On the Characterization of a Bivariate Geometric Distribution

Sunil K. Dhar and Srinivasan Balaji

Department of Mathematical Sciences and
Center for Applied Mathematics and Statistics

New Jersey Institute of Technology, Newark, NJ 07102

CAMS Report 0203-010, Fall 2002

Center for Applied Mathematics and Statistics

NJIT

ON THE CHARACTERIZATION OF A BIVARIATE GEOMETRIC DISTRIBUTION

SUNIL K. DHAR AND SRINIVASAN BALAJI
Department of Mathematical Sciences and the Center for Applied Mathematics and Statistics
New Jersey Institute of Technology, *Newark*, *New Jersey* 07102

ABSTRACT

The discrete analogue of the bivariate Freund's models is generalized by way of characterization. This characterization is achieved by extending a key lemma in this area and is in terms of constant total failure rate and mixture geometric marginals. The generalized bivariate Freund's model includes a bivariate geometric distribution previously characterized.

Key words and Phrases: Bivariate geometric distribution; Failure rate; Characterization

1. INTRODUCTION

In this paper the discrete analogue of the Freund's model has been generalized and characterized. Arriving at a discrete analogue of bivariate continues lifetime random variables has been a tradition as old as Nair and Nair (1988), if not older. In their paper the Nairs' characterized a bivariate geometric distribution that is the discrete analogue of the Gumblel's bivariate exponential distribution. Dhar (1998a) arrived at another bivariate geometric distribution that was developed using ideas from Freund (1961) reliability models. Though this newly obtained geometric model is named after the author Freund (1961), not all the distribution properties are the same. In Dhar (1998b) the assumptions leading to the discrete analogue of the Freund's model are further interpreted.

Many authors have characterized different bivariate geometric distributions (BVG), giving additional insight into these models.

The bivariate distribution characterized by Sun and Basu (1995) is also the bivariate geometric distribution which was generalized to its multivariate case by Basu and Dhar (1995). This multivariate geometric distribution is the same discrete analogue of Marshall-Olkin's (1967) multivariate exponential (MVE) distribution with probability parameters strictly between 0 and 1. To see the equality of the joint distribution considered by Basu and Dhar (1995) with the Sun and Basu (1995, equation 1.1), one can re-parametrize the latter by taking the $p_{10} = (1-p_2)p_1p_{12}$ and $p_{01} = (1-p_1)p_2\, p_{12}$, and $p_{11} = p_1p_2p_{12}$. The paper Dhar (1998a) was also developed around the same time as Sun and Basu(1995) but was not characterized.

2. Generalization and characterization of the Freund's model

Consider the discrete analogue of the Freund's model from Dhar (1998a or b) for the sake of completeness:

$$P(M=m,N=n) = \frac{q_2q_3}{p_2p_3}\left[\frac{p_1p_2}{p_3}\right]^n p_3^m, \qquad \text{if } n<m,\ m,\ n=1,\ 2,\bullet\bullet\bullet,$$
$$= \frac{q_1q_4}{p_1p_4}\left[\frac{p_1p_2}{p_4}\right]^m p_4^n, \qquad \text{if } m<n,\ m,\ n=1,\ 2,\bullet\bullet\bullet, \qquad 2.1$$
$$= \frac{q_1q_2q_{12}}{1-p_1p_2}p_{12,}^{m-1} \qquad \text{if } m=n=1,\ 2,\bullet\bullet\bullet.$$

In order for the above function to be a density, we need $p_1p_2 < p_3$ and $p_1p_2 < p_4$, where $q's$ are $= 1-p's$. In Dhar(1998a) it is seen that the survival function corresponding to (2.1) satisfies the loss of memory property iff $p_{12} = p_1p_2$. The bivariate geometric model with the additional assumption $p_{12} = p_1p_2$, in (2.1) will be referred to as BVG($p_1,\ p_2,\ p_3,\ p_4$). The BVG ($p_1,\ p_2,\ p_3,\ p_4$) model from Dhar(1998a, Lemma 3.2), has its marginals as a mixture of two geometric distributions. The marginal density of M is

$$\frac{p_1-p_1p_2}{p_3-p_1p_2}(1-p_3)(p_3)^{m-1} + \frac{p_3-p_1}{p_3-p_1p_2}(1-p_1p_2)(p_1p_2)^{m-1},\ m \ge 1,$$

$0 < p_1 < p_3 < 1$ and that of N is 2.2

$$\frac{p_2-p_1p_2}{p_4-p_1p_2}(1-p_4)(p_4)^{n-1} + \frac{p_4-p_2}{p_4-p_1p_2}(1-p_1p_2)(p_1p_2)^{n-1},\ n \ge 1,$$

$$0 < p_2 < p_4 < 1.$$

From the proof of the following lemma we shall see that it is the generalization of the key result obtained by Sun and Basu(1995, Lemma 2.1). To state the following lemma we need the notations $g_1(t) = P\{M > N / \min(M,N) = t\}$, $g_2(t) = P\{M < N / \min(M,N) = t\}$ and $g_3(t) = P\{M = N / \min(M,N) = t\}$. Please note that $g_1 + g_2 + g_3 \equiv 1$.

Lemma 2.1 If (M,N) has constant total failure rate say $(q,\ q_1,\ q_2)$ and marginal density of M is mixture geometric $ab_1(1-b_1)^{m-1} + (1-a)b_2(1-b_2)^{m-1},\ m \ge 1,$.where $0 < b_1 \le b_2 < 1$ (otherwise replace a by $1-a$) and take the marginal density of N to be $cd_1(1-d_1)^{n-1} + (1-c)d_2(1-d_2)^{n-1},\ n \ge 1,$ where $0 < d_1 \le d_2 < 1$. Then, $g_1(t)$ or $g_2(t)$ are constant in t, if, either M or N is distributed geometric, also if $q = b_2$ or $q = d_2$, respectively. On the other hand, if either M or N is strictly mixture distributed geometric, with $q_1 \ne b_1$ or $q_2 \ne d_2$, then g_1 or g_2 is a linear combination of a geometric distribution function in t, with $0 < b_1 < q < b_2 < 1$ or $0 < d_1 < q < d_2 < 1$, respectively.

Proof. The proof of this lemma is along the lines of Sun and Basu (1995). Therefore, let us take the total failure rate to be $(q,\ q_1,\ q_2)$, where $0 < q,\ q_1,\ q_2 < 1$. Denote $p(m,n) = P(M=m,N=n) = P\{M=m,\ \min(M,N)=n,\ M>N\} + P\{N=n,\ \min(M,N)=m, M<N\} + P\{M=m,\ \min(M,N) =$

$$\begin{cases} P\{\min(M,N) = n\}P\{M > N/\min(M,N) = n\} \times \\ P\{M = m/M > N,\ \min(M,N) = n\},\ \text{for } m > n \ge 1, \\ P\{\min(M,N) = m\}P\{M < N/\min(M,N) = m\} \times \\ P\{N = n/M < N,\ \min(M,N) = m\},\ \text{for } n > m \ge 1, \\ P\{\min(M,N) = m\}P\{M = N/\min(M,N) = m\} \times \\ P\{N = m/M = N,\ \min(M,N) = m\},\ \text{for } m = n \ge 1 \end{cases} . \qquad 2.3$$

The last probability on the right hand side of the preceding statement is 1. Constant bivariate total failure rate implies that $\min(M,N)$ is geometric with density $q(1-q)^{m-1}$, $m \ge 1$, and for $m > n \ge 1$, $M/N = n$, has geometric density given by $P(M = m/N = n) = P(M = m/N = n,\ M > n) = q_1(1-q_1)^{m-n-1}$, and for $n > m \ge 1$, $N/M = m$, has geometric density given by $P(N = n/M = m) = P(N = n/M = m,\ N > m) = q_2(1-q_2)^{n-m-1}$. These equations in turn simplify (2.3) to

$$\begin{cases} q(1-q)^{n-1}g_1(n)q_1(1-q_1)^{m-n-1}, & \text{for } m > n \ge 1 \\ q(1-q)^{m-1}g_2(m)q_2(1-q_2)^{n-m-1}, & \text{for } n > m \ge 1 \\ q(1-q)^{m-1}g_3(m), & \text{for } m = n \ge 1 \end{cases} . \qquad 2.4$$

Summing up (2.4) over all values of $n \ge 1$, yields the following for the case $m = 1$, $ab_1 + (1-a)b_2 = q\{1 - g_1(1)\}$, which in turn gives $g_1(1) = \frac{q-[ab_1+(1-a)b_2]}{q}$ and for the case $m \ge 2$,

$$P(M = m) = \sum_{n=1}^{m-1} q(1-q)^{n-1}g_1(n)q_1(1-q_1)^{m-n-1} +$$

$$\sum_{n=m+1}^{\infty} q(1-q)^{m-1}g_2(m)q_2(1-q_2)^{n-m-1} + q(1-q)^{m-1}g_3(m).$$

The last equation in turn gives for $m \ge 2$

$$\frac{ab_1(1-b_1)^{m-1} + (1-a)b_2(1-b_2)^{m-1}}{qq_1(1-q_1)^m} = \qquad 2.5$$

$$\sum_{n=1}^{m-1} \frac{(1-q)^{n-1}g_1(n)}{(1-q_1)^{n+1}} + \frac{(1-q)^{m-1}}{q_1(1-q_1)^m}\{1 - g_1(m)\}$$

Now, subtract from equation (2.5) with m replaced by $m+1$, equation (2.5).This yields after some algebraic simplification, for $m \ge 2$,

$$q(1-q)^{m-1}g_1(m) - q(1-q)^m g_1(m+1) = \qquad 2.6$$

$$ab_1(1-b_1)^{m-1}(q_1 - b_1) +$$

$$(1-a)b_2(1-b_2)^{m-1}(q_1 - b_2) + q(q - q_1)(1-q)^{m-1}.$$

Notice the left hand side of (2.6) is essentially the lag one difference of $q(1-q)^{t-1}g_1(t)$. Therefore, applying the operation $\sum_{t=2}^{m}$ to both sides of (2.6) yields

$$q(1-q)^m g_1(m+1) = q(1-q)g_1(2) - a(q_1-b_1)(1-b_1)[1-(1-b_1)^{m-1}] - (1-a)(q_2-b_2)(1-b_2)[1-(1-b_2)^{m-1}] - (q-q_1)(1-q)[1-(1-q)^{m-1}]. \quad 2.7$$

In equation (2.5) substituting $m = 2$ yields

$$q(1-q)g_1(2) = a(q_1-b_1)(1-b_1) + (1-a)(q_1-b_2)(1-b_2) + (q-q_1)(1-q). \quad 2.8$$

Substituting (2.8) in (2.7) gives, for $m \geq 0$,

$$g_1(m+1) = \frac{a(q_1-b_1)(1-b_1)^m + (1-a)(q_1-b_2)(1-b_2)^m + (q-q_1)(1-q)^m}{q(1-q)^m}. \quad 2.9$$

Going through similar steps as in (2.4) to (2.9), yields, for $n \geq 0$,

$$g_2(n+1) = \frac{c(q_2-d_1)(1-d_1)^n + (1-c)(q_2-d_2)(1-d_2)^n + (q-q_2)(1-q)^n}{q(1-q)^n}. \quad 2.10$$

Please notice that, $P(M \geq 2, N \geq 1) > P(M \geq 2, N \geq 2)$ implies

$$a(1-b_1) + (1-a)(1-b_2) > 1-q. \quad 2.11$$

Now, if $a = 1$ or $b_1 = b_2$, then in either case

$$1-b_1 > 1-q \quad 2.12$$

and from equation (2.9) we can see that the right hand side of this equation is unbounded unless $q_1 = b_1$. This in turn implies that g_1 is a constant function. Similarly, in case $a = 0$ from (2.11) and (2.9) we can see that g_1 is a constant function. Again, similar reasoning as in (2.11) gives for the cases $c = 0$ or 1 or $d_1 = d_2$, g_2 is a constant function. Thus we have the same result as in Sun and Basu (1995). Let us now consider the case when $0 < a < 1$, and $b_1 < b_2$. Then from equation (2.11) we see that $b_1 < q$. We now claim that $q \leq b_2$. We shall prove this by the method of contradiction. Suppose that $q > b_2$, this implies $\frac{1-b_2}{1-q} > 1$. This, along with (2.12) and $b_1 < b_2$, yield the following possibilities: $q_1 - b_1 > q_1 - b_2 > 0$ is not possible because that would mean from (2.9) that g_1 is unbounded and also $0 > q_1 - b_1 > q_1 - b_2$ is not possible because for large enough m, $g_1(m+1)$ will be negative. Therefore, $q_1 - b_1 > 0 > q_1 - b_2$ must be true, because if either $q_1 - b_1$ or $q_1 - b_2 = 0$, would imply from (2.9) that for large enough m, $g_1(m+1)$ will be negative or g_1 is unbounded, respectively. Rewrite (2.9) as follows

$$\frac{g_1(m+1)}{\frac{(1-b_1)^m}{q(1-q)^m}} = \frac{(1-a)(q_1-b_2)(1-b_2)^m}{(1-b_1)^m} + a(q_1-b_1) + \frac{(q-q_1)}{\frac{(1-b_1)^m}{(1-q)^m}} \quad 2.13$$

Taking limit as $m \to \infty$ in the preceding equation we see that the left hand side of (2.13) goes to $a(q_1 - b_1)$, because $b_1 < b_2$. However, this impossible because we are in the case when a and $q_1 - b_1$ are both > 0, which would imply that g_1 is unbounded. Hence the contradiction. We have thus shown $b_1 < q \leq b_2$. Consider the case when $q = b_2$. Then, equation (2.9) reduces to

$$g_1(m+1) - \frac{a(q-q_1)}{q} = \frac{a(q_1-b_1)(1-b_1)^m}{q(1-q)^m}. \qquad 2.14$$

Please note that from (2.12) and (2.14), we see that $g_1(m+1) - \frac{a(q-q_1)}{q}$ is an unbounded function. Unboundedness of the preceding function is impossible unless $q_1 = b_1$. Therefore, $g_1 \equiv \frac{a(q-q_1)}{q}$. Please note that this is not surprising because $P(M \geq 2, N \geq 2) > P(M \geq 2)$ or $P(N \geq 2)$ gives $1 > q > q_i > 0,\ i = 1,2$. Now, consider the case $b_1 < q < b_2$. Rewrite (2.9) as follows

$$g_1(m+1) - \frac{(q-q_1)}{q} - (1-a)(q_1-b_2)\frac{(1-b_2)^m}{q(1-q)^m} = \frac{a(q_1-b_1)(1-b_1)^m}{q(1-q)^m}. \qquad 2.15$$

Notice that the right hand side of equation (2.15) diverges, which impossible because the left hand side is bounded. Unless, $q_1 = b_1$. This proves the desired lemma because the study of the case $g_2(n+1)$ is analogous to $g_1(m+1)$.■

To summarize the above result from the perspective of obtaining a generalized density, please note that, (M,N) has constant total failure rate say $(q,\ q_1,\ q_2)$, implies that the joint density satisfies

$$p(m,n) = \begin{cases} q(1-q)^{n-1}g_1(n)q_1(1-q_1)^{m-n-1}, & \text{for } m > n \geq 1, \\ q(1-q)^{m-1}g_2(m)q_2(1-q_2)^{n-m-1}, & \text{for } n > m \geq 1, \\ q(1-q)^{m-1}g_3(m), & \text{for } m = n \geq 1, \end{cases} \qquad 2.16$$

$$\text{where } g_1 + g_2 + g_3 \equiv 1.$$

Consider the additional assumption, M, N each mixture geometric $ab_1(1-b_1)^{m-1} + (1-a)b_2(1-b_2)^{m-1},\ m \geq 1$, where $0 < b_1 \leq b_2 < 1$ and $cd_1(1-d_1)^{n-1} + (1-c)d_2(1-d_2)^{n-1},\ n \geq 1$, where $0 < d_1 \leq d_2 < 1$,.respectively. Then, under the possibility $a = 0$ or $a = 1$, or $b_1 = b_2$, ($c = 0$ or $c = 1$, or $d_1 = d_2$) $g_1 \equiv \frac{q-q_1}{q}\left(g_2 \equiv \frac{q-q_2}{q}\right)$. Under the possibility $q = b_2$($q = d_2$) then

$$g_1 \equiv \frac{a(q-q_1)}{q}\left\{g_2 \equiv \frac{c(q-q_2)}{q}\right\}. \qquad 2.17$$

Further, under the possibility $b_1 < q < b_2$ ($d_1 < q < d_2$)

$$g_1(m+1) = \frac{(q-q_1)}{q} + (1-a)(q_1-b_2)\frac{(1-b_2)^m}{q(1-q)^m},\ m \geq 0$$

$$\left\{g_2(n+1) = \frac{(q-q_2)}{q} + (1-c)(q_2-d_2)\frac{(1-d_2)^n}{q(1-q)^n},\ n \geq 0\right\}.$$

Please note that (2.16) with conditions described around (2.17) gives rise to a generalized version of the BVG $(p_1,\ p_2,\ p_3,\ p_4)$ defined in (2.1) with $p_{12} = p_1 p_2$.

Theorem 2.2 The bivariate random variable (M,N) has constant total failure rate say

$(1-p_1p_2,\ q_3,\ q_4)$ [these parameters are from equation 2.1], and marginal densities of *M* and *N* are mixture geometric as in (2.2) iff (*M*,*N*) has joint density as in (2.1) with loss of memory property and marginals mixture geometric as in (2.2).

Proof. Let (*M*,*N*) have constant total failure rate say $(1-p_1p_2,\ q_3,\ q_4)$ (please see equation 2.1) with marginals as in (2.2). In (2.16) and the parameters of the Lemma 2.1.please take $b_1 = q_1 = [1-p_3]$, $d_1 = q_2 = [1-p_4]$, $b_2 = d_2 = q = [1-p_1p_2]$, $a = \left[\frac{p_1-p_1p_2}{p_3-p_1p_2}\right]$ and $c = \left[\frac{p_2-p_1p_2}{p_4-p_1p_2}\right]$, where all the parameters in the square bracket are from (2.1). Since, in this case $b_2 = d_2 = q$,

$$g_1 \equiv \left[\frac{p_1-p_1p_2}{1-p_1p_2}\right],\ g_2 \equiv \left[\frac{p_2-p_1p_2}{1-p_1p_2}\right] \text{ and } g_3 = \left[\frac{p_1-p_1p_2}{p_3-p_1p_2}\right]. \qquad 2.18$$

Substituting (2.18) in (2.16) gives the bivariate discrete analogue of Freund's model described in (2.1) and (2.2), with $p_{12} = p_1p_2$.
On the other hand let (*M*,*N*) have joint density as in (2.1) with loss of memory property. Then simple algebra shows that (*M*,*N*) has total failure rate given by $(1-p_1p_2,\ q_3,\ q_4)$.■

REFERENCES

Basu, A. P. and Dhar, S. K. "Bivariate geometric distribution." *Journal of Applied Statistical Science* Vol. 2 No. 1, (1995) pp. 33-44.

Dhar, Sunil K. "Data Analysis with discrete analogue of Freund's model." *Journal Applied Statistical Science*, Volume 7,(1998a) pp. 169-183.

Dhar, Sunil K. "Modeling with a bivariate geometric distribution." *Advances on Methodological and Applied Aspects of Probability and Statistics*, Edited by N. Balakrishnan, Gordon and Breach Science publishers, IISA 98 McMaster Conference Proceedings, Volume 1(1998b), pp. 101-109.

Freund, J. E. "A bivariate extension of the exponential distribution." *J. Amer. Statist. Assoc.* Vol. 56, (1961) pp. 971-977.

Marshall, A. W. and Olkin, I. "A multivariate exponential distribution." *J. Amer. Statist. Assoc.* Vol. 62, (1967) pp.30-44.

Sun, K. and Basu A. P. "A characterization of a bivariate geometric distribution " *Statistic & Probability Letters* Vol. 23, (1995) pp.307-311.