DISCUSSION BY STEPHEN W. PHILBRICK

Mr. Dropkin's paper consists of two parts. The first is a discussion of the "importance of the negative binomial distribution as a valuable instrument in its own right." Second, this tool is used to comment on the use of the number of traffic violations to "split up the total heterogeneous group into homogeneous groups."

The author succeeds admirably in his first endeavor. A concise explanation of the rationale for the use of the negative binomial distribution is given. The arguments are intuitively appealing, since the choice of a Poisson distribution for an individual risk is desirable, and the notion that the parameters for the individuals vary from person to person is certainly more reasonable than the assumption that all drivers have identical accident propensities. The author is also to be commended for the algorithm for the calculation of the probabilities of $N(x)$, which is much more convenient than evaluating the traditional formula.

The second section contains some problems. The author concludes, "the fact that the negative binomial fits the data for the total group indicates that there is a real spread, that is, a distribution, of the probability of having an accident." Unfortunately, this conclusion cannot be supported by this argument. To demonstrate this, I randomly sampled from a Poisson distribution and attempted to fit both a Poisson and a negative binomial to the sample data. A Poisson distribution would have absolutely no "spread" of the parameter since the parameter is a fixed constant. The distribution of the parameter should not be confused with the distribution of the number of accidents. As Dropkin correctly points out, "It is important to emphasize here that there are two distributions which enter into our considerations. On the one hand, there is the distribution of the probability of having an accident. On the other hand, there is the distribution of risks by number of accidents. If the first distribution is a constant, then the second is a Poisson."

The example in Table 1 shows the result of 10,000 trials from a Poisson distribution with the parameter of .274. The parameter value was selected to equal the mean of the group with two violations. The sample mean and variance are shown, as well as the values of a and r as calculated by Dropkin's formulas. Comparison of the actual results with the expected results of a Poisson indicates a reasonably good fit (as expected), which is further corroborated by calculating the chi-square statistic and noting that it is significant at the 5% level. Note, however, that the negative binomial provides an even better fit. This should not be completely surprising, since the negative binomial is a two-parameter distribution and, more importantly, the Poisson can be thought of as the "limit" of a negative binomial. (Let a and r go to infinity such that r/a remains constant, and the result is a Poisson with parameter r/a.) This concept is given added intuitive appeal by examining the formula for a; the denominator is $(\sigma^2 - M)$ whose expected value is zero. Hence, calculations of the parameter a for samples from a Poisson would be expected to produce large values, which is borne out by observation.

TABLE 1

FIXED VALUE OF M

Expected Mean	.274		
Number of Trials	10,000		
Sample Mean	.2749	Poisson Chi-square	9.75
Sample Variance	.2801		
Sample a	52.5622	Negative Binomial Chi-square	5.58
Sample r	14.4494		

Number of Claims	Actual	Expected Poisson	Expected Negative Binomial
0	7613	7603.32	7616.12
1	2061	2083.31	2054.58
2	296	285.41	296.31
3	24	26.07	30.33
4 or more	6	1.89	2.65

Therefore, we see that a good fit of a negative binomial does *not* imply a *real* spread of the parameter, since a good fit is expected when there is no spread.

It should not be inferred that I disagree that there is a real spread of the parameter. I merely disagree with his proof. Indeed, the sample variance of .193 is too much larger than the mean of .163 to be accounted for by process variance.[1]

[1] This could be shown mathematically, but my statement is made upon empirical observations. A sample of 95,000 trials from a Poisson distribution with a mean of .163 produced variances generally no more than .002 higher than the mean.

He next suggests that "the function of a segregating system is to split the total heterogeneous group into homogeneous groups." I generally agree with this except I would prefer to replace "homogeneous" with "more homogeneous."

He states, "If the system we are dealing with here accomplished this purpose totally, then the distributions by number of accidents of the individual groups should be describable by Poisson curves." This statement is too strong. He has hypothesized that the accident propensities are describable by a Type III curve, which is continuous. He proposed to partition this curve into six discrete groups and measure the results against a standard (the Poisson) which requires that each group have a single-valued accident propensity. This is clearly impossible with a discrete partitioning. I would prefer that he would test to see if the result were *closer* to a Poisson curve.

His test is to compare the sample variance to the binomial variance. I am at a loss as to the reasoning behind this. If the results were Poisson, I would expect the variance to be close to the mean, not to the binomial variance which is always less than the mean.

He then concludes, "since a Poisson distribution is not indicated for the distributions by number of accidents, a negative binomial is indicated." This statement does not follow at all. This statement is equivalent to the following reasoning: "I have shown that the total group is negative binomial. This means that the distribution of parameters, $T(m)$, is describable by a Type III curve. The segregating system can be thought of as assigning individuals, hence their particular parameter, to various groups. Define $T_i(m)$ as the resulting distribution of m for the i^{th} group. If the distribution of *accidents* for each group is Poisson, then the associated $T_i(m)$ is a constant. *If the distribution is not Poisson, then the associated* $T_i(m)$ *is Type III*." It should be clear that this is not true. Even if one accepts that the distribution of parameters of the total group is Type III, it is unreasonable to assume that the only possible partitions of $T(m)$ into $T_i(m)$ are either constants or Type III curves. This error is serious, since he uses it to draw conclusions about the overlap of parameter between groups.

He has made two errors of implication:

1. If the underlying $T_i(m)$ are not constants, they must be Type III. (This is equivalent to the statement that if the accident distributions are not Poisson, they must be negative binomial.)

   If we could analyze the actual distribution of accidents within each group and find that, indeed, it is closely fit by a negative binomial, then the above problem would be moot. But he still could not draw his conclusions. To see this, we have to examine the second (and most critical) error of implication.

2. If a distribution is closely fit by a negative binomial, then the distribution of parameters, $T(m)$, is closely fit by a Type III curve. Furthermore, the parameters of the Type III curve can be estimated from the mean and variance of the accident distribution.

This is basically a sensitivity question. How sensitive is the resulting distribution to the form of $T(m)$? How "close" to a Type III must $T(m)$ be to cause the accident distribution to be "close" to a negative binomial? The fact is that many reasonable forms of $T(m)$ other than a Type III curve will produce a distribution which is fit very well by a negative binomial. Table 2 shows the result of 10,000 trials from a Poisson distribution whose parameter is uniformly distributed between .194 and .354 (hence, has mean .274). Notice that the result is fit quite well by a negative binomial.

TABLE 2

M IS UNIFORMLY DISTRIBUTED OVER (.194, .354)

Expected Mean	.274		
Number of Trials	10,000		
Sample Mean	.2667	Poisson Chi-square	10.71
Sample Variance	.2776		
Sample a	24.5329	Negative Binomial Chi-square	.23
Sample r	6.5429		

Number of Claims	Actual	Expected Poisson	Expected Negative Binomial
0	7698	7603.32	7699.67
1	1978	2083.31	1973.08
2	287	285.41	291.44
3	33	26.07	32.50
4 or more	4	1.89	3.31

Table 3 is another example where the parameter could take on the value .184 or .364 with equal probability. Again, the negative binomial fits well.

TABLE 3

M HAS EQUAL PROBABILITY OF BEING .184 OR .364

Expected Mean	.274		
Number of Trials	10,000		
Sample Mean	.2690	Poisson Chi-square	4.57
Sample Variance	.2806		
Sample *a*	23.1120	Negative Binomial Chi-square	2.60
Sample *r*	6.2171		

Number of Claims	Actual	Expected Poisson	Expected Negative Binomial
0	767	760.33	768.48
1	203	208.33	198.15
2	24	28.54	29.65
3	6	2.80	3.72

These examples were not chosen arbitrarily. Note that in the example used in Table 2, there is no overlap as defined by Dropkin, i.e., no value of the parameter falls outside the mean of the neighboring groups. On the other hand, also in the example in Table 2, the possible values of the parameter are always outside the means of the neighboring groups. Hence, the conclusions he reaches concerning overlap are not well-founded.

Let me reemphasize: Although the form of the distribution of $T(m)$ needs to be Type III for the negative binomial to follow, a distribution of $T(m)$ which is significantly different from Type III will produce an accident distribution which can be fit very closely by a negative binomial. Hence, it is improper to conclude that a good fit of a negative binomial necessarily implies that the underlying $T(m)$ is Type III.

This result is certainly unfortunate, particularly with the recent furor over classifications. To my knowledge, the questions of overlap are currently unresolved, since the true accident propensities are unknown and only the resulting accident distributions are known. For a particular individual, the expected frequency is so low that process variance overpowers the information contained in the results. Dropkin's paper provides a novel approach to the problem. His approach, in brief, is to observe the distribution of accidents and, together with an assumed knowledge of the accident producing process, make inferences about the underlying distribution of accident propensities. *The concept is theoretically sound; unfortunately, the low sensitivity of the resulting distribution to the form of $T(m)$ makes it impossible to draw meaningful conclusions about* $T(m)$. The approach, however, should not be quickly discarded. Is there another way of looking at our data? Can we find some function of our data that is dependent on the form of the distribution of the accident propensities *and is highly sensitive to the form?* If so, then we could draw valid conclusions about accident propensities.

In conclusion, this paper has given an excellent discussion of the propensities of the negative binomial, and an interesting approach to the solution of a knotty problem, although this specific application of the approach was less than conclusive.