LOSS RATIO DISTRIBUTIONS A MODEL

C. C. HEWITT, JR.

1. INTRODUCTION AND SUMMARY

1.1 Historical

Traditionally in casualty insurance loss ratio distributions have been obtained empirically and often at great expense and with great labor [for the most recent such effort see (13)]. Associated with collecting masses of raw data have been serious problems of fitting such data [see (8), (12) and 14)]. The end-product of all of these efforts has been non-analytical; and of value only for use in linear retrospective rating and as a rough guide to loss ratio distributions by size.

1.2 An Analytical Model

In this author's review (11) of (13) he indicated the successful fitting of a mathematical model—the gamma distribution—to actual loss ratio distributions. Also, there was indicated a relationship among the significant parameters for loss ratio distributions at various premium sizes.

1.3 Purpose and Results of this Paper

In this paper are set forth some important mathematical properties of the gamma distribution (Chapter 2) including the very important characteristic of *reproductivity* and *divisibility.* In most instances the development of formulas and lemmas is left to the reference texts, or the reader. The gamma distribution is applied directly to loss ratio distributions as a model (Chapter 3) and a single parameter form is asserted.

In Chapter 4 the method of fitting actual data is explained and the goodness of the fit is discussed. A relationship among parameters at various premium sizes is also asserted. As a corollary it becomes evident that for actual data, loss ratio distributions for larger premium sizes are not equivalent to loss ratio distributions that might have been obtained by taking random samples from smaller premium sizes. An attempt is made to account for this phenomenon.

Finally (Chapter 5) the utility of the new model is discussed for:

(1) Linear retrospective rating.

(2) Non-linear retrospective rating.

(3) Competitive "retro" dividend plans.

1.4 *Prospectus*

Although significant results are obtained within this paper, the implications go far beyond the answers. For example, the gamma distribution as a loss ratio model for larger risk premium sizes must be the synthesis of:

(1) Distributions of a single loss,
(2) Distributions of occurrence of one or more losses, and
(3) Inherent risk heterogeneity.

It would be interesting to see this analyzed further; such analysis would undoubtedly explain why the goodness-of-fit tests fail for smaller premium sizes. Hopefully, then, this paper will not be an end but merely a beginning.

2. THE GAMMA DISTRIBUTION

2.1 *The Gamma Function*

2.11 *The (Complete) Gamma Function*

(a) Definition:

$$\Gamma(r) = \int_0^\infty x^{r-1} e^{-x} dx \,;\ [r>0] \qquad \text{............(2.111)}$$

(b) The (complete) gamma function has the recursive property:

$$\Gamma(r+1) = r\Gamma(r) \qquad \text{............(2.112)}$$

(c) If r is integral,

$$\Gamma(r+1) = r! \qquad \text{............(2.113)}$$

(d) The (complete) gamma function has a *minimum* when r is approximately 1.4616; the minimum is approximately 0.8856.

As r approaches zero, or increases without limit, the (complete) gamma function increases without limit.

(e) $\Gamma(r) \sim \sqrt{2\pi r}\, e^{-r}\, r^{r-1}$

for large r ....(2.114)

(f) For intermediate, non-integral values of r, use may be made of the recursive property (2.112) and published tables [(1), p. 316].

(g) $$\Gamma(\tfrac{1}{2}) = \sqrt{\pi} \qquad \text{............(2.115)}$$

(h) $$\Gamma(r) = \lim_{n\to\infty} \frac{n!\, n^{r-1}}{r(r+1)(r+2)\ldots(r+n-1)} \quad\ldots\ldots(2.116)$$

[(3), p. 697]

(i) $$\frac{1}{\Gamma(r)} = re^{\gamma r} \prod_{k=1}^{\infty} \left(1 + \frac{r}{k}\right) e^{-\frac{r}{k}} \quad\ldots\ldots(2.117)$$

where γ is the Euler-Mascheroni constant.

$$\gamma = \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \log_e n\right) = -\int_0^\infty e^{-x} \log_e x\,dx,$$

$\gamma \sim 0.5772157$ [(3), p. 697]

(j) The k^{th} derivative of $\Gamma(r)$ is:

$$\Gamma^{(k)}(r) = \int_0^\infty x^{r-1} (\log_e x)^k e^{-x}\,dx \quad\ldots\ldots(2.118)$$

2.12 *The Incomplete Gamma Function*

(a) Definition:

$$I(u, p) = \frac{\int_0^{u\sqrt{p+1}} x^p e^{-x} dx}{\Gamma(p+1)}; \; (p + 1 = r, \therefore p > -1) \quad\ldots\ldots(2.121)$$

where $u = \dfrac{x_0}{\sqrt{p+1}}$ ......(2.121a)

(b) In the gamma *distribution* a scale parameter, a, is introduced; in this case:

$$u = \frac{ax_0}{\sqrt{p+1}}; \; (a > o) \quad\ldots\ldots(2.121b)$$

(c) Use may be made of published tables, (4). Also see [(5) p. 223] for adaptations from other published tables.

(d) For p near -1, values of the incomplete gamma function may be approximated to a desired degree of accuracy by iteration of the following series [(4, p. xv]:

$$I(u, p) = \frac{\zeta^{(p+1)}}{\Gamma(p+2)} \left\{ 1 - \frac{\zeta(p+1)}{1!(p+2)} + \frac{\zeta^2(p+1)}{2!(p+3)} - \frac{\zeta^3(p+1)}{3!(p+4)} + \ldots \right\}$$

where $\zeta = u\sqrt{p+1}$ ......(2.122)

2.2 *The Gamma Distribution*

2.21 *Basic Properties*

(a) Definition:

$$\Gamma_{a,r}(x) = \frac{a^r}{\Gamma(r)} x^{r-1} e^{-ax}; \begin{bmatrix} x \geqq o \\ a > o \\ r > o \end{bmatrix} \qquad \text{............(2.211)}$$

(b) The gamma distribution is a special case of the more general Pearson Type III distribution:

$f(x) = A(x - \mu)^{r-1} e^{-a(x-\mu)}; \ [x > \mu]$

when $\mu = o$. ....(2.212)

(c) a is the trivial scale parameter [(2) II, p. 46]; r is the important parameter.

(d) The *mode* of $\Gamma_{a,r}(x)$ occurs at:

$$\frac{r-1}{a} \qquad ;(r >) \qquad \text{............(2.213)}$$

(e) The characteristic function is

$$\left(1 - \frac{it}{a}\right)^{-r} \qquad \text{............(2.214)}$$

(f) The exponential distribution:

$$ae^{-ax} \qquad \text{............(2.215)}$$

is a special case of $\Gamma_{a,r}(x)$
when $r = 1$

(g) The gamma distribution is the continuous analogue of the negative binomial [(2) II, p. 10].

(h) Estimators [(6), p. 39.]:
If r is given, the maximum likelihood estimator for a is:

$$\hat{a} = \frac{r}{\bar{x}} \ ; \bar{x} = \text{sample mean} \qquad \text{............(2.216)}$$

For large $\underline{n}$, the p. d. f. of $\underline{\hat{a}}$ approaches normality with mean $\underline{a}$ and variance $\frac{a^2}{nr}$. Also the p. d. f. of $\sqrt{r} \log \hat{a}$ approaches normality with mean $\sqrt{r} \log a$ and variance $\frac{1}{n}$.

2.22 *Reproductivity & Divisibility (Theoretical)*

(a) Convolutions:

If x_1 and x_2 are *independent* with p. d. f. Γ_{a,r_1} and Γ_{a,r_2} respectively, then

$$X = x_1 + x_2$$

is gamma-distributed with p. d. f. Γ_{a,r_1+r_2}

[(2) II. p. 46, (6), p. 121, and (5), p. 225]

Similarly $\sum x_i$ under the same conditions would have p. d. f. $\Gamma_{a,\Sigma r_i}$. This is often expressed:

$$\Gamma_{a,r_1} * \Gamma_{a,r_2} * \ldots * \Gamma_{a,r_n} = \Gamma_{a,\sum_1^n r_i} \qquad \ldots\ldots\ldots\ldots(2.221)$$

Consequently the sum of the values of a random sample of size n from a gamma-distributed population,

$$z = \sum_1^n x_i \text{ has a p. d. f. } \Gamma_{a,nr}.$$

(b) Divisibility:

The "inverse" of this reproductive property of the gamma distribution is infinite divisibility, i.e. $\Gamma_{a,r}$ is the distribution of the sum of n independent random variables with a common p. d. f. $\Gamma_{a,\frac{r}{n}}$

[(2) II, p. 173]

2.23 *Exponential Polynomials — Moments of* x *and* e^x

(a) Functions of x of the form:

$$Ax^n e^{-bx}\, \Gamma_{a,r}(x)\,; \begin{pmatrix} n > -r \\ b > -a \end{pmatrix}$$

are themselves gamma-distributions, $\Gamma_{a+b,r+n}$.

(b) For $\Gamma_{a,r}$

$$E(x^n e^{-bx}) = \left(\frac{a}{a+b}\right)^r \frac{\Gamma(r+n)}{(a+b)^n\,\Gamma(r)} \qquad \ldots\ldots\ldots\ldots(2.231)$$

(c) Thus the k^{th} moment of x about the origin is:

$$E(x^k) = \frac{(r+k-1)(r+k-2)\ldots.(r)}{a^k} \qquad \ldots\ldots\ldots\ldots(2.231a)$$

$$E(x) = \frac{r}{a} \qquad \ldots\ldots\ldots\ldots(2.231b)$$

$$E(x^2) = \frac{(r+1)r}{a^2} \qquad \text{............(2.231c)}$$

etc.

and $$D^2(x) = E(x^2) - \overline{E(x)}^2 = \frac{r}{a^2} \qquad \text{............(2.231d)}$$

(d) Similarly, the k^{th} moment of e^x is:

$$E(e^x)^k = \left(\frac{a}{a-k}\right)^r \quad ;(a>k) \qquad \text{............(2.231e)}$$

and $$E(e^x) = \left(\frac{a}{a-1}\right)^r \quad ;(a>1) \qquad \text{............(2.231f)}$$

etc.

(e) This latter situation is helpful if the logarithm of a variable, y, is gamma-distributed, i.e.,

if, $\underline{x = \log_e y}$ has p. d. f. $\Gamma_{a,r}(x)$

Since $\underline{y = e^x}$, the moments of y are given in (2.231e). [See (10)]

3. GAMMA-DISTRIBUTION AS A MODEL FOR LOSS RATIO DISTRIBUTIONS

3.1 *Definition of terms:*

L = actual (risk) losses ($)

ϵ = expected loss ratio

P = (risk) premium ($)

r^1 = actual (risk) loss ratio $= \dfrac{L}{P}$

$$R = \frac{\text{actual loss ratio}}{\text{expected loss ratio}} = \frac{r^1}{\epsilon}$$

3.2 *The distribution form of R:*

If R is gamma-distributed, its p. d. f. would take the form [see (2.211)]:

$$f(R) = \frac{a^r}{\Gamma(r)} R^{r-1} e^{-aR} ; \begin{pmatrix} a > o \\ r > o \end{pmatrix} \qquad \text{............(3.21)}$$

with $E(R) = \dfrac{r}{a}$, but if *total actual losses* balance with *total expected losses,*

$E(R) = 1$ by definition,

and $a = r$; therefore

$$f^*(R) = \frac{r^r}{\Gamma(r)} R^{r-1} e^{-rR} \qquad \text{............(3.21*)}$$

with r as its own scale parameter.

3.3 *The distribution form of r':*

From (3.21), (3.21*) and $r' = \epsilon R$
it follows that:

$$g(r') = \frac{\left(\frac{a}{\epsilon}\right)^r}{\Gamma(r)} r'^{r-1} e^{-\frac{a}{\epsilon} r'} \qquad \text{............(3.31)}$$

and

$$g^*(r') = \frac{\left(\frac{r}{\epsilon}\right)^r}{\Gamma(r)} r'^{r-1} e^{-\frac{r}{\epsilon} r'} \qquad \text{............(3.31*)}$$

with $\frac{r}{\epsilon}$ as the scale parameter.

3.4 *The distribution form of L:*

From (3.31), (3.31*) & L = $r'P$
it follows that:

$$h(L) = \frac{\left(\frac{a}{\epsilon P}\right)^r}{\Gamma(r)} L^{r-1} e^{-\left(\frac{a}{\epsilon P}\right)L} \qquad \text{............(3.41)}$$

and

$$h^*(L) = \frac{\left(\frac{r}{\epsilon P}\right)^r}{\Gamma(r)} L^{r-1} e^{-\left(\frac{r}{\epsilon P}\right)L} \qquad \text{............(3.41*)}$$

with $\frac{r}{\epsilon P}$ as the scale parameter.

3.5 *The distribution forms of random samples:*

If *random* samples of size nP are taken (or go to make up a sample of size P) from the risk-population, it follows from the reproductivity and infinite divisibility property of the gamma distribution (see Section 2.22) that L_n, the losses in such random samples, are distributed

$$h(L_n) = \Gamma_{\frac{a}{\epsilon P},\, nr} \qquad \text{............(3.51)}$$

and

$$h^*(L_n) = \Gamma_{\frac{r}{\epsilon P},\, nr} \qquad \text{............(3.51*)}$$

where $(n > o)$

Since $r'_n = \frac{L_n}{nP}$, it follows that

$$g(r'_n) = \Gamma_{\frac{na}{\epsilon},\, nr} \qquad \text{............(3.52)}$$

and

$$g^*(r'_n) = \Gamma_{\frac{nr}{\epsilon},\, nr} \qquad \text{............(3.52*)}$$

Also

$$R_n = \frac{r'_n}{\epsilon}, \text{ therefore}$$

$$f(R_n) = \Gamma_{na,nr} \qquad \text{............(3.53)}$$

and

$$f^*(R_n) = \Gamma_{nr,nr} \qquad \text{............(3.53*)}$$

3.6 *Fitting loss-ratio data to the Gamma-distribution:*

(a) As r increases $\Gamma_{a,r}$ approaches the form of the normal distribution.

(b) Furthermore the function:

$$r'g(r')$$

is itself of the gamma distribution form [see 2.23(a)]

$$\Gamma_{\frac{a}{\epsilon},\, r+1}$$

(c) Also for *fixed P* (premium size) $r'g(r')$ is proportional to:

$$Lh(L)$$

(d) This combination of (1) increased "normality", (2) primary interest in the distribution of amount ($'s) of loss (as opposed to number of risks), and (3) the convenience of "generated" gamma-distributions of amounts of loss from gamma-distributions by number of risks suggests the following method for determining the important parameter, r:

(1) Use $\bar{r}'$ (= sample mean loss ratio) as an estimator for r'.

(2) Use $\bar{r}''$ (= sample mean loss ratio-weighted by amount of loss in each loss ratio interval—rather than by number of risks) as an estimator for:

$$r'' = E(r'g(r')) = E(\Gamma_{\frac{a}{\epsilon},\, r+1})$$

(3) Then from (2.231b)

$$\bar{r}' = r' = \frac{\hat{r}}{a}\epsilon$$

$$\bar{r}'' = r'' = \frac{\hat{r}+1}{a}\epsilon$$

$$\frac{\bar{r}''}{\bar{r}'} = \frac{\hat{r}+1}{\hat{r}}, \text{ and}$$

$$\hat{r} = \frac{\bar{r}'}{\bar{r}'' - \bar{r}'} \qquad \text{............(3.61)}$$

and $\hat{r}$ can be used as an estimator for r.

(e) This uniquely determines the distribution—$f^*(R)$—for a particular premium size. The other distributions in the (*) family, (3.31*) and (3.41*), are known if ϵ is known. Furthermore the non-starred distributions, (3.21), (3 31) and (3.41) can then be determined by using (2.216) to determine a.

(f) From (2.231a) it follows that for $g(r')$:

$$\frac{E(x^{k+1})}{E(x^k)} = \frac{(r+k \ldots . (r)\epsilon^{k+1}}{a^{k+1}} \cdot \frac{a^k}{(r+k-1) \ldots . (r)\epsilon^k}$$

$$= \frac{r+k}{a}\epsilon$$

which, for $k = 1$, gives another way of obtaining $\bar{r}''$.

(g) In general the approach for estimating r which is described in this section is particularly appropriate for highly-skewed loss ratio distributions, since it emphasizes higher-moment-weighted distributions that are more nearly normal.

However, great care must be exercised in fitting higher-moment means to small samples because of the increased effect of infrequent and often erratic large losses and larger loss ratios upon the estimators.

4. FITTING THE GAMMA-DISTRIBUTION TO ACTUAL LOSS-RATIO DISTRIBUTIONS

4.1 *The Data*

(a) Fortunately, data of unsual homogeneity for its large amount was obtained by using workmen's compensation insurance experience for the single state of California—now our largest state, not only in population, but also in workmen's compensation premium volume. The data is contained in a special unpublished report of the California Inspection Rating Bureau dated January 31, 1963 entitled "California Experience Rating Statistics—Series II—By Interval of Subject Premium Loss Ratio."

(b) The raw data is tabulated in a series of fourteen exhibits by subject premium size:

Exhibit	Subject Premium	Number of Risks
A	Less than $500	2,946
B	$ 500 - $ 749	3,126
C	750 - 999	2,889
D	1,000 - 1,499	3,718
E	1,500 - 2,499	3,576
F	2,500 - 4,999	2,891
G	5,000 - 7,499	939
H	7,500 - 9,999	465
I	10,000 - 14,999	454
J	15,000 - 24,999	316
K	25,000 - 49,999	256
L	50,000 - 99,999	91
M	More than $99,999	55
N	Total	21,722

and is for policies effective in the first nine months of 1958. An extract of the significant portions of the data contained in Exhibit K is shown in Table 1 in the Appendix as an example.

(c) Determining the estimator for r:
Using (3.61) and the data in Table 1 for illustration of the method:

$$\bar{r}' = .578 \text{ (in Table 1, } \frac{\Sigma\,(5)}{\Sigma\,(3)})$$

$$\bar{r}'' = .931 \text{ (in Table 1, } \frac{\Sigma\,(5) \times (6)}{\Sigma\,(5)})$$

$$\hat{r} = \frac{0.578}{0.931 - 0.578} = 1.639$$ [rounded to 1.6 for use as an entry in (4)]

For other premium sizes:

Subject Premium Interval	$\hat{r}$ (rounded)
$5,000 - $7,499	0.45
7,500 - 9,999	0.65
10,000 - 14,999	0.85
15,000 - 24,999	1.3
25,000 - 49,999	1.6
50,000 - 99,999	2.9
More than $99,999	6.2

(d) Goodness of Fit:

In some instances the raw data was adjusted for "contamination," but such changes were minor. Despite the broadness of some of the premium intervals Chi-square tests were met for all of the premium sizes in (c). An example is given in Table 2 in the Appendix.

Premium sizes below \$5,000 can *not* be made to satisfy Chi-square tests even with minor smoothing.

(e) Relationship between r and Premium Size:

The results set forth in (c) suggest that there should be a relationship between premium size and the key gamma-distribution parameter, r. Although the goodness of fit for premiums below \$5,000 leaves something to be desired, estimators for r were calculated and a logarithmic curve of the form:

$$\log r = \alpha + \beta \log P \qquad \text{............(4.11)}$$

was fitted by least squares. The results are tabulated below:

Average Subject Premium Size (in interval)	$\hat{r}$ Raw	r Using (4.11)*
\$ 296	.038	.044
628	.081	.079
869	.096	.102
1,223	.132	.132
1,924	.187	.188
3,481	.326	.298
6,050	.472	.457
8,652	.627	.601
12,265	.868	.787
18,944	1.336	1.104
33,455	1.639	1.710
68,758	2.898	2.985
220,786	6.145	7.362

*(For logarithms to base 10, $\alpha = -3.264$ and $\beta = 0.773$)

The use of average premiums as representative of an entire premium interval is crude particularly in the \$100,000-and-over interval. Nevertheless a relationship indicating some predictability does exist. A hasty application of the above methods to the new Table M raw data [see (13)] supports both the use of the gamma-distribution as a model for loss ratio distributions and the use of a logarithmic curve to determine $\acute{r}$.

(f) Reproductivity and Divisibility (Actual):

In Section 3.5 it was shown that, if r is the key parameter for premium size P, then a larger $(n > 1)$ or smaller $(o < n < 1)$ *random* sample from the same risk-population would have the key parameter nr. If the *actual* loss ratio distributions discussed above followed a random sampling pattern then from (4.11)

(A) $$\log r = \alpha + \beta \log P$$

(B) $$\log nr = \alpha + \beta \log nP$$

but subtracting (A) from (B)

(C) $$\log \frac{nr}{r} = \beta \log \frac{nP}{P}$$

(D) $$\beta = 1$$

However, β was found to be *0.773* for the California data (and logarithm-base 10). Thus, it can be inferred that larger-risk loss ratio distributions can not be obtained by a randomized pyramiding of smaller-risk loss ratio distributions and vice versa. Putting it bluntly—for loss ratio distribution purposes—*two $50,000 risks don't make a $100,000 risk.* Nor is a $100,000 risk for one year the same as a $50,000 risk for two years.

This result challenges formerly-used methods of arriving at loss ratio distributions for large risks [see (9)]. Also challenged is the present method of equating insurance charges for three-year retrospective rating plans with insurance charges for a one-year plan on a risk three times as large. Similarly, there would appear to be some inaccuracy in calculating the insurance charges (contained in the basic premium) in Retrospective Rating Plan D for premium sizes 50% of, and 200% of, the estimated standard premium by equating such charges to those of risks one-half, and twice, the estimated size of the risk in question.

Since the insurance charge for larger premium sizes is a small portion of the total premium and, since the margin of error in previous and current methods of computing insurance charges would also seem to be small, it is doubtful if any great harm has been, or is being, done by the methods here impugned.

(g) Size Characteristics of Actual Loss Ratio Distributions:

Let r_n be the value of r for a sample of size nP as determined by (4.11), i.e., from actual loss ratio distributions. Then

(A) $\log r = \infty + \beta \log P$

(B) $\log r_n = \infty + \beta \log nP$

(C) $r_n = n^{\beta} r$

On the other hand, if ρ_n is the value of r for a random sample of size nP:

(D) $\rho_n = nr$

From (C) and (D) it follows that:

(E) $\rho_n = n^{(1-\beta)} r_n$

For $o \leqq \beta < 1$,

where $n > 1$, $\rho_n > r_n$ ....(4.12a)

where $o < n < 1$, $\rho_n < r_n$ ....(4.12b)

But the variance of $f^*(R)$ [see (2.231d) and (3.21*)] is:

$$\frac{1}{r}$$

Thus a loss ratio distribution of larger risk size obtained by pyramiding a loss ratio distribution of smaller risk size on a *random* basis has a smaller variance $\frac{1}{\rho_n}$ than the *actual* loss ratio distribution $\frac{1}{r_n}$.

There are a number of possible explanations for this conclusion. One such explanation, which would seem logical, would run as follows:

Risk	(1) Exposure (Units)	(2) Frequency (Units)	(3) Severity (Units)	(4) Premium (Units) (1) × (2) × (3)
A	1	1	1	1
B	1	1	2	2

When the premium for Risk A is made equal to that of Risk B by doubling the exposure units of Risk A, it seems clear that the variance of loss ratios for double-units of Risk A would be less than the variance of loss ratios for single-units of Risk B. This is so because the severity for Risk A is only one-half the severity of Risk B.

It is clear, *a fortiori,* that, all other things being equal, risks with larger severities would be in the larger premium size intervals.

5. UTILIZING THE MODEL

5.1 Linear Retrospective Rating [see (12) pp. 52-56]

Let S_o = insurance charge for loss ratios exceeding R_o

Then

$$S\hat{o} = \frac{\int_{R_o}^{\infty} (R - R_o)f^*(R)dR}{\int_o^{\infty} Rf^*(R)dR} \qquad \text{............(5.11)}$$

but

$$\int_o^{\infty} Rf^*(R)dR = E(R) = 1,$$

and

$$\int_o^{\infty} f^*(R)dR = 1$$

therefore,

$$S_o = 1 - \int_o^{R_o} Rf^*(R)dR - R_o[1 - \int_o^{R_o} f^*(R)dR] \qquad \text{............(5.11a)}$$

but

$$\int_o^{R_o} f^*(R)dR = I(u_o,p)$$

$$\int_o^{R_o} Rf^*(R)dR = I(u_1,p + 1)$$

and

$$\int_o^{R_o} R^if^*(R)dR = I(u_i,p + i)$$

where

$$r = p + 1$$

$$u^i = \frac{rR_o}{\sqrt{r + i}}$$

see section 2.12 a & b

Thus

$$S_o = 1 - I(u_1,p+1) - R_o[1 - I(u_o,p)] \qquad \text{............(5.116)}$$

Similarly if S_o' = insurance saving from loss ratios less than R_o

$$S_o' = \frac{\int_0^{R_o} (R_o - R)f^*(R)dR}{\int_0^{\infty} Rf^*(R)dR} \qquad \text{............(5.12)}$$

$$S_o' = R_o\int_0^{R_o} f^*(R)dR - \int_0^{R_o} Rf^*(R)dR \qquad \text{............(5.12a)}$$

and

$$S_o' = R_oI(u_o,p) - I(u_1,p+1) \qquad \text{............(5.12b)}$$

also

$$S_o' = S_o + R_o - 1 \qquad \text{............(5.13)}$$

The advantages of being able to compute insurance charges (or savings) by a relatively simple formula which requires only one parameter, when the parameter is a simple logarithmic function of premium size, are many and obvious. It should be sufficient to point out that parameterization of loss ratio distributions would eliminate huge tables of ratios and charges, would lend itself to computerization and would permit different and more appropriate insurance charges among various lines of insurance, geographical territories, classifications of risk, and even between one year and the next.

5.2 *Non-Linear Retrospective Rating*

Inflexibility with respect to arriving at insurance charges is not the only rigidity imposed by linear forms of retrospective rating. Linear retrospective rating implies a minimum premium and a maximum premium with the intermediate values expressed as a linear function of risk losses, i.e.,

$$P = c_1R + c_o\text{, but}$$

$$c_1 = \frac{\overline{P} - \underline{P}}{\overline{R} - \underline{R}} \text{ and } c_o = \frac{\overline{R}\underline{P} - \underline{R}\overline{P}}{\overline{R} - \underline{R}}$$

where $\underline{P}$ = minimum premium

$\overline{P}$ = maximum premium

With $\underline{R}$ and $\overline{R}$ corresponding to the respective P's

These requirements limit the insured and insurer in their choice of values for linear plans.

Why not

$$P = c_1 R^{n_1} e^{-b_1 R} + c_2 R^{n_2} e^{-b_2 R} + \ldots \quad ?$$

As long as $n_i > -r$ and $b_i > -r$

[see Section 2.23 (a)], insurance charges are calculable with the knowledge of r.

Of course there are common sense restrictions on

$$P = F(R)$$

such as $F'(R) \geq o \quad ; \quad (o < \underline{R} \leq R \leq \overline{R})$

[see (7)].

5.3 *Competitive "Retro" Dividend Plans*

Finally, it now becomes possible to design a retrospective dividend scale to be most competitive for the most desirable risks. This is not the same as saying most competitive for risks with a zero loss-ratio, since for larger premium sizes there are very few risks with near-zero loss ratios. Rather, if a competitor's dividend formula produces a net premium:

$$P = c_1 R + c_o \quad ; \quad (\underline{P} \leq P \leq \overline{P})$$

then choose a

$$P' = F(R) \qquad \text{such that}$$

$$\psi = \int_{\underline{R}}^{\overline{R}} [c_1 R + c_o - F(R)] f^* R dR + (\underline{P} - \overline{P'}) I(\underline{u}, p) + (\overline{P} - \overline{P'})[1 - I(\bar{u}, p)]$$

is a maximum. Of course, the correct insurance charges must be made. (The expression for ψ is a deliberate over-simplification since $\underline{R}$, $\underline{R}'$, $\overline{R}$, and $\overline{R}'$ will almost certainly not be equal. However, the principle is the same).

BIBLIOGRAPHY

BOOKS

(1) ———— *Standard Mathematical Tables,* Chemical Rubber Publishing Co., (12th Edition)

(2) Feller, W. *An Introduction to Probability Theory and Its Applications*
[Wiley, Vol. I — 1957 (2nd edition)
Vol. II — 1966]

(3) Korn & Korn *Mathematical Handbook for Scientists and Engineers* (McGraw-Hill, 1961)

(4) Karl Pearson, et.al. *Tables of the Incomplete Gamma Function* (Cambridge University Press, 1957)

(5) Raiffa & Schlaiffer *Applied Statistical Decision Theory* (Harvard University, 1961)

(6) Wilks, S. *Mathematical Statistics* (Wiley, 1962)

PAPERS

(7) Charles W. Crouse "On Non-Linear Retrospective Rating," *PCAS* Vol. XXXVI, pp. 35-62

(8) Paul Dorweiler "On Graduating Excess Pure Premium Ratios," *PCAS* Vol. XXVIII, p. 132 et seq.

(9) Paul Dorweiler "On the Use of Synthetic Risks in Determining Pure Premium Excess Ratios for Large Compensation and Liability Risks," *PCAS* Vol. XXIII, pp. 77-90

(10) C. C. Hewitt, Jr. "Distribution by Size of Risk—A Model," *PCAS* LIII, pp. 106-115

(11) C. C. Hewitt, Jr. "The 1965 Table M"—(Discussion), *PCAS,* Vol. LII, pp. 49-51

(12) Lewis H. Roberts "Graduation of Excess Ratio Distributions by the Method of Moments," *PCAS* Vol. XLIV, pp. 45-62

(13) LeRoy J. Simon "The 1965 Table M," *PCAS* Vol. LII, pp. 1-45

(14) N. M. Valerius "Risk Distributions Underlying Insurance Charges in the Retrospective Rating Plan," *PCAS* Vol. XXIX, p. 96 et seq.

CALIFORNIA WORKMEN'S COMPENSATION
EXPERIENCE RATING STATISTICS
Series II - By Interval of Subject Premium Loss Ratio
Exhibit K - Subject Premium 25,000 - 49,999

Policies Effective 1/1/58 - 9/30/58

(1) Subject Premium Loss Ratio Interval	(2) # of Risks	(3) Subject Premium	(4) Avg. Subj. Prem. Size (3) ÷ (2)	(5) Incurred Losses	(6) Loss Ratio (Subj. Prem.) (5) ÷ (3)
.000	2	67,253	33,627	0	.000
.001 - .199	44	1,430,618	32,514	187,072	.131
.200 - .299	33	1,069,378	32,405	262,067	.245
.300 - .399	26	884,548	34,021	313,728	.355
.400 - .499	29	909,445	31,360	406,078	.447
.500 - .599	24	811,911	33,830	443,257	.546
.600 - .699	17	577,790	33,988	373,189	.646
.700 - .799	22	808,761	36,762	601,346	.744
.800 - .899	18	623,852	34,658	530,364	.850
.900 - .999	10	353,662	35,366	338,464	.957
1.000 - 1.249	16	529,073	33,067	586,367	1.108
1.250 - 1.499	8	280,824	35,103	392,025	1.396
1.500 - 1.749	2	65,603	32,802	106,069	1.617
1.750 - 1.999	1	27,352	27,352	48,252	1.764
2.000 - 2.999	2	70,361	35,181	183,178	2.603
3.000 - 3.999	2	54,063	27,032	178,850	3.308
Total	256	8,564,494	33,455	4,950,306	.578

APPENDIX
TABLE 2

CALIFORNIA WORKMEN'S COMPENSATION
EXPERIENCE RATING STATISTICS
Series II - By Interval of Subject Premium Loss Ratio

Exhibit K - Subject Premium 25,000 - 49,999

Policies Effective 1/1/58 - 9/30/58

(1) Subject Premium Loss Ratio Interval	Number of Risks (2) Actual	(3) Theoretical	(4) χ^2 $\frac{[(2) - (3)]^2}{(3)}$
.000 - .199	46	50	0.32
.200 - .299	33	32	.03
.300 - .399	26	29	.31
.400 - .499	29	26	.35
.500 - .599	24	22	.18
.600 - .799	39	34	.74
.800 - .999	28	23	1.09
1.000 & up	31	40	2.03
Total	256	256	5.05

For 7 degrees of freedom:

Level	χ^2
pr= 0.95	2.17
pr= 0.05	14.07