'Stuttering-Poisson' Distributions

By C. D. KEMP

(Read before the Society in Belfast on May 1st, 1967)

' Stuttering-Poisson ' (Galliher *et al.*[10]) aptly describes a very general class of discrete distributions which appears under a variety of names in the literature and which has many practical applications in operational research, e.g. in queueing and inventory problems. In Gurland's[11] terminology these are generalised Poisson distributions. After reviewing definitions and derivations, some general properties are noted. A very simple method is used to derive recurrence relationships between the probabilities; derivatives of the probabilities w.r.t. the parameters are also obtained. Some individual members of the class are considered, particularly one examined by Galliher *et al.*—this is identified as a Polya–Aeppli distribution and has an especially simple recurrence relationship. Finally, methods of approximating to these distributions are briefly mentioned.

INTRODUCTION

The probability generating function (PGF) is defined as

$$g(s) = E(s^x) = \sum_{x=0} P_x s^x, \tag{1}$$

where P_x is the probability that the (discrete) random variable takes the value x.

The class of distributions we consider here has PGF

$$G(s) = \exp\left[\sum_{i=1}^{\infty} a_i(s^i - 1)\right],\ a_i \geq 0. \tag{2}$$

Galliher *et al.*[10] used the name ' stuttering-Poisson ' to describe distributions arising from the following model: suppose that ' bursts ' of demand are distributed as a Poisson, whilst numbers of demands per burst are independently distributed as some other distribution, then total demands per unit time are distributed as the ' stuttering-Poisson '. Cox[5] called this type of process a ' cumulative process associated with a Poisson process '. Galliher *et al.* were particularly concerned with the case where number of demands per burst follows a geometric distribution, and Adelson[1] has pointed out that the PGF of the resultant distribution can be expressed in the form (2). He showed this by considering the distribution as arising from an intermingling of Poisson streams. Another way of showing the result in the general case is to utilise the following well-known result[2,3,9]: if $S_n = X_1 + X_2 + \ldots + X_n$, where the X_i are independently identically distributed with PGF $\varnothing(s)$, and if n itself is distributed with PGF $g(s)$, then S_n is distributed with PGF

$$G(s) = g[\varnothing(s)]. \tag{3}$$

In particular, if n is distributed as a Poisson with parameter λ, then $g(s) = \exp\,[\lambda(s-1)]$ and

$$G(s) = \exp\{\lambda[\varnothing(s) - 1]\}, \tag{4}$$

and, since $\varnothing(s)$ is itself a PGF, $G(s)$ can be written in the form (2). Hence the general ' stuttering-Poisson ' distribution has PGF of form (2).

There is considerable confusion in the literature concerning the use of the terms ' compound ' and ' generalised '. Some authors[1,2,3,9] call (4) a compound Poisson distribution whilst others[8,11,21] call it a generalised Poisson distribution. We adopt Gurland's[11] terminology and notation. According to this, a *compound* Poisson results when we allow the Poisson parameter to be distributed, whilst a *generalised* Poisson results when we replace the generating variable s in the PGF of the Poisson by a function $\varnothing(s)$, which is the PGF of the generalising distribution, as in (4). These two processes are quite distinct. Many individual distributions are both compound and generalised Poissons and thus can result from two or more entirely different and often contradictory physical models. A good example is the negative binomial, which is both a Poisson compounded by a gamma (Poisson $\wedge$ gamma in Gurland's notation) and a Poisson generalised through a logarithmic (Poisson $\vee$ logarithmic). To avoid some of this confusion, it might perhaps be advantageous to use the name ' stuttering-Poisson ', rather than either compound or generalised Poisson, when the distribution is generated by the process leading to (4), which is, of course, ' generalisation ' in Gurland's terminology.

However, in addition to the use of ' compound Poisson ' and ' generalised Poisson ' to describe (4) and hence (2), the latter appears in the literature under a variety of names, e.g. the (Pollaczek–Geiringer) distribution of the multiple occurrences of rare events[13,20,24], the Poisson power series distribution,[19] the Poisson distribution with events in clusters,[4,26] the multiple Poisson distribution.[9] (In some definitions we are restricted to PGF's with a finite number of non-zero a_i.) It should be noted that the list of references at the end of this paper makes no attempt to be exhaustive, but we hope that it gives some indication of the wide variety of published work on these distributions.

Several well-known distributions are particular members of this class, e.g. negative binomial (Pascal), Neyman types A, B and C and Beall and Rescia's generalisation of these, Poisson Pascal, Poisson binomial. If in (3) both n and X_i are independently Poisson distributed, then S_n has Neyman's type A distribution. This distribution and its convolution with another independent Poisson distribution (which results in a new stuttering-Poisson—the ' Short ' distribution) have, for example, been applied to road accident data.[6,17] Again, if n is Poisson and X_i is positive binomial (with index 2) then S_n has the Hermite distribution[18] which, as we note later, is a useful approximating distribution. Galliher *et al.* consider the case when n is Poisson and X_i is geometric—in Gurland's notation this is Poisson $\vee$ geometric. In these various examples, of course, the a_i are all functions of the (usually 2 or 3) parameters of the Poisson and the generalising distribution.

SOME PROPERTIES OF THE DISTRIBUTIONS

Several authors[1,12,19] have derived a general recurrence relationship for the probabilities by repeatedly differentiating the PGF w.r.t. s and then placing $s = 0$. This method, with minor modifications, has also been used on individual stuttering-Poisson distributions by various authors.

However, by making use of the obvious property of (2) that

$$\frac{\partial G(s)}{\partial s} = \theta(s)G(s), \tag{5}$$

we can derive the recurrence relationship by only differentiating once thus:

$$\Sigma i p_i s^{i-1} = \frac{\partial G}{\partial s} = G(s)\Sigma i a_i s^{i-1} = \Sigma i a_i s^{i-1}\, \Sigma p_j s^j. \tag{6}$$

Equating coefficients of s^x in (6) we have

$$(x+1)P_{x+1} = \sum_{i=0}^{x}(i+1)a_{i+1}\, P_{x-i}, \tag{7}$$

which is the desired recurrence relationship. $P_0 = G(0) = \exp(-\Sigma a_i)$ and all succeeding probabilities can be computed recursively.

Now suppose that we differentiate $G(s)$ w.r.t. a_k:

$$\Sigma \frac{\partial P_i}{\partial a_k} s^i = \frac{\partial G(s)}{\partial a_k} = (s^k - 1)G(s) = (s^k - 1)\Sigma P_i s^i. \tag{8}$$

Again equating coefficients of s^x we have

$$\frac{\partial P_x}{\partial a_k} = P_{x-k} - P_x. \tag{9}$$

Equation (9) may be used to set up maximum-likelihood equations for estimating the a_i when we are interested in estimating only the first two or three a_i (and assume the rest negligible).

These equations are obtained immediately by substituting for $\partial P_x/\partial a_k$ in

$$\frac{\partial L}{\partial a_k} \equiv \sum_x f_x \frac{1}{P_x}\frac{\partial P_x}{\partial a_k} = 0, \tag{10}$$

where $L \equiv \Sigma f_x \log P_x$ is the log likelihood and f_x is the observed frequency of x. Except where we are only fiitting a_1 (the Poisson), equations (10) have to be solved iteratively (for details of methods and examples see, e.g.

Rao[23] or Kemp[17]). A computer programme for fitting a_1 up to and including a_3 has been written in Algol by the author and run on the S.R.C. atlas computer; results will be published elsewhere.

Interesting relationships hold between the parameters of the Poisson and the generalising distribution and those of the resultant stuttering-Poisson. In particular, if we denote the mean and variance of the generalising distribution by μ and σ^2 respectively, then the mean and variance of the stuttering-Poisson (4) are $\lambda\mu$ and $\lambda(\sigma^2 + \mu^2)$ respectively[9] (λ is the Poisson parameter), whilst in terms of (2) the mean and variance are $\Sigma i a_i$ and $\Sigma i^2 a_i$[1,20]. One way of obtaining these results is by differentiating the PGF repeatedly w.r.t. s and evaluating at $s = 1$ in order to obtain the factorial moments.

All distributions of class (4) have the property of being infinitely divisible.[9] This means that, in terms of the demand model which led to the stuttering-Poisson, if we consider the number of bursts in time t to be Poisson with parameter tt so that

$$G(s;\, t) = \exp\{\lambda t[\varnothing(s) - 1]\}, \tag{11}$$

then

$$G(s;\, t_1 + t_2) = G(s;\, t_1)G(s;\, t_2) \tag{12}$$

for two non-overlapping time periods t_1 and t_2, i.e. the number of demands in any time period is independent of that in any other time period. This simple property has led to the widespread use of stuttering-Poisson distributions (and in particular the Poisson itself) in queueing and inventory theory for, amongst integer-valued random variables, only the stuttering-Poisson has this property.

THE POLYA-AEPPLI (POISSON $\smile$ GEOMETRIC) DISTRIBUTION

The stuttering-Poisson distribution which Galliher *et al.* dealt with in detail has number of demands distributed geometrically $[P_x = (1-\psi)\psi^{x-1}$ $x > 0]$ and hence is Poisson $\smile$ geometric. Their Poisson and geometric distributions had means $(1 - \psi)\rho$ and $1/(1 - \psi)$ respectively, thus giving mean ρ and variance $\rho(1 + \psi)/(1 - \psi)$ for the Poisson $\smile$ geometric.

It is well known that the geometric distribution is a particular case of the negative binomial distribution. However, a variety of parameterisations of these distributions may be found in the literature and care needs to be taken in identifying the parameters. There are two main dichotomies —between the points at which the distribution effectively starts, i.e. $x = 0$ or $x = \mathrm{r}$ (where $\mathrm{r} = 1$ for the geometric) and between parameter bounds. These lead to four main types of PGF for each of the negative binomial and geometric distributions as follows:

	Negative Binomial	Geometric	
(i)	$p^k(1 - qs)^{-k}$	$p(1 - qs)^{-1}$	$\left\{\begin{array}{l} p + q = 1, \\ 0 \le p \le 1, \end{array}\right.$
(ii)	$p^k s^k(1 - qs)^{-k}$	$ps(1 - qs)^{-1}$	
(iii)	$(Q - Ps)^{-k}$	$(Q - Ps)^{-1}$	$\left\{\begin{array}{l} Q = 1 + P, P \ge 0 \\ \text{i.e. } Q = \frac{1}{p}, P = \frac{q}{p} \end{array}\right.$ (13)
(iv)	$s^k(Q - Ps)^{-k}$	$s(Q - Ps)^{-1}$	

Galliher's geometric is of type (ii) with $p = (1 - \psi)$, leading to a Poisson $\smallsmile$ geometric with PGF

$$G(s) = \exp[(1 - \psi)\rho\{(1 - \psi)s(1 - \psi s)^{-1} - 1\}] \qquad (14)$$
$$= \exp\left[\frac{(1 - \psi)\rho(s - 1)}{1 - \psi s}\right].$$

Hence $P_0 = \exp[-\rho(1 - \psi)]$ and the general recurrence relationship (7) becomes

$$(x + 1)P_{x+1} = (1 - \psi)^2\rho \sum_{i=0}^{x} (i + 1)\psi^i P_{x-i}. \qquad (15)$$

Bearing in mind the variations (13), we can immediately identify (14) as the PGF of the Polya–Aeppli distribution[22,25], which has been applied to a variety of biological data. Various authors have described its properties: in particular, Evans[7] points out that a recurrence relationship exists between the probabilities as follows (in our notation)

$$(x + 1)P_{x+1} = P_x[(1 - \psi)^2\rho + 2\psi x] - \psi^2(x - 1)P_{x-1}. \qquad (16)$$

Evans also notes that individual probabilities can be expressed in terms of confluent hypergeometric functions. Galliher *et al.* gave the probabilities in terms of Laguerre polynomials (which are, of course, particular confluent hypergeometrics). The recurrence relationship (16) can be derived as a consequence of the standard recurrence relationship between confluent hypergeometric functions. It is, however, interesting to note that our differentiation method yields both recurrence relationships as follows:

$$\Sigma i P_i s^{i-1} = \frac{\partial G}{\partial s} = \frac{(1 - \psi)^2}{(1 - \psi s)^2}\rho G(s) = (1 - \psi)^2\rho(1 + 2\psi s + 3\psi^2 s^2 + \ldots)\Sigma P_i s^i \qquad (17)$$

leads directly to (15), whilst rearranging this as

$$(1 - \psi s)^2 \Sigma i P_i s^{i-1} = (1 - \psi s)^2 \frac{\partial G}{\partial s} = (1 - \psi)^2\rho\Sigma P_i s^i \qquad (18)$$

leads to (16).

Differentiating the PGF also gives us the derivatives of the probabilities w.r.t. ρ and ψ:

$$\frac{\partial P_x}{\partial \rho} = \frac{1}{\rho}\{[x - \rho(1 - \psi)]P_x - \psi(x - 1)P_{x-1}\}, \qquad (19)$$

$$\frac{\partial P_x}{\partial \psi} = \frac{1}{(1 - \psi)}\{[\rho(1 - \psi) - 2x]P_x + (1 + \psi)(x - 1)P_{x-1}\}. \qquad (20)$$

Maximum-likelihood equations may thus be set up (as in the previous section, equation (10)) and solved iteratively to obtain estimates of ρ and ψ.

If demands per burst are distributed as a negative binomial instead of a geometric distribution, the resultant stuttering-Poisson is known both as the generalised Polya–Aeppli[25] and as the Poisson Pascal[15] distribution (note that Katti and Gurland's[15] treatment is based on parameterisation (iii) of (12)).

Jewell[14] and Kaufman and Cruon[16] have studied various properties of the stuttering-Poisson process based on Galliher's distribution.

LIMITING FORMS OF THE DISTRIBUTION

If we consider the Poisson $\vee$ geometric in terms of the a_i of (1), it is clear that provided ψ is reasonably small, the magnitude of a_i drops off rapidly with increasing i, since $a_1 = (1-\psi)^2\rho$ and $a_i = \psi^{i-1}a_1$, $(i > 1)$. This situation often occurs with stuttering-Poisson distributions. Kemp and Kemp[18] have shown that in such cases even if the ordinary Poisson distribution is a poor approximation, the Hermite distribution, whose PGF is

$$G(s) = \exp\,[a_1(s-1) + a_2(s^2-1)], \tag{21}$$

is often an acceptable approximation. Since it is of general applicability to distributions of class (4), the Hermite is particularly useful for empirical description of stuttering-Poisson data when there is uncertainty as to the exact form of $\varnothing(s)$. Properties of the Hermite are discussed in Kemp and Kemp[18] and the computer programme mentioned earlier fits by maximum likelihood a Poisson, a Hermite and a distribution (tentatively called the tri-Poisson), with PGF

$$G(s) = \exp\,[a_1(s-1) + a_2(s^2-1) + a_3(s^3-1)], \tag{22}$$

to any set of data believed to come from a stuttering-Poisson distribution.

REFERENCES

[1]ADELSON, R. M. (1966). 'Compound Poisson Distributions', *Opl. Res. Quart.* **17**, 73–75.

[2]BAILEY, N. T. J. (1964). *The Elements of Stochastic Processes*, Wiley, New York.

[3]BHARUCHA-REID, A. T. (1960). *Elements of the Theory of Markov Processes and Their Applications*, McGraw-Hill, New York.

[4]CASTOLDI, L. (1963). 'Poisson Processes with Events in Clusters', *Rend. Seminar Fac. Sci. Univ. Cagliari*, **33**, 433–437.

[5]COX, D. R. (1962). *Renewal Theory*, Methuen, London.

[6]CRESSWELL, W. L. and FROGGATT, P. (1963). *The Causation of Bus Driver Accidents: An Epidemiological Study*, Oxford University Press, London.

[7]EVANS, D. A. (1953). 'Experimental Evidence Concerning Contagious Distributions in Ecology', *Biometrika*, **40**, 186–211.

[8]FELLER, W. (1943). 'On a General Class of Contagious Distributions', *Ann. Math. Statist.* **14**, 389–400.

[9]FELLER, W. (1957). *An Introduction to Probability Theory and its Applications*, Vol. 1, 2nd Edition, Wiley, New York.

[10]GALLIHER, H. P., MORSE, P. M. and SIMOND, M. (1959). 'Dynamics of Two Classes of Continuous-Review Inventory Systems', *Opns. Res.* **7**, 362–383.

[11]GURLAND, J. (1957). 'Some Interrelations Among Compound and Generalised Distributions', *Biometrika*, **44**, 265–268.

[12]GURLAND, J. (1958). 'A Generalised Class of Contagious Distributions', *Biometrics*, **14**, 229–249.

[13]Haight, F. A. (1961). 'Index to the Distributions of Mathematical Statistics', *J. Res. N.B.S.*, **65B**, 23–60.

[14]Jewell, W. S. (1960). 'The Properties of Recurrent-Event Processes', *Opns. Res.* **8**, 446–472.

[15]Katti, S. K. and Gurland, J. (1961). 'The Poisson Pascal Distribution', *Biometrics*, **17**, 527–538.

[16]Kaufmann, A. and Cruon, R. (1959). 'Le Processus de Galliher', *Rev. Franc. de Recherche Operationelle*, **3**, 137–144.

[17]Kemp, C. D. (1967). 'On a Contagious Distribution Suggested for Accident Data', *Biometrics*, **23**, 241-255.

[18]Kemp, C. D. and Kemp, Adrienne W. (1965). 'Some Properties of the "Hermite" Distribution', *Biometrika*, **52**, 381–394.

[19]Khatri, C. G. and Patel, I. R. (1961). 'Three Classes of Univariate Distribution', *Biometrics*, **17**, 567–575.

[20]Luders, Von Rolf (1934). 'Die Statistik der Seltenen Ereignisse', *Biometrika*, **26**, 108–128.

[21]Maceda, E. Cansado (1948). 'On the Compound and Generalised Poisson Distributions', *Ann. Math. Statist.*, **19**, 414–416.

[22]Polya, G. (1930). 'Sur Quelques Points de la Theorie des Probabilities', *Ann. de L'Inst. Henri Poincare*, **1**, 117–161.

[23]Rao, C. R. (1952). *Advanced Statistical Methods in Biometric Research*, Wiley, New York.

[24]Saaty, T. L. (1959). *Mathematical Methods of Operations Research*, McGraw-Hill, New York.

[25]Skellam, J. G. (1952). 'Studies in Statistical Ecology I, Spacial Pattern', *Biometrika*, **39**, 346–362.

[26]Thyrion, P. (1960). 'Note sur les Distributions "par Grappes"', *Bull. Assoc. Roy. des Act. Belges*, **60**, 49–66.