TRANSFORMED BETA AND GAMMA DISTRIBUTIONS AND AGGREGATE LOSSES

GARY VENTER

Abstract

Distribution functions are introduced based on power transformations of beta and gamma distributions, and properties of these distributions are discussed. The gamma, beta, F, Pareto, Burr, Weibull and loglogistic distributions are special cases. The transformed gamma mixed with a gamma yields a transformed beta.

The transformed gamma is used to model aggregate distributions by matching moments. The transformed beta is used to account for parameter uncertainty in this model. Calculation procedures are discussed and *APL* program listings are included.

The transformed gamma is compared to exact methods of computing the aggregate distribution function based on the entire frequency and severity distributions.

INTRODUCTION

For pricing aggregate covers it is useful on occasion to have a way to estimate the distribution function for aggregate losses from the moments of this distribution. The usual approximation methods are designed primarily to calculate percentiles of the far right tail for mildly skewed distributions (e.g., see Pentikainen [9]). The gamma distribution has been suggested for this purpose (e.g., Hewitt [7]). However, the skewness of the gamma is always twice the coefficient of variation (see Hastings & Peacock [6]). Adding a third parameter to the gamma has been suggested by Seal [10], but the added parameter shifts the origin, sometimes resulting in the possibility of negative losses, which is often unsatisfactory. The transformed gamma distribution offers an alternative third parameter that affects the shape of the distribution but not its location.

The transformed beta and its special cases could be tried in this regard, also. However, its principal application herein is to deal with one kind of parameter uncertainty in the transformed gamma. The distributions are introduced below and then applications are discussed for each.

TRANSFORMED GAMMA AND TRANSFORMED BETA DISTRIBUTIONS

Transformed Gamma

The gamma function at r is defined as $\Gamma(r) = \int_0^\infty t^{r-1}e^{-t}dt$. The percentage of this integral reached by integrating up to some point x defines a probability distribution, i.e., the probability of being less than or equal to x. The gamma distribution usually is given by adding a scalar transformation of the variable; i.e., the probability of being less than or equal to x is given by the percentage of the integral that occurs up to λx for some positive number λ. The transformed gamma distribution adds a power transformation; i.e., the cumulative probability is given by:

$$G(x;r,\alpha,\lambda) = \frac{\int_0^{(\lambda x)^\alpha} t^{r-1}e^{-t}dt}{\Gamma(r)}$$

This distribution will be considered below as a model for aggregate losses although it may be a reasonable candidate for severity distributions as well. As it has three parameters it can match three moments of the distribution being modeled.

The gamma and exponential distributions are special cases given by $\alpha = 1$ and $\alpha = r = 1$ respectively. The Weibull distribution is also reached by taking $r = 1$. Thus the transformed gamma distribution provides a common generalization of the gamma and Weibull distributions and offers the possibility of improved fits whenever either have been found approximately suitable.

The moments are given by

$$E(X^n) = \frac{\Gamma(r + (n/\alpha))}{\lambda^n\,\Gamma(r)}$$

and the moment distributions

$$\frac{\int_0^a x^n dG_x}{E(X^n)}$$

are given by $G(a; (r + (n/\alpha),\alpha,\lambda)$. The probability density function is

$$g(x,r,\alpha,\lambda) = \frac{\alpha\lambda}{\Gamma(r)}\,(\lambda x)^{\alpha r-1}e^{-(\lambda x)^\alpha}.$$

These formulas require $n > -\alpha r$ but not necessarily an integer.

Finding parameters r, α, and λ from data involves the solution of non-linear equations whether matching moments or maximum likelihood is used. These equations can be quite readily solved by numerical means, e.g., Newton-Raphson iteration, as discussed more fully in Appendices A and B.

To match moments it has proven quite practical to solve for α and r using the known (e.g., known from sampling or calculated from frequency and severity) coefficients of variation and skewness, which do not depend on λ, in a system of two equations in two unknowns, and then to solve for λ using the mean. Handy equations are:

$$CV^2 + 1 = \Gamma(r + 2/\alpha)\,\Gamma(r) \div \Gamma(r + 1/\alpha)^2, \quad \text{and}$$
$$(SK \times CV^3) + 3CV^2 + 1 = \Gamma(r + 3/\alpha)\,\Gamma(r)^2 \div \Gamma(r + 1/\alpha)^3,$$

where CV is the coefficient of variation and SK skewness. See Appendix A for a discussion of how to solve this system.

Maximum likelihood techniques are discussed in Appendix B.

Once the parameters r,α, and λ have been estimated, the expected losses, higher moments, and percentiles of the aggregate layer from a to b can be read from the distribution. For example, expected losses for the layer are expected losses excess of a less expected losses excess of b. Define $R(a)$ to be the ratio of expected losses excess of a to all expected losses, i.e.,

$$R(a) = \frac{\int_a^\infty (x - a)dG_x}{E(X)}\,.$$

It is not difficult to show that

$$R(a) = 1 - \frac{\int_0^a x\, dG_x}{E(X)} - \frac{a}{E(X)}\,(1 - G(a)).$$

So far this is valid for any positive distribution G. Now using the moment ratio property of the transformed gamma:

$$R(a) = 1 - G(a;(r + (1/\alpha)), \alpha,\lambda) - \frac{a\lambda\Gamma(r)}{\Gamma(r + (1/\alpha))}\,(1 - G(a;r,\alpha,\lambda)).$$

Thus, if we knew how to compute the probability distribution function G, the aggregate layer expected losses would follow immediately. G can be calculated using numerical integration, but there is a series expansion for the incomplete gamma function that is also fairly quick to use. The incomplete gamma function is defined as

$$IG(x;r) = \int_0^x t^{r-1} e^{-t}\, dt \div \Gamma(r).$$

Then $G(x;r,\alpha,\lambda) = IG((\lambda x)^\alpha;r)$. From Abramowitz and Stegun [1] formula 6.5.29 (page 262) the expansion

$$IG(x;r) = \frac{e^{-x}x^{r-1}}{\Gamma(r)} \sum_{i=0}^{\infty} \prod_{k=0}^{i} \frac{x}{r+k}$$

can be derived. From 30 to 200 terms of this sum generally give acceptable accuracy. Exhibit 1 lists an *APL* program for *IG*.

For cases where the expected number of losses is low, there is a non-negligible probability that no losses will occur. The transformed gamma can not account for this because it is an entirely positive distribution. An alternative is a point mass at zero with the conditional probability on losses greater than zero being modeled by a transformed gamma. The probability of no losses can be computed from the frequency distribution. Formulas for computing the moments of the positive (conditional) distribution from the moments of the entire loss distribution and the probability of having a loss are given in Appendix C, along with standard formulas for computing aggregate moments from those for frequency and severity.

Example

Professional liability losses limited to $1 million per occurrence for a small group of hospitals are believed to have expected losses of $219,316 with coefficients of variation and skewness of 1.550 and 2.510 respectively and a probability of .123 of no losses. The aggregate expected losses excess of $1 million will be calculated by the above method.

By the formulas in Appendix C the positive portion of the aggregate distribution has expected losses of 250,000 and coefficients of variation and skewness of 1.409 and 2.344. Using the method in Appendix A gives parameters $r = .2478$, $\alpha = 1.470$, and $\lambda = 1.144 \times 10^{-6}$ for the positive portion. Thus the entire distribution has the cumulative probability function $\Pr(L < x) = .123 + .877\, G(x;\ .2478,\ 1.470,\ 1.144 \times 10^{-6})$. The excess ratio at $a = \$1{,}000{,}000$ can be calculated by the methods above to be .0728 for the conditional positive distribution, so the excess expected losses are $18,200 = .0728 × $250,000 for this piece and .877 × 18,200 = $16,000 for the entire distribution.

Transformed Beta

The beta function $B(r,s)$ may be defined as

$$B(r,s) = \frac{\int_0^\infty t^{r-1}\, dt}{(t + 1)^{r+s}} .$$

This is a transformation of the more usual definition

$$B(r,s) = \int_0^1 u^{r-1}(1 - u)^{s-1}\, du$$

accomplished by taking $t = u \div (1 - u)$ or $u = t \div (t + 1)$. The beta is related to the gamma by

$$B(r,s) = \frac{\Gamma(r)\Gamma(s)}{\Gamma(r + s)} .$$

As in the gamma case a distribution function F may be defined by the partial integral, i.e.,

$$F(x;r,s,\alpha,\beta) = \int_0^{(x/\beta)^\alpha} \frac{t^{r-1}\, dt}{(t + 1)^{r+s}} \div B(r,s).$$

This will be called the transformed beta distribution. Its density is

$$f(x;r,s,\alpha,\beta) = \frac{(\alpha/\beta)(x/\beta)^{\alpha r-1}}{B(r,s)(1 + (x/\beta)^\alpha)^{r+s}} .$$

For $r = 1$ the closed form

$$F(x;1,s,\alpha,\beta) = 1 - ((x/\beta)^\alpha + 1)^{-s}$$

results. This is coming to be known as the Burr distribution, and in turn has two special cases, namely $\alpha = 1$ which is the Pareto, and $s = 1$ which gives the log transform of the logistic. As the logistic is like a heavy-tailed normal the loglogistic can be thought of as being like a lognormal with heavier right and left tails. Its distribution function

$$F(x;\ 1,1,\alpha,\beta) = 1 - \frac{\beta^\alpha}{x^\alpha + \beta^\alpha}$$

is of particularly simple form.

The case $\alpha = 1$, i.e., $F(x;r,s,1,\beta)$ is a version of the transformed beta that has been investigated for severity applications. This will be called the generalized-F as its special case $\alpha = 1$, $\beta = s/r$ gives the F distribution where $2r$ and $2s$ are integers. The Pareto is also a special case of the generalized-F given by $r = 1$.

There is an interesting mixture property of the transformed gamma that generates a transformed beta, namely that with a population of transformed gamma random variables with fixed r and α, and with the transformed scale parameter λ^{α} itself gamma-distributed across the population, the compound process of picking a variable from the population, then taking a realization of that variable, is a transformed beta process. This is proved in Appendix D. Several corollary statements follow by taking the special cases of the transformed gamma (i.e., Weibull, gamma, and exponential) and mixing by a gamma, viz.,

(a) Weibull mixed by gamma yields Burr;
(b) Gamma mixed by gamma yields generalized-F;
(c) Exponential mixed by gamma yields Pareto;
(d) Weibull mixed by exponential yields loglogistic.

Exhibit 2 diagrams this situation.

Robert Hogg proved (a), (b), and (c) separately and Gary Patrik independently proved (c). The transformed beta and gamma distributions originally were developed in order to unify these results. Robert Miccolis pointed out that the generalized-F is a ratio of two gamma variates. This suggested the result, proved in Appendix E, that if X is transformed beta with parameters r, s, α, β, then $1/X$ is also, with parameters s, r, α, β^{-1}.

If X is transformed beta in r, s, α, β then

$$E(X^n) = \beta^n B(r + n/\alpha, s - n/\alpha) \div B(r,s)$$

if $-\alpha r < n < \alpha s$ and non-existent otherwise. This is an example of a distribution with unbounded moments for $n \geq \alpha s$ which arises in a natural way as a combination of distributions with all moments finite. For $\alpha = 1$ (generalized-F, Pareto) the moments simplify to

$$\beta^{-n}E(X^n) = (r) \frac{(r + 1) \times \cdots (r + n - 1)}{(s - 1) \times (s - 2) \cdots (s - n)} = \prod_{i=1}^{n} \frac{r + i - 1}{s - i} .$$

This makes methods of moments parameter estimation quite simple for this special case. Maximum likelihood parameter estimation for the transformed beta is similar to that for the transformed gamma as covered in Appendix H. Loss severity distributions also have been fit by the transformed beta and gamma distributions by matching sample and formula values of the excess ratio $R(a)$ in a manner similar to that in Harwayne [5].

As with the transformed gamma, the moment distributions are of the same form as the original distribution, in fact

$$\int_0^a x^n \, dF_x \div E(X^n) = F(a;r + n/\alpha, s - n/\alpha, \alpha, \beta).$$

Thus, as with the transformed gamma, a calculation of excess losses can be made if the cumulative distribution can be calculated. This has proven to be most practically accomplished through numerical integration. Appendix F discusses one method. The moment distribution formulas for the transformed beta and gamma show that the Burr and Weibull moment distributions do not maintain the original form, i.e., $r = 1$.

The mixture derivation of the transformed beta provides an interesting way to deal with so called "parameter risk." It is fairly plausible that aggregate losses for a given company (insured or insurer) are distributed transformed gamma and that the shape parameters r and α are fairly well known and stable but because of uncertain trend (or other factors) there is substantial uncertainty about the scale parameter λ, which relates to the overall level of expected results. If λ^α is gamma distributed in s and γ then the overall aggregate distribution is transformed beta in r, s, α, β where $\beta = \gamma^{1/\alpha}$. It also is not difficult to show that λ^α is gamma in s, α means that λ is transformed gamma in s, α, β (see Appendix G). Thus it can be concluded that if aggregate losses are transformed gamma in r, α, Λ where Λ is unknown but is itself transformed gamma in s, α, β (same α) then the aggregate losses are transformed beta in r, s, α, β.

In theory it would be a great coincidence if the uncertainty about λ had the same parameter α as did the aggregate losses themselves. As a practical technique for quantifying this uncertainty, however, it should not be too burdensome to use the α already in hand for aggregate losses. There will still be two parameters, s and β, available to match to the uncertainty the analyst feels is inherent.

There are several ways in which s and β could be determined. Different values could be tried and the 25th, 50th, and 75th percentile λ calculated for each, with the corresponding percentile of aggregate expected losses $\Gamma(r + 1/\alpha) \div \lambda\Gamma(r)$ following. These can be compared with the uncertainty that seems inherent in the overall level of losses. The latter uncertainty can be estimated by trying to combine the uncertainties in the trend, development, and other factors used to estimate the overall level. The regression statistics used in developing these factors may be helpful if regression was used.

Another approach to measuring the distribution of λ is using industry loss ratios. Expected losses for an aggregate loss distribution with cdf $G(x;r,\alpha,\lambda)$

are $\Gamma(r + (1/\alpha)) \div \lambda\Gamma(r)$. Thus, for fixed r,α, the reciprocal of the aggregate losses, and thus the reciprocal of the loss ratio, is proportional to λ. Therefore if λ is unknown but is a realization of a random variable Λ which is transformed gamma in s, α, β, where α is fixed, the shape parameter s can be estimated by looking at the historical distribution of loss ratio reciprocals. This would measure some of the variation that would occur even if Λ were known, however. An alternative is to look at some broader base of comparable experience, such as the line for the industry or state or class in question where the process variance is minimal and hence the principal source of variation is the parameter uncertainty. Depending on the similarity between the company in question and the broader base as to projection methods for trend and loss development, the stability of the historical data base, and so forth, this approach may give a reasonable estimate of the parameter uncertainty.

Estimating β then could proceed by matching the formula $E(1/\Lambda)$ for the transformed gamma distribution to the expected value of $1/\Lambda$ calculated for the year and company in question. For Λ with cdf $G(\lambda;s,\alpha,\beta)$ the $E(1/\Lambda)$ is $\beta\,\Gamma(s - 1/\alpha) \div \Gamma(s)$ from the transformed gamma moment formula.

Borrowing loosely from our earlier example, suppose a malpractice risk has aggregate losses distributed according to the transformed gamma with $r = .2478$, $\alpha = 1.470$ and $E(1/\Lambda) = 1 \div (1.144 \times 10^{-6})$, where Λ is transformed gamma in s, 1.470, β. Suppose the previous four years of industry malpractice experience produced loss ratios of .505, .750, 1.001, and 1.357, which have reciprocals 1.980, 1.333, .999, and .737. The reciprocals average 1.262 and have an unbiased sample standard deviation estimate of .5370 for an estimated CV of .4255. The formula

$$1 + CV^2 = \Gamma(s + 2/\alpha)\,\Gamma(s) \div \Gamma(s + 1/\alpha)^2 \text{ then becomes}$$

$$1.181 = \Gamma(s + 1.36)\,\Gamma(s) \div \Gamma(s + .68)^2,$$

which can be solved numerically to find $s = 2.597$. Then

$$\begin{aligned} 1 \div 1.144 \times 10^{-6} = E(1/\Lambda) &= \beta\,\Gamma(s - 1/\alpha) \div \Gamma(s) \\ &= \beta\,\Gamma(2.597 - .68) \div \Gamma(2.579) \end{aligned}$$

can be solved directly to yield $\beta = 1{,}288{,}500$. From the transformed beta in $r = .2478$, $s = 2.597$, $\alpha = 1.470$, $\beta = 1{,}288{,}500$ expected losses of

$$\frac{\beta\,\Gamma(r + 1/\alpha)\,\Gamma(s - 1/\alpha)}{\Gamma(r)\,\Gamma(s)} = 250{,}000$$

can be calculated, confirming the calculation of β.

The expected losses excess of $1 million in the aggregate increase substantially when this additional uncertainty is included. For this transformed beta an excess ratio of .1348 can be computed at $1,000,000 which yields excess expected losses of $33,700 compared to .0728 and $18,200 for the transformed gamma.

The great disparity between these figures comes from the wide divergence in loss ratios in the period studied. If the uncertainty in Λ really is so great that next year's loss ratio for the whole industry can come out anywhere in the range .50 to 1.35, then there is a much greater chance that total losses for a small segment of the industry will exceed the target $1 million.

For other more stable lines a similar analysis would show a much smaller difference. In those cases there is a danger that the potential variation in level would be understated by looking at industry loss ratios. Swings in calendar year ratios may be dampened by reserve changes. Also, a particular sector of the industry would probably have wider variation than the total industry in the degree to which the proper level could be projected. This would be important if the company under study were concentrated in one area. The selection of the parameter s probably should be made with a good deal of judgement because of these considerations.

SUMMARY AND EXTENSIONS

The above gives a method of approximating the distribution function of aggregate losses from the moments of that distribution, based on the transformed beta and gamma distributions. Since a distributional assumption is involved, the method is likely to be less precise than the exact methods of Adelson [11], Panjer [12] and Heckman and Meyers [13]. Those methods do, however, require more input information, namely the underlying frequency and severity distribution functions, and they also require substantially more computation. As computing becomes faster and less expensive and as good parameterized frequency and severity distributions become available those methods become increasingly viable, and the assumption of a distributional form for aggregate losses becomes more avoidable. Methods based on moments only are nonetheless of definite value at present.

The transformed beta distribution is a good candidate for casualty loss severity distributions, because it generalizes the Pareto and Burr which have been used with moderate success. The problems of trend and development by layer of loss have yet to be settled entirely in casualty lines, however, especially with regard to having factors that are independent of distributional assumptions. Thus, there currently is a fair amount of uncertainty as to casualty severity distributions.

The transformed gamma may be useful in loss severity, for example, in workers' compensation. Also, the inverse transformed gamma, i.e., the distribution of Y when $X = 1 \div Y$ is transformed gamma, is a heavy-tailed distribution which may have application to casualty loss severity. This distribution function is:

$$G'(y) = \int_0^{(y/\lambda)^\alpha} \frac{t^{-r-1} e^{-1 \div t}}{\Gamma(r)}\, dt$$

and $E(X^n) = \lambda^n\, \Gamma(r - n/\alpha) \div \Gamma(r)$ for $n < r\alpha$.

A problem that sometimes arises with maximum likelihood estimation with these distributions is that no maximum exists. Usually this happens because the maximum likelihood, given α, increases as α decreases. After some point the increase becomes negligible however. One alternative in this case is to pick a "low enough" value of α and maximize the likelihood fixing that value. This usually gives much better fits than the Weibull, gamma, Burr, etc., in these cases.

Another alternative is that there may be other functions that are limiting values of these distributions. For instance, in the Burr case, $F(x) = 1 - ((x/\beta)^\alpha + 1)^{-s}$, small α often leads to large β but with $(x/\beta)^\alpha$ near zero for the range of interest, so $1 + (x/\beta)^\alpha$ is close to $e^{(x/\beta)^\alpha}$ and $F(x)$ is approximately $1 - e^{-s(x/\beta)^\alpha}$ which is a Weibull. Conversely, small β and large α make (x/β) very close to $(x/\beta)^\alpha + 1$, relatively speaking, so $F(x)$ is approximately $1 - (x/\beta)^{-\alpha s}$, which is a non-shifted Pareto. Similar relationships may occur for the general cases.

A limitation of the above methods is that the transformed gamma does not seem able to take on any combination of moments. For example, it appears that the coefficient of skewness must be greater than the coefficient of variation (CV) if $CV > 1.25$. In the gamma case the coefficient of skewness is always twice the CV. Thus, the transformed gamma allows a fair amount of departure from gamma-ness but not complete latitude. Appendix J discusses this problem and suggests alternate approaches.

Much of the interest in the gamma stems from a 1940 theorem of Lundberg [14] which shows that under certain conditions the negative binomial frequency leads to an approximately gamma aggregate distribution. Since aggregate distributions seem to be positively skewed for the most part, but do not always have the skewness double the *CV*, gamma-like distributions allowing some deviation from the gamma are thus appealing candidates for this purpose.

Exhibit 3 gives the results of a test of the transformed gamma against an exact calculation of an aggregate distribution using the characteristic function method. The severity distribution is piecewise linear. Approximating the severity by a discrete distribution also permits a comparison to the recursive method of Adelson and Panjer. Intervals of $500 were chosen for this discrete approximation. Details are provided in Exhibit 3. The results show that the two exact methods are extremely similar, indicating that not much is lost by the discrete approximation to severity. The transformed gamma also is reasonably close over a wide range of loss sizes, confirming, at least in this one case, the usefulness of this simplifying approximation.

EXHIBIT 1

```
      ∇IG[⎕]∇
       ∇ E←V IG I;R;X;D
[1]   ⍝INCOMPLETE GAMMA FCT 0 TO X, PARAM R; I IS PRECISION SUGGEST~35 TO 350
[2]    X←V[1]
[3]    R←V[2]
[4]    →((R>55)∨(175<X)∨X>7E75*÷R+1)/BIG
[5]    D←((X*(R-1))×*(-X))÷!(R-1)
[6]    →END
[7]   ⍝BIG:R←1E¯12×⌊0.5+R×1000000000000
[8]   ⍝SOMETIMES ABOVE LINE NEEDED TO AVOID TRUNCATION PROBLEMS
[9]   BIG:D←(X*1|R)×(×/X÷(*X÷⌊R-1)×R-⍳⌊R-1)÷!1|R
[10]  END:E←D×+/×\X÷R+¯1+⍳I
       ∇
```

EXHIBIT 2

TRANSFORMED GAMMA MIXED BY GAMMA WITH SPECIAL CASES

$$\begin{array}{ccc}
\dfrac{1}{\Gamma(r)}\displaystyle\int_0^{\theta x^\alpha} t^{r-1}e^{-t}\,dt & & \dfrac{1}{\Gamma(r)}\displaystyle\int_0^{\theta x} t^{r-1}e^{-t}\,dt \\
\text{(Transformed Gamma)} & \xrightarrow[\alpha=1]{} & \text{(Gamma)} \\
\Big\downarrow r=1 & & \Big\downarrow r=1 \\
1-e^{-\theta x^\alpha} & & 1-e^{-\theta x} \\
\text{(Weibull)} & \xrightarrow{\alpha=1} & \text{(Exponential)}
\end{array}$$

If θ is distributed Gamma in s, γ:

$$\begin{array}{ccc}
\dfrac{1}{B(r,s)}\displaystyle\int_0^{(x/\beta)^\alpha} \dfrac{t^{r-1}\,dt}{(t+1)^{r+s}} & & \dfrac{1}{B(r,s)}\displaystyle\int_0^{(x/\beta)} \dfrac{t^{r-1}\,dt}{(t+1)^{r+s}} \\
\text{(Transformed Beta)} & \xrightarrow{\alpha=1} & \text{(Generalized-}F\text{)} \\
\Big\downarrow r=1 & & \Big\downarrow r=1 \\
1-((x/\beta)^\alpha+1)^{-s} & \xrightarrow[\alpha=1]{} & 1-(x/\beta+1)^{-s} \\
\text{(Burr)} & & \text{(Pareto)}
\end{array}$$

where $\beta = \gamma^{1/\alpha}$

EXHIBIT 3
PART 1

AGGREGATE LOSS DISTRIBUTIONS
COMPARATIVE SUMMARY

Aggregate Loss ($000)	Characteristic Function Method		Recursive Method		Transformed Gamma	
	Cumulative Probability	Excess Ratio	Cumulative Probability	Excess Ratio	Cumulative Probability	Excess Ratio
25	.0508	.9016	.0516	.9016	.0621	.9031
50	.1291	.8107	.1298	.8107	.1260	.8125
75	.2009	.7273	.2015	.7272	.1895	.7283
100	.2676	.6507	.2683	.6507	.2520	.6503
125	.3289	.5806	.3295	.5806	.3129	.5786
150	.3843	.5163	.3848	.5163	.3717	.5129
175	.4341	.4573	.4346	.4573	.4280	.4529
200	.4788	.4030	.4793	.4029	.4817	.3984
225	.5189	.3529	.5193	.3529	.5324	.3491
250	.5548	.3066	.5552	.3066	.5801	.3047
275	.6034	.2642	.6040	.2642	.6245	.2650
300	.6556	.2273	.6561	.2273	.6658	.2295
325	.7008	.1951	.7013	.1951	.7039	.1981
350	.7405	.1672	.7408	.1672	.7388	.1702
375	.7749	.1431	.7752	.1431	.7707	.1457
400	.8047	.1221	.8049	.1221	.7995	.1243
425	.8303	.1039	.8305	.1039	.8255	.1055
450	.8524	.0880	.8526	.0880	.8488	.0893
475	.8714	.0742	.8716	.0742	.8696	.0752
500	.8878	.0622	.8879	.0622	.8881	.0631
525	.9045	.0518	.9047	.0518	.9043	.0528
550	.9201	.0430	.9203	.0430	.9186	.0439
575	.9332	.0357	.9333	.0357	.9310	.0364
600	.9442	.0296	.9443	.0296	.9418	.0301
625	.9534	.0245	.9535	.0245	.9511	.0247
650	.9611	.0202	.9611	.0202	.9592	.0203
675	.9675	.0167	.9675	.0167	.9660	.0165
700	.9728	.0137	.9729	.0137	.9718	.0134
725	.9773	.0112	.9773	.0112	.9768	.0109
750	.9810	.0091	.9810	.0091	.9809	.0088
775	.9844	.0074	.9844	.0074	.9844	.0070
800	.9873	.0060	.9873	.0060	.9873	.0056
825	.9897	.0048	.9897	.0048	.9897	.0045
850	.9916	.0039	.9916	.0039	.9917	.0035

EXHIBIT 3
PART 2

AGGREGATE LOSS DISTRIBUTIONS
COMPARATIVE ASSUMPTIONS

Frequency: Poisson $\lambda = 13.7376$

Piecewise Linear *CDF*

Limit (000)	Cumulative Probability	Limit (000)	Cumulative Probability
1	.38935	25	.85690
5	.77870	35	.87927
6	.78438	50	.90280
7	.78981	75	.92739
8	.79498	100	.94256
9	.79993	125	.95277
10	.80466	150	.96009
12.5	.81564	175	.96556
15	.82553	200	.96979
17.5	.83449	225	.97316
20	.84264	250	.97590

Discrete *PDF*

Amount	Probability
500	.38326640625
1000	.03041796875
1500 to 4000	.04866875 each 500
4500	.054731628
5000	.019691497
5500 to 249,000 at each N = 500k	Piecewise linear probability from $N - 250$ to $N + 250$
249,500	.0000685
250,000	.0241137

Moments

	Mean	Coefficient of Variation	Coefficient of Skewness
Severity	18,198	2.6600	3.6746
Aggregate	250,000	.7667	1.0744

Transformed Gamma Parameters

r : .5613125

α : 1.8300318

λ : $1 \div 417896.414$

APPENDIX A

SOLVING TWO EQUATIONS

Many systems of two equations in two unknowns, including the transformed gamma moment system in the text, can be solved by Newton-Raphson iteration, with the partial derivatives taken numerically. The numerical partial derivative of $f(x,y)$ with respect to y, for example, is $(f(x,y(1 + \Delta)) - f(x,y)) \div y\Delta$, where Δ is a small number; e.g., 10^{-7}. Because of limits to computer accuracy, Δ should not be too small, e.g., $\Delta = 10^{-50}$ would be too small for most computer installations. This method is quite useful when the partials are not available in closed form or are excessively intricate.

Given $f(x,y)$ and $g(x,y)$, initial estimates x_0 and y_0 and derivatives f_x, f_y, g_x, g_y the iteration proceeds by setting

$$x_{i+1} = x_i - (fg_y - gf_y) \div (f_x g_y - g_x f_y)$$
$$y_{i+1} = y_i - (gf_x - fg_x) \div (f_x g_y - g_x f_y)$$

where the functions and derivatives are evaluated at (x_i,y_i). See Conte and de Boor [3] page 86 for details.

Exhibit A1 gives an *APL* system for this procedure. The user interactively defines the equations to be solved. Any user-defined functions may be called in this process. A sample run of the system is shown in Exhibit A2.

EXHIBIT A1
PAGE 1

```
∇DELUXENR[⎕]∇
      ∇ DELUXENR;ΔA;ΔB;LOOPTOL;DELTOL;MODEFLAG;PFQA;PFQB;PGQA;PGQB
[1]   ⍝WRITTEN BY STAN STIEFEL
[2]    'SPECIFY ONE FUNCTIONAL RELATION. . .'
[3]    'USE THE VARIABLE NAMES A AND B FOR THE UNKNOWNS.'
[4]    'FQ' MAKEFX ⍞
[5]    'SPECIFY THE OTHER RELATION'
[6]    'GQ' MAKEFX ⍞
[7]    'ENTER INITIAL VALUE FOR A'
[8]    A←⎕
[9]    'ENTER INITIAL VALUE FOR B'
[10]   B←⎕
[11]   MODEFLAG←1↑,DELTOL-DELTOL←LOOPTOL←0.00001
[12]   'WOULD YOU LIKE TO USE DEFAULT CONDITIONS (0)'
[13]   'OR SEE A MENU OF OPTIONS (1). . .0 OR 1'
[14]   ⍎⎕/'MENU'
[15]  LP:PARTIALS DELTOL
[16]   A←A-ΔA←(DET(2 2 ρ(A FQ B),PFQB,(A GQ B),PGQB))÷DET(2 2 ρPFQA,PFQB,PGQA,PGQB)
[17]   ⍎MODEFLAG/'PARTIALS DELTOL'
[18]   B←B-ΔB←(DET(2 2 ρPFQA,(A FQ B),PGQA,(A GQ B)))÷DET(2 2 ρPFQA,PFQB,PGQA,PGQB)
[19]   →(∨/LOOPTOL<|(ΔA,ΔB)÷(A,B)+0=A,B)/LP
[20]   'A:  ';A;'     B:    ';B
[21]   ⎕WA←⎕EX 2 2 ρ'FQGQ'
       ∇


      ∇MAKEFX[⎕]∇
      ∇ NAME MAKEFX RELAT;X;TITLE
[1]    →(0='='∊RELAT)/DID
[2]    RELAT[RELATι'=']←'-'
[3]   DID:TITLE←'RSLT',NAME,'←A ',NAME,' B'
[4]    RELAT←'RSLT',NAME,'←',RELAT
[5]    RELAT←RELAT,(0.5×X+|X←(ρTITLE)-ρRELAT)ρ' '
[6]    TITLE←TITLE,((ρRELAT)-(ρTITLE))ρ' '
[7]    ⎕WA←⎕FX TITLE,[0.5] RELAT
       ∇
```

EXHIBIT A1
PAGE 2

```
∇MENU[⎕]∇
      ∇ MENU
[1]     'FOR PURPOSES OF TAKING NUMERICAL DERIVATIVES, FUNCTIONS WILL BE EVALUATED AT A,  A-ΔA, B, B-ΔB.'
[2]     'ΔA AND ΔB ARE SPECIFIED AS FRACTIONS OF A AND B. . .1E¯5 IS THE DEFAULT.    PLEASE SPECIFY THE FRACTION.'
[3]     DELTOL←⎕
[4]     'ITERATION WILL BE CONSIDERED COMPLETE WHEN BOTH A AND B HAVE CHANGED BY LESS THAN SOME FRACTION OF THEMSELVES'
[5]     'DEFAULT IS 1E¯5.   PLEASE SPECIFY THE FRACTION.'
[6]     LOOPTOL←⎕
[7]     'SEQUENCE OF CALCULATION CAN BE EITHER OF TWO OPTIONS'
[8]     '(0) GET PARTIALS, GET NEW A, GET NEW B.'
[9]     '(1) GET PARTIALS, GET NEW A, GET PARTIALS, GET NEW B.'
[10]    'DEFAULT IS 0.  PLEASE SPECIFY 0 OR 1.'
[11]    MODEFLAG←⎕
      ∇


      ∇PARTIALS[⎕]∇
      ∇ PARTIALS XXXX;Z
[1]     PFQA←((A FQ B)-((A-Z) FQ B))÷Z←1E¯10⌈|Z←XXXX×A
[2]     PGQA←((A GQ B)-((A-Z) GQ B))÷Z
[3]     PFQB←((A FQ B)-(A FQ(B-Z)))÷Z←1E¯10⌈|Z←XXXX×B
[4]     PGQB←((A GQ B)-(A GQ(B-Z)))÷Z
      ∇


      ∇DET[⎕]∇
      ∇ Y←DET X
[1]     Y←(X[1;1]×X[2;2])-X[1;2]×X[2;1]
      ∇
```

EXHIBIT A2

```
      ∇CV[⎕]∇
     ∇ Y←A CV R
[1]   Y←(!¯1+R)×(!¯1+R+2÷A)+(!¯1+R+÷A)*2
[2]   Y←(Y-1)*0.5
     ∇



      ∇SKW[⎕]∇
     ∇ Y←A SKW R
[1]   Y←((!¯1+R)*2)×(!¯1+R+3÷A)÷(!¯1+R+÷A)*3
[2]   Y←Y+2-3×(!¯1+R)×(!¯1+R+2÷A)÷(!¯1+R+÷A)*2
[3]   Y←Y÷(A CV R)*3
     ∇


      DELUXENR
SPECIFY ONE FUNCTIONAL RELATION. . .
USE THE VARIABLE NAMES A AND B FOR THE UNKNOWNS.
(A CV B)=1.409
SPECIFY THE OTHER RELATION
(A SKW B)=2.344
ENTER INITIAL VALUE FOR A
⎕:
      1.2
ENTER INITIAL VALUE FOR B
⎕:
      .3

WOULD YOU LIKE TO USE DEFAULT CONDITIONS (0)
OR SEE A MENU OF OPTIONS (1). . .0OR 1
⎕:
      0
A:  1.47     B:  0.2478

      A CV B
1.409
      A SKW B
2.344



      ∇SKEW2[⎕]∇
     ∇ Y←A SKEW2 R
[1]   N←⌊¯1+R+÷A
[2]   M←(-⍳N)+R+÷A
[3]   O←!(-N+1)+R+÷A
[4]   S←(-⍳N)+R+3÷A
[5]   T←!(-N+1)+R+3÷A
[6]   U←(×/S+M)×(T+O)
[7]   Y←((!¯1+R)*2)×U÷(!¯1+R+÷A)*2
[8]   Y←Y+2-3×(!¯1+R)×(!¯1+R+2÷A)÷(!¯1+R+÷A)*2
[9]   Y←Y÷(A CV R)*3
     ∇
```

APPENDIX B

MAXIMUM LIKELIHOOD FOR THE TRANSFORMED GAMMA

Maximum likelihood in the case where there are no problems of truncation or censorship of the sample reduces to one non-linear equation to solve for α, then linear equations for r and λ. The α equation is somewhat intricate but is solved easily numerically. Given a sample y_i, $i = 1$ to n, the likelihood function is

$$L(r,\alpha,\lambda) = \prod_{i=1}^{n} \alpha\lambda^{\alpha r} y_i^{\alpha r-1} e^{-(\lambda y_i)^\alpha} \div \Gamma(r) \text{ and}$$

$$\ln L(r,\alpha,\lambda) = n \ln \alpha + n\, \alpha\, r \ln \lambda - n \ln \Gamma(r)$$
$$+ (\alpha r - 1) \Sigma \ln y_i - \lambda^\alpha \sum_{i=1}^{n} y_i^\alpha .$$

Setting the partial derivatives of this to zero, and denoting the derivative of $\ln \Gamma(r)$ by $\psi(r)$ yields the likelihood equations:

(a) $\psi(r) - \ln r = \alpha\, \overline{\ln y} - \ln \overline{y^\alpha}$
(b) $r = \overline{y^\alpha} \div \alpha(\overline{y^\alpha \ln y} - \overline{y^\alpha}\; \overline{\ln y})$
(c) $\lambda = (\overline{y^\alpha} \div r)^{1/\alpha}$

Substituting for r in (a) via (b) gives a single equation for α which when solved allows r and λ to be calculated from (b) and (c). This is a generalization of the method found in Hachemeister [4] for the gamma distribution. Note that to solve (a),

$$\overline{y^\alpha} = \frac{1}{n} \sum_{i=1}^{n} y_i^\alpha, \quad \overline{\ln y} = \frac{1}{n} \sum_{i=1}^{n} \ln y_i,$$

$$\text{and } \overline{y^\alpha \ln y} = \frac{1}{n} \sum_{i=1}^{n} y_i^\alpha \ln y_i,$$

must be calculated from the sample at each iteration.

As suggested on page 152 of Aquino [2], differentiating Abramowitz and Stegun's [1] formula 6.1.34 (page 256) gives the series approximation

$$\psi(z) = \Gamma(z) \sum_{k=1}^{26} kc_k\, z^{k-1},$$

where c_1 to c_{26} are as shown in Exhibit B1. This expansion gives more than 13 place accuracy on [1,2] and the recursive relation $\psi(1 + z) = \psi(z) + 1/z$ can be used outside of this interval.

To solve equation (a) with (b) substituted for r we have an equation $f(\alpha) = 0$ where f is calculable by computer or calculator. This can be solved iteratively by numerical Newton-Raphson:

Start with a guess α_0. Then let

$$\alpha_{i+1} = \alpha_i - \frac{f(\alpha_i)}{\dfrac{f(\alpha_i\,(1 + \Delta)) - f(\alpha_i)}{\alpha_i}}$$

$$\text{i.e. } \alpha_{i+1} = \alpha_i \left(1 - \frac{\Delta}{\dfrac{f(\alpha_i\,(1 + \Delta))}{f(\alpha_i)} - 1} \right)$$

where Δ is small, e.g. 10^{-7}.

A reasonable starting value α_0 usually is given by calculating the sample ratio of the coefficient of variation over half the coefficient of skewness, as this is greater, less than, or equal to 1 when α is.

As an alternative, the secant method

$$\alpha_{i+1} = \alpha_i - \frac{f(\alpha_i)\,(\alpha_i - \alpha_{i-1})}{f(\alpha_i) - f(\alpha_{i-1})}$$

can be used to solve for α. This involves only one computation of f each iteration, so it may be faster than Newton-Raphson iteration.

EXHIBIT B1

SERIES EXPANSION FOR $\psi(z)$

$$\psi(z) = \Gamma(z) \sum_{k=1}^{26} k c_k z^{k-1}$$

k	c_k
1	−1.00000 00000 000000
2	−0.57721 56649 015329
3	0.65587 80715 202538
4	0.04200 26350 340952
5	−0.16653 86113 822915
6	0.04219 77345 555443
7	0.00962 19715 278770
8	−0.00721 89432 466630
9	0.00116 51675 918591
10	0.00021 52416 741149
11	−0.00012 80502 823882
12	0.00002 01348 547807
13	0.00000 12504 934821
14	−0.00000 11330 272320
15	0.00000 02056 338417
16	−0.00000 00061 160950
17	−0.00000 00050 020075
18	0.00000 00011 812746
19	−0.00000 00001 043427
20	−0.00000 00000 077823
21	0.00000 00000 036968
22	−0.00000 00000 005100
23	0.00000 00000 000206
24	0.00000 00000 000054
25	−0.00000 00000 000014
26	−0.00000 00000 000001

APPENDIX C

AGGREGATE MOMENTS

A. In terms of frequency and severity moments, assume individual claim sizes are independent, identically distributed, and independent of the number of claims.

Let N denote number of claims, X claim size, L aggregate losses, μ the mean, σ the standard deviation, γ the coefficient of skewness, c the coefficient of variation, and

$$N_i = \frac{E\,(N - \mu_N)^i}{\mu_N}.$$

Then

$$\mu_L = \mu_X \mu_N$$

$$\sigma_L^2 = \mu_N \sigma_X^2 + (\mu_X \sigma_N)^2$$

$$\gamma_L \sigma_L^3 = \mu_N \gamma_X \sigma_X^3 + 3\mu_X \sigma_X^2 \sigma_N^2 + \mu_X^3 \gamma_N \sigma_N^3$$

$$\sigma_L^2 = \mu_X^2 \mu_N (c_X^2 + N_2)$$

$$\gamma_L = (\gamma_X c_X^3 + 3c_X^2 N_2 + N_3) \div \sqrt{\mu_N (c_X^2 + N_2)^3}$$

$$c_L^2 = (c_X^2 + N_2) \div \mu_N$$

B. Moments of conditional (positive) distribution in terms of moments of entire distribution and probability of losses being non-zero

$$F(a) = \Pr\,(L \le a) = \begin{cases} 1 - p \text{ when } a = 0 \\ 0 \qquad \text{when } a < 0 \\ (1 - p) + pG(a) \text{ when } a > 0 \end{cases}$$

Then

$$\mu_F = p\mu_G$$

$$\mu_G = \mu_F \div p$$

$$c_G^2 = p\, c_F^2 + p - 1$$

$$\gamma_G = \frac{p^2 \gamma_F\, c_F^3 + (p - 1)(3pc_F^2 + p - 2)}{c_G^3}$$

EXHIBIT C1

```
      ∇CONDITMO[⎕]∇
      ∇ X←P CONDITMO Q;TERM1;TERM2;COEFVAR;GAMMA
[1]   ⍝WRITTEN BY VICTOR PUGLISI
[2]   ⍝  THIS PROGRAM CALCULATES CONDITIONAL MOMENTS IN THE FORM OF THE COEFFICIENT OF VARIATION (CV) AND THE SKEWNESS
[3]   ⍝  (GAMMA) BASED UPON RISKMODEL OUTPUT FOR THE PART OF THE DISTRIBUTION GREATER THAN 0.
[4]   ⍝  IT TAKES AS LEFT-HAND ARGUMENT THE PROBABILITY OF CLAIMS BEING LARGER THAN 0, CURRENTLY FOUND AT THE TOP OF THE
[5]   ⍝  RISKMODEL OUTPUT FOR EACH LAYER DENOTED BY 'PROBABILITY OF LOSS' AND FOR RIGHT-HAND ARGUMENT REQUIRES A TWO
[6]   ⍝  ELEMENT VECTOR CONSISTING OF THE COEFFICIENT OF VARIATION AND THE COEFFICIENT OF SKEWNESS FOR EACH MAJOR GROUP.
[7]   ⍝  THESE ARE FOUND IN COLUMNS 8 AND 9 RESPECTIVELY OF THE RISKMODEL OUTPUT.
[8]    COEFVAR←((P×Q[1]*2)+P-1)*0.5
[9]    TERM1←(P*2)×Q[2]×Q[1]*3
[10]   TERM2←(P-1)×(3×P×Q[1]*2)+P-2
[11]   GAMMA←(TERM1+TERM2)÷COEFVAR*3
[12]   X←COEFVAR,GAMMA
      ∇
```

APPENDIX D

TRANSFORMED BETA IS TRANSFORMED GAMMA MIXED BY A GAMMA

The transformed gamma density function

$$g(x;r,\alpha,\lambda) = \frac{\alpha\lambda^{\alpha r}\, x^{\alpha r-1}\, e^{-\lambda^\alpha x^\alpha}}{\Gamma(r)}$$

can also be parameterized as $\alpha\, \theta^r x^{\alpha r-1} e^{-\theta x^\alpha} \div \Gamma(r)$, taking $\theta = \lambda^\alpha$. Given a family of such random variables with α and r fixed and θ itself gamma distributed with parameters s and γ, i.e., having density $\gamma^s\, \theta^{s-1}\, e^{-\gamma\theta} \div \Gamma(s)$, then the compound process is transformed beta.

To demonstrate this the density for the compound distribution will be calculated. This is the probability-weighted average of the densities of the family, that at x equals:

$$\int_0^\infty \frac{\alpha\, \theta^r x^{\alpha r-1}\, e^{-\theta x^\alpha}}{\Gamma(r)} \, \frac{\gamma^s \theta^{s-1}\, e^{-\gamma\theta}\, d\theta}{\Gamma(s)}$$

$$= \frac{\alpha\gamma^s\, x^{\alpha r-1}}{\Gamma(r)\,\Gamma(s)} \int_0^\infty \theta^{r+s-1}\, e^{-\theta(x^\alpha+\gamma)}\, d\theta$$

which, after the change of variable $\phi = \theta(x^\alpha + \gamma)$, becomes

$$\frac{\alpha\gamma^s\, x^{\alpha r-1}}{\Gamma(r)\,\Gamma(s)} \int_0^\infty \left(\frac{\phi}{x^\alpha + \gamma}\right)^{r+s-1} e^{-\phi}\, \frac{d\phi}{x^\alpha + \gamma}$$

$$= \frac{\alpha\gamma^s\, x^{\alpha r-1}}{\Gamma(r)\,\Gamma(s)(x^\alpha + \gamma)^{r+s}} \int_0^\infty \phi^{r+s-1}\, e^{-\phi}\, d\phi$$

$$= \frac{\alpha\gamma^s\, x^{\alpha r-1}}{\Gamma(r)\,\Gamma(s)(x^\alpha + \gamma)^{r+s}}\, \Gamma(r + s)$$

$$= \frac{\alpha\gamma^s\, x^{\alpha r-1}}{B(r,s)(x^\alpha + \gamma)^{r+s}}\,.$$

Now defining β by $\gamma = \beta^\alpha$ gives for the compound density

$$\frac{\alpha\beta^{\alpha s}\, x^{\alpha r-1}}{B(r,s)(x^\alpha + \beta^\alpha)^{r+s}} = \frac{\beta^{-\alpha(r+s)}\,\alpha\beta^{\alpha s}\, x^{\alpha r-1}}{B(r,s)((x/\beta)^\alpha + 1)^{r+s}}$$

$$= (\alpha/\beta)(x/\beta)^{\alpha r-1} \div B(r,s)((x/\beta)^\alpha + 1)^{r+s}$$

which is the transformed beta density.

APPENDIX E

RECIPROCAL OF TRANSFORMED BETA VARIATE IS TRANSFORMED BETA

Let $Y = \dfrac{1}{X}$ where X has cdf $F(x;r,s,\alpha,\beta)$.

Now $Y \le a \rightleftharpoons X \ge (1/a)$ so $\Pr(Y \le a) = 1 - \Pr(X < (1/a)$

$$= 1 - \frac{1}{B(r,s)} \int_0^{(a\beta)^{-\alpha}} \left[\frac{t^{r-1}}{(1 + t)^{r+s}}\right] dt.$$

Let $u = (1/t)$; $t = (1/u)$; $dt = -du/u^2$.

$$\text{Then } \Pr(Y \le a) = 1 + \frac{1}{B(r,s)} \int_\infty^{(a\beta)^\alpha} \frac{u^{1-r}\, du}{(1 + (1/u)^{r+s}\, u^2}$$

$$= 1 - \frac{1}{B(r,s)} \int_{(a\beta)^\alpha}^{\infty} \left[\frac{u^{1+s-2}}{(u + 1)^{r+s}}\right] du$$

$$= \frac{1}{B(r,s)} \int_0^{(a\beta)^\alpha} \left[\frac{u^{s-1}}{(u + 1)^{r+s}}\right] du.$$

Therefore Y has cdf $f(y;s,r,\alpha,1/\beta)$.

APPENDIX F

NUMERICAL INTEGRATION BY GAUSSIAN QUADRATURE

Gaussian quadrature is a method of numerical integration that estimates the integral by taking a weighted sum of the value of the function being integrated at several points. In general

$$\int_a^b f(y)dy \approx \frac{b-a}{2} \sum_{i=1}^{n} W_i f(y_i),$$

where $2y_i = (b - a)x_i + b + a$ and W_i and x_i are somewhat complex to calculate. Exhibits F1 and F2 give W_i and x_i for a few values of n. See Abramowitz [1] pages 916–919 for others. Hildebrand [8] discusses the mathematical background.

This approach works best for functions that can be closely approximated by polynomials of degree n.

The integration of the transformed beta distribution function is more accurate if two transformations are made. First the mapping $u = t/(t + 1)$ transforms the integral to

$$F(x;r,s,\alpha,\beta) = \int_0^{x^\alpha/x^\alpha+\beta^\alpha} [u^{r-1} (1 - u)^{s-1}]\, du \div B(r,s)$$

$$= IB\left(\frac{x^\alpha}{x^\alpha + \beta^\alpha}\, ; r,s\right),$$

which can be taken as the definition of the function IB. However, the approximation of this integral by the above quadrature formula is not close for small values of r and s, e.g., below 1. A recurrence relation was derived to express $IB(x;r,s)$ as a function of $IB\,(x;r + 1, s + 1)$, putting the integral to be solved in a more satisfactory area. This relationship is $rsIB(x;r,s) = x^r(1 - x)^s\,(s - (r + s)x) + (r + s + 1)(r + s)\, IB(x;r + 1, s + 1)$, and was derived by George Phillips from Abramowitz's [1] formulas 26.5.2 and 26.5.16 on page 944. In practice this formula is applied thrice to get to the $r + 3$, $s + 3$ level. Exhibit F3 gives a series of *APL* programs which performs the calculation of $F(x;r,s,\alpha,\beta)$.

EXHIBIT F1

ABSCISSAS AND WEIGHTS FOR n POINT GAUSSIAN QUADRATURE

$n=6$

x_i			W_i		
±0.23861	91860	83197	0.46791	39345	72691
±0.66120	93864	66265	0.36076	15730	48139
±0.93246	95142	03152	0.17132	44923	79170

$n=10$

±0.14887	43389	81631	0.29552	42247	14753
±0.43339	53941	29247	0.26926	67193	09996
±0.67940	95682	99024	0.21908	63625	15982
±0.86506	33666	88985	0.14945	13491	50581
±0.97390	65285	17172	0.06667	13443	08688

$n=24$

±0.06405	68928	62606	0.12793	81953	46752
±0.19111	88674	73616	0.12583	74563	46828
±0.31504	26796	90163	0.12167	04729	27803
±0.43379	35076	26045	0.11550	56680	53726
±0.54542	14713	88840	0.10744	42701	15966
±0.64809	36519	36976	0.09761	86521	04114
±0.74012	41915	78554	0.08619	01615	31953
±0.82000	19859	73903	0.07334	64814	11080
±0.88641	55270	04401	0.05929	85849	15437
±0.93827	45520	02733	0.04427	74388	17420
±0.97472	85550	71309	0.02853	13886	28934
±0.99518	72199	97021	0.01234	12297	99987

EXHIBIT F2

$n = 96$

	X_i	W_i		X_i	W_i
1	−.999689503883231	.000796792065552	49	.016276744849603	.032550614492363
2	−.998364375863182	.001853960788947	50	.048812985136050	.032516118713869
3	−.995981842987209	.002910731817935	51	.081297495464426	.032447163714064
4	−.992543900323763	.003964554338445	52	.113695850110666	.032343822568576
5	−.988054126329624	.005014202742928	53	.145973714654897	.032206204794030
6	−.982517263563015	.006058545504236	54	.178096882367619	.032034456231993
7	−.975939174585136	.007096470791154	55	.210031310460567	.031828758894411
8	−.968326828463264	.008126876925698	56	.241743156163840	.031589330770727
9	−.959688291448743	.009148671230783	57	.273198812591049	.031316425596861
10	−.950032717784438	.010160770535008	58	.304364944354496	.031010332586314
11	−.939370339752755	.011162102099838	59	.335208522892625	.030671376123669
12	−.927712456722309	.012151604671088	60	.365696861472314	.030299915420828
13	−.915071423120898	.013128229566962	61	.395797649828909	.029896344136328
14	−.901460635315852	.014090941772315	62	.425478988407301	.029461089958168
15	−.886894517402420	.015038721026995	63	.454709422167743	.028994614150555
16	−.871388505909297	.015970562902562	64	.483457973920596	.028497411065085
17	−.854959033434601	.016885479864245	65	.511694177154668	.027970007616848
18	−.837623511228187	.017782502316045	66	.539388108324357	.027412962726029
19	−.819400310737932	.018660679627411	67	.566510418561397	.026826866725592
20	−.800308744139141	.019519081140145	68	.593032364777572	.026212340735672
21	−.780369043867433	.020356797154333	69	.618925840125469	.025570036005349
22	−.759602341176647	.021172939892191	70	.644163403784967	.024900633222484
23	−.738030643744400	.021966644438744	71	.668718310043916	.024204841792365
24	−.715676812348968	.022737069658329	72	.692564536642172	.023483399085926
25	−.692564536642172	.023483399085926	73	.715676812348968	.022737069658329
26	−.668718310043916	.024204841792365	74	.738030643744400	.021966644438744
27	−.644163403784967	.024900633222484	75	.759602341176647	.021172939892191
28	−.618925840125469	.025570036005349	76	.780369043867433	.020356797154333
29	−.593032364777572	.026212340735672	77	.800308744139141	.019519081140145
30	−.566510418561397	.026826866725592	78	.819400310737932	.018660679627411
31	−.539388108324357	.027412962726029	79	.837623511228187	.017782502316045
32	−.511694177154668	.027970007616848	80	.854959033434601	.016885479864245
33	−.483457973920596	.028497411065085	81	.871388505909297	.015970562902562
34	−.454709422167743	.028994614150555	82	.886894517402420	.015038721026995
35	−.425478988407301	.029461089958168	83	.901460635315852	.014090941772315
36	−.395797649828909	.029896344136328	84	.915071423120898	.013128229566962
37	−.365696861472314	.030299915420828	85	.927712456722309	.012151604671088
38	−.335208522892625	.030671376123669	86	.939370339752755	.011162102099838
39	−.304364944354496	.031010332586314	87	.950032717784438	.010160770535008
40	−.273198812591049	.031316425596861	88	.959688291448743	.009148671230783
41	−.241743156163840	.031589330770727	89	.968326828463264	.008126876925698
42	−.210031310460567	.031828758894411	90	.975939174585136	.007096470791154
43	−.178096882367619	.032034456231993	91	.982517263563015	.006058545504236
44	−.145973714654897	.032206204794030	92	.988054126329624	.005014202742928
45	−.113695850110666	.032343822568576	93	.992543900323763	.003964554338445
46	−.081297495464426	.032447163714064	94	.995981842987209	.002910731817935
47	−.048812985136050	.032516118713869	95	.998364375863182	.001853960788947
48	−.016276744849603	.032550614492363	96	.999689503883231	.000796792065552

EXHIBIT F3

```
∇TBXR[⎕]∇
      ∇ Y←X TBXR AGD;B;A;G;D
[1]   ⍝TRANSFORMED BETA XS RATIO AT X TIMES MEAN
[2]    A←AGD[1]
[3]    G←AGD[2]
[4]    D←AGD[3]
[5]    B←(G CBETA D)÷(G+÷A) CBETA D-÷A
[6]    Y←X TBETAXR A,B,G,D
      ∇

      ∇TBETAXR[⎕]∇
      ∇ Y←X TBETAXR ABGD;A;B;G;D;H;L
[1]   ⍝TRANSFORMED BETA XS RATIO PARAMS A B G D
[2]    A←ABGD[1]
[3]    B←ABGD[2]
[4]    G←ABGD[3]
[5]    D←ABGD[4]
[6]    L←L÷1+L←(X÷B)*A
[7]    Y←(H←(G+÷A) CBETA D-÷A)-L IB(G+÷A),D-÷A
[8]    Y←Y-X×(÷B)×(G CBETA D)-L IB G,D
[9]    Y←Y÷H
      ∇

      ∇CBETA[⎕]∇
      ∇ Y←V CBETA W
[1]   ⍝COMPLETE BETA OF V AND W
[2]    Y←÷(V!V+W)×V×W÷V+W
      ∇

      ∇IB[⎕]∇
      ∇ R←X IB AB;Y1;Y2;Y3;Y4;Y5;A;B
[1]   ⍝ WRITTEN BY GEORGE PHILLIPS
[2]    A←AB[1]×1 ⎕CC 0
[3]    B←AB[2]
[4]    Y1←¯1↑×\(X,1-X)*AB
[5]    Y2←((B-1)+⍳3)-X×(A+B-2)+2×⍳3
[6]    Y3←(X×1-X)* 0 1 2
[7]    Y4←1,×\(A+B-1)+⍳6
[8]    Y5←×\(1,(A+1),A+2)×1,(B+1),B+2
[9]    R←(÷A×B)×(Y1×+/Y2×Y3×Y4[1 3 5]÷Y5)+(Y4[7]÷Y5[3])×(X INCBETA(A+3),B+3)
      ∇

      ∇INCBETA[⎕]∇
      ∇ RST←A INCBETA VW
[1]   ⍝WRITTEN GREGG EVANS
[2]    RST←1 GSQD '(X*VW[1]-1)×(1-X)*VW[2]-1/DX'
      ∇

      ∇GSQD[⎕]∇
      ∇ RST←X GSQD Y;A̲;B̲;C̲;D̲;E̲;F̲
[1]   ⍝WRITTEN BY GREGG EVANS
[2]    F̲←(E̲←⍎,(22⍴ 1 0)\(0 11)↓(11 0)⍕0,[1.5] D̲+(10*10)×~D̲←C̲≠¯1↑Y)\C̲←(∧\Y≠'/')/Y
[3]    C̲←'(X÷2)×',F̲,0↑F̲[D̲]←(⍴D̲←,(¯1+D̲[;1]),D̲←(B̲,(⍴D̲)÷B̲←(+/~E̲)÷10)⍴D̲←(0=E̲)/⍳⍴,E̲)⍴'(0.5×X×A̲+1)'
[4]    RST←+/(2÷+/GSQDVAR[2;])×GSQDVAR[2;]×⍎C̲,0↑A̲←GSQDVAR[1;]
      ∇
```

APPENDIX G

RELATIONSHIP BETWEEN GAMMA AND TRANSFORMED GAMMA

To show: Λ^α is gamma in s, γ if and only if Λ is transformed gamma in s,α,β where $\beta = \gamma^{1/\alpha}$.

Note that $\Pr(\Lambda \leq \lambda) = \Pr(\Lambda^\alpha \leq \lambda^\alpha)$

$$= G(\lambda^\alpha;s,1,\gamma)$$

$$= \int_0^{\gamma\lambda^\alpha} t^{s-1}\, e^{-t}\, dt$$

$$= \int_0^{(\beta\lambda)^\alpha} t^{s-1}\, e^{-t}\, dt = G(\lambda;s,\alpha,\beta).$$

APPENDIX H

MAXIMUM LIKELIHOOD ESTIMATORS FOR TRANSFORMED BETA PARAMETERS

Given a sample $x_1, \ldots, x_n$, fitting the parameters r, s, α, and β of the transformed beta by maximum likelihood involves finding the maximum of the log-likelihood function

$$\ln L(r, s, \alpha, \beta) = n \ln \Gamma(r + s) + n \ln \alpha + (\alpha r - 1) \sum_{i=1}^{n} \ln x_i$$

$$- (n\alpha r \ln \beta + n \ln \Gamma(r) + n \ln \Gamma(s) + (r + s) \sum_{i=1}^{n} \ln(1 + x_i/\beta)^{\alpha}.$$

As with the transformed gamma let the derivative of $\ln \Gamma(x)$ be denoted $\psi(x)$. Dividing the partials of $\ln L$ by n and setting to zero gives the following 4 equations:

(r): $\psi(r + s) = \psi(r) + \overline{\ln(1 + \beta/x_i)^{\alpha})}$

(s): $\psi(r + s) = \psi(s) + \overline{\ln(1 + x_i/\beta)^{\alpha})}$

(α): $1/\alpha + r\,\overline{\ln(x_i/\beta)} = (r + s)\overline{(\ln(x_i/\beta))(\beta/x_i)^{\alpha} + 1)^{-1}}$

(β): $r = (r + s)\overline{(1 + (\beta/x_i)^{\alpha})^{-1}}$

where the bar denotes the average over the sample of the barred function.

The (α) and (β) equations are linear in r and s, so they can be solved to yield r and s as functions of α and β. These can be substituted into the (r) and (s) equations to give two non-linear equations in two unknowns (α,β) which can be solved by the methods of Appendix A.

An *APL* system for solving these equations is shown in Exhibit H1 and a run with sample data in Exhibit H2.

EXHIBIT H1

APL PROGRAMS FOR TRANSFORMED BETA MLE

```
    ∇NRFN[⎕]∇
      ∇ AB1←V NRFN AB;YA;YB;J;Z
     ⍝WRITTEN BY GARY VENTER
     ⍝NEWTON RAPHSON ITERATION FOR TRBET PARAMS, SAMPLE IN V
      AB1←AB
      Z←1E¯7
     TOP:AB←AB1
      Y←V FN AB
      YA←V FN(AB[1]×1+Z),AB[2]
      YB←V FN AB[1],AB[2]×1+Z
      YA←(YA-Y)÷Z×AB[1]
      YB←(YB-Y)÷Z×AB[2]
      J←(YA[1]×YB[2])-YA[2]×YB[1]
      AB1←AB[1]-((Y[1]×YB[2])-Y[2]×YB[1])÷J
      AB1←AB1,AB[2]-((Y[2]×YA[1])-Y[1]×YA[2])÷J
      '2 OLD         TOLERANCES        2 NEW'
      AB,Y,AB1
      'R.S:';R,S
      →(2E¯7<+/|¯1+AB1÷AB)/TOP
      ⎕←Y←V FN AB1
      'R,S,ALPHA,BETA'
      R,S,AB1
      ∇


      ∇FN[⎕]∇
      ∇ Y←V FN AB;D;F;G;H;N;PS;PR;PRS;DL;LL
     ⍝R AND S ARE GLOBALS
     ⍝V A VECTOR OF OBSERVATIONS, AB IS ALPHA,BETA
     ⍝Y IS A 2 VECTOR TRYING TO GET TO 0,0 FOR TBET MLE
      N←÷⍴V
      G←V÷AB[2]
      H←⍟G
      D←1+G*-AB[1]
      F←N×+/H÷D
      H←AB[1]×N×+/H
      D←N×+/÷D
      R←-÷H-AB[1]×F÷D
      S←R×-1-÷D
      G←N×+/⍟1+G*AB[1]
      PS←SI S
      Y←H+PS-PR←SI R
      Y←Y,G+PS-PRS←SI R+S
      ∇
```

EXHIBIT H1 PAGE 2

```
  ∇SI[⎕]∇
      ∇ PSIX←SI X;Z;PSIZ;Y;M;N
[1]    ⍝PSI FUNCTION IE DERIVATIVE LOG GAMMA FUNCTION
[2]    ⍝WRITTEN HARRY SOUL
[3]     Z←X-⌊X
[4]     →(L1,L2)[1+Z=0]
[5]    L1:PSIZ←-(!-1-Z)×+/(⍳26)×CEE×Z*¯1+⍳26
[6]     Y←1000|⌊X
[7]     N←0
[8]     M←⌊X÷1000
[9]     PSIX←PSIZ++/÷Z+¯1+⍳Y
[10]    →(M=0)/0
[11]   LT:N←N+1
[12]    PSIX←PSIX++/÷Z+(1000×N-1)+Y+¯1)+Y+-1+⍳1000
[13]    →(N<M)/LT
[14]    →0
[15]   L2:PSIZ←-(!Z)×+/(⍳26)×CEE×(Z+1)*¯1+⍳26
[16]    PSIX←PSIZ++/÷⍳X-1
      ∇
```

60	CEE
61	1
62	0.5772156649015329
63	−0.6558780715202538
64	−0.0420026350340952
65	0.1665386113822915
66	−0.0421977345555443
67	−0.009621971527877
68	0.007218943246663
69	−0.0011651675918591
70	−0.0002152416741149
71	0.0001280502823882
72	−2.01348547807E−5
73	−1.2504934821E−6
74	1.133027232E−6
75	−2.056338417E−7
76	6.116095E−9
77	5.0020075E−9
78	−1.1812746E−9
79	1.043427E−10
80	7.7823E−12
81	−3.6968E−12
82	5.1E−13
83	−2.06E−14
84	−5.4E−15
85	1.4E−15
86	1E−16

EXHIBIT H2

SAMPLE RUN OF TRANSFORMED BETA MLE WITH GOOD STARTING ESTIMATES

```
      V
2.201825487277711   1.747798995989603   1.555619456471727   1.434261861491408
1.345898293955564   1.276532762732432   1.219472497925706   1.171009053335359
1.128878212884788   1.091598560297855   1.058149169964544   1.027797375655266
0.9999999999999999  0.9743434501286376  0.9505056924135983  0.9282311847924588
0.9073138148639067  0.8875849650558165  0.8689049641059015  0.8511568358547295
0.8342416436253031  0.81807496609125    0.8025841905289312  0.7877064064383325
0.7733867467937893  0.7595770676095318  0.7462348863847357  0.733322520898723
0.7208063846790024  0.7086564061646645  0.6968455463959924  0.6853493958275812
0.674145835167939   0.6632147483965766  0.6525377785832587  0.6420981190348407
0.6318803337678372  0.6218702024548765  0.6120545858977834  0.6024213087964627
0.5929590571538267  0.5836572881149439  0.5745061504078917  0.5654964138531555
0.5566194066522674  0.5478669593658831  0.5392313546553542  0.53070528199689
0.5222817966889851  0.5139542825661741  0.5057164179087162  0.4975621441012036
0.4894856366454348  0.4814812781759202  0.4735436331613866  0.4656674240036885
0.4578475082673738  0.4500788567894164  0.4423565324296311  0.434675669228307
0.4270314517386136  0.4194190942972299  0.4118338199870745  0.404270839030415
0.3967253263281841  0.3891923978309076  0.3816670853866948  0.3741443096602311
0.3666188506509329  0.3590853152547595  0.3515381012078956  0.3439713566152265
0.3363789340936847  0.328754338338611   0.3210906656344104  0.3133805334572376
0.305615997826835   0.2977884554141212  0.2898885265394574  0.2819059140149433
0.273829231162068   0.2656457900782352  0.2573413380345932  0.248899725301121
0.2403024809791267  0.2315282633825993  0.2225521361535894  0.2133445971882582
0.2038702484700424  0.1940859297219838  0.1839380254588229  0.1733584487947235
0.1622584092416313  0.1505182593963702  0.1379699089521555  0.1243638396796979
0.1093001477080087  0.09205965646857106 0.07106750819518526 0.04089307909136584

V NRFN   1.521 1.553

2 OLD                                   TOLERANCES                                   2 NEW
1.521 1.553                             1.456996026050206E-6 1.088012693967189E-7   1.520915599542439 1.553092179774157
R, S: 1.441569975759713 6.476705211863293
2 OLD                                   TOLERANCES                                   2 NEW
1.520915599542439 1.553092179774157     2.314481939436064E-11 2.418509836843441E-12 1.520915600822739 1.553092175281865
R, S: 1.441699580189243 6.477401387277938
4.440892098500626E-16 2.775557561562891E-16
R, S, ALPHA, BETA
1.441699614500499 6.477400647693872 1.520915600822739 1.553092175281865
```

APPENDIX J

TRANSFORMED GAMMA
RELATION BETWEEN COEFFICIENTS OF VARIATION AND SKEWNESS

Empirical investigations suggest that not all pairs of positive real numbers can be realized as the coefficients of variation (*CV*) and skewness (*SKW*) of a transformed gamma distribution. For example, as mentioned in the text, for *CV*s of 1.25 and greater the *SKW* always seems to exceed the *CV*.

While not proven analytically, observation suggests the following:

(1) For fixed *r* the ratio *SKW/CV* is a decreasing function of alpha.
(2) If the ratio *SKW/CV* is held constant (by increasing alpha), then the *CV* and *SKW* increase as *r* decreases.
(3) These increases are asymptotic to some finite value as *r* goes to zero.

Thus for a fixed *SKW/CV* ratio, the *CV* and *SKW* can not exceed a maximum. The following table gives these approximate maximum values for selected ratios.

SKW/CV	*Maximum CV*	*Maximum SKW*
1.4	11.1	15.5
1.3	3.9	5.1
1.2	2.0	2.4
1.1	1.51	1.66
1.0	1.25	1.25
.9	1.09	.98
.83	1.00	.83
.8	.97	.78
.7	.88	.62
.6	.81	.49
.5	.76	.38
.4	.71	.28
.3	.67	.20
.2	.64	.13
0.00	.58	0.00

This relationship thus restricts the values which the *CV*, *SKW* pairs can take on. As the maxima seem to be increasing functions of the ratio *SKW/CV*, each maximum is an upper bound over all lower values of that ratio. For example, if the *SKW* is less than or equal to .83*CV*, then the *CV* does not exceed 1.0. Conversely, if the *CV* is above 1.0 the *SKW* is .83 or greater.

It is interesting to note that the skewness can be negative. This seems possible for any value of r. For small r, SKW reaches zero at about an alpha of $1/r$. In the Weibull case ($r = 1$) zero skewness occurs for α just above 3.6.

The use of empirical studies in mathematical investigations is of course subject to pitfalls. The findings in this appendix should thus be regarded as hypotheses until more rigorous demonstrations can be provided.

Further investigation has also revealed that matching transformed gamma moments is not possible if the CV is very small and the SKW is large. In this case, it has been possible to match transformed beta moments. The case $\alpha = 1$ often suffices, and this yields closed form solutions for the parameters as follows:

Define $M_j = E(X^j)/E(X)^j$ for any random variable X. Then the transformed beta parameters r and s are:

$$r = 2\,\frac{M_3 - M_2^2}{M_2^2 + M_2M_3 - 2M_3}$$

$$s = \frac{r + 1 - 2M_2r}{r + 1 - M_2r}$$

Unfortunately, those equations sometimes yield negative parameters. In that case the transformed beta with $r = 20$ ($\alpha \neq 1$) has seemed to give satisfactory fits.

Using the transformed beta to match moments in this way would seem to give up the parameter uncertainty. This is not necessary, however, as the moments of the combined process-parameter system can be found by combining the process and parameter moments. In fact,

$$M_j(\text{combined}) = M_j(\text{process})\, M_j(\text{parameter}).$$

Thus the combined moments can be used to calculate the transformed beta or gamma parameters. This, in fact, allows for greater freedom in selecting the parameter distribution moments, in that the skewness need not be strictly determined by the CV.

BIBLIOGRAPHY

[1] M. Abramowitz, and I. A. Stegun, *Handbook of Mathematical Functions,* National Bureau of Standards Applied Mathematics Series .55 (1969).

[2] R. Q. Aquino, "Contributions to the Unified Treatment of Nonlife Insurance Problems," unpublished Ph.D. thesis, Rensselaer Polytechnic Institute (1979).

[3] S. D. Conte and C. de Boor, *Elementary Numerical Analysis: An Algorithmic Approach,* McGraw-Hill, New York (1972).

[4] C. A. Hachemeister, Discussion of "Loss Ratio Distributions—A Model," *PCAS* LIV (1967), p. 89.

[5] F. Harwayne, "Accident Limitations for Retrospective Rating," *PCAS* LXIII (1976), p. 1.

[6] N. A. J. Hastings and J. B. Peacock, *Statistical Distributions,* John Wiley and Sons, Inc., New York (1975).

[7] C. C. Hewitt, Jr., "Loss Ratio Distributions—A Model," *PCAS* LIV (1967), p. 70.

[8] F. B. Hildebrand, *Introduction to Numerical Analysis,* McGraw-Hill, New York (1974).

[9] T. Pentikainen, "On the Approximation of the Total Amount of Claims," ASTIN *Bulletin,* vol. IX, (1977), page 281.

[10] H. L. Seal, "Approximations to Risk Theory's $F(x,t)$ By Means of the Gamma Distribution," ASTIN *Bulletin,* vol. IX (1977), page 213.

[11] R. M. Adelson, "Compound Poisson Distributions," *Operations Research Quarterly,* 17, pp. 73–75 (1966).

[12] H. H. Panjer, "Recursive Evaluation of a Family of Compound Distributions," ASTIN *Bulletin,* vol. 12, no. 1 (1981), page 22.

[13] P. E. Heckman and G. G. Meyers, "The Calculation of Aggregate Loss Distributions from Claim Severity and Claim Count Distributions," *PCAS* LXX (1983).

[14] O. Lundberg, "On Random Processes and Their Applications to Sickness and Accident Statistics," University of Stockholm, thesis, Uppsala (1940). 2nd edition. Uppsala: Almquist & Wiksells (1964).