ON NON-LINEAR RETROSPECTIVE RATING

BY

CHARLES W. CROUSE

Introduction. A retrospective rating agreement is an agreement providing that the premium u ultimately to be considered as earned for a certain insurance or set of insurances, shall be determined as a function of the losses actually incurred by the insurer on account thereof, valued (in accordance with a set of prescribed rules) *after* the events causing such losses have occurred.

All of the plans for the making of such agreements which have ever been explicitly approved by any of the insurance commissioners in the United States, are plans under which u becomes a continuous function of the sum of the losses so valued, which is linear within some interval and constant outside that interval. The range of practicable possibilities under the condition that u shall be such a linear or sectionally linear function of the sum of the actual values of the losses, has been very well explored by T. O. Carlson **[4]** and F. S. Perryman **[10]**.*

Now, the following question naturally arises. *What practicable retrospective rating possibilities (if any) lie beyond the domain governed by that condition?* In seeking an answer to this question, it is necessary first to determine precisely what we mean by the term "practicable retrospective rating possibility." Section B of this paper is therefore devoted to a formulation of the conditions which it is either necessary or desirable that any particular functional relationship shall satisfy in order that it may be employed in practice as a formula for the determination of u. In Section C, continuous functions linear in each interval of a set consisting of any finite number of contiguous intervals covering the entire range of possible values of the sum of incurred losses, are examined as possible retrospective rating formulae and found to leave unsatisfied at least one of the conditions which it is desirable that such formulae should fulfill. In the first paragraph of Section D, a class of non-linear functions is defined, and it is asserted that the members of that class are fundamentally better adapted to serve as retrospective rating formulae than those belonging to the sectionally linear class; and the remainder of the paper is devoted to the defense of that thesis.

* Bold face numerals in square brackets refer to the Bibliography on page 62.

A. Terminology and Notation

Let x be a one-dimentional random variable, i.e., a random variable of which each possible value is a single real number. If, for every real number x, the probability that x will not exceed x (i.e., the probability of the event: $\mathrm{x} \leqq x$) is given as the value of a real function $F(x)$, then $F(x)$ will be called the *distribution function* of x and may be regarded as defining the *probability distribution* of x. Such a function is necessarily single-valued and non-decreasing. Moreover, $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to +\infty} F(x) = 1$.

Hence the Stieltjes integral, $\displaystyle\int_{-\infty}^{+\infty} dF(x) = 1.$ (1)

Let $\mathrm{z} = \psi(\mathrm{x})$ represent any function of x. Then the *mathematical expectation* or *mean value* of z, which will be denoted by the symbol $\mathbf{E}(\mathrm{z})$, is defined by the equation:

$$\mathbf{E}(\mathrm{z}) = \int_{-\infty}^{+\infty} \psi(x)\,dF(x) \tag{2}$$

and will be said to exist if the Lebesgue-Stieltjes integral on the right exists and has a finite value. If $\mathrm{z} = (\mathrm{x} - c)^k$ where k is a positive integer, then $\mathbf{E}(\mathrm{z})$ is the kth *moment of the distribution* of x *about the point c*. Obviously, for $k = 1$ and $c = 0$, $\mathbf{E}(\mathrm{z})$ becomes $\mathbf{E}(\mathrm{x})$, the mean value of x. The moments about $\mathbf{E}(\mathrm{x})$, often called the *central moments*, will be denoted by the customary symbol, $\mu_k(\mathrm{x})$.

Thus: $$\mu_k(\mathrm{x}) = \mathbf{E}\left[\{\mathrm{x} - \mathbf{E}(\mathrm{x})\}^k\right] = \int_{-\infty}^{+\infty} \{x - \mathbf{E}(\mathrm{x})\}^k dF(x) \tag{3}$$

If $F(x)$ is continuous in the interval $(-\infty, +\infty)$ and has a continuous differential coefficient, $F'(x) = f(x)$, at every point of that interval, with the admissible exception of points of which any finite interval contains at most a finite number, then the probability distribution defined by $F(x)$ will be said to be of the *continuous type*, and $f(x)$ will be called the *frequency function* of x.

For distributions of the continuous type, $F(x) = \displaystyle\int_{-\infty}^{x} f(x)dx$ for every real value of x, and the Lebesgue-Stieltjes integral in equation (2) is equivalent to an ordinary Riemann integral, $$\int_{-\infty}^{+\infty} \psi(x)f(x)dx = \mathbf{E}(\mathbf{z}), \qquad (2')$$

provided $\psi(\mathbf{x})$ is almost everywhere * continuous and the integral on the left side of (2′) is absolutely convergent.

Each random variable with which we shall have occasion to deal, will be represented by a lower-case letter in bold-face Roman type, and the same letter in italic type of ordinary face will be used as the variable in its distribution function and its frequency function. Regardless of what random variable it is to which we are referring, the letter F will be used to denote its distribution function and the letter f to denote its frequency function. Thus $F(x)$ will mean the distribution function of $\mathbf{x}$; $F(y)$ will mean the distribution function of $\mathbf{y}$; and these may be two quite different functions, i.e., $x = c$ and $y = c$ will *not* imply that $F(x) = F(y)$. This departure from the usual convention under which $F(y)$ means the value of $F(x)$ when x has the value y, should not be confusing, and it will avoid the introduction of a number of different letters, or of subscripts or other distinguishing marks upon the basic letters F and f.†

Now let J be any set of one or several insurances, and let $\mathbf{n}$ be the number of events (hereinafter called *loss-events*) each of which results in some loss of positive value being incurred by the insurer under one or several of the contracts belonging to J. From a purely theoretical viewpoint, it seems unnecessary to make any assumptions about J. It may include Workmen's Compensation insurances, liability insurances of one or several kinds, group life or group disability insurances, and even property insurances of certain forms. However, there are a great many sets of insurances to which it would be impracticable to apply a retrospective rating agreement; we assume that J is not one of them.

Let the $\mathbf{n}$ loss-events be simply ordered; let $\mathbf{a}_k$ be the actual value of the loss to the insurer resulting from the kth loss-event; and let

$$\mathbf{s} = \mathbf{a}_1 + \mathbf{a}_2 + \ldots\ldots + \mathbf{a}_\mathbf{n}. \qquad (4)$$

Now, suppose that $\mathbf{u}$, the premium ultimately to be considered earned for all of the insurances included in J, is to be determined in accordance with a retrospective rating agreement as a function of $\mathbf{v}_1, \mathbf{v}_2, \ , \ , \mathbf{v}_\mathbf{n}$, where for each positive integer $k \leqq \mathbf{n}$, $\mathbf{v}_k$ is the value placed upon the loss resulting from the

* I.e., with the admissible exception of points forming a set of Lebesgue measure zero.

† A very beautiful exposition of random variables and probability distributions will be found in Cramér [5, Second Part, or 6].

kth loss-event in accordance with some prescribed set of rules. The possibility that $\mathbf{v}_k = \mathbf{a}_k$ for some or all values of k, is — of course — not excluded.

Under any of the approved plans for the making of retrospective rating agreements, the functional relationship which $\mathbf{u}$ would bear to $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, would be of the following form:

$$\mathbf{u} = \begin{cases} H & \text{if } \mathbf{t} \leqq h \\ B + C\mathbf{t} & \text{if } h \leqq \mathbf{t} \leqq g \\ G & \text{if } \mathbf{t} \geqq g \end{cases} \tag{5}$$

where $\mathbf{t} = \mathbf{v}_1 + \mathbf{v}_2 + \ldots\ldots + \mathbf{v}_n$; $H = B + Ch$; $G = B + Cg$; and B, C, h and g are parameters independent of $\mathbf{n}$, $\mathbf{a}_k$, $\mathbf{v}_k$ and every other characteristic or consequence of any particular loss-event. H is usually called the *Minimum Retrospective Premium*; G, the *Maximum Retrospective Premium*; B, the *Basic Premium*; and C, the *Loss Conversion Factor*; though in most of the plans at present in use these two latter terms refer respectively to B/τ and C/τ, where τ is a factor dependent solely upon the ratio to $\mathbf{u}$ of the sum of all taxes which will be levied directly upon $\mathbf{u}$ or upon some part of $\mathbf{u}$.

The range of possible values of H, G, h, g, B and C, and the determination of any three of them for any set of insurances when the remaining three are given, have been thoroughly discussed by Carlson and Perryman in the papers, [4] and [10], to which we have previously referred. Though, for each particular set of insurances, the class of sectionally linear functions defined by (5) includes a three-fold infinitude of members each of which represents a practicable retrospective rating possbiility, it is nevertheless a very restricted class of functions.

Therefore, let us suppose that $\mathbf{u}$ is to be some single-valued function of $\mathbf{t}$, say $\mathbf{u} = R(\mathbf{t})$, where $R(\mathbf{t})$ is not necessarily of the form (5), but

$$\mathbf{t} = \mathbf{v}_1 + \mathbf{v}_2 + \ldots\ldots + \mathbf{v}_n \tag{6}$$

and every parameter or variable other than $\mathbf{t}$ which may be involved in the determination of $\mathbf{u}$, is—as in (5)—independent of $\mathbf{n}$, $\mathbf{a}_k$, $\mathbf{v}_k$, and every other characteristic or consequence of any particular loss-event. We have then to inquire: *What conditions must be imposed upon the function $R(\mathbf{t})$ in order that $\mathbf{u} = R(\mathbf{t})$ shall be a practicable retrospective rating formula?*

These conditions are in general dependent upon the set of rules which have been prescribed for the determination of the values, $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. At the outset of our inquiry we shall assume those rules to be of the following form.

$$\text{For every } k\text{: } \mathbf{v}_k = \mathbf{a}_k \text{ if } \mathbf{a}_k < \lambda_k; \qquad \mathbf{v}_k = \lambda_k \text{ if } \mathbf{a}_k \geqq \lambda_k; \tag{7}$$

where λ_k is a constant depending only upon characteristics of the kth loss-event other than the actual value $\mathbf{a}_k$ of the loss resulting therefrom, e.g., upon the number of persons injured in consequence thereof, upon the number or the nature of the insurances involved, or simply upon the index k. Every set of rules for determining $\mathbf{v}_1, \mathbf{v}_2, , , , \mathbf{v}_n$ appearing in any of the retrospective rating plans which have ever been approved in any of the States, can be shown to be of the form (7) under appropriate definitions of the terms *loss* and *loss-event*. In those plans which provide that actual values of losses without limitation shall be used in the computation of $\mathbf{u}$, $\lambda_k = \infty$ and $\mathbf{v}_k = \mathbf{a}_k$ for every k, so that $\mathbf{t} = \mathbf{s}$. In subsequent sections of this paper, we shall refer to $\mathbf{v}_1, \mathbf{v}_2, , , , \mathbf{v}_n$ as the *modified losses*.

In the course of our inquiry, we shall need to consider the following random variable in addition to those already introduced:

$\mathbf{w}$ = the sum of all expenses (other than losses) * incurred by the insurer and directly allocable to the insurances included in J, plus an appropriate portion of the insurer's expenses (such as home office rent) not directly allocable to any particular set of insurances, but rather to the totality of the insurer's business.

We assume throughout that the probability distributions of $\mathbf{s}$ and $\mathbf{w}$ are such that both $\mathbf{E}(\mathbf{s})$ and $\mathbf{E}(\mathbf{w})$ exist, for otherwise the risk covered by J could hardly be regarded as insurable. From the existence of $\mathbf{E}(\mathbf{s})$, it follows—in view of (4), (6) and (7)—that $\mathbf{E}(\mathbf{t})$ exists and $\mathbf{E}(\mathbf{t}) \leqq \mathbf{E}(\mathbf{s})$. (8)

From the existence of $\mathbf{E}(\mathbf{w})$, it follows † that there is a single-valued real function $W(t)$ such that, if t' and t'' be any two real numbers for which $F(t'') - F(t') > 0$, then $\dfrac{1}{F(t'') - F(t')} \displaystyle\int_{t'}^{t''} W(t)\, dF(t)$ is the mathematical expectation of $\mathbf{w}$ under the condition $t' < \mathbf{t} \leqq t''$. Hence $\mathbf{E}\{W(\mathbf{t})\} = \mathbf{E}(\mathbf{w})$. (9)

In fact, there may be an infinite class of such functions. But if $W_1(t)$ and $W_2(t)$ be any two members of that class, then $W_1(t) = W_2(t)$ for all values of t except those belonging to a set Z such that the probability that $\mathbf{t}$ will fall in Z, is zero. Although the probability that no loss whatever will be incurred

* The term *expenses (other than losses)* includes taxes other than net income taxes, and also includes all expenses for investigation and adjustment of claims except such as may be included by definition within the denotation of the term *losses*.

† Kolmogoroff [9, Fünftes Kapitel, § 4].

under J may be very small, we assume that it is not zero. Hence, zero does not belong to any such set Z. Therefore, $W(0)$ is an uniquely determined number, which is quite properly to be regarded as the mathematical expectation of $\mathbf{w}$ under the condition $\mathbf{t} = 0$. Since $\mathbf{w}$ will certainly exceed zero even though no loss whatever be incurred under J, $W(0) > 0$. Since the expenses* allocable to any particular insurance never decrease in consequence of any increase in the losses incurred thereunder, we assume that every function having the character ascribed to $W(t)$ is monotone non-decreasing throughout the domain of non-negative real numbers with the possible exception of those belonging to a set Z of the aforesaid character.

By their definitions, none of the random variables: $\mathbf{n}$, $\mathbf{a}_1$, $\mathbf{a}_2$, , , , $\mathbf{v}_1$, $\mathbf{v}_2$, , , , $\mathbf{s}$, $\mathbf{t}$, $\mathbf{w}$, can have a negative value. Therefore, their distribution functions, $F(n)$, $F(a_1)$, etc., are all identically zero for negative values of n, a_1, etc., so that every integral of the form (1), (2), (2′) or (3) may be written with 0 as its lower limit, in place of $-\infty$, without changing its value, so long as $\mathbf{x}$ is one of the random variables listed above. Moreover, since $\mathbf{t}$ cannot be negative, $R(\mathbf{t})$ need not be defined for $\mathbf{t} < 0$. All that is hereinafter said about $R(\mathbf{t})$ is to be understood as referring to $R(\mathbf{t})$ for $\mathbf{t} \geqq 0$.

It may be argued that, since $\mathbf{t}$ will be expressed as an integral multiple of the smallest current fraction of some monetary unit, $R(\mathbf{t})$ need be defined only for values of $\mathbf{t}$ which are such multiples of that fraction, e.g., integral multiples of .01 if $\mathbf{t}$ is to be expressed in dollars. However, we shall regard the domain of possible values of losses, expenses and premiums as a continuum identical with the set of all non-negative real numbers. Accordingly, $R(\mathbf{t})$ must be defined for every non-negative real value of $\mathbf{t}$.

B. Conditions of Practicability

Let t' and t'', $t' < t''$, be any pair of possible values of $\mathbf{t}$. From (4), (6) and (7), it is clear that an increase in $\mathbf{t}$ from t' to t'' can not occur except in consequence of an increase in $\mathbf{n}$ or in $\mathbf{a}_k$ for some one or several values of k. Therefore, $R(t'')$ must not be less than $R(t')$ for, if it were, the insurer in effect would be offering the insured a reward in the form of a reduction in ultimate premium for an increase either in the number of loss-events or in the actual value of the loss to the insurer resulting from some one or several of them. Thus, one of the conditions imposed upon $R(\mathbf{t})$ by the requirement of practicability, is that

$$R(\mathbf{t}) \textit{ be monotone non-decreasing.} \qquad \text{(I)}$$

But $R(t'')$ must not so greatly exceed $R(t')$ that an increase in $\mathbf{t}$ from t' to t'' would be likely to result in an increase in the insurer's underwriting profit, $\mathbf{u} - (\mathbf{s} + \mathbf{w})$, for then the retrospective rating agreement would operate to encourage carelessness in the adjustment of claims, neglect of salvage possibilities and laxity in loss prevention, on the part of the insurer. Nor should $R(t'')$ so greatly exceed $R(t')$ that the insured might save money by withhold-

* Except commissions under certain contingent commission agreements, which we leave out of account since such agreements are not properly applicable to retrospectively rated insurances.

ing reports of loss-events from the insurer and bearing the ensuing losses and expenses himself, for—at least in the case of liability insurances—such withholding of reports may prove to be very costly to both the insured and the insurer. * Apparently, then, the following condition must be satisfied:

$$R(t'') - R(t') \leqq \delta(t', t'') + \alpha(t', t'') \quad \textit{for every pair, } t', t'', \textit{ such that } t' < t'', \tag{II}$$

where $\delta(t'\ t'')$ is the *minimum* increment in s which would effect an increase in t from t' to t'', and $\alpha(t', t'')$ is some *non-munificent* allowance for an increase in w concomitant with such an increase in t. From (4), (6) and (7), it follows that $\delta(t', t'') = t'' - t'$. Precisely what the allowance $\alpha(t', t'')$ should be, we need not decide. It is sufficient to observe that at most $\alpha(t', t'')$ should certainly not exceed the mathematical expectation of the increment in w concomitant with an increase in t from t' to t''; and that expectation is given as the difference $W(t'') - W(t')$ where $W(t)$ is any member of the class of functions to which the context of (9) refers, provided that neither t' nor t'' belongs to a certain set Z, dependent upon the function $W(t)$, whereof the probability that t will fall in Z is zero. It follows that

$$R(t'') - R(t') \leqq t'' - t' + W(t'') - W(t') \tag{10}$$

must be satisfied for every pair, t', t'', such that $t' < t''$, with the possible exception of any pair of which at least one member belongs to Z.

It may be argued that, if (II) were violated only for values t', t'', such that $t' < t'' < \zeta$, where ζ is some number such that the event: $\mathbf{t} \geq \zeta$, is practically certain to occur, then the probability of any ultimate benefit to the insurer arising out of its own carelessness, or of any ultimate saving to the insured as a result of withholding reports, would be so small as to be of no practical importance. Hence, with respect to increments in $R(\mathbf{t})$, practicability demands only that (II) be satisfied for every pair, t', t'', such that $t' < t''$ and $t'' \geq \zeta'$, where ζ' is the least upper bound of the numbers ζ. Likewise, it may be argued that, in respect to monotonicity of $R(t)$, practicability demands only that $R(t') \leq R(t'')$ be satisfied for every pair, t' t'', such that $t' < t''$ and $t'' \geq \zeta'$.

But it is well known that there are sets of insurances under which the probability that no loss whatsoever will be incurred is by no means inconsiderable, yet to which it would not be impracticable to apply a retrospective rating agreement. J may be such a set, for in the interest of generality we have refrained from making any assumption about the probability distribution of $\mathbf{s}$, except that $\mathbf{E}(\mathbf{s})$ exists. If J is such a set, then zero is the only non-negative number ζ of which it may be said that the event: $\mathbf{t} \geq \zeta$, is practically certain to occur. In that case, $\zeta' = 0$, and the conditions (I) and (II) as originally stated must be satisfied if $\mathbf{u} = R(\mathbf{t})$ is to be a practicable retrospective rating formula.

By definition, the underwriting profit to the insurer on the insurances included in J, is $\mathbf{u} - \mathbf{s} - \mathbf{w}$, which is a random variable.† Clearly, $\mathbf{E}(\mathbf{u} - \mathbf{s} - \mathbf{w})$ must exist, that is to say, the Lebesgue-Stieltjes integral

$\int_0^\infty (u - s - w)\, dF(u - s - w)$ must exist and have a finite value, for otherwise

* Cf. Perryman [10, p. 7].

† Cramér [5, p. 154, 155, 162-164].

the Law of Large Numbers would not hold—in fact, it could not even be meaningfully stated—with respect to the net underwriting profit on any portfolio of insurances of which those included in J might form a part.* Since $\mathbf{u} = \mathbf{s} + \mathbf{w} + (\mathbf{u} - \mathbf{s} - \mathbf{w})$ and since, by assumption, both $\mathsf{E}(\mathbf{s})$ and $\mathsf{E}(\mathbf{w})$ exist, it follows † that

> $\mathsf{E}(\mathbf{u})$ *must exist; and the number p defined by the equation*
>
> $$p = \mathsf{E}(\mathbf{u}) - \mathsf{E}(\mathbf{s}) - \mathsf{E}(\mathbf{w}),$$
>
> *must satisfy such criteria of reasonableness as may be applicable in practice to the mathematical expectation of underwriting profit under the retrospective rating agreement in accordance with which* $\mathbf{u}$ *is to be determined.* **(III)**

As a necessary consequence of (10) and (III), we have the following proposition: $R(0)$ *must not be less than* $\mathsf{E}(\mathbf{s} - \mathbf{t}) + W(0) + p$. (11)

Proof: Substituting t for t'' and 0 for t' in (10), we have

$$R(t) - R(0) \leqq t + W(t) - W(0),$$

which must hold for all values of $t \geqq 0$, with the possible exception of a set Z whereof $\int_Z dF(t) = 0$. Therefore, since $F(t)$ is non-decreasing,

$$\int_0^\infty R(t)dF(t) - R(0)\int_0^\infty dF(t) \leqq \int_0^\infty t\,dF(t) + \int_0^\infty W(t)dF(t) - W(0)\int_0^\infty dF(t).$$

But, since $R(\mathbf{t}) = \mathbf{u}$, $\int_0^\infty R(t)dF(t) = \mathsf{E}(\mathbf{u}) = \mathsf{E}(\mathbf{s}) + \mathsf{E}(\mathbf{w}) + p$, by (2) and (III).

By (1), $\int_0^\infty dF(t) = 1$; by (2) and (8), $\int_0^\infty t\,dF(t) = \mathsf{E}(\mathbf{t})$;

* Uspensky [12, Chp. X, esp. Par. 7, p. 191].

† Cramér [5, p. 172-173]. The existence of $\mathsf{E}(\mathbf{u})$ can also be shown to be a necessary consequence of (8), (9), (10), (I) and the fact that $R(t')$ and $W(t')$ must be uniquely defined for $t' = 0$.

$$\text{and by (2) and (9),} \quad \int_0^\infty W(t)dF(t) = \mathbf{E}(\mathbf{w}).$$

Hence: $\mathbf{E}(\mathbf{s}) + \mathbf{E}(\mathbf{w}) + p - R(0) \leqq \mathbf{E}(\mathbf{t}) + \mathbf{E}(\mathbf{w}) - W(0)$, from which (11) follows.

Since, by (8), $\mathbf{E}(\mathbf{s} - \mathbf{t}) \geqq 0$, and since $W(0) + p$ is certainly positive, (11) implies that $R(0)$ must be positive. Therefore, in view of (I), $R(\mathbf{t})$ *must be positive for every possible value of* $\mathbf{t}$.

It may appear that this conclusion is an obvious condition of the practicability of $\mathbf{u} = R(\mathbf{t})$ as a retrospective rating formula, independent of any of the conditions (I), (II) and (III). However, this is not the case. If it were practicable to violate (I) or (II), then a practicable formula could be devised under which $\mathbf{u}$ would be negative for some values of $\mathbf{t}$. Moreover, if there is a number $\zeta > 0$ such that $\mathbf{t}$ is practically certain to be not less than ζ, then under certain conditions relative to the probability distributions of $\mathbf{s}$ and $\mathbf{t}$, it would be possible to devise a practicable formula under which $\mathbf{u}$ would be negative for some values of $\mathbf{t}$, without violating either (I) or (II) for any pair, t', t'', such that $t' < t''$ and $t'' \geq \zeta$.

In addition to the conditions (I), (II), (III) and their logical consequences, which the function $R(\mathbf{t})$ *must* satisfy in order that it shall be a practicable retrospective rating formula, there is at least one condition not implied by (I), (II) and (III) which it is desirable (though not necessary) that $R(\mathbf{t})$ satisfy, to wit: *that* $R(\mathbf{t})$ *be bounded.*

If $R(\mathbf{t})$ is bounded, then it has a least upper bound G and, since $R(\mathbf{t})$ must be monotone non-decreasing, either (i) there exists a number g such that $R(\mathbf{t}) = G$ for every $\mathbf{t} \geqq g$, and $R(\mathbf{t}) < G$ for every $\mathbf{t} < g$, or (ii) no such number g exists, but $\lim R(\mathbf{t}) = G$ as $\mathbf{t} \to \infty$. In accordance with custom, we shall call G the *Maximum Retrospective Premium*; and in case (i) we shall say that G is attained at the point g, while in case (ii) G is never attained.

If $\mathbf{E}(\mathbf{n})$, i.e., the expected number of loss-events, is large—say—more than 500, and if $\mu_2(\mathbf{n})$, i.e., the variance of $\mathbf{n}$, is not much greater than $\mathbf{E}(\mathbf{n})$, the insured in negotiating the retrospective rating agreement may be much less interested in a Maximum Retrospective Premium than he is in the limits $\lambda_1, \lambda_2, \ldots$, placed upon $\mathbf{v}_1, \mathbf{v}_2, \ldots$, by rules of the form (7) in accordance with which the value of $\mathbf{t}$ shall be determined. In fact, if for every k, $\lambda_k = \Lambda$ then—even if $R(\mathbf{t})$ were unbounded—the insured would be exposed to no greater economic risk under the retrospective rating agreement, than that which he would have retained under a fixed-premium excess insurance against loss greater than Λ resulting from any one loss-event, provided—of course—that $R(\mathbf{t})$ satisfies conditions (I), (II) and (III). By proper selection of the value of Λ, the latter risk could in any case be so limited as to be quite bearable by the insured. These facts constitute a sufficient reason why $R(\mathbf{t})$ need not be bounded in order to be practicable as a retrospective rating formula.

Nevertheless, however great $\mathbf{E}(\mathrm{n})$ may be and however much the insured may be interested in prescribing the limits λ_1, λ_2, , , , he will certainly prefer an agreement in which a definite limit G is placed upon the possible values of $\mathbf{u}$, to one under which there is no such limit. Moreover, it should be observed that if $R(\mathbf{t})$ is bounded, then the existence of $\mathbf{E}(\mathbf{u})$ and of all of the moments of the distribution of $\mathbf{u}$ about any point whatever, is assured, and every analytical difficulty which might arise if some moment of that distribution were not known to exist, is thus avoided.

Finally, there are two conditions which cannot very well be expressed in categorical terms, but which it is desirable that $R(\mathbf{t})$ satisfy in some degree depending upon the circumstances of its proposed application. The first of these conditions is that, *given any particular value of* $\mathbf{t}$, *the corresponding value of* $\mathbf{u}$ *shall be computable to a certain number of significant figures with relatively small expenditure of time and effort.* For example, a function $R(\mathbf{t})$ such that the computation of its value to six significant figures for a particular value of $\mathbf{t}$, would require ten hours of labor by a skilled computer working with all available tables and computing devices, would certainly not be a very suitable formula for the rating of risks on each of which $\mathbf{E}(\mathbf{u})$ is \$5,000. Yet it might be quite satisfactory as a formula for the rating of a risk on which $\mathbf{E}(\mathbf{u})$ is \$500,000.

The second such condition is that, *in the case of any particular set of insurances, it shall be possible with relatively small expenditure of time and effort to determine the values of whatever parameters appear in the formula* $\mathbf{u} = R(\mathbf{t})$ *so that there is only a small probability that* $|\mathbf{E}(\mathbf{u}) - \bar{u}|$ *exceeds* η, where $\bar{u}$ is some predetermined value which it is desired that $\mathbf{E}(\mathbf{u})$ shall take, and η is some number which is relatively small in comparison with $\bar{u}$, e.g., $.05\bar{u}$. That $R(\mathbf{t})$ should satisfy this condition arises out of the fact that, in practice, estimates of $\mathbf{E}(\mathbf{s})$, $\mathbf{E}(\mathbf{w})$ and p [which we will represent by the symbols $\overline{\mathbf{E}(\mathbf{s})}$, $\overline{\mathbf{E}(\mathbf{w})}$ and $\bar{p}$] are given, * and it is required to determine the parameters in $R(\mathbf{t})$ so that condition (III) shall be satisfied. Thus, the sum of those three estimates is a predetermined value $\bar{u}$ which it is desired that $\mathbf{E}(\mathbf{u})$ shall take, for if $\mathbf{E}(\mathbf{u})$ does indeed take that value, then that condition will be satisfied provided the algebraic sum of $\bar{p}$ and the errors of the estimates $\overline{\mathbf{E}(\mathbf{s})}$ and $\overline{\mathbf{E}(\mathbf{w})}$ is a number which satisfies the criteria of reasonableness referred to therein. But in the course of any determination of the aforesaid parameters, $\mathbf{E}(\mathbf{u})$ can only be expressed in terms of characteristics of the probability distribution of $\mathbf{t}$. In practice, none of these characteristics is ever exactly known. Estimates of their values must be made from data relative to past experience on a properly selected sample of insurances, and such estimates are—of course—subject to error. Hence it is never possible to say (after the values of those parameters have

* These estimates are usually called the *Expected Losses*, the *Expense Allowance* and the *Profit or Contingency Loading*, respectively.

been determined) precisely what value $\mathbf{E}(u)$ *does indeed* take. At best, one can only say how small is the probability that $\mathbf{E}(u)$ differs from $\bar{u}$ by more than η, which is the reason why the condition under discussion in this paragraph has been phrased as it is in the first sentence hereof.

C. Sectionally Linear Retrospective Rating Formulae in General

In many categories of insurance, the losses which insurers undertake to bear, though they do indeed have about them a certain air of fortuity, are nevertheless subject to control by the insured to an extent which is often surprisingly great, through one or several of the following means: (i) rigorous inspection of premises, material and equipment; (ii) installation of special devices designed to prevent the occurrence of loss-events or to minimize the losses resulting therefrom; (iii) revision of working methods so as to eliminate hazardous operations; (iv) careful selection and training of employees; (v) cooperation with insurers in the adjustment of claims, in efforts toward salvage, and in proceedings to recover damages from negligent third parties; and (vi) in the case of Compensation losses, the re-employment and rehabilitation of injured workmen.

However, with the possible exception of (v), all of these means are costly to the insured. Their employment frequently appears to be in conflict with his immediate interests. Under non-retrospective methods of rating his risks for insurance, any reduction in premium rates which he may obtain as a result of their employment, is generally delayed or spread piecemeal over a period of years, even though their employment brings about an immediate reduction in loss rates.

The fundamental premise upon which the most widely applicable argument in favor of retrospective rating is based, is that the insured will be much more likely to employ those means vigorously and persistently if—in addition to the rather indefinite moral and long-term economic incentives which are always present—he has a very definite and immediate economic incentive to do so. Under a retrospective rating agreement made in accordance with any of the approved plans, such an incentive exists so long as the sum of the modified values of the losses incurred under the insurances which are subject to the agreement, is less than the "allowance for losses in the Maximum Retrospective Premium," i.e., the lower bound (g) of the values of t for which $u = G$ by formula (5).

But if that premise is sound, then *it would be well to have such an incentive persist throughout the entire term of the agreement, regardless of how great the sum of the losses becomes.*

Now the incentive provided by an agreement under which **u** is determined by formula (5) arises out of the fact that, for every increment (Δt) in the sum of the modified losses after that sum has attained the value h and until it attains the value g, the cost to the insured of the very insurances covering the events causing such an increment, is forthwith increased by an amount equal to $C\Delta t$. Obviously, if C is a constant, such an incentive cannot be made to persist regardless of how great the sum of the modified losses becomes, unless **u** is to be unbounded, i.e., unless there is to be no Maximum Retrospective Premium. Yet, as we have previously observed, the insured will certainly prefer an agreement under which **u** is bounded. In fact, except under extraordinary circumstances, he will refuse to be interested in any proposal in which G is greater than $4\mathbf{E}(\mathbf{s})$.

In many cases, given values of $\lambda_1, \lambda_2, , ,$ e.g., $\lambda_k = \$10{,}000$ for every k, it is possible to select a value of g so great that **t**, the sum of the modified losses, is practically certain not to exceed g, and then to select some two of the other parameters, B, C, h, H and G in formula (5), so that—after the remaining three have been determined—C is found to be not too small—say—not less than .30, and neither H nor G is so large that the insured would refuse to be interested in any proposal in which they were set forth as the minimum and maximum values of **u**. In any *such* case, (5) may be regarded as a satisfactory formula, for an agreement drawn in accordance with (5) wherein the parameters have been so selected and determined, will be an agreement under which a definite and immediate economic incentive for the insured to prevent losses is practically certain to persist throughout its term. However, it must be admitted that "practically certain" is a term having no precise meaning.

But in other cases [unless the upper bound of the sequence $\lambda_1, \lambda_2, \ldots$ is so small that the retrospective rating agreement would have little value as a device to encourage the insured's efforts to hold down the value of each $\mathbf{a}_k$, or unless that sequence is constructed in some special way, e.g., monotonically decreasing so that λ_k is very small for values of k which are large relative to $\mathbf{E}(\mathbf{n})$, in consequence of which evaluation of the parameters in (5) would be very difficult] it turns out that, as soon as one selects g so great that **t** is practically certain not to exceed g, then either H or G must be too large to be of any interest to the insured, or else C must be so small that the prospect of a saving of $C\Delta t$ in premium for the insurances in question, could hardly be considered (at least for small values of Δt) as an incentive for the insured to take actions he would not otherwise be inclined to take which might prevent an increase of Δt in losses. In any of *these* cases, or—for that matter—*in any case*, an agreement under which a considerable incentive of the kind under discussion will persist until the sum of the modified losses has attained a certain value, say g_1 or g_2, while some lesser incentive will persist until that sum has attained any value g_m which one chooses to specify, may be drawn in accordance with the following formula:

$$\mathbf{u} = \begin{cases} H & \text{if } \mathbf{t} \leqq h \\ B_1 + C_1 \mathbf{t} & \text{if } h \leqq \mathbf{t} \leqq g_1 \\ B_2 + C_2 \mathbf{t} & \text{if } g_1 \leqq \mathbf{t} \leqq g_2 \\ \cdots\cdots & \cdots\cdots\cdots \\ \cdots\cdots & \cdots\cdots\cdots \\ B_m + C_m \mathbf{t} & \text{if } g_{m-1} \leqq \mathbf{t} \leqq g_m \\ G & \text{if } \mathbf{t} \geqq g_m \end{cases} \tag{12}$$

in which:

$$\left.\begin{array}{ll} B_1 + C_1 h = H & \cdots\cdots\cdots \quad \cdots\cdots\cdots \\ B_2 + C_2 g_1 = B_1 + C_1 g_1 & B_m + C_m g_{m-1} = B_{m-1} + C_{m-1} g_{m-1} \\ \cdots\cdots \quad \cdots\cdots & G = B_m + C_m g_m \end{array}\right\} \tag{13}$$

It is apparent that, when the relationship between **u** and **t** is of the form (12), the graph of **u** plotted against **t** is a segmental arc composed of $m + 2$ straight line segments joined together in succession so that the right end-point of any one of them is the left end-point of the next. The slope of the first and the last is zero; while the slopes of the m intervening segments are $C_1, C_2, \;, \;, C_m$ respectively. For the purpose of drawing an agreement under which the incentives shall be as stated in the preceding paragraph, it is generally sufficient to have $m = 2$ or 3. However, in the interest of generality we shall treat the case in which m may be any finite positive integer. Clearly, (5) is only a special case of (12), i.e., the case in which $m = 1$.

In (12) there are $3m + 3$ parameters (m B's, m C's, m g's, h, H and G) and in (13) there are $m + 1$ independent equations relating them to each other. Hence $2m + 2$ of the parameters must be determined by means other than the system (13). Values for $2m + 1$ of them may be selected arbitrarily—with a view toward the design of a formula under which the insured will have an adequate incentive to prevent losses, provided—however—that the values selected must be such that **u** as a function $R(\mathbf{t})$ shall satisfy the conditions (I) and (II) set forth in Section B. But **u** must also satisfy condition (III), i.e., $\mathbf{E}(\mathbf{u})$ must equal $\mathbf{E}(\mathbf{s}) + \mathbf{E}(\mathbf{w}) + p$, where p is some number satisfying the criteria referred to in (III). Accordingly, after integrating the function of **t** defined by (12) with respect to $F(t)$ over the interval $(0, \infty)$, and collecting terms, we have:

$$H + C_1 \Xi(h) + \sum_{j=1}^{m-1} (C_{j+1} - C_j)\, \Xi(g_j) - C_m \Xi(g_m) = \mathbf{E}(\mathrm{s}) + \mathbf{E}(\mathrm{w}) + p, \qquad (14)$$

in which $\Xi(g_j) = \int_{g_j}^{\infty} (t - g_j)\, dF(t)$ and $\Xi(h)$ is of course defined by the same equation with h in place of g_j.

In case $\lambda_k = \infty$ for every k, so that $\mathrm{t} = \mathrm{s}$, the function of two variables, r and P, first studied by Paul Dorweiler [7] and now generally called the *Excess Pure Premium Ratio*, when defined in terms of operations such as summation performed upon a finite number of observed values of losses or loss-ratios, may be regarded as an *estimate* of the value of $\Xi(g)/\mathbf{E}(\mathrm{s})$ for $g = rP$ and $\mathbf{E}(\mathrm{s}) = aP$, where a is a certain constant usually called the "permissible loss-ratio." But when defined as by Stefan Peters [**11**, p. 589] or as by A. L. Bailey [**1**, p. 67] in terms of a frequency function, the Excess Pure Premium Ratio is the *same* as $\Xi(g)/\mathbf{E}(\mathrm{s})$ for $g = rP$ and $\mathbf{E}(\mathrm{s}) = aP$, provided $\mathrm{t} = \mathrm{s}$ and provided the probability distribution of s is of the continuous type. For $m = 1$, the left side of (14) reduces to an expression equivalent to the product of P by the left side of Perryman's fundamental equation [**10**, p. 7] when the symbols $H'p$ and $G'p$ appearing therein are interpreted to mean $\Xi(h)/P$ and $\Xi(g)/P$, respectively.

Equation (14) and the $m + 1$ equations in (13) form a system of $m + 2$ independent equations; so that after a value for p and values for $2m + 1$ of the parameters in (12) have been selected, the values of the remaining $m + 2$ parameters are uniquely determined by that system, as soon as a particular set J of insurances has been specified and values of $\lambda_1, \lambda_2, \ldots$, prescribed, so that $\mathbf{E}(\mathrm{s})$, $\mathbf{E}(\mathrm{w})$ and $F(t)$ are theoretically fixed.

However, the actual computation of the numerical values of the remaining $m + 2$ parameters may be very difficult. In fact, unless a table is at hand from which a reliable estimate of $\Xi(g)$ for any particular value of g may be obtained, the time and effort required to complete a satisfactory approximation to the values uniquely determined as stated above, may be so great as to prohibit the use of a formula of the type (12). Even when such tables are at hand, one generally has to resort to trial and error methods unless every one but one of the $2m + 2$ parameters appearing in (14) is either included in the set of $2m + 1$ parameters to which values have been pre-assigned by selection, or else its value can be determined from the pre-assigned values and the equations in (13).

A reliable estimate of $\Xi(g)$ for any particular value of g may be obtained by the proper use of a table of Excess Pure Premium Ratios *provided* the distribution of total losses underlying or implied by that table for risks of "standard premium size" P equal to $\mathbf{E}(t)/a$, is approximately the same as $F(t)$. Otherwise, estimates of $\Xi(g)$ obtained by use of that table cannot be regarded as reliable.* Now, sets of insurances to which it is practicable and desirable to apply retrospective rating agreements differ widely among themselves in respect to the characteristics of the probability distributions of total modified losses incurred thereunder, even though we confine our attention to a single particular prescription of values for λ_1, λ_2, , , e.g., λ_k = \$10,000 for every k, and to sets for each of which $\mathbf{E}(t)$ has the same value under that prescription. Therefore, if one is always to have at hand an appropriate table for use in drawing a retrospective rating agreement under which the ultimate premium $\mathbf{u}$ for a set J of insurances is to be determined by a formula of the type (12), regardless of the nature of the set J, then a great many different tables [*either* of Excess Pure Premium Ratios or of values of some other function from which reliable estimates of $\Xi(g)$ may be obtained] will have to be constructed. But the construction of any one such table for any one distribution function $F(t)$ e.g., the construction of any one column in a table of Excess Pure Premium Ratios, requires a very considerable expenditure of time and effort; † and the construction of a number of such tables for a number of different distribution functions all having the same first moment, $\mathbf{E}(t)$, about the point $t = 0$, would be a rather formidable task. When one considers that this would have to be done for each of a series of values of $\mathbf{E}(t)$ ranging from \$3,000 to \$500,000, it is clear that the entire project would be a prodigious undertaking indeed. Consequently, (12) can hardly be said to satisfy the condition expressed in the last paragraph of Section B. All of the statements relative to (12) made in this paragraph or in the preceding one, apply also to (5), since (5) is only a special case of (12).

D. A Class of Non-linear Retrospective Rating Formulae

The facts set forth in the preceding Section, impel us to inquire whether there may not be a class of functions which are fundamentally better adapted to serve as retrospective rating formulae than those belonging to the class of sectionally linear functions defined by (12). The principal thesis of this paper is that there is indeed at least one such class, to wit, the class defined by the equation:

* Cf. Bailey [2, Part VI, Section A].

† For methods of constructing tables of Excess Pure Premium Ratios, see Dorweiler [7 and 8] Valerius [13], Peters [11] and Bailey [2, Part VI, Section B]. Each of the methods outlined or exemplified in these papers, involves a very considerable expenditure of time and effort, not by reason of any defect in the method, but by reason of the character of the function $\Xi(g)$.

$$u = \mathfrak{R}(t) = G - Ae^{-\beta t} \qquad \text{for } t \geqq 0, \tag{15}$$

in which G, A and β are parameters having positive real values independent of every characteristic or consequence of any particular loss-event, and t is defined by (6). Thus, all that was said of $R(t)$ in Section A applies to $\mathfrak{R}(t)$; $\mathfrak{R}(t)$ is defined for every non-negative real value of t; and for $t < 0$, $\mathfrak{R}(t)$ is not defined.

Since A and β are real and positive, $\mathfrak{R}(t)$ is monotone non-decreasing, and so satisfies condition (I) of Section B.

If t', t'' be any pair of possible values of t, such that $t' < t''$, it can easily be shown * that $\mathfrak{R}(t'') - \mathfrak{R}(t') < A\beta(t'' - t')$. Hence, it is always possible by a suitable restriction upon the values of $A\beta$ to ensure that $\mathfrak{R}(t)$ shall satisfy condition (II) of Section B. For example, (II) will certainly be satisfied if $A\beta \leqq 1$.

$\mathfrak{R}(t)$ has a lower bound, $G - A$, which it attains only at the point $t = 0$, and an upper bound, $G = \lim_{t \to \infty} \mathfrak{R}(t)$, which it never attains. Hence, under (15), $G - A$ may be called the *Minimum Retrospective Premium*; G, the *Maximum Retrospective Premium;* and A, the *amplitude* or "swing."

Since $\mathfrak{R}(t)$ is bounded, the existence of $\mathbf{E}(u)$ under (15) is assured. Hence, $\mathfrak{R}(t)$ will satisfy (III) of Section B as soon as the values of the parameters G, A and β have been determined so that

$$\mathbf{E}(u) = \int_0^\infty (G - Ae^{-\beta t})\, dF(t) = \mathbf{E}(s) + \mathbf{E}(w) + p, \tag{16}$$

where p is some number satisfying the criteria referred to in (III).

In any case in which numerical values have been assigned to G, A and β, the value of $u = \mathfrak{R}(t)$ for any particular value of t may be computed to six significant figures in a very few minutes with the help of a seven-place table of common logarithms and an ordinary calculating machine, by means of the formula

$$u = G - \operatorname{antilog}_{10}\left\{\log_{10} A - \beta\, t \log_{10} e\right\}, \tag{15'}$$

which follows at once from (15). Six significant figures will be sufficient in almost all cases encountered in practice. But of course, greater accuracy may be attained by the use of more elaborate tables, or by expanding $e^{-\beta t}$ in a power series and computing the sum of a sufficiently large number of terms.

* See (17).

At every point in the domain of t, $\mathfrak{R}(t)$ possesses a positive first differential coefficient, $\mathfrak{R}'(t) = A\beta e^{-\beta t}$. Hence, under an agreement drawn in accordance with (15), the insured would always have a definite and immediate economic incentive to prevent losses, regardless of the height to which the sum of the losses already incurred may have arisen. For at every point in the stochastic process by which the sum of the modified losses attains its ultimate value t, the premium which the insured will be obligated ultimately to pay to the insurer is increasing (or the return to which the insured will be entitled out of the amount he has already paid to the insurer is decreasing) at the rate of $A\beta e^{-\beta t}$ per monetary unit of increase in the sum of the modified losses, where t is the attained value of that sum at that point.

Thus, $A\beta e^{-\beta t}$ may be regarded as an index of such incentive under such an agreement. It has its maximum value $A\beta$ at $t = 0$; it decreases as t increases, approaching 0 asymptotically as $t \to \infty$. Admittedly, in any particular case, for very large values of t the incentive will be so small as to be inconsiderable. But, in every case in which the making of a retrospective rating agreement could be at all justified, it is possible by proper choice of the values of λ_1, λ_2, , , A and β, to construct such an agreement in accordance with (15), under which the incentive will be quite considerable, say $A\beta e^{-\beta t} \geq .30$, for every $t \leq t'$, where t' is some number of which the a priori probability that t will exceed t' is small, say $1 - F(t') < .05$, while some lesser incentive (smaller and continuously diminishing as t increases) will persist for every $t > t'$, without making G so great that the insured will not accept the agreement, or λ_k so small that the agreement would have little value as a device to encourage the insured's efforts to hold down the value of a_k.

Thus, we have shown that under a proper choice of values of G, A and β, $\mathfrak{R}(t)$ satisfies all of the necessary conditions of practicability as a retrospective rating formula, and that $\mathfrak{R}(t)$ also possesses all of the properties which it is desirable that such a formula should possess, except that we have not yet shown that $\mathfrak{R}(t)$ satisfies the condition expressed in the last paragraph of Section B. That it does indeed satisfy that condition in a great many cases, is the theme of the following Section.

E. Evaluation of the Parameters G, A and β.

In (5) there are three parameters, G, A and β, to which numerical values must be assigned; and (16) is the only equation relating their values one to another which must be satisfied. Hence, values for two of them may be selected arbitrarily—or with a view toward the design of a formula which meets the insured's requirements and at the same time offers the insured an adequate incentive to prevent losses, provided—however—that the two values selected and the third which is consequently determined by (16) must, when taken together, be such that $\mathfrak{R}(\mathbf{t})$ satisfies condition (II) of Section B. [As previously remarked, $\mathfrak{R}(\mathbf{t})$ always satisfies condition (I) since A and β are real and positive by definition.]

If we set $\dfrac{\alpha(t', t'')}{t'' - t'} = \rho(t', t'')$, then, since $\delta(t', t'') = t'' - t'$,

condition (II) with respect to $\mathfrak{R}(\mathbf{t})$ may be expressed thus:

$$\frac{\mathfrak{R}(t'') - \mathfrak{R}(t')}{t'' - t'} \leqq 1 + \rho(t', t'')$$

for every pair, t', t'', such that $t' < t''$.

But, by the Theorem of the Mean and the fact that $\mathfrak{R}'(\mathbf{t}) = A\beta e^{-\beta \mathbf{t}}$ is a decreasing function,

$$\frac{\mathfrak{R}(t'') - \mathfrak{R}(t')}{t'' - t'} < A\beta e^{-\beta t'} \tag{17}$$

for every such pair. Therefore, if $\quad A\beta e^{-\beta t'} \leqq 1 + \rho(t', t'') \qquad (18)$

for every such pair, (II) will certainly be satisfied. In the case of any particular set J of insurances, all of the facts upon which the allowance $\alpha(t', t'')$ depends (e.g., the past expense experience on similar insurances, the tax rates applicable to $\mathbf{u}$, whether or not the term *losses* is to be defined to include allocated claim expenses, etc.) will be known or can be ascertained; whereupon a value can be assigned to $\alpha(0, \bar{t})$ where $\bar{t}$ is some relatively small possible value of $\mathbf{t}$, e.g., .01 $\mathbf{E}(\mathbf{t})$, from which the value of $\rho(0, \bar{t})$ can be computed. $\rho(t', t'')$ can be considered constant for small values of t' and t''. Hence, if $A\beta \leqq 1 + \rho(0, \bar{t})$, then (18) will be satisfied for small values of t' and t''; and when it is satisfied for small values, it will usually be found that $A\beta e^{-\beta t}$, decreases fast enough as t increases so that (18) is satisfied for all values of t' and t'', $t' < t''$.

In accordance with (11), in order that the values of the parameters G, A and β shall be such that $\mathfrak{R}(\mathbf{t})$ satisfies both (II) and (III), it is necessary that $\mathfrak{R}(0)$ be not less than $\mathbf{E}(\mathbf{s} - \mathbf{t}) + W(0) + p$, which is positive. Now, $\mathfrak{R}(0) = G - A$. Therefore, the choice of values for two of the parameters must be governed not only by (18), but also by the following relationship:

$$G - A \geqq \mathbf{E}(\mathbf{s} - \mathbf{t}) + W(0) + p > 0 \tag{19}$$

Of course, any set of values of G, A and β which satisfies (16) and (18), will also satisfy (19). But (19) is useful as a guide in selecting the values for two of the parameters which are to be substituted in (16) in order to obtain a value for the third.

Let us turn our attention now to the integral which appears in (16).

$$\int_0^\infty (G - Ae^{-\beta t})dF(t) = G\int_0^\infty dF(t) - A\int_0^\infty e^{-\beta t}\, dF(t) \tag{20}$$

The integral appearing in the second term on the right is the Laplace-Stieltjes transform of $F(t)$, which we shall denote by $\mathfrak{L}\{F(t),\beta\}$.

In general, if x be any random variable, we shall write

$$\int_0^\infty e^{-\beta x}\, dF(x) = \mathfrak{L}\{F(x),\beta\}. \tag{21}$$

If the probability distribution of x is of the continuous type, the Lebesgue-Stieltjes integral in (21) has the same value as the ordinary Riemann integral,

$\int_0^\infty e^{-\beta x} f(x)dx$, in which $f(x) = F'(x)$. We shall denote this latter integral,

which is the Laplace transform of the frequency function $f(x)$, by $\mathfrak{L}\{f(x),\beta\}$. These notations emphasize the fact that the Laplace-Stieltjes transform is a function of β, of which the character is determined as soon as the probability distribution of the random variable x is known.

If in (21) we were to replace 0 as the lower limit of integration by $-\infty$ we would have the bilateral Laplace-Stieltjes transform, $\mathfrak{L}_2\{F(x), \beta\}$, which — in view of (2) — is the same as $\mathbf{E}(e^{-\beta x})$. $\mathfrak{L}_2\{F(x), -i\theta\}$ regarded as a function of θ, is the familiar *characteristic function* of the distribution of x; and $\mathfrak{L}_2\{F(x), -\theta\}$ regarded as function of θ, is often called the *moment-generating function* of the distribution. If x is a random variable (like **n**, $\mathbf{a}_k$, $\mathbf{v}_k$, **s**, **t**) which cannot take a negative value, then of course $\mathfrak{L}_2\{F(x), \beta\} = \mathfrak{L}\{F(x), \beta\}$. Furthermore, if x is any such variable, then

$\int_0^\infty dF(x) = 1$, from which it follows (by virtue of certain fundamental theorems

on Lebesgue-Stieltjes integrals, * and since $0 < e^{-\beta x} \leq 1$ for $x \geq 0$) that $\mathfrak{L}\{F(x),\beta\}$ exists and $0 < \mathfrak{L}\{F(x),\beta\} \leq 1$ for every real value of $\beta \geq 0$. In

* Cramér [5, p. 63].

particular, $\mathfrak{L}\{F(x),0\} = 1$. It can also be shown * that $\mathfrak{L}\{F(x),\beta\}$ is continuous at every point in the interval $0 \leqq \beta < \infty$, and that it decreases monotonically from 1 at $\beta = 0$ and approaches $F(0)$ as $\beta \to \infty$.

In view of (20), (21) and the fact that $\int_0^\infty dF(t) = 1$, equation (16) may be rewritten as follows:

$$G - A\,\mathfrak{L}\{F(t),\beta\} = \mathbf{E}(\mathrm{s}) + \mathbf{E}(\mathrm{w}) + p \tag{22}$$

In the structure of retrospective rating formulae of the non-linear type (15), this single equation (22) is the analogue of the system of $m + 2$ equations consisting of (14) and the $m + 1$ equations (13) in the structure of retrospective rating formulae of the sectionally linear type (12). Or perhaps it would be better to say that (22) is the analogue of (14) and that in the structure of formulae of the type (15) there is no analogue of the $m + 1$ equations (13), for (13) is really a part of the definition of the class of functions from which sectionally linear retrospective rating formulae are chosen, whereas the class of non-linear functions under discussion is defined by the single equation (15) and the assertion that G, A and β are real and positive.

When a particular set J of insurances has been specified and values of $\lambda_1, \lambda_2, , ,$ prescribed, $\mathbf{E}(\mathrm{s})$, $\mathbf{E}(\mathrm{w})$ and $F(t)$ are theoretically fixed, so that as soon as values of p and two of the parameters G, A and β have been selected, the value of the third is uniquely determined by (22).

If $F(t)$ is (or can be closely approximated by) a function whose Laplace-Stieltjes transform can be expressed in a few simple terms easy to evaluate numerically, then the value of the third parameter can be computed without difficulty. For example, if the probability distribution of t is of the continuous type, and if the frequency function of t is:

$$\left.\begin{aligned} f(t) &= 0 && \text{for } t \leqq 0, \\ f(t) &= \frac{c^b}{\Gamma(b)}\, t^{b-1} e^{-ct} && \text{for } t > 0, \end{aligned}\right\} \tag{23}$$

then
$$\mathfrak{L}\{F(t),\beta\} = \mathfrak{L}\{f(t),\beta\} = \left(1 + \frac{\beta}{c}\right)^{-b} \tag{24}$$

* By application of the first two theorems in Section 7.3 of Cramér [5].

On substituting this last expression for $\mathfrak{L}\{F(t),\beta\}$ in (22), one sees at once that the value of any one of the three parameters, G, A or β, may be computed without much effort as soon as values for the other two have been selected and the values of b, c and $\mathbf{E}(\mathrm{s}) + \mathbf{E}(\mathrm{w}) + p$ are known.

Frequency functions of the class defined by (23), which is a sub-class of those of Pearson Type III, will be found to fit distributions of t very satisfactorily in many cases; * and in those cases it is already clear that non-linear functions of the form (15) are better adapted than sectionally linear functions of the form (5) or (12) to serve as retrospective rating formulae.

Of course, it may be argued that if $F(t)$ is such that, for any particular g, the numerical value of $\Xi(g)$ can easily be found, then the values of any $m + 2$ of the parameters in (12) can be computed without difficulty as soon as values for the other $2m + 1$ of them have been selected. But the functions which can be satisfactorily fitted to distributions of t without extraordinary labor, do not have that property. They are of such a character that integrals of the form $\Xi(g) = \int_g^\infty (t - g)dF(t)$ are not easy to evaluate for $g > 0$; whereas some of them [notably those corresponding to the frequency functions defined by (23)] have Laplace-Stieltjes transforms which are very easy to evaluate for any given β, and which in addition are such that, if the value of the transform be given, the corresponding value of β may very quickly be computed.

In any case, whether or not $F(t)$ can be satisfactorily approximated by a function whose Laplace-Stieltjes transform has the virtue discussed above, $\mathfrak{L}\{F(t),\beta\}$ itself possesses two properties by reason of which the non-linear functions defined by (15) are especially well adapted to serve as retrospective rating formulae. The first of these is set forth in the following paragraph.

By a MacLaurin expansion of the factor $e^{-\beta x}$ appearing in (21), we have:

$$\mathfrak{L}\{F(x),\beta\} = \int_0^\infty \left\{ 1 - \beta x + \frac{(\beta x)^2}{\underline{|2}} - \frac{(\beta x)^3}{\underline{|3}} + \ldots + q_m(\beta,x) \right\} dF(x)$$

in which $q_m(\beta, x)$, the remainder after the term of degree $m - 1$, is given by the equation: $q_m(\beta, x) = (-1)^m e^{-\beta c x} \dfrac{(\beta x)^m}{\underline{|m}}$, where $0 < c < 1$. Now, if x is any random variable (like n, a_k, v_k, s, t) which cannot take a negative

* Cf. A. L. Bailey [1, p. 78 and Table 9].

value, then $\int_0^\infty dF(x) = 1$ and $\int_0^\infty \frac{(\beta x)^k}{\underline{|k}} dF(x) = \frac{\beta^k}{\underline{|k}} \mathbf{E}(\mathrm{x}^k)$ for every $k \geqq 0$,

in which $\mathbf{E}(\mathrm{x}^k)$ is the kth moment of the distribution of x about the point $\mathrm{x} = 0$. Therefore, if $\mathbf{E}(\mathrm{x}^m)$ exists, then

$$\mathfrak{L}\{F(x),\beta\} = 1 - \beta\mathbf{E}(\mathrm{x}) + \frac{\beta^2}{\underline{|2}}\mathbf{E}(\mathrm{x}^2) - \frac{\beta^3}{\underline{|3}}\mathbf{E}(\mathrm{x}^3) + \ldots + Q_m(\beta),$$

in which $Q_m(\beta) = \int_0^\infty q_m(\beta, x)dF(x)$. So $|Q_m(\beta)| \leq \left|\frac{\beta^m}{\underline{|m}}\mathbf{E}(\mathrm{x}^m)\right|$

Thus, if the first $m - 1$ moments of the distribution of t about the point $\mathrm{t} = 0$ are known, then for any particular value of β, an estimate of the transform $\mathfrak{L}\{F(t),\beta\}$ is immediately given by the polynomial

$$1 - \beta\mathbf{E}(\mathrm{t}) + \frac{\beta^2}{\underline{|2}}\mathrm{E}(\mathrm{t}^2) - \ldots + (-1)^{m-1}\frac{\beta^{m-1}}{\underline{|m-1}}\mathbf{E}(\mathrm{t}^{m-1}); \qquad (25)$$

and the absolute difference between that estimate and the true value of the transform is less than $\left|\frac{\beta^m}{\underline{|m}}\mathbf{E}(\mathrm{t}^m)\right|$, which is very small for sufficienty small values of β.

Now, in the case of any particular set J of insurances, all of the moments, $\mathbf{E}(\mathrm{t}^k)$ for $k = 1, 2, 3, \ldots$, certainly exist, since the range of possible values of t is bounded by the total amount of wealth in the world. But none of them is ever exactly known. However. estimates of their values may be made from data relative to past experience on a properly selected sample of insurances; and although the variances of the sampling distributions of moments of order $k > 3$ are very great, the contribution to the probable error of (25) resulting from the substitution therein of such an estimate of $\mathbf{E}(\mathrm{t}^k)$ in place of $\mathbf{E}(\mathrm{t}^k)$ itself, will be very small for small values of β, provided—of course—that the sample from which such estimate was derived, was very large and properly chosen. Thus, both the absolute difference $|Q_m(\beta)|$ and the probable error introduced into (25) by the substitution of estimates of the moments therein appearing, in place of the moments themselves, will be small in comparison with the true value of the transform $\mathfrak{L}\{F(t),\beta\}$ for sufficiently small values of β.

Fortunately, in many cases, the only values of β which are of any practical interest, are ones which are very small indeed. For, in general, $A\beta$ can not be greater than 1.25 without violating condition (II) of Section B, at least for small values of t''; and in retrospective rating formulae of the type (15), the amplitude or "swing" A can be quite high, e.g., 3 times the Expected Losses, without in effect shifting almost all of the risk back to the insured. In many cases, values of $A < 2\mathbf{E}(t)$ will be of no interest; and in those cases, only values of $\beta < 1.25/2\mathrm{E}(t)$ will be of any practical importance.

For example, suppose that $\mathbf{E}(t)$ has been estimated to be \$50,000 with a probable error of ϵ_1, that A is to be \$100,000, and that $A\beta$ is to be 1.00. Then $\beta = .00001$. Suppose further that it is known that $\mathbf{E}(t^4) < (\$10)^{19}$, and that estimates, $\overline{\mathbf{E}(t^2)}$ and $\overline{\mathbf{E}(t^3)}$, of $\mathbf{E}(t^2)$ and $\mathbf{E}(t^3)$ are at hand, of which the probable errors are ϵ_2 and ϵ_3, respectively. Then

$1 - \beta(50{,}000) + \frac{\beta^2}{\underline{|2}}\,\overline{\mathbf{E}(t^2)} - \frac{\beta^3}{\underline{|3}}\,\overline{\mathbf{E}(t^3)}$ gives an estimate (L) of the value of $\mathfrak{L}\left\{F(t), \beta\right\}$ for $\beta = .00001$, of which the error made by neglecting all moments of order greater than 3, is less than .0042, and the probable error ($\bar{\epsilon}$) due to the use of estimates of $\mathbf{E}(t)$, $\mathbf{E}(t^2)$ and $\mathbf{E}(t^3)$ is less than

$.00001\epsilon_1 + \frac{(.00001)^2\,\epsilon_2}{\underline{|2}} + \frac{(.00001)^3\,\epsilon_3}{\underline{|3}}$. When a value $\bar{u}$ for the right side of equation (22) has been given, L may be substituted for $\mathfrak{L}\left\{F(t), \beta\right\}$ and \$100,000 substituted for A in the left side, and the equation solved for G. With G equal to the value thus obtained, $A = \$100{,}000$, and $\beta = .00001$, $\mathbf{E}(u)$ will fall short of $\bar{u}$ by less than \$420 with a probable error less than $\bar{\epsilon}$ times \$100,000.

Thus, it is apparent that—in many cases—the parameters in (15) can be determined in a manner satisfying the condition set forth in the last paragraph of Section B, without any knowledge of the shape of the probability distribution of the sum of the modified losses, that is to say—without any knowledge of the function $F(t)$, provided only that one has reliable estimates of the first three or four moments of that distribution about the point $t = 0$. But in no case can any such statement be made relative to the parameters in (5) or (12). In order to determine them, one must have at hand reliable estimates of the value of the function $\Xi(g)$ for at least some values of g; and in order to obtain such estimates, one must have knowledge of the shape of the aforesaid distribution.

The second of the two properties of the transform $\mathfrak{L}\left\{F(t), \beta\right\}$ by reason of which the non-linear functions defined by (15) are especially well adapted to serve as retrospective rating formulae, is expressed in the following

Theorem. *The Laplace-Stieltjes transform of the distribution function of the sum of any finite number (N) of independent random variables, none of which can take a negative value, is equal to the product of the Laplace-Stieltjes transforms of the distribution functions of the N terms in that sum.*

The following proof for the case in which $N = 2$ can easily be extended to the case in which N is any finite number.

Let x and y be any two independent random variables, neither of which can take a negative value; and let $\mathrm{t} = \mathrm{x} + \mathrm{y}$. Then $e^{-\beta \mathrm{x}}$ and $e^{-\beta \mathrm{y}}$ are independent random variables, and

$$\mathbf{E}\,(e^{-\beta \mathrm{t}}) = \mathbf{E}\,(e^{-\beta(\mathrm{x}+\mathrm{y})}) = \mathbf{E}\,(e^{-\beta \mathrm{x}} e^{-\beta \mathrm{y}}) = \mathbf{E}\,(e^{-\beta \mathrm{x}}) \times \mathbf{E}\,(e^{-\beta \mathrm{y}}).\ *$$

But, by (2) and (21), in view of the hypothesis that neither x nor y can take a negative value,

$$\mathbf{E}(e^{-\beta \mathrm{t}}) = \mathfrak{L}\{F(t), \beta\}; \quad \mathbf{E}(e^{-\beta \mathrm{x}}) = \mathfrak{L}\{F(x), \beta\}; \quad \mathbf{E}(e^{-\beta \mathrm{y}}) = \mathfrak{L}\{F(y), \beta\}.$$

Hence: $$\mathfrak{L}\{F(t), \beta\} = \mathfrak{L}\{F(x), \beta\} \times \mathfrak{L}\{F(y), \beta\}.$$

Now, suppose that the set J to which a single retrospective rating agreement is to be applied, consists of N insurances, $j_1, j_2, \ , \ , j_N$, such that the sums, $\mathrm{t}_1, \mathrm{t}_2, \ , \ , \mathrm{t}_N$, of the modified losses incurred under $j_1, j_2, \ , \ , j_N$, respectively, are independent random variables. Then
$\mathrm{t} = \mathrm{t}_1 + \mathrm{t}_2 + \ldots. + \mathrm{t}_N$; and by the foregoing theorem,

$$\mathfrak{L}\{F(t), \beta\} = \mathfrak{L}\{F(t_1), \beta\} \times \mathfrak{L}\{F(t_2), \beta\} \times \ldots. \times \mathfrak{L}\{F(t_N), \beta\}. \qquad (26)$$

Thus, if for each of the insurances $j_1, j_2, \ , \ , j_N$, we know the Laplace-Stieltjes transform of the distribution function of the sum of the modified losses incurred thereunder, then $\mathfrak{L}\{F(t), \beta\}$ is immediately given by this simple equation (26), the right side of which may then be substituted for $\mathfrak{L}\{F(t), \beta\}$ in (22) in order to determine the value of any one of the parameters G, A and β as soon as values for the other two have been selected.

The analysis of J into components, $j_1, j_2, \ , \ , j_N$, need not follow traditional lines of insurance classification. For example, suppose J consists of group life insurance, non-occupational disability insurance, and Workmen's Compensation insurance covering all employees of some one corporation. Then j_1 may be defined to be the first and the last of these insurances in so far as they apply to events resulting in the death of one or several employees by reason of which an obligation arises under the Workmen's Compensation Law to pay benefits in addition to those afforded by the group life insurance; j_2

* Cramér [5, 14.5 and 15.3.4].

may be defined to be the group life insurance in so far as it applies to deaths not resulting from events to which j_1 applies; j_3 may be defined to be the non-occupational disability insurance; and j_4 may be defined to be the Workmen's Compensation insurance in so far as it applies to events other than those to which j_1 applies. Under such definitions, the random variables $\mathbf{t}_1$, $\mathbf{t}_2$, $\mathbf{t}_3$ and $\mathbf{t}_4$ may – with negligible error – be considered independent.

Now, let us consider any single one (j_i) of the insurances, $j_1, j_2, \ , \ , j_N$, of which the set J is composed. Let $\mathbf{m}$ be the number of loss-events under j_i, and $\mathbf{v}_1, \mathbf{v}_2, \ , \ , \mathbf{v}_\mathbf{m}$ be the modified losses consequent thereto. Suppose (i) that $\mathbf{v}_1, \mathbf{v}_2, \ , \ , \mathbf{v}_\mathbf{m}$ are independent and (ii) that—prior to the actual occurrence of any loss-event—their distribution functions are identical, so that the distribution function of any one of them may be represented by $F_i(v)$. Let $\mathbf{t}_i = \mathbf{v}_1 + \mathbf{v}_2 + \ldots . + \mathbf{v}_\mathbf{m}$, and let $\mathbf{t}_i^{(m)}$ represent the value of $\mathbf{t}_i$ on the hypothesis that $\mathbf{m}$ takes the value m.

Then $\mathfrak{L}\{F(t_i), \beta\} = \mathbf{E}(e^{-\beta \mathbf{t}_i}) = \int_0^\infty \underset{\mathbf{m}=m}{\mathbf{E}}(e^{-\beta \mathbf{t}_i}) \ dF(m)$, where $\underset{\mathbf{m}=m}{\mathbf{E}}(e^{-\beta \mathbf{t}_i})$ is the *conditional mathematical expectation* (or *conditional mean value*) of $e^{-\beta \mathbf{t}_i}$) on the hypothesis that $\mathbf{m} = m$,* given by the integral $\int_0^\infty e^{-\beta \mathbf{t}_i^{(m)}} dF(t_i^{(m)})$, which is the Laplace-Steiltjes transform of $F(t_i^{(m)})$. But $\mathbf{t}_i^{(m)}$ is the sum of m independent random variables each of which has the distribution function $F_i(v)$.

Hence, by the Theorem stated above, $\underset{\mathbf{m}=m}{\mathbf{E}}(e^{-\beta \mathbf{t}_i}) = \left[\mathfrak{L}\{F_i(v), \beta\}\right]^m$.

Therefore:
$$\mathfrak{L}\{F(t_i), \beta\} = \int_0^\infty \left[\mathfrak{L}\{F_i(v), \beta\}\right]^m dF(m). \tag{27}$$

Now, suppose (iii) that $\mathbf{m}$ has a Poisson distribution, that is to say, suppose

$F(m) = \sum_{k=0}^{m} \frac{(\omega_i)^k}{\underline{|k}} e^{-\omega_i}$, where $\omega_i = \mathbf{E}(\mathbf{m})$. Then the Lebesgue-Stietljes integral on the right side of (27) becomes equivalent to the infinite series

* Kolmogoroff [9, Fünftes Kapitel, § 4].

$\sum_{k=0}^{\infty} \frac{(\omega_i)^k}{\underline{|k}} e^{-\omega_i} \Big[\mathfrak{L}\{F_i(v), \beta\} \Big]^k$, which converges for all real non-negative values of ω_i and β, to e^{θ_i} where $\theta_i = \omega_i \Big[\mathfrak{L}\{F_i(v), \beta\} - 1 \Big]$. (28)

Given any particular set of insurances to which it would be practicable and desirable to apply a retrospective rating agreement, it is generally possible to define the members $j_1, j_2, , , j_N$, and the terms *loss* and *loss-event* in such a way that (a) the conditions (i), (ii) and (iii) are realized, or can be considered as realized for all practical purposes, in the case of each of the insurances $j_1, j_2, , , j_N$; * and that (b) the sums, $t_1, t_2, , , t_N$, of the modified losses incurred under them, respectively, are independent random variables, or can for all practical purposes be so considered. Then each of the transforms on the right side of (26) may be expressed in the form (28), so that we have

$$\mathfrak{L}\{F(t), \beta\} = e^{\theta_1} \times e^{\theta_2} \times \ldots . \times e^{\theta_N} = e^{\Phi}$$

$$\text{where } \Phi = \sum_{i=1}^{N} \omega_i \Big[\mathfrak{L}\{F_i(v), \beta\} - 1 \Big]. \tag{29}$$

Thus we have shown that, if for each insurance (j_i) included in the set J, we have a realiable estimate $(\bar{\omega}_i)$ of the mathematical expectation of the number of loss-events, and reliable estimates, $L_i(\beta)$, of the Laplace-Stieltjes transform of the distribution function of the modified losses each resulting from a single loss-event, for the values of β in which we may be interested, then a reliable estimate of $\mathfrak{L}\{F(t), \beta\}$ for each of those values of β may be obtained by a fairly simple computation, indicated by (29). If, for any particular value of β, Φ' be the estimate of Φ obtained by substituting the estimates $\bar{\omega}_i$ and $L_i(\beta)$ for ω_i and $\mathfrak{L}\{F_i(v), \beta\}$ in (29), and if ϵ be the *actual* error of Φ', then the absolute difference, $| \mathbf{E}(u) - \bar{u} |$ will be $Ae^{\Phi'} \times | 1 - e^{-\epsilon} |$, when the parameters G and A have been determined so as to satisfy equation (22) with $e^{\Phi'}$ substituted in place of $\mathfrak{L}\{F(t), \beta\}$ and $\bar{u}$ in place of $\mathbf{E}(s) + \mathbf{E}(w) + p$. It is clear, therefore, that the parameters in (15) can be determined in a manner eminently satisfying the condition set forth in the last paragraph of Section B in every case in which sufficiently reliable estimates, $\bar{\omega}_i$ and $L_i(\beta)$ for $i = 1, 2, , , N$ and for the values of β in which we may be interested, can be obtained with relatively small expenditure of time and effort.

* Concerning (iii), see Bailey [1, Part I, Section A]. A very interesting discussion of imperfect realizations of condition (iii) will be found in Carleton [3]. Note that when (iii) is not realized even approximately, the transform of $F(t_i)$ may still be evaluated in terms of the transform of $F_i(v)$ and the distrition of m by means of (27), provided (i) and (ii) are satisfied.

Such estimates can indeed be so obtained from statistics which are regularly compiled by insurers and their rating organizations, in the cases of a great many sets of insurances to which it is practicable and desirable to apply retrospective rating agreements. For example, let us consider Workmen's Compensation and Employers' Liability insurances, with respect to which we may assume that only those rules of the form (7) in which $\lambda_k = \Lambda$ for every k, will be of any interest. Each such insurance with respect to operations falling within several Manual classes may be considered to be several differennt insurances, one (j_i) for each such class. Then $\bar{\omega}_i$ may be obtained by multiplying the exposures of the ith class for the risk in question, by the observed mean accident frequency per unit of exposure for all exposures of that class within—say—the three latest Policy Years, modified in accordance with an appropriate individual risk experience rating procedure.

For each jurisdiction, all Manual classes could be divided into a small number of groups, e.g., five groups, such that the distribution function of actual losses each resulting from a single accident arising out of operations of any one class, could be shown to be approximately the same as the distribution function of actual losses resulting from a single accident arising out of operations of any other class in the same group. For each such group, a table of values of the transform $\mathfrak{L}\{F(v), \beta\}$ [where $F(v)$ is the distribution function of modified losses each resulting from a single accident, which is characteristic of that group] could be constructed for values of β ranging from 0 to .01 and for each of several values of Λ, e.g., \$5,000, \$10,000, \$15,000, \$20,000 and ∞. Values of $\beta > .01$ can be shown to be of no practical interest. For small values of β, $\mathfrak{L}\{F(v), \beta\}$ could be very satisfactorily approximated by polynomials of the form (25) with v in place of t. The construction of such tables would not be a very laborious task, especially if it were found that for some groups the distribution of v could be satisfactorily approximated by a frequency function of the form (23). From such tables, the estimtes $L_i(\beta)$ could be read directly or determined by interpolation.

Finally, it may be asserted that no rule comparable in simplicity to either (26) or (29) can be stated for the evaluation of $\Xi(g)$ in terms of statistics relative to individual members of a set consisting of several different insurances. Therefore, when any such set is to be the subject of a single retrospective rating agreement, the determination of the parameters in a sectionally linear formula of the type (5) or (12) to a prescribed degree of accuracy, will necessarily involve a greater expenditure of time and effort than the determination of the parameters in a non-linear formula of the type (15), to the same degree of accuracy.

BIBLIOGRAPHY

1. A. L. Bailey, *Sampling Theory in Casualty Insurance, Introduction and Parts I and II*, P.C.A.S. Vol. XXIX, No. 59 (1942) pp. 50-93.

2. ———— *Sampling Theory in Casualty Insurance, Parts III-VII*, P.C.A.S. Vol. XXX, No. 60 (1943) pp. 31-65.

3. John Carleton, *Non-Random Accident Distributions and the Poisson Series*, P.C.A.S. Vol. XXXII, No. 62 (1945) pp. 21-26.

4. T. O. Carlson, *An Actuarial Analysis of Retrospective Rating*, P.C.A.S. Vol. XXVIII, No. 58 (1942) pp. 283-323.

5. Harald Cramér, *Mathematical Methods of Statistics*, Princeton: Princeton University Press, 1946.

6. ———— *Random Variables and Probability Distributions*, London: Cambridge University Press, 1937.

7. Paul Dorweiler, *Observations on Making Rates for Excess Compensation Insurance*, P.C.A.S. Vol. XIII, No. 28 (1927) pp. 154-180.

8. ———— *On Graduating Excess Pure Premium Ratios*, P.C.A.S. Vol. XXVIII, No. 57 (1941) pp. 132-157.

9. A. Kolmogoroff, *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Berlin: 1933; New York: Chelsea Publishing Co., 1946.

10. F. S. Perryman, *Possible Values for Retrospective Rating Plans*, P.C.A.S. Vol. XXXI, No. 61 (1944) pp. 5-34.

11. Stefan Peters, *Discussion of Dorweiler* [8], P.C.A.S. Vol. XXVIII, No. 58 (1942) pp. 588-590.

12. J. V. Uspensky, *Introduction to Mathematical Probability*, New York and London: McGraw-Hill, 1937.

13. N. M. Valerius, *Risk Distributions Underlying Insurance Charges in the Retrospective Rating Plan*, P.C.A.S. Vol. XXIX, No. 59 (1942) pp. 94-121.