1999

Credibility Calculations Using Analysis of Variance Computer Routines

Dennis H. Tolley
Brigham Young University, tolley@byu.edu

Michael D. Nielsen
University of Pennsylvania's Wharton School of Business, mdn2@wharton.upenn.edu

Robert Bachler
Educators Mutual Insurance Association, bachlero@educatorsmutual.com

Follow this and additional works at: http://digitalcommons.unl.edu/joap

Part of the Accounting Commons, Business Administration, Management, and Operations Commons, Corporate Finance Commons, Finance and Financial Management Commons, Insurance Commons, and the Management Sciences and Quantitative Methods Commons

http://digitalcommons.unl.edu/joap/87

This Article is brought to you for free and open access by the Finance Department at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Journal of Actuarial Practice 1993-2006 by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Credibility Calculations Using Analysis of Variance Computer Routines

H. Dennis Tolley, Michael D. Nielsen, and Robert Bachler

Abstract

In this paper we present a method of calculating Bühlmann-Straub credibility factors using standard statistical techniques developed for the analysis of variance. Emphasis is placed on using readily available statistical packages such as SAS and SPSS. Additionally many other computational tools such as EXCEL can be programmed to make such calculations. An example and some sample SAS programs are provided.

Key words and phrases: Bühlmann-Straub credibility factors, empirical Bayes, borrowing strength, random ANOVA model

*Copyright 1999 by the Authors.

†Permission to make digital/hard copy of part or all of this material is granted without fee provided that copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication and its date appear, and notice is given that copying is by permission of the/owners. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires specific permission and/or fee.

Copyright 1999 by American Society of Actuaries.
1 Introduction

Casualty actuaries long have recognized the use of the methods of credibility theory as important in assisting them when setting premiums for (i) renewing business, (ii) blocks of new business, and (iii) determining experience-based refunds. The value of these methods also is gaining recognition among health actuaries.\(^1\) Implementation of these credibility methods, however, is varied. Although formal methods of calculating credibility rates are well established, their implementation varies mathematically from ad hoc computations to simple approximations to detailed estimation of the model parameters. One of the reasons for this is the differences in computational complexity. Despite the fact that company experience is maintained in well-documented databases, use of computer programs on these databases to form credibility estimates is far from seamless and may be too complex to warrant the effort.

We present a method of calculating credibility factors under the Buhlmann-Straub (1970) model using readily available statistical software.\(^2\) The Buhlmann-Straub model is one of a variety of credibility models and is based on a least squares argument. Though the least squares basis for credibility is adequate justification for the procedure, it has been shown that the Buhlmann-Straub method of calculating credibility is identical to the empirical Bayes method when the distribution of losses is a member of the linear exponential family, the loss is quadratic, and when the Bayesian prior used is the conjugate prior for this distribution (Ericson, 1970). Although software programs do not explicitly identify the credibility factors in the software documentation and are not part of the traditional statistical reports generated by these packages, Buhlmann-Straub credibility factors can be calculated from such packages with minimal effort. This paper illustrates these procedures.

A credibility premium uses data from two sources: the estimate of the pure premium based only on the data from a specific group of interest at a specific time and an estimate of the pure premium based on the other data sources and/or prior information. This second estimate may be the overall average of observed rates taken from samples of other groups of policies or the historical average of the group of policies of interest.

\(^1\)There is an extensive literature on credibility in general (see, e.g., Longley-Cook, 1962; Norberg, 1979; Hossack et al., 1983; Herzog, 1996; Goulet, 1998).

\(^2\)For other papers on the Buhlmann-Straub model see, for example, Morris and Slyke, (1978), and Venter (1985, 1990), and Klugman (1987).
The credibility premium classically takes the form

\[ C = ZR + (1 - Z)H, \quad 0 \leq Z \leq 1, \]  

(1)

where \( C \) is the credibility premium; \( R \) is the estimate of pure premium using the data from the group of interest; \( H \) is a global premium (i.e., an exogenous estimate or assumed value of the average of observations); and \( Z \) is the credibility factor and denotes the weight assigned to \( R \). If \( Z = 1 \) then the data are said to be fully credible, and no compromise estimate is needed.

Although the simple form given in equation (1) is found in most of the literature, there are many different approaches to calculate the credibility factor. Bühlmann (1967) arrives at a credibility premium by finding the linear estimator that minimizes the expected squared error. The resulting credibility premium follows the form of the model shown in equation (1), with the credibility factor, \( Z \), given as

\[ Z = \frac{n \times VHM}{n \times VHM + EPV}, \]  

(2)

where \( EPV \) is the expected value of the process variance and refers to the value of the variance of the pure premium within each group, averaged across all groups; and \( VHM \) is the variance of the hypothetical means, which is the mean square distance between the mean of the pure premium in each group and the mean over all groups. Bühlmann (1967) proposes this estimate of credibility for cases when the \( n_i \) are equal. The extension to the case where the \( n_i \) are not equal is presented by Bühlmann-Straub (1970).

2 The Analysis of Variance (ANOVA) Approach

The connection between credibility methods and analysis of variance (ANOVA) has been alluded to in several papers. For example, both Venter (1990) and Morris and Van Slyke (1978) describe a model similar

---

3 Morris and Van Slyke (1978) determine \( Z \) using a Bayesian framework to obtain a form of equation (1). Bühlmann (1970) suggests an alternative method that is also related to the empirical Bayes approach. Herzog (1996), Philbrick (1981), and Venter (1990) also describe this method.

4 Analysis of variance is a standard statistical technique in the design and analysis of experiments. For more on analysis of variance, see, for example, Scheffé (1959) and Neter, Wasserman, and Craig (1990, Part 3.)
to the random one-way analysis of variance model. Dannenburg (1995) uses a one-way random effects model in a cross-classification credibility model that determines the credibility score using estimated variance components. Dannenburg et al. (1996) use the general variance components models of which this is a special case. (See also Goulet, 1998.)

Analysis of variance can be put into the context of the insurance model as follows: Consider an insurance company with $I$ groups of policies. Suppose further that there are $n_i$ individuals from group $i$ who have a claim within a single period (a month, quarter, or year, say). For $i = 1, 2, \ldots, I$, the claim amount, $Y_{iu}$, associated with individual $u$ in group $i$, is modeled as

$$Y_{iu} = \mu + \alpha_i + e_{iu}, \quad u = 1, \ldots, n_i,$$

where $\mu$ represents the mean over all groups and $\alpha_i$ represents the amount that the mean of the $i$th group varies from this overall mean, $\alpha_i$s are mutually independent random variables mean zero and variance $\sigma^2_i$, and the $e_{iu}$s are mutually independent random variables mean zero and variance $\sigma^2_0$. We also assume that $\alpha_i$ and $e_{iu}$ are mutually independent.

If an assumption of normality of the distribution of $\alpha_i$ and $e_{iu}$ were added to equation (3), this would be the standard formulation of the random one-way ANOVA model. This assumption is unnecessary to form the Bühlmann-Straub credibility premium.

Equation (3) implies that the unconditional expected value of $Y_{iu}$ is $\mu$. Conditional on $\alpha_i$, however, the expected value of $Y_{iu}$ is $\mu + \alpha_i$. It is the past experience that provides the basis for improving our estimate of the expected value of $Y_{iu}$, for each group by providing information regarding $\alpha$.

In the ANOVA model of equation (3), the credibility factor is easy to estimate if we use the method of moments estimate of the variance components as suggested by Venter (1990). The method of moments estimate of $\sigma_i^2$ is referred to in the European literature as $\hat{\alpha}$. Other than simplicity and unbiasedness, this method of estimation has no known optimality properties. Other estimates of $\sigma_i^2$ exist with optimality properties, however (see Goulet, 1998; and DeVylder and Goovaerts, 1992). We will use the simple method of moments estimator.
The following notation is used:

\[ N = \sum_{i=1}^{t} n_i; \quad (4) \]

\[ \bar{Y}_i = \text{Average of all observations in group } i; \]
\[ = \frac{\sum_{u=1}^{n_i} Y_{iu}}{n_i}; \quad (5) \]

\[ \bar{Y}_. = \text{Average of all observations, across all groups;} \]
\[ = \frac{1}{N} \sum_{i=1}^{t} \sum_{u=1}^{n_i} Y_{iu}; \quad (6) \]

\[ s_i^2 = \frac{1}{n_i - 1} \sum_{u=1}^{n_i} (Y_{iu} - \bar{Y}_i)^2 \quad (7) \]

\[ \text{MSE} = \frac{1}{N - t} \sum_{i=1}^{t} (n_i - 1)s_i^2, \quad (8) \]

\[ \text{MS}(\alpha) = \frac{1}{t - 1} \sum_{i=1}^{t} n_i(\bar{Y}_i - \bar{Y}_.)^2. \quad (9) \]

The last two expressions are referred to as the mean square for error (MSE) and the mean square for groups (MS(\alpha)), respectively. The expected values of these mean squares are:\footnote{For a derivation of \( E[\text{MSE}] \) and \( E[\text{MS}(\alpha)] \) see Scheffé (1959, Chapter 3) or Neter, Wasserman, and Craig (1990, Chapters 14, pages 543-546).}

\[ E[\text{MSE}] = \sigma_0^2 \]

and

\[ E[\text{MS}(\alpha)] = \sigma_0^2 + n_0 \sigma_1^2, \]

where

\[ n_0 = \frac{N^2}{t - 1} \left( 1 - \sum_{i=1}^{t} \frac{n_i^2}{N^2} \right). \quad (10) \]

In Bühlmann's notation, \( \sigma_0^2 \) is the expected value of the process variance and \( \sigma_1^2 \) is the variance of the hypothetical means. Thus, Bühlmann's \( k \) is given as
\[ k = \frac{n_0 \times \text{MSE}}{\text{MS}(\alpha) - \text{MSE}}. \]

From these expectations we can calculate the following method of moments estimators for the variance components:

\[ \hat{\sigma}_0^2 = \text{MSE}, \]

and

\[ \hat{\sigma}_1^2 = \frac{\text{MS}(\alpha) - \text{MSE}}{n_0}. \] (11)

Thus, for the simple one-way model in equation (3), the Bühlmann-Straub credibility factor, \( Z \), given in equation (2) becomes

\[ Z_i = \frac{n_i}{n_i + k}, \]

\[ = \frac{n_i}{n_i + \hat{\sigma}_0^2 / \hat{\sigma}_1^2}, \]

\[ = \frac{n_i \hat{\sigma}_1^2}{n_i \hat{\sigma}_1^2 + \hat{\sigma}_0^2}, \] (12)

which can be rewritten as

\[ Z_i = \frac{\text{MS}(\alpha) - \text{MSE}}{\text{MS}(\alpha) + (\frac{n_0}{n_i} - 1) \times \text{MSE}}. \] (13)

Most analysis of variance routines calculate MSE and MS(\( \alpha \)). Only the number of observations in the ith group, \( n_i \), and the value of \( n_0 \) need to be determined.

The credibility factor is different for each group depending on the value of \( n_i \). As \( n_i \) increases, \( Z_i \) goes to unity and the group becomes fully credible. On the other hand, as \( \sigma_i^2 \) increases, indicating a high degree of variability from group to group, \( Z_i \) approaches unity and the group becomes fully credible. When \( \sigma_i^2 \) is small relative to \( \sigma_0^2 \) and/or \( n_i \) is small relative to \( n_0 \), \( Z_i \) drops below unity and the group experience is less credible. In this case the compromise estimate borrows more strength from the experience of other groups.
Equation (13) provides a simple calculation of the credibility factor using output from ANOVA routines. Many times, however, the data have been summarized so that for each group \( i \) only the observed pure premium, say \( \hat{Y}_i \), the number insured, \( n_i \), and the standard deviation, \( s_i \), are known. In this case the formulas can be used by first observing that

\[
\bar{Y} = \frac{\sum_{i=1}^{t} n_i \hat{Y}_i}{N}
\]  

(14)

Thus, MS(\( \alpha \)) is calculated as given in equation (9) using \( \bar{Y} \) as given in equation (14). Rearranging the terms in equation (9) yields a formula that is often easier to use. Explicitly,

\[
MS(\alpha) = \frac{1}{t-1} \left( \sum_{i=1}^{t} n_i \hat{Y}_i^2 - N\bar{Y}^2 \right)
\]  

(15)

Second, MSE is calculated as in equation (8).

The credibility factors \( Z_i \) can be calculated using equation (13) where the MSE is given by equation (8) and MS(\( \alpha \)) is calculated using equation (15) with \( \bar{Y} \) as defined in equation (14).

### 3 Calculation of \( Z \) via Computer Programs

#### 3.1 Individual Data Case

To illustrate the formulas and computer programs we consider the hypothetical data given in Table 1. The data sets are small and would not be seriously considered as reliable insurance experience. With such small data sets, however, the details of calculations are more apparent. The data in Table 1 represent four hypothetical groups with claims for each group. We wish to determine the credibility factors for each group assuming that the four groups represent the entire experience of interest for the insurer.

Table 2 gives the EXCEL\(^6\) output for a one-way analysis of variance of the data in Table 1. To obtain this analysis we perform the following steps:

---

\(^6\)EXCEL is a registered trademark of: Microsoft Corporation, One Microsoft Way, Redmond WA 98052-6399, USA.
Step 1: Click the *Data Analysis* menu selection under *Tools*;

Step 2: We then click *One-Way*;

Step 3: As each column represents a different group, we indicate the *Grouped by Columns* option and then proceed.

The output consists of one table (Table 2) with two panels, Panel A and Panel B. The first column in Panel A lists the group name. The second gives the value of $n_i$ for group $i$, where $i$ indicates the column of the group data. The fourth column gives $\bar{Y}_i$ for group $i$ as given by equation (5). The fourth column of Panel B lists the MS($\alpha$) in the first row and the MSE in the second row.

Using the second column of Table 2, Panel A we calculate $n_0$ using equation (10). For this equation $t - 1 = 4 - 1 = 3$. The other components of the equation are given as:

$$N = 22$$
$$\sum n_i^2 = 126, \quad \text{and}$$
$$\quad n_0 = \frac{(22^2 - 126)}{(22 \times 3)} = 5.4242.$$
Table 2
Output from Excel Program of the
One-Way ANOVA Analysis of the Data in Table 1

Panel A: ANOVA Single Factor (Summary)

<table>
<thead>
<tr>
<th>Groups</th>
<th>Count ($n_i$)</th>
<th>Sum</th>
<th>Average ($\bar{Y}_i$)</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>5</td>
<td>8254</td>
<td>1650.800</td>
<td>109582.70</td>
</tr>
<tr>
<td>Group 2</td>
<td>6</td>
<td>13736</td>
<td>2289.333</td>
<td>140929.50</td>
</tr>
<tr>
<td>Group 3</td>
<td>7</td>
<td>12933</td>
<td>1847.571</td>
<td>68661.60</td>
</tr>
<tr>
<td>Group 4</td>
<td>4</td>
<td>5223</td>
<td>1305.750</td>
<td>52624.92</td>
</tr>
</tbody>
</table>

Panel B: ANOVA

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>df</th>
<th>MSE</th>
<th>F-Value</th>
<th>P-Value</th>
<th>F-Crit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between Groups</td>
<td>2527409</td>
<td>3</td>
<td>842469.6</td>
<td>8.853487</td>
<td>0.000805</td>
<td>3.159911</td>
</tr>
<tr>
<td>Within Groups</td>
<td>1712823</td>
<td>18</td>
<td>95156.81</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>4240231</td>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: SS = Sum of Squares; *MSE(α) = Between Groups MSE; F-value = Test statistic to test whether mean costs are the same across groups under the ANOVA assumptions; P-value = Probability of a value greater than or equal to the F-value assuming the means are the same; F-Crit = The value which, if it is exceeded by the F-value, there is statistical evidence that the mean costs differ from between groups.
Using these values we calculate the $Z_i$ for each group using equation (13). Explicitly, for group 1 we have

$$Z_1 = \frac{842469.6 - 95156.81}{842469.6 + (\frac{5.4242}{5} - 1) \times 95156.81} = 0.878631$$

Thus, the credibility score for group 1 is about 87.9 percent. Relative to the complete set of data available, the data on group 1 are relatively credible—there is little difference between the compromise estimate of the group pure premium and the estimate using the observed average of the group.

### 3.2 Grouped Data Case

Suppose that only the summary data consisting of $n_i, \bar{Y}_i,$ and $s_i^2$ for each group are available (columns (2), (4), and (5) of Table 2, Panel A). In this case we can use equations (15) and (8) to calculate the components of equation (13). Explicitly we make the following calculations. First from equation (14) we have

$$\bar{Y}_n = \frac{(5 \times 1650.8 + 6 \times 2289.333 + 7 \times 1847.571 + 4 \times 1305.75)}{22} = 1824.818182.$$  

Using these in equation (15) we obtain

$$MS(\alpha) = \frac{75786559.41 - 73259150.73}{3} = \frac{2527408.68}{3} = 842469.56.$$  

This is close to the value given in Table 2, Panel B (row (1), column (4)). The difference is due to roundoff error.

Calculation of MSE follows similarly using equation (8). Explicitly, we get

$$MSE = \frac{1712822.66}{18} = 95156.81.$$
These results can be used to calculate the credibility scores as before. Computer code for the same calculations using SAS are given in the appendix; no code is provided for SPSS.\(^7\)

## 4 Discussion

We have illustrated how the Bühlmann-Straub credibility factors can be calculated using one-way ANOVA statistical routines common in many computer programs. In order to form such scores the mean squares reported in the ANOVA tables must be used as given in equation (13). Under certain situations estimated $\text{MS}(X)$ can be negative. In this case the value of $Z_i = 0$ is used. This reduces the bias of the compromise estimate as shown by Morris (1983).

## References


\(^7\)SAS is a registered trademark of: SAS Institute Inc., Cary, NC 27512-8000, USA; and SPSS is a registered trademark of: SPSS Inc., 444 North Michigan Avenue, Chicago IL 60611, USA.


Venter, G.G. "Structured Credibility in Application—Hierarchical, Multidimensional and Multivariate Models." *Actuarial Research Clearing...*
Appendix

The codes for making credibility calculations using SAS for the data in Table 1 are given below. First we use the individual data. We have used the cards option. In practice one would read a data file. Below we give the code for grouped data. In both cases the amount of work to get the SAS code seems long relative to the simple problem considered. For longer, more practical problems, however, the benefits of SAS routines are more apparent.

DATA costs;
  INFILE cards;
  INPUT cost group;
  CARDS;
  1550 1
  1325 1
  1417 1
  1824 1
  2138 1
  1879 2
  2028 2
  2150 2
  2245 2
  2516 2
  2918 2
  1440 3
  1601 3
  1790 3
  1852 3
  1998 3
  2081 3
  2171 3
/*** Getting number of individuals per group ***/
PROC SQL;
CREATE TABLE counts AS
SELECT DISTINCT group,count(group) AS number
FROM costs
GROUP BY group;
RUN;

/*** Calculating n_not ***/
PROC SQL;
SELECT (sum(number)-(sum(number**2)/sum(number)))/(count(number)-1)
INTO :n_not
FROM counts;

/*** Calculating MSE, MSA ***/
PROC ANOVA DATA=costs OUTSTAT=results NOPRINT;
CLASS group;
MODEL cost=group;
RUN;

DATA _null_
SET results;
mean_sqr=ss/df;
SELECT (_source_);
WHEN ("ERROR") CALL SYMPUTC("MSE",mean_sqr);
WHEN ("GROUP") CALL SYMPUTC("MSA",mean_sqr);
END;
RUN;

/*** Calculating credibilities ***/
DATA creds;
SET counts;
cred=(&MSA-&MSE)/(&MSA+(&n_not/number-1)*&MSE);
KEEP group cred;
RUN;
PROC PRINT NOOBS DATA=creds;
  TITLE 'Credibility Factors for Individual Data';
RUN;

/**************************
USING GROUPED DATA
**************************/

DATA grouped;
  INFILE cards;
  INPUT group number avg_cost var_cost;
CARDS;
  1 58 1666 49597893
  2 115 5051 216276545
  3 81 4670 193990984
  4 108 8966 757144094
;RUN;

/*** Calculating n_not and the overall mean ***/
PROC SQL;
  SELECT (sum(number)-(sum(number)**2)/sum(number))
    /(count(number)-1),
    sum(avg_cost*number)/sum(number)
INTO :n_not,:y_bar2
FROM grouped;

/*** Calculating MSE, MSA ***/
PROC SQL;
  SELECT 1/(count(group)-1)*(sum(number*avg_cost**2)
    -sum(number)**2*year_bar2**2),
    1/(sum(number)-count(group))*sum((number-1)
    *var_cost)
INTO :msa,:mse
FROM grouped;

/*** Calculating credibilities ***/
DATA creds;
  SET grouped;
  cred=(&msa-&mse)/(&msa+(&n_not/number-1)*&mse);
KEEP group cred;
time period. Recent examples of collective risk modeling in insurance include Butler, Gardner, and Gardner (1998); Butler and Worall (1991); and Cummins and Tennyson (1996).

The stochastic structure is two-pronged: both the size of the individual claims and the number of claims are considered random variables. Specifically, let $S$ denote the aggregate claims random variable, i.e.,

$$S = \sum_{i=1}^{N} X_i$$

where $N$ is the number of claims and $X_i$ is the size of the $i$th individual claim. The $X_i$s are assumed to be mutually independent and identically distributed (i.i.d.) and are mutually independent of $N$. In the literature equation (1) is referred to as a compound random variable; see, for example, Bowers et al. (1997, Chapter 12).

Theoretically, the distribution of $S$ can be obtained from equation (1) as follows:

$$\Pr[S \leq s] = \sum_{n=0}^{\infty} p_n F^{*n}(s)$$

where $p_n = \Pr[N = n]$ and $F^{*n}(s) = \Pr[X_1 + \ldots + X_n \leq s]$, i.e., $F^{*n}(s)$ is the $n$th convolution of the $X_i$s, with $F(x) = F^{*1}(x)$ being the cumulative distribution function of $X_1$.

The difficulty in using equation (2), however, often lies in calculating $F^{*n}(s)$. Thus, approximations are frequently used. There are several approximations used by actuaries, including discretizing the claim size distribution (Panjer 1981); using the Wilson-Hilferty approximation or Haldane Type A approximation (Pentikäinen, 1987); and, of course, the normal approximation. See Panjer and Willmot (1992, Chapter 6) and Bowers et al. (1997, Chapters 2 and 12) for a discussion of the actuarial approaches. Other methods such as the Edgeworth expansion (Feller, 1971) or the conjugate density method (Esscher, 1932) have been applied.

The methods mentioned above provide good approximations near the center of the distribution but can be slow or inaccurate for calculating tail probabilities of the form $\Pr[S > s]$ (for large values of $s$). For a discussion of the tail behavior of aggregate distributions; see Panjer and Willmot (1992, Chapter 10). A fairly quick and accurate method of calculating tail probabilities is the so-called saddlepoint approximation.
Since their introduction by Daniels (1954) saddlepoint approximations have been utilized to approximate tail probabilities in a variety of situations; see, for example, Goutis and Casella (1999), Huzurbazar (1999), Butler and Sutton (1998), Tsuchiya and Konishi (1997), and Wood, Booth, and Butler (1993). Field and Ronchetti (1990) document the accuracy of these procedures for small sample sizes (even of sample size one). In this paper a saddlepoint approximation is developed for $\Pr[S > s]$ and is applied to specific examples.

2 The Saddlepoint Approximation

The key assumption in the saddlepoint approximation is the assumption of the existence of the moment-generating functions corresponding to $X_i$ and $N$, which are denoted by $M_X(\theta)$ and $M_N(\theta)$, respectively, where $\theta$ is a real valued parameter.\(^1\) The moment-generating function of $S$, $M_S(\theta)$, is then given by

\[
M_S(\theta) = E[e^{\theta S}]
= E[E[e^{\theta S}|N]]
= M_N(\log(M_X(\theta))).
\] (3)

Equation (3) can be used to derive the well-known results on the moments of compound sums of i.i.d. random variables:

\[
\mu_S = E[S] = E[N]E[X_1] \tag{4}
\]

\[
\sigma_S^2 = \text{Var}[S] = \text{Var}[N](E[X_1])^2 + E[N]\text{Var}[X_1]. \tag{5}
\]

The saddlepoint approximation for the tail probability $\Pr[S > s]$ is adapted from Field and Ronchetti (1990) for sample size one. First let $T$ denote the standardized random variable

\[
T = \frac{S - \mu_S}{\sigma_S}
\]

\(^1\)The moment-generating function of a random variable $Z$ is defined as

\[
M_Z(\theta) = E[e^{\theta Z}], \ \ \theta > 0.
\]
where $\mu_S$ and $\sigma_S$ and the mean and standard deviation of $S$ respectively (which can be obtained from equations (4) and (5)). The moment-generating function for $T$ is easily seen to be:

$$M_T(\theta) = e^{-\mu_S/\sigma_S} M_S(\theta/\sigma_S). \quad (6)$$

For a fixed value of $s$, let $t = (s - \mu_S)/\sigma_S$ and let $\beta$ be the solution to the equation

$$M_T'(\beta) = t M_T(\beta) \quad (7)$$

where the $'$ denotes differentiation with respect to $\theta$. Note that $\beta$ is a function of $t$. Further, let

$$c = \frac{e^{\beta t}}{M_T(\beta)} \quad (8)$$

and

$$\sigma^2 = \frac{M_T''(\beta)}{M_T(\beta)} - t^2. \quad (9)$$

The saddlepoint approximation for $P(S > s)$ is:

$$Pr(S > s) \approx 1 - \Phi(\sqrt{2\ln(c)}) + \frac{1}{c\sqrt{2\pi}} \left[ \frac{1}{\beta \sigma} - \frac{1}{\sqrt{2\ln(c)}} \right] \quad (10)$$

where $\Phi(\cdot)$ is the standard normal distribution function, and $c$ and $\sigma$ are defined in equations (8) and (9).

In practice, once $s$ is chosen and $t$ is computed, equation (7) is solved numerically using a technique such as Newton’s method or the secant method; see, for example, Burden and Faires (1997, Chapter 2).

3 Examples

The saddlepoint approximations of tail probabilities are now applied to several specific collective risk models. These saddlepoint approximations are compared to the Haldane Type A and the normal approximations, and the exact probabilities. The exact calculations are found
by simulation using 10,000 repetitions, which gives accuracy to four decimal places.

If $X$ has mean $\mu_X$, standard deviation $\sigma_X$, and coefficient of skewness $\gamma_X$, then the Haldane Type A approximation is as follows:

$$Pr[X \leq x_0] \approx \Phi \left[ \left( (1 + r \bar{x}_0)^h - \mu(h, r) \right) / \sigma(h, r) \right]$$

(11)

where

$$\bar{x}_0 = \frac{(x_0 - \mu_X)}{\sigma_X}$$

(12)

$$r = \frac{\sigma_X}{\mu_X}$$

$$h = 1 - \frac{\gamma_X}{3r}$$

(13)

$$\mu(h, r) = 1 - \frac{1}{2} h (1 - h) (1 - \frac{1}{4} (2 - h)(1 - 3h)r^2) r^2$$

(14)

$$\sigma(h, r) = hr \sqrt{1 - \frac{1}{2} (1 - h)(1 - 3h)r^2}.$$  

(15)

The Haldane approximation is chosen because Pentikainen's (1987) results show it to be, under certain circumstances, an accurate approximation. Recall that the normal approximation is

$$Pr[X \leq x_0] \approx \Phi[\bar{x}_0].$$

(16)

The relative errors shown in the tables are calculated as:

$$\text{Relative Error} = \frac{|\text{Approximation} - \text{Exact}|}{\text{Exact}}.$$

3.1 Light and Medium Tailed Claim Size Distributions

Example 1: $X_1$ is normal random variables with mean $\mu_X = 100$ and standard deviation $\sigma_X = 10$ while $N$ is Poisson with mean $\lambda = 10$. From equation (3)

$$M_S(\theta) = \exp \left[ \lambda \left( \exp \left( \mu_X \theta + \frac{1}{2} \sigma_X^2 \theta^2 \right) - 1 \right) \right].$$

(17)
Table 1
Approximating Tail Probabilities for
The Compound Normal–Poisson Model

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\beta$</th>
<th>Exact</th>
<th>Normal</th>
<th>HALDA</th>
<th>SADP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.4637</td>
<td>0.2964</td>
<td>0.0411</td>
<td>0.0039</td>
<td>0.0034</td>
</tr>
<tr>
<td>1.0</td>
<td>0.8672</td>
<td>0.1575</td>
<td>0.0074</td>
<td>0.0077</td>
<td>0.0070</td>
</tr>
<tr>
<td>1.5</td>
<td>1.2243</td>
<td>0.0750</td>
<td>0.1089</td>
<td>0.0062</td>
<td>0.0087</td>
</tr>
<tr>
<td>2.0</td>
<td>1.5445</td>
<td>0.0303</td>
<td>0.2498</td>
<td>0.0125</td>
<td>0.0082</td>
</tr>
<tr>
<td>2.5</td>
<td>1.8347</td>
<td>0.0112</td>
<td>0.4469</td>
<td>0.0019</td>
<td>0.0089</td>
</tr>
<tr>
<td>3.0</td>
<td>2.1001</td>
<td>0.0036</td>
<td>0.6351</td>
<td>0.0091</td>
<td>0.0084</td>
</tr>
</tbody>
</table>

In this setting the central limit theorem is known to hold for large $\lambda$.

Example 2: $X_1$ is a gamma random variable with a mean of $\mu_X = 100$ and standard deviation $\sigma_X = 10$. $N$ is a negative binomial random variable with mean of $\alpha = 10$ and standard deviation $\gamma = 20$. Here

$$M_S(\theta) = \left[\frac{1 - q}{1 - q(1 - \beta - \delta)}\right]^r \tag{18}$$

where $q = 0.5$, $\mu_X = \beta \delta$, $\sigma_X = \beta \sqrt{\delta}$, $\alpha = r q / (1 - q)$ and $\gamma = r q / (1 - q)^2$.

Table 2
Approximating Tail Probabilities for
The Compound Gamma–Negative Binomial

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\beta$</th>
<th>Exact</th>
<th>Normal</th>
<th>HALDA</th>
<th>SADP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.4284</td>
<td>0.2684</td>
<td>0.1494</td>
<td>0.0961</td>
<td>0.0417</td>
</tr>
<tr>
<td>1.0</td>
<td>0.7502</td>
<td>0.1548</td>
<td>0.0252</td>
<td>0.0284</td>
<td>0.0032</td>
</tr>
<tr>
<td>1.5</td>
<td>1.001</td>
<td>0.0796</td>
<td>0.1608</td>
<td>0.0515</td>
<td>0.0050</td>
</tr>
<tr>
<td>2.0</td>
<td>1.203</td>
<td>0.0375</td>
<td>0.3920</td>
<td>0.6907</td>
<td>0.0027</td>
</tr>
<tr>
<td>2.5</td>
<td>1.369</td>
<td>0.0166</td>
<td>0.6265</td>
<td>0.3012</td>
<td>0.0084</td>
</tr>
<tr>
<td>3.0</td>
<td>1.508</td>
<td>0.0070</td>
<td>0.8086</td>
<td>0.4571</td>
<td>0.0100</td>
</tr>
</tbody>
</table>
Example 3: $X_1$ is an inverse Gaussian random variable with mean $\mu_X = 100$ and standard deviation $\sigma_X = 10$. $N$ is Poisson with mean $\lambda = 10$. The moment-generating function for the inverse Gaussian distribution is

$$M_X(\theta) = \exp \left[ \left( \frac{\mu_X}{\sigma_X} \right)^2 \left( 1 - \left( 1 - \frac{2\sigma_X^2 \theta}{\mu_X} \right) \right) \right],$$

see Johnson and Kotz (1970, Chapter 15). Hence

$$M_S(\theta) = \exp \left[ \lambda \left( \left( \frac{\mu_X}{\sigma_X} \right)^2 \left( 1 - \left( 1 - \frac{2\sigma_X^2 \theta}{\mu_X} \right) \right) - 1 \right) \right].$$

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\beta$</th>
<th>Exact</th>
<th>Normal</th>
<th>HALDA</th>
<th>SADP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.4537</td>
<td>0.2998</td>
<td>0.0290</td>
<td>0.0153</td>
<td>0.0147</td>
</tr>
<tr>
<td>1.0</td>
<td>0.8671</td>
<td>0.1629</td>
<td>0.0258</td>
<td>0.0258</td>
<td>0.0264</td>
</tr>
<tr>
<td>1.5</td>
<td>1.2242</td>
<td>0.0775</td>
<td>0.1381</td>
<td>0.0387</td>
<td>0.0413</td>
</tr>
<tr>
<td>2.0</td>
<td>1.5444</td>
<td>0.03 16</td>
<td>0.2785</td>
<td>0.0285</td>
<td>0.0348</td>
</tr>
<tr>
<td>2.5</td>
<td>1.8345</td>
<td>0.0119</td>
<td>0.4790</td>
<td>0.0588</td>
<td>0.0672</td>
</tr>
<tr>
<td>3.0</td>
<td>2.0998</td>
<td>0.0038</td>
<td>0.6474</td>
<td>0.0526</td>
<td>0.0526</td>
</tr>
</tbody>
</table>

These examples show that the saddlepoint approximation is superior to the central limit theorem, but seems to be on par with the Hal-dane approximation in calculating tail probabilities. Next we consider a more difficult setting involving heavy tailed distributions.

4 Heavy Tailed Claim Size Distributions

The saddlepoint approximation requires the existence of the moment-generating function of the claim variable. For heavy tailed distributions, such as the Pareto (the moment-generating function does not exist)
and lognormal (the moment-generating function is not in convenient a closed form), an approximation is required. For these problem cases a censoring limit is imposed on the claim distribution.

For cases where the moment-generating function does not exist, the distribution of the claim variable is approximated utilizing an upper tail censoring limit. For small $\epsilon$ the censoring limit, $L$, satisfies $\Pr[X_i > L] = \epsilon$. Let us define the censored claim random variable as

$$Y_i = \begin{cases} X_i & \text{if } X_i \leq L \\ L & \text{if } X_i > L. \end{cases}$$

The distribution function for the $Y_i$'s is now

$$F_Y(x) = \begin{cases} F(x) & \text{if } x < L \\ 1 & \text{if } x \geq L. \end{cases}$$

The corresponding moment-generating function is

$$M_Y(\theta) = \int_{x=0}^{L} e^{\theta x} dF(x) + e^{\theta L}. \quad \text{(20)}$$

The saddlepoint approximation is applied using the censoring moment-generating function in equation (20). This technique is now demonstrated on two examples of heavy tailed claim distributions. In both cases the number of claims is assumed to be Poisson with mean 5.

**Example 4:** Claims are assumed to follow a lognormal distributed with probability density function (pdf) of $X_1$ is

$$f(x) = \frac{1}{\sqrt{2\pi} \sigma} \exp\left[-\frac{1}{2} \left(\frac{\ln(x) - \mu}{\sigma}\right)^2\right] \quad -\infty < x < \infty. \quad \text{(21)}$$

where $\mu = 0$ and $\sigma = 1$. We assume that $\epsilon = 0.001$, which produces a censoring limit of $L = 59.7697$.

**Example 5:** Here we assume the claim size follows a Pareto distribution with distribution function given by

$$F(x) = 1 - \frac{1}{(1 + x)^3}. $$
Table 4
Approximating Tail Probabilities for The Compound Lognormal-Poisson Model

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\beta$</th>
<th>Exact</th>
<th>Normal</th>
<th>HALDA</th>
<th>SADP</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>0.7251</td>
<td>0.1628</td>
<td>0.8950</td>
<td>0.5565</td>
<td>0.0498</td>
</tr>
<tr>
<td>1.0</td>
<td>0.9501</td>
<td>0.0630</td>
<td>1.5190</td>
<td>1.3016</td>
<td>0.0825</td>
</tr>
<tr>
<td>1.5</td>
<td>1.0512</td>
<td>0.0241</td>
<td>1.7718</td>
<td>2.3361</td>
<td>0.0622</td>
</tr>
<tr>
<td>2.0</td>
<td>1.2001</td>
<td>0.0108</td>
<td>1.1111</td>
<td>1.0463</td>
<td>0.1574</td>
</tr>
<tr>
<td>2.5</td>
<td>1.4211</td>
<td>0.0047</td>
<td>0.3191</td>
<td>5.5319</td>
<td>0.3404</td>
</tr>
</tbody>
</table>

Again, $\epsilon = 0.001$, and this produces a censoring limit of $L = 9.0$.

As in the previous section, normalized tail probabilities and the saddlepoint approximations are compared to the exact values as obtained by simulation. These computations are listed in Tables 4 and 5.

Table 5
Approximating Tail Probabilities for The Compound Pareto-Poisson Model

<table>
<thead>
<tr>
<th>$t$</th>
<th>$\beta$</th>
<th>Exact</th>
<th>Relative Error</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Normal</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6959</td>
<td>0.1664</td>
<td>0.8540</td>
</tr>
<tr>
<td>1.0</td>
<td>0.9880</td>
<td>0.0688</td>
<td>1.3067</td>
</tr>
<tr>
<td>1.5</td>
<td>1.1623</td>
<td>0.0327</td>
<td>1.0428</td>
</tr>
<tr>
<td>2.0</td>
<td>1.2842</td>
<td>0.0165</td>
<td>0.3818</td>
</tr>
<tr>
<td>2.5</td>
<td>1.3772</td>
<td>0.0094</td>
<td>0.3404</td>
</tr>
</tbody>
</table>

For the heavy tailed distributions, the saddlepoint approximation is superior to the central limit theorem and the Haldane approximation in calculating tail probabilities.
References


