1.388	1.850	2.775
Empirical $\hat{R}$:	0.258	0.547	0.884	1.309

Table 2. Estimates of R and of Var($\hat{R}$), based on the sample of Table 1.

Premium rate c	1.110	1.388	1.850	2.775
500 bootstrap runs:				
Average R^*	0.25487	0.54720	0.88711	1.31548
s.d.(R^*)	0.26224	0.24316	0.22539	0.20930
under truncation:	0.28168	0.54949	0.88721	no
	0.21410	0.23641	0.22497	change
Asymptotic estimate of s.d.($\hat{R}$)				
based on data	0.25729	0.24004	0.22358	0.20838
based on true distribution	0.23600	0.22052	0.20632	0.19367
True value of R	0.15261	0.45585	0.80737	1.24875
500 Monte Carlo runs:				
Average $\hat{R}$	0.13776	0.44498	0.80062	1.24655
s.d.($\hat{R}$)	0.24173	0.22516	0.20997	0.19654
under truncation:	0.18428	0.44759	no	no
	0.17449	0.21922	change	change
Asymptotic estimate of bias	–.01315	–.00948	–.00576	–.00191

If the premium is inadequate, the solutions of equation (3) are $r = 0$ and some negative number. One option is to say that R^* is not defined for these cases; however, simply ignoring them could seriously distort our appreciation of the situation. At a minimum we must keep track of the number of times this occurs. Another tactic is to set $R^* = 0$ (the larger solution); this at least reflects faithfully the effect on ruin probability, which tends to 1 as the adjustment coefficient tends to 0 through positive values and equals 1 when premium is inadequate. However, such truncation results in a larger estimate of the mean of R^*, and a smaller estimate of $\mathrm{Var}(R^*)$, than if we take R^* to be the negative (nontrivial) solution. For the smaller values of c, the difference can be quite important. In this paper R^* will be taken to be the nonzero solution of (3), but results under truncation will also be reported.

Bootstrap estimates of $\mathrm{Var}(\hat{R})$ for various premium rates are given in Table 2. It is clear that there is considerable relative variation in $\hat{R}$, unless the loading factor is rather large. If this is accurate, we must conclude that the estimate of $\hat{R}$ has little value. Since the bootstrap may be a somewhat controversial method, it will be useful to consider some alternatives.

3. Asymptotic methods

As noted earlier, $\hat{R}$ belongs to the class of M–estimators. Such estimators have been treated extensively in the statistics literature; see, for example, [Serfling 1980], chapter 7. In particular, under some mild regularity conditions, M–estimators are consistent (asymptotically unbiased) and asymptotically normal. For the present example,

THEOREM: $\sqrt{n}\,(\hat{R}_n - R)$ converges in distribution as $n \to \infty$ to $N(0,\sigma^2)$, where

$$\sigma^2 = \frac{\mathrm{Var}(e^{RX} - cRW - 1)}{E[Xe^{RX} - cW]^2}.$$

(This is essentially Serfling's Theorem B, page 253. Its proof is based on a first–order Taylor series expansion. Since we will present below a second–order expansion of the same sort, we omit the proof here.)

Thus an asymptotic approximation to $\mathrm{Var}(\hat{R})$ is

$$\frac{1}{n}\frac{\mathrm{Var}(e^{RX} - cRW - 1)}{E[Xe^{RX} - cW]^2} \tag{4}$$

Here R is the true value of the adjustment coefficient, and the variance and expectation are taken under the true joint distribution of (X,W). Not knowing these, we can only estimate, using the data we have. The data of Table 1 lead us to estimates[2] given in Table 2. These are similar in size to the bootstrap estimates made earlier; thus, our suspicion of the unreliability of $\hat{R}$ is reinforced.

There are several difficulties with asymptotic estimates like (4). One is that it is not clear how sensitive they may be to small errors in estimating characteristics of the true distribution. Second, asymptotic estimates are based on finite–order approximations to a Taylor series expansion, which necessarily involve some truncation error. This can be very significant. Finally, the series expansions involve calculations that quickly become tedious, and the possibility of human error is considerable. By comparison, the bootstrap minimizes the chance of mistakes by substituting automatic, computer–assisted calculations for the difficult analysis. In the process it avoids truncation error. (See [Beran 1982] for examples in which the bootstrap is considerably better than competitors for essentially this reason.) Even when (as in this case) the first–order asymptotic formula gives rather good results, the bootstrap accomplishes the same thing more easily.

4. The true distribution of $\hat{R}$

The advantage of dealing with synthetically simulated rather than actual data is that, having suggested several statistical estimators, we can compare their values to the true parameters. As the reader may well have guessed, the X values in Table 1 were generated to be u(0,2) and the W to be exponential with mean 1. The true values of R for these distributions and various premium rates are given in Table 2. Since our estimate of net premium, 0.925 based on the sample of Table 1, was 7.5% low, of course the empirical estimates $\hat{R}$ were high, for every premium level c. The considerable difference between the true values R and our estimates $\hat{R}$ should not be surprising, considering the large estimates we have for $\text{Var}(\hat{R})$. Using the exact characteristics of the joint distribution of (X,W) in the asymptotic formula, we obtain an estimate of the variance that differs little from the previously calculated values.

Direct calculation of $\text{Var}(\hat{R})$ would still be very difficult; but again we can do a Monte Carlo simulation. We draw blocks of 50 pairs (X,W), where X is u(0,2), W is exponential with mean 1, and X and W are independent. For each of many blocks we calculate a value of $\hat{R}$, and then examine their distribution. Summary data for various values of c are included in Table 2.

Two things are shown quite clearly by the Monte Carlo results. First, even though our original $\hat{R}$ was very far from the true value of R, our data–based estimates of $\text{Var}(\hat{R})$ were rather good. Second, the estimator $\hat{R}$ is biased; the difference between the average Monte Carlo value and the true value is much greater than can be attributed to random fluctuation.

Figure 1 shows the empirical distribution function of $\hat{R}$ for $c = 1.110$, based on 500 Monte Carlo runs, and for comparison the empirical distribution function of R^* based on 500 bootstrap runs. Note that the two curves are nearly parallel. This provides visual evidence that, apart from a shift caused by the difference between the empirical and theoretical means, the distribution of R^* is a good approximation to that of $\hat{R}$.

5. Approximating the ruin probability

Consider, for example, setting $c = 1.850$, so that we originally estimate the loading factor at 2. Our original, empirical estimate $\hat{R}$ was 0.88383. For any particular initial surplus level u we might then estimate the ruin probability at $\psi(u) \cong k\, e^{-\hat{R}u}$, using $k = 1$ or some more sophisticated estimate; more to the point, we might have a target figure for $\psi(u)$ and choose u accordingly.

Since we have estimated the standard deviation of $\hat{R}$ at approximately $\frac{1}{4}\hat{R}$, there is a small but not insignificant chance that the true R is only about $\frac{1}{2}\hat{R}$. Then the true ruin probability would be the square root of our estimate. For example, if we chose u in an effort to have $\psi(u) \cong .05$, we would actually have a ruin probability of about 22%. We would need to double u in order to have the desired ruin probability. This is hardly a tolerable situation.

The criticism may well be made that this situation is a consequence of choosing a simple but unrealistic distribution function for X. Tables 3 and 4 give the numerical results when the above analysis is repeated with X distributed as a truncated exponential,

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1 - e^{-x} & 0 \leq x < 10 \\ 1 & 10 \leq x \end{cases}$$

(Truncation is necessary in order that the moment–generating function exist in all the places we need it for the asymptotic approximations.) The relative variation in $\hat{R}$ is still unacceptably large. Note also that $\mathrm{Var}(\hat{R})$ increases with c, whereas in the uniform case it was roughly constant. It should be clear that the exercise of estimating $\mathrm{Var}(\hat{R})$ is worthwhile whenever the adjustment coefficient is being used to estimate the ruin probability.

Empirical Distribution Functions

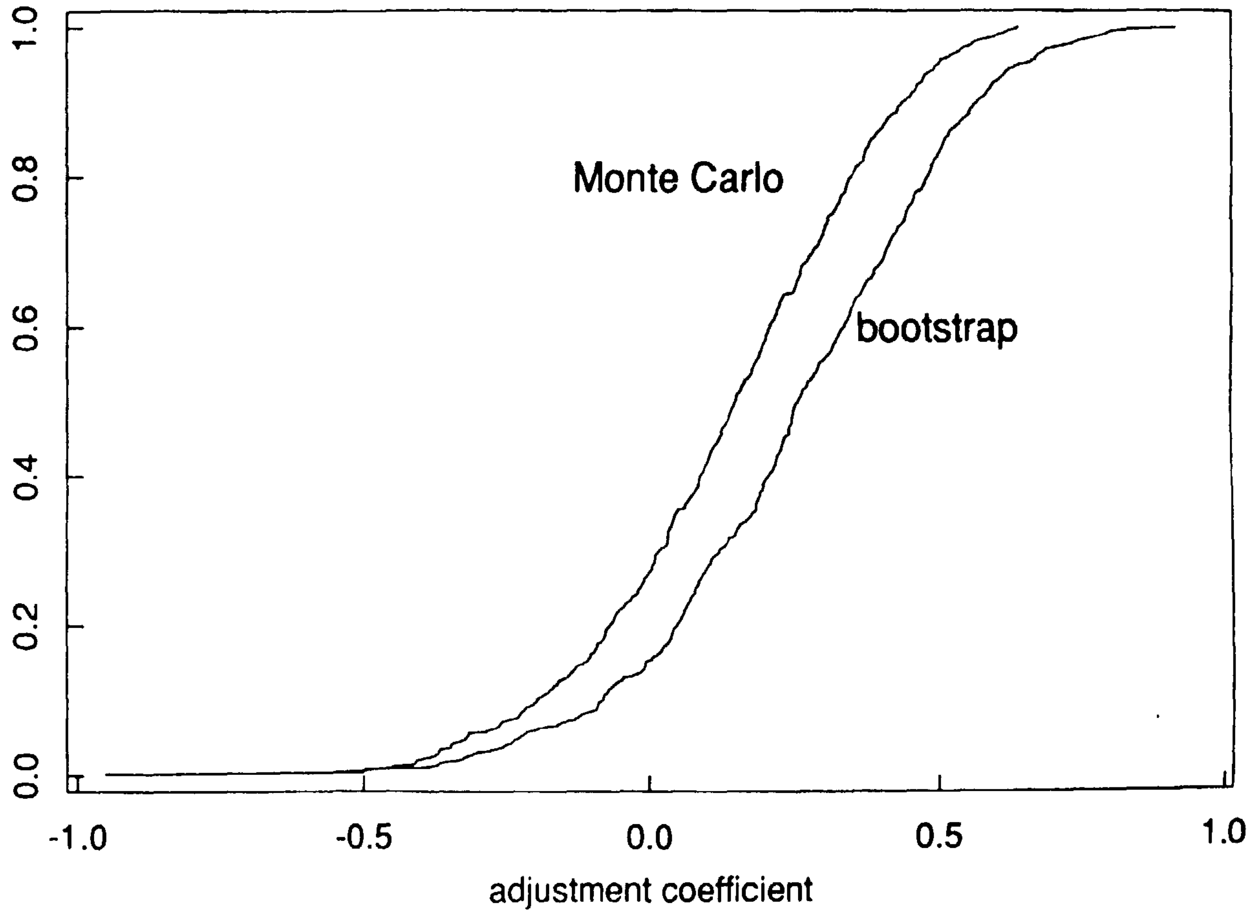

Table 3. A different pseudo–sample of claim sizes and waiting times:

(5.260,0.476)	(3.359,1.623)	(0.058,0.367)	(0.894,0.586)	(0.408,0.988)
(2.421,0.490)	(0.780,1.033)	(1.214,0.275)	(0.048,1.145)	(1.127,1.666)
(0.956,2.174)	(0.086,0.329)	(2.197,1.466)	(0.079,0.136)	(0.972,0.702)
(1.712,1.590)	(1.092,3.772)	(0.428,0.131)	(1.203,0.863)	(0.192,0.699)
(1.291,1.395)	(1.057,1.702)	(0.339,0.350)	((0.085,0.108)	(0.631,0.513)
(0.846,3.455)	(0.961,0.609)	(1.036,1.507)	(1.826,0.889)	(0.932,0.831)
(0.251,3.400)	(1.027,1.209)	(1.766,1.837)	(1.154,0.400)	(0.558,0.423)
(0.021,0.744)	(3.775,0.042)	(1.056,0.252)	(1.964,0.244)	(0.515,2.182)
(2.028,2.698)	(1.509,2.752)	(2.797,1.705)	(0.866,2.415)	(1.009,0.966)
(0.294,0.507)	(0.484,0.683)	(0.016,1.212)	(0.611,0.066)	(1.183,0.313)

Premium rate:	1.210	1.512	2.016	3.025
Empirical $\hat{R}$:	0.164	0.332	0.508	0.708
800 bootstrap runs:				
Average R^*	0.190	0.348	0.546	0.791
s.d.(R^*)	0.179	0.174	0.197	0.236
Asymptotic estimate of sd($\hat{R}$)	0.160	0.158	0.160	0.166

Table 4. Truncated exponential claim amounts:

Premium rate c	1.210	1.512	2.016	3.025
Value of R:	0.174	0.340	0.508	0.683
Asymptotic estimate of sd($\hat{R}$):	0.186	0.184	0.211	0.296
800 Monte Carlo runs:				
average $\hat{R}$	0.199	0.385	0.587	0.822
s.d.($\hat{R}$)	0.211	0.212	0.226	0.257
truncated	0.215	0.386	no	no
	0.187	0.210	change	change
Asymptotic estimate of bias:	0.014	0.046	0.134	0.398

6. Bias of the estimator

On the subject of bias the bootstrap method is essentially worthless; its strength is in estimates of dispersion. The asymptotic expansion, on the other hand, can be developed to give us an estimate of bias, which agrees well with the Monte Carlo results.

Define $\Psi_n(r) = \frac{1}{n}\Sigma\,(e^{rX_i} - 1 - crW_i)$. Note that $\Psi_n(\hat{R}) = 0$ is the defining equation for $\hat{R}$; that $E[\Psi_n(R)] = 0$ for any n; and that

$$\mathrm{Var}(\Psi_n(r)) = \frac{1}{n}[M_X(2r) - M_X(r)^2 + c^2r^2/\lambda^2]$$

for any n, and any r for which $M_X(2r)$ exists. We now write

$$0 = \Psi_n(\hat{R}) = \Psi_n(R) + (\hat{R} - R)\,\Psi_n{}'(R) + \frac{1}{2}(\hat{R} - R)^2\,\Psi_n{}''(\rho),$$

where ρ is between R and $\hat{R}$. Solving the quadratic for $\hat{R} - R$ and doing some algebraic rearranging, we get

$$\hat{R} - R = -\frac{\Psi_n{}'(R)}{\Psi_n{}''(\rho)}\{1 - \sqrt{1 - 2Y}\}$$

$$= -\frac{\Psi_n(R)}{\Psi_n{}'(R)}\left\{1 + \frac{1}{2}Y + \frac{1}{2}Y^2 + \frac{5}{8}Y^3 + \dots\right\}$$

$$\text{where } Y = \frac{\Psi_n{}''(\rho)\,\Psi_n(R)}{\Psi_n{}'(R)^2};$$

$$\frac{1}{\Psi_n{}'(R)} = \frac{1}{E[\Psi_n{}'(R)]}\left\{1 - Z + Z^2 - Z^3 + \dots\right\}$$

$$\text{where } Z = \frac{\Psi_n{}'(R)}{E[\Psi_n{}'(R)]} - 1;$$

$$\text{thus } \hat{R} - R = -\frac{\Psi_n(R)}{E[\Psi_n{}'(R)]}\left\{1 + (\tfrac{1}{2}Y - Z) + (\tfrac{1}{2}Y^2 - \tfrac{1}{2}YZ + Z^2) + \dots\right\}.$$

Both Y and Z are $\mathcal{O}_p(\frac{1}{\sqrt{n}})$. Estimate (4) is obtained by truncating the series above at the first term; we can get better and better estimates by including more terms.

For the moment, let us obtain an asymptotic estimate of the bias, $E[\hat{R} - R]$. The one–term truncation

$$\hat{R} - R \cong -\frac{\Psi_n(R)}{E[\Psi_n'(R)]}$$

is the basis for the statement that $\hat{R}$ is consistent, since $E[\Psi_n(R)] = 0$. Extending the approximation to another term,

$$\hat{R} - R \cong -\frac{\Psi_n(R)}{E[\Psi_n'(R)]}\left\{1 + (\tfrac{1}{2}Y - Z)\right\},$$

we obtain the estimate

$$\text{bias} = E[\hat{R} - R] \cong -\frac{1}{E[\Psi_n'(R)]} E[\Psi_n(R)\cdot(\tfrac{1}{2}Y - Z)]$$

$$\cong -\frac{1}{n\ E[\Psi_n'(R)]^3}\left\{\tfrac{1}{2} M_X'(R)[M_X(2R) - M_X(R)^2 + c^2R^2/\lambda^2] \right. \quad (5)$$

$$\left. - E[\Psi_n'(R)][M_X'(2R) - M_X'(R)M_X(R) + c^2R/\lambda^2]\right\}$$

(At the second " $\cong$ ", additional terms of order $1/n^2$ are omitted.) Substituting into this formula the characteristics of the true distribution[3] we obtain bias estimates given in Table 2. Note that they agree rather well with the difference between the actual value of R and the average of Monte Carlo values.

Lest one assume that the bias will always be negative, as it is in this first example, note that the bias in the truncated–exponential case described in Tables 3 and 4 is positive, and that it increases as c increases.

[1]The fact that $c > \lambda \cdot E[X]$ does not necessarily imply $c > \bar{x}/\bar{w}$ for a particular sample $\{(x_i,w_i)\}$; however, for any fixed c the probability of this event tends to 1 as $n \to \infty$. The function $f(r) = \frac{1}{n} \Sigma (e^{rx_i} - 1 - crw_i)$ vanishes at 0 and has a strictly positive second derivative; thus there is at most one other zero. If $c > \bar{x}/\bar{w}$ then $f'(0) < 0$ and $f(r) \to \infty$ as $r \to \infty$, so there is a unique positive root (and no negative root). If $c < \bar{x}/\bar{w}$, then $f'(0) > 0$ and $f(r) \to \infty$ as $r \to -\infty$, so there is a unique negative root (and no positive root). The case $c = \bar{x}/\bar{w}$, which is technically of probability zero but given the necessity of rounding may have a small positive probability, gives 0 as the only root.

[2]In making these estimates, use was not made of possible simplifications based on the assumptions that W is exponential and that X and W are independent; that is, the values

$$e^{Rx_i} - cRw_i - 1$$

were calculated for each pair (x_i,w_i), using the empirical estimate $\hat{R}$ obtained earlier; and the sample variance was calculated directly. The difference between the value obtained in this way and the quantity

$$\text{sample var } (e^{Rx_i}) + c^2R^2 \text{ sample var } (w_i)$$

was small but nonzero.

[3]One cannot obtain a meaningful estimate of bias from formula (5) without knowledge of the true distribution. If we substitute $\hat{R}$ for R in the asymptotic bias formula, and use empirical estimates for the moment–generating function and for λ, the number obtained is essentially 0.

REFERENCES

Beran, R. (1982) "Estimating sampling distributions: the bootstrap and competitors," Annals of Statistics 10, 212–225.

Efron, B. (1979) "Bootstrap methods: another look at the jackknife," Annals of Statistics 7, 1–26.

________ (1982) The Jackknife, the Bootstrap and Other Resampling Plans. Philadelphia: SIAM.

Serfling, Robert J. (1980) Approximation Theorems of Mathematical Statistics. New York: Wiley.