

H. PAPAGEORGIU (Lawrence, Kansas)

## ON A BIVARIATE POISSON-GEOMETRIC DISTRIBUTION

A bivariate Poisson-geometric model is introduced with marginals a Poisson and a Poisson-geometric distribution, respectively. Various properties of this model are examined, including the derivation of conditional distributions and regression functions. Potential applications are also indicated.

**1. Introduction.** Univariate generalized (randomly stopped) discrete distributions with the structure

$$(1) \quad Z = Z_1 + Z_2 + \dots + Z_X,$$

where the  $Z_i$  are independent identically distributed random variables, also independent of  $X$ , arise in several areas of application, e.g. risk theory, ecology, biology, genetics, physics, operations research, etc., and for relevant references one can refer to the books by Johnson and Kotz [10] and Douglas [7]. Park [14] suggested that bivariate distributions  $(X, Z)$  with the structure (1) can also be used as appropriate models for many problems of social, economic and physical nature. Leiter and Hamdan [13] and Cacoullos and Papageorgiou ([2]–[4]) discussed a variety of bivariate  $(X, Z)$  discrete models with the structure (1) and examined the possible use of these models for analysing accident data.

In this paper a bivariate Poisson-geometric model is introduced by assuming that  $X$  follows a Poisson distribution and the  $Z_i$ ,  $i = 1, \dots, X$ , follow a geometric distribution. Various properties of this model are examined, including the derivation of the conditional distribution of  $X$  given  $Z$ . It is worth noting that in the univariate case (marginal distribution of  $Z$ ) the model was used by Galliher *et al* [8] in inventory policy problems.

**2. The bivariate Poisson-geometric model.** Let  $X$  be a Poisson random variable with probability function

$$f(x) = e^{-\lambda} \frac{\lambda^x}{x!}, \quad x = 0, 1, 2, \dots$$

Further, suppose that the probability function of the  $Z_i$  has a geometric distribution with parameter  $p$ . That is

$$(2) \quad P(Z_i = k) = pq^k, \quad k = 0, 1, 2, \dots, \quad 0 < p < 1, \quad q = 1 - p.$$

Now, since the  $Z_i$  are mutually independent, the conditional distribution of

$$Z = Z_1 + Z_2 + \dots + Z_X$$

given that  $X = x$  is a Pascal distribution with probability generating function (p.g.f.)

$$(3) \quad G_{Z|X=x}(v) = \left( \frac{p}{1 - qv} \right)^x.$$

The joint p.g.f. of  $(X, Z)$  is given by

$$(4) \quad G_{X,Z}(u, v) = \exp \left[ \lambda \left( u \frac{p}{1 - qv} - 1 \right) \right],$$

or

$$(5) \quad G_{X,Z}(u, v) = \exp \left[ \sum_{i=0}^{\infty} a_i (uv^i - 1) \right],$$

where  $a_i = \lambda pq^i$ .

The latter representation indicates possible applications of the model in operations research studies, and in particular in inventory problems ([8], [1]).

It is worth noting that when  $a_2, a_3, \dots$  become negligible compared with  $a_0, a_1$ , the p.g.f. (5) tends to

$$G(u, v) = \exp [a_0(u - 1) + a_1(uv - 1)]$$

which is the p.g.f. of a bivariate Poisson-Bernoulli model studied by Leiter and Hamdan [13] and Cacoullos and Papageorgiou [2].

The marginal p.g.f. of  $Z$  is

$$G_Z(v) = \exp \left[ \lambda \left( \frac{p}{1 - qv} - 1 \right) \right]$$

which is a special case of the Poisson-Pascal distribution ([15], p. 49) or,

$$G_Z(v) = \exp \left[ \sum_{i=0}^{\infty} a_i (v^i - 1) \right]$$

which is called a ‘stuttering’ or a composed or a multiple Poisson distribution ([1]; [10], p. 194; [11]).

The distribution of the sum  $W = X + Z$ ,

$$G_W(s) = \exp \left[ \lambda \left( \frac{ps}{1-qs} - 1 \right) \right]$$

is a Polya–Aeppli distribution ([15], p. 48).

### 3. Properties.

3.1. Probabilities. The joint probability function is given by

$$p(x, z) = f(x) p(z|x) = e^{-\lambda} \frac{\lambda^x}{x!} \binom{z+x-1}{z} p^x q^z$$

with

$$p(0, 0) = G(0, 0) = \exp(-\lambda).$$

Recurrences for the probabilities can be derived by differentiating the p.g.f. once with respect to a generating variable and then equating coefficients. Thus from equation (4)

$$(x+1)p(x+1, z) = \lambda p \cdot p(x, z) + (x+1)q \cdot p(x+1, z-1)$$

and from equation (5)

$$(x+1)p(x+1, z) = \lambda p \sum_{k=0}^z q^k p(x, z-k).$$

Similarly

$$(z+1)p(x, z+1) = 2zq \cdot p(x, z) + \lambda pq \cdot p(x-1, z) - (z-1)q^2 p(x, z-1)$$

and

$$(z+1)p(x, z+1) = \lambda p \sum_{k=0}^z (k+1)q^{k+1} p(x-1, z-k).$$

The probability function of the marginal distribution of  $Z$  is [5]

$$(6) \quad p(z) = \exp(\lambda q) \frac{q^z}{z!} C_{z,-1}(\lambda q),$$

where

$$C_{z,-1}(s) = \sum_{k=0}^z |C(z, k, -1)| s^k,$$

and the numbers  $|C(z, k, -1)|$  have been tabulated by Charalambides [6].

3.2. *Moments, cumulants, etc.* From the cumulant generating function

$$K(t_1, t_2) = \lambda \left( e^{t_1} \frac{p}{1 - qe^{t_2}} - 1 \right)$$

the expression

$$k_{i,j} = \lambda \mu'_j$$

relating the bivariate cumulants  $k_{i,j}$  to the moments about the origin  $\mu'_j$  of the geometric distribution (2), can be easily derived.

A corresponding expression for the factorial cumulants  $k_{(i,j)}$  can also be obtained, namely

$$k_{(i,j)} = \begin{cases} \lambda \mu_{(j)} = j! \lambda \left( \frac{q}{p} \right)^j, & \text{for } i = 0, 1, \\ 0, & \text{for } i \geq 2, \end{cases}$$

where  $\mu_{(j)}$  is the factorial moment of the geometric distribution (2) of order  $j$ .

As expected, [4]

$$E(Z) = \text{Cov}(X, Z) = \lambda \frac{q}{p}$$

and

$$\text{Corr}(X, Z) = \left( \frac{q}{q+1} \right)^{1/2}$$

independent of  $\lambda$ .

3.3. *Conditional distributions.* As regards the conditional distribution of  $Z$  given  $X = x$  this is clearly the  $x$ -fold convolution of the distribution of  $Z_i$ , with p.g.f. given by equation (3). The p.g.f. of the conditional distribution of  $X$  given  $Z = z$  can be derived using the formula (see [16]),

$$G_{X|Z=z}(u) = G_v^{(z)}(u, 0) / G_v^{(z)}(1, 0),$$

where we set

$$G_v^{(z)}(\alpha, \beta) = \frac{\partial^\alpha G(u, v)}{\partial v^z} \Big|_{v=\beta}^{u=\alpha},$$

and after some algebra

$$(7) \quad G_{X|Z=z}(u) = \exp \{ \lambda p(u-1) \} \frac{C_{z,-1}(\lambda p u)}{C_{z,-1}(\lambda p)}.$$

Note that the first factor on the right-hand side of (7) is the p.g.f. of the conditional distribution of  $X$  given  $Z = 0$ . The second factor represents the

p.g.f. of a combinatorial distribution in the sense of Harper [9] with probability function

$$P_k(z, -1, \lambda p) = \frac{|C(z, k, -1)|(\lambda p)^k}{C_{z,-1}(\lambda p)}.$$

From equation (3) it is clear that

$$E(Z|X = x) = \frac{q}{p}x.$$

Since the regression of  $X$  on  $Z = z$  can be written as

$$E(X|Z = z) = \sum_x \frac{x p(z|x) f(x)}{p(z)} = \sum_x \frac{x \binom{z+x-1}{z} p^x q^z f(x)}{p(z)}$$

by making use of the combinatorial identity

$$x \binom{x+z-1}{z} = (z+1) \binom{x+z}{z+1} - z \binom{x+z-1}{z},$$

we obtain

$$E(X|Z = z) = (z+1)q \frac{p(z+1)}{p(z)} - z.$$

Hence, by using equation (6) we have

$$E(X|Z = z) = \frac{C_{z+1,-1}(\lambda q)}{C_{z,-1}(\lambda q)} - z.$$

**4. Estimation.** Maximum likelihood estimates for the parameters  $\lambda$  and  $p$  can be derived by considering the likelihood function  $L$  given by

$$L = \prod_{j=1}^N e^{-\lambda} \frac{\lambda^{x_j}}{x_j!} \binom{z_j+x_j-1}{z_j} p^{x_j} q^{z_j},$$

where  $N$  is the size of a bivariate sample  $(x_j, z_j)$ ,  $j = 1, \dots, N$ , from the bivariate Poisson-geometric model.

The maximum likelihood estimators obtained are

$$\hat{p} = \frac{\bar{x}}{\bar{x} + \bar{z}}, \quad \hat{\lambda} = \bar{x}.$$

These estimates are identical to those obtained by the method of moments since

$$E(X) = \lambda, \quad E(Z) = \lambda \frac{q}{p}.$$

**5. Applications.** The Poisson-geometric distribution may serve as a proper model for several real problem situations. For example, see [8], we can assume that 'bursts' of demand are distributed as a Poisson, whilst the number of demands per burst is distributed according to the geometric distribution (2). Then, if we consider as bursts of demand the number of customers per day in a department store, the Poisson-geometric model is an appropriate one for analyzing the joint distribution of the total number of items required and the number of customers.

Another area of application is in data collected on the occurrence of illnesses in medicine. For instance, suppose that the number of patients contacting a physician in the general practice of medicine follows a Poisson distribution and that the number of episodes of illness experienced by a person in a given period of time (say a year) is given by the geometric (2). Defining an episode as 'the occurrence of a specific problem or illness in a patient extending over a period of time from its onset to resolution' (Kilpatrick [12]), equation (4) gives the joint distribution of the number of patients and the total number of medical cases treated by the physician.

#### References

- [1] R. M. Adelson, *Compound Poisson distributions*, *Opl. Res. Quart.* 17 (1966), p. 73-75.
- [2] T. Cacoullos and H. Papageorgiou, *On some bivariate probability models applicable to traffic accidents and fatalities*, *Int. Stat. Rev.* 48 (1980), p. 345-356.
- [3] —, —, *On bivariate discrete distributions generated by compounding*, in: *Statistical Distributions in Scientific Work*, C. Taillie et al. (eds.), Vol. 4, Reidel, Dordrecht 1981, p. 197-212.
- [4] —, —, *Bivariate negative binomial-Poisson and negative binomial-Bernoulli models with an application to accident data*, in: *Statistics and Probability: Essays in Honor of C. R. Rao*, G. Kallianpur et al. (eds.), North-Holland, 1982, p. 155-168.
- [5] Ch. A. Charalambides, *On the generalized discrete distributions and the Bell polynomials*, *Sankhyā B39* (1977), p. 36-44.
- [6] —, *On a restricted occupancy model and its applications*, *Biometrical Journal* 23 (1981), p. 601-610.
- [7] J. B. Douglas, *Analysis with standard contagious distributions*, International Co-operative Publishing House, Maryland 1980.
- [8] H. P. Galliher, R. M. Morse and M. Simond, *Dynamics of two classes of continuous-review inventory systems*, *Opns. Res.* 7 (1959), p. 362-383.
- [9] L. H. Harper, *Stirling behavior is asymptotically normal*, *Ann. Math. Statist.* 38 (1967), p. 410-414.
- [10] N. L. Johnson and S. Kotz, *Distributions in statistics: Discrete distributions*, Wiley, New York 1969.
- [11] C. D. Kemp, *'Stuttering-Poisson' distributions*, *J. Statist. and Social Inquiry Society of Ireland* 21 (1967), p. 151-157.
- [12] S. J. Kilpatrick, *The distribution of episodes of illness: A research tool in general practice*, *J. Roy. Coll. Gen. Pract.* 25 (1975), p. 686-690.

- [13] E. Leiter and M. A. Hamdan, *Some bivariate probability models applicable to traffic accidents and fatalities*, Int. Stat. Rex. 41 (1973), p. 87-100.
- [14] S. H. Park, *Regression and correlation in a bivariate distribution with different marginal densities*, Technometrics 12 (1970), p. 687-691.
- [15] G. P. Patil and S. W. Joshi, *A dictionary and bibliography of discrete distributions*, Oliver and Boyd, Edinburgh 1968.
- [16] K. Subrahmaniam, *A test for 'intrinsic' correlation in the theory of accident proneness*, J. Roy. Statist. Soc. B 28 (1966), p. 180-189.

DEPARTMENT OF MATHEMATICS  
THE UNIVERSITY OF KANSAS  
LAWRENCE, KANSAS 66045-2142, USA

*Received on 5. 10. 1983;*  
*revised version on 13. 12. 1983*

---