NON-RANDOM ACCIDENT DISTRIBUTIONS AND THE POISSON SERIES

BY

JOHN CARLETON

In recent years several papers have appeared in the *Proceedings* in which the Poisson formula has been used to obtain the theoretical distribution of accident frequencies in casualty insurance. In casualty terminology this formula with its required conditions may be expressed somewhat as follows:

> If accidents are distributed individually and collectively at random in a large exposure, the probability of exactly n accidents occurring in a portion of that exposure: i.e., in an individual risk, is given by
>
> $$P_n = \frac{c^n e^{-c}}{n!}$$
>
> where c is the accident expectation for the portion under consideration.

The requisite of individual randomness is that the timing of each accident must be independent of the timing of each other accident.

The requisite of collective randomness is that the accident frequency in any interval within the portion must be independent of the accident frequency in any other such interval not overlapping the first interval.

The formula is applicable even if the accident producing potentialities are changed from interval to interval within the risk period, provided these changes are not made in such a way as to violate either of the two requisites of randomness. To illustrate: the formula applies to risks in which the accident producing potentialities are different during business hours than they are during other hours, but it does not apply to risks in which the occurrence of an accident modifies the accident producing potentialities for an interval immediately following such occurrence. The formula applies to risks for which the hazards are greater during a particular season, but it does not apply to risks in which a high accident frequency early in the policy period has led to effective safety engineering the results of which are operative in the balance of the policy period.

Consideration of the characteristics of casualty insurance risks might lead to the surmise that they do not meet the two requisites of randomness with perfection.

Accidents occurring in connection with defective equipment may remove the equipment from the exposure until such time as it can be repaired. A serious collision involving a particularly antiquated truck of a trucking fleet will remove, at least temporarily, that piece of equipment from the road and correspondingly modify the hazard per unit of exposure. Similarly, if there is validity in the contention that a disproportionate part of the workmen's compensation hazard inheres in accident prone employees, disability of one of these employees modifies the hazard per unit of payroll for the period of disability.

Again in the workmen's compensation business, in almost any reasonably large risk there will be safety rules and practices, some formal, some informal. It is not unreasonable to believe that the interest in and enforcement of such rules and practices will vary from time to time. It is likely that a peak of interest and enforcement will be immediately following a spectacular accident.

In several lines of casualty insurance significant accidents of an unusual nature will prompt either the insured or the carrier to initiate efforts to prevent the recurrence of that particular type of accident. A catastrophe involving a bus line may result in a rerouting to avoid certain intersections.

In each of these illustrations an accident operates to reduce the risk hazard for a period of time following its occurrence, and, accordingly, the general requirement that accidents must be independent is not perfectly met. The questions are (1) to what extent do these and other similar imperfections actually exist, and (2) to what degree do they modify the theoretical distribution of accidents given by the Poisson formula and thus impair its use, say, in prognosticating loss ratio distributions from accident cost distributions, or in evaluating the accuracy of a rate making system. Preliminary judgment probably would be that the actual influence of accidents upon one another is very small and can be ignored in other than exceptional circumstances. However, effort to explore this problem seems justified, even if only to confirm this preliminary judgment.

The exploratory efforts made in this paper are very limited. To answer the first question would require an analysis of accident timing in actual risk experience, an analysis which would be made very difficult by a lack of precise figures for true inherent hazards of individual risks and a lack of any information on hazard changes during policy periods. Only the second question has been touched, and that only in a way to provide some assistance in forming personal judgments.

The situations cited as examples obviously do not lend themselves to very much precise mathematical calculation. It may be helpful, however, to define an idealized risk in which the requisite of individual randomness is violated in a nice uniform way and to calculate for several sets of conditions the properties of the theoretical accident frequency distributions for such risks. The variances (second moments about the mean) of these distributions can be compared with the variances of Poisson distributions for the same accident expectations. To utilize the results of these calculations, it will be necessary, of course, to subjectively appraise the lack of independence of accidents occurring in real risks and to compare that appraisal with the conditions assumed for the idealized risks.

The idealized risks selected are assumed to have the following properties: An accident hazard of constant magnitude is operative from the start of the policy period until the occurrence of the first accident. Immediately following the first accident, and immediately following each subsequent accident, the accident hazard becomes negligible for a specific interval. At the end of each such interval the accident hazard immediately resumes its original magnitude and continues at this level until the following accident.

To give these idealized models a touch of reality, they may be considered as corresponding to that portion of a workmen's compensation risk which relates to fatal cases. Whenever a fatal accident happens, factors are brought into play which practically preclude the occurrence of another fatal accident until some time has elapsed. Obviously, in actual risks the hypothetical accident spreading factors would be neither so uniform nor so completely effective. However, they would operate in the general direction of creating a risk with properties similar to those of the mathematical models.

It might be well to mention that this paper is concerned only with the results of a lack of randomness. If the Poisson formula is used with accident expectations derived from casualty insurance rates, discrepancies between actual and theoretical results are usually observed, mainly because casualty insurance rate making does not accurately give the accident expectation of each individual risk. There are a multitude of possible reasons for this theoretical deficiency in casualty rate making, including, conceivably, the possibility that some accidents may modify the risks without correspondingly modifying the rated premiums. In such a situation the discrepancies would arise from at least two sources: inaccurate figures for accident expectations and a lack of randomness in the accidents themselves. This paper is concerned only with the latter.

Let the period of exposure be unity. Let the accident expectation which would prevail if the peak hazard were operative without interruption be a. Let b be the portion of the exposure period for which the hazard is rendered inoperative by an accident.

The average or overall accident expectation which contemplates these intervals of zero hazard will obviously be less than a. It is this average expectation which is involved in what is usually considered to be the inherent hazard of the risk and which is assumed to be reflected in the risk rating. Call this average expectation c.

While the peak hazard is operative, the probability of an accident in any infinitesimal of time dx is adx, and the probability of no accidents occurring for a period of time d is e^{-ad}.

The probability (P_n) of exactly n accidents in a risk as defined above will be obtained by taking the product of a succession of probabilities which account for the n accidents and the exposure time and integrating n times between suitable limits. The product comprises the probability of zero accidents up to point of time t, the probability of an accident at t, the certainty of no accidents between t and $(t+b)$, the probability of zero accidents from $(t+b)$ to u, the probability of an accident at u, etc. The limits of integration are complicated somewhat by the circumstance that the interval of zero hazard occasioned by the last accident is not necessarily confined to the exposure period. This circumstance seems to necessitate $(n+1)$ sets of integrals.

For $nb \lesseqgtr 1$

$$P_0 = e^{-a}$$

$$P_1 = \int_0^{1-b} ae^{-a(1-b)}\,dy + \int_{1-b}^{1} ae^{-ay}\,dy$$

$$P_2 = \int_0^{1-2b}\int_{x+b}^{1-b} a^2 e^{-a(1-2b)}\,dx\,dy + \int_0^{1-2b}\int_{1-b}^{1} a^2 e^{-a(y-b)}\,dx\,dy + \int_{1-2b}^{1-b}\int_{x+b}^{1} a^2 e^{-a(y-b)}\,dx\,dy$$

$$P_3 = \int_0^{1-3b}\int_{w+b}^{1-2b}\int_{x+b}^{1-b} a^3 e^{-a(1-3b)}\,dw\,dx\,dy + \int_0^{1-3b}\int_{w+b}^{1-2b}\int_{1-b}^{1} a^3 e^{-a(y-2b)}\,dw\,dx\,dy$$
$$+ \int_0^{1-3b}\int_{1-2b}^{1-b}\int_{x+b}^{1} a^3 e^{-a(y-2b)}\,dw\,dx\,dy + \int_{1-3b}^{1-2b}\int_{w+b}^{1-b}\int_{x+b}^{1} a^3 e^{-a(y-2b)}\,dw\,dx\,dy$$

$$P_4 = \int_0^{1-4b}\int_{v+b}^{1-3b}\int_{w+b}^{1-2b}\int_{x+b}^{1-b} a^4 e^{-a(1-4b)}\,dv\,dw\,dx\,dy$$
$$+ \int_0^{1-4b}\int_{v+b}^{1-3b}\int_{w+b}^{1-2b}\int_{1-b}^{1} F + \int_0^{1-4b}\int_{v+b}^{1-3b}\int_{1-2b}^{1-b}\int_{x+b}^{1} F$$

$$+\int_0^{1-4b}\int_{1-3b}^{1-2b}\int_{w+b}^{1-b}\int_{x+b}^{1} F+\int_{1-4b}^{1-3b}\int_{v+b}^{1-2b}\int_{w+b}^{1-b}\int_{x+b}^{1} F$$

where $F = a^4\, e^{-a(y-3b)}\, dv\, dw\, dx\, dy$.

Integration reveals that for $0 < n \leqq \frac{1}{b}$ the probability of exactly n accidents takes the form:

$$P_n = \left[1 + a\,(1-nb) + \frac{a^2\,(1-nb)^2}{2} \cdots + \frac{a^n\,(1-nb)^n}{n!}\right] e^{-a(1-nb)}$$
$$- \left[1 + a\,(1-(n-1)b) + \cdots \frac{a^{n-1}\,[1-(n-1)\,b]^{n-1}}{(n-1)!}\right] e^{-a\,[1-(n-1)b]}$$

It will be noted that when b is made zero, P_n reduces to the Poisson formula. It will also be noted that

$$\sum_{n=0}^{n=s} P_n = 1 \text{ when } sb = 1.$$

The latter result could be anticipated, as under the conditions given it is impossible to have more than s accidents.

Several distributions have been calculated from the above formula and the results are exhibited in the following table:

		$a=2$ $b=0.2$	$a=2$ $b=0.1$	$a=2$ $b=0.05$	$a=10$ $b=0.2$	$a=10$ $b=0.1$
(1)	c or $\Sigma\, P_n X$	1.4693	1.6805	1.8223	3.5539	5.1509
(2)	$\Sigma\, P_n\, x^2$	.7975	1.1877	1.5134	.4889	1.4933
(3)	$\frac{\Sigma\, P_n\, x^2}{c}$	.5427	.7067	.8305	.1376	.2899
(4)	$\sqrt{(3)}$	.7375	.8407	.9113	.3709	.5384

Line 1 gives the accident expectations of the risks as defined. These figures approach $\frac{a}{1+ab}$ as b is made smaller, but again because the zero hazard interval following the last accident is not confined to the policy period, they are somewhat larger.

In the Poisson distribution, the variance (second moment about the mean) is equal to the expectation (first moment from the origin). Consequently, line 1 also gives the variances of Poisson distributions most nearly comparable with the distributions found for the idealized risks.

Line 2 gives the variances calculated for the idealized risk distributions. Line 3 shows the ratios of the two variances, and line 4 shows the square roots of these ratios. Since the standard deviation is the square root of the variance, the figures in line 4 are the ratios of the standard deviations and are intended to show the degree to which these departures from randomness affect dispersion. As might be anticipated, if the values of either a or b are increased, the dispersion is moved further from the value it would have if random conditions prevailed.

It is noted that the conditions assumed: i.e., a reduction of the accident producing potentialities following the occurrence of an accident, tend to reduce the dispersion. Conversely, it may be assumed that the dispersion would be increased by conditions under which an accident renders more likely the occurrence of other accidents. Something analogous to the latter situation is found in the use of claim data as the only available approximation for accident data in estimating loss ratio distributions. When there is the possibility of multiple claim accidents, the instant of time coincident with or immediately following a claim may be considered to have been endowed with a special hazard for other related claims.

With respect to the influences operating in the direction of the idealized risks, it is necessary to inject personal opinion of their magnitude if any effort is to be made to relate the mathematical models to reality. There should be general agreement that this magnitude is sufficiently small so that actual risks, or even the serious portion of actual risks, could not be considered comparable with any of the idealized risks with the possible exception of the one in which $a = 2$ and $b = 0.05$.

If a workmen's compensation risk is of such a size that an average of 1.8 serious accidents occur per year, it is conceivable that the occurrence of one of these would bring forces into play which would preclude a recurrence for a little over two weeks, or at least greatly reduce the hazard for a correspondingly longer period. Such a departure from randomness would reduce the standard deviation of the serious accident distribution by about 9%. Unfortunately, 9% is neither large enough to be startling nor small enough to be dismissed as insignificant.

It is a matter of conjecture just how many risks are subject to such departures from randomness. If it is agreed that these influences are normally much less effective, then it can be concluded that concern over the applicability of the Poisson distribution to casualty insurance accidents can be confined to special situations in which accidents are known definitely to be other than independent.