

The Bivariate Generalized Waring Distribution and its Application to Accident Theory

BY

EVDOKIA XEKALAKI

Reprinted from

THE JOURNAL OF THE ROYAL STATISTICAL SOCIETY

SERIES A (GENERAL)

Volume 147, Part 3, 1984

(pp. 488–498)



PRINTED FOR PRIVATE CIRCULATION

1984

The Bivariate Generalized Waring Distribution and its Application to Accident Theory

By EVDOKIA XEKALAKI

University of Missouri, Columbia, USA

SUMMARY

The univariate generalized Waring distribution was shown by Irwin (1968, 1975) to provide a useful accident model which enables one to split the variance into three additive components due to randomness, proneness and liability. The two non-random variance components, however, cannot be separately estimated.

In this paper a way of tackling this problem is suggested by defining a bivariate extension of the generalized Waring distribution. Using this it is possible to obtain distinguishable estimates for the variance components and hence inferences can be made about the role of the underlying accident factors. The technique is illustrated by two examples.

Keywords: BIVARIATE GENERALIZED WARING DISTRIBUTION; ACCIDENT THEORY; PRONENESS; LIABILITY; WARING'S EXPANSION

1. INTRODUCTION

The phenomenon of accident causation has always attracted much interest and various hypotheses have been developed towards its interpretation. Among them the idea of accident proneness has stimulated much interesting statistical theory and method. One important contribution in this direction is Irwin's (1968) "proneness-liability" model giving rise to a three parameter discrete distribution, the univariate generalized Waring distribution (UGWD) with probability generating function (p.g.f.)

$$G(s) = \frac{\rho(k)}{(a + \rho)_{(k)}} {}_2F_1(a, k; a + k + \rho; s).$$

Here ${}_2F_1$ denotes the Gauss hypergeometric series defined by

$${}_2F_1(\alpha, \beta; \gamma; z) = \sum_{r=0}^{\infty} \frac{\alpha_{(r)}\beta_{(r)}}{\gamma_{(r)}} \frac{z^r}{r!}$$

and $h_{(l)} = \Gamma(h + l)/\Gamma(h)$, $h > 0$, $l \in R$. (For a more detailed account regarding the structure, models for, properties, extensions and applications of this distribution see Irwin (1963, 1975), Xekalaki (1981, 1983a–g, 1984a) and Xekalaki and Panaretos (1983).)

Irwin's accident model assumes that all non-random factors can be further split into psychological and external factors. So, the term "accident proneness" (denoted by ν) refers to a person's idiosyncratic predisposition to accidents and the term "accident liability" (denoted by $\lambda | \nu$, i.e. λ for given ν) refers to a person's exposure to external risk of accident. The UGWD then arises from a Poisson distribution of accidents with p.g.f. $\exp\{(\lambda | \nu)(s - 1)\}$ for individuals with proneness ν and liability $\lambda | \nu$ when $\lambda | \nu$ has a gamma (Pearson Type III) distribution and ν has a beta distribution of the second kind (Pearson Type VI).

Present address: University of Crete, Department of Mathematics, Iraklio, Crete, Greece.

The distribution was applied by Irwin (1968, 1975) to data on accidents sustained by men in a soap factory, providing an improved fit as compared to the negative binomial. But this fact alone is not very important. The innovation brought by this model in accident theory does not lie in the better fit but in the possibility of partitioning the total variance (σ^2) into three additive components due to proneness (σ_ν^2), liability (σ_λ^2) and randomness (σ_R^2) thus†,

$$\sigma^2 = \sigma_\lambda^2 + k^2 \sigma_\nu^2 + \sigma_R^2, \quad (1.1)$$

where

$$\begin{aligned} \sigma_\lambda^2 &= ak(a+1)(\rho-1)^{-1}(\rho-2)^{-1}, & \sigma_\nu^2 &= a(a+\rho-1)(\rho-1)^{-2}(\rho-2)^{-1}, \\ \sigma_R^2 &= ak(\rho-1)^{-1}, & \sigma^2 &= ak(a+\rho-1)(k+\rho-1)(\rho-1)^{-2}(\rho-2)^{-1}. \end{aligned} \quad (1.2)$$

By estimating the parameters of the distribution and hence the variance components, one is in a position to draw conclusions regarding the contribution of each factor to the given accident situation.

There is, however, a problem arising from the fact that the UGWD is symmetrical in a, k (UGWD($a, k; \rho$) \sim UGWD($k, a; \rho$)). As a result, two solutions are obtained for a and k when fitting the distribution and hence distinguishable estimates for σ_λ^2 and σ_ν^2 cannot be obtained. Dealing with the soap factory accident data, Irwin (1968) considered k to be the bigger value, as that made the variance component for proneness larger than that for liability. The assumption that the larger component measured proneness was based on his knowledge of departmental differences. In general, however, although one may consider that $\sigma_\lambda^2 + k^2 \sigma_\nu^2$ represents the variance component due to all non-random factors, the mathematics alone do not tell us whether σ_λ^2 represents the liability component and $k^2 \sigma_\nu^2$ the proneness component or vice versa. Thus, if the observations are arranged as a univariate distribution, the choice of the proneness or liability component will, as Irwin points out, depend upon extra information concerning the group of individuals under observation. Can we, then, rearrange the observations in such a way as to throw some light on this problem?

A sensible method would appear to be to divide the whole period of observation into two non-overlapping sub-periods and then study the resulting bivariate accident distribution. So, we need to introduce a bivariate theoretical distribution with marginals as UGWD's to describe our data in the light of a proneness liability hypothesis. The aim is to estimate the variance components based on the mathematics alone (objective information) rather than on personal (subjective) judgement.

In the next section a bivariate extension of the UGWD is defined and its structural properties are studied in the Appendix. Section 3 goes on to examine this bivariate form in relation to accident theory and to show that, using it, it is indeed possible to obtain distinguishable estimates for the variance components due to proneness and liability. Finally, in Section 4 the bivariate model is applied to data on road accidents and conclusions are drawn with regard to the underlying contributing factors.

2. THE BIVARIATE GENERALIZED WARING DISTRIBUTION

To construct a bivariate extension of the UGWD, a natural approach would be to go back to Irwin's (1963) original definition of the univariate version. His starting point was Waring's expansion

$$\frac{1}{x-a} = \sum_{r=0}^{\infty} \frac{a(r)}{x(r+1)}.$$

This he then generalized to

† The difference of (1.1) from Irwin's (1968) original relationship is only notational and due to a different scale parameter in the distribution of $\lambda | \nu$.

$$\frac{1}{(x-a)_{(k)}} = \sum_{r=0}^{\infty} \frac{a_{(r)}k_{(r)}}{x_{(k+r)}} \frac{1}{r!}, \quad a, k > 0.$$

Hence letting $\rho = x - a > 0$ and multiplying both sides by $\rho_{(k)}$ the successive terms of the resulting series defined his generalized Waring distribution UGWD $(a, k; \rho)$ with probability function (p.f.).

$$p_r = \frac{\rho_{(k)}}{(a+\rho)_{(k)}} \frac{a_{(r)}k_{(r)}}{(a+k+\rho)_{(r)}} \frac{1}{r!}, \quad a, k, \rho > 0, \quad r = 0, 1, 2, \dots$$

It is our aim now to define the bivariate version of this distribution. In doing so we will first generalize further Waring's expansion.

Let $\Delta f(x) = f(x+1) - f(x)$. Then, for $k, m, a > 0$.

$$\begin{aligned} \frac{1}{(x-a)_{(k+m)}} &= (1+\Delta)^{-a} \frac{1}{x_{(k+m)}} = \sum_{l=0}^{\infty} \frac{a_{(l)}(-1)^l}{l!} \Delta^l \left[\frac{1}{x_{(k)}} \frac{1}{(x+k)_{(m)}} \right] \\ &= \sum_{l=0}^{\infty} \sum_{r=0}^l \frac{a_{(l)}(-1)^l}{l!} \binom{l}{r} \Delta^r \frac{1}{x_{(k)}} \Delta^{l-r} \frac{1}{(x+k+r)_{(m)}} \\ &= \sum_{r=0}^{\infty} \sum_{l=0}^{\infty} \frac{a_{(r+l)}(-1)^{r+l}}{r! l!} \Delta^r \frac{1}{x_{(k)}} \Delta^l \frac{1}{(x+k+r)_{(m)}} \\ &= \sum_{r=0}^{\infty} \sum_{l=0}^{\infty} \frac{a_{(r+l)}k_{(r)}m_{(l)}}{x_{(k+m+r+l)}} \frac{1}{r!} \frac{1}{l!}. \end{aligned}$$

Provided that $x > a$, the double series in the right-hand side of this equation is convergent. If we now let $\rho = x - a > 0$ and multiply both sides by $\rho_{(k+m)}$, we will obtain a double series of positive terms which converges to unity and thus its general term can be regarded as defining a bivariate discrete probability distribution with p.f.

$$p_{r,l} = \frac{\rho_{(k+m)}}{(a+\rho)_{(k+m)}} \frac{a_{(r+l)}k_{(r)}m_{(l)}}{(a+k+m+\rho)_{(r+l)}} \frac{1}{r!} \frac{1}{l!}, \quad a, k, m, \rho > 0; \quad r, l = 0, 1, 2, \dots$$

In the sequel, we refer to this distribution as the *bivariate generalized Waring distribution* with parameters a, k, m and ρ and we denote it by BGWD $(a, k, m; \rho)$.

3. THE BGWD IN RELATION TO ACCIDENT THEORY

Let us now see how the bivariate model defined in Section 2 can arise in an accident situation in the context of accident proneness and accident liability.

Consider individuals of proneness ν and liability $\lambda_i | \nu$ for a period i of observation. Assume that over two non-overlapping time periods the numbers X, Y of accidents incurred by these individuals follow a double Poisson distribution with p.g.f.

$$G(X, Y | \lambda_1, \lambda_2, \nu)(s, t) = \exp\{(\lambda_1 | \nu)(s-1) + (\lambda_2 | \nu)(t-1)\}, \quad \lambda_1, \lambda_2 > 0.$$

Let the liability parameters $\lambda_1 | \nu, \lambda_2 | \nu$ be independently distributed as gamma $(k; \nu)$ and gamma $(m; \nu)$ respectively, i.e. $\lambda_i | \nu \sim (\Gamma(\theta_i)\nu^{\theta_i})^{-1} e^{-\lambda_i/\nu} \lambda_i^{\theta_i-1}$, $\theta_1 \equiv k, \theta_2 \equiv m, \nu > 0$. Then, for individuals with the same proneness but varying liabilities the joint distribution of accidents over

the two periods is double negative binomial, i.e.

$$G_{(X, Y)|\nu}(s, t) = \{1 + \nu(1-s)\}^{-k} \{1 + \nu(1-t)\}^{-m}.$$

Assume further that the proneness parameter ν has a beta distribution of the second type with parameters a and ρ , i.e. $\nu \sim \Gamma(a + \rho)\nu^{a-1}(1 + \nu)^{-(a+\rho)}/[\Gamma(a)\Gamma(\rho)]$. Then, the joint distribution of the numbers of accidents over the two periods is

$$\begin{aligned} G_{(X, Y)}(s, t) &= \frac{\Gamma(\rho + a)}{\Gamma(\rho)\Gamma(a)} \int_0^{+\infty} \nu^{a-1}(1 + \nu)^{-(a+\rho)} \{1 + \nu(1-s)\}^{-k} \{1 + \nu(1-t)\}^{-m} d\nu \\ &= \frac{\Gamma(\rho + a)}{\Gamma(\rho)\Gamma(a)} \sum_{r=0}^{\infty} \sum_{l=0}^{\infty} \frac{k(r)}{r!} \frac{m(l)}{l!} \int_0^{+\infty} \nu^{a+r+l-1}(1 + \nu)^{-(a+\rho+k+m+r+l)} d\nu \\ &= \frac{\rho(k+m)}{(a + \rho)(k+m)} F_1(a; k, m; a + k + m + \rho; s, t) \sim \text{BGWD}(a; k, m; \rho). \end{aligned}$$

An interesting feature of this model is that it allows not only for differences in the exposure conditions from person to person within each sub-period, but also for differences in the exposure conditions of the same individual between the two periods. The person's proneness is regarded as constant throughout the entire period of observation. Taking into account the fact that the term proneness reflects the individual's inherent tendency to incur accidents, the latter assumption is reasonable, at least for a limited period of time. (By no means do we imply that the person's proneness remains constant throughout the person's lifetime.) The choice of a gamma form for the distribution of the liability parameter and of a beta form for the distribution of the proneness parameter is completely analogous to Irwin's choice in the univariate case.

Let us now come to the question of assessing the significance of the contribution of proneness, liability and randomness in a given accident situation. Assume that a period of observation is split into two non-overlapping sub-intervals and that the BGWD $(a; k, m; \rho)$ fits the resulting bivariate accident distribution. Then we have for the total variance of the overall period that $\sigma^2 = \sigma_\lambda^2 + (k + m)^2 \sigma_\nu^2 + \sigma_R^2$ where $\sigma_\lambda^2 = a(k + m)(a + 1)(\rho - 1)^{-1}(\rho - 2)^{-1}$ (liability component),

$$\sigma_\nu^2 = a(a + \rho - 1)(\rho - 1)^{-2}(\rho - 2)^{-1} \quad (\sigma_\nu^2 \propto \text{proneness component})$$

and $\sigma_R^2 = a(k + m)(\rho - 1)^{-1}$ (random component). Note that the BGWD is not symmetrical in a , k or a , m , so any method of parameter estimation will provide separate estimates for a , ρ and $k + m$. Hence, distinguishable estimates of the variance components can be obtained using the above equations. Moreover, one can infer the contribution that proneness and liability have had in each of the two sub-periods. Starting from the obvious relationships $\sigma_X^2 = \sigma_{\lambda_1}^2 + k^2 \sigma_\nu^2 + \sigma_{R_1}^2$ and $\sigma_Y^2 = \sigma_{\lambda_2}^2 + m^2 \sigma_\nu^2 + \sigma_{R_2}^2$ we arrive at estimators for the variance components. The results can be summarized in Table 1.

It should be noted that the BGWD is symmetrical in k and m (BGWD $(a; k, m; \rho) \sim$ BGWD $(a; m, k; \rho)$). Thus, fitting the distribution by any method of parameter estimation will yield two solutions. Judging from the fact that under the proneness-liability hypothesis the random components of the marginal variances $\sigma_{R_1}^2$ and $\sigma_{R_2}^2$ are (from (1.2) and (1) in the Appendix) identical to μ_X and μ_Y respectively. It would seem reasonable to choose the solution that yields estimates of these two components within a neighbourhood of the corresponding observed marginal means. This would eliminate any ambiguity arising from interchanging the roles of k and m with regard to the estimation of the components of the two marginal variances.

TABLE 1
Estimators of the components of the variance of the generalized Waring distribution

Component due to	Marginal variance of X	Marginal variance of Y	Variance of X + Y
Random factors	$\frac{\hat{a}\hat{k}}{\hat{\rho}-1}$	$\frac{\hat{a}\hat{m}}{\hat{\rho}-1}$	$\frac{\hat{a}(\hat{k}+\hat{m})}{\hat{\rho}-1}$
Proneness	$\frac{\hat{k}^2\hat{a}(\hat{a}+\hat{\rho}-1)}{(\hat{\rho}-1)^2(\hat{\rho}-2)}$	$\frac{\hat{m}^2\hat{a}(\hat{a}+\hat{\rho}-1)}{(\hat{\rho}-1)^2(\hat{\rho}-2)}$	$\frac{(\hat{k}+\hat{m})^2\hat{a}(\hat{a}+\hat{\rho}-1)}{(\hat{\rho}-1)^2(\hat{\rho}-2)}$
Liability	$\frac{\hat{a}\hat{k}(\hat{a}+1)}{(\hat{\rho}-1)(\hat{\rho}-2)}$	$\frac{\hat{a}\hat{m}(\hat{a}+1)}{(\hat{\rho}-1)(\hat{\rho}-2)}$	$\frac{\hat{a}(\hat{k}+\hat{m})(\hat{a}+1)}{(\hat{\rho}-1)(\hat{\rho}-2)}$
Total	$\frac{\hat{a}\hat{k}(\hat{\rho}+\hat{k}-1)(\rho+a-1)}{(\hat{\rho}-1)^2(\hat{\rho}-2)}$	$\frac{\hat{a}\hat{m}(\hat{\rho}+\hat{m}-1)(\hat{\rho}+\hat{a}-1)}{(\hat{\rho}-1)^2(\hat{\rho}-2)}$	$\frac{\hat{a}(\hat{k}+\hat{m})(\hat{\rho}+\hat{k}+\hat{m}-1)(\hat{\rho}+\hat{a}-1)}{(\hat{\rho}-1)^2(\hat{\rho}-2)}$

4. APPLICATIONS

In this section we consider applications of the BGWD ($a; k, m; \rho$) to road accident data. Table 2 contains Cresswell and Froggatt's (1963) data on accidents incurred by 183 bus drivers in Northern Ireland during the period 1952-55. The data are arranged in a bivariate distribution whose margins correspond to the 1952-53 and 1954-55 periods (upper entries).

TABLE 2(*)
Accidents to Belfast Corporation Transport bus drivers

Y \ X		1954-55										MARG X
		0	1	2	3	4	5	6	7	8	9	
1952-53	0	8 9.03	9 10.72	4 7.52	3 4.06	5 1.87	0 0.77	0 0.29	0 0.10	0 0.04	0 0.01	29 34.41
	1	13 10.26	15 14.33	14 11.56	7 7.05	4 3.62	1 1.64	0 0.68	0 0.27	0 0.10	0 0.03	54 49.54
	2	6 6.89	16 11.06	10 10.08	8 6.86	2 3.88	3 1.93	2 0.87	0 0.36	1 0.14	0 0.05	48 42.13
	3	0 3.57	8 6.47	7 6.57	3 4.93	5 3.05	1 1.64	0 0.80	0 0.36	1 0.15	0 0.06	25 27.59
	4	2 1.57	4 3.18	1 3.56	3 2.92	3 1.96	1 1.14	0 0.59	0 0.28	0 0.13	0 0.05	14 15.37
	5	1 0.62	1 1.38	0 1.69	1 1.51	1 1.09	1 0.68	0 0.38	1 0.19	0 0.09	0 0.04	6 7.67
	6	0 0.23	1 0.55	1 0.73	0 0.70	1 0.54	1 0.36	0 0.21	0 0.12	0 0.06	1 0.03	5 3.53
	7	0 0.08	0 0.21	0 0.29	0 0.30	0 0.25	0 0.18	0 0.11	0 0.06	0 0.03	0 0.02	0 1.53
	8	0 0.03	1 0.07	0 0.11	0 0.12	0 0.11	0 0.08	0 0.05	0 0.03	0 0.02	0 0.01	1 0.63
	9	0 0.01	0 0.02	1 0.04	0 0.05	0 0.04	0 0.03	0 0.02	0 0.02	0 0.01	0 0.00	1 0.25
MARG Y		30 32.27	55 47.99	38 42.15	25 28.50	21 16.40	8 8.45	2 4.01	1.00 1.79	2 0.76	1 0.31	183 182.64

$\chi^2 = 9.4144$ with 16 degrees of freedom. $P(\chi_{16}^2 \geq 9.4144) = 0.894$.
(*) Broken lines indicate the grouping adopted in the application of the chi-square test.

The table shows a correlation coefficient of 0.259 between X and Y . It would therefore appear reasonable to assume that the population is non-homogeneous. This might also be supported by the fact that a Poisson distribution can be fitted to the observations within the first sub-interval quite well ($P > 0.3$), but not to the observations within the second sub-interval ($0.01 < P < 0.05$) and certainly not to those over the whole period ($P < 0.001$, Cresswell and Froggatt, 1963, pp. 155, 165, 166). Moreover, over the entire period of observation 63 drivers out of the 183 (34.4 per cent) were responsible for 433 out of 732 (59.2 per cent) accidents. Also, of those who had 4 or more accidents during the first period, 51.9 per cent had at least 3 accidents during the second period. It is perhaps worth noting at this point that the 183-driver group did not include inexperienced drivers, drivers with little experience, or near retirement. As Cresswell and Froggatt (1963) point out, this was done in an attempt to eliminate as much as possible any influence of age and experience. Attempts were also made to ensure as far as possible homogeneity in the environmental risk. Cresswell and Froggatt state (p. 18): "It seems reasonably certain that so far as can be ascertained, all drivers within each grouping selected were indeed equally exposed to the risk of incurring an accident." However, they cautiously add (p. 62): "Although the relevant variables of age and experience can be controlled to some extent, an equal-risk environment cannot be exactly realized in practice, . . ." Therefore, even though all the drivers that were included in the group worked in the same district, the encountered hazards (other than the road condition) cannot be ascertained as having been exactly the same for all the drivers during the 4-year period, nor for the same driver in the two consecutive 2-year sub-periods. Considering all these points, we conclude that the observed differences in individual behaviour cannot be explained by chance alone. Hence the proneness-liability model discussed in the previous section seems to be appropriate.

The BGWD has been fitted to this bivariate accident distribution by a method analogous to the one employed by Irwin (1968) for the UGWD using the first- and second-order factorial moments of the distribution. In particular, if X, Y denote the numbers of accidents during the first and second period respectively, the estimating equations used are

$$\bar{X} = \frac{\hat{a}\hat{k}}{\hat{\rho}-1}, \quad \bar{Y} = \frac{\hat{a}\hat{m}}{\hat{\rho}-1}, \quad \bar{T} = \frac{\hat{a}(\hat{a}+1)\hat{k}\hat{m}}{(\hat{\rho}-1)(\hat{\rho}-2)}, \quad \bar{W} + \bar{Z} = \frac{\hat{a}(\hat{a}+1)[\hat{k}(\hat{k}+1) + \hat{m}(\hat{m}+1)]}{(\hat{\rho}-1)(\hat{\rho}-2)}, \quad (4.1)$$

where $\bar{W} = \sum_{i,j} j(j-1)f_{ij}/n$, $\bar{Z} = \sum_{i,j} i(i-1)f_{ij}/n$, $\bar{T} = \sum_{i,j} ijf_{ij}/n$ and f_{ij} is the observed joint frequency.

The asymptotic variance-covariance matrix V of these parameter estimators can be obtained as follows.

Let $\theta \equiv (\theta_1, \theta_2, \theta_3, \theta_4)$ denote the parameter vector (a, k, m, ρ) , $\hat{\theta} \equiv (\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4)$ denote the vector $(\hat{a}, \hat{k}, \hat{m}, \hat{\rho})$ and let $\tau \equiv (\tau_1, \tau_2, \tau_3, \tau_4)$ and $t \equiv (t_1, t_2, t_3, t_4)$ denote the vectors $(\mu_x, \mu_y, \mu(2, 0) + \mu(0, 2), \mu(1, 1))$ and $(\bar{X}, \bar{Y}, \bar{W} + \bar{Z}, \bar{T})$ respectively. (Here $\mu(r, l) = E(X^{(r)} Y^{(l)})$ where $X^{(r)} = X(X-1) \dots (X-r+1)$, $X^{(0)} = 1$). The parameter estimators $\hat{\theta}_i$ are functions of the sample factorial moments t_i , say $\hat{\theta}_i = \phi_i(t_1, t_2, t_3, t_4)$ $i = 1, 2, 3, 4$. Then, from the general theory we have

$$V = \mathbf{S} \mathbf{J}' \quad (4.2)$$

where

$$J = (J_{ij}) = \left(\frac{\partial \phi_i^{-1}}{\partial \theta_j} \right)_{\hat{\theta} = \theta}^{-1} = \left(\frac{\partial \phi_i}{\partial t_j} \right)_{t = \tau} \quad \text{and} \quad S = (S_{ij}) = (\text{cov}(t_i, t_j)), \quad i, j = 1, 2, 3, 4$$

with ' for transpose.

Then, the asymptotic variances of the variance component estimators will be given by

$$V(\sigma_{ij}^2) = J_{ij} V I'_{ij} \quad (4.3)$$

with i for component type, i.e. random ($i = 1$), proneness ($i = 2$), liability ($i = 3$) and j for time period, i.e. first ($j = 1$), second ($j = 2$) overall ($j = 3$). Here

$$I_{ij} = \left(\frac{\partial \hat{\sigma}_{ij}^2}{\partial \hat{a}}, \frac{\partial \hat{\sigma}_{ij}^2}{\partial \hat{k}}, \frac{\partial \hat{\sigma}_{ij}^2}{\partial \hat{m}}, \frac{\partial \hat{\sigma}_{ij}^2}{\partial \hat{\rho}} \right)_{\hat{\theta} = \theta}, \quad i, j = 1, 2, 3.$$

Explicit formulae for J , S and I_{ij} in terms of a , k , m and ρ can be obtained after much algebra, (see Xekalaki (1984b)).

The means and second-order moments estimated from the data of Cresswell and Froggatt are $\bar{x} = 1.9563$, $\bar{y} = 2.0437$, $\bar{z} = 4.4372$, $\bar{w} = 5.0492$, $\bar{t} = 4.7159$. Hence using equations (5.1) we find $\hat{a} = 5.6338$, $\hat{k} = 202.2860$, $\hat{m} = 211.3270$, $\hat{\rho} = 583.5520$.

The upper and lower entries of Table 2 show the observed and expected cell frequencies of accidents. As judged by the Pearson chi-squared goodness-of-fit criterion (χ^2), the BGWD fits the data well ($\chi_{16}^2 = 9.4144$, $P = 0.894$) and so do the marginals as well as the distribution for the total period (Table 3). Estimates of the components of σ^2 , σ_X^2 and σ_Y^2 can now be obtained using Table 1 and are given in Table 6.

TABLE 3
Accidents to Belfast Corporation Transport bus drivers (entire period 1952-55)

u	0	1	2	3	4	5	6	7	8	9	10	11	> 12	Total
Observed	8	22	25	33	32	24	8	12	8	3	3	3	2	183
Expected	9.03	20.98	28.74	30.25	27.04	21.62	15.91	11.00	7.26	4.57	2.77	1.63	1.84	182.64

It seems that in this particular accident situation, chance and proneness were by far the dominating causing factors. In both sub-periods as well as in the total period the contribution of random factors to the total variance is above 50 per cent. Of the remaining percentage, proneness appears to have the largest share in all the three cases. The exposure of the individuals to external risk of accident does not seem to have had any significant contribution. This indeed confirms Cresswell and Froggatt's assertion of an "as far as possible" equal-risk environment. Most important of all, it indicates that introducing a parameter that accounts for differences in the external conditions has not led to contradictions. Quite to the contrary, it is noteworthy that the model produces an estimate of the effect of the external factors that is as small as, by the driver selection procedure, it was expected to be. In fact, the results tell us more. They imply that it can be ascertained from the model that not only did those external factors that were controlled (e.g. homogeneity of road conditions) have little influence on accident causation, but, also, that impossible-to-control external factors (e.g. careless driving of other road users) made no significant contribution either. One can, therefore, say that, among the non-random factors adduced to account for the fact that some drivers have had more accidents than others, proneness had a dominating effect—a conclusion that is not at variance with what was anticipated.

The model is next applied to a set of accident data from a different driver population for which the need for allowing for differences in the environmental risk would be more obvious. The motor-vehicle accident data of Table 4 represent the accident experience of a sample of 29 531 general drivers in the State of Connecticut, USA licensed and driving through the entire 6-year period from 1931 to 1936 (US Bureau of Public Roads, 1938). The data are presented as a bivariate frequency distribution whose marginals refer to the periods 1931-33 and 1934-36. The drivers were selected from the files of applicants for renewal of operator's licences who had been licensed in every year from 1931 to 1936. There was no selection whatsoever on the basis of an "equal risk environment". The reported number of accidents per driver refers to the specified 6-year period "without regard to the mileage covered or to the number of hazards encountered by each driver", as indicated on page 18 of the above-mentioned document. Hence the fact that the exposure of different drivers to external risk may be different should be taken into consideration.

TABLE 4 (*)
Accidents to Connecticut general drivers

X \ Y	1931-33					MARG X	
	0	1	2	3	4		
1934-36	0	23 881 23 881.17	2117 2144.91	242 213.43	17 23.26	2 2.75	26 259 26 265.52
	1	2 386 2 374.65	419 422.20	57 62.36	9 8.97	3 1.31	2 874 2 869.49
	2	275 258.90	64 68.33	12 13.32	5 2.37	1 0.41	357 343.33
	3	22 30.65	5 10.68	2 2.57	2 0.54	0 0.11	31 44.56
	4	5 3.91	4 1.68	0 0.48	1 0.12	0 0.03	10 6.22
MARG Y	26 569 26 549.29	2609 2647.80	313 292.17	34 35.26	6 4.61	29 531 29 529.13	

$\chi^2 = 17.9978$ with 10 degrees of freedom $P(\chi_{10}^2 \geq 17.9978) = 0.056$
 (*) Broken lines indicate the grouping adopted in the application of the Pearson chi-square test.

The 29 531 drivers collectively reported 7082 accidents. Of these, there were 1147 drivers, 3.9 per cent of the total driver-population, who had 2 or more accidents each and who together had 2579 accidents, i.e. 36.4 per cent of all the accidents. Also, 211 drivers having 3 or more accidents who constituted about 0.7 per cent of the population were responsible for 10 per cent of the accidents. On the other hand, the majority of the drivers, 80.9 per cent, had no accidents. Note also that, in the period 1934-36, the total number of accidents incurred by the accident-free drivers of 1931-33 was 2660 yielding a rate of 0.101 accidents per driver, while the corresponding rate for the drivers with 1 or more accidents in 1931-33 was $701/3272 = 0.214$ accidents per driver. The BGWD has been fitted to this set of data by the same method as before. The values of the sample statistics involved are $\bar{x} = 0.1260$, $\bar{y} = 0.1138$, $\bar{w} = 0.0305$, $\bar{z} = 0.0345$, $\bar{t} = 0.0596$ yielding $\hat{a} = 0.9992$, $\hat{k} = 9.2774$, $\hat{m} = 8.3798$, $\hat{p} = 74.5709$.

The value of the Pearson chi-square statistic is $\chi^2 = 17.9978$ with 10 degrees of freedom indicating a quite reasonable fit ($P = 0.06$). The fit is also quite good for the total period of observation (Table 5). The estimates of the variance components for this set of data are given in

TABLE 5
Accidents to Connecticut general drivers (entire period (1931-36))

u	0	1	2	3	4	5	6	7	8	Total
Observed	23 881	4503	936	160	33	14	3	1	0	29 531
Expected	23 881.17	4519.56	894.53	184.60	39.63	7.93	1.43	0.23	0.03	29 529.13

Table 6. Again the random factors seem to have significantly contributed to this accident situation in both sub-periods as well as in the total period. The corresponding effects of the external factors are higher than those of the previous application but certainly much lower than the effects of proneness.

As far as the asymptotic standard errors of the parameter estimators \hat{a} , \hat{k} , \hat{m} and \hat{p} are concerned, it may be pointed out that in both cases they come out to be uncomfortably large.

TABLE 6
Estimates of the components of the variance of the generalized Waring distribution

Component	Connecticut drivers			Belfast Corporation transport bus drivers		
	1931-33	1934-36	1931-36	1952-53	1954-55	1952-55
Random	0.1260 (86.4%)	0.1138 (87.4%)	0.2398 (78.5%)	1.9563 (73.4%)	2.0437 (72.6%)	4.0000 (57.8%)
Proneness	0.0163 (11.2%)	0.0133 (10.2%)	0.0591 (19.3%)	0.6871 (25.8%)	0.7498 (26.6%)	2.8724 (41.5%)
Liability	0.0035 (2.4%)	0.0031 (2.4%)	0.0066 (2.2%)	0.0223 (0.8%)	0.0233 (0.8%)	0.0456 (0.7%)
Total	0.1458 (100%)	0.1302 (100%)	0.3055 (100%)	2.6657 (100%)	2.8169 (100%)	6.9180 (100%)

One explanation for this may be the fact that the determinant of the matrix $(\partial\phi_i^{-1}/\partial\theta_j)$ is very small in both cases, $(-9.4)10^{-10}$ and $(-2.5)10^{-10}$ respectively, which implies an almost singular matrix. Hence, the asymptotic theory seems to offer not much help in the evaluation of standard errors for these two cases. Maybe the answer lies in the fact that (4.2) provides exact variances and covariances as long as the parameter estimators are linear functions of the sample statistics. Here the form of the estimators is far from being linear and, thus, (4.2) may not be appropriate for a reasonable representation of the variance-covariance matrix of \hat{a} , \hat{k} , \hat{m} and \hat{p} . On the other hand, the large standard errors may be due to the inherent weaknesses of the moment estimators. All these signal the need for a separate investigation into the properties of the suggested method of estimation and possibly of alternative methods of estimation. At this stage, evaluating the standard errors of the variance components through (4.3) will not be meaningful as these will incorporate any inaccuracies in the standard errors of the parameters.

In conclusion, it should be noted that the BGWD introduced in this paper should not be thought of as limited to applications in the field of accident studies. It may well provide a useful model for other practical situations for which data can be arranged in a bivariate form, especially in cases where splitting the variance into meaningful components and estimating their importance might provide further valuable information.

ACKNOWLEDGEMENTS

The authoress would like to thank Professor C. D. Kemp for encouraging her to pursue research in the field of accident theory. She would also like to express her appreciation to her husband Dr J. Panaretos for numerous enlightening discussions and thoughtful suggestions.

REFERENCES

- Cresswell, W. L. and Froggatt, P. (1963) *The Causation of Bus Driver Accidents*. London: Oxford University Press.
- Irwin, J. O. (1963) The place of mathematics in medical and biological statistics. *J. R. Statist. Soc. A*, **126**, 1-44.
- (1968) The generalized Waring distribution applied to accident theory. *J. R. Statist. Soc. A*, **131**, 205-225.
- (1975) The generalized Waring distribution. *J. R. Statist. Soc. A*, **138**, 18-31 (Part I), 204-227 (Part II), 374-384 (Part III).
- US Bureau of Public Roads. Motor vehicle traffic conditions in the United States: the accident-prone driver. (House document No. 462, Part 6, 75th Congress, Third Session). Washington: US Government Printing Office, 1938, Pp. x + 52.
- Xekalaki, E. (1981) Chance mechanisms for the univariate generalized Waring distributions. In *Statistical Distributions in Scientific Work, Vol. 4 (Models, Structures and Characterizations)* (C. Taillie, G. P. Patil and B. Baldessari, eds), pp. 157-171. Dordrecht, Holland: D. Reidel.
- (1983a) The univariate generalized Waring distribution in relation to accident theory: proneness, spells or contagion? *Biometrics*, **39** (3), 887-895.
- (1983b) Infinite divisibility, completeness and regression properties of the generalized Waring distribution. *Ann. Inst. Statist. Math.*, **35** (A), 161-171.
- (1983c) A property of the Yule distribution and its applications. *Commun. Statist. A*, **12**, 1181-1189.
- (1983d) Hazard functions and life distributions in discrete time. *Commun. Statist. A*, **12**, 2503-2509.
- (1983e) Some bivariate extensions of the generalized Waring distribution *Studia Sci. Math. Hungar.* (to appear).

- (1983f) Models leading to the bivariate generalized Waring distribution. Technical report 93, Department of Statistics and Actuarial Science, University of Iowa. (to appear in *Utilitas Mathematica*.)
- (1983g) Some identifiability problems involving generalized Waring distributions. Technical report 94, Department of Statistics and Actuarial Science, University of Iowa. (to appear in *Publicationes Mathematicae*.)
- (1984a) Linear regression and the Yule distribution. *J. Econometrics*, **24**, 397–403.
- (1984b) Factorial moment estimation for the bivariate generalized Waring distribution. *Statistische Hefte* (to appear).
- Xekalaki E. and Panaretos J. (1983) Identifiability of compound Poisson distributions. *Scand. Actuarial J.*, 1983 (1), 39–45.

APPENDIX

The p.g.f. of the bivariate generalized Waring distribution can be expressed in terms of the Appell function of the first kind, i.e.

$$G(s, t) = \frac{\rho_{(k+m)}}{(a+\rho)_{(k+m)}} F_1(a; k, m; a+k+m+\rho; s, t),$$

where

$$F_1(a; \beta, \beta'; \gamma; u, v) = \sum_{r=0}^{\infty} \sum_{l=0}^{\infty} \frac{\alpha_{(r+l)} \beta_{(r)} \beta'_{(l)}}{\gamma_{(r+l)}} \frac{u^r v^l}{r! l!}.$$

The fact that $\rho > 0$ implies that $G(s, t)$ is convergent for all the values of a, k, m within the area $[-1, 1] \times [-1, 1]$ and by Gauss's theorem

$$F_1(a; k, m; a+k+m+\rho; 1, 1) = (a+\rho)_{(k+m)} / \rho_{(k+m)}.$$

The successive probabilities of the BGWD ($a; k, m; \rho$) are related by first-order recurrence relationships, namely

$$\frac{p_{i+1, j}}{p_{i, j}} = \frac{(a+i+j)(k+i)}{(a+k+m+\rho+i+j)(i+1)}, \quad \frac{p_{i, j+1}}{p_{i, j}} = \frac{(a+i+j)(m+j)}{(a+k+m+\rho+i+j)(j+1)}.$$

Properties of the BGDW

Both the unconditional and the conditional marginal distributions as well as the convolution of the two marginals have a UGWD form.

Theorem. Consider a random vector (X, Y) of non-negative and integer-valued components such that $(X, Y) \sim \text{BGWD}(a; k, m; \rho)$. Then

- (i) $X \sim \text{UGWD}(a, k; \rho)$, $Y \sim \text{UGWD}(a, m; \rho)$.
- (ii) $X + Y \sim \text{UGWD}(a, k+m; \rho)$.
- (iii) $X | (Y=y) \sim \text{UGWD}(a+y, k; \rho+m)$, $Y | (X=x) \sim \text{UGWD}(a+x, m; \rho+k)$.

Proof: Let $G_Z(s)$ denote the p.g.f. of a random variable Z . Then

$$\begin{aligned} \text{(i)} \quad G_X(s) &= \frac{\rho_{(k+m)}}{(a+\rho)_{(k+m)}} F_1(a; k, m; a+k+m+\rho; s, 1) \\ &= \frac{\rho_{(k+m)}}{(a+\rho)_{(k+m)}} \sum_{x=0}^{\infty} \frac{a_{(x)} k_{(x)}}{(a+k+m+\rho)_{(x)}} \frac{s^x}{x!} {}_2F_1(a+x, m; a+k+m+\rho+x; 1) \\ &= \frac{\rho_{(k)}}{(a+\rho)_{(k)}} {}_2F_1(a, k; a+k+\rho; s) \sim \text{UGWD}(a, k; \rho). \end{aligned}$$

$$\begin{aligned}
 \text{(ii)} \quad G_{X+Y}(s) &= \frac{\rho(k+m)}{(a+\rho)(k+m)} F_1(a; k, m; a+k+m+\rho; s, s) \\
 &= \frac{\rho(k+m)}{(a+\rho)(k+m)} {}_2F_1(a, k+m; a+k+m+\rho; s) \sim \text{UGWD}(a, k+m; \rho). \\
 \text{(iii)} \quad G_{X|Y=y}(s) &= \frac{(\partial y / \partial t^y) F_1(a; k, m; a+k+m+\rho; s, 0)}{(\partial y / \partial t^y) F_1(a; k, m; a+k+m+\rho; 1, 0)} \\
 &= \frac{F_1(a+y; k, m+y; a+k+m+y+\rho; s, 0)}{F_1(a+y; k, m+y; a+k+m+y+\rho; 1, 0)} \\
 &= \frac{(\rho+m)(k)}{(a+y+\rho+m)(k)} {}_2F_1(a+y, k; a+k+m+\rho+y; s) \sim \text{UGWD}(a+y, k; \rho+m).
 \end{aligned}$$

Remark: Note that proposition (iii) implies that the regressions of X on Y and of Y on X are linear.

The fact that both marginals as well as their convolution are UGWD's has important consequences in the field of accident studies where arbitrarily splitting a period of time into two sub-periods requires the marginal distributions of accidents and the distribution of accidents for the whole period to have the same form.

The factorial moments of the BGWD $(a; k, m; \rho)$ are given by the formula

$$\mu_{(i,j)} = \frac{a^{(i+j)} k^{(i)} m^{(j)}}{(\rho-1)(\rho-2)\dots(\rho-i-j)}, \quad i, j = 0, 1, 2, \dots$$

It can be seen that for $\rho \leq i+j$ the factorial moments and, hence, the moments become infinite. Through the known transformation formulae we obtain for the moments of the BGWD

$$\begin{aligned}
 \mu'_{1,0} \equiv \mu_X &= \frac{ak}{\rho-1}, \quad \mu'_{0,1} \equiv \mu_Y = \frac{am}{\rho-1}, \quad \mu_{1,1} \equiv \sigma_{XY} = \frac{akm(a+\rho-1)}{(\rho-1)^2(\rho-2)}, \\
 \mu_{2,0} \equiv \sigma_X^2 &= \frac{ak(\rho+k-1)(\rho+a-1)}{(\rho-1)^2(\rho-2)}, \quad \mu_{0,2} \equiv \sigma_Y^2 = \frac{am(\rho+m-1)(\rho+a-1)}{(\rho-1)^2(\rho-2)}. \quad (1)
 \end{aligned}$$

Since it is necessary that $\rho > 2$ in order that all the second-order central moments exist it follows that the covariance σ_{XY} is positive which implies that X, Y are always positively correlated.