ACTUARIAL RESEARCH CLEARING HOUSE
1992 VOL. 1

A METHOD TO DETERMINE CONFIDENCE INTERVALS FOR TREND

by William A. Bailey

ABSTRACT

The method involves resampling (with replacement but without random numbers), numerical convolutions for sums and quotients, and the estimating of confidence intervals for trend in average size claim. Starting with an original sample of comprehensive major medical claims (per claimant) for each of two calendar years, we use numerical convolutions for sums to generate distributions of average size claim (per claimant) for resamples of various sizes from each of the two calendar years, use numerical convolutions for quotients to generate distributions of trend (in average size claim per claimant) from the first to the second of the two calendar years, note certain stabilities in standardized versions of these distributions, and estimate confidence intervals for the underlying trends.

TABLE OF CONTENTS

ABSTRACT

TABLE OF CONTENTS

INTRODUCTION

BOOTSTRAPPING FOR TREND IN AVERAGE SIZE CLAIM

Resampling (With Replacement) Using Random Numbers

Table #0

OPERATIONAL BOOTSTRAPPING FOR TREND IN AVERAGE SIZE CLAIM

Resampling (With Replacement) Without Random Numbers

CONFIDENCE INTERVALS FOR TREND - LARGE SIZED RESAMPLES

Table #1

A Numerical Example of Determining a Confidence Interval for Trend Using Large Resamples

Test of Normality Assumptions

CONFIDENCE INTERVALS FOR TREND - MEDIUM SIZED RESAMPLES

A Numerical Example of Determining a Confidence

Interval for Trend Using Medium Sized Resamples

Table #2

Test of Normality Assumptions

Table #3

RECAP AND CONCLUSIONS

APPENDIX - UNIVARIATE GENERALIZED NUMERICAL CONVOLUTIONS

A Univariate Generalized Numerical Convolution Algorithm

REFERENCES

ACKNOWLEDGMENTS

INTRODUCTION

Suppose for a given accident year we have n claims with severities

$$x'_1,\ x'_2,\ x'_3, \ldots,\ x'_n;$$

and suppose for a later accident year we have m claims with severities

$$y'_1,\ y'_2,\ y'_3, \ldots,\ y'_m;$$

If the coverage is a type for which inflationary trends are significant, we might want to estimate the trend from the earlier to the later of the given accident years.

An estimate $\hat{t}$ of the true trend ϑ in severity could be obtained from the ratio of the average claim severities in the later accident year to the average claim severities in the earlier accident year; namely,

$$\hat{t} = \frac{\frac{1}{m} \cdot \sum_{i=1}^{m} y_i'}{\frac{1}{n} \cdot \sum_{i=1}^{n} x_i'} - 1.$$

If the given accident years are s years apart, then the annual trend might be estimated by

$$(1+\hat{t})^{1/s} - 1.$$

While useful, $\hat{t}$ is a single point estimate for the true severity trend ϑ and gives no indication of the uncertainty involved in the estimate. In order to try to measure the degree of statistical uncertainty involved in this estimate, we begin by reinterpreting our data.

Instead of considering the set of values

$$\left\{ \; x_i' \; \right\}_{i=1,n}$$

to be the experience for the earlier of the given accident years, we treat it as a sample † of n claims drawn from the population of all claims that could have occurred in that accident year.

---

† *This sample will be referred to as the original sample for this accident year.*

We let the empirical distribution f_X of severities X for the earlier of the two given accident years be expressed as

$$f_X = \left(x_i \quad p_i \right)_{i=1,n'}$$

where n′ is the number of different severities in the set

$$\left(x'_i \right)_{i=1,n}$$

and p_i is the relative frequency of x_i for $i=1,2,\ldots,n'$. Clearly, n >= n′.

Similarly, the set of values

$$\left(y'_i \right)_{i=1,m}$$

can be treated as a sample † of m claims drawn from the population of all claims that could have occurred in the later of the two given accident years; and we let the empirical distribution f_Y of the severities Y for that accident year be expressed as

$$f_Y = \left(y_i \quad \tilde{p}_i \right)_{i=1,m'}.$$

---

† This sample will be referred to as the original sample for the later of the two given accident years.

where m′ is the number of different severities in the set

$$\left\{ y_i' \right\}_{i=1,m}$$

and $\tilde{p}_i$ is the frequency of y_i for i=1,2,...,m′. Clearly, m >= m′.

We can estimate the distribution $f_{\hat{T}}$ of resample point estimates $\hat{T}$ for the true severity trend ϑ as follows:

(1) sample n times from the distribution f_X, summing the results and dividing by n, to obtain a possible average size claim (say a) from the earlier of the two given accident years;

(2) sample m times from the distribution f_Y, summing the results and dividing by m, to obtain a possible average size claim (say b) from the later of the two given accident years;

(3) calculate $\hat{t} = \frac{b}{a} - 1$, which is a trial resample point estimate of the true severity trend ϑ;

Repeating steps (1) through (3) many (say ν) times produces an approximation to the distribution $f_{\hat{T}}$ of possible sample point estimates $\hat{T}$ of the true severity trend ϑ.

We now describe this classical simulation process in more detail. Afterward, we will offer a more efficient method (the generalized numerical convolution) for simulating the distribution of resample point estimates.

BOOTSTRAPPING FOR TREND IN AVERAGE SIZE CLAIM

Resampling (With Replacement) Using Random Numbers

The cumulative empirical distributions for the two given accident years are

$$\left(x_i \quad \sum_{k=1}^{i} p_k \right)_{i=1,n'} \quad \text{and} \quad \left(y_j \quad \sum_{k=1}^{j} \tilde{p}_k \right)_{j=1,m'}$$

respectively. The resampling (with replacement) from the original samples would involve the following steps:

(1) Generate a random number, say r, and determine i such that

$\sum_{k=1}^{i} p_k$ is the cumulative probability nearest to r.

Look up x_i and add it to an accumulator.

(2) Repeat Step (1) n times.

(3) Divide the resulting accumulation by n, to obtain the average size loss per claimant, and call the result a;

(4) Perform Steps (1) through (3) again, but using $\sum_{k=1}^{j} \tilde{p}_k$ instead of $\sum_{k=1}^{i} p_k$ and y_j instead of x_i in Step (1), and m instead of n in Steps (2) and (3), and call the result b;

(5) Calculate $\hat{t} = \frac{b}{a} - 1$, which is a possible point estimate $\hat{t}$ for the true trend ϑ;

(6) Repeat Steps (1) through (5), say, ν times.

Let the frequency distribution of the resulting values of $1+\hat{T}$ be labelled as $f_{1+\hat{T}}$ and represented as

$$\left\{ 1+\hat{t}_k \quad r_k \right\}_{k=1,\nu'}$$

where ν' is the number of different point estimates $\hat{t}$ obtained in Step (5), and r_k is the frequency of $1+\hat{t}_k$ for $k=1,2,\ldots,\nu'$. Now $f_{1+\hat{T}}$, once generated, could be used to estimate the standard error in trend or other such statistics. This procedure is referred to as bootstrapping.†

If we are going to use this approach, it would be helpful to know how large ν should be in order to produce a reasonably good representation of what the distribution of $1+\hat{T}$ would be if ν were chosen to be infinity. **Table #0** shows results of this approach using $\nu=10^3$, 10^4 and 10^5 trial resample point estimates, m=n=64, and the accident year pair is 1983-84. The last column of **Table #0** shows results from an almost exact representation of what the distribution $f_{1+\hat{T}}$ of $1+\hat{T}$ would be if ν were chosen to be infinity.‡

† For a detailed description of bootstrapping see Efron and Tibshirani ref[1]. Efron coined the term "bootstrapping" in the late 1970's.

‡ For the method used to obtain this distribution, see section "OPERATIONAL BOOTSTRAPPING FOR TREND IN AVERAGE SIZE CLAIM."

Table #0

$$f_{1+\hat{T}}$$

cumula-tive	$\nu=10^3$ $1+\hat{T}$	$\nu=10^4$ $1+\hat{T}$	$\nu=10^5$ $1+\hat{T}$	$\nu=$ infinity $1+\hat{T}$
.000001				0.109
.00001			.146	0.146
.0001			.206	0.179
.001			.264	0.247
.01	0.378	0.392	.389	0.375
.025	0.472	0.467	.462	0.457
.05	0.544	0.535	.532	0.526
.1	0.636	0.622	.622	0.618
.2	0.748	0.737	.742	0.737
.3	0.838	0.834	.840	0.837
.4	0.938	0.925	.930	0.932
.5	1.022	1.017	1.025	1.028
.6	1.125	1.123	1.130	1.135
.7	1.236	1.243	1.256	1.261
.8	1.418	1.409	1.424	1.431
.9	1.723	1.687	1.706	1.723
.95	2.037	1.968	2.000	2.027
.975	2.398	2.323	2.331	2.370
.99	2.964	2.871	2.833	2.940
.999	4.068	4.433	4.274	4.844
.9999		6.458	5.618	6.635
.99999			7.299	8.360
.999999				10.239
mean	1.121	1.061	1.055	1.066
$var \cdot 10^{+3}$	.244	.204	.204	.204
n	64	64	64	64
m	64	64	64	64

This resampling procedure is practical if ν=n and m are each small. However, as ν, n and/or m increase, this procedure becomes impractical. So, we turn to a method which we call "Operational Bootstrapping."

OPERATIONAL BOOTSTAPPING FOR TREND IN AVERAGE SIZE CLAIM

Resampling (With Replacement) Without Random Numbers

In contrast to classical bootstrapping, where random numbers are used to do the resampling, we can use numerical convolutions to generate the distributions without using any random numbers. For example, consider the distribution

$$f_{X_1+X_2}$$

of

$$X_1 + X_2$$

where X_1 and X_2 are independent identically distributed random variables, each distributed as

$$\begin{bmatrix} x_i & p_i \end{bmatrix}_{i=1,n'}.$$

Symbolically, we can express $f_{X_1+X_2}$ in terms of the distributions of X_1 and X_2, as follows:

$$f_{X_1+X_2} = \ = f_{X_1} + f_{X_2}{}^{\dagger}$$

$$= \begin{bmatrix} x_i & p_i \end{bmatrix}_{i=1,n'} + \begin{bmatrix} x_j & p_j \end{bmatrix}_{j=1,n'}$$

---

† The symbol + between two distributions is being used here to mean convolute for sums.

$$
= \begin{bmatrix}
x_1+x_1 & p_1 \cdot p_1 \\
x_1+x_2 & p_1 \cdot p_2 \\
& \vdots \\
x_1+x_{n'} & p_1 \cdot p_{n'} \\
x_2+x_1 & p_2 \cdot p_1 \\
x_2+x_2 & p_2 \cdot p_2 \\
& \vdots \\
x_2+x_{n'} & p_2 \cdot p_{n'} \\
& \vdots \\
x_{n'}+x_1 & p_{n'} \cdot p_1 \\
x_{n'}+x_2 & p_{n'} \cdot p_2 \\
& \vdots \\
x_{n'}+x_{n'} & p_{n'} \cdot p_{n'}
\end{bmatrix}
$$

which we might express as

$$
\begin{bmatrix} x_i+x_j & p_i \cdot p_j \end{bmatrix}_{i=1,n' ;\, j=1,n'}
$$

Using this distribution as our prototype and assuming that $X_1, X_2, \ldots, X_n$ are independent identically distributed random variables each distributed as

$$
\begin{bmatrix} x_i & p_i \end{bmatrix}_{i=1,n'}
$$

we can proceed recursively to generate

$$f_{X_1+X_2+X_3+X_4}$$

where, since $f_{X_3+X_4} = f_{X_1+X_2}$, we can write

$$f_{X_1+X_2+X_3+X_4} = f_{X_1+X_2} + f_{X_3+X_4} = f_{X_1+X_2} + f_{X_1+X_2} ;$$

and continue to perform convolutions between the results of other convolutions until we have obtained the desired result; namely,

$$f_{X_1+X_2+\ldots+X_n} .$$

Proceeding naively in this manner, the number of lines in the resulting distributions could become prohibitively large from the standpoints of both computer storage and computing time. The **APPENDIX - UNIVARIATE GENERALIZED NUMERICAL CONVOLUTIONS** describes a method of overcoming this problem. This method (after dividing the amounts by n) produces a distribution having mean equal to the mean of the original sample and variance equal to the variance of the original sample.

We can similarly generate the distribution

$$f_{Y_1+Y_2+\ldots+Y_m}$$

of

$$Y_1 + Y_2 + \ldots + Y_m,$$

where

$$Y_1 \ , \ Y_2 \ , \ \ldots \ , \ Y_m$$

are independent identically distributed random variables, each distributed as

$$\left[\begin{array}{cc} y_i & \tilde{p}_i \end{array} \right]_{i=1,m'} .$$

To generate the distribution

$$f_{1+\hat{T}} = f_{(1/m)(Y_1+Y_2+\ldots+Y_m)} \ / \ f_{(1/n)(X_1+X_2+\ldots+X_n)}$$

of

$$(1/m)\cdot(Y_1 + Y_2 + \ldots + Y_m) \ / \ \left((1/n)\cdot(X_1 + X_2 + \ldots + X_n) \right),$$

we can first generate

$$f_{Y_1+Y_2+...+Y_m} / f_{X_1+X_2+...+X_n}{}^{\dagger} \quad .$$

Letting

$f_{X_1+X_2+...+X_n}$ be represented as $\begin{bmatrix} u_i & p_i \end{bmatrix}_{i=1,n}$.

and

$f_{Y_1+Y_2+...+Y_m}$ be represented as $\begin{bmatrix} v_i & \tilde{p}_i \end{bmatrix}_{i=1,m}$.

we have

$$f_{(Y_1+Y_2+...+Y_m)/(X_1+X_2+...+X_n)}$$

$$= \begin{bmatrix} v_j & \tilde{p}_j \end{bmatrix}_{j=1,m} / \begin{bmatrix} u_i & p_i \end{bmatrix}_{i=1,n}.$$

---

† The symbol / between two distributions is being used to mean convolute for quotients, dividing the first random variable by the second.

$$f_{V/U} = \left[\quad v_j/u_i \quad p_i \cdot \tilde{p}_j \quad \right]_{i=1,n^*;\, j=1,m^*}$$

Then, $f_{1+\hat{T}} = f_{(1/m)(Y_1+Y_2+...+Y_m)/(\ (1/n)(X_1+X_2+...+X_n)\ )}$ would be obtained by multiplying the amounts (not the probabilities) in the distribution $f_{(Y_1+Y_2+...+Y_m)/(X_1+X_2+...+X_n)}$ by n/m.

The distributions $f_{1+\hat{T}}$ generated by the methods of this section are representations of the distribution $f_{1+\hat{T}}$ which would have been generated by the method of the previous section if we could have generated an infinite number of random numbers.† For this reason we would expect the distributions shown in **Table #0** in the columns headed $\nu=10^3$, $\nu=10^4$ and $\nu=10^5$ to approach the distribution shown in the column headed "ν=infinity" as ν increases.

---

† See **APPENDIX - Univariate Generalized Numerical Convolutions.**

CONFIDENCE INTERVALS FOR TREND - LARGE SIZED RESAMPLES

Lowrie and Lipsky† presented group major medical expense claims by claimant per accident year for each of the five years 1983 to 1987. Their distributions are shown separately for adult or child combined with either comprehensive or supplemental coverage. We will focus on adult comprehensive coverage only, noting that the deductible is \$100 per calendar year and the coinsurance is 20%.

We considered the random variable

$$W = \frac{1+\hat{T}}{E[1+\hat{T}]} - 1.$$

Note that E[1+W] is clearly equal to unity. We were interested to find that f_{1+W} shows a remarkable degree of stability as we vary the accident year pairs. Using the operational bootstrapping approach described in the previous section, the distributions $f_{1+\hat{T}}$ and f_{1+W} were generated for each of the accident year pairs 1983-84, 1984-85, 1985-86 and 1986-87 and are shown in **Table #1**. In **Table #1** the numbers of claims in the resamples varied from 66,260 to 111,263. We concluded that,

---

† See ref[2].

provided the numbers of claims are of this order of magnitude, f_{1+W} can be used as a pivotal distribution; that is, that for any true trend ϑ_0 the point estimates $1+\hat{T}|\vartheta_0$ can be considered to be distributed as $f_{(1+\vartheta_0)\cdot(1+W)}$.

Table #1

$$f_{1+\hat{T}} = \sum_{i=1}^{m} f_{(1/m)\cdot Y_i} \Big/ \sum_{i=1}^{m} f_{(1/n)\cdot X_i}{}^{\dagger}$$

$$f_{1+W} = f_{(1+\hat{T})/E[1+\hat{T}]}$$

cumula-tive	1983-84 $1+\hat{T}$	1983-84 $1+W$	1984-85 $1+\hat{T}$	1984-85 $1+W$	1985-86 $1+\hat{T}$	1985-86 $1+W$	1986-87 $1+\hat{T}$	1986-87 $1+W$
.000001	.965	.934	.982	.936	.979	.938	1.001	.938
.00001	.972	.940	.989	.942	.986	.945	1.007	.945
.0001	.979	.948	.996	.949	.993	.951	1.015	.951
.001	.988	.956	1.005	.958	1.001	.959	1.023	.959
.01	.999	.967	1.016	.968	1.011	.969	1.034	.969
.025	1.004	.972	1.021	.973	1.016	.974	1.039	.974
.05	1.009	.976	1.026	.977	1.021	.978	1.043	.978
.1	1.014	.982	1.031	.982	1.026	.983	1.048	.983
.2	1.021	.988	1.037	.988	1.032	.989	1.054	.989
.3	1.025	.992	1.042	.993	1.036	.993	1.059	.993
.4	1.029	.996	1.046	.996	1.040	.997	1.063	.997
.5	1.033	1.000	1.049	1.000	1.043	1.000	1.066	1.000
.6	1.037	1.004	1.053	1.003	1.047	1.003	1.070	1.003
.7	1.041	1.007	1.057	1.007	1.051	1.007	1.074	1.007
.8	1.046	1.012	1.062	1.012	1.055	1.011	1.078	1.011
.9	1.053	1.019	1.068	1.018	1.062	1.017	1.085	1.017
.95	1.058	1.024	1.073	1.023	1.067	1.022	1.090	1.022
.975	1.063	1.029	1.078	1.027	1.071	1.027	1.095	1.027
.99	1.069	1.034	1.084	1.033	1.077	1.032	1.100	1.032
.999	1.081	1.046	1.095	1.043	1.088	1.042	1.112	1.042
.9999	1.091	1.055	1.104	1.052	1.097	1.051	1.121	1.051
.99999	1.099	1.064	1.113	1.060	1.105	1.059	1.129	1.059
.999999	1.109	1.073	1.122	1.070	1.115	1.068	1.139	1.068
mean	1.033	1.000	1.049	1.000	1.044	1.000	1.066	1.000
$var\cdot 10^{+3}$	.209	.223	.211	.192	.195	.179	.204	.179
n	66260		76857		83457		88977	
m	76857		83457		88977		111263	

$\dagger$ $\sum_{i=1}^{m} f_{(1/m)\cdot Y_i} = f_{(1/m)\cdot Y_1} + f_{(1/m)\cdot Y_2} + \ldots + f_{(1/m)\cdot Y_m}$ is being used to mean convolute $f_{(1/m)\cdot Y_1}$, $f_{(1/m)\cdot Y_2}$, ... and $f_{(1/m)\cdot Y_m}$

We note that

f_{1+W} is approximately $N(E[1+W],Var[1+W])$ and
$f_{1+\hat{T}}$ is approximately $N(E[1+\hat{T}],Var[1+\hat{T}])$.†

A Numerical Example of Determining a Confidence Interval for Trend Using Large Resamples

We now turn our attention to determining a confidence interval for the true trend ϑ. We wish to determine

$$\vartheta_1 \text{ such that } \Pr\{\ \vartheta_1 < \vartheta\ \} = 1 - \alpha/2$$

and

$$\vartheta_2 \text{ such that } \Pr\{\ \vartheta < \vartheta_2\ \} = 1 - \alpha/2,$$

so that the random interval $(\vartheta_1,\vartheta_2)$ encloses the true trend ϑ at the desired level $(1-\alpha)$ of confidence.

From Table #1 we can select a value of W (say w_1) such that
$1 - \alpha/2 = \Pr\{\ w_1 < W\ \}$; that is, such that
$1 - \alpha/2 = \Pr\{\ 1+w_1 < 1+W\ \}$

---

† An expression such as $N(\mu,\sigma^2)$ is being used, as is customary, to indicate a normal distribution with mean μ and variance σ^2.

$= \Pr(\ 1+w_1 < (1+\hat{T})/(1+\vartheta)\ )$

$= \Pr(\ (1+w_1)\cdot(1+\vartheta) < 1+\hat{T}\ )$

$= \Pr(\ 1+\vartheta < (1+\hat{T})/(1+w_1)\ )$

$= \Pr(\ \vartheta < (1+\hat{T})/(1+w_1) - 1\ )$

so we choose $\vartheta_1 = (1+\hat{T})/(1+w_1) - 1$.

Similarly, from Table #1 we can select a value of W (say w_2) such that

$1 - \alpha/2 = \Pr(\ W < w_2\ )$; that is, such that

$1 - \alpha/2 = \Pr(\ 1+W < 1+w_2\ )$

$= \Pr(\ (1+\hat{T})/(1+\vartheta) < 1=w_2)$

$= \Pr(\ 1+\hat{T} < (1+w_2)\cdot(1+\vartheta)\ )$

$= \Pr(\ (1+\hat{T})\cdot(1+w_2) < 1+\vartheta\ )$

$= \Pr(\ (1+\hat{T})/(1+w_2)-1 < \vartheta\ )$

so we choose $\vartheta_2 = (1+\hat{T})/(1+w_2) - 1$.

If $1-\alpha = 95\%$, then referring to Table #1 (1983-84) we can let $1+w_1 = .972$ and $1+w_2 = 1.029$ and find that

$$\vartheta_1 = 1.033/1.029 -1 = .004 \text{ and}$$

$$\vartheta_2 = 1.033/.972 -1 = .063.$$

Therefor, the confidence interval for the true trend ϑ is

$$(\vartheta_2, \vartheta_1) = (\ 0.4\%,\ 6.3\%\ );$$

3.3% was the corresponding point estimate. This result and the corresponding results for the other calendar year pairs are shown in the following table:

$(1+\vartheta)$

Mean and† 50th Percentile	95% Confidence Interval	Calendar Year	n	m
1.033	(1.004,1.063)	1983-84	66260	76857
1.049	(1.021,1.078)	1984-85	76857	83457
1.044	(1.016,1.071)	1985-86	83457	88977
1.066	(1.039,1.095)	1986-87	88977	111263

Test of Normality Assumptions

In order to see whether we could produce equally good confidence intervals making use of some normality assumptions, we assumed that $f_X{}^{*n}$ and $f_Y{}^{*m}$ could be approximated by the normal distributions $N(n\cdot E[X],n\cdot Var[X])$ and $N(m\cdot E[Y],m\cdot Var[Y])$, respectively. $f_{1+\hat{T}}$ was then obtained by generating

$$N(m\cdot E[Y],m\cdot Var[Y]) \;/\; N(n\cdot E[X],n\cdot Var[X])$$

and transforming the resulting distribution by multiplying the amounts (not the probabilities) by n/m. The resulting figures turned out to agree exactly with the figures shown in **Table #1.**‡

---

† The mean and the median could turn out to be different, but here they happen to be identical to the number of decimal places shown.

‡ A referee pointed out that if X and Y are asymptotically normal

In the following section we investigate the corresponding situation where n and m are equal and medium sized, say 64 to 16,384.

CONFIDENCE INTERVALS FOR TREND - MEDIUM SIZED RESAMPLES

So far we have been dealing with resamples of size n or m from an original sample of size n or m, respectively, either using or not using random numbers. But even though the original samples are of size n or m, we can generate resamples of, say, size $\tilde{n}$ (<n)

---

random variables and $\hat{T} = \bar{Y} / \bar{X} - 1$, then $\hat{T}$ is asymptotically $N(\mu,\sigma^2)$ with

$$\mu = \mu_Y / \mu_X - 1 \text{ and}$$

$$\sigma^2 = \mu_Y^2 \cdot \sigma_X^2 / \mu_X^4 \cdot n + \sigma_Y^2 / \mu_X^2 \cdot m;$$

and that these can be approximated by replacing the population quantities with the sample values.

If we had available (and used) the detailed data underlying the loss distributions presented by Lowrie and Lipsky (ref [2]), our confidence intervals would be slightly wider. Using the calendar year pair 1987-1988 and the above formula for σ^2 we find that the ratio of σ^2 based on the detailed data to σ^2 based on the grouped data is 1.016; that is, a 1.6% deficiency in the variance. The data for 1988 was not shown in reference [2]; however, Professor Lowrie was kind enough to furnish that data to me for the purposes of this paragraph. Professor Lowrie said that the "STANDARD DEVIATION" figures shown in reference [2] were calculated by an incorrect formula and should not be used.

and $\tilde{m}$ (<m); in particular, we can choose $\tilde{n} = \tilde{m}$ (< min{n,m}). The purpose of doing this would be to see what confidence intervals for trend might look like if the resamples were of medium (rather than large) size.

Consider

$$f_{1+\hat{T}} = f_{(1/\tilde{n})}(Y_1+Y_2+ \ldots + Y_{\tilde{n}}) \,/\, f_{(1/\tilde{n})}(X_1+X_2+ \ldots + X_{\tilde{n}})$$

$$= \sum_{i=1}^{\tilde{n}} f_{(1/\tilde{n})\cdot Y_i} \,/\, \sum_{i=1}^{\tilde{n}} f_{(1/\tilde{n})\cdot X_i}$$

where $\tilde{n}$ takes on the value 64, 128, ..., or 1024 and the X_i and Y_i are based on calendar years 1983 and 1984, respectively, 1984 and 1985, respectively, 1985 and 1986, respectively, or 1986 and 1987, respectively. The distributions $f_{1+\hat{T}}$ are shown in Table #2, along with the corresponding standardized distributions $f_{1+W} = f_{(1+\hat{T})/E[1+\hat{T}]}$.

For determining confidence intervals for trend where the resamples are of medium size, we wish to assume for given $\tilde{n}$ that f_{1+W} does not differ significantly as we vary the calendar year pairs. The reasonableness of making this assumption seems to be confirmed by the fact that for fixed $\tilde{n} = \tilde{m}$ the standardized distributions f_{1+W} in Table #2 vary as little as they do by calendar year pair, at least in the portion of the distributions between cumulatives of .025 and .975.

A Numerical Example of Determining a Confidence Interval for Trend Using Medium Resized Samples

Suppose a trend factor of 1.15 has been observed from one year to another and the number of claims is 64 in each of the two accident years. We will now determine a 95% confidence interval for the true severity trend ϑ, again using the formulas shown in the previous numerical example.

Referring to Table #1 (1983-84) we can let $1+w_1=2.108$ and $1+w_2=.406$ if $1-\alpha=.95$; so $\vartheta_1 = 1.15/2.108 = .546$ and $\vartheta_2 = 1.15/.406 = 2.83$. Thus the estimated 95% confidence interval for the underlying trend factor $1+\vartheta$ would be

$$(.546,2.83).$$

This result and the corresponding results for the other calendar year pairs are shown in the following table:

mean	50th percentile	95% confidence interval	calendar year	$\tilde{n}$	$\tilde{m}$
1.125	1.028	(.546,2.83)	1983-84	64	64
1.147	1.050	(.542,2.91)	1984-85	64	64
1.142	1.037	(.533,2.96)	1985-86	64	64
1.171	1.059	(.524,3.05)	1986-87	64	64

Table #2 includes distributions for $\tilde{n}=\tilde{m}=64,128,256,512$ and 1024 for calendar year pairs 1983-84, 1984-85, 1985-86 and 1986-87; and distributions for $\tilde{n}=\tilde{m}=2048$, 4096, 8192 and 16384 for calendar year pair 1983-84.

Table #2

$$f_{1+\hat{T}} = \sum_{i=1}^{64} f_{(1/64)\cdot Y_i} \,/\, \sum_{i=1}^{64} f_{(1/64)\cdot X_i}$$

$$f_{1+W} = f_{(1+\hat{T})/E[1+\hat{T}]}$$

cumulative	1983-84 $1+\hat{T}$	1983-84 1+W	1984-85 $1+\hat{T}$	1984-85 1+W	1985-86 $1+\hat{T}$	1985-86 1+W	1986-87 $1+\hat{T}$	1986-87 1+W
.000001	.109	.097	.101	.088	.123	.107	.101	.087
.00001	.146	.129	.124	.108	.150	.131	.122	.104
.0001	.179	.159	.161	.140	.193	.169	.159	.135
.001	.247	.220	.220	.191	.253	.221	.225	.192
.01	.375	.334	.366	.319	.369	.323	.360	.308
.025	.457	.406	.453	.395	.444	.389	.441	.377
.05	.526	.468	.531	.463	.518	.453	.520	.444
.1	.618	.550	.623	.543	.610	.535	.614	.524
.2	.737	.655	.749	.652	.737	.646	.749	.639
.3	.837	.744	.854	.744	.839	.735	.853	.728
.4	.932	.829	.949	.827	.937	.821	.951	.812
.5	1 028	.914	1.050	.915	1.037	.909	1.059	.904
.6	1.135	1.009	1.159	1.010	1.151	1.008	1.175	1.003
.7	1.261	1.121	1.294	1.127	1.283	1.124	1.316	1.123
.8	1.431	1.272	1.474	1.284	1.464	1.282	1.502	1.283
.9	1.723	1.532	1.775	1.547	1.770	1.551	1.826	1.559
.95	2.027	1.802	2.094	1.825	2.099	1.839	2.172	1.854
.975	2.370	2.108	2.435	2.122	2.465	2.159	2.570	2.194
.99	2.940	2.614	2.924	2.548	3.011	2.637	3.218	2.747
.999	4.844	4.308	4.237	3.693	4.714	4.129	5.435	4.640
.9999	6.635	5.900	5.578	4.861	6.461	5.659	9.069	7.742
.99999	8.360	7.434	6.972	6.076	8.138	7.128	12.108	10.336
.999999	10.239	9.104	8.453	7.366	9.926	8.695	14.811	12.644
mean	1.125	1.000	1.147	1.000	1.142	1.000	1.171	1.000
var	.267	.211	.264	.201	.283	.217	.338	.246
$\tilde{n}$	64		64		64		64	
$\tilde{m}$	64		64		64		64	

Table #2 - Continued

$$f_{1+\hat{T}} = \sum_{i=1}^{128} f_{(1/128)\cdot Y_i} \Big/ \sum_{i=1}^{128} f_{(1/128)\cdot X_i}$$

$$f_{1+W} = f_{(1+\hat{T})/E[1+\hat{T}]}$$

cumula-tive	1983-84 $1+\hat{T}$	1983-84 $1+W$	1984-85 $1+\hat{T}$	1984-85 $1+W$	1985-86 $1+\hat{T}$	1985-86 $1+W$	1986-87 $1+\hat{T}$	1986-87 $1+W$
.000001	.209	.193	.187	.170	.221	.202	.192	.171
.00001	.246	.227	.223	.202	.255	.233	.227	.202
.0001	.295	.272	.269	.244	.305	.278	.274	.244
.001	.366	.338	.339	.307	.376	.343	.347	.309
.01	.489	.452	.474	.429	.489	.446	.475	.423
.025	.560	.518	.555	.503	.556	.507	.551	.490
.05	.624	.576	.626	.568	.618	.564	.619	.551
.1	.702	.649	.710	.643	.697	.636	.704	.627
.2	.803	.742	.816	.740	.804	.737	.813	.724
.3	.881	.814	.899	.814	.886	.808	.902	.802
.4	.954	.881	.975	.883	.963	.878	.980	.872
.5	1.028	.949	1.051	.953	1.039	.947	1.060	.943
.6	1.107	1.023	1.135	1.028	1.122	1.024	1.147	1.021
.7	1.200	1.108	1.231	1.116	1.220	1.112	1.249	1.111
.8	1.319	1.219	1.356	1.229	1.347	1.228	1.383	1.231
.9	1.515	1.399	1.559	1.413	1.556	1.419	1.602	1.426
.95	1.717	1.586	1.758	1.593	1.766	1.610	1.832	1.630
.975	1.936	1.788	1.958	1.774	1.983	1.809	2.081	1.851
.99	2.271	2.098	2.224	2.015	2.291	2.089	2.452	2.181
.999	3.168	2.927	2.887	2.616	3.129	2.854	3.689	3.282
.9999	3.982	3.678	3.544	3.211	3.930	3.584	5.261	4.681
.99999	4.813	4.446	4.211	3.817	4.733	4.316	6.456	5.743
.999999	5.703	5.268	4.893	4.434	5.577	5.087	7.627	6.785
mean	1.082	1.000	1.103	1.000	1.097	1.000	1.124	1.000
var	.126	.108	.127	.104	.135	.112	.160	.127
$\tilde{n}$	128		128		128		128	
$\tilde{m}$	128		128		128		128	

Table #2 - Continued

$$f_{1+\hat{T}} = \sum_{i=1}^{256} f_{(1/256 \cdot Y_i} / \sum_{i=1}^{256} f_{(1/256) \cdot X_i}$$

$$f_{1+W} = f_{(1+\hat{T})/E[1+\hat{T}]}$$

cumula-tive	1983-84 $1+\hat{T}$	1983-84 1+W	1984-85 $1+\hat{T}$	1984-85 1+W	1985-86 $1+\hat{T}$	1985-86 1+W	1986-87 $1+\hat{T}$	1986-87 1+W
.000001	.327	.309	.301	.279	.338	.315	.309	.281
.00001	.369	.349	.342	.317	.378	.353	.350	.319
.0001	.423	.399	.397	.369	.430	.402	.405	.369
.001	.495	.468	.472	.438	.501	.467	.480	.437
.01	.602	.569	.589	.546	.603	.563	.594	.542
.025	.660	.623	.656	.608	.660	.616	.656	.598
.05	.713	.673	.715	.663	.711	.664	.714	.651
.1	.776	.732	.784	.727	.776	.724	.783	.714
.2	.857	.809	.871	.808	.859	.802	.871	.794
.3	.919	.867	.936	.868	.923	.862	.939	.856
.4	.974	.920	.995	.923	.982	.916	1.001	.913
.5	1.029	.971	1.052	.975	1.040	.970	1.061	.967
.6	1.088	1.027	1.113	1.032	1.102	1.028	1.126	1.026
.7	1.154	1.089	1.183	1.097	1.172	1.094	1.201	1.094
.8	1.239	1.170	1.270	1.178	1.262	1.178	1.296	1.181
.9	1.375	1.298	1.406	1.303	1.403	1.309	1.448	1.319
.95	1.508	1.423	1.531	1.420	1.537	1.435	1.597	1.456
.975	1.644	1.552	1.650	1.430	1.670	1.559	1.750	1.595
.99	1.829	1.726	1.803	1.671	1.846	1.723	1.963	1.789
.999	2.278	2.151	2.167	2.010	2.279	2.127	2.654	2.419
.9999	2.271	2.555	2.518	2.335	2.696	2.516	3.324	3.029
.99999	3.137	2.962	2.864	2.655	3.111	2.904	3.904	3.558
.999999	3.570	3.370	3.206	2.973	3.528	3.293	4.522	4.121
mean	1.059	1.000	1.078	1.000	1.097	1.000	1.097	1.000
var	.062	.055	.063	.055	.135	.058	.079	.066
$\tilde{n}$	256		256		256		256	
$\tilde{m}$	256		256		256		256	

Table #2 - Continued

$$f_{1+\hat{T}} = \sum_{i=1}^{512} f_{(1/512)\cdot Y_i} \Big/ \sum_{i=1}^{512} f_{(1/512)\cdot X_i}$$

$$f_{1+W} = f_{(1+\hat{T})/E[1+\hat{T}]}$$

cumula-	1983-84		1984-85		1985-86		1986-87	
tive	$1+\hat{T}$	1+W	$1+\hat{T}$	1+W	$1+\hat{T}$	1+W	$1+\hat{T}$	1+W
.000001	.457	.436	.432	.405	.465	.440	.440	.406
.00001	.498	.476	.475	.446	.505	.478	.483	.446
.0001	.549	.525	.529	.497	.555	.525	.536	.495
.001	.614	.587	.599	.563	.618	.585	.606	.560
.01	.702	.671	.696	.654	.705	.666	.701	.648
.025	.748	.715	.748	.702	.750	.709	.751	.694
.05	.789	.754	.793	.745	.791	.748	.796	.735
.1	.838	.801	.847	.796	.841	.795	.850	.785
.2	.901	.860	.914	.859	.906	.856	.919	.849
.3	.947	.905	.965	.906	.954	.902	.971	.897
.4	.989	.945	1.009	.948	.998	.943	1.017	.940
.5	1.030	.984	1.052	.988	1.041	.984	1.063	.982
.6	1.073	1.025	1.097	1.030	1.086	1.027	1.110	1.026
.7	1.121	1.071	1.146	1.076	1.137	1.075	1.164	1.075
.8	1.182	1.129	1.207	1.134	1.200	1.134	1.232	1.138
.9	1.274	1.217	1.298	1.219	1.295	1.224	1.336	1.234
.95	1.361	1.300	1.379	1.295	1.381	1.306	1.433	1.324
.975	1.443	1.379	1.454	1.365	1.463	1.323	1.529	1.413
.99	1.549	1.480	1.546	1.452	1.566	1.481	1.660	1.533
.999	1.798	1.718	1.760	1.653	1.811	1.712	2.017	1.863
.9999	2.035	1.944	1.959	1.840	2.042	1.930	2.350	2.171
.99999	2.266	2.165	2.149	2.019	2.266	2.142	2.667	2.464
.999999	2.492	2.381	2.334	2.192	2.485	2.349	3.000	2.767
mean	1.047	1.000	1.065	1.000	1.058	1.000	1.082	1.000
var	.031	.028	.032	.028	.033	.029	.039	.034
$\tilde{n}$	512		512		512		512	
$\tilde{m}$	512		512		512		512	

Table #2 - Continued

$$f_{1+\hat{T}} = \sum_{i=1}^{1024} f_{(1/1024)\cdot Y_i} \Big/ \sum_{i=1}^{1024} f_{(1/1024)\cdot X_i}$$

$$f_{1+W} = f_{(1+\hat{T})/E[1+\hat{T}]}$$

cumulative	1983-84 $1+\hat{T}$	1983-84 $1+W$	1984-85 $1+\hat{T}$	1984-85 $1+W$	1985-86 $1+\hat{T}$	1985-86 $1+W$	1986-87 $1+\hat{T}$	1986-87 $1+W$
.000001	.549	.559	.565	.532	.549	.559	.524	.531
.00001	.592	.594	.607	.570	.593	.593	.573	.567
.0001	.641	.636	.655	.615	.643	.634	.627	.611
.001	.699	.687	.713	.670	.702	.684	.693	.665
.01	.774	.754	.787	.743	.778	.751	.776	.736
.025	.812	.788	.824	.780	.816	.784	.817	.772
.05	.844	.819	.857	.812	.850	.814	.854	.804
.1	.884	.854	.897	.851	.890	.851	.898	.842
.2	.933	.899	.947	.898	.940	.897	.953	.890
.3	.970	.933	.984	.934	.978	.931	.995	.927
.4	1.002	.963	1.017	.965	1.011	.962	1.031	.959
.5	1.033	.992	1.049	.994	1.043	.992	1.066	.990
.6	1.065	1.022	1.082	1.024	1.076	1.022	1.103	1.022
.7	1.101	1.055	1.119	1.058	1.113	1.057	1.143	1.058
.8	1.144	1.096	1.164	1.098	1.157	1.098	1.192	1.103
.9	1.206	1.156	1.229	1.157	1.222	1.160	1.263	1.169
.95	1.261	1.210	1.287	1.207	1.278	1.214	1.325	1.229
.975	1.311	1.260	1.339	1.253	1.330	1.263	1.382	1.286
.99	1.371	1.322	1.404	1.209	1.392	1.324	1.453	1.360
.999	1.510	1.464	1.555	1.434	1.537	1.464	1.615	1.540
.9999	1.639	1.595	1.697	1.547	1.673	1.591	1.769	1.710
.99999	1.766	1.720	1.839	1.652	1.806	1.713	1.921	1.874
.999999	1.892	1.841	1.984	1.752	1.941	1.829	2.076	2.036
mean	1.040	1.000	1.058	1.000	1.051	1.000	1.075	1.000
var	.016	.014	.017	.015	.017	.015	.021	.017
$\tilde{n}$	1024		1024		1024		1024	
$\tilde{m}$	1024		1024		1024		1024	

Table #2 - Continued

$$f_{1+\hat{T}} = \sum_{i=1}^{\tilde{m}} f_{(1/\tilde{m})\cdot Y_i} / \sum_{i=1}^{\tilde{n}} f_{(1/\tilde{n})\cdot X_i}$$

$$f_{1+W} = f_{(1+\hat{T})/E[1+\hat{T}]}$$

cumula-tive	1983-84 $1+\hat{T}$	1983-84 $1+W$	1983-84 $1+\hat{T}$	1983-84 $1+W$	1983-84 $1+\hat{T}$	1983-84 $1+W$	1983-84 $1+\hat{T}$	1983-84 $1+W$
.000001	.674	.650	.769	.743	.841	.813	.894	.865
.00001	.707	.682	.794	.767	.859	.831	.907	.878
.0001	.744	.718	.822	.794	.880	.851	.923	.893
.001	.788	.760	.855	.826	.904	.875	.941	.910
.01	.844	.814	.897	.866	.935	.904	.963	.932
.025	.872	.841	.917	.886	.950	.919	.974	.942
.05	.897	.865	.935	.903	.963	.931	.983	.951
.1	.925	.893	.956	.924	.978	.946	.994	.962
.2	.961	.927	.982	.949	.997	.964	1.007	.974
.3	.988	.953	1.001	.967	1.010	.977	1.017	.984
.4	1.011	.975	1.018	.983	1.022	.988	1.025	.992
.5	1.033	.997	1.033	.998	1.033	.999	1.033	1.000
.6	1.056	1.018	1.049	1.014	1.045	1.010	1.041	1.007
.7	1.080	1.042	1.066	1.030	1.057	1.022	1.050	1.016
.8	1.110	1.071	1.087	1.050	1.071	1.036	1.060	1.025
.9	1.153	1.112	1.116	1.079	1.091	1.055	1.074	1.039
.95	1.189	1.147	1.141	1.102	1.108	1.072	1.086	1.050
.975	1.222	1.178	1.163	1.124	1.123	1.086	1.096	1.060
.99	1.261	1.216	1.189	1.149	1.141	1.103	1.108	1.072
.999	1.347	1.299	1.246	1.203	1.179	1.140	1.134	1.097
.9999	1.424	1.374	1.294	1.251	1.211	1.171	1.156	1.118
.99999	1.496	1.443	1.339	1.293	1.240	1.199	1.175	1.137
.999999	1.564	1.509	1.380	1.333	1.266	1.224	1.193	1.154
mean	1.037	1.000	1.035	1.000	1.034	1.000	1.034	1.000
var	.008	.007	.004	.004	.002	.002	.001	.001
$\tilde{n}$	2048		4096		8192		16384	
$\tilde{m}$	2048		4096		8192		16384	

If $\tilde{n}$ and $\tilde{m}$ are sufficiently large, we can avoid performing the convolutions to produce

$$f_{Y_1+Y_2+...+Y_{\tilde{m}}} \text{ and } f_{X_1+X_2+...+X_{\tilde{n}}}.$$

That is, if $f_{Y_1+Y_2+...+Y_{\tilde{m}}}$ and $f_{X_1+X_2+...+X_{\tilde{n}}}$ are close to being normal distributions, we can assume that

$$f_{Y_1+Y_2+...+Y_{\tilde{m}}} \text{ is } N(\ E(Y_1+Y_2+\ ...\ +\ Y_{\tilde{m}}),\ Var(Y_1+Y_2+\ ...\ +\ Y_{\tilde{m}})\ )^{\dagger}$$

and

$$f_{X_1+X_2+...+X_{\tilde{n}}} \text{ is } N(\ E(X_1+X_2+\ ...\ +\ X_{\tilde{n}}),\ Var(X_1+X_2+\ ...\ +\ X_{\tilde{n}})\ ).$$

and do only a single convolution for quotients; namely,

---

†A good discretized version of a normal distribution can be obtained by *generating* a binomial distribution $b(n;p)$, where n is large and $p = .5$; and then a discretized version of $n(\mu,\sigma^2)$ can be obtained by performing the usual type of transformation

$$z = \sigma \cdot \frac{x-n \cdot p}{n \cdot p \cdot q} + \mu.$$

$$N(E((1/\tilde{m}) \cdot (Y_1+Y_2+...Y_{\tilde{m}}), Var((1/\tilde{m}) \cdot (Y_1+Y_2+...+Y_{\tilde{m}}))$$

/

$$N(E((1/\tilde{n}) \cdot (X_1+X_2+...+X_{\tilde{n}})), Var((1/\tilde{n}) \cdot (X_1+ X_2+...+X_{\tilde{n}})).$$

Based on the underlying adult comprehensive major medical claim samples and the generated distributions, we can draw the following conclusions for this data:

1. For resample sizes of 256 or less, the assumption of normality for distributions of average size claims may not be particularly useful; this is because such assumption produces negative average size claim per claimant with appreciable probability.

2. From **Table #3** it can be ascertained how good the assumption of normality for distributions of average size claim per claimant are for generating distributions of point estimates of trend for resamples of size $\tilde{n}=\tilde{m}=512$.

3. **Table #3** for $\tilde{m}=\tilde{n}=1024$ (not shown) demonstrated that the assumption of normality for distributions of average size claim distributions for resamples of size 1024 produces point estimate of trend distributions shown in **Table #2** (for $\tilde{n}=\tilde{m}=1024$), to an accuracy of at least 3 decimal places in $1+\hat{T}$. This does not imply that the point estimate of trend distributions themselves are normal.

4. **Table #2** can be used almost directly to determine how large the samples need to be in order for the trend distributions themselves to be essentially normal; that is, whether

$f_{1+\hat{T}}$ is approximately $N(E[1+\hat{T}],Var[1+\hat{T}])$ or

f_{1+W} is approximately $N(E[1+W],Var[1+W])$.

Of course, it is easy to see that such normality is lacking if the median is not equal to the mean or if symmetry is lacking. If the median is close to the mean and a fair degree of symmetry exists, then you may want to compare

$N(E[1+\hat{T}],Var[1+\hat{T}])$ with $f_{1+\hat{T}}$ or

$N(E[1+W],Var[1+W])$ with f_{1+W}

at selected cumulative probabilities, e.g. .025, .05, .95 and .975.

Table #3

$$f_{1+\hat{T}} = \sum_{i=1}^{512} f_{(1/512)\cdot Y_i} \Big/ \sum_{i=1}^{512} f_{(1/512)\cdot X_i}$$

$$F_{1+\hat{T}'} = N(E((1/512)\cdot(Y_1+Y_2+\ldots Y_{512}), Var((1/512)\cdot(Y_1+Y_2+\ldots+Y_{512}))$$
$$/$$
$$N(E((1/512)\cdot(X_1+X_2+\ldots+X_{512})), \ Var((1/512)\cdot(X_1+ X_2+\ldots+X_{512}))$$

cumula-	1983-84		1984-85		1985-86		1986-87	
tive	$1+\hat{T}$	$1+\hat{T}'$	$1+\hat{T}$	$1+\hat{T}'$	$1+\hat{T}$	$1+\hat{T}'$	$1+\hat{T}$	$1+\hat{T}'$
.000001	.457	.384	.432	.405	.465	.381	.440	.338
.00001	.498	.440	.475	.460	.505	.439	.483	.403
.0001	.549	.506	.529	.523	.555	.504	.536	.475
.001	.614	.582	.599	.597	.618	.583	.606	.563
.01	.702	.680	.696	.694	.705	.683	.701	.671
.025	.748	.730	.748	.743	.750	.734	.751	.726
.05	.789	.774	.793	.787	.791	.778	.796	.776
.1	.838.	.826	.847	.840	.841	.832	.850	.835
.2	.901	.894	.914	.907	.906	.900	.919	.909
.3	.947	.944	.965	.959	.954	.953	.971	.966
.4	.989	.990	1.009	1.004	.998	.998	1.017	1.017
.5	1.030	1.033	1.052	1.050	1.041	1.043	1.063	1.066
.6	1.073	1.079	1.097	1.096	1.096	1.091	1.110	1.117
.7	1.121	1.130	1.146	1.149	1.137	1.143	1.164	1.176
.8	1.182	1.193	1.207	1.214	1.200	1.208	1.232	1.248
.9	1.274	1.286	1.298	1.315	1.295	1.305	1.336	1.355
.95	1.361	1.371	1.379	1.404	1.381	1.393	1.433	1.452
.975	1.443	1.450	1.454	1.490	1.463	1.475	1.529	1.545
.99	1.549	1.550	1.546	1.598	1.566	1.579	1.660	1.662
.999	1.798	1.792	1.760	1.869	1.811	1.834	2.017	1.953
.9999	2.035	2.038	1.959	2.155	2.042	2.097	2.350	2.259
.99999	2.266	2.303	2.149	2.472	2.266	2.382	2.667	2.599
.999999	2.492	2.598	2.334	2.840	2.485.	2.703	3.000	2.367
mean	1.047	1.048	1.065	1.066	1.058	1.059	1.082	1.084
var	.031	.034	.032	.036	.033	.036	.039	.044
$\tilde{n}$	512		512		512		512	
$\tilde{m}$	512		512		512		512	

The columns headed $1+\hat{T}$ in this table are taken from Table #2.

RECAP AND CONCLUSIONS

We started with original samples of comprehensive major medical claims per claimant, one sample for each of two calendar years. By resampling with replacement (using numerical convolutions) from the corresponding empirical distributions, we generated distributions of average size claim per claimant where the number of resamples was a power of 2 from 6 to 15 (i.e. 64, 128, 256, 512, 1024, 2048, 4096, 8192 or 16384). Assuming an equal number of resamples in each of two calendar years, we convoluted these latter distributions for quotients to obtain distributions of point estimates for trend in average size claim per claimant from the one calendar year to the other. The results are shown in **Table #2**.

Table #1 presents similar distributions of resample point estimates for trend in average size claim per claimant where the numbers of resamples in adjacent calendar years are those of the original experience during the observation period (1983 to 1987, inclusive). The distributions in **Table #1** are close to normal, which is perhaps not unexpected in view of the fact that the numbers of claims lie in the range from 66,260 to 111,263. Standardizing these trend distributions by dividing the amounts (not the probabilities) by their respective mean values, we find a high degree of stability as we move from one pair of calendar years to another. This enables us to use the distributions in **Table #1** for determining confidence intervals for trend in average size claim per claimant, where we are dealing with such large resample sizes.

We show how we might use **Table #2** to estimate 95% confidence intervals for trend where medium-sized samples of comprehensive major medical losses per claimant are available. Of course, since the underlying experience data involves $100 deductible/20% coinsurance and essentially no maximum, **Table #2** should be used with caution if the major medical plan deviates significantly from this. **Table #2** shows considerable stability† in the standardized distributions of resample point estimates for trend, as we move from one pair of calendar years to another. This enables us to use the distributions in **Table #2** for determining 95% confidence intervals for trend in average size claims per claimant , where we are dealing with resamples of medium sizes.

The numerical convolutions (for sums and quotients) used in producing the figures in **Table #1, #2, and #3** were generated using the methods described in **APPENDIX - UNIVARIATE GENERALIZED NUMERICAL CONVOLUTIONS** using $\varepsilon=10^{-15}$ and $nax=1000$. For any one convolution the total of the discarded probability products did not exceed $5\cdot10^{-7}$; choosing a smaller value for ε would make this

† At least where the cumulative is in the range from .025 to .975.

figure even smaller.

APPENDIX - UNIVARIATE GENERALIZED NUMERICAL CONVOLUTIONS

If f_X and f_Y are independent distributions of the discrete finite univariate random variables X and Y, respectively, then the distribution f_{X+Y} of the sum W=X+Y is the convolution $f_X + f_Y$ of f_X and f_Y for sums.†

Let f_X be expressed in element notation as

$$\begin{bmatrix} x1_1 & p1_1 \\ & \vdots & \\ x1_{n_1} & p1_{n_1} \end{bmatrix}$$

which we will also express as

$$\begin{pmatrix} x1_i & p1_i \end{pmatrix}_{i=1,n_1}.$$

---

† We are using the operation + instead of * between two distributions to indicate convolution for sums; that is, $f_X + f_Y$ instead of $f_X * f_Y$. We use the notation f_X/f_Y for the convolution of f_X and f_Y for quotients X/Y.

Similarly, let f_Y be $\left(x2_j \quad p2_j \right)_{j=1,n_2}$

Then $f_{1+W} = f_{X+Y} = f_X + f_Y =$

$$\begin{bmatrix} x1_1+x2_1 & p1_1\cdot p2_1 \\ x1_1+x2_2 & p1_1\cdot p2_2 \\ & \vdots \\ x1_1+x2_{n_2} & p1_1\cdot p2_{n_2} \\ x1_2+x2_1 & p1_2\cdot p2_1 \\ X1_2+X2_2 & P1_2\cdot p2_2 \\ & \vdots \\ x1_2+x2_{n_2} & p1_2\cdot p2_{n_2} \\ x1_3+x2_1 & p1_3\cdot p2_1 \\ x1_3+x2_2 & p1_3\cdot p2_2 \\ & \vdots \\ x1_3+x2_{n_2} & p1_3\cdot p2_{n_2} \\ & \vdots \\ x1_{n_1}+x2_1 & p1_{n_1}\cdot p2_1 \\ x1_{n_1}+x2_2 & p1_{n_1}\cdot p2_2 \\ & \vdots \\ x1_{n_1}+x2_{n_2} & p1_{n_1}\cdot p2_{n_2} \end{bmatrix} \qquad \text{Matrix (1)}$$

which we might also express as

$$\left[\begin{array}{cc} x1_i+x2_j & p1_i\cdot p2_j \end{array} \right]_{i=1,n_1;\, j=1,n_2} \dagger$$

If n_1 and n_2 are (say) 1000, then generating this matrix would involve 10^6 lines.‡ This would be practical if we do not intend to use f_{1+W} in further convolutions. But, if (for example) we want to generate the distribution $f_U = f_W + f_Z$ of $U = (X+Y)+Z$ where

$$f_Z = \left[\begin{array}{cc} x3_k & p3_k \end{array} \right]_{k=1,1000}$$

then we would be dealing with 10^9 lines. And further convolutions would become impractical, because of both the amount of computer storage and the amount of computing time required. The following algorithm has been designed to overcome these

---

† For a generalized convolution of f_{X_1} and f_{X_2} to generate the distribution f_{X_1/X_2} of the random variable X_1/X_2 this expression would be replaced by

$$\left[\begin{array}{cc} x1_i/x2_j & p1_i\cdot p2_j \end{array} \right]_{i=1,n_1;\, j=1,n_2}$$

‡ There may be some collapsing due to identical amounts on different lines. The number of lines produced is reduced by representing on a single line all lines with identical amounts; on that line is the amount and the sum of the original probabilties.

problems.

The Univariate Generalized Numerical Convolution Algorithm

Choose $\varepsilon>0$. Typically ε is chosen to be 10^{-10} or 10^{-15}.

Loop #1:

Perform the calculations indicated in Matrix (1) above, discarding any lines for which the resulting probability is less than ε; that is discard lines for which
$p1_i \cdot p2_j < \varepsilon$.

The purpose of this is to avoid underflow problems and to increase the fineness of the partitions (meshes) to be imposed.

Calculate

$low_x = \min\{ x1_i+x2_j \neq 0 \mid p1_i \cdot p2_j < \varepsilon \}$ † ‡

---

† In many applications we replace $x1_i+x2_j$ by $\log(x1_i+x2_j)$, which will allow finer subintervals at the low end of the range. Of course, to be able to use logs the range of X+Y should not include values less than one (to avoid theoretical and numerical problems).

‡ For a generalized convolution of f_{X_1} and f_{X_2} to generate the

$$i=1,2,\ldots,n_1$$
$$j=1,2,\ldots,n_2$$

and

$$high_x = \max_{\substack{i=1,2,\ldots,n_1 \\ j=1,2,\ldots,n_2}}\{\ x1_i+x2_j\neq 0 \ \mid\ p1_i\cdot p2_j < \varepsilon\ \}.^{\dagger\ddagger}$$

Let *nax* be a positive integer selected for the purpose of creating the following partition:

$$\text{let } \Delta = \frac{high_x - low_x}{nax/2-1};$$

partition the interval $(low_X-\Delta, high_x+\Delta)$ into $nax/2+1$ subintervals:

---

distribution f_{X_1/X_2} of the random variable X_1/X_2 these expressions would be replaced by

$$low_x = \min_{\substack{i=1,2,\ldots,n_1 \\ j=1,2,\ldots,n_2}}\{\ x1_i/x2_j\neq 0 \ \mid\ p1_i\cdot p2_j < \varepsilon\ \}$$

and

$$high_x = \max_{\substack{i=1,2,\ldots,n_1 \\ j=1,2,\ldots,n_2}}\{\ x1_i/x2_j\neq 0 \ \mid\ p1_i\cdot p2_j < \varepsilon\ \}.$$

$$\text{let } \Delta = \frac{high_x - low_x}{nax/2-1};$$

partition the interval $(low_x-\Delta, high_x+\Delta)$ into $nax/2+1$ subintervals:

r	Subinterval $I(r)$
1	$[0,0]$
2	$(low_x-\Delta,\ low_x)$
3	$[low_x,\ low_x+1\cdot\Delta)$
4	$[low_x+1\cdot\Delta,\ low_x+2\cdot\Delta)$
.	.
.	.
.	.
$nax/2$	$[low_x+(nax/2-3)\cdot\Delta,\ high_x-\Delta)$
$nax/2+1$	$[high_x, high_x+\Delta)$

Subinterval I_1 is the degenerate interval consisting of 0 alone. If for some $r_0>1$ $0\in I_{r_0}$, then 0 is deleted from I_{r_0}; that is, that particular subinterval has a hole at 0.

Loop #2:

For each r ($r=1,2,\ldots,nax/2+1$) set to zero the initial value of each of the accumulators

$$m_0[I_{(r)}]$$

$$m_1[I_{(r)}]$$

$$m_2[I_{(r)}] \text{ and}$$

$$m_3[I_{(r)}].$$

For each i ($1=1,2,\ldots,n_1$) and j ($j=1,2,\ldots,n_2$) for which

$$x1_i+x2_j > \varepsilon,$$

determine the positive integer r for which

$$x1_i+x2_j \in I_{(r)}$$

and perform the accumulations

$$m_0[I_{(r)}] = m_0[I_{(r)}] + p1_i \cdot P2_j$$
$$m_1[I_{(r)}] = m_1[I_{(r)}] + (x1_i\ x2_j)^1 \cdot P1_i \cdot p2_j$$
$$m_2[I_{(r)}] = m_2[I_{(r)}] + (x1_i+x2_j)^2 \cdot p1_i \cdot P2_j$$
$$m_3[I_{(r)}] = m_3[I_{(r)}] + (x1_i+x2_j)^3 \cdot p1_i \cdot p2_j$$

That is, we generate the probability and the 0^{th} through 3^{rd} moments for each mesh interval $I_{(r)}$ ($r=1,2,\ldots,nax/2+1$).

LOOp #3:

Von Mises Theorem and algorithm† guarantee that for each r ($r=1,2,\ldots,nax/2+1$) there exist and we can find two pairs of real numbers

$$(x_{1(r)},p_{1(r)}) \text{ and } (x_{2(r)},p_{2(r)})\ ^{\dagger}$$

such that

$$x_{1(r)}\in I_{(r)} \text{ and } x_{2(r)}\in I_{(r)}$$

and such that the following relationships hold:

---

† See pages 269-270 of Ref[3].

† In some cases $x_1=x_2$ and what would otherwise be two pairs $(x_1(r),p_1(r))$ and $(x_2(r),p_2(r))$ collapse into one pair $(x_1(r),p_1(r)+p_2(r))$. This would happen, for example, where the values of $x1_i+x2_j$ that fall into $I(r)$ are all identical.

Moment	Relationship	"C" Notation
0	$\sum_{i=1}^{2} p_i(r) = m_0[I(r)]$	probability
1	$\sum_{i=1}^{2} x_i(r)^1 \cdot p_i(r) = m_1[I(r)]$	1st moment
2	$\sum_{i=1}^{2} x_i(r)^2 \cdot p_i(r) = m_2[I(r)]$	2nd moment
3	$\sum_{i=1}^{2} x_i(r)^3 \cdot p_i(r) = m_3[I(r)]$	3rd moment

The "C" program *VONMISES* accepts the 0^{th} through 3^{rd} moments and produces two points† and associated probabilities, with the feature that these moments are accurately retained.

Having kept accurately the 0^{th} through 3^{rd} moments of X+Y within each mesh interval, we have automatically kept accurately the corresponding global moments.

We can then express the full distribution f_{X+Y} of the univariate random variable X+Y as

$$\left\{ \begin{matrix} x_k(r) & p_k(r) \end{matrix} \right\}_{r=1,nax/2+1;k=1,2}$$

---

† Ibid

We will now describe how we actually obtain the number pairs

$$(x_1(r), p_1(r)) \text{ and } (x_2(r), p_2(r))$$

for any given value of r (r=1,2,...,nax/2+1). To simplify the notation somewhat in this description we will replace the symbols

$$m_0[I(r)],\ m_1[I(r)],\ m_2[I(r)] \text{ and } m_3[I(r)]$$

by

$$m_0,\ m_1,\ m_2 \text{ and } m_3,$$

respectively.

If $m_1=0$ and $m_0\neq0$, then we let

$$x_1(r) = 0 \qquad p_1(r) = m_0$$

$$x_2(r) = 0 \qquad p_2(r) = 0;$$

otherwise,

if $m_0 \cdot m_2 - m_1 \cdot m_1 < 10^{-10} \cdot |m_1|$, we let

$$x_1(r) = m_1/m_0 \qquad p_1(r) = m_0$$

$$x_2(r) = 0 \qquad p_2(r) = 0;$$

that is, in effect, use a single number pair rather than two number pairs if the variance in $I(r)$ is close to zero.†

---

† We treat this situation differently in order to avoid exceeding the limits of precision of the numbers being held by the computer.

otherwise, perform the following calculations:

$$c_0 = \frac{m_1 \cdot m_3 - m_2 \cdot m_2}{m_0 \cdot m_2 - m_1 \cdot m_1}$$

$$c_1 = \frac{m_1 \cdot m_2 - m_0 \cdot m_3}{m_0 \cdot m_2 - m_1 \cdot m_1}$$

$$a_1 = \frac{1}{2} \cdot (-c_1 - |c_1 \cdot c_1 - 4 \cdot c_0|^{.5})$$

$$a_2 = \frac{1}{2} \cdot (-c_1 + |c_1 \cdot c_1 - 4 \cdot c_0|^{.5})$$

$$s_1 = \frac{m_0 \cdot a_2 - m_1}{a_2 - a_1}$$

$$s_2 = \frac{m_1 - m_0 \cdot a_1}{a_2 - a_1}$$

$$x_1(r) = a_1 \qquad p_1(r) = s_1$$

$$x_2(r) = a_2 \qquad p_2(r) = s_2;$$

We check that $x_1(r)$ and $x_2(r)$ both lie in $I(r)$; and, if not, then

if $I(r)$ is a degenerate interval (i.e. consists of a single point), then we let

$$x_1(r) = m_1/m_0 \qquad p_1(r) = m_0$$

$$x_2(r) = 0 \qquad p_2(r) = 0;$$

otherwise,† we let

---

† This situation will occur only when the accuracy of the numbers being held by the computer is being impaired by the fact that the computer can hold numbers to only a limited degree of precision; since this situation occurs only where the associated probability is extremely small, the fact that not all of the first three

$$\sigma = |(-m_1/m_0)\cdot(m_1/m_0)+(m_2/m_0)|^{.5}$$

$$k = \left| \frac{m_1/m_0 - \text{left endpoint of } I_{(r)}}{\text{right endpoint of } I_{(r)} - m_1/m_0} \right|$$

$$x_{1(r)} = -\sigma\cdot|k|^{.5} + m_1/m_0 \qquad p_{1(r)} = m_0/(1+k)$$

$$x_{2(r)} = \sigma/|k|^{.5} + m_1/m_0 \qquad p_{2(r)} = p_{1(r)}\cdot k;$$

It is desirable to use double precision floating point numbers in performing these calculations; otherwise, you may run into numerical difficulties.

REFERENCES

[1] Efron, B. and Tibshirani, R., "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy," Statistical Science 1986, Vol.1, No. 1, 54-77.

[2] Lowrie, W. and Lipsky, L., "Power Tail Distributions and Group Medical Expense Insurance Payments,", ARCH 1990.2 ISSUE, 209-260.

[3] Springer, M.D., The Algebra of Random Variables, John Wiley & Sons, New York, 1979, 269-270.

---

moments are being retained in this situation is not of practical significance.

ACKNOWLEDGMENTS

I want to thank Robert Bender of Kemper National Insurance Company for editorial assistance which was truly helpful. Josh Zirin, formerly of Kemper Group, designed a practical way of implementing the Von Mises calculations. Larry Hickey of PolySystems, Inc. called my attention to Von Mises Theorem and has been helpful to me over a number of years as programmer, mathematician and codesigner. Anne Hoban, also of Kemper National, has been helpful with the selection of notation and in using the scientific/multifont word processor CHIWRITER.