

ON THE REPRESENTATION OF LOSS AND INDEMNITY DISTRIBUTIONS

YOONG-SIN LEE

Abstract

In this paper many relations and equations relating pure premium or expected value quantities are presented in terms of random variables. This is made possible by the use of the indicator function so that awkward representations of functions of loss are simplified. Relations and formulas on such topics as basic limits losses, excess of loss coverages and retrospective rating are presented in stronger, more primitive forms. The related mathematics is often simplified and, in particular, an effective technique for handling trend is presented.

I. INTRODUCTION

The loss distribution is an essential component in actuarial work but because of the various limitations of payment of loss in an insurance contract, the indemnity is not always identical to the loss. Hence the indemnity often has a rather complicated representation in terms of the original loss. Actuarial formulas and expressions become less tractable and more difficult to understand. For this reason the treatment of basic limits losses, excess of loss coverages and retrospective rating, for example, are replete with complicated mathematical relations, and the formulas and equations presented in the literature provide little insight into their meanings. This paper uses the indicator function to give the indemnity a single representation as a random variable. Many of the mathematical relations connecting expected value pure premiums are now turned into more primitive, stronger relations between random variables. The algebra in manipulating the mathematical relations is often reduced, and the relations themselves become more transparent when viewed this way. Mathematical notations are known to have revolutionized mathematics and science in the long history of these disciplines;

witness the invention of zero, the use of Arabic numerals in place of the Roman numerals, and the introduction of vectors and matrices in modern mathematics. While it is not pretended that the use of the indicator function will have such portentous effects in actuarial science, it does simplify actuarial mathematics, add new insight in many areas, and lead more easily to some new results.

2. DEFINITIONS

It will clarify matters if we distinguish between the original loss incurred by the insured and the indemnity paid by the insurer. Let us represent the original loss by the random variable X , which follows the loss distribution. We also assume that the indemnity depends solely on the loss so that, being a function of the random variable X , it is itself a random variable with distribution called the indemnity distribution. The indemnity relates to the loss X in many ways, depending on the nature of the insurance contract. Typically, the indemnity as a function of the loss assumes different functional forms over different ranges of the size of loss. This contributes to the unwieldiness in the mathematics of the indemnity distribution. For example, if the original loss is X and the indemnity is the basic limits loss with limit k , then the indemnity $g(X;k)$ may be described as

$$g(X;k) = \begin{cases} X & 0 < X \leq k \\ k & X > k. \end{cases}$$

Note that $g(X;k)$ is a random variable, but is represented above and elsewhere in the *Proceedings* by two separate expressions. This representation has obstructed the view of the users of this random variable, making it awkward to work with. It can be represented in a single expression using the indicator function:

$$g(X;k) = XI_{(0, k]}(X) + kI_{(k, \infty)}(X),$$

where $I_S(x)$ is an indicator function defined as follows:

$$I_S(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{otherwise} \end{cases}$$

with S representing a set of possible values of x . The inclusion of $X = k$

in the lower line segment instead of in the upper agrees with the convention in defining the distribution function $F(x) = \text{Prob}(X \leq x)$. Although in the above definition of $I(\cdot)$ the argument is a numerical variable, the definition extends easily to the case of a random variable argument in the usual way. When expressed in this form many expressions involving the random variable $g(X;k)$ can be manipulated more easily. The following simple properties of the indicator function contribute much to its power:

$$I_{S_1}(x) I_{S_2}(x) = I_{S_1 \cap S_2}(x),$$

$$I_{S_1}(x) + I_{S_2}(x) = I_{S_1 \cup S_2}(x) \quad \text{if } S_1 \cap S_2 = \emptyset.$$

It can be easily deduced that

$$I_{S_1}(x) - I_{S_2}(x) = I_{S_1 \cap S_2^c}(x) \quad \text{if } S_2 \subseteq S_1$$

and

$$I_{S_1}(x) I_{S_2}(x) = 0 \quad \text{if } S_1 \cap S_2 = \emptyset.$$

The prescribed statistical text, Mood, Graybill and Boes [4], uses the indicator function quite freely. However, examples in casualty actuarial science where the indicator function is applicable are much more interesting and richer, and its use could also be more sophisticated. LaRose [2] presents the following notations for expected values of certain functions of the loss. In his notation

$$X1(k) = \frac{1}{\alpha} \int_0^k t dF(t),$$

where $\alpha = \int_0^\infty t dF(t)$ is the mean loss,

$$X2(k) = \frac{1}{\alpha} \int_0^k t dF(t) + \frac{k}{\alpha} \int_k^\infty dF(t),$$

and

$$X3(k) = \frac{1}{\alpha} \int_k^\infty (t - k) dF(t).$$

He shows that many actuarial functions expressed in a variety of forms can be represented in terms of these three quantities. It is clear that $X1(k)$, $X2(k)$ and $X3(k)$ are respectively the expected values of the following functions of loss, measured in units of the mean loss:

$$XI_{(0, k]}(X)$$

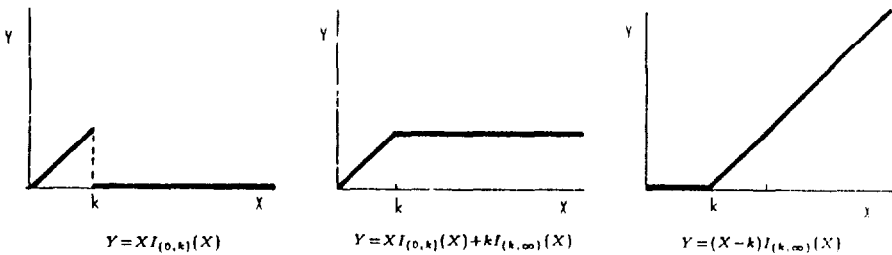
$$XI_{(0, k]}(X) + kI_{(k, \infty)}(X)$$

and

$$(X - k)I_{(k, \infty)}(X).$$

As functions of X , they are shown graphically in Figure 1. These quantities are more closely related to the loss and are more easily understood in their random variable forms. Most of the relations treated by LaRose [2] can be generalized to random variable versions; some of them are presented in the rest of this paper. A glance through Figures 1-4 shows visually the underlying similarity of many quantities derived from the loss function. Some quantities known by different names are the same function of the loss, and some bear simple relationships to others.

FIGURE 1
SOME FUNCTIONS OF THE LOSS



$$Y = XI_{(0, k]}(X)$$

$$Y = XI_{(0, k]}(X) + kI_{(k, \infty)}(X)$$

$$Y = (X - k)I_{(k, \infty)}(X)$$

3. EXCESS OF LOSS COVERAGE

An excess of loss coverage pays the amount of loss in excess of r for losses exceeding r but not greater than s , and the amount $j = s - r$ for losses exceeding s :

$$h(X;r,j) = \begin{cases} X - r & \text{if } r < X \leq s, \\ j & \text{if } s < X. \end{cases}$$

In terms of the indicator function

$$h(X;r,j) = (X - r)I_{(r,s]}(X) + jI_{(s,\infty)}(X).$$

Miccolis [3] shows that

$$h(X;r,j) = g(X;s) - g(X;r),$$

where $r + j = s$. As an example of algebraic manipulation with the indicator function representation, we derive this result as follows:

$$\begin{aligned} g(X;s) - g(X;r) &= XI_{(0,s]}(X) + sI_{(s,\infty)}(X) - \{XI_{(0,r]}(X) + rI_{(r,\infty)}(X)\} \\ &= X\{I_{(0,s]}(X) - I_{(0,r]}(X)\} + sI_{(s,\infty)}(X) \\ &\quad - rI_{(r,s]}(X) - rI_{(s,\infty)}(X) \\ &= XI_{(r,s]}(X) - rI_{(r,s]}(X) + sI_{(s,\infty)}(X) - rI_{(s,\infty)}(X) \\ &= (X - r)I_{(r,s]}(X) + (s - r)I_{(s,\infty)}(X). \end{aligned}$$

Hence

$$g(X;s) - g(X;r) = h(X;r,j).$$

(See equations 10–11 of Miccolis [3]). Miccolis derives a result on the expectation of $h^2(X;r,j)$ (see his equation 13), which can also be conveniently derived as follows:

$$\begin{aligned} g^2(X;s) &= \{g(X;r) + h(X;r,j)\}^2 \\ &= g^2(X;r) + h^2(X;r,j) + 2g(X;r)h(X;r,j). \end{aligned}$$

But

$$\begin{aligned}
 g(X;r)h(X;r,j) &= \{XI_{(0,r]}(X) + rI_{(r,\infty)}(X)\} \{(X-r)I_{(r,s]}(X) + (s-r)I_{(s,\infty)}(X)\} \\
 &= r\{(X-r)I_{(r,s]}(X) + (s-r)I_{(s,\infty)}(X)\} \\
 &= rh(X;r,j).
 \end{aligned}$$

(This result, in terms of random variables, is also given by Miccolis [3] in his equation 39.) We have the random variable version of his equation 13:

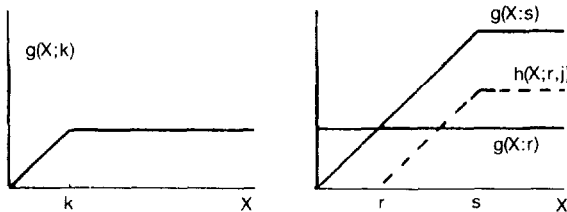
$$h^2(X;r,j) = g^2(X;s) - g^2(X;r) - 2rh(X;r,j).$$

In statistics the calculus of expectations is made easier by manipulating the random variables or their functions, rather than dealing with integrals directly. For example,

$$\begin{aligned}
 \text{Var} [XI_{(0,k]}(X) + kI_{(k,\infty)}(X)] &= \text{Var} [XI_{(0,k]}(X)] + \text{Var} [kI_{(k,\infty)}(X)] \\
 &\quad + 2\text{Cov} [XI_{(0,k]}(X), kI_{(k,\infty)}(X)] \\
 &= \text{Var} [XI_{(0,k]}(X)] + \text{Var} [kI_{(k,\infty)}(X)] \\
 &\quad - 2kE \{(XI_{(0,k]}(X)) E \{I_{(k,\infty)}(X)\}\},
 \end{aligned}$$

since in the covariance the cross-product is zero. It is easier to perform manipulations such as this than to work with integrals. Figure 2 shows some relations between these quantities, which are treated as functions of the loss X .

FIGURE 2
BASIC LIMITS AND EXCESS OF LOSS COVERAGES



An Example on Moments

As an example using the expectation operator, consider first the h^{th} moment of $g(X;k)$. We have

$$E \{g^h(X;k)\} = E \{[XI_{(0,k)}(X) + kI_{(k,\infty)}(X)]^h\}.$$

Expanding the power on the left hand side we see easily that all the cross terms are zero and so the h^{th} moment of $g(X;k)$ is

$$E \{g^h(X;k)\} = E \{X^h I_{(0,k)}(X)\} + k^h [1 - F(k)].$$

Now consider the h^{th} central moment of $g(X;k)$. We have the general result

$$\mu_h[g(X;k)] = \sum_{j=0}^{h-2} \binom{h}{j} (-1)^j \mu'_{h-j} \mu^j - (-1)^h (h-1) \mu^h,$$

where μ'_j is the h^{th} moment of $g(X;k)$ and $\mu = \mu'_1$ is its mean. We may then make use of our result for the moments of $g(X;k)$ to obtain the h^{th} central moment of $g(X;k)$

$$\mu_h[g(X;k)] = \sum_{j=0}^{h-2} \binom{h}{j} (-1)^j [\alpha_{h-j} + \beta_{h-j}] \mu^j - (-1)^h (h-1) \mu^h,$$

where

$$\alpha_j = E \{[X^j I_{(0,k)}(X)]\}, \beta_j = k^j [1 - F(k)] \text{ and} \\ \mu = E \{X I_{(0,k)}(X)\} + k [1 - F(k)].$$

4. TREND

First we derive a result which will considerably simplify and clarify the treatment of trend effect. Let y be a monotone function of x . To fix the idea, we assume the function to be increasing in x :

$$y = \alpha(x),$$

so that the transformation is invertible:

$$x = \alpha^{-1}(y).$$

Then

$$\begin{aligned} I_{(a, b]}(y) &= I_{(a, b]}(\alpha(x)) \\ &= I_{(\alpha^{-1}(a), \alpha^{-1}(b)]}(x) \end{aligned} \quad (1)$$

because the following statements are equivalent:

$$\begin{aligned} I_{(a, b]}(y) &= 1, \\ y &\in (a, b], \\ \alpha^{-1}(y) &\in (\alpha^{-1}(a), \alpha^{-1}(b)], \\ x &\in (\alpha^{-1}(a), \alpha^{-1}(b)], \\ I_{(\alpha^{-1}(a), \alpha^{-1}(b)]}(x) &= 1; \end{aligned}$$

and similarly whenever one of the indicator functions assumes the value 0, the other does also. The result is also true for a decreasing function, in which case the terminal points of the intervals are reversed.

Consider a loss X being subject to inflation. Suppose that at a future time point the loss becomes

$$Y = \alpha(X)$$

with $\alpha(\cdot)$ increasing. Then the basic limits loss becomes

$$g(\alpha(X); k) = \alpha(X)I_{(0, k]}(\alpha(X)) + kI_{(k, \infty)}(\alpha(X)).$$

Using (1) we have

$$g(\alpha(X); k) = \alpha(X)I_{(0, \alpha^{-1}(k)]}(X) + kI_{(\alpha^{-1}(k), \infty)}(X) \quad (2)$$

This can be rewritten in the form of a rescaled g function:

$$\begin{aligned} g(\alpha(X); k) &= \frac{k}{\alpha^{-1}(k)} \left\{ \frac{\alpha^{-1}(k)\alpha(X)}{k} I_{(0, \alpha^{-1}(k)]}(X) + \alpha^{-1}(k)I_{(\alpha^{-1}(k), \infty)}(X) \right\} \\ &= \{k/\alpha^{-1}(k)\} g\{\alpha^{-1}(k)\alpha(X)/k; \alpha^{-1}(k)\}. \end{aligned}$$

In this representation we have a rescaled g function of the form $g(w(X); b)$ which takes the value $w(X) = \alpha^{-1}(k)\alpha(X)/k$ over the interval $(0, b]$ and the value $b = \alpha^{-1}(k)$ over the interval (b, ∞) , with $w(b) = b$.

It is easily verified that when X takes the value $\alpha^{-1}(k)$, the first argument of the g function above also takes the value $\alpha^{-1}(k)$. The manipulation of random variables in this manner is a more general method of treating the trend effect and could be of practical use if the assumption of uniform inflation rate over the range of the size of loss is too serious a deviation from reality. Similarly, the effect of inflation on an excess of loss coverage can be represented as

$$h(\alpha(X); r, j) = [(\alpha(X) - r) I_{(\alpha^{-1}(r), \alpha^{-1}(s))}(X) + (s - r) I_{(\alpha^{-1}(s), \infty)}(X)];$$

or alternatively described in the form of a rescaled h function:

$$\begin{aligned} h(\alpha(X); r, j) &= c' h(X'; r', j') \\ &= \frac{s - r}{\alpha^{-1}(s) - \alpha^{-1}(r)} \left\{ \frac{\alpha^{-1}(s) - \alpha^{-1}(r)}{s - r} [(\alpha(X) - r) I_{(\alpha^{-1}(r), \alpha^{-1}(s))}(X) \right. \\ &\quad \left. + [\alpha^{-1}(s) - \alpha^{-1}(r)] I_{(\alpha^{-1}(s), \infty)}(X) \right\} \\ &= c' \{ (X' - r') I_{(r', s)}(X') + (s' - r') I_{(s', \infty)}(X') \}, \end{aligned}$$

where

$$c' = \frac{s - r}{\alpha^{-1}(s) - \alpha^{-1}(r)},$$

$$X' = \alpha(X)/c',$$

$$r' = r/c', \quad s' = s/c',$$

$$j' = s' - r' = (s - r)/c'$$

In the rescaled form the function $h(X'; r', j')$, where X' is a function of X , takes the value $X' - r'$ in the interval $(X' = r', X' = s')$ and the value j' in the interval $(X' = s', \infty)$, with $X' - r' = 0$ when $X' = r'$ and $X' - r' = j'$ when $X' = s'$. When $\alpha(\cdot)$ is the identity function, the original definition of h is recovered. The quantity of interest is the expectation of the h function. In general the functional form would not be easy to obtain, but the numerical computation should not be much more difficult than the computation of the untransformed h function.

If the trend function $\alpha(\cdot)$ is not monotone, then it can be broken up into pieces each of which forms a monotone function. The above analysis can then be carried out piecemeal. It is more usual to assume that the inflation rate is uniform over the loss size:

$$\alpha(X) = aX.$$

Venter [8] gives an extensive treatment for this case. It can be shown by straightforward substitution in the rescaled versions of the formulas above that

$$g(aX; k) = ag(X; k/a),$$

and

$$h(aX; r, j) = ah(X; r/a, j/a).$$

Formulas involving trend could cause difficulties because of the lack of suitable tools for handling the transformed loss payment resulting from inflation. Bickerstaff [1] describes a model for automobile physical damage loss in which there is a deductible D representable by the random variable

$$XI_{(0, D]}(X) + DI_{(D, \infty)}(X)$$

and an upper limitation of loss payment L such that the reduction in loss payment can be represented by the random variable

$$(X - L)I_{(L, \infty)}(X) = XI_{(L, \infty)}(X) - LI_{(L, \infty)}(X).$$

The total reduction due to deductible and limitation in year 1 is then

$$XI_{(0, D]}(X) + DI_{(D, \infty)}(X) + XI_{(L, \infty)}(X) - LI_{(L, \infty)}(X),$$

with $D < L$. The loss payment limitation L is subject to a discount at an annual rate of $1 - d$ and the loss incurred is subject to inflation at an annual rate of r , simultaneously, so that in the n^{th} year X becomes $X(1 + r)^{n-1}$ and L becomes Ld^{n-1} . Consequently the total reduction becomes

$$\begin{aligned} & (1 + r)^{n-1}XI_{(0, D]}((1 + r)^{n-1}X) + DI_{(D, \infty)}((1 + r)^{n-1}X) \\ & + (1 + r)^{n-1}XI_{(Ld^{n-1}, \infty)}((1 + r)^{n-1}X) - Ld^{n-1}I_{(Ld^{n-1}, \infty)}((1 + r)^{n-1}X) \end{aligned}$$

which is obtained by simply applying the appropriate factors to X and D . Using equation (1), we can reduce this to

$$(1+r)^{n-1}XI_{(0, D(1+r)^{1-n})}(X) + DI_{(D(1+r)^{1-n}, \infty)}(X) \\ + (1+r)^{n-1}XI_{(Ld^{n-1}(1+r)^{1-n}, \infty)}(X) - Ld^{n-1}I_{(Ld^{n-1}(1+r)^{1-n}, \infty)}(X)$$

Thus with the results developed earlier in this section the effect of trend can be handled in a fairly formal way. Taking expectations leads to the pure premium version of Bickerstaff's [1] formula. The original formula given by Bickerstaff is incorrect; Philbrick [5] corrects the error.

A Numerical Example

We now give an example in the calculation of trend when the inflation rate varies with the amount of loss. There does not seem to be any theory on the specific functional form for rates of inflation which vary over the range of the amount of loss. In any case such rates cannot be precisely determined in practice. One way is to break up the range of loss value into sub-intervals and for each of the sub-intervals to approximate the rate of inflation by a linear function. This would often lead to mathematically tractable solutions and is quite satisfactory for handling practical problems.

We assume that the inflation rate i of the loss X increases with the value of X . Specifically, we assume that i starts at 10%, increases linearly until it reaches 20% at $X = 20,000$, and thereafter remains at 20%. Thus $i = i(X)$ can be described by the formula

$$i(X) = (0.10 + 0.10X/20000)I_{(0, 20000)}(X) + 0.20I_{(20000, \infty)}(X).$$

The loss after inflation is then described by the formula

$$\alpha(X) = [1 + i(X)]X \\ = (1.10X + 0.10X^2/20000)I_{(0, 20000)}(X) \\ + 1.20XI_{(20000, \infty)}(X). \quad (3)$$

The overall effect of inflation over the full range of loss can be described by the average rate of inflation $E\{\alpha(X)\}/E\{X\} - 1$, where

$$E\{\alpha(X)\} = E\{(1.10X + 0.10X^2/20000)I_{(0, 20000)}(X)\} + 1.20E\{XI_{(20000, \infty)}(X)\}. \quad (4)$$

In this example we have for the purpose of illustration divided the range of loss into only two sub-intervals; there is actually no difficulty in dividing the range into any finite number of sub-intervals. In this approach the main task in the calculation of the effect of inflation is the evaluation of the incomplete first and second moments of the distribution of loss, as is clear from the preceding formula. In fact, formula (4) can be rewritten as follows:

$$E\{\alpha(X)\} = 1.10E_{(s)}\{X\} + (0.10/s)E_{(s)}\{X^2\} + 1.20[E\{X\} - E_{(s)}\{X\}]$$

where we have written s for 20,000 and where

$$E_{(s)}\{X^k\} = \int_0^s x^k dF(x)$$

is the incomplete k^{th} moment of X up to s . For most distributions commonly used to model loss data, explicit formulas for such moments are available.

Suppose we are interested in the effect of inflation on the basic limits loss $g(X;k)$ with $k = 10,000$. From (2) the indemnity after inflation would be

$$\begin{aligned} g(\alpha(X);k) &= \alpha(X)I_{(0, \alpha^{-1}(k))}(X) + kI_{(\alpha^{-1}(k), \infty)}(X) \\ &= [1.10X + (0.10/20,000)X^2]I_{(0, \alpha^{-1}(k))}(X) \\ &\quad + 10,000I_{(\alpha^{-1}(k), \infty)}(X). \end{aligned}$$

The value of $\alpha^{-1}(k)$ is determined by the quadratic equation

$$1.10x + (0.10/20,000)x^2 = k = 10,000,$$

and the solution is easily found to be $x = \alpha^{-1}(10,000) = 8,743$.

Thus the expected value of the basic limits indemnity would be

$$\begin{aligned} E\{g(\alpha(X);k)\} &= E\{1.10X + (0.10/20,000)X^2\}I_{(0, 8743]}(X) \\ &\quad + 10,000[1 - F(8,743)]. \end{aligned} \quad (5)$$

Suppose the limit is $k = 25,000$; the solution would be somewhat different. A close look at equation (3) reveals that the inverse $x = \alpha^{-1}(k)$ is given by $k/1.20 = 20,833$. The expected value of the indemnity under this contract would then be

$$\begin{aligned} E\{g(\alpha(X);k)\} &= E(1.10X + (0.10/20,000)X^2)I_{(0, 20000]}(X) \\ &\quad + 1.2E\{XI_{(20000, 20833]}(X)\} \\ &\quad + 25,000[1 - F(20,833)], \end{aligned} \quad (6)$$

remembering the change in functional form at $X = 20,000$. The actual calculation can then be carried through by evaluation of the distribution function and the appropriate incomplete moments.

Now let us take the specific functional form for the loss distribution to be lognormal with parameters

$$\mu = 7.6 \text{ and } \sigma = 1.8$$

so that the mean loss is

$$E\{X\} = 10,097 \text{ with CV} = 4.953,$$

where CV stands for the coefficient of variation (standard deviation / mean). For the lognormal the k^{th} moment is

$$E\{X^k\} = \exp[k\mu + (1/2)k^2\sigma^2]$$

and the incomplete k^{th} moment is

$$E_{(x)}\{X^k\} = E\{X^k\}\Phi[(\ln x - \mu)/\sigma - k\sigma]$$

where $\Phi(\cdot)$ denotes the distribution function of the standard normal variable. Table 1 shows the incomplete moments with orders shown in column (1) up to the various values shown in the first row. Order zero means the value of the distribution function.

TABLE 1

INCOMPLETE MOMENTS

(1) <u>k</u>	(2) <u>20,000</u>	(3) <u>8743</u>	(4) <u>10,000</u>	(5) <u>20,833</u>	(6) <u>25,000</u>
0	0.8997	0.7939	0.8145	0.9036	0.9198
1	3,044	1,651	1,844	3,124	3,493
2	26,455,454	7,075,407	8,880,632	28,093,176	36,534,773

First we will calculate the overall inflation; formula (4) and the numbers in column (2) of Table 1 give

$$1.10 \times 3044 + .10 \times 26,455,454 / 20,000 \\ + 1.20 \times 20,000 \times (10,097 - 3,044) = 11,945,$$

corresponding to an overall rate of $11,945/10,097 - 1 = 18.3\%$. Next consider the basic limits indemnity with a limit at 10,000. Noting that $\alpha^{-1}(k)$ in formula (5) has the value 8743 and using the numbers in column (3), we have the expected basic limits indemnity

$$1.10 \times 1651 + .10 \times 7,075,407 / 20,000 \\ + 10,000 \times (1 - .7939) = 3,912.$$

From column (4) we can easily obtain the limited indemnity without inflation:

$$1,844 + 10,000(1 - .8145) = 3,699$$

so that the effective inflation rate is $3,912/3,699 - 1 = 5.76\%$.

Similarly, we can calculate the inflation rate for the basic limits indemnity with a limit of 25,000. From formula (6) and the numbers in columns (2) and (5) we find this to be

$$1.10 \times 3044 + .10 \times 26,455,454 / 20,000 \\ + 1.2 \times (3,124 - 3,044) + 25,000 \times (1 - .9036) = 5,987.$$

Again, we can calculate the expected indemnity without inflation for this case with the numbers in column (6):

$$3,493 + 2,500 \times (1 - .9198) = 5,498$$

so that the effect of inflation is to increase the expected indemnity by $5,987/5,498 - 1 = 8.89\%$.

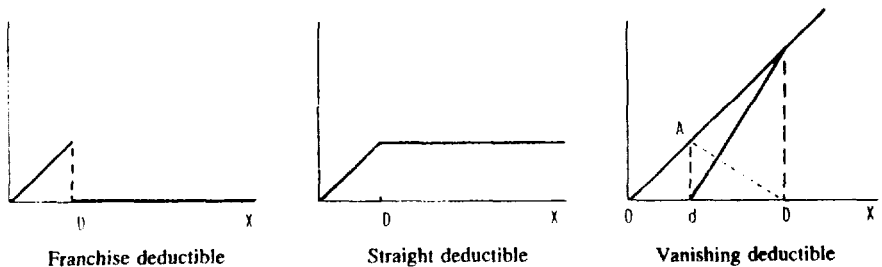
5. DEDUCTIBLES

The common type of deductible, called straight deductible, has the simple representation

$$XI_{(0, D]}(X) + DI_{(D, \infty)}(X).$$

A franchise deductible is represented as $XI_{(0, D]}(X)$. These are depicted as functions of the loss X in Figure 3. A more complicated form of deductible is the vanishing deductible, which equals the loss up to the amount d , but thereafter reducing linearly to 0 when the loss becomes $D > d$ (Snader [7]); see Figure 3 for a pictorial description.

FIGURE 3
DEDUCTIBLES



It is easier to describe the indemnity after the deductible. The geometry in Figure 3 shows easily that the equation for the indemnity is

$$y = \frac{D}{D-d}(x-d)$$

over the range $(d, D]$. Thus the indemnity can be written as

$$Y = \frac{D}{D-d}(X-d)I_{(d, D]}(X) + XI_{(D, \infty)}(X).$$

The deductible itself can be found by taking the difference:

$$X - Y = XI_{(0, d)}(X) + \frac{d}{D-d}(D-X)I_{(d, D]}(X).$$

The expectation of the deductible can then be obtained as

$$\begin{aligned} E\{X - Y\} &= E\{XI_{(0, d)}(X)\} + \frac{d}{D-d} E\{(D-X)I_{(d, D]}(X)\} \\ &= \int_0^d x dF(x) + \frac{d}{D-d} \int_d^D (D-x) dF(x). \end{aligned}$$

6. RETROSPECTIVE RATING

In retrospective rating, the charge over rE , where E stands for the expected loss, can be represented as a random variable:

$$\Phi(r)E = (X - rE)I_{(rE, \infty)}(X).$$

$\Phi(r)$ may be interpreted as the excess of the loss X over rE , measured in units of E , and

$$E\{\Phi(r)\} = \phi(r)$$

where $E\{\}$ means expectation and $\phi(r)$ has the usual meaning of charge over rE as an average. Similarly, we represent the random variable savings as

$$\Psi(r)E = (rE - X)I_{(0, rE]}(X)$$

with

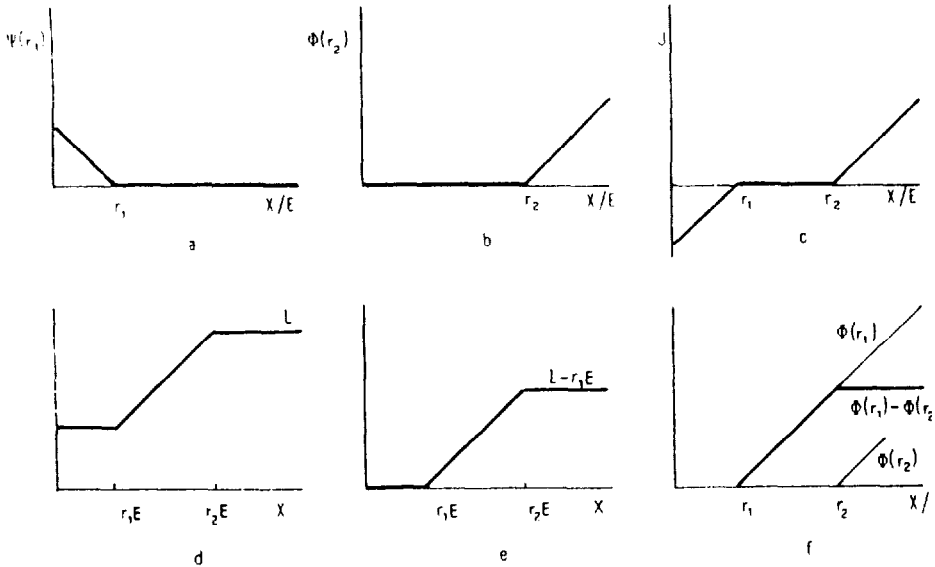
$$E\{\Psi(r)\} = \psi(r).$$

The graphs of $\Phi(r)$ and $\Psi(r)$ as functions of X/E are shown in Figures 4a and 4b. The loss portion of the limited pure premium paid by the insured, L , is a random variable under the insurance contract:

$$L = r_1 EI_{(0, r_1 E]}(X) + XI_{(r_1 E, r_2 E]}(X) + r_2 EI_{(r_2 E, \infty)}(X). \quad (7)$$

The insured pays the minimum premium $r_1 E$ if the actual loss incurred is not more than $r_1 E$, the actual loss if it is greater than $r_1 E$ but not greater than $r_2 E$, and the maximum premium $r_2 E$ if the actual loss exceeds $r_2 E$, as far as the loss portion of the premium is concerned. The graph of L as a function of X is shown in Figure 4d.

FIGURE 4
RETROSPECTIVE RATING



By adding and subtracting $XI_{(0, r_1E]}(X)$ and $XI_{(r_2E, \infty)}(X)$ in the formula (7) for L we have

$$\begin{aligned} L &= r_1EI_{(0, r_1E]}(X) - XI_{(0, r_1E]}(X) + XI_{(0, r_1E]}(X) + XI_{(r_1E, r_2E]}(X) \\ &\quad + XI_{(r_2E, \infty)}(X) - XI_{(r_2E, \infty)}(X) + r_2EI_{(r_2E, \infty)}(X) \\ &= (r_1E - X)I_{(0, r_1E]}(X) + X - (X - r_2E)I_{(r_2E, \infty)}(X) \end{aligned}$$

Thus

$$\begin{aligned} L &= X - \{\Phi(r_2)E - \Psi(r_1)E\} \\ &= X - J \end{aligned}$$

where

$$J = \Phi(r_2)E - \Psi(r_1)E$$

is the random variable version of the net charge I as defined, for example, in Skurnick [6] and, of course, $E\{J\} = I$. Note that J is equal to $-(r_1E - X)$ if the loss X is less than r_1E , zero if the loss X is between r_1E and r_2E , and $X - r_2E$ if the loss X exceeds r_2E . Its graph as a function of X is shown in Figure 4c.

The following identity is interesting:

$$\begin{aligned} \Phi(r)E - \Psi(r)E &= (X - rE)I_{(rE, \infty)}(X) - (rE - X)I_{(0, rE]}(X) \\ &= X - rE, \end{aligned}$$

being the random variable version of the well-known identity $\phi(r) - \psi(r) = 1 - r$. The graph of $\Phi(r)E - \Psi(r)E$ as a function of X is simply a 45-degree line through the point $(rE, 0)$.

The following equality is easily derived from equation (7):

$$\begin{aligned} L - r_1E &= r_1EI_{(0, r_1E]}(X) + XI_{(r_1E, r_2E]}(X) + r_2EI_{(r_2E, \infty)}(X) - r_1E \\ &= (X - r_1E)I_{(r_1E, r_2E]}(X) + (r_2E - r_1E)I_{(r_2E, \infty)}(X). \quad (8) \end{aligned}$$

This equality is illustrated in Figure 4e, and its relationship to L is clearly visualized by comparison to Figure 4d. Whereas

$$\begin{aligned}\Phi(r_1)E - \Phi(r_2)E &= (X - r_1E)I_{(r_1E, \infty)}(X) - (X - r_2E)I_{(r_2E, \infty)}(X) \\ &= (X - r_1E)I_{(r_1E, r_2E]}(X) + (X - r_1E)I_{(r_2E, \infty)}(X) \\ &\quad - (X - r_2E)I_{(r_2E, \infty)}(X) \\ &= (X - r_1E)I_{(r_1E, r_2E]}(X) \\ &\quad + (r_2E - r_1E)I_{(r_2E, \infty)}(X).\end{aligned}\tag{9}$$

See Figure 4f for an illustration. Equations (8) and (9) show that

$$\Phi(r_1)E - \phi(r_2)E = L - r_1E,\tag{10}$$

and both are the excess of loss function $h(X; r_1E, r_2E - r_1E)$ as described in Section 2. Let BP be the basic premium of the retrospective plan. Adding and subtracting this on the right-hand side of (10) yields

$$\Phi(r_1) - \Phi(r_2) = [\{BP + CL\} - \{BP + Cr_1E\}]/CE,\tag{11}$$

where C is the loss conversion factor to be applied to the loss to obtain the premium. Equation (11) is useful in determining the exact entry ratios as well as the minimum and maximum premiums in a retrospective rating plan. Noting that

$$E\{BP + CL\} = P(1 - D),$$

which is the premium after adjustment for expense gradation D , and

$$E\{BP + r_1CE\} = H,$$

where H is the minimum premium, we have by taking expectations on both sides of (11) the familiar identity (Skurnick [6])

$$\phi(r_1)E - \phi(r_2)E = (P - PD - H)/CE.$$

7. CONCLUSION

An alternative method of representing a function assuming different functional forms over the range of its argument is by the Heaviside or delta function, which is defined as

$$H(x) = \begin{cases} 1 & \text{if } 0 \leq X \\ 0 & \text{otherwise.} \end{cases}$$

Thus the function

$$g(X;k) = \begin{cases} X & 0 < X \leq K \\ k & X > k. \end{cases}$$

can be represented in terms of the Heaviside function as

$$\begin{aligned} g(X;k) &= XH(X) - XH(X - k) + kH(X - k) \\ &= XH(X) - (X - k)H(X - k). \end{aligned} \tag{12}$$

The Heaviside function obviates the explicit use of the set and so is more parsimonious in notation. Most people, however, would take a relatively long time to picture the shape of the function represented by (12). While the indicator function representation visually shows the sets of points where the g function assumes different forms, it is not so with the Heaviside function. Thus, although at times clumsy in form, the indicator function representation is preferred here.

In mathematics and, more generally, scientific work, given relations are to be made as general as possible. Relations between random variables are certainly more general than those derivable from these relations but pertaining to their expectations only. In actuarial work, the most important quantity related to a loss is the expected value. It is natural that much of the work on the topics described in this paper has focused on expected values. This paper has shown that it is often possible to express the relations in terms of the random variables, thus strengthening the existing mathematical results. The results are stronger in the sense that when a relation holds for random variables, it is true for each realization, whereas a relation for expected values holds only on the average.

Such an approach allows us to look at the results in another way. The stronger results give not only the expectation relationship, but also relationships pertaining to other characteristics of the indemnity distribution, such as higher order moments. Also, quite often the mathematics become simpler and easier to understand. In particular, the treatment of trend in this fashion is more effective than techniques hitherto available.

REFERENCES

- [1] D. R. Bickerstaff, "Automobile Collision Deductible and Repair Cost Groups: The Lognormal Model," *PCAS* LIX, 1972.
- [2] J. G. LaRose, "A Note on Loss Distributions," *PCAS* LXIX, 1982.
- [3] R. Miccolis, "On the Theory of Increased Limits and Excess of Loss Pricing," *PCAS* LXIV, 1977.
- [4] A. M. Mood, F. A. Graybill, and D. C. Boes, *Introduction to the Theory of Statistics* (third edition), McGraw-Hill, New York, 1974, p. 20.
- [5] S. W. Philbrick, Discussion on J. G. LaRose, "A Note on Loss Distributions," *PCAS* LXXI, 1984.
- [6] D. Skurnick, "The California Table L," *PCAS* LXI, 1974.
- [7] R. H. Snader, "Fundamentals of Individual Risk Rating and Related Topics," Casualty Actuarial Society Study Note, Part III.
- [8] G. G. Venter, "Scale Adjustments to Excess Expected Losses," *PCAS* LXIX, 1982.