AUTOMOBILE MERIT RATING and INVERSE PROBABILITIES

BY

LESTER B. DROPKIN

Introduction

The previous paper by this writer[1] started from the fact that the negative binomial distribution provides an analytical expression for risk distribution by number of accidents. The expressions used in that paper did not explicitly introduce the time factor – e.g. the 3-year experience of the California study[2] was treated as a unit entity.

In the discussion of my paper by Mr. R. A. Bailey[3] the negative binomial was utilized to analytically develop an expression for the average frequency of a sub-group having n accident-free years.

It is the purpose of the present paper to set forth (a) the negative binomial distribution with the time element explicitly introduced and (b) a general expression for the probability of x accidents in subsequent years, knowing that a specified number of accidents have occurred in a given time period. This involves the classic problem of inverse probability and its solution is afforded by recourse to Bayes Theorem.

This general expression of risk distribution should find particular application to those automobile merit rating systems which determine credits and debits on the basis of fixed experience periods. However, the required expressions for a system based on the number of accident-free years also fall out in a simple manner.

The Negative Binomial, N(x; t)

We assume, as before, that to each member of the population may be coordinated a measure, m, of inherent accident potential which remains constant for the individual throughout the period involved. Further, we assume that the distribution of m in the population is given by the two parameter function, T(m):

$$T(m) = a^r m^{r-1} e^{-am} / \Gamma(r) \tag{1}$$

This function is independent of the time, has a mean equal to r/a and a variance equal to r/a^2.

For a given time period t, the (forward) probability of the number of accidents equaling x where $x=0, 1, \ldots$ is denoted by $N(x; t)$ and is given by:

$$N(x; t) = \int_0^\infty P(x; mt) T(m) dm \tag{2}$$

[1] "Some Considerations on Automobile Rating Systems Utilizing Individual Driving Records," CAS XLVI, p. 165.

[2] See Harwayne, F., "Merit Rating in Private Passenger Automobile Liability Insurance and the California Driver Record Study," CAS XLVI, p. 189, for a description of the study and its results.

[3] Page 152.

where $P(x; mt)$ is the Poisson frequency function:

$$(3) \qquad P(x; mt) = (mt)^x e^{-mt}/x!$$

Upon substituting (1) and (3) in (2), integrating and simplifying we have that[4]

$$(4) \qquad N(x; t) = \left(\frac{a}{a+t}\right)^r \left(\frac{t}{a+t}\right)^x (-1)^x \binom{-r}{x}$$

This distribution, the negative binomial, depends upon t, has a mean equal to rt/a and a variance[5] equal to $\left(\frac{rt}{a}\right)\left(\frac{a+t}{a}\right)$.

Inverse Probabilities and Bayes Theorem

It will be recalled that Bayes Theorem is properly applicable in the following kind of situation. Let us suppose that various mutually exclusive conditions, represented by A, B, C, . . . may exist. Also suppose that the probability of condition A existing is known and given by a, the probability of condition B by b, etc. Further, let us suppose that if condition A exists, then the probability of the happening of an event in which we are interested is known and given by E_a; that if condition B exists, the probability of the happening of this event is E_b; etc. The problem of inverse probabilities arises when we know that the event has occurred but we do not know what condition caused it. We ask, for example, what is the probability that this event arose out of condition A? The probability of condition A existing, knowing that the event has happened is denoted by Prob. (A/event). Bayes Theorem says that

$$(5) \qquad \text{Prob. (A/event)} = \frac{a \cdot E_a}{a \cdot E_a + b \cdot E_b + \ldots}$$

$$\text{Similarly, Prob. (B/event)} = \frac{b \cdot E_b}{a \cdot E_a + b \cdot E_b + \ldots}, \text{ etc.}$$

Inverse Probabilities and Automobile Merit Rating

In the model which we have been utilizing, we have supposed the existence of various mutually exclusive conditions. That is, for each individual we have assumed the existence of a particular measure of inherent accident potential, which has a probability given by $T(m)$. These values of $T(m)$ correspond to the a's and b's mentioned above. Now the probability of a given number of accidents occurring, say c accidents, in a time period s, under the condition of an inherent accident potential of m, is given by $P(c; ms)$.

$$(6) \qquad P(c; ms) = (ms)^c e^{-ms}/c!$$

These values of $P(c; ms)$ correspond to the E_a's and the E_b's above.

[4]See Appendix A.

[5]Means and variances are neatly determined by the use of moment generating functions. See Simon, L. J., "The Negative Binomial and Poisson Distributions Compared." Page 20.

Now suppose that we have observed the event: c accidents in s years. We ask the question, what is the probability that this event arose out of a particular measure, m? By Bayes Theorem

$$(7) \qquad T(m/c, s) = \frac{P(c; ms) \cdot T(m)}{\int_0^\infty P(c; ms) \cdot T(m)\,dm}$$

$$(8) \qquad T(m/c, s) = \frac{P(c; ms) \cdot T(m)}{N(c; s)}$$

When the operations indicated by (8) are carried out, we have that

$$(9) \qquad T(m/c,s) = (a+s)^{r+c}\, m^{r+c-1}\, e^{-m(a+s)} / \Gamma(c+r)$$

This function has a mean equal to $(r+c)/(a+s)$ and a variance equal to $(r+c)/(a+s)^2$.

This function, $T(m/c,s)$, is the distribution of inherent accident potential for a particular sub-group, viz. those who have been observed to have c accidents in s years. But we are now in a position to determine the (forward) probability of x accidents in the next t years for this sub-group. Denoting this probability by $N(x; t/c; s)$ we have that

$$(10) \qquad N(x; t/c; s) = \int_0^\infty P(x; mt) \cdot T(m/c, s)\,dm$$

$$(11) \qquad = \int_0^\infty \frac{P(x; mt) \cdot P(c; ms) \cdot T(m)\,dm}{N(c; s)}$$

Upon substituting, integrating and simplifying we have that[6]

$$(12) \quad N(x; t/c; s) = \left(\frac{a+s}{t+a+s}\right)^{r+c} \left(\frac{t}{t+a+s}\right)^{x} (-1)^{x} \binom{-(r+c)}{x}$$

This distribution has a mean equal to $\dfrac{t(r+c)}{a+s}$ and a variance equal to

$$\frac{t(r+c)\,(a+s+t)}{(a+s)^2}.$$

[6]See Appendix B.

APPENDIX A

$$N(x;t)=\int_0^\infty P(x;mt)T(m)dm$$

$$=\frac{a^r t^x}{x!\Gamma(r)}\int_0^\infty m^{x+r-1}e^{-m(a+t)}dm$$

$$=\frac{a^r\,t^x}{x!\,\Gamma(r)}\,\frac{\Gamma(x+r)}{(a+t)^{x+r}}=\left(\frac{a}{a+t}\right)^r\left(\frac{t}{a+t}\right)^x\frac{\Gamma(x+r)}{x!\,\Gamma(r)}$$

Since $\Gamma(x+r)=(x+r-1)(x+r-2)\ldots r\Gamma(r)$,

$$\frac{\Gamma(x+r)}{x!\,\Gamma(r)}=\frac{r(r+1)\ldots(r+x-1)}{x!}$$

$$=(-1)^x\binom{-r}{x}, \text{ and}$$

$$N(x;t)=\left(\frac{a}{a+t}\right)^r\left(\frac{t}{a+t}\right)^x(-1)^x\binom{-r}{x}$$

APPENDIX B

$$N(x;t/c;s)=\int_0^\infty P(x;mt)\cdot T(m/c,s)dm$$

$$=\int_0^\infty\frac{P(x;mt)\cdot P(c;ms)\cdot T(m)}{N(c;s)}dm$$

$$=\int_0^\infty\frac{(mt)^x e^{-mt}}{x!}\,\frac{(ms)^c e^{-ms}}{c!}\,\frac{a^r m^{r-1}e^{-am}}{\Gamma(r)}\cdot\frac{1}{N(c;s)}dm$$

$$=\frac{t^x a^r s^c}{x!\Gamma(r)c!N(c;s)}\int_0^\infty m^{x+r-1+c}e^{-m(s+a+t)}dm$$

$$=\frac{t^x a^r s^c}{x!\Gamma(r)c!N(c;s)}\cdot\frac{\Gamma(x+r+c)}{(s+a+t)^{x+r+c}}$$

Since $N(c;s)=\left(\frac{a}{a+s}\right)^r\left(\frac{-s}{a+s}\right)^c\binom{-r}{c}=\frac{a^r\,s^c\,\Gamma(c+r)}{c!\,\Gamma(r)(a+s)^{c+r}}$

$$N(x;t/c;s)=\frac{t^x\,(a+s)^{c+r}}{(s+a+t)^{x+c+r}}\cdot\frac{\Gamma(x+r+c)}{x!\,\Gamma(c+r)}$$

$$=\left(\frac{a+s}{s+a+t}\right)^{r+c}\left(\frac{t}{s+a+t}\right)^x(-1)^x\binom{-(r+c)}{x}$$