PARAMETER UNCERTAINTY IN (LOG)NORMAL DISTRIBUTIONS

RODNEY E. KREPS

Abstract

The modeling of parameter uncertainty due to sample size in normal and lognormal distributions with diffuse Bayesian priors is solved exactly and compared to the large-sample approximation. Large-scale simulation results are presented. The results suggest that (1) the large-sample approximation is not very good in this case; and (2) estimates of reserve uncertainty may be considerably understated. A consequence is that intrinsic risk loads and reinsurance premiums may also be considerably understated. An example is given from Best's Homeowners paid data, where the mean estimate of IBNR hardly changes: it is \$9.96B without parameter uncertainty and \$10.01B with it, but the corresponding distribution standard deviations are 6.9% and 24.9% of the respective means.

1. INTRODUCTION

One of the most ubiquitous sources of parameter uncertainty is the fact that samples in real life are never infinite. Thus, when using a sample to estimate parameters of a presumed underlying distribution, the size of the sample must play a role in the uncertainty in the derived values of the parameters. In general, this uncertainty goes to zero as the sample size gets large. The converse, that the uncertainty can be large and even infinite when the sample size is small, is generally unappreciated.

For large samples the parameter distributions can be approximated by normal distributions, using the inverse of the matrix of second derivatives of the negative log-likelihood as the covariance matrix[1]. This is what is usually done for all sample sizes. What is not often understood is how wrong this approximation can be for small samples, say less than 10 data points.

The present paper is an attempt to give both an exact theoretical underpinning[2] and the practical cumulative distribution functions for use with these distributions. Section 2 is the theory; Section 3 is a numerical description of the actual distributions; and Section 4 is a reserving application to stable paid data. Of course, these results apply to any use of normal or lognormal distributions on empirical data. Claim severity distributions would be one example, and especially for reinsurance data the claim volume can be very small.

The general approach here will be to assume that we know the form of the distribution, thus ignoring what is in practice a very real source of parameter uncertainty. What is treated here is only the effect of finite sample size. What is desired is the probability of the parameters, given the observed sample. Given that, the predictive distribution of the variable itself may be obtained by summing over different parameter probabilities. In the present case, this is done using simulation.

The method of treatment is to use a Bayesian approach. The likelihood function gives the probability of the sample actually seen, given the parameters of the underlying distribution. Bayes' theorem says that the desired parameter probability distribution is, up to a normalization, the product of the likelihood function and an assumed prior distribution of the parameters. The assumed prior is here taken to be "diffuse," meaning that it contains as little information as possible in some sense.

---

[1]This results essentially from taking just a second-order Taylor expansion of the negative log-likelihood in the neighborhood of the minimum, as will be done in the special case below. See [1, Section 18.26, page 675].

[2]This particular case is simple enough that it must have been solved many times. However, I am not aware of an actuarial application, and the derivation is instructive.

2. THEORY

We will do the lognormal case, as the normal case is essentially the same with the substitution of x for $\ln(x)$. We are given a sample of data x_i with $i = 1,2,\ldots,n$. The probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi}x\sigma}\exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right\}. \tag{2.1}$$

The corresponding negative log-likelihood (NLL) is, up to constant terms,

$$NLL = \frac{1}{2}\sum_{i=1}^{n}\frac{(\ln(x_i)-\mu)^2}{\sigma^2} + \sum_{i=1}^{n}\ln(x_i) + n\ln(\sigma) + cst. \tag{2.2}$$

The analysis begins by constructing the partial derivatives

$$\frac{\partial NLL}{\partial\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(\mu - \ln(x_i)) = \frac{1}{\sigma^2}\left\{n\mu - \sum_{i=1}^{n}\ln(x_i)\right\} \tag{2.3}$$

and

$$\frac{\partial NLL}{\partial\sigma} = -\frac{1}{\sigma^3}\sum_{i=1}^{n}(\mu - \ln(x_i))^2 + \frac{n}{\sigma}. \tag{2.4}$$

The maximum likelihood estimators are obtained by finding μ_0 and σ_0 such that these partial derivatives are both zero:

$$\mu_0 = \frac{1}{n}\sum_{i=1}^{n}\ln(x_i), \qquad \text{and} \tag{2.5}$$

$$\sigma_0 = \sqrt{\frac{1}{n}\sum_{i=1}^{n}[\ln(x_i) - \mu_0]^2}. \tag{2.6}$$

The usual large sample approximation continues by creating the second partial derivatives:

$$\frac{\partial^2 NLL}{\partial \mu^2} = \frac{n}{\sigma^2}, \tag{2.7}$$

$$\frac{\partial^2 NLL}{\partial \mu \partial \sigma} = \frac{-2}{\sigma^3}\left\{n\mu - \sum_{i=1}^{n} \ln(x_i)\right\} \tag{2.8}$$

$$= \frac{-2n}{\sigma^3}(\mu - \mu_0), \qquad \text{and} \tag{2.9}$$

$$\frac{\partial^2 NLL}{\partial \sigma^2} = \frac{3}{\sigma^4}\sum_{i=1}^{n}(\mu - \ln(x_i))^2 - \frac{n}{\sigma^2} \tag{2.10}$$

$$= \frac{3n}{\sigma^4}\{(\mu - \mu_0)^2 + \sigma_0^2\} - \frac{n}{\sigma^2}. \tag{2.11}$$

Evaluating them at the maximum likelihood (minimum of the NLL),

$$\frac{\partial^2 NLL}{\partial \mu^2}(\mu_0, \sigma_0) = \frac{n}{\sigma_0^2}, \tag{2.12}$$

$$\frac{\partial^2 NLL}{\partial \mu \partial \sigma}(\mu_0, \sigma_0) = 0, \qquad \text{and} \tag{2.13}$$

$$\frac{\partial^2 NLL}{\partial \sigma^2}(\mu_0, \sigma_0) = \frac{2n}{\sigma_0^2}. \tag{2.14}$$

We note in passing that the mixed partial derivative is zero only on the line $\mu = \mu_0$. This means (as will shortly be made explicit) that in general the variables μ and σ are correlated.

The matrix of second-order partial derivatives evaluated at the minimum is

$$\begin{Bmatrix} \dfrac{\partial^2 NLL}{\partial \mu^2} & \dfrac{\partial^2 NLL}{\partial \mu \partial \sigma} \\ \dfrac{\partial^2 NLL}{\partial \mu \partial \sigma} & \dfrac{\partial^2 NLL}{\partial \sigma^2} \end{Bmatrix} = \frac{n}{\sigma_0^2}\begin{Bmatrix} 1 & 0 \\ 0 & 2 \end{Bmatrix}. \tag{2.15}$$

The inverse of this matrix is the covariance matrix for μ and σ around the minimum when they are expressed as a bivariate normal distribution:

$$\begin{Bmatrix} \mathrm{var}(\mu) & \mathrm{cov}(\mu,\sigma) \\ \mathrm{cov}(\mu,\sigma) & \mathrm{var}(\sigma) \end{Bmatrix} = \frac{\sigma_0^2}{n} \begin{Bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{Bmatrix}. \qquad (2.16)$$

A simulation consists of drawing three deviates z, z_1, and z_2 from a standard normal distribution and setting

$$\ln(x) = \mu + \sigma z \qquad (2.17)$$

with

$$\mu = \mu_0 + z_1 \frac{\sigma_0}{\sqrt{n}} \qquad (2.18)$$

and

$$\sigma = \sigma_0 + z_2 \frac{\sigma_0}{\sqrt{2n}}. \qquad (2.19)$$

Equivalently,

$$\ln(x) = \mu_0 + \sigma_0 z_{\mathrm{app}} \qquad (2.20)$$

where the effective z in the large sample approximation, z_{app}, is given by

$$z_{\mathrm{app}} = \frac{z_1}{\sqrt{n}} + z\left(1 + \frac{z_2}{\sqrt{2n}}\right). \qquad (2.21)$$

We note that the distribution for z_{app} is symmetric about the origin, which implies a mean of zero, and that the variance is given by

$$\mathrm{var}(z_{\mathrm{app}}) = 1 + \frac{3}{2n}. \qquad (2.22)$$

It has been pointed out to the author[3] that another approach to a large sample approximation is to use $\ln(\sigma)$ as a variable in place of σ in the *NLL*. Following the same procedure through,[4]

[3]By the reviewer, to whom thanks are given for this remark.

[4]Although the derivation is straightforward, it is somewhat tedious and not relevant for the rest of the paper. Interested readers are invited to contact the author.

Equation 2.20 remains the same but Equations 2.21 and 2.22 become

$$z_{\text{app}} = \frac{z_1}{\sqrt{n}} + z\exp\left(\frac{z_2}{\sqrt{2n}}\right) \tag{2.23}$$

and

$$\text{var}(z_{\text{app}}) = \frac{1}{n} + \exp\left(\frac{1}{n}\right). \tag{2.24}$$

The variable z_{app} has zero mean, but is no longer symmetric. The lack of symmetry is disturbing to the author. However, the variance is larger than before [Equation 2.22] at any n. The increased dispersion of this large sample approximation will be closer to reality.

The underlying technique for the large sample approximation is to approximate the NLL by its Taylor series to second order around the minimum and to take the Bayesian prior to be one (i.e., not dependent on the parameters). However, the resulting simple quadratic form for the NLL is exactly what one gets from a normal (Gaussian) distribution. Hence the remark that, for large samples, the parameter distribution is taken to be normal. The hope is that by the time the NLL deviates significantly from the approximation, its value is sufficiently large that it represents a very small probability.[5]

However, in the present instance this hope is not fulfilled. Returning to the exact problem, the NLL may be rewritten as

$$NLL = n\left\{\frac{\sigma_0^2 + (\mu - \mu_0)^2}{2\sigma^2} + \mu_0 + \ln(\sigma)\right\}. \tag{2.25}$$

Rescale the problem by defining normalized variables v and y such that

$$\mu = \mu_0 + v\sigma_0 \tag{2.26}$$

and

$$\sigma = y\sigma_0. \tag{2.27}$$

[5]The justification for this technique is essentially the same as for the central limit theorem. For a heuristic approach, see the discussion after Equation 2.28.

Then the NLL becomes

$$NLL = n\left\{\frac{1+v^2}{2y^2} + \ln(y) + \mu_0 + \ln(\sigma_0)\right\}. \tag{2.28}$$

The range of v and y is the same as that for μ and σ: $-\infty < v < \infty$ and $0 \leq y < \infty$. It is perhaps not quite obvious, but easy to prove, that the minimum NLL is at $v = 0$ and $y = 1$.

Although the NLL is exactly quadratic in v, it is not so in y. In fact, it is the rather extreme asymmetry in y around the minimum which results in the inadequacies of the large sample approximation. The large sample approximation results from noticing that, as n gets large, only values of v and y which get nearer to the minimum will give NLL values near its minimum. Specifically, one could take NLL of, say, 20 plus the minimum to be the largest value of interest. This corresponds to assuming a probability for the parameters involved of $\exp(-20)$ to be effectively zero. Then as n gets larger the values of v and y which give $NLL = \text{minimum} + 20$ get closer and closer to their minimum values, approximately inversely with the square root of n. This approximation gets better as n increases. In this approximation, terms in the Taylor series expansion of order higher than the second all have contributions to the NLL which decrease as n increases, and the NLL is better and better represented by just the second order term.

We take a Bayesian approach and use diffuse prior distributions for v and y. Since v runs along the full axis from minus infinity to infinity, the prior used is just 1. Since y runs along the semi-axis, the suggested prior is proportional to $1/y^\theta$ where θ is either 0 or 1, depending on one's preference[6]. The choice $\theta = 1$ emphasizes small values of y and corresponds to the assumption that the prior distribution of $\ln(y)$ is flat; the choice $\theta = 0$ assumes that the prior distribution of y is flat. Venter[7] has

---

[6][1, Section 8.28 p. 304]. A reference is made to an article by Jeffries, advocating $\theta = 1$.
[7]Gary Venter, private communication. He points out that on a semi-axis a flat prior corresponds to assuming that it is as likely for the variable to lie between a million and

emphasized that any choice of prior has strong implications. Ideally, the nature of the data being fitted would give some clues as to proper priors.

The joint distribution of v and y is, up to normalization factors (we use the symbol $\sim$), given by the product of the Bayesian priors and the likelihood:

$$f(v,y) \sim \frac{\exp\left\{-n\left(\frac{1+v^2}{2y^2}\right)\right\}}{y^{n+\theta}}. \tag{2.29}$$

We now change variables from y to w by

$$y = \sqrt{\frac{n(1+v^2)}{w}} \tag{2.30}$$

so that

$$\frac{\partial y}{\partial w} = -\frac{1}{2}\sqrt{\frac{n(1+v^2)}{w^3}}, \tag{2.31}$$

and for the variables v and w, the joint distribution behaves as

$$f(v,w) \sim \left\{w^{(n+\theta-3)/2}\exp\left(-\frac{w}{2}\right)\right\}\left\{(1+v^2)^{-(n+\theta-1)/2}\right\}. \tag{2.32}$$

This transformation does several nice things. First, since the joint distribution is a product, the variables are independent (and therefore uncorrelated) and may be simulated separately. A corollary of this is that v and y, and hence μ and σ, are correlated. Second, we can recognize the variable distributions as well known.

The variable w is chi-squared distributed[8] with parameter $(n+\theta-1)$. Equivalently, $w/2$ is gamma distributed [2, p. 104] with parameter $(n+\theta-1)/2$. Both of the inverse functions exist

---

a million and one as it is for the variable to lie between zero and one, and that it is infinitely more likely to be excess of any finite amount than to be less than that amount.

[8]Almost any text on statistics has the chi-squared and t-distributions, e.g., [2, p. 107].

in Excel,[9] and can be used in simulations. The mean value of w is $(n+\theta-1)$, and its variance is $2(n+\theta-1)$. Thus w/n has a mean of $(1+(\theta-1)/n)$ and a standard deviation of $\sqrt{2(n+\theta-1)}/n$. As n, the sample size, becomes large these go respectively to 1 and 0.

The variable $v\sqrt{n+\theta-2}$ is t-distributed [2, p. 145] with parameter $(n+\theta-2)$. Therefore the mean value of v is zero, and its standard deviation[10] is $1/\sqrt{n+\theta-4}$. The standard deviation does not exist if $n+\theta \leq 4$, but goes to zero as the sample size increases.

In simulation situations if the underlying distribution does not have a finite variance then the mean of the simulation will not converge, because the mean of the simulation itself will have an infinite standard deviation. In practice, this shows up as occasional large jumps in the mean, even with millions of simulations (in fact, no matter how many simulations are done). If the simulation is being done in a situation where the upper end is limited—for example in a ceded layer of reinsurance—then the variance will always be finite. However, "finite" does not mean the same as "of reasonable size." In some numerical modeling the author has come across cases where a distribution with finite variance and a theoretical mean of a million dollars was producing an occasional value of a trillion dollars. Clearly, very many millions of simulations would be necessary to get a reasonable amount of convergence. It is recommended that actuaries should try to avoid small sample sizes and/or at least work with lognormal distributions which are truncated at the upper end.

Equation 2.32 shows that as far as v and w are concerned taking $\theta = 1$ is the same as assuming that there is one more data point than actually exists and taking $\theta = 0$. The results in Sections 3 and 4 and Appendix A are all done with $\theta = 0$. If one can convince oneself that an appropriate value of θ is 4, then all

[9]Microsoft Excel 5.0. These functions may also be found elsewhere.

[10]The variance of the Student's t distribution with parameter n is $n/(n-2)$.

worries of convergence are over and as little as one data value can be used. Trying to justify this may take some doing—not to mention getting both a mean and standard deviation from one value!

Another representation for v can be obtained by changing to

$$u = \frac{v^2}{1 + v^2}. \tag{2.33}$$

Clearly, the support of this variable runs from 0 to 1, rather than from $-\infty$ to ∞, but

$$v = \pm\sqrt{\frac{u}{1 - u}} \tag{2.34}$$

can be obtained from a u deviate by another random choice to get the sign. Since

$$\frac{dv}{du} = \frac{1}{2\sqrt{u(1 - u)^3}} \tag{2.35}$$

then

$$f(u) \sim u^{-1/2}(1 - u)^{(n+\theta-2)/2} \tag{2.36}$$

which is recognizable as the beta distribution with parameters 1/2 and $(n + \theta - 2)/2$. Random deviates for the beta distribution can be obtained either from the inverse function in Excel or as a ratio of gamma deviates. Specifically, a beta(α, β) deviate can be obtained [2, p. 139] as $x/(x + y)$ where x is gamma distributed with parameter α and y is gamma distributed with parameter β.

Returning to the simulation methodology, if we let z be a deviate from the standard normal distribution, then in parallel with Equations 2.17, 2.18, and 2.19 for the large sample approximation we have the exact results

$$\ln(x) = \mu + \sigma z \tag{2.37}$$

with

$$\mu = \mu_0 + v\sigma_0, \tag{2.38}$$

and

$$\sigma = \sigma_0 \sqrt{\frac{n(1+v^2)}{w}}. \tag{2.39}$$

Combining Equations 2.37, 2.38, and 2.39

$$\ln(x) = \mu_0 + \sigma_0 z_{\text{eff}} \tag{2.40}$$

where the effective deviate z_{eff}, is given by

$$z_{\text{eff}} = v + z\sqrt{\frac{n(1+v^2)}{w}}. \tag{2.41}$$

Equation 2.41 for z_{eff} is the exact result for which z_{app} of Equation 2.21 is an approximation. Like z_{app}, z_{eff} is symmetric about the origin and has mean zero. This effective deviate generally has a much broader tail than the large sample approximation. However, in the limit of large n (as mentioned earlier) v goes to zero and w goes to n, so that z_{eff} goes to z. In fact, z_{eff} goes to z_{app} to order $1/n$ and they both go to z.

In order to get the variance of z_{eff}, the expectation of $1/w$ is needed. To obtain this, use the fact that for any variable x which is gamma distributed with parameter α, the expectation of any power p of x is

$$\mathrm{E}(x^p) = \frac{\Gamma(\alpha+p)}{\Gamma(\alpha)} \tag{2.42}$$

so

$$\mathrm{E}\left(\frac{1}{x}\right) = \frac{\Gamma(\alpha-1)}{\Gamma(\alpha)} = \frac{1}{\alpha-1}. \tag{2.43}$$

Since $w/2$ is gamma distributed with parameter $(n+\theta-1)/2$,

$$\mathrm{E}\left(\frac{2}{w}\right) = \frac{2}{n+\theta-3}. \tag{2.44}$$

Since the mean of z_{eff} is zero, its variance is just the expectation of its square

$$\mathrm{var}(z_{\text{eff}}) = \mathrm{E}([z_{\text{eff}}]^2). \tag{2.45}$$

Because of the independence of the variables, this implies

$$\mathrm{var}(z_{\mathrm{eff}}) = \mathrm{var}(v) + \mathrm{var}(z) n \mathrm{E}\left(\frac{1}{w}\right)[1 + \mathrm{var}(v)] \tag{2.46}$$

$$= \frac{1}{n+\theta-4} + \frac{n}{n+\theta-3}\left(1 + \frac{1}{n+\theta-4}\right) \tag{2.47}$$

$$= \frac{n+1}{n+\theta-4}. \tag{2.48}$$

In the end, this is a remarkably simple result. Although this variance clearly goes to 1 as n becomes large, for $n = 5$ and $\theta = 0$ its value is 6! Of course, for $n + \theta \leq 4$ it is infinite. This formula also tempts one to choose $\theta = 5$ so that $\mathrm{var}(z_{\mathrm{eff}}) = 1$ for all n.

3. PRACTICE

All of the results for z_{eff} were done using Equation 2.41 with $\theta = 0$ and different values of n. The tables and graphs are useful for getting a feel for how the distributions change with n. If one is uncomfortable with the diffuse prior used, then it is recommended to generate one's own values. It may in the course of simulations be faster to look up values in tables rather than generate them on the fly, but as a matter of general preference the author would rather generate than look up, especially in someone else's tables.

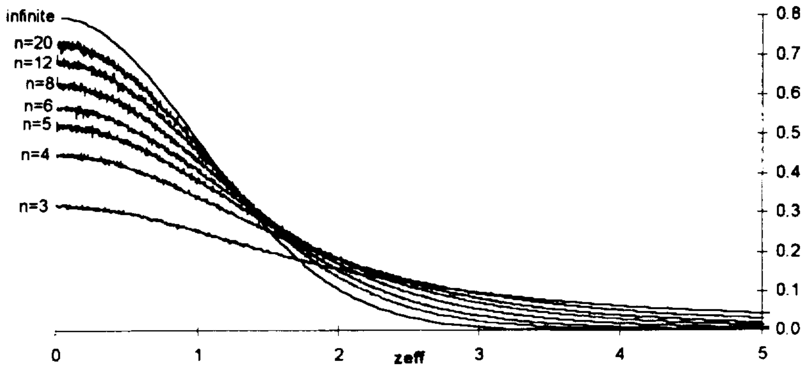
For various values of n, the density function of z_{eff} was simulated in two stages. In the first, 10,000,000 simulations were run to get the range from 50% to 90% on the cumulative distribution function (CDF). Then for values of $n < 10$, 50,000,000 simulations were run to get 5,000,000 simulations of values greater than the 90% level[11] in order to get the tails of the distributions.

Let us look first at the general shape of the density functions. As usual, the effect of parameter uncertainty is to push probability away from the mean out into the tail, and the effect is more

[11]This was not done in a spreadsheet, but in a C++ program.

FIGURE 1

PROBABILITY DENSITY FUNCTIONS FOR DIFFERENT SAMPLE SIZES



pronounced with increasing parameter uncertainty (i.e., decreasing sample size). See Figure 1.

The differences begin to show up dramatically when we look at the Cumulative Distribution Function (CDF) for various sample sizes. Because of the symmetry, only the portion from 50% to 100% is shown in Figure 2.

The extension to even larger z_{eff} is shown in Figure 3. The conclusion from these graphs is at least that the effect of sample size can be substantial even for what might be thought to be relatively large samples.

It is also of interest to compare for a fixed sample size the normal distribution (infinite sample size, no parameter variation), the large sample approximation, and the exact result. Figure 4 displays this comparison for sample size $N = 3$.

Clearly, the large sample approximation is not very good. On the other hand, we didn't expect it to be. However, sample size $N = 8$ shows a similar pattern. See Figure 5.

FIGURE 2

CUMULATIVE DISTRIBUTION FUNCTION FOR VARIOUS SAMPLE SIZES

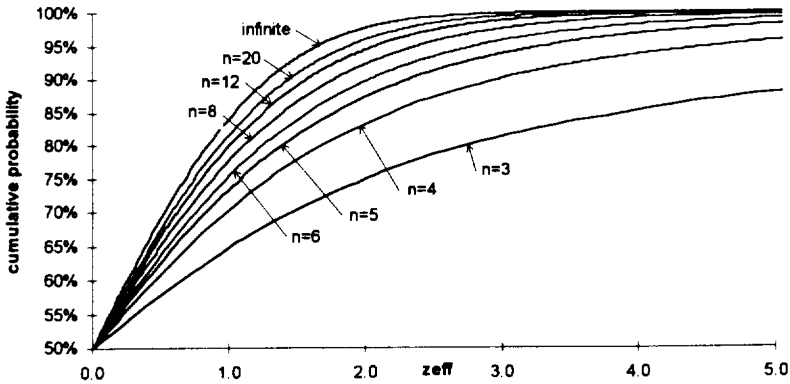


FIGURE 3

CUMULATIVE DISTRIBUTION FUNCTION FOR VARIOUS SAMPLE SIZES

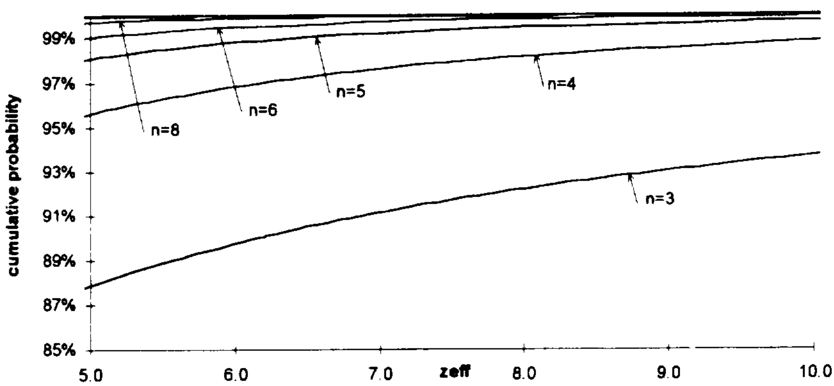


FIGURE 4

n = 3 CUMULATIVE DISTRIBUTION FUNCTION FOR VARIOUS TECHNIQUES

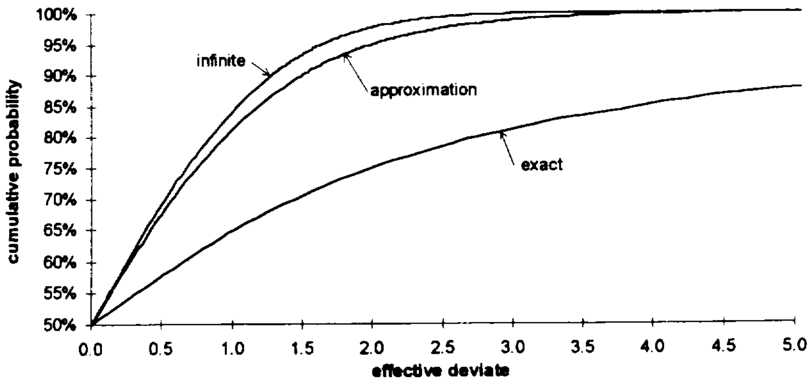


FIGURE 5

n = 8 CUMULATIVE DISTRIBUTION FUNCTION FOR VARIOUS TECHNIQUES

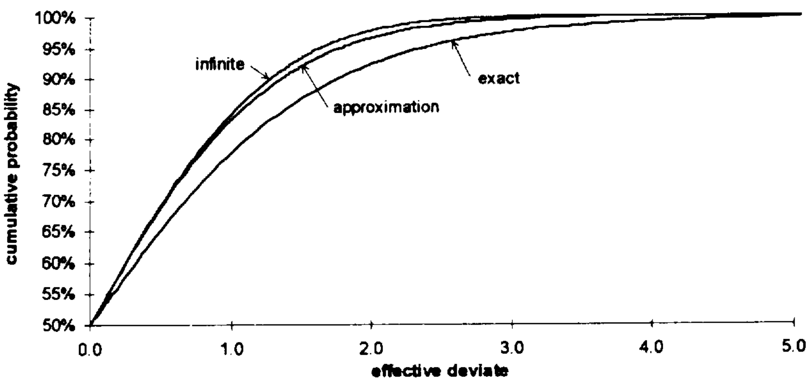


TABLE 1

EFFECTIVE z BY SAMPLE SIZE FOR SOME KEY CDF VALUES

Sample Size	3	4	5	6	8	12	20	Infinite
CDF								
50%	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
60%	0.650	0.456	0.391	0.358	0.325	0.297	0.278	0.253
70%	1.454	0.976	0.826	0.752	0.677	0.618	0.576	0.524
80%	2.752	1.677	1.384	1.245	1.109	1.002	0.931	0.842
90%	6.159	2.981	2.315	2.028	1.762	1.564	1.436	1.282
95.0%	12.62	4.617	3.327	2.819	2.380	2.067	1.873	1.645
97.5%	25.41	6.802	4.501	3.672	2.998	2.541	2.269	1.960
98.0%	31.81	7.666	4.923	3.965	3.200	2.691	2.391	2.054
99.0%	63.71	11.02	6.422	4.956	3.850	3.151	2.757	2.326
99.5%	127.5	15.71	8.260	6.089	4.542	3.614	3.108	2.576
99.9%	**639.3**	**35.35**	**14.47**	**9.526**	**6.392**	**4.726**	**3.897**	**3.090**
99.95%	1,308	50.13	18.46	11.49	7.328	5.247	4.232	3.290
99.99%	6,476	130.4	32.58	17.53	9.822	6.513	5.023	3.719
99.995%	12,470	164.1	42.64	20.75	11.18	7.108	5.371	3.891
99.999%	57,550	345.4	67.41	31.11	14.73	9.353	6.158	4.265

Even here, the large sample approximation is much closer to the pure normal than it is to the exact result, especially in the region of high cumulative probability. The approximation has essentially the same tail behavior as a normal, while the exact result has a much fatter tail. This suggests that the approximation does not hold well for these sample sizes, which are, unfortunately, typical of those usable in chain-ladder reserving.

A complete set of appropriate effective deviates for various CDF values and various sample sizes all at $\theta = 0$ is given in Appendix A. That set is intended for use in simulations if the reader does not want to generate directly the underlying distributions. A subset for some key values of the CDF is given in Table 1.

If we look, for example, at the 99.9% level (in **bold** type), then for n infinite we recognize $z_{\text{eff}} = 3.090$ as a familiar friend from the normal distribution. As the sample size decreases, the location of the 99.9% level increases from 3.09. For $n = 8$ it has more than doubled to 6.4; for $n = 5$ it has almost quintupled to 14.5; and for $n = 3$ it is up to 639! In general, in order to reach any CDF level one must go to increasingly higher multiples of the sigma estimator as the sample size decreases, and the effect is more pronounced as the CDF level increases.

All of the above indicates that the tails are much fatter than one might have thought when using either the large sample approximation to the parameter uncertainty or no parameter uncertainty at all.

4. RESERVING

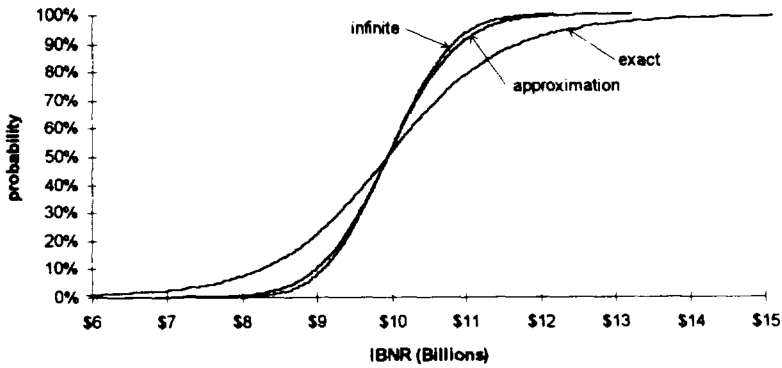
Typically in chain-ladder reserving, the age-to-age factors are implicitly or explicitly taken to be normal or lognormal. For example, the not atypical procedure which we will use here starts by taking the most recent five calendar years of data and averages the logs of the appropriate age-to-age factors in the data to get the log of the projected age-to-age factor. This gives point estimates of the age-to-age factors, which generate the age-to-ultimate factors, which give the IBNR.

Five years is chosen as an intuitive compromise between wanting to stabilize the results by having lots of data and wanting to use only data which is close enough to the current business to be relevant. Clearly there will always be judgment calls of some sort.

In order to go beyond a point estimate of IBNR, the next step is to explicitly assume that the age-to-age factors are lognormally distributed independently at each age. Then we have a sample of five for each age-to-age factor and can calculate the maximum likelihood estimators for both μ and σ. Since the product of lognormal variables is also lognormal, the age-to-ultimate factors are lognormal and their parameters can be easily calculated. This allows the representation of IBNR as a distribution, rather than just a single value.

FIGURE 6

CDF FOR HOMEOWNERS PAID IBNR



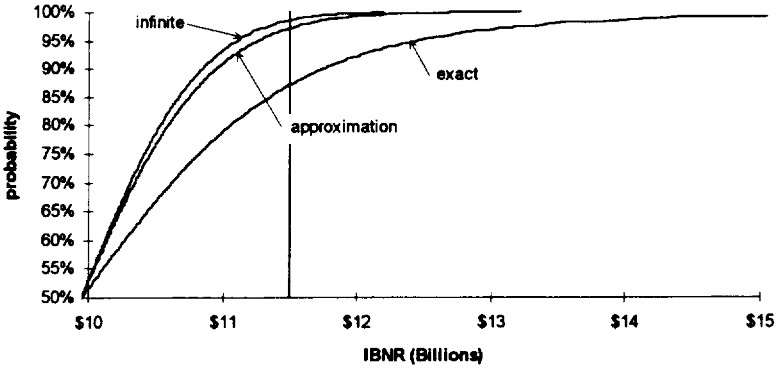
However, this procedure corresponds to using the infinite sample size approximation for the parameter variation—i.e., assuming that there isn't any. Given the above discussion, it will be no surprise that we recommend that the z_{eff} for $n = 5$ be used. It does mean that the distributions for the age-to-age factors must be numerically rather than analytically generated, but this is a relatively minor difficulty.

For a concrete example, we use industry data from Best's 1995 Aggregates and Averages. The original data is Homeowners-Farmowners Schedule P paid data from accident years 1985 to 1994 inclusive, which is displayed in Appendix B. The CDFs are shown in Figure 6, and the labels "infinite," "approximation," and "exact" refer as before to the situations with no parameter variation (infinite sample), the large sample approximation, and the exact result.

An expansion of the dangerous half of the distribution is shown in Figure 7. A line has been put in at $11.5 billion to guide the eye. The probability of exceeding that value is 1.39% for the "infinite" calculation, which would seem a conservative

FIGURE 7

CDF FOR HOMEOWNERS PAID IBNR



reserving level. However, for the approximation the probability is 2.81%, and for the exact result it is 12.78%. To get to the exact 1.39% level, it is necessary to reserve $14.1 billion! These differences are clearly important for a reinsurer. Even for an insurer who reserves at the mean value, the unexpectedly large variability will show up either as an increased risk load cost—probably as cost of liquidity—or as a nasty surprise.

The main simulation results[12] are summarized in Table 2.

It should be noted that even these results are somewhat optimistic (in the sense of providing a small coefficient of variation) in that all factors were taken to have $n = 5$ and in reality the tail of the triangle did not have that much data.

Since this is industry data on a relatively stable line, the 24.9% coefficient of variation for the exact result may be indicative of the minimum reserve variation to be expected.

---

[12]For 1,000,000 simulations in each case. Run times were 10 minutes, 20 minutes, and 40 minutes.

TABLE 2

SIMULATION RESULTS

CDF	Infinite	Approximation	Exact
20%	\$9,377,999	\$9,323,378	\$8,889,821
40%	\$9,775,408	\$9,762,350	\$9,638,914
60%	\$10,121,909	\$10,137,497	\$10,267,019
80%	\$10,530,213	\$10,586,319	\$11,050,725
90%	\$10,839,277	\$10,944,285	\$11,743,068
95%	\$11,097,636	\$11,257,818	\$12,453,300
98%	\$11,393,344	\$11,637,681	\$13,550,822
99.0%	\$11,590,893	\$11,912,637	\$14,599,413
99.5%	\$11,769,344	\$12,172,122	\$16,014,574
99.9%	\$12,144,913	\$12,745,661	\$21,581,916
mean	\$9,956,034	\$9,959,629	\$10,007,938
standard deviation	\$685,580	\$782,023	\$2,489,269
coefficient of variation	**6.9%**	**7.9%**	**24.9%**

REFERENCES

[1] Stuart, Alan and J. Keith Ord, *Kendall's Advanced Theory of Statistics*, Fifth Edition, Volume 2, 1991, Wiley and Sons.

[2] Hogg, Robert V. and Allen T. Craig, *Introduction to Mathematical Statistics*, Fourth Edition, 1978, Macmillan, New York.

APPENDIX A

TABLE OF EFFECTIVE z FOR $\theta = 0$ BY CDF VALUE BY SAMPLE SIZE

Size	3	4	5	6	7	8	10	12	20	Infinite
CDF										
50%	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
51%	0.0628	0.0448	0.0384	0.0354	0.0335	0.0320	0.0304	0.0292	0.0274	0.0251
52%	0.1259	0.0895	0.0769	0.0706	0.0668	0.0641	0.0608	0.0585	0.0550	0.0502
53%	0.1892	0.1344	0.1155	0.1060	0.1002	0.0962	0.0911	0.0879	0.0825	0.0753
54%	0.2531	0.1795	0.1542	0.1414	0.1336	0.1283	0.1216	0.1175	0.1100	0.1004
55%	0.3173	0.2247	0.1931	0.1770	0.1673	0.1606	0.1523	0.1470	0.1376	0.1257
56%	0.3821	0.2702	0.2322	0.2129	0.2010	0.1929	0.1830	0.1766	0.1652	0.1510
57%	0.4476	0.3161	0.2715	0.2488	0.2349	0.2255	0.2138	0.2062	0.1931	0.1764
58%	0.5143	0.3624	0.3109	0.2848	0.2689	0.2583	0.2448	0.2361	0.2210	0.2019
59%	0.5816	0.4089	0.3507	0.3212	0.3032	0.2913	0.2760	0.2662	0.2492	0.2275
60%	0.6503	0.4560	0.3909	0.3579	0.3378	0.3245	0.3074	0.2966	0.2776	0.2533
61%	0.7207	0.5038	0.4315	0.3950	0.3727	0.3580	0.3391	0.3271	0.3061	0.2793
62%	0.7925	0.5525	0.4728	0.4324	0.4079	0.3918	0.3711	0.3579	0.3348	0.3055
63%	0.8665	0.6018	0.5143	0.4703	0.4435	0.4259	0.4034	0.3891	0.3637	0.3319
64%	0.9422	0.6519	0.5566	0.5087	0.4794	0.4603	0.4360	0.4203	0.3929	0.3585
65%	1.0198	0.7030	0.5996	0.5477	0.5161	0.4952	0.4691	0.4522	0.4226	0.3853
66%	1.1003	0.7549	0.6433	0.5873	0.5532	0.5307	0.5023	0.4844	0.4525	0.4125
67%	1.1835	0.8079	0.6876	0.6274	0.5908	0.5664	0.5360	0.5170	0.4829	0.4399
68%	1.2703	0.8623	0.7328	0.6683	0.6290	0.6028	0.5703	0.5501	0.5136	0.4677
69%	1.3598	0.9182	0.7791	0.7100	0.6678	0.6399	0.6053	0.5837	0.5446	0.4958
70%	1.4535	0.9755	0.8262	0.7524	0.7074	0.6775	0.6407	0.6176	0.5762	0.5244
71%	1.5518	1.0345	0.8747	0.7959	0.7476	0.7159	0.6768	0.6522	0.6082	0.5534
72%	1.6547	1.0953	0.9244	0.8402	0.7889	0.7553	0.7135	0.6875	0.6409	0.5828
73%	1.7634	1.1579	0.9754	0.8855	0.8310	0.7953	0.7510	0.7235	0.6742	0.6128
74%	1.8784	1.2229	1.0280	0.9321	0.8743	0.8365	0.7893	0.7602	0.7082	0.6433
75%	2.0002	1.2902	1.0821	0.9800	0.9186	0.8787	0.8285	0.7980	0.7431	0.6745
76%	2.1291	1.3607	1.1379	1.0292	0.9642	0.9221	0.8689	0.8365	0.7788	0.7063
77%	2.2681	1.4341	1.1957	1.0804	1.0112	0.9666	0.9103	0.8762	0.8152	0.7388
78%	2.4171	1.5110	1.2557	1.1331	1.0598	1.0125	0.9530	0.9169	0.8526	0.7722
79%	2.5778	1.5918	1.3184	1.1878	1.1101	1.0601	0.9971	0.9588	0.8911	0.8064
80%	2.7521	1.6767	1.3838	1.2448	1.1623	1.1092	1.0429	1.0021	0.9306	0.8416
81%	2.9420	1.7666	1.4524	1.3042	1.2165	1.1603	1.0901	1.0470	0.9716	0.8779
82%	3.1509	1.8623	1.5245	1.3667	1.2734	1.2136	1.1390	1.0937	1.0140	0.9154
83%	3.3808	1.9642	1.6004	1.4318	1.3329	1.2691	1.1900	1.1422	1.0580	0.9542
84%	3.6372	2.0735	1.6809	1.5008	1.3954	1.3275	1.2433	1.1929	1.1039	0.9945
85%	3.9250	2.1914	1.7669	1.5736	1.4611	1.3888	1.2993	1.2461	1.1519	1.0364
86%	4.2498	2.3191	1.8592	1.6515	1.5313	1.4537	1.3586	1.3021	1.2023	1.0803
87%	4.6222	2.4593	1.9582	1.7343	1.6056	1.5228	1.4210	1.3613	1.2556	1.1264
88%	5.0524	2.6146	2.0665	1.8240	1.6853	1.5967	1.4881	1.4243	1.3121	1.1750
89%	5.5561	2.7866	2.1844	1.9217	1.7717	1.6764	1.5597	1.4915	1.3721	1.2265
90%	6.1588	2.9813	2.3153	2.0283	1.8658	1.7624	1.6376	1.5640	1.4362	1.2816

TABLE OF EFFECTIVE z FOR $\theta = 0$ BY CDF VALUE BY SAMPLE SIZE (*Continued*)

Size	3	4	5	6	7	8	10	12	20	Infinite
CDF										
90.1%	6.2199	3.0024	2.3297	2.0389	1.8757	1.7727	1.6456	1.5720	1.4435	1.2873
90.2%	6.2876	3.0234	2.3437	2.0503	1.8858	1.7819	1.6538	1.5797	1.4503	1.2930
90.3%	6.3568	3.0448	2.3579	2.0618	1.8959	1.7911	1.6621	1.5875	1.4571	1.2988
90.4%	6.4270	3.0665	2.3722	2.0734	1.9061	1.8005	1.6704	1.5953	1.4639	1.3047
90.5%	6.4988	3.0885	2.3867	2.0852	1.9164	1.8099	1.6788	1.6032	1.4709	1.3106
90.6%	6.5723	3.1108	2.4014	2.0970	1.9269	1.8194	1.6873	1.6111	1.4779	1.3165
90.7%	6.6473	3.1336	2.4163	2.1090	1.9374	1.8291	1.6959	1.6191	1.4849	1.3225
90.8%	6.7235	3.1567	2.4313	2.1212	1.9481	1.8388	1.7046	1.6272	1.4920	1.3285
90.9%	6.8017	3.1800	2.4466	2.1335	1.9588	1.8486	1.7134	1.6354	1.4992	1.3346
91.0%	6.8814	3.2037	2.4621	2.1458	1.9696	1.8585	1.7222	1.6436	1.5065	1.3408
91.1%	6.9627	3.2277	2.4777	2.1583	1.9806	1.8686	1.7311	1.6520	1.5138	1.3469
91.2%	7.0456	3.2520	2.4935	2.1710	1.9918	1.8787	1.7401	1.6604	1.5212	1.3532
91.3%	7.1312	3.2768	2.5096	2.1838	2.0031	1.8889	1.7492	1.6688	1.5286	1.3595
91.4%	7.2189	3.3018	2.5258	2.1969	2.0144	1.8994	1.7584	1.6774	1.5361	1.3658
91.5%	7.3084	3.3272	2.5423	2.2101	2.0259	1.9098	1.7678	1.6861	1.5436	1.3722
91.6%	7.3998	3.3531	2.5590	2.2235	2.0375	1.9204	1.7772	1.6948	1.5513	1.3787
91.7%	7.4930	3.3795	2.5759	2.2369	2.0493	1.9311	1.7867	1.7037	1.5590	1.3852
91.8%	7.5892	3.4063	2.5931	2.2506	2.0613	1.9419	1.7964	1.7126	1.5669	1.3917
91.9%	7.6873	3.4335	2.6106	2.2644	2.0734	1.9528	1.8061	1.7217	1.5749	1.3984
92.0%	7.7878	3.4612	2.6284	2.2785	2.0856	1.9639	1.8159	1.7307	1.5828	1.4051
92.1%	7.8913	3.4895	2.6464	2.2926	2.0980	1.9752	1.8258	1.7399	1.5908	1.4118
92.2%	7.9968	3.5182	2.6646	2.3071	2.1106	1.9865	1.8359	1.7492	1.5990	1.4187
92.3%	8.1049	3.5476	2.6831	2.3216	2.1233	1.9980	1.8460	1.7587	1.6072	1.4255
92.4%	8.2158	3.5772	2.7020	2.3364	2.1360	2.0096	1.8563	1.7682	1.6155	1.4325
92.5%	8.3294	3.6076	2.7211	2.3514	2.1490	2.0213	1.8667	1.7778	1.6240	1.4395
92.6%	8.4467	3.6385	2.7404	2.3666	2.1621	2.0331	1.8773	1.7876	1.6325	1.4466
92.7%	8.5661	3.6700	2.7600	2.3821	2.1754	2.0452	1.8879	1.7975	1.6411	1.4538
92.8%	8.6892	3.7020	2.7799	2.3978	2.1888	2.0575	1.8987	1.8075	1.6498	1.4611
92.9%	8.8153	3.7347	2.8003	2.4137	2.2026	2.0699	1.9096	1.8176	1.6586	1.4684
93.0%	8.9453	3.7680	2.8210	2.4299	2.2164	2.0825	1.9207	1.8279	1.6675	1.4758
93.1%	9.0797	3.8019	2.8420	2.4463	2.2304	2.0952	1.9320	1.8384	1.6764	1.4833
93.2%	9.2163	3.8365	2.8634	2.4630	2.2447	2.1082	1.9434	1.8489	1.6855	1.4909
93.3%	9.3579	3.8720	2.8852	2.4800	2.2593	2.1214	1.9549	1.8596	1.6948	1.4985
93.4%	9.5033	3.9081	2.9072	2.4972	2.2740	2.1347	1.9666	1.8704	1.7041	1.5063
93.5%	9.6537	3.9449	2.9295	2.5148	2.2890	2.1482	1.9784	1.8813	1.7136	1.5141
93.6%	9.8088	3.9825	2.9524	2.5326	2.3043	2.1619	1.9904	1.8925	1.7232	1.5220
93.7%	9.9690	4.0211	2.9759	2.5506	2.3198	2.1759	2.0026	1.9037	1.7329	1.5301
93.8%	10.134	4.0605	2.9997	2.5690	2.3356	2.1901	2.0150	1.9152	1.7427	1.5382
93.9%	10.305	4.1007	3.0240	2.5877	2.3516	2.2045	2.0275	1.9267	1.7527	1.5464
94.0%	10.481	4.1420	3.0488	2.6068	2.3678	2.2191	2.0402	1.9386	1.7628	1.5548
94.1%	10.663	4.1843	3.0741	2.6262	2.3842	2.2340	2.0532	1.9504	1.7731	1.5632
94.2%	10.852	4.2276	3.0999	2.6460	2.4009	2.2490	2.0664	1.9624	1.7835	1.5718
94.3%	11.047	4.2721	3.1262	2.6661	2.4179	2.2643	2.0797	1.9748	1.7940	1.5805

TABLE OF EFFECTIVE z FOR $\theta = 0$ BY CDF VALUE BY SAMPLE SIZE (*Continued*)

Size	3	4	5	6	7	8	10	12	20	Infinite
CDF										
94.4%	11.247	4.3174	3.1530	2.6867	2.4354	2.2800	2.0933	1.9872	1.8048	1.5893
94.5%	11.455	4.3643	3.1803	2.7076	2.4533	2.2960	2.1071	2.0000	1.8158	1.5982
94.6%	11.670	4.4121	3.2085	2.7289	2.4713	2.3123	2.1212	2.0129	1.8269	1.6072
94.7%	11.895	4.4614	3.2372	2.7507	2.4898	2.3288	2.1355	2.0260	1.8381	1.6164
94.8%	12.127	4.5119	3.2667	2.7729	2.5087	2.3456	2.1501	2.0394	1.8495	1.6258
94.9%	12.369	4.5636	3.2966	2.7957	2.5279	2.3627	2.1650	2.0530	1.8610	1.6352
95.0%	12.621	4.6168	3.3275	2.8189	2.5477	2.3802	2.1800	2.0668	1.8728	1.6449
95.1%	12.883	4.6718	3.3593	2.8428	2.5684	2.3981	2.1955	2.0809	1.8849	1.6546
95.2%	13.155	4.7282	3.3915	2.8671	2.5889	2.4163	2.2112	2.0953	1.8972	1.6646
95.3%	13.439	4.7863	3.4247	2.8920	2.6098	2.4348	2.2272	2.1100	1.9096	1.6747
95.4%	13.736	4.8458	3.4587	2.9176	2.6313	2.4539	2.2436	2.1250	1.9224	1.6849
95.5%	14.045	4.9072	3.4938	2.9437	2.6532	2.4734	2.2603	2.1402	1.9354	1.6954
95.6%	14.371	4.9711	3.5297	2.9706	2.6758	2.4933	2.2774	2.1558	1.9486	1.7060
95.7%	14.708	5.0372	3.5669	2.9980	2.6988	2.5136	2.2949	2.1717	1.9619	1.7169
95.8%	15.064	5.1054	3.6047	3.0263	2.7224	2.5343	2.3128	2.1880	1.9756	1.7279
95.9%	15.438	5.1759	3.6438	3.0554	2.7464	2.5556	2.3310	2.2046	1.9897	1.7392
96.0%	16.830	5.2490	3.6843	3.0851	2.7710	2.5775	2.3497	2.2217	2.0040	1.7507
96.1%	16.241	5.3246	3.7260	3.1158	2.7964	2.6000	2.3689	2.2390	2.0188	1.7624
96.2%	16.673	5.4032	3.7688	3.1474	2.8226	2.6229	2.3886	2.2570	2.0337	1.7744
96.3%	17.131	5.4848	3.8130	3.1800	2.8494	2.6465	2.4088	2.2752	2.0489	1.7866
96.4%	17.611	5.5698	3.8589	3.2134	2.8771	2.6709	2.4295	2.2940	2.0647	1.7991
96.5%	18.117	5.6577	3.9065	3.2480	2.9055	2.6960	2.4507	2.3133	2.0807	1.8119
96.6%	18.652	5.7498	3.9556	3.2836	2.9348	2.7219	2.4726	2.3332	2.0971	1.8250
96.7%	19.219	5.8455	4.0063	3.3205	2.9652	2.7484	2.4951	2.3536	2.1140	1.8384
96.8%	19.824	5.9463	4.0595	3.3588	2.9967	2.7760	2.5183	2.3744	2.1315	1.8522
96.9%	20.472	6.0513	4.1146	3.3983	3.0290	2.8044	2.5422	2.3958	2.1495	1.8663
97.0%	21.160	6.1607	4.1723	3.4394	3.0626	2.8338	2.5669	2.4180	2.1679	1.8808
97.1%	21.892	6.2756	4.2322	3.4822	3.0974	2.8645	2.5924	2.4409	2.1868	1.8957
97.2%	22.674	6.3965	4.2945	3.5268	3.1336	2.8960	2.6187	2.4647	2.2064	1.9110
97.3%	23.518	6.5244	4.3603	3.5732	3.1711	2.9286	2.6461	2.4892	2.2263	1.9268
97.4%	24.431	6.6590	4.4286	3.6216	3.2101	2.9624	2.6743	2.5147	2.2474	1.9431
97.5%	25.414	6.8024	4.5006	3.6721	3.2507	2.9977	2.7037	2.5410	2.2691	1.9600
97.6%	26.481	6.9538	4.5760	3.7251	3.2934	3.0345	2.7343	2.5686	2.2916	1.9774
97.7%	27.640	7.1139	4.6558	3.7805	3.3379	3.0730	2.7661	2.5972	2.3149	1.9954
97.8%	28.902	7.2850	4.7399	3.8388	3.3848	3.1136	2.7997	2.6272	2.3394	2.0141
97.9%	30.284	7.4690	4.8287	3.9003	3.4342	3.1560	2.8347	2.6583	2.3648	2.0335
98.0%	31.809	7.6656	4.9230	3.9655	3.4863	3.2004	2.8714	2.6908	2.3912	2.0537
98.1%	33.482	7.8775	5.0239	4.0344	3.5411	3.2472	2.9102	2.7251	2.4191	2.0748
98.2%	35.355	8.1065	5.1315	4.1081	3.5996	3.2969	2.9507	2.7612	2.4482	2.0969
98.3%	37.446	8.3553	5.2479	4.1868	3.6613	3.3497	2.9936	2.7993	2.4788	2.1201
98.4%	39.785	8.6267	5.3734	4.2710	3.7274	3.4059	3.0392	2.8396	2.5113	2.1444
98.5%	42.451	8.9249	5.5087	4.3615	3.7983	3.4660	3.0881	2.8826	2.5454	2.1701

TABLE OF EFFECTIVE z FOR $\theta = 0$ BY CDF VALUE BY SAMPLE SIZE (*Continued*)

Size	3	4	5	6	7	8	10	12	20	Infinite
CDF										
98.6%	45.492	9.2531	5.6560	4.4599	3.8742	3.5306	3.1400	2.9282	2.5816	2.1973
98.7%	48.998	9.6177	5.8178	4.5669	3.9573	3.6007	3.1960	2.9771	2.6209	2.2262
98.8%	53.097	10.027	5.9985	4.6837	4.0473	3.6765	3.2571	3.0302	2.6626	2.2571
98.9%	57.907	10.489	6.1984	4.8125	4.1458	3.7592	3.3235	3.0877	2.7077	2.2904
99.0%	63.707	11.019	6.4222	4.9558	4.2560	3.8501	3.3963	3.1507	2.7573	2.3263
99.1%	70.795	11.635	6.6779	5.1181	4.3794	3.9518	3.4770	3.2210	2.8114	2.3656
99.2%	79.707	12.356	6.9732	5.3036	4.5194	4.0677	3.5679	3.2993	2.8716	2.4089
99.3%	91.036	13.230	7.3215	5.5182	4.6799	4.2011	3.6714	3.3882	2.9391	2.4573
99.4%	106.30	14.315	7.7411	5.7747	4.8698	4.3566	3.7915	3.4911	3.0164	2.5121
99.5%	127.54	15.709	8.2600	6.0891	5.1006	4.5425	3.9348	3.6137	3.1085	2.5758
99.6%	159.43	17.595	8.9434	6.4912	5.3913	4.7766	4.1131	3.7633	3.2206	2.6521
99.7%	212.56	20.367	9.8927	7.0398	5.7817	5.0854	4.3460	3.9576	3.3629	2.7478
99.8%	318.87	25.001	11.384	7.8785	6.3601	5.5437	4.6817	4.2374	3.5614	2.8782
99.9%	639.32	35.346	14.466	9.5264	7.4639	6.3924	5.2836	4.7261	3.8973	3.0902
99.91%	710.10	37.358	15.013	9.7723	7.6388	6.5205	5.3910	4.8041	3.9464	3.1214
99.92%	802.48	39.704	15.607	10.091	7.8452	6.6830	5.5026	4.8898	4.0060	3.1560
99.93%	924.73	42.935	16.388	10.453	8.0855	6.8630	5.6142	4.9882	4.0662	3.1947
99.94%	1071.4	46.165	17.200	10.871	8.3560	7.0571	5.7557	5.0962	4.1476	3.2390
99.95%	1308.2	50.132	18.463	11.487	8.6942	7.3285	5.9347	5.2472	4.2324	3.2905
99.96%	1603.7	55.886	19.756	12.104	9.2022	7.6146	6.1392	5.4053	4.3366	3.3528
99.97%	2129.8	64.578	21.719	13.037	9.7135	8.0344	6.4159	5.6194	4.4780	3.4319
99.98%	3195.1	79.327	24.905	14.473	10.582	8.6579	6.8130	5.9433	4.6819	3.5402
99.99%	6476.5	130.35	32.577	17.533	12.262	9.8218	7.5466	6.5133	5.0235	3.7195
99.991%	7155.0	137.10	34.590	17.863	12.547	10.020	7.6674	6.6017	5.0765	3.7462
99.992%	8105.2	143.86	36.603	18.424	13.066	10.219	7.7953	6.7262	5.1366	3.7742
99.993%	9463.0	150.61	38.617	19.038	13.677	10.539	7.9589	6.8534	5.1956	3.8091
99.994%	10820.	157.36	40.630	19.774	14.288	10.861	8.1225	6.9806	5.2681	3.8464
99.995%	12470.	164.12	42.644	20.754	14.899	11.182	8.3533	7.1077	5.3707	3.8906
99.996%	15513.	177.26	44.657	22.133	15.510	11.558	8.6325	7.3289	5.4865	3.9442
99.997%	20567.	204.65	47.133	23.752	16.122	12.166	8.9968	7.5696	5.6387	4.0140
99.998%	30128.	249.52	53.817	26.086	17.333	13.066	9.5083	8.0240	5.8325	4.1071
99.999%	57549.	345.39	67.405	31.115	20.156	14.730	11.495	9.3525	6.1575	4.2655
99.9991%	123124	528.79	89.391	38.011	23.330	16.773	13.385	11.975	6.5736	4.2841
99.9992%	139934	558.35	93.151	39.277	23.970	17.150	13.557	12.113	6.6238	4.3213
99.9993%	152587	606.41	99.113	40.437	24.688	17.528	13.729	12.330	6.6901	4.3400
99.9994%	174499	652.78	104.17	42.105	25.339	17.931	13.901	12.396	6.8068	4.3772
99.9995%	200226	703.32	109.14	43.939	26.214	18.529	14.072	12.651	6.9205	4.4145
99.9996%	237470	771.56	118.18	46.238	27.656	19.410	14.244	12.824	7.0653	4.4703
99.9997%	281976	856.72	128.12	49.570	28.945	20.215	14.416	13.351	7.2843	4.5449
99.9998%	354906	1079.4	145.55	56.462	31.573	21.907	14.588	13.818	7.5316	4.6194
99.9999%	566663	1407.4	176.46	67.100	36.220	24.042	15.525	14.616	8.0696	4.7684
100%	4870750	8260.6	958.00	156.91	121.20	57.244	27.447	19.958	10.705	#NUM!

The last line of the table may seem surprising, as the values should all be infinite, as indicated in the last column. However, in doing simulations it is necessary to have some way of creating very large values. The best way is simply to generate deviates as one needs them. If one is going to use a table such as the above, then a theoretically correct possibility is to create a tail distribution, and simulate off that. A possibility which also works is to have explored the high end in enough detail and to include a value for 100%, in order to interpolate. The values shown here are the largest obtained during the 50,000,000 simulations. Here, the table is reasonably accurate to the one chance in a million level at the high end. If this is not good enough for the problem at hand, then other procedures must be used. This could happen, for example, if many million simulations are to be used, or if results are sensitive to the very high end of the distribution.

APPENDIX B

SCHEDULE P PART 3 HOMEOWNERS-FARMOWNERS PAID DATA FROM BEST'S 1995 AGGREGATES AND AVERAGES

Years in Which Losses Were Incurred	1 12 Months	2 24 Months	3 36 Months	4 48 Months	5 60 Months
1. Prior	0	961,195	1,539,215	1.853,854	2,162,283
2. 1985	7,122,424	9,387,076	9,733,306	9,975,586	10,142,891
3. 1986	6,540,125	8,549,792	8,959,180	9,210,201	9,363,385
4. 1987	6,549,833	9,431,522	9,348,973	9,606,804	9,757,094
5. 1988	7,387,876	9,934,924	10,367,041	10,614,036	10,736,491
6. 1989	9,159,289	12,691,762	13,200,544	13,558,787	13,670,011
7. 1990	9,204,653	12,321,906	12,859,522	13,155,938	13,337,299
8. 1991	10,631,838	13,987,066	14,667,645	15,022,004	
9. 1992	17,421,697	22,112,982	22,871,006		
10. 1993	11,304,871	14,537,267			
11. 1994	13,181,700				

Years in Which Losses Were Incurred	6 72 Months	7 84 Months	8 96 Months	9 108 Months	10 120 Months
1. Prior	2,275,182	2,340,769	2,390,115	2,415,395	2,432,657
2. 1985	10,226,434	10,270,069	10,301,410	10,327,519	10,339,393
3. 1986	9,456,400	9,505,716	9,530,693	9,546,517	
4. 1987	9,858,142	9,914,405	9,943,700		
5. 1988	10,832,847	10,889,518			
6. 1989	13,778,348				
7. 1990					
8. 1991					
9. 1992					
10. 1993					
11. 1994					

job no. 1908 ☐ casualty actuarial society ☐ CAS journal ☐ 1908d03 [2] ☐ 09-17-98 ☐ 12:19 am

SCHEDULE P PART 3 HOMEOWNERS-FARMOWNERS PAID DATA FROM BEST'S 1995 AGGREGATES AND AVERAGES (*Continued*)

Years in Which Losses Were Incurred	LN (Age-to-Age Factors) 1–2	2–3	3–4	4–5	5–6
2. 1985	0.27608573	0.03621976	0.02458710	0.01663236	0.00820287
3. 1986	0.26795068	0.04677175	0.02763297	0.01649520	0.00988489
4. 1987	0.36461793	−0.0087910	0.02720510	0.01552301	0.01030310
5. 1988	0.29621595	0.04257541	0.02354564	0.01147104	0.00893459
6. 1989	0.32618457	0.03930492	0.02677678	0.00816963	0.00789392
7. 1990	0.29166954	0.04270589	0.02278867	0.01369133	
8. 1991	0.27427996	0.04751100	0.02387201		
9. 1992	0.23844847	0.03370514			
10. 1993	0.25148180				

Years in Which Losses Were Incurred	LN (Age-to-Age Factors) 6–7	7–8	8–9	9–10
2. 1985	0.00425781	0.00304704	0.00253130	0.00114908
3. 1986	0.00520154	0.00262413	0.00165894	
4. 1987	0.00569104	0.00295043		
5. 1988	0.00521777			
6. 1989				
7. 1990				
8. 1991				
9. 1992				
10. 1993				

Taking the last five calendar years, which are shaded in the previous table, the results for the maximum likelihood estimators are:

Period	μ_0	σ_0
1 to 2	0.27641	0.03091
2 to 3	0.04116	0.00456
3 to 4	0.02484	0.00180
4 to 5	0.01307	0.00299
5 to 6	0.00904	0.00093
6 to 7	0.00509	0.00052
7 to 8	0.00287	0.00018
8 to 9	0.00210	0.00044
9 to ultimate	0.00115	0.00100

The sigma estimator for 9 to ultimate is, of course, a guess. In the actual calculation, all estimators were taken to have come from a sample of size five calendar years, whereas the last four really have less than that. In reserving practice, since there is always judgment involved in the tail factor and its standard deviation, it seems a good idea to use only estimators which are from at least five calendar years. At least this way the assumptions are made explicit, rather than hidden in factors whose standard deviation is actually infinite due to parameter variation.