

## The Generalized Waring Distribution applied to Accident Theory

By J. O. IRWIN

*The Galton Laboratory*

### SUMMARY

The generalized Waring distribution provides a model for accident distributions in which the variance can be split into *three* components. The first of these is the usual random component of classical accident theory and the other two will often be identifiable as the separate variances due to proneness and liability. The sum of the last two components is the single component which in the classical theory is identified with variation in susceptibility.† This new model is illustrated by an example taken from Newbold's original work.

The negative binomial is a limiting case of the generalized Waring distribution; so that the classical theory of those particular compound Poisson distributions which are of negative binomial form is a particular case of that presented here. Just as a compound Poisson distribution *need not* be of negative binomial form (it will only be so when "susceptibility" has a gamma distribution) so the "three component distribution" is not necessarily a generalized Waring distribution, but will only be so under the conditions explained in the paper. This more general case, while alluded to in a paragraph on p. 223, is not dealt with in this paper.

The generalized Waring distribution (of the number  $A$  of accidents) depends on three parameters,  $\rho, a, k$  and is symmetrical in  $a, k$ . In Part II it is shown how these parameters may be expressed in terms of three parameters directly relevant to a real accident situation. These are  $\bar{A}$  the mean number of accidents in the exposure period,  $\omega = (\sigma_A^2 - \bar{A})/\bar{A}^2$  which is the square of the coefficient of variation of "susceptibility", and a parameter  $d$  which is the percentage difference between the two components into which the variance of "susceptibility" is divisible.  $a, k, \rho$  are single valued functions of  $\bar{A}, \omega$  and  $d$ , and vice versa. Table 4 gives values of  $a, k, \rho$  for given  $\bar{A}, \omega, d$ . Interchanging  $a$  and  $k$  does not alter the Waring distribution; thus  $\bar{A}$  and  $\rho$  remain the same but  $d$  is replaced by a value  $d'$ . Over a large range of values of  $a$  and  $k$ ,  $d$  and  $d'$  do not differ greatly. They are equal with opposite sign, when  $\bar{A} = 1$ , and with the same sign when  $a = k$  which implies  $\bar{A} = \bar{A}_1$ , a value which is a function of  $d$  and  $\omega$ .

Table 4 gives  $a, k, \rho, d'$  for given values of  $\bar{A}, \omega$  and  $d$ . In addition to giving a conspectus of the whole situation over the possible range of values of the parameters, it should be useful to analysts of accident data, as is explained in a note accompanying the table.

It is realized that this is only a first study of a new technique. It is hoped that younger statisticians, with greater computing facilities than the author, will be stimulated to explore it.

† I have retained Newbold's term "susceptibility" for a person's tendency to accident due to all non-random factors, "liability" to mean his tendency to accident due to external factors, and "proneness" that due to internal factors. I use "proneness" in the same sense as Farmer and Chambers (1939); they used "liability" as synonymous with Newbold's "susceptibility". I think this can be confusing. For some further discussion of this question see Irwin (1964).

PART I—GENERAL THEORY AND AN EXAMPLE

1. IN A PAPER read to the Royal Statistical Society (Irwin, 1963), I drew attention to the generalized Waring distribution. Its generating function can be written in the form

$$\frac{(x-a)_{[k]}}{x_{[k]}} F(a, k, x+k, \theta) \tag{1}$$

where  $F$  is the hypergeometric series,  $\theta$  the generating symbol and  $x_{[k]}$  denotes  $x(x+1)\dots(x+k-1)$ . However, it is in general more convenient to put  $x-a = \rho$  and write (1) in the form

$$\frac{\Gamma(\rho+a)\Gamma(\rho+k)}{\Gamma(\rho)\Gamma(\rho+a+k)} \left\{ 1 + \frac{ak\theta}{\rho+a+k} + \frac{a(a+1)k(k+1)\theta^2}{(\rho+a+k)(\rho+a+k+1)2!} + \frac{a_{[r]}k_{[r]}\theta^r}{(\rho+a+k)_{[r]}r!} \right\} \tag{2}$$

Here  $\rho > 0, a > 0, k > 0$  and  $k$  need not be an integer, in fact the distribution is symmetrical in  $a$  and  $k$ .

Originally I was interested in the distribution because, for certain values of the parameters, it can have both a mode and a very long tail, whereas the simple Waring distribution (which has  $k = 1$ ), while possessing the latter property, is always J-shaped. This in turn has been suggested by coming across actual biological distributions which had exceptionally long tails. Theoretical distributions (with one exception) so far discussed in the literature, were totally inadequate to deal with this situation. The exception was the Yule distribution (Yule, 1924); this is in fact the particular case of (2) when  $a = 1, k = 1$ .

However, the tail is not exceptionally long for *all* values of the parameters; in fact the continuous analogue of the distribution is Pearson’s Type VI, with some relatively rare exceptions when it is Type IV.† In the application of it which we shall make to accident theory the tails will not usually be especially long.

The expression in (2) can be written

$$\begin{aligned} & \frac{\Gamma(\rho+a)\Gamma(\rho+k)}{\Gamma(\rho+a+k)\Gamma(\rho)} F\{a, k, \rho+a+k, \theta\} \\ &= \frac{\Gamma(\rho+k)}{\Gamma(\rho)\Gamma(k)} \int_0^1 u^{\rho-1}(1-u)^{k-1} \left\{ \frac{1}{u} - \theta \left( \frac{1}{u} - 1 \right) \right\}^{-a} du. \end{aligned} \tag{3}$$

The integral in (3) is a well-known representation of the hypergeometric function (Forsyth, 1914; Whittaker and Watson, 1935).

2. The negative binomial distribution often fits observed accident distributions. The classical theory (Greenwood and Yule, 1920; Newbold, 1925, 1927) supposes that we have a population of individuals all subject to the same external risk of accident, but with unequal *proneness* to accident. An individual’s *proneness* is represented by  $\lambda$  the mean number of accidents incurred under the given exposure conditions by individuals with the same proneness as he has.  $\lambda$  is assumed to have the continuous distribution

$$\frac{1}{\Gamma(a)} e^{-c\lambda}(c\lambda)^{a-1} d(c\lambda) \tag{4}$$

† Since 1964 I have studied the general properties of the distribution, moments, shape, existence and position of mode, etc., and the way in which these vary with the parameters; and extensive calculations have been made on an electronic computer. I hope to publish these results shortly.

while the distribution of accidents among individuals with the same proneness is the Poisson distribution  $\exp\{\lambda(\theta-1)\}$ . Writing  $\theta = 1 + \alpha$  we find that the factorial moment generating function of (5) is  $e^{\lambda\alpha}$ . Thus the f.m.g.f. of the resulting accident distribution is

$$\frac{1}{\Gamma(a)} \int_0^\infty e^{-(c-\alpha)\lambda} (c\lambda)^{a-1} d(c\lambda) = \left(1 - \frac{\alpha}{c}\right)^{-a} \tag{5}$$

Now putting  $\alpha = \theta - 1$  we find for the g.f. of the final accident distribution

$$\left(\frac{c+1}{c} - \frac{\theta}{c}\right)^{-a}, \tag{6}$$

a negative binomial distribution with mean  $a/c$ . Apart from the fact that this combination of different Poisson distributions is not the only interpretation of a negative binomial distribution, there is another difficulty in *applying* the theory. We cannot know that all individuals are exposed to exactly the same external risk of accident. Differences in exposure to external risk of accident from person to person are known as differences in accident *liability* as distinguished from constitutional or internal differences which are known as differences in *proneness*. In practice, effects of proneness and liability are *confounded*, that is inseparable, when the negative binomial is fitted. The combination may be called *susceptibility*, the term originally used by Newbold. This is true whatever the distribution of  $\lambda$ . The original choice (4) for the  $\lambda$  distribution of susceptibility, was largely a matter of convenience. It provided a simple form of distribution which worked; and led to a simple distribution of accidents—the negative binomial. Certain results are true, whatever the distribution of  $\lambda$ , as long as the distribution for fixed  $\lambda$  is a Poisson with mean  $\lambda$ . In particular,

(i) The following results are true:

$$\left. \begin{aligned} \bar{\lambda} &= \bar{A} \\ \sigma_\lambda^2 &= \sigma_A^2 - \bar{A} \\ r_{A\lambda} &= (\sigma_\lambda / \sigma_A) = \{1 - (\bar{A} / \sigma_A^2)\}^{\frac{1}{2}} \end{aligned} \right\} \tag{7}$$

where  $\bar{A}$  and  $\sigma_A$  are the (theoretical) mean and S.D. of the final accident distribution.

(ii) The factorial moment generating function of the accident distribution is the ordinary moment generating function of the  $\lambda$  distribution. For the f.m.g.f. of the Poisson distribution with mean  $\lambda$  is  $e^{\lambda\alpha}$ , whence the f.m.g.f. of the accident distribution is

$$\int_0^\infty p(\lambda) e^{\lambda\alpha} d\lambda$$

which is the m.g.f. of the  $\lambda$  distribution. If  $\lambda$  has the distribution (4), then

$$\bar{A} = \bar{\lambda} = a/c \quad \sigma_\lambda^2 = a/c^2 \quad r_{A\lambda} = \{1/(c+1)\}^{\frac{1}{2}}$$

3. The generalized Waring distribution may be used to allow for proneness and liability separately. In the negative binomial

$$\left\{ \frac{1}{u} - \theta \left( \frac{1}{u} - 1 \right) \right\}^{-a}, \quad \frac{1}{u} = \frac{c+1}{c} = 1 + \frac{v}{a}$$

can be regarded as a parameter fixed for individuals with the same proneness. The negative binomial, which may be written

$$\left(1 + \frac{\nu}{a} - \frac{\nu\theta}{a}\right)^{-a}$$

can be regarded as the distribution for individuals with proneness  $\nu$  arising from a Poisson distribution  $\exp\{(\lambda|\nu)\theta - 1\}$  for individuals with liability  $(\lambda|\nu)$  and fixed proneness  $\nu$ , where  $(\lambda|\nu)$ , that is  $\lambda$  for given  $\nu$ , has the usual Pearson Type III (or gamma) distribution

$$\frac{1}{\Gamma(a)} e^{-a\lambda/\nu} \left(\frac{a\lambda}{\nu}\right)^{a-1} d\left(\frac{a\lambda}{\nu}\right). \tag{8}$$

If now  $u$  has the Pearson Type I (or Beta) distribution

$$\frac{\Gamma(\rho+k)}{\Gamma(\rho)\Gamma(k)} u^{\rho-1}(1-u)^{k-1} du \tag{9}$$

it follows from (3) that the generalized Waring distribution will give the final distribution of accidents.

In terms of  $\nu$ , (9) becomes

$$\frac{\Gamma(\rho+k)}{\Gamma(\rho)\Gamma(k)} \left(\frac{\nu}{a}\right)^{k-1} \left(1 + \frac{\nu}{a}\right)^{-(\rho+k)} \frac{d\nu}{a} \tag{10}$$

which has the Pearson Type VI form.

The original choice of (8) for the  $\lambda$  distribution of proneness or liability, as has been already noted, was largely a matter of convenience.

A beta distribution for  $u$  or  $\nu$  (Type VI is of course also closely related to a beta-distribution) is more general but equally plausible as an assumption; (8) is in fact one limiting form of Type I or Type VI.

The variance of the Waring distribution is

$$\sigma_A^2 = \frac{ak(\rho+k-1)(\rho+a-1)}{(\rho-1)^2(\rho-2)}. \tag{11}$$

The variance of the Poisson distribution corresponding to  $(\lambda|\nu)$  is  $\nu$ , thus the variance of the random component in the final distribution is  $\bar{\nu} = ak/\rho - 1$ . The variance of  $\nu$  is  $a^2 k(\rho+k-1)/(\rho-1)^2(\rho-2)$ , thus if

$$\begin{aligned} \sigma_{\lambda|\nu}^2 &= E_\nu\{V(\lambda|\nu)\} \\ \sigma_{\lambda\nu}^2 &= E_\nu(\nu^2/a) = ak(k+1)/(\rho-1)(\rho-2). \end{aligned} \tag{12}$$

Thus finally we have the situation shown in Table 1 and the first two columns of Table 2.

If the generalized Waring distribution is found to fit the data, we have a means of estimating separately accident proneness and accident liability (but see p. 222). There is nothing in the above mathematical analysis, however, to tell us whether  $\lambda$  is liability and  $\nu$  proneness or vice versa. If, however, we know that external exposure

was approximately constant for the individuals for whom we have data and if one of the variance components (say  $\sigma_{\nu}^2$ ) is very much larger than the other (say  $\sigma_{\lambda|\nu}^2$ ), we shall be justified in attributing the larger component to proneness.

TABLE 1  
*The component distributions and their resultant†*

Persons with	
(1) Same proneness and liability (Poisson)	$\exp \{(\lambda   \nu) (\theta - 1)\}$
(2) Distribution of $(\lambda   \nu)$ , i.e., $\lambda$ for fixed $\nu$	$\frac{1}{\Gamma(a)} \exp \left\{ - \left( \frac{a\lambda}{\nu} \right) \left( \frac{a\lambda}{\nu} \right)^{a-1} d \left( \frac{a\lambda}{\nu} \right) \right\}$
(3) Resulting <i>negative binomial</i>	$\left( 1 + \frac{\nu}{a} - \frac{\nu\theta}{a} \right)^{-a}$
(4) Proneness distribution of $\nu$	$\frac{\Gamma(\rho+k)}{\Gamma(\rho)\Gamma(k)} \left( \frac{\nu}{a} \right)^{k-1} \left( 1 + \frac{\nu}{a} \right)^{-(\rho+k)} \frac{d\nu}{a}$
(5) Final resulting distribution	$\frac{\Gamma(a+\rho)\Gamma(k+\rho)}{\Gamma(a+k+\rho)\Gamma(k)} F\{a, k, a+k+\rho, \theta\}$

† The generating function is given for discrete distributions, and the probability density for continuous distributions.

$F$  is the hypergeometric series  $\left\{ 1 + \frac{ak}{(a+k+\rho)}\theta + \frac{a(a+1)k(k+1)}{(a+k+\rho)(a+k+\rho+1)}\frac{\theta^2}{2!} + \dots \right\}$ .

TABLE 2  
*The variance of the generalized Waring distribution and its components with estimates made from the observed distribution*

Component	Theoretical	Variance	
		Estimated from the observed accident distribution (given in Table 3)	
		(1) $a = 6.0480$ $k = 1.0595$	(2) $k = 6.0480$ $a = 1.0595$
(1) Random $Ev(\lambda   \nu) = \bar{\nu} = \bar{A}$	$\frac{ak}{(\rho-1)}$	0.9776 (37.9%)	0.9776 (37.9%)
(2) Liability component $Ev\{V(\lambda   \nu)\} = \sigma_{\lambda \nu}^2 = \sigma_{\nu}^2$	$\frac{ak(k+1)}{(\rho-1)(\rho-2)}$	0.3625 (14.1%)	1.2405 (48.1%)
(3) Proneness component	$\frac{a^2 k(\rho+k-1)}{(\rho-1)^2(\rho-2)}$	1.2366 (48.0%)	0.3586 (13.9%)
Total	$\frac{ak(\rho+k-1)(\rho+a-1)}{(\rho-1)^2(\rho-2)}$	2.5767 (100%)	2.5767 (100%)

4. An example, given by Newbold in her original papers (1925, 1927), may be used here. Table 3 gives the observed frequencies of accidents among men in a soap factory. The period of exposure was 5 months. Miss Newbold fitted a negative

TABLE 3  
*Accidents to men in a soap factory (5 months' exposure)*

	Observed	Expected Negative binomial (Newbold)	Expected Generalized (Waring)
0	239	251	240
1	98	93	105
2	57	46	49
3	33	25	24
4	9	14	12
5	2	8	7
6	2	4	4
7	1		
8	—		
9	4		
10	1	6	6
11	—		
12	—		
13	1		
Total	447	447	447
		$\chi^2_5 = 13.7$ $P = 0.018$	$\chi^2_4 = 10.6$ $P = 0.030$

binomial distribution, by equating the observed values of the mean and variance to their theoretical values. Observed and expected frequencies for seven or more accidents were combined into one group, after fitting. Thus there were eight frequency groups giving five degrees of freedom for which  $\chi^2_5 = 13.7$ ,  $P = 0.018$  (Miss Newbold, as was not unusual at the time, used seven degrees of freedom which gave  $P(\chi^2_7) = 0.058$ ). I have fitted the generalized Waring distribution, using three factorial moments and find  $\rho = 7.55446$ ,  $ak = 6.40784$ ,  $a+k = 7.10748$ , whence  $a = 6.04798$ ,  $k = 1.05950$ , or vice versa. The Generalized Waring Distribution is symmetrical in  $a$  and  $k$ , so we obtain two solutions. I have chosen  $a$  to be the bigger value, because this makes the variance component for proneness  $\sigma_p^2$  much bigger than that for liability  $\sigma_\lambda^2$ . This gives  $\chi^2_4 = 10.6$ ,  $P = 0.030$ .

The distribution fitted by this method, as well as Newbold's original negative binomial are both given in Table 3. The generalized Waring distribution fits rather better than the negative binomial, but this is not very important. The advantage of the former lies in the possibility of dividing the total variance into three additive components as in Table 2, where column (3) gives the results for  $a = 6.0480$ ,  $k = 1.0895$ , and the last column shows the effect of interchanging  $a$  and  $k$ . Two features of these results are noteworthy. The sum of the second and third components is the same in columns (3) and (4); this must be so because their sum is  $\sigma_A^2 - \bar{A}$ , both of which are symmetrical in  $a$  and  $k$ . Furthermore, the effect of interchanging  $a$  and  $k$  is almost the same as interchanging the roles of  $\lambda$  and  $\nu$ , as far as the variance

components are concerned. It is easy to see from column (1) that this would be exactly true if  $\rho = ak + 1$  in which case  $\bar{A} = 1$ . In fact the estimated values are  $\rho = 7.55446$ ,  $ak + 1 = 7.04784$ ,  $\bar{A} = 0.9776$ . This is further discussed in Section 5.

Column (3) of Table 2 also shows that accident liability accounts for about 14 per cent of the total variance. Newbold remarked that there was some heterogeneity between different departments in the factory; the correlation ratio,  $\eta_{A,x}$ , where  $A$  is the number of accidents and  $x$  the department, being 0.36. The proportion of the total variance due to departmental differences, in terms of  $\eta$ , depends on the number of departments but is not very different from  $\eta^2$  or 13 per cent. For five departments it is about 16 per cent and for 10 departments about 12 per cent. Thus the observed 14 per cent is just about what we should expect, if there were variation in liability between but not within departments.

5. The assumptions made above can be stated a little more explicitly. Let  $A$  be the number of accidents incurred under the given exposure conditions. Let  $\nu$  be the expected number of accidents for a person with a given fixed degree of proneness and average liability, relative to the given exposure conditions. Let  $\lambda|\nu$  be the expected number of accidents incurred by an individual of proneness  $\nu$  and liability  $\lambda$ .

It is assumed that whatever the distributions of  $\lambda$  and  $\nu$  may be,

$$A = \nu + (\lambda|\nu) - (\bar{\lambda}|\nu) + \epsilon \tag{13}$$

where  $\epsilon$  is a Poisson variable with mean  $\bar{\lambda}|\nu$ . Thus

$$\left. \begin{aligned} A &= \lambda|\nu + \epsilon \\ \bar{A} &= \bar{\nu} \end{aligned} \right\} \tag{14}$$

$$\left. \begin{aligned} \sigma_A^2 &= \sigma_\nu^2 + E_\nu\{V(\lambda|\nu)\} + \bar{\nu} \\ &= \sigma_\nu^2 + \sigma_{\lambda|\nu}^2 + \bar{\nu} \end{aligned} \right\} \tag{15}$$

and clearly,

$$r_{A\nu} = (\sigma_\nu/\sigma_A) \quad r_{A\lambda} = (\sigma_{\lambda|\nu}/\sigma_A) = (\sigma_\lambda/\sigma_A) \tag{16}$$

while the multiple correlation coefficient is

$$R_{A,\lambda\nu} = \{1 - (\bar{A}/\sigma_A^2)\}^{\frac{1}{2}} \tag{17}$$

It is shown in Section (6) that  $\sigma_{\lambda|\nu} = \sigma_\lambda$ . These are the properties analogous to those expressed by (i), p. 207 (equation (7)) in the classical accident theory.

Further, the f.m.g.f. of the Poisson with fixed  $\lambda$  and  $\nu$  is  $e^{(\lambda|\nu)\alpha}$  and the f.m.g.f. of the accident distribution is

$$\sum p(\nu)p(\lambda|\nu) e^{(\lambda|\nu)\alpha} \tag{18}$$

But this is the m.g.f. of the distribution of  $(\lambda|\nu)$ . This is completely analogous to (ii), p. 207.

It is interesting to verify (18) for the particular distributions used here. The m.g.f. of  $\lambda|\nu$  for the distribution (8) is  $(1 - \alpha/a)^{-a}$  whence

$$\mu'_r(\lambda|\nu) = a(a+1) \dots (a+r-1) \left(\frac{\nu}{a}\right)^r \tag{19}$$

while for the distribution (10),

$$\mu'_r\left(\frac{\nu}{a}\right) = \frac{k(k+1) \dots (k+r-1)}{(\rho-1)(\rho-2) \dots (\rho-r)} \tag{20}$$



Hence the factorial moment  $\mu_{[r]}(A)$  is given by

$$E_{\nu}\{\mu'_{[r]}(\lambda|\nu)\} = \frac{a_{[r]}k_{[r]}}{(\rho-1)(\rho-2)\dots(\rho-r)} \tag{21}$$

which is known to be the  $r$ th factorial moment of the generalized Waring distribution. 6. It is important to notice that the average over  $\nu$  of the conditional distribution of  $\lambda-\nu$ , is the same as the unconditional distribution of  $\lambda-\bar{A}$  for if  $x = \lambda-\nu$ ,  $x$  has means zero and conditional distribution

$$p(x|\nu)dx = \frac{1}{\Gamma(a)} e^{-a(1+x/\nu)}[a\{1+x/\nu\}^{a-1}](a/\nu) dx \tag{22}$$

The average unconditional distribution of  $\lambda-\nu$  is therefore

$$\int_0^\infty p(\nu)p(x|\nu) dx = \frac{1}{\Gamma(a)} \int_0^\infty p(\nu) e^{-a\lambda/\nu} \left(\frac{a\lambda}{\nu}\right)^{a-1} d\left(\frac{a\lambda}{\nu}\right) \tag{23}$$

but this is the unconditional distribution of  $\lambda$ , with unconditional mean  $\bar{A}$ . It follows that  $\sigma_{\lambda,\nu}^2 = \sigma_{\lambda}^2$ , where  $\sigma_{\lambda,\nu}^2$  is defined as  $E_{\nu}\{V(\lambda|\nu)\}$  and  $\sigma_{\lambda}^2$  is the unconditional variance of  $\lambda$ .

In view of this and the other results so far obtained, it might be tempting to suppose that interchanging  $a$  and  $k$  is equivalent to interchanging the roles of  $\lambda$  and  $\nu$ . That is if  $\lambda(a, k)$  denotes liability and  $\nu(a, k)$  proneness;  $\nu(k, a)$  denotes liability and  $\lambda(k, a)$  proneness. This is exactly true, as far as variance components are concerned, when  $\rho = ak + 1$  which implies  $\bar{A} = 1$ . However, it is not true in general. For if we carry out these interchanges we find for the probability density of  $\nu$  for fixed  $\lambda$

$$\frac{1}{\Gamma(k)} e^{-(k\nu/\lambda)}(k\nu/\lambda)^{k-1} d(k\nu/\lambda) \tag{24}$$

and for that of  $\lambda$

$$\frac{\Gamma(\rho+a)}{\Gamma(\rho)\Gamma(a)} \left(\frac{\lambda}{k}\right)^{a-1} \left(1+\frac{\lambda}{k}\right)^{-(\rho+a)} \frac{d\lambda}{k} \tag{25}$$

Now either (8) and (10) or (24) and (25) can be true; and either pair leads to the same generalized Waring distribution of accidents. But both pairs cannot be true simultaneously, even if  $\rho = ak + 1, \bar{A} = 1$ . For consistency the four probability densities should satisfy

$$p(\lambda|\nu)p(\nu) = p(\nu|\lambda)p(\lambda)$$

and they clearly do not. If one assumes (8) and (10) to be true, one finds a series of confluent hypergeometric type for the unconditional density of  $\lambda$ . The two pairs of assumptions might be approximately true for the sort of values of  $\rho, a, k$  which occur in practice, or for some range of these values. To verify this, one method would be to show that (25) does not differ much from the unconditional distribution of  $\lambda$  given by (8) and (10), and that (10) does not differ very much from the unconditional distribution of  $\nu$  given by (24) and (25). If one pair of assumptions is not even approximately true, the other pair may still provide a solution and knowledge of the data would usually tell us whether  $\sigma_{\lambda}$  was proneness and  $\sigma_{\nu}^2$  liability or vice versa.

Alternatively, the matter can be examined by comparing the two solutions over a suitable range of values of the parameters. This is what has been done. The object of Part II, which follows, is to examine critically the degree of agreement or disagreement between the two pairs of solutions.



PART II.—CRITIQUE OF THE THEORY

1. The best way of examining the relation between the two sets of solutions obtained by interchanging  $a$  and  $k$ , is to choose three parameters which are clearly descriptive of the actual accident situation and to express  $a, k, \rho$  in terms of these; then to select a series of values for these parameters such as might occur in practice (or are otherwise interesting) and (i) find the corresponding values of  $a, k, \rho$  (ii) interchange  $a$  and  $k$  and find the new set of parameters. We shall then be able to compare the values of  $\bar{A}, \sigma_v^2, \sigma_\lambda^2$  before and after the interchange;  $\bar{A}$  and  $\rho$  will remain the same.

2. (i) One obvious parameter to choose is the mean  $\bar{A}$  of the whole accident distribution. Other conditions, external and internal, remaining the same  $\bar{A}$  will increase directly as exposure time. Large values of  $\bar{A}$ , which would never occur in practice, can tell us much about what would happen to our accident phenomena if exposure time were greatly prolonged.

(ii) The early workers on the subject used the “ $r_{A\lambda}$ ” of (i) (p. 207) to measure the importance of “susceptibility”. This is the same as our  $R_{A,\lambda v} = \{1 - (\bar{A}/\sigma_A^2)\}^{\frac{1}{2}}$ . Though Newbold did not do so, they often tended to identify “proneness” with “susceptibility”, assuming that variation in the external conditions, under which individual subjects worked, was negligible. “ $r_{A\lambda}$ ” was not a very good measure to use, for when exposure time increases  $\bar{A}/\sigma_A^2$ , that is the relative size of the random component of the variance, tends to fall and “ $r_{A\lambda}$ ” to increase.

A better measure to use is  $\omega = (\sigma_A^2 - \bar{A})/\bar{A}^2$  as was implicitly suggested by Chambers and Yule (1941). For this is the square of the coefficient of variation of the distribution of total susceptibility, and should be independent of exposure time *per se*. We then have  $\sigma_A^2 = \bar{A} + \omega\bar{A}^2$  where  $\bar{A}$  is the random component. Although “ $r_{A\lambda}$ ” is not a good measure to use, there is much experience of accident data to suggest that values of “ $r_{A\lambda}$ ” below 0.5 are of little use in discriminating individual variation and that values above 0.95 are seldom obtained. Now

$$r_{A\lambda}^2 = 1 - \frac{\bar{A}}{\sigma_A^2}$$

Thus values of  $\omega = \frac{1}{2}, 1, 2, 3, 4$  will be more than sufficient to cover the practical range; for we then have, for example, for  $\bar{A} = 0.1, 1, 10$  and  $\omega = \frac{1}{2}, 1, 2, 4$ , the following values of Newbold’s  $r_{A\lambda}$  or our  $R_{A,\lambda v}$ .

	$\bar{A}$		
	0.1	1	10
$\omega = \frac{1}{2}$	0.218	0.577	0.913
1	0.302	0.707	0.953
2	0.409	0.816	0.976
4	0.535	0.894	0.988

(iii) The third parameter should express the magnitude of the differences between  $\sigma_v^2$  and  $\sigma_\lambda^2$ . We take for this  $d = (\sigma_v^2 - \sigma_\lambda^2)/\sigma^2$  where  $\sigma^2 = \sigma_v^2 + \sigma_\lambda^2 = \omega\bar{A}^2$ .  
 3. We must now express  $\rho, a, k$  in terms of  $\bar{A}, \omega$  and  $d$ . Clearly  $\sigma_v^2/\sigma^2 = \frac{1}{2}(1+d)$ ,  $\sigma_\lambda^2/\sigma^2 = \frac{1}{2}(1-d)$ ; so  $d$  can vary from  $-1$  to  $+1$ . Interchanging  $a$  and  $k$ , will give a new set of Waring parameters,  $\rho, k, a$  ( $\rho$  remaining the same). We proceed as follows.

From the results in Table 2 we have:

$$\sigma_v^2 = \frac{\bar{A}(\bar{A}+a)}{\rho-2} \quad \sigma_\lambda^2 = \frac{\bar{A}(k+1)}{\rho-2} \tag{1}$$

$$d = \frac{\sigma_v^2 - \sigma_\lambda^2}{\omega \bar{A}^2} = \frac{(\bar{A}-1)+(a-k)}{\omega \bar{A}(\rho-2)} \tag{2}$$

Now  $\omega = (\sigma_A^2 - \bar{A})/\sigma_A^2$ . Substituting the values from Table 2, and from (2) above we find:

$$\left. \begin{aligned} a+k &= \omega \bar{A}(\rho-2) - (\bar{A}+1) \\ a-k &= d\omega \bar{A}(\rho-2) - (\bar{A}-1) \end{aligned} \right\} \tag{3}$$

Whence

$$\left. \begin{aligned} a &= \bar{A}\left\{\frac{1}{2}(1+d)\omega(\rho-2) - 1\right\} \\ k &= \bar{A}\left\{\frac{1}{2}(1-d)\omega(\rho-2)\right\} - 1 \end{aligned} \right\} \tag{4}$$

Thus

$$\begin{aligned} 4ak &= (a+k)^2 - (a-k)^2 \\ &= \omega^2 \bar{A}^2(1-d)^2(\rho-2)^2 - 2\omega \bar{A}\{\bar{A}(1-d) + 1 + d\}(\rho-2) + 4\bar{A} \end{aligned} \tag{5}$$

But

$$4ak = 4\bar{A}(\rho+1) = 4\bar{A}(\rho-2) + 4\bar{A}$$

and thus either  $\rho = 2$ , leading to a distribution with infinite variance but finite mean, not applicable to accident data,† or

$$4 = \omega^2 \bar{A}(1-d^2)(\rho-2) - 2\omega \bar{A}(1+d) - 2\omega(1-d) \tag{6}$$

giving

$$(\rho-2) = \{4 + 2\omega \bar{A}(1+d) + 2\omega(1-d)\} / \omega^2 \bar{A}(1-d^2) \tag{7}$$

$$= \frac{2}{\omega(1+d)} + \frac{2}{\omega \bar{A}(1-d)} + \frac{4}{\omega^2 \bar{A}(1-d^2)} \tag{8}$$

and

$$(\rho-1) = \left(1 + d + \frac{2}{\omega}\right) \left(1 - d + \frac{2}{\omega \bar{A}}\right) / (1-d^2) \tag{9}$$

Thus

$$a = \bar{A}\left\{\frac{1}{2}(1+d)\omega(\rho-2) - 1\right\} = \left(1 + d + \frac{2}{\omega}\right) / (1-d) \tag{10}$$

and

$$k = \left\{\bar{A}(1-d) + \frac{2}{\omega}\right\} / (1+d) \tag{11}$$

Equations 9, 10 and 11 enable us to express  $\rho, a, k$  as simple, single valued functions of  $\bar{A}, d$  and  $\omega$ , or conversely.

† The solution is given by Part I, equation (2), with  $a = (1+d)/(1-d), k = \bar{A}(1-d)/(1+d)$  and  $\omega \rightarrow \infty, \bar{A} \rightarrow \infty$ . It is of interest as a limiting case. It is easily shown that, in this case,  $d \rightarrow 0, d' \rightarrow -1$  or vice versa.

We can now choose any value of  $d$ , for example  $d = \frac{1}{2}$ , for which  $\sigma_v^2/\sigma^2 = 0.75$  and  $\sigma_\lambda^2/\sigma^2 = 0.25$  (where  $\sigma^2 = \omega A^2$ ) and calculate  $\rho, a, k$ . The result of interchanging  $a$  and  $k$  will yield a new value  $d'$ , for which  $\sigma_v^2(d')/\sigma^2 = \frac{1}{2}(1+d')$  and  $\sigma_\lambda^2(d')/\sigma^2 = \frac{1}{2}(1-d')$  and

$$d' = \frac{k - 1 - \frac{2}{\omega}}{k + 1} = \frac{\bar{A} + \frac{2}{\omega} - a}{\bar{A} + a} \tag{12}$$

The two forms of  $d'$  provide a check on the accuracy of the calculations. Thus we can compare the two solutions obtained by interchanging  $a$  and  $k$ , keeping  $\rho$  the same.

This has been done for

$$d = 0.1, 0.3, 0.5, 0.7, 0.9,$$

$$\omega = \frac{1}{2}, 1, 2, 3, 4$$

and

$$\bar{A} = 0.1, 0.2, 0.5, 1, 2, 5, 10.$$

It would be rare to find accident data for which  $\bar{A}$  was as large as 10, but Newbold's 1927 paper quotes a value 6.4 for  $\bar{A}$ . However, consideration of the form of the distribution for large values of  $\bar{A}$  is of interest, as showing what would happen if exposure time were greatly increased.

The results are shown in Table 4. It will be observed that as  $\bar{A}$  increases from 0 to  $\infty$ ,  $d'$  increases from  $(2/\omega a - 1)$  to 1; that is from  $-(\omega d + \omega - 2)/(\omega d + \omega + 2)$  to 1. The lower limit is necessarily  $> -1$ . If  $d$  is positive the lower limit for  $d'$  is negative. Thus for some  $\bar{A} = \bar{A}_0$ ,  $d' = 0$  and  $\sigma_v^2(d') = \sigma_\lambda^2(d') = \frac{1}{2}\sigma^2$ . For  $\bar{A} > \bar{A}_0$ ,  $d'$  is positive and both solutions give  $\sigma_v^2 > \sigma_\lambda^2$ . The value of  $\bar{A}_0$  is easily found by putting  $d' = 0$  in (12).

This gives

$$\bar{A}_0 = a - \frac{2}{\omega} = \left\{ 1 + d \left( 1 + \frac{2}{\omega} \right) \right\} / (1 - d) \tag{13}$$

e.g. for  $d = \frac{1}{2}$  we have for  $\omega = \frac{1}{2}, 1, 2, 3, 4$ ;  $\bar{A}_0 = 7, 5, 4, 3\frac{2}{3}, 3\frac{1}{2}$ . The values of  $\bar{A}_0$  are also included in the tables.

In general, the reciprocal property that interchange of  $a$  and  $k$  gives much the same result as the interchange of  $\sigma_\lambda^2$  and  $\sigma_v^2$  holds approximately for a certain range of  $\bar{A}$  in the neighbourhood of  $\bar{A} = 1$ .

It is also clear that when  $d'$  is large enough, there will be a value  $\bar{A}_1$ , for which  $d' = d$ , and there is a unique solution. This happens when  $a = k$ , which gives

$$\bar{A}_1 = \left\{ (1+d)^2 + \frac{4d}{\omega} \right\} / (1-d)^2 \tag{14}$$

For instance, when  $d = 0.1, \omega = \frac{1}{2}, 1, 2, 3, 4$ :

$$\bar{A}_1 = \frac{1}{81}(201, 161, 141, 134\frac{1}{3}, 130) \text{ or } 2.481, 1.988, 1.740, 1.658, 1.605$$

and when  $d = \frac{1}{2}, \omega = \frac{1}{2}, 1, 2, 3, 4$ :

$$\bar{A}_1 = 25, 17, 13, 11\frac{2}{3}, 11.$$

Thus in the neighbourhood of  $\bar{A}_1$ , the two solutions for  $\sigma_v^2, \sigma_\lambda^2$  agree *without* transposal.

Thus we have two regions in which for different reasons, there is approximately a unique solution. Sometimes these two regions overlap forming one region.

Changes in the value of  $\omega$  for constant  $d$  and  $\bar{A}$  result in relatively little change in the ratio  $\sigma_v^2(d')/\sigma_\lambda^2(d')$ . This is particularly noticeable between  $\omega = 3$  and  $\omega = 4$ . The regions in which the two solutions are almost the same for one reason or the other (allowing a difference of about 10 per cent of  $\sigma^2$  in  $\sigma_v^2$  or  $\sigma_\lambda^2$ ) are:

	$d$				
	0.1	0.3	0.5	0.7	0.9
$\omega = 0.5$	$<0.1 \leq \bar{A} \leq 5$	$<0.1 \leq \bar{A} \leq 10+$	$<0.1 \leq \bar{A} \leq 4$ and around $\bar{A} = 25$	$<0.1 \leq \bar{A} \leq 4$ and around $\bar{A} = 22.7$	$<0.1 \leq \bar{A} \leq 8$ and around $\bar{A} = 253$
$\omega = 1$	$0.2 \leq \bar{A} \leq 4$	$0.3 \leq \bar{A} \leq 10+$	$0.1 \leq \bar{A} \leq 3$ and around $\bar{A} = 17$	$<0.1 \leq \bar{A} \leq 3$ and around $\bar{A} = 19.6$	$<0.1 \leq \bar{A} \leq 6$ and around $\bar{A} = 217$
$\omega = 2$	$0.5 \leq \bar{A} \leq 4$	$0.4 \leq \bar{A} \leq 9$	$0.2 \leq \bar{A} \leq 2$ $8 \leq \bar{A} \leq 13+$	$<0.1 \leq \bar{A} \leq 3$ and around $\bar{A} = 18.1$	$<0.1 \leq \bar{A} \leq 5$ and around $\bar{A} = 199$
$\omega = 3$	$0.6 \leq \bar{A} \leq 4$	$0.5 \leq \bar{A} \leq 8$	$0.4 \leq \bar{A} \leq 2$ $7 \leq \bar{A} \leq 11.6+$	$0.1 \leq \bar{A} \leq 2$ and around $\bar{A} = 17.592$	$<0.1 \leq \bar{A} \leq 4$ and around $\bar{A} = 195$
$\omega = 4$	$0.6 \leq \bar{A} \leq 3$	$0.6 \leq \bar{A} \leq 7$	$0.4 \leq \bar{A} \leq 2$ $6 \leq \bar{A} \leq 11+$	$0.1 \leq \bar{A} \leq 2$ and around $\bar{A} = 17.3$	$<0.1 \leq \bar{A} \leq 4$ and around $\bar{A} = 189$

For low values of  $d$ , (e.g. 0.1 and 0.3) the two regions overlap, forming one continuous region. When  $d = 0.5$  and  $\omega = 2, 3, 4$  the second region in which the two solutions are approximately the same, have a lower limit  $< 10$  for  $\bar{A}$ . When  $d = 0.7$  and  $d = 0.9$  the second region lies far beyond  $\bar{A} = 10$ . In these cases the exact value of  $\bar{A}_1$  (when  $d = d'$ ) has been given above. These values of  $\bar{A}$  are too high to be of much practical importance. What is regarded as approximate agreement is of course a subjective matter. The reader can form his own opinion by looking at the Tables. He will also notice from the tables that, when  $\bar{A}$  is small, the total systematic component forms a very small percentage of the total variance.

4. Thus there are considerable ranges of  $\bar{A}$ ,  $\omega$  and  $d$  in which two solutions are in approximate agreement. Nevertheless, it must be admitted that, in general, for given  $a, k, \rho$  there are two distinct solutions, obtained by interchanging  $a$  and  $k$ .

NOTE TO ACCOMPANY TABLE 4. The analyst of actual accident statistics will be able to get an estimate of  $\bar{A}$  and  $\omega = (\sigma_A^2 - \bar{A})/\bar{A}^2$  without difficulty. If he knows his subject matter reasonably well, he should sometimes be able to make at least a rough estimate of  $d = (\sigma_v^2 - \sigma_\lambda^2)/(\sigma_v^2 + \sigma_\lambda^2)$ . Let him assign  $\nu$  and  $\lambda$  to proneness and liability, so that  $d$  is positive. Then interpolation in the table will give him rough estimates of  $a, k, \rho$  and  $d'$ .

For complete information the values of  $a, k$  and  $d'$  for  $0 \leq \bar{A} < \infty, 0 \leq \omega < \infty$  and  $0 \leq d \leq 1$  are sufficient. Corresponding values of  $d$  and  $d'$  lie on the part of the rectangular hyperbola

$$(\bar{A} - 1)(1 + dd') - \{\bar{A} + 1 + (2/\omega)\}(d + d') = 0$$

which lies between the points (1, -1) and (-1, 1). The hyperbola becomes a straight line if  $\bar{A} = 1$ . It is clear from this expression that for given  $\bar{A}$  and  $\omega > 0$ , there is only one pair of corresponding values of  $d, d'$ . One member of the pair can be taken as positive; suppose this is  $d$ . Then (subject to interpolation) the table gives  $d'$ , which has an overall range of -1 to 1. There is no need to give the table for negative as well as positive values of  $d$ .

TABLE 4  
The two solutions for given  $\bar{A}$ ,  $d$ ,  $\omega$

		$d = 0.1$	$100\sigma_y^2/\sigma^2 = 55$	$100\sigma_x^2/\sigma^2 = 45$				
$\omega = \frac{1}{2}, a = 5.6$	$\bar{A} = 0.1$	0.2	0.5	1	1.6	2	5	10
$\sigma^2 = \omega \bar{A}^2$		0.005	0.02	0.125	0.5	1.38	2	12.5
$\sigma_{\bar{A}}^2$		0.105	0.22	0.625	1.5	3.05	4	17.5
$\sigma^2/\sigma_{\bar{A}}^2$ %		4.76	9.09	20	33.3	45.45	50	71.43
$k$		3.718	3.800	4.045	4.455	5	5.273	7.727
$\rho$		211.69	108.67	46.843	26.245	17	15.940	7.697
$d'$		-0.272	-0.250	-0.189	-0.100	0	0.044	0.312
$\sigma_y^2(d')/\sigma^2$ %		36.42	37.50	40.54	45	50	52.18	65.62
$\sigma_x^2(d')/\sigma^2$ %		63.58	62.50	59.46	55	50	47.82	34.38
$\omega = 1, a = 3.4$	$\bar{A} = 0.1$	0.2	0.5	1	1.4	2	5	10
$\sigma^2 = \omega \bar{A}^2$		0.01	0.04	0.25	1	1.96	4	25
$\sigma_{\bar{A}}^2$		0.11	0.24	0.75	2	3.36	6	30
$\sigma^2/\sigma_{\bar{A}}^2$ %		9.09	16.67	33.33	50	58.33	66.67	83.33
$k$		1.900	1.982	2.227	2.636	2.964	3.455	5.909
$\rho$		66.444	35.134	16.342	10.080	8.292	6.950	5.071
$d'$		-0.379	-0.342	-0.239	-0.100	0	0.102	0.421
$\sigma_y^2(d')/\sigma^2$ %		31.04	32.92	38.03	45	50	55.10	71.06
$\sigma_x^2(d')/\sigma^2$ %		68.96	67.08	61.97	55	50	44.90	28.94
$\omega = 2, a = 2.3$	$\bar{A} = 0.1$	0.2	0.5	1	1.3	2	5	10
$\sigma^2 = \omega \bar{A}^2$		0.02	0.08	0.50	2	3.5	8	50
$\sigma_{\bar{A}}^2$		0.12	0.28	1	3	4.8	10	55
$\sigma^2/\sigma_{\bar{A}}^2$ %		16.67	28.57	50	66.67	72.73	80	90.91
$k$		0.991	1.073	1.318	1.727	2	2.545	5
$\rho$		24.123	13.518	7.151	5.030	4.500	3.969	3.333
$d'$		-0.507	-0.447	-0.294	-0.100	0	0.154	0.500
$\sigma_y^2(d')/\sigma^2$ %		24.66	27.63	35.30	45	50	57.69	75
$\sigma_x^2(d')/\sigma^2$ %		75.34	72.37	64.70	55	50	42.31	25
$\omega = 3, a = 1.963$	$\bar{A} = 0.1$	0.2	0.5	1	1.296	2	5	10
$\sigma^2 = \omega \bar{A}^2$		0.03	0.12	0.75	3	5.041	12	75
$\sigma_{\bar{A}}^2$		0.13	0.32	1.25	4	6.337	14	80
$\sigma^2/\sigma_{\bar{A}}^2$ %		23.08	37.50	60	75	79.55	85.71	93.75
$k$		0.688	0.770	1.015	1.424	1.667	2.242	4.697
$\rho$		14.505	8.558	4.985	3.795	3.524	3.201	2.844
$d'$		-0.580	-0.507	-0.323	-0.100	0	0.216	0.532
$\sigma_y^2(d')/\sigma^2$ %		21	24.66	33.84	45	50	60.80	76.60
$\sigma_x^2(d')/\sigma^2$ %		79	75.34	66.16	55	50	39.20	23.40
$\omega = 4, a = 1.7$	$\bar{A} = 0.1$	0.2	0.5	1	1.27	2	5	10
$\sigma^2 = \omega \bar{A}^2$		0.04	0.16	1	4	6.53	16	100
$\sigma_{\bar{A}}^2$		0.14	0.36	1.50	5	7.81	18	105
$\sigma^2/\sigma_{\bar{A}}^2$ %		28.57	44.44	66.67	80	83.63	88.89	95.24
$k$		0.536	0.618	0.864	1.273	1.500	2.091	4.545
$\rho$		10.529	6.493	4.072	3.263	3.087	2.859	2.616
$d'$		-0.627	-0.545	-0.342	-0.100	0	0.191	0.549
$\sigma_y^2(d')/\sigma^2$ %		18.64	22.76	32.92	45	50	59.56	77.46
$\sigma_x^2(d')/\sigma^2$ %		81.36	77.24	67.08	55	50	40.44	22.54

TABLE 4—continued

	$d = 0.03$	$100\sigma_v^2/\sigma^2 = 65$	$100\sigma_\lambda^2/\sigma^2 = 35$					
$\omega = \frac{1}{2}, a = 7.571$	$\bar{A} = 0.1$	0.2	0.5	1	2	3.571	5	10
$\sigma^2 = \omega \bar{A}^2$	0.005	0.02	0.125	0.5	2	4.592	12.5	50
$\sigma_A^2$	0.105	0.22	0.625	1.5	4	8.163	17.5	60
$\sigma^2/\sigma_A^2$ %	4.76	9.09	20	33.33	50	56.25	71.43	83.33
$k$	3.131	3.185	3.346	3.615	4.154	5	5.769	8.462
$\rho$	238.05	121.56	51.671	28.374	16.725	11.600	9.736	7.407
$d'$	-0.452	-0.434	-0.380	-0.300	-0.164	0	0.114	0.366
$\sigma_v^2(d')/\sigma^2$ %	27.38	28.31	30.98	35	41.79	50	55.68	68.30
$\sigma_\lambda^2(d')/\sigma^2$ %	72.62	71.69	69.02	65	58.21	50	44.32	31.70
$\omega = 1, a = 4.714$	$\bar{A} = 0.1$	0.2	0.5	1	2	2.714	5	10
$\sigma^2 = \omega \bar{A}^2$	0.01	0.04	0.25	1	4	7.367	25	100
$\sigma_A^2$	0.11	0.24	0.75	2	6	10.081	30	110
$\sigma^2/\sigma_A^2$ %	9.09	16.67	33.33	50	66.67	73.08	83.33	90.91
$k$	1.592	1.646	1.808	2.077	2.615	3	4.231	6.923
$\rho$	76.066	39.799	18.047	10.792	7.164	6.211	4.989	4.264
$d'$	-0.543	-0.512	-0.425	-0.300	-0.106	0	0.235	0.495
$\sigma_v^2(d')/\sigma^2$ %	22.85	24.42	28.76	35	44.68	50	61.76	74.76
$\sigma_\lambda^2(d')/\sigma^2$ %	77.15	75.58	71.24	65	55.32	50	38.24	25.24
$\omega = 2, a = 3.286$	$\bar{A} = 0.1$	0.2	0.5	1	2	2.286	5	10
$\sigma^2 = \omega \bar{A}^2$	0.02	0.08	0.50	2	8	10.45	50	200
$\sigma_A^2$	0.12	0.28	1.00	3	10	12.73	55	210
$\sigma^2/\sigma_A^2$ %	16.67	28.57	50	66.67	80	82.05	90.91	95.24
$k$	0.823	0.877	1.038	1.308	1.846	2	3.462	6.154
$\rho$	28.041	15.408	7.821	5.298	4.033	3.875	3.275	3.022
$d'$	-0.646	-0.598	-0.472	-0.300	-0.054	0	6.328	0.581
$\sigma_v^2(d')/\sigma^2$ %	17.72	20.08	26.42	35	47.30	50	66.38	79.03
$\sigma_\lambda^2(d')/\sigma^2$ %	82.28	79.92	73.58	65	52.70	50	33.62	20.97
$\omega = 3, a = 2.810$	$\bar{A} = 0.1$	0.2	0.5	1	2	2.143	5	10
$\sigma^2 = \omega \bar{A}^2$	0.03	0.12	0.75	3	12	13.775	75	300
$\sigma_A^2$	0.13	0.32	1.25	4	14	15.918	80	310
$\sigma^2/\sigma_A^2$ %	23.08	37.50	60	75	85.71	86.54	93.75	96.77
$k$	0.567	0.621	0.782	1.051	1.590	1.667	3.205	5.897
$\rho$	33.921	9.724	5.394	3.953	3.234	3.185	2.801	2.657
$d'$	-0.702	-0.646	-0.496	-0.300	-0.030	0	0.366	0.613
$\sigma_v^2(d')/\sigma^2$ %	14.90	17.72	25.18	35	48.52	50	68.30	80.67
$\sigma_\lambda^2(d')/\sigma^2$ %	85.10	82.28	74.82	65	51.48	50	31.70	19.33
$\omega = 4, a = 2.571$	$\bar{A} = 0.1$	0.2	0.5	1	2	2.071	5	10
$\sigma^2 = \omega \bar{A}^2$	0.04	0.16	1	4	16	17.163	100	400
$\sigma_A^2$	0.14	0.36	1.50	5	18	19.234	105	410
$\sigma^2/\sigma_A^2$ %	28.57	44.44	66.67	80	88.89	89.23	95.24	97.56
$k$	0.438	0.492	0.654	0.923	1.462	1.500	3.077	5.769
$\rho$	12.263	7.326	4.363	3.373	2.880	2.862	2.582	2.483
$d'$	-0.738	-0.675	-0.512	-0.300	-0.016	0	0.387	0.631
$\sigma_v^2(d')/\sigma^2$ %	13.10	16.24	24.42	35	49.22	50	69.34	81.54
$\sigma_\lambda^2(d')/\sigma^2$ %	86.90	83.76	75.58	65	50.78	50	30.66	18.46

TABLE 4—continued

	$d = 0.5$	$100\sigma_v^2/\sigma^2 = 75$	$100\sigma_\lambda^2/\sigma^2 = 25$					
$\omega = \frac{1}{2}, a = 11$	$\bar{A} = 0.1$	0.2	0.5	1	2	5	7	10
$\sigma^2 = \omega \bar{A}^2$	0.005	0.02	0.125	0.5	2	12.5	24.5	50
$\sigma_A^2$	0.105	0.22	0.625	1.5	4	17.5	31.5	60
$\sigma^2/\sigma_A^2 \%$	4.76	9.09	20	33.33	50	71.43	77.78	83.33
$k$	2.700	2.733	2.833	3	3.333	4.333	5	6
$\rho$	298	151.33	63.333	34	19.333	10.533	8.857	7.600
$d'$	-0.622	-0.607	-0.565	-0.500	-0.385	-0.125	0	0.143
$\sigma_v^2(d')/\sigma^2 \%$	18.90	19.65	21.75	25	30.75	43.75	50	57.15
$\sigma_\lambda^2(d')/\sigma^2 \%$	81.10	80.35	78.26	75	69.30	56.25	50	42.85
$\omega = 1, a = 7$	$\bar{A} = 0.1$	0.2	0.5	1	2	5		10
$\sigma^2 = \omega \bar{A}^2$	0.01	0.04	0.25	1	4	25		100
$\sigma_A^2$	0.11	0.24	0.75	2	6	30		110
$\sigma^2/\sigma_A^2 \%$	9.09	16.67	33.33	50	66.67	83.33		90.91
$k$	1.367	1.400	1.500	1.667	2	3		4.667
$\rho$	96.667	50	22	12.667	8	5.200		3.267
$d'$	-0.690	-0.667	-0.600	-0.500	-0.333	0		0.294
$\sigma_v^2(d')/\sigma^2 \%$	15.50	16.67	20	25.00	33.33	50		64.70
$\sigma_\lambda^2(d')/\sigma^2 \%$	84.50	83.33	80	75.00	66.67	50		35.30
$\omega = 2, a = 5$	$\bar{A} = 0.1$	0.2	0.5	1	2	4	5	10
$\sigma^2 = \omega \bar{A}^2$	0.02	0.08	0.50	2	8	32	50	200
$\sigma_A^2$	0.12	0.28	1.00	3	10	36	55	210
$\sigma^2/\sigma_A^2 \%$	16.67	28.57	50	66.67	80	88.89	90.91	95.24
$k$	0.700	0.733	0.833	1	1.333	2	2.333	4
$\rho$	36	19.333	9.333	6	4.333	3.500	3.333	3
$d'$	-0.765	-0.731	-0.636	-0.500	-0.286	0	0.100	0.400
$\sigma_v^2(d')/\sigma^2 \%$	11.76	13.46	18.18	25	35.72	50	55	70
$\sigma_\lambda^2(d')/\sigma^2 \%$	88.24	86.54	81.82	75	64.28	50	45	30
$\omega = 3, a = 4.3$	$\bar{A} = 0.1$	0.2	0.5	1	2	3.667	5	10
$\sigma^2 = \omega \bar{A}^2$	0.03	0.12	0.75	3	12	40.333	75	300
$\sigma_A^2$	0.13	0.32	1.25	4	14	44	80	310
$\sigma^2/\sigma_A^2 \%$	23.08	37.50	60	75	85.71	91.67	93.75	96.77
$k$	0.478	0.511	0.611	0.778	1.111	1.667	2.111	3.778
$\rho$	21.704	11.074	6.296	4.370	3.407	2.978	2.830	2.637
$d'$	-0.804	-0.765	-0.655	-0.500	-0.263	0	0.143	0.442
$\sigma_v^2(d')/\sigma^2 \%$	9.78	11.76	17.24	25	36.84	50	57.14	72.10
$\sigma_\lambda^2(d')/\sigma^2 \%$	90.22	88.24	82.76	75	63.16	50	42.86	27.90
$\omega = 4, a = 4$	$\bar{A} = 0.1$	0.2	0.5	1	2	3.5	5	10
$\sigma^2 = \omega \bar{A}^2$	0.04	0.16	1	4	16	49	100	400
$\sigma_A^2$	0.14	0.36	1.50	5	18	52.5	105	410
$\sigma^2/\sigma_A^2 \%$								
$k$	0.367	0.400	0.500	0.667	1	1.500	2	3.667
$\rho$	15.667	9	5	3.667	3	2.714	2.600	2.467
$d'$	-0.829	-0.786	-0.667	-0.500	-0.250	0	0.167	0.464
$\sigma_v^2(d')/\sigma^2 \%$	8.54	10.72	16.67	25	37.50	50	58.33	73.22
$\sigma_\lambda^2(d')/\sigma^2 \%$	91.46	89.28	83.33	75	62.50	50	41.67	26.78



TABLE 4—continued

	$d = 0.7$	$100\sigma_v^2/\sigma^2 = 85$	$100\sigma_\lambda^2/\sigma^2 = 15$					
$\omega = \frac{1}{2}, a = 19$	$\bar{A} = 0.1$	0.2	0.5	1	2	5	10	15
$\sigma^2 = \omega\bar{A}^2$	0.005	0.02	0.125	0.5	2	12.5	50	112.5
$\sigma_A^2$	0.105	0.22	0.625	1.5	4	17.5	60	127.5
$\sigma^2/\sigma_A^2$ %	4.76	9.09	20	33.33	50	71.43	83.33	88.24
$k$	2.371	2.388	2.441	2.529	2.706	3.235	4.118	5
$\rho$	451.49	227.86	93.758	49.051	26.707	13.293	8.824	7.333
$d'$	-0.780	-0.771	-0.744	-0.700	-0.619	-0.417	-0.172	0
$\sigma_v^2(d')/\sigma^2$ %	11	11.46	12.82	15	19.05	29.17	41.38	50
$\sigma_\lambda^2(d')/\sigma^2$ %	89	88.54	87.18	85	80.95	70.83	58.62	50
$\omega = 1, a = 12.3$	$\bar{A} = 0.1$	0.2	0.5	1	2	5	10	10.3
$\sigma^2 = \omega\bar{A}^2$	0.01	0.04	0.25	1	4	25	100	106.778
$\sigma_A^2$	0.11	0.24	0.75	2	6	30	110	117.111
$\sigma^2/\sigma_A^2$ %	9.09	16.67	33.33	50	66.67	83.33	90.91	91.18
$k$	1.194	1.212	1.265	1.353	1.529	2.059	2.941	3
$\rho$	148.26	75.740	32.203	17.687	10.429	3.612	4.627	3.581
$d'$	-0.823	-0.808	-0.766	-0.700	-0.581	-0.308	-0.015	0
$\sigma_v^2(d')/\sigma^2$ %	8.84	9.58	11.69	15	20.93	34.62	49.26	50
$\sigma_\lambda^2(d')/\sigma^2$ %	91.16	90.42	88.31	85	79.07	65.38	50.74	50
$\omega = 2, a = 9$	$\bar{A} = 0.1$	0.2	0.5	1	2	5	8	10
$\sigma^2 = \omega\bar{A}^2$	0.02	0.08	0.50	2	8	50	128	200
$\sigma_A^2$	0.12	0.28	1	3	10	55	136	210
$\sigma^2/\sigma_A^2$ %	16.67	28.57	50	66.67	80	90.91	94.12	95.24
$k$	0.606	0.624	0.676	0.765	0.941	1.471	2	2.353
$\rho$	55.540	29.080	13.168	7.885	5.234	3.648	2.250	3.118
$d'$	-0.868	-0.848	-0.790	-0.700	-0.546	-0.214	0	0.105
$\sigma_v^2(d')/\sigma^2$ %	6.60	7.61	10.52	15	22.72	39.28	50	55.26
$\sigma_\lambda^2(d')/\sigma^2$ %	93.40	92.39	89.48	85	77.28	60.72	50	44.74
$\omega = 3, a = 7.8$	$\bar{A} = 0.1$	0.2	0.5	1	2	5	7.222	10
$\sigma^2 = \omega\bar{A}^2$	0.03	0.12	0.78	3	12	75	156.48	300
$\sigma_A^2$	0.13	0.32	1.25	4	14	80	163.70	310
$\sigma^2/\sigma_A^2$ %	23.08	37.50	60	75	85.71	93.75	95.59	96.77
$k$	0.410	0.427	0.480	0.569	0.745	1.275	1.667	2.157
$\rho$	33.344	17.843	8.573	5.489	3.939	3.012	2.821	2.702
$d'$	-0.892	-0.868	-0.801	-0.700	-0.528	-0.172	0	0.155
$\sigma_v^2(d')/\sigma^2$ %	5.42	6.60	9.94	15	23.60	41.38	50	57.76
$\sigma_\lambda^2(d')/\sigma^2$ %	94.58	93.40	90.06	85	76.40	58.62	50	42.24
$\omega = 4, a = 7.3$	$\bar{A} = 0.1$	0.2	0.5	1	2	5	6.833	10
$\sigma^2 = \omega\bar{A}^2$	0.04	0.16	1	4	16	100	186.778	400
$\sigma_A^2$	0.14	0.36	1.50	5	18	105	193.611	410
$\sigma^2/\sigma_A^2$ %	28.57	44.44	66.67	80	88.89	95.24	96.47	97.56
$k$	0.312	0.329	0.382	0.471	0.647	1.176	1.500	2.059
$\rho$	23.880	13.063	6.603	4.454	3.372	2.725	2.610	2.510
$d'$	-0.906	-0.880	-0.808	-0.700	-0.518	-0.149	0	0.183
$\sigma_v^2(d')/\sigma^2$ %	4.71	5.98	9.58	15	24.10	42.57	50	59.14
$\sigma_\lambda^2(d')/\sigma^2$ %	95.29	94.02	90.42	85	75.90	57.43	50	40.86

TABLE 4—continued

	$d = 0.9$	$100\sigma_v^2/\sigma^2 = 95$	$100\sigma_\lambda^2/\sigma^2 = 5$					
$\omega = \frac{1}{2}, a = 59$	$\bar{A} = 0.1$	0.2	0.5	1	2	5	10	55
$\sigma^2 = \omega \bar{A}^2$	0.005	0.02	0.125	0.5	2	12.5	50	1512.5
$\sigma_A^2$	0.105	0.22	0.625	1.5	4	17.5	60	1567.5
$\sigma^2/\sigma_A^2$ %	4.76	9.09	20	33.33	50	71.43	83.33	96.50
$k$	2.111	2.116	2.132	2.158	2.211	2.368	2.632	5
$\rho$	1246.49	625.22	252.58	128.32	66.225	28.942	16.529	6.364
$d'$	-0.929	-0.926	-0.916	-0.900	-0.869	-0.781	-0.652	0
$\sigma_v^2(d')/\sigma^2$ %	3.56	3.72	4.20	5	6.56	10.94	17.39	50
$\sigma_\lambda^2(d')/\sigma^2$ %	96.44	96.28	95.80	95	93.44	89.06	82.61	50
$\omega = 1, a = 39$	$\bar{A} = 0.1$	0.2	0.5	1	2	5	10	37
$\sigma^2 = \omega \bar{A}^2$	0.01	0.04	0.25	1	4	25	100	1369
$\sigma_A^2$	0.11	0.24	0.75	1	6	30	110	1406
$\sigma^2/\sigma_A^2$ %	9.09	16.67	33.33	50	66.67	83.33	90.91	97.37
$k$	1.058	1.063	1.079	1.105	1.158	1.316	1.579	3
$\rho$	413.62	208.28	85.162	44.095	23.581	11.265	7.158	4.162
$d'$	-0.944	-0.939	-0.924	-0.900	-0.854	-0.727	-0.551	0
$\sigma_v^2(d')/\sigma^2$ %	2.82	3.06	3.80	5	7.32	13.64	22.45	50
$\sigma_\lambda^2(d')/\sigma^2$ %	97.18	96.94	96.20	95	92.68	86.36	77.55	50
$\omega = 2, a = 29$	$\bar{A} = 0.1$	0.2	0.5	1	2	5	10	28
$\sigma^2 = \omega \bar{A}^2$	0.02	0.08	0.50	2	8	50	200	1568
$\sigma_A^2$	0.12	0.28	1	3	10	55	210	1596
$\sigma^2/\sigma_A^2$ %	16.67	28.57	50	66.67	80	90.91	95.24	98.25
$k$	0.532	0.537	0.553	0.579	0.632	0.790	1.053	2
$\rho$	155.16	78.836	33.051	17.788	10.158	5.579	4.053	3.071
$d'$	-0.959	-0.952	-0.932	-0.900	-0.839	-0.676	-0.462	0
$\sigma_v^2(d')/\sigma^2$ %	2.06	2.40	3.39	5	8.06	16.18	26.92	50
$\sigma_\lambda^2(d')/\sigma^2$ %	97.94	97.60	96.61	95	91.94	83.82	73.08	50
$\omega = 3, a = 25.6$	$\bar{A} = 0.1$	0.2	0.5	1	2	5	10	25
$\sigma^2 = \omega \bar{A}^2$	0.03	0.12	0.75	3	12	75	300	1875
$\sigma_A^2$	0.13	0.32	1.25	4	14	80	310	1900
$\sigma^2/\sigma_A^2$ %	23.08	37.50	60	75	85.71	93.78	96.77	98.68
$k$	0.356	0.361	0.377	0.404	0.456	0.614	0.877	1.667
$\rho$	92.399	47.380	20.363	11.357	6.853	4.152	3.251	2.711
$d'$	-0.966	-0.959	-0.936	-0.900	-0.831	-0.652	-0.421	0
$\sigma_v^2(d')/\sigma^2$ %	1.68	2.06	3.18	5	8.44	17.39	28.97	50
$\sigma_\lambda^2(d')/\sigma^2$ %	98.32	97.94	96.82	95	91.56	82.61	71.03	50
$\omega = 4, a = 24$	$\bar{A} = 0.1$	0.2	0.5	1	2	5	10	23.5
$\sigma^2 = \omega \bar{A}^2$	0.04	0.16	1	4	16	100	400	2209
$\sigma_A^2$	0.14	0.36	1.50	5	18	105	410	2232.5
$\sigma^2/\sigma_A^2$ %	28.57	44.44	66.67	80	88.89	95.24	97.56	98.95
$k$	0.268	0.274	0.290	0.316	0.368	0.526	0.790	1.500
$\rho$	65.416	33.844	14.896	8.579	5.421	3.526	2.895	2.532
$d'$	-0.971	-0.963	-0.939	-0.900	-0.827	-0.638	-0.397	0
$\sigma_v^2(d')/\sigma^2$ %	1.45	1.86	3.06	5	8.66	18.10	30.14	50
$\sigma_\lambda^2(d')/\sigma^2$ %	98.55	98.14	96.94	95	91.34	81.90	69.86	50

The statistician will usually know from studying the data in various ways whether the proneness or the liability component should be the greater, and in certain ranges of  $\bar{A}$  particularly in the neighbourhood of  $\bar{A} = 1$ , it will not matter greatly which solution he chooses. However, if the model is valid (which of course involves the assumption that one or other of the components *does* represent accident proneness or liability), the following method should, given good enough data, enable him to decide which solution is the correct one. He can divide up the exposure period into several sub-periods and fit the Waring distribution separately to the first sub-period, the first two sub-periods and so on, cumulatively. As these cumulative sub-periods increase  $\bar{A}$  will increase. This process will give him for each cumulative sub-period and the whole period the values of  $d$  and  $d'$ . For one of the two choices of  $a$  and  $k$ ,  $d$  will be positive. Take  $\sigma_p^2(d) = \frac{1}{2}(1+d)$ ,  $\sigma_\lambda^2(d) = \frac{1}{2}(1-d)$  for this choice; and then find the corresponding values of  $d'$ . If  $d$  is the correct solution,  $d$  should, apart from sampling errors, remain constant while  $d'$  should increase—and vice versa.

The problems of estimation will not be considered here, but a few remarks may not be out of place. The sampling errors of this method, for some of the studies made in the past, might well be considerable. It is conjectured that it would be better to make estimates of  $\bar{A}$ , and  $\omega$  from the estimates of  $a, k, \rho$  by using the formulae given in Part I, Table 2, rather than from the observed  $\bar{A}$  and  $\sigma_{\bar{A}}^2$ , where the two sets differ.  $d$  can be estimated from equation (10) or (11) (and the results will agree) and  $d'$  from equation (11). Judging from experience of accident data in the past, it should be reasonably efficient to estimate  $a, k, \rho$  from the first three factorial moments. This will be computationally more convenient than using ordinary central moments, but will give the same answer. The maximum likelihood solution could, if necessary, be examined.

5. Certain limiting forms of the Waring distribution are of interest in connection with the model considered.

If  $k \rightarrow \infty$  and  $\rho \rightarrow \infty$  so that  $q_k = k/(k + \rho)$  remains constant the generalized Waring tends to the negative binomial

$$\left\{ \frac{1 - q_k \theta}{p_k - p_k} \right\}^{-a} \tag{15}$$

In the case considered here

$$\left. \begin{aligned} a+k &= \omega \bar{A}(\rho-2) - (\bar{A}+1), \\ ak &= \bar{A}(\rho-1), \\ a-k &= d\omega \bar{A}(\rho-1) - (\bar{A}-1). \end{aligned} \right\} \tag{16}$$

If  $\rho \rightarrow \infty$ ,  $\bar{A}$  remaining finite then

$$\left. \begin{aligned} (a+k) &\sim \omega \bar{A} \rho & \text{(i),} \\ (a-k) &\sim d\omega \bar{A} \rho & \text{(ii),} \\ (ak) &\sim \bar{A} \rho & \text{(iii).} \end{aligned} \right\} \tag{17}$$

Thus either  $a \rightarrow \infty$  or  $k \rightarrow \infty$ . Suppose  $k \rightarrow \infty$  then from (i) and (iii)  $a = 1/\omega$  and

$$q_k = \frac{k}{k + \rho} \sim \frac{\omega \bar{A}}{\omega \bar{A} + 1} \quad \text{and} \quad p_k \sim \frac{1}{\omega \bar{A} + 1} \tag{18}$$

So for large  $\rho$  the Waring distribution takes the negative binomial form

$$(\omega \bar{A} + 1 - \omega \bar{A} \theta)^{-1/\omega} \tag{19}$$

In this case, from (ii),  $k \sim -d\omega \bar{A} \rho$ ; so we must have  $d \rightarrow -1$ . Conversely if  $d \rightarrow -1$ , from (ii)  $k \rightarrow \infty$ . Similarly if and only if  $d \rightarrow +1$ ,  $a \rightarrow \infty$  and the Waring distribution becomes the negative binomial.

If  $\bar{A} \rightarrow \infty$  we have

$$\begin{aligned} k &\rightarrow \infty \\ a &\rightarrow \frac{1+d+(2/\omega)}{1-d}, \\ \rho &\rightarrow \frac{2+d+(2/\omega)}{1+d}. \end{aligned} \tag{20}$$

So the distribution becomes, formally,  $C(1-\theta)^{-a}$  where ultimately  $C$  is an infinite constant. We may write this  $(K+1-K\theta)^{-a}$  where  $K \rightarrow \infty$  and all terms in this distribution  $\rightarrow 0$ . So the distribution has an infinitely long tail.

Since  $aK = \bar{A}$ , the limiting form of the distribution for very large  $\bar{A}$  is

$$\left(1 + \frac{\bar{A}}{a} - \frac{\bar{A}\theta}{a}\right)^{-a}$$

with variance

$$\bar{A} + \frac{\bar{A}^2}{a}$$

where

$$a = \left(1 + d + \frac{2}{\omega}\right) / (1 + d).$$

There are certain limiting forms as  $\bar{A} \rightarrow 0$ , but as no accidents can give no information on susceptibility, these are not discussed here.

6. Even if the Waring distribution fitted the data, the model might under certain circumstances fail. Every frequency in the Waring distribution is a function of  $(a+k)$  and  $ak$ . So any reasonable method of estimation must in fact estimate  $(a+k)$  and  $ak$ ;  $a$  and  $k$  are then obtained from a quadratic equation. Conceivably the estimate of  $a$  or  $k$  or both might turn out to be real and negative. We may dismiss this case, because with one exception, even it could possibly occur, some of the fitted frequencies would be infinite or negative and the distribution could not be said to fit the data. The exception is when  $a = k = -\gamma$  where  $\gamma > 0$ . In this case the series converges, (for  $\rho = \{1 + \bar{A}/ak\} > 0$ ) but on our model we should have

$$d = d' = \frac{a-1-(2/\omega)}{a+1} = 0$$

which gives  $a = (1+2/\omega) > 0$ , a contradiction. So here too the model would fail.

But the estimates of  $a$  and  $k$  might turn out to be complex conjugates  $a = \alpha + i\beta$  and  $k = \alpha - i\beta$ . In this case every frequency in the Waring distribution is real and positive and the series still converges, for  $\rho$  is still real and  $> 0$ .

The properties of this generalization of the Waring distribution have not so far been studied. Its continuous analogue is presumably Pearson Type IV; but if it fitted the data, the model would fail because  $d, d', \sigma_\lambda^2, \sigma_\nu^2$  would be complex quantities.

I do not know whether, for the Generalized Waring Distribution, the sub-division which I have given of  $\sigma^2$  into two independent components is unique. One margin of the  $(\nu, \lambda)$  distribution is assumed to be Type VI and the arrays Type III. Suppose, for definiteness, that this is the  $\nu$  margin; the other margin is then of more complex form. Other sub-divisions might be possible but then the distribution of  $\lambda \cdot \nu + \epsilon$  would not, I think, be negative binomial.

Again, the  $\nu$  distribution might have any form and so might the  $\lambda \cdot \nu$  distribution, the other margin ( $\lambda$  distribution) would then be determined. The resulting distribution of accidents would not (in general) then have the Waring distribution form and the distribution of  $\lambda \cdot \nu + \epsilon$  would not in general be negative binomial.

7. It may be emphasized once more that accident proneness means the sum total of the *internal* factors in the individual, predisposing towards accidents, *liability* on the other hand means the sum total of external factors. It does not by any means follow that an individual's *proneness* remains the same for the whole of his lifetime. This is a matter for research and analysis of accident records. Some writers have supposed that the term *proneness* implies constancy throughout life and criticized it accordingly. This seems to me like trying to knock down a skittle which they have erected themselves. I have said elsewhere (Irwin, 1964) that I believe that *proneness* does play a small but important part in the causation of accidents.

Anyone who believes in the classical accident theory at all will admit that  $\sigma_\lambda^2 + \sigma_\nu^2$  represents susceptibility but the mathematics alone cannot tell us that  $\sigma_\lambda^2$  represents proneness and  $\sigma_\nu^2$  liability (or vice versa). By his knowledge of the survey and by methods of analysis involving sub-division of his subjects into sub-groups with different exposure conditions (e.g. exposure time, heating, lighting, ventilation, kind of task performed, danger of the machinery used etc.) the analyst may be able to get an independent estimate of the *liability* component. No doubt it would be harder to get an independent estimate of the *proneness* component but maybe this could sometimes be done by using aesthetokinetic tests (dotting, reaction times etc.) such as Farmer and Chambers (1939) used. If one component can be independently estimated, the other follows by subtraction.

In this paper a new method has been suggested which envisages the possibility of estimating the importance of accident proneness and liability separately, in the study of individual accidents. It is little more than a survey of the field. However, it is a field which seems to me worth exploration by young and energetic researchers. In particular the method should be tried out with bodies of data as extensive as possible—for which I regret I have not the time or facilities. It is with these hopes that I am publishing this study.

#### ACKNOWLEDGEMENT

I am indebted to the Medical Research Council for a grant in support of this and other studies. I must express my gratitude to Professor C. A. B. Smith for arranging that I could do this work in the Galton Laboratory and to Professor H. Harris for his kindness in granting these facilities; also to Miss Irene Allen for computing the final table.

#### REFERENCES

- CHAMBERS, E. G. and YULE, G. U. (1941). Theory and observation in the investigation of accident causation. *J.R. Statist. Soc. Suppl.*, 7, 101–7.  
 FARMER, E. and CHAMBERS, E. G. *Industrial Health Board Report*, No. 84.  
 FORSYTH, A. R. (1914). *A Treatise on Differential Equations*. Chap. VII. London: Macmillan.

- GREENWOOD, M. and YULE, G. U. (1920). An enquiry into the nature of frequency distributions representative of multiple happenings, with special reference to multiple attacks of disease or repeated accidents. *J.R. Statist. Soc.*, **83**, 255–279.
- IRWIN, J. O. (1963). The place of mathematics in medical and biological statistics. *J.R. Statist. Soc.*, A **126**, 1–44. Appendix II.
- (1964). The personal factor in accidents. *J.R. Statist. Soc.*, A **127**, 438–451.
- (1966). Inverse factorial series as frequency distributions. *Classical and Contagious Discrete Distributions. Proc. Int. Symp. Montreal, 1963*. Calcutta: Statistical Publications Co. Distributed by Pergamon Press.
- NEWBOLD, E. M. (1925). A contribution to the study of the human factor in the causation of accidents. *Industrial Health Research Board Report*, No. 34.
- (1927). Applications of the statistics of repeated events, particularly to industrial accidents. *J.R. Statist. Soc.*, **90**, 487–547.
- WHITTAKER, E. T. and WATSON, G. N. (1935). *Modern Analysis* (4th ed.), chap. XIV. Cambridge: University Press.
- YULE, G. U. (1924). A mathematical theory of evolution based on the conclusions of Dr J. C. Willis, F.R.S. *Philos. Trans. B* **243**, 21–87.
-