Noname manuscript No.
(will be inserted by the editor)

Properties and estimation of a bivariate geometric model with locally constant failure rates

Alessandro Barbiero

Received: date / Accepted: date

Abstract Stochastic models for correlated count data have been attracting a lot of interest in the recent years, due to their many possible applications: for example, in quality control, marketing, insurance, health sciences, and so on. In this paper, we revise a bivariate geometric model, introduced by Roy (1993), which is very appealing, since it generalizes the univariate concept of constant failure rate - which characterizes the geometric distribution within the class of all discrete random variables - in two dimensions, by introducing the concept of "locally constant" bivariate failure rates. We mainly focus on four aspects of this model that have not been investigated so far: 1) pseudo-random simulation, 2) attainable Pearson's correlations, 3) stress-strength reliability parameter, and 4) parameter estimation. A Monte Carlo simulation study is carried out in order to assess the performance of the different estimators proposed and application to real data, along with a comparison with alternative bivariate discrete models, is provided as well.

Keywords attainable correlations · correlated counts · failure rate · Gumbel-Barnett copula · method of moments · mean residual life · stress-strength model

1 INTRODUCTION

In recent years, the construction of bivariate (and multivariate) discrete distributions has attracted much interest, since stochastic models for correlated count data find application in many fields. For example, in marketing, modeling the number of purchases of different products is of special interest for

A. Barbiero
Department of Economics, Management and Quantitative Methods,
Università degli Studi di Milano, via Conservatorio 7, 20122 Milan, Italy
Tel.: +390250321533
Fax: +3902503111
E-mail: alessandro.barbiero@unimi.it

predicting sales in the future or examining the behavior of different types of buyers. In insurance risk applications, the numbers of claims in different classes are frequently statistically dependent. In health economics, different components of health care demand, such as the number of consultations with a doctor or specialist and the number of consultations with non-doctor health professionals, can be jointly modeled. Multivariate discrete distributions are also developed to explicitly model the influence of correlated failures in multi-version software (Fondella and Zeephongsekul 2016).

Many authors have discussed the problem of constructing a bivariate version of a given univariate distribution, although there is no universally accepted criterion for producing a unique distribution which can unequivocally be called the bivariate version of a univariate distribution. In Galambos and Kotz (1978), however, bivariate generalization of probability distributions through characterization results is recommended. Here we focus on the geometric distribution, which along with the Poisson is perhaps the most popular discrete distribution used for modeling counts, especially in the reliability context. The probability mass function (p.m.f.) of a geometric distribution with parameter $\theta \in (0,1)$ is $p(x) = (1-\theta)\theta^x$, for $x = 0, 1, \ldots$; its cumulative distribution function (c.d.f.) is $F(x) = P(X \leq x) = 1 - \theta^{x+1}$; its failure rate function $r(x) = p(x)/P(X \geq x)$ is constant and equal to $1 - \theta$. The constancy of the latter function characterizes the geometric among all the discrete distributions (Xekalaki 1983). Expected value and variance are equal to $\theta/(1-\theta)$ and $\theta/(1-\theta)^2$, respectively.

There have been several proposals for constructing a bivariate version of the geometric distribution. Hawkes (1972) proposed a bivariate discrete distribution as “the most natural generalization of the geometric distribution”, which Sun and Basu (1995) later proved to be characterized by a constant “total failure rate” (a 3-dimensional vector whose components are (conditional) failure rates). The model is defined through the joint survival function (s.f.):

$$P(X \geq x, Y \geq y) = \begin{cases} p_{11}^y (p_{10} + p_{11})^{x-y} \text{ for } x \geq y \\ p_{11}^x (p_{01} + p_{11})^{y-x} \text{ for } y \geq x \end{cases} \tag{1}$$

where $p_{00} + p_{10} + p_{01} + p_{11} = 1$, $p_{10}, p_{01}, p_{11} > 0$, $p_{00} \geq 0$. From (1), the joint p.m.f. can be derived as:

$$P(X = x, Y = y) = \begin{cases} p_{11}^y (p_{10} + p_{11})^{x-y-1} p_{10}(p_{01} + p_{00}) & \text{for } x > y \\ p_{11}^y p_{00} & \text{for } x = y \\ p_{11}^x (p_{01} + p_{11})^{y-x-1} p_{01}(p_{10} + p_{00}) & \text{for } y > x \end{cases} \tag{2}$$

It can be shown that the marginal distributions of X and Y are still geometric with parameters $p_{11}+p_{10}$ and $p_{11}+p_{01}$, respectively. Paulson and Uppuluri (1972), through a characteristic-functional equation, introduced a five-parameter bivariate geometric distribution whose p.m.f. can be computed only through recursive formulas. Phatak and Sreehari (1981) introduced a two-parameter bivariate geometric distribution with an easy expression of its p.m.f.; it was later studied also by Krishna and Pundir (2009). In Basu and Dhar (1995) a bivariate geometric distribution is proposed by a discrete analog of the continuous bivariate distribution of Marshall and Olkin (Marshall and Olkin 1967); in Dhar (1998) another bivariate geometric model, which is a discrete analog to Freund's model (Freund 1961), is introduced. Roy (1993) extended the univariate concept of failure rate in two dimensions, by introducing the bivariate failure rates, and proposed a three-parameter bivariate geometric distribution enjoying an analogous property to its univariate version, i.e., locally constant bivariate failure rates.

Here we will examine other relevant features of this latter model that have not been investigated so far neither by Roy (1993) nor by subsequent works, namely, simulation, dependence structure and correlations, and methods for point estimation. These aspects are essential if a researcher wants to apply the model to real data with discernment. The paper is structured as follows: in the next section, we briefly revise the bivariate geometric model introduced by Roy (1993); in Section 3 we address pseudo-random number simulation; Section 4 computes Pearson's correlation and outlines its lower and upper bounds as a function of marginal parameters; in Section 5 the stress-strength reliability parameter is computed and analysed; Section 6 describes possible methods for point estimation; Section 7 illustrate a Monte Carlo simulation study aimed at empirically comparing the point estimators of Section 6; Section 8 describes an application to real data, and Section 9 provides some final remarks.

2 THE BIVARIATE GEOMETRIC DISTRIBUTION WITH LOCALLY CONSTANT FAILURE RATES

The joint p.m.f. of the bivariate geometric distribution introduced by Roy (1993) is given by

$$p(x,y) = \theta_1^x \theta_2^y \theta_3^{xy} \left[1 - \theta_1\theta_3^y - \theta_2\theta_3^x + (\theta_1\theta_2\theta_3)\theta_3^{x+y}\right], \qquad x, y \in \mathbb{Z}_0^+ \tag{3}$$

with $0 < \theta_1 < 1$, $0 < \theta_2 < 1$, $0 < \theta_3 \leq 1$ such that $\theta_3 \geq (\theta_1 + \theta_2 - 1)/(\theta_1\theta_2)$. The corresponding bivariate s.f., defined as $S(x, y) = P(X \geq x, Y \geq y)$, takes the following compact expression:

$$S(x,y) = \theta_1^x \theta_2^y \theta_3^{xy}. \tag{4}$$

The joint c.d.f. can be computed as

$$F(x,y) = 1 - \theta_1^{x+1} - \theta_2^{y+1} + \theta_1^{x+1}\theta_2^{y+1}\theta_3^{(x+1)(y+1)}.$$

It has been shown that this bivariate geometric distribution admits marginal geometric distributions for X and Y with parameters θ_1 and θ_2, respectively; so θ_1 and θ_2 have a natural and clear interpretation. The other parameter θ_3 may be viewed as a parameter of dependency, where $\theta_3 = 1$ ensures independence between X and Y.

For this model, the bivariate failure rates take the following expressions:

$$\begin{aligned}\lambda_1(x,y) &:= P(X=x, Y\geq y)/P(X\geq x, Y\geq y) = 1-\theta_1\theta_3^y\\ \lambda_2(x,y) &:= P(Y=y, X\geq x)/P(Y\geq y, X\geq x) = 1-\theta_2\theta_3^x,\end{aligned}$$

which allows us to claim that this bivariate geometric distribution possesses "locally constant" failure rates, since $\lambda_1(x,y)$ is only a function of y and $\lambda_2(x,y)$ is only a function of x; Roy (1993) proved it is the unique bivariate discrete distribution enjoying this property.

An analogous property is met by the "bivariate mean residual lives", defined by Roy (1993) as

$$\begin{aligned}\mu_1(x,y) &:= \mathbb{E}(X>x|X>x, Y>y)\\ \mu_2(x,y) &:= \mathbb{E}(Y>y|X>x, Y>y),\end{aligned}$$

which represent a bivariate counterpart to univariate mean residual lives. The following relationship linking bivariate failure rates and mean residual holds:

$$\lambda_1(x,y)\mu_1(x,y-1) = 1 = \lambda_2(x,y)\mu_2(x-1,y) \quad x,y\in\mathbb{Z}_0^+. \tag{5}$$

3 PSEUDO-RANDOM SIMULATION

In order to simulate a pseudo-random pair (x,y) from the bivariate geometric distribution (3) with parameters θ_1, θ_2, and θ_3, one can first simulate from the marginal distribution of X and then, conditionally on the sampled x, simulate from Y.

The conditional p.m.f. of Y given $X=x$ can be computed as

$$\begin{aligned}p_{y|x}(y) = p(x,y)/p(x) &= \theta_1^x\theta_2^y\theta_3^{xy}\left[1-\theta_1\theta_3^y-\theta_2\theta_3^x+(\theta_1\theta_2\theta_3)\theta_3^{x+y}\right]/[\theta_1^x(1-\theta_1)]\\ &= \theta_2^y\theta_3^{xy}\left[1-\theta_1\theta_3^y-\theta_2\theta_3^x+(\theta_1\theta_2\theta_3)\theta_3^{x+y}\right]/(1-\theta_1);\end{aligned}$$

then, the conditional c.d.f. of Y given $X=x$ is provided as

$$\begin{aligned}F_{y|x}(y) &= \sum_{s=0}^{y} p_{s|x}(s|x) = \sum_{s=0}^{y}\theta_2^s\theta_3^{xs}\left[1-\theta_1\theta_3^s-\theta_2\theta_3^x+(\theta_1\theta_2\theta_3)\theta_3^{x+s}\right]/(1-\theta_1)\\ &= \sum_{s=0}^{y}[(\theta_2\theta_3^x)^s-\theta_1(\theta_2\theta_3^{x+1})^s-(\theta_2\theta_3^x)^{s+1}+\theta_1(\theta_2\theta_3^{x+1})^{s+1}]/(1-\theta_1)\\ &= \sum_{s=0}^{y}[(1-\theta_2\theta_3^x)(\theta_2\theta_3^x)^s-\theta_1(1-\theta_2\theta_3^{x+1})(\theta_2\theta_3^{x+1})^s]/(1-\theta_1)\\ &= \left[(1-\theta_2\theta_3^x)\frac{1-(\theta_2\theta_3^x)^{y+1}}{1-\theta_2\theta_3^x}-\theta_1(1-\theta_2\theta_3^{x+1})\frac{1-(\theta_2\theta_3^{x+1})^{y+1}}{1-\theta_2\theta_3^{x+1}}\right]/(1-\theta_1)\\ &= [1-(\theta_2\theta_3^x)^{y+1}-\theta_1(1-(\theta_2\theta_3^{x+1})^{y+1})]/(1-\theta_1). \end{aligned}\tag{6}$$

Note that if $\theta_3 = 1$, corresponding to the independence case, then the expression of the conditional c.d.f. (6) above correctly reduces to

$$F_{y|x}(y) = [1 - \theta_2^{y+1} - \theta_1(1 - \theta_2^{y+1})]/(1 - \theta_1) = 1 - \theta_2^{y+1},$$

which corresponds to the marginal c.d.f. of Y (geometrically distributed with parameter θ_2). Straightforward but tedious calculations lead to the following expression for the conditional expectation $\mathbb{E}(Y|x)$:

$$\mathbb{E}(Y|x) = \frac{\theta_2\theta_3^x}{(1-\theta_2\theta_3^x)(1-\theta_2\theta_3^{x+1})}\frac{1-\theta_2\theta_3^{x+1}-\theta_1\theta_3+\theta_1\theta_2\theta_3^{x+1}}{1-\theta_1}.$$

If we set $\theta_3 = 1$, the conditional moment above consistently reduces to $\theta_2/(1-\theta_2)$.

The simulation procedure then translates into implementing the following steps:

1. Let u, v be two independent uniform random numbers in $(0, 1)$
2. Set $x = G_{\theta_1}^{\leftarrow}(u)$, with G_{θ_1} denoting the c.d.f. of a geometric random variable with parameter θ_1 and $G_{\theta_1}^{\leftarrow}$ its generalized inverse: $G_{\theta_1}^{\leftarrow}(u) := \inf\left\{z \in \mathbb{Z}_0^+ : G_{\theta_1}(z) \geq u\right\} = \left\lceil \frac{\ln(1-u)}{\ln\theta_1} \right\rceil - 1$, with $\lceil\cdot\rceil$ denoting the ceiling function
3. Set $y = F_{y|x}^{\leftarrow}(v)$. Since $F_{y|x}(y)$ is not analytically invertible, one needs to resort to some root-finding routine in order to find the non-negative integer value y such that $F_{y|x}(y) \geq v$ and $F_{y|x}(y-1) < v$. This can be carried out numerically by solving the equation $F_{y|x}(y) - v = 0$ in $\mathbb{R}$, and then rounding its unique solution y^* to $y = \lceil y^* \rceil$
4. (x, y) is a random pair from the bivariate geometric distribution (3) with parameters θ_1, θ_2, and θ_3.

4 PEARSON'S CORRELATION AND DEPENDENCE STRUCTURE

For the bivariate model at study, it would be useful to determine the range of dependence in terms of linear correlation ρ_{xy}. Recalling that $\rho_{xy} = (\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y))/\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}$, in order to compute ρ_{xy}, we need to calculate the mixed moment $\mathbb{E}(XY)$. We have

$$
\begin{aligned}
\mathbb{E}(XY) &= \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} xyp(x,y) = \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} xy\theta_1^x\theta_2^y\theta_3^{xy}\left[1-\theta_1\theta_3^y-\theta_2\theta_3^x+(\theta_1\theta_2\theta_3)\theta_3^{x+y}\right] \\
&= \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} xy\theta_1^x(\theta_2\theta_3^x)^y - \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} x\theta_1^{x+1}y(\theta_2\theta_3^{x+1})^y + \\
&- \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} x\theta_1^x y(\theta_2\theta_3^x)^{y+1} + \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} x(\theta_1\theta_3)^{x+1}y\theta_2(\theta_2\theta_3^{x+1})^y \\
&= \theta_2\sum_{x=0}^{\infty}\frac{x(\theta_1\theta_3)^x}{(1-\theta_2\theta_3^x)^2} - \theta_2\sum_{x=0}^{\infty}\frac{x(\theta_1\theta_3)^{x+1}}{(1-\theta_2\theta_3^{x+1})^2} - \theta_2^2\sum_{x=0}^{\infty}x\frac{(\theta_1\theta_3^2)^x}{(1-\theta_2\theta_3^x)^2} + \theta_2^2\sum_{x=0}^{\infty}\frac{x(\theta_1\theta_3^2)^{x+1}}{(1-\theta_2\theta_3^{x+1})^2} \\
&= \theta_2\sum_{x=0}^{\infty}\frac{(\theta_1\theta_3)^{x+1}}{(1-\theta_2\theta_3^{x+1})^2} - \theta_2^2\sum_{x=0}^{\infty}\frac{(\theta_1\theta_3^2)^{x+1}}{(1-\theta_2\theta_3^{x+1})^2} \\
&= \sum_{x=0}^{\infty}\frac{(\theta_1\theta_3)^{x+1}\theta_2[1-\theta_2\theta_3^{x+1}]}{(1-\theta_2\theta_3^{x+1})^2} = \theta_1\theta_2\theta_3\sum_{x=0}^{\infty}\frac{(\theta_1\theta_3)^x}{1-\theta_2\theta_3^{x+1}},
\end{aligned}
$$

and then, recalling the expressions of expected value and variance for the univariate geometric distribution,

$$
\begin{aligned}
\rho_{xy} &= \frac{\theta_1\theta_2\theta_3\sum_{x=0}^{\infty}\frac{(\theta_1\theta_3)^x}{1-\theta_2\theta_3^{x+1}} - \frac{\theta_1\theta_2}{(1-\theta_1)(1-\theta_2)}}{\theta_1\theta_2/[(1-\theta_1)(1-\theta_2)]^2} \\
&= \sqrt{\theta_1\theta_2}\left[\theta_3(1-\theta_1)(1-\theta_2)\sum_{x=0}^{\infty}\frac{(\theta_1\theta_3)^x}{1-\theta_2\theta_3^{x+1}} - 1\right]. \qquad (7)
\end{aligned}
$$

Note how ρ_{xy} depends not only on the dependence parameter θ_3, but also on the two marginal parameters θ_1 and θ_2. Since ρ_{xy} can be also rewritten as

$$
\rho_{xy} = \sqrt{\theta_1\theta_2}(1-\theta_1)(1-\theta_2)\sum_{x=0}^{\infty}\left[\frac{\theta_3(\theta_1\theta_3)^x}{1-\theta_2\theta_3^{x+1}} - \frac{\theta_1^x}{1-\theta_2}\right]
$$

and being

$$
\frac{\theta_3(\theta_1\theta_3)^x}{1-\theta_2\theta_3^{x+1}} \le \frac{\theta_1^x}{1-\theta_2\theta_3^{x+1}} \le \frac{\theta_1^x}{1-\theta_2}, \forall x \ge 0,
$$

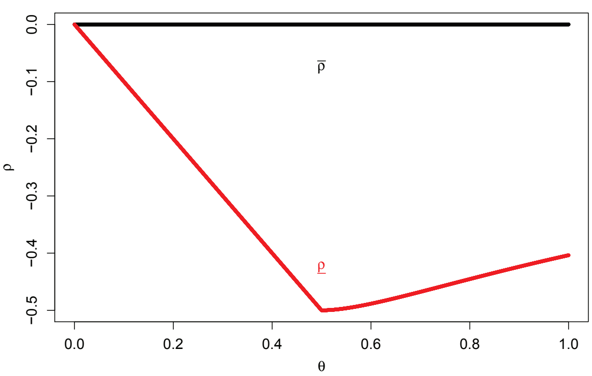
we have that $\rho_{xy} \le 0$ for any possible choice of θ_3, the equality holding if and only if $\theta_3 = 1$. Thus the bivariate geometric distribution can model only negative dependence. Furthermore, being $f(\theta_3;\theta_1,\theta_2,x) = \theta_3(\theta_1\theta_3)^x/(1-\theta_2\theta_3^{x+1})$ an increasing function of θ_3 for any fixed θ_1, θ_2, and $x \in \mathbb{Z}_0^+$; we can state that ρ_{xy} is an increasing function of θ_3, with $0 < \theta_3 \le 1$ and $\theta_3 \ge (\theta_1+\theta_2-1)/(\theta_1\theta_2)$, for any given $0 < \theta_1 < 1$, $0 < \theta_2 < 1$. Table 1 reports the values of ρ_{xy} for several parameter combinations. If we consider the particular case $\theta_1 = \theta_2 = \theta$ (identical marginal distributions), the constraint on θ_3 becomes $0 < \theta_3 \le 1$

Table 1 Pearson's correlation for the bivariate geometric model (3) for several parameters' combinations, computed by approximating the infinite series sum in Eq. (7)

θ_1	θ_2	θ_3	ρ_{xy}
0.3	0.3	0.1	-0.284
0.3	0.7	0.1	-0.448
0.5	0.5	0.1	-0.486
0.3	0.3	0.3	-0.247
0.3	0.7	0.3	-0.419
0.5	0.5	0.3	-0.449
0.3	0.3	0.6	-0.171
0.3	0.7	0.6	-0.342
0.5	0.5	0.6	-0.355
0.6	0.6	0.6	-0.470
0.3	0.3	0.9	-0.055
0.3	0.7	0.9	-0.153
0.5	0.5	0.9	-0.148
0.6	0.6	0.9	-0.231
0.7	0.7	0.9	-0.354

and $\theta_3 \geq (2\theta-1)/\theta^2$; i.e, simply $0 < \theta_3 \leq 1$ if $\theta \leq 1/2$ and $(2\theta-1)/\theta^2 \leq \theta_3 \leq 1$ if $\theta > 1/2$. For any value of the common marginal parameter θ, we can evaluate which is the range of the correlation coefficient ρ_{xy}, obtained by varying the dependence parameter θ_3. In Figure 1, we plot the lower bound of Pearson's correlation ρ as a function of θ, $\underline{\rho}(\theta) = \inf_{\theta_3} \rho_{xy}$ (the upper bound being $\overline{\rho}(\theta) = 0$ for any choice of θ, which is obtained by setting $\theta_3 = 1$). Note that from (7), by setting $\theta_1 = \theta_2 = \theta \leq 1/2$ and letting θ_3 go to zero, we obtain that $\underline{\rho} = -\theta$ if $\theta \leq 1/2$. Moreover, the absolute infimum of $\underline{\rho}(\theta)$, $\inf_\theta \inf_{\theta_3} \rho_{xy}$, is just obtained for $\theta = 1/2$ letting θ_3 go to zero and is equal to $\underline{\underline{\rho}} = -1/2$.

Fig. 1 Lower and upper bounds for Pearson's correlation for the bivariate geometric model with identical geometric margins with parameter θ



We must remark that for two identically distributed geometric margins with common parameter θ the maximum attainable linear correlation is 1, which is achieved by linking them through the comonotonicity copula $M(u,v) = \min\{u,v\}$, whereas the minimum attainable correlation, which is achieved by linking them through the countermonotonicity copula $W(u,v) = \max\{u+v-1,0\}$, is equal to $-\theta$, if $\theta \leq 1/2$, whereas for $\theta > 1/2$ it can be numerically computed by resorting to the algorithm provided in Huber and Maric (2014) and is in any case smaller than -0.5.

We close this section with some remarks about the dependence structure of the bivariate geometric model. Since for the geometric margins the s.f.s are given by $S_1(x) = \theta_1^x$ and $S_2(y) = \theta_2^y$, we can write the joint s.f. $S(x,y)$ (4) as

$$S(x,y) = \hat{C}(S_1(x), S_2(y)), \tag{8}$$

with $\hat{C}$ being the survival copula

$$\hat{C}(u,v) = uv \exp\left(-\theta \ln u \ln v\right), \tag{9}$$

where $\theta = -\ln\theta_3/(\ln\theta_1 \ln\theta_2)$. The copulas (9), with the parameter θ in $[0,1)$ (see also Eq.(4.2.9) in Nelsen (1999)), are the survival copulas associated with Gumbel's bivariate exponential distribution (Gumbel 1960). These copulas are also known as Gumbel-Barnett family, as Barnett (1980) first discussed it as a family of copulas. Requiring that the parameter θ in (9) lies in $[0,1)$ results in $\theta_3 \leq 1$ and $\ln\theta_3 > -\ln\theta_1 \ln\theta_2$; this constraint is more stringent than the actual constraint on θ_3 for the bivariate geometric model. Note that for $\theta = 0$ (i.e., $\theta_3 = 1$), the function in (9) reduces to the independence copula $\Pi(u,v)$.

Recall that Sklar's theorem (Sklar 1959) establishes that for a bivariate model the survival copula $\hat{C}$ that can be extracted from the joint s.f. is unique if and only if the marginal s.f.s are both continuous; if not, as in this case, there are infinite survival copulas associated with the joint s.f.. All these copulas $\hat{C}(u,v)$ should however satisfy Eq.(8) for any $(u,v) \in \mathrm{Ran}(S_1) \times \mathrm{Ran}(S_2)$, where $\mathrm{Ran}(S_i)$ denotes the range of S_i, $i = 1,2$.

Since $F(x,y) \leq F_1(x)F_2(y) = (1-\theta_1^{x+1})(1-\theta_2^{y+1}) = 1 - \theta_1^{x+1} - \theta_2^{y+1} + \theta_1^{x+1}\theta_2^{y+1}$, for every $(x,y) \in \mathbb{Z}_0^+ \times \mathbb{Z}_0^+$, we conclude that the bivariate geometric model possesses negative quadrant dependence (NQD, see Mari and Kotz (2001, p.34)). We recall that negative quadrant dependence is shown to be a stronger notion of dependence than the negative (Pearson) correlation.

Indeed, it can be proved that the bivariate geometric distribution possesses a stronger form of dependence, namely negative regression dependence (NRD). A bivariate random variable (X,Y) is said to enjoy the NRD property if its conditional c.d.f. $F_{y|x}(y)$ is an increasing function of x for all y (Mari and Kotz 2001, p.38). The conditional c.d.f. (6), for any y, is an increasing function of x: in fact, if we compute its first order derivative with respect to x we obtain

$$\frac{\partial F_{y|x}(y)}{\partial x} = \frac{\ln\theta_3 (y+1)\theta_2^{y+1}(\theta_3^{y+1})^x(-1+\theta_1\theta_3^{y+1})}{1-\theta_1}$$

which is larger than zero for any choice of x and y if $\theta_3 < 1$, and equal to zero if $\theta_3 = 0$. Thus we can conclude that for $\theta_3 \neq 1$ the function (6) is increasing in x for all y and thus the bivariate geometric distribution possesses the NRD property.

5 STRESS-STRENGTH RELIABILITY PARAMETER

The probability $R = P(X \leq Y)$ (and $P(X < Y)$) has been often investigated in the statistical literature. If Y is the strength of a (mechanical, electrical,...) system which is subject to a stress X, and the system regularly operates unless the stress exceeds the strength, then R is the probability that the system works, i.e. a measure of system performance. Most of the papers studying the computation and estimation of R deal with independent continuous probability distributions for stress and strength. However, in some real life situations, stress or strength can be modeled by discrete distribution. For example, when the stress is the number of the products that customers want to buy and the strength is the number of the products that factory produces (Jovanović 2017). Furthermore, stress and strength may be modeled as non-independent random variables; this can be justified since a system that have to resist to higher levels of stress is designed to have higher levels of strength (thus implying a positive dependence/correlation between stress and strength).

For the proposed bivariate geometric distribution, the stress-strength parameter R can be computed as

$$R = P(X \leq Y) = \sum_{y=0}^{\infty} \sum_{x=0}^{y} p(x, y),$$

with $p(x, y)$ given by Eq.(3), and thus assumes the following expression:

$$\begin{aligned}
R &= \sum_{y=0}^{\infty} \sum_{x=0}^{y} \theta_1^x \theta_2^y \theta_3^{xy} \left[1 - \theta_1 \theta_3^y - \theta_2 \theta_3^x + (\theta_1 \theta_2 \theta_3) \theta_3^{x+y}\right] \\
&= \sum_{y=0}^{\infty} \sum_{x=0}^{y} \theta_2^y (\theta_1 \theta_3^y)^x - \theta_2 (\theta_1 \theta_3^y)^{x+1} - \theta_2^{y+1} (\theta_1 \theta_3^{y+1})^x + \theta_2^{y+1} (\theta_1 \theta_3^{y+1})^{x+1} \\
&= \sum_{y=0}^{\infty} \theta_2^y \frac{1 - (\theta_1 \theta_3^y)^{y+1}}{1 - \theta_1 \theta_3^{y+1}} (1 - \theta_1 \theta_3^{y+1}) - \theta_2^{y+1} \frac{1 - (\theta_1 \theta_3^{y+1})^{y+1}}{1 - \theta_1 \theta_3^{y+1}} (1 - \theta_1 \theta_3^{y+1}) \\
&= \sum_{y=0}^{\infty} \theta_2^y (1 - \theta_2) - \sum_{y=0}^{\infty} \theta_2^y [(\theta_1 \theta_3^y)^{y+1} - \theta_2 (\theta_1 \theta_3^{y+1})^{y+1}] \\
&= 1 - \sum_{y=0}^{\infty} \theta_1^{y+1} \theta_2^y \theta_3^{y(y+1)} (1 - \theta_3^{y+1} \theta_2) \qquad (10)
\end{aligned}$$

In general, the infinite sum in (10) cannot be expressed in a closed form, but has to be evaluated numerically. However, if $\theta_3 = 1$ (independence case), its expression reduces to the following:

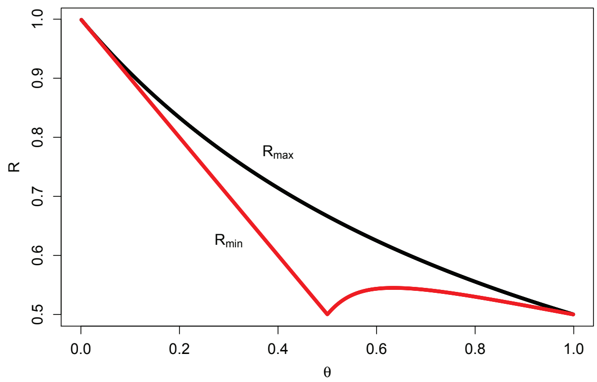
$$R = 1 - \sum_{y=0}^{\infty} \theta_1^{y+1} \theta_2^y (1 - \theta_2) = 1 - \theta_1 (1 - \theta_2) \sum_{y=0}^{\infty} (\theta_1 \theta_2)^y = 1 - \frac{\theta_1 (1 - \theta_2)}{1 - \theta_1 \theta_2} = \frac{1 - \theta_1}{1 - \theta_1 \theta_2}$$

which precisely coincides with that derived by Maiti (1995) for two independent geometric variables. For identically distributed margins, with common parameter $\theta = \theta_1 = \theta_2$, the expression of R in (10) becomes:

$$R = 1 - \sum_{y=0}^{\infty} \theta^{2y+1} \theta_3^{y(y+1)} (1 - \theta_3^{y+1}\theta);$$

the graph of lower and upper bounds ($\underline{R}(\theta) = \inf_{\theta_3} R(\theta, \theta_3)$ and $\overline{R}(\theta) = \sup_{\theta_3} R(\theta, \theta_3)$) of the reliability parameter as a function of θ is displayed in Figure 2. It can be empirically shown that $R(\theta, \theta_3)$, for a fixed value of θ, is an increasing function of θ_3, so that its upper bound is obtained for $\theta_3 = 1$ (independent margins), whereas the lower bound is obtained when θ_3 tends to its minimum feasible value. Note that whereas the upper bound $\overline{R}(\theta)$ is a monotone decreasing function of θ, the lower bound $\underline{R}(\theta)$ is not monotonic in $(0, 1)$.

Fig. 2 Lower and upper bounds for the reliability parameter R for the bivariate geometric model with identical geometric margins with parameter θ



6 ESTIMATION

In this section, we propose possible methods for point estimation of the parameters of the bivariate geometric model, based on a i.i.d. sample (x_i, y_i), $i = 1, 2 \ldots, n$. We start with the standard maximum likelihood method and then move to several original variants/mixtures of the method of moments and method of proportion, and to a least-squares method, which are strictly grounded on the mathematical properties of the bivariate distribution. All the three parameters are always assumed to be unknown.

6.1 Maximum Likelihood

For the model at study with p.m.f. expressed by (3), the log-likelihood function is given by

$$\ell(\theta_1,\theta_2,\theta_3;(x_1,y_1),\dots,(x_n,y_n)) = \ln\theta_1\sum_{i=1}^n x_i + \ln\theta_2\sum_{i=1}^n y_i + \ln\theta_3\sum_{i=1}^n x_i y_i \\ + \sum_{i=1}^n \ln(1-\theta_1\theta_3^{y_i}-\theta_2\theta_3^{x_i}+\theta_1\theta_2\theta_3\theta_3^{x_i+y_i}).$$

The three normal equations derived from equating the log-likelihood derivatives (with respect to θ_1, θ_2, and θ_3) to zero are given by:

$$\frac{\partial\ell(\theta_1,\theta_2,\theta_3)}{\partial\theta_1} = \frac{\sum_{i=1}^n x_i}{\theta_1} + \sum_{i=1}^n \frac{-\theta_3^{y_i}+\theta_2\theta_3\theta_3^{x_i+y_i}}{1-\theta_1\theta_3^{y_i}-\theta_2\theta_3^{x_i}+\theta_1\theta_2\theta_3\theta_3^{x_i+y_i}} = 0$$

$$\frac{\partial\ell(\theta_1,\theta_2,\theta_3)}{\partial\theta_2} = \frac{\sum_{i=1}^n y_i}{\theta_2} + \sum_{i=1}^n \frac{-\theta_3^{x_i}+\theta_1\theta_3\theta_3^{x_i+y_i}}{1-\theta_1\theta_3^{y_i}-\theta_2\theta_3^{x_i}+\theta_1\theta_2\theta_3\theta_3^{x_i+y_i}} = 0$$

$$\frac{\partial\ell(\theta_1,\theta_2,\theta_3)}{\partial\theta_3} = \frac{\sum_{i=1}^n x_i y_i}{\theta_3} + \sum_{i=1}^n \frac{-\theta_1 y_i\theta_3^{y_i-1}-\theta_2 x_i\theta_3^{x_i-1}+\theta_1\theta_2(x_i+y_i+1)\theta_3^{x_i+y_i}}{1-\theta_1\theta_3^{y_i}-\theta_2\theta_3^{x_i}+\theta_1\theta_2\theta_3\theta_3^{x_i+y_i}} = 0.$$

The system above cannot be solved analytically yet; the MLEs can be derived only numerically, by using some proper routine in any mathematical/statistical environment implementing (constrained) optimization (e.g., `optim` in R).

6.2 Method of Moments and Proportion

A standard method of moments should equate the two marginal moments and the mixed moment of the geometric distribution to the corresponding sample quantities and then solve this system of three equations for the three unknown parameters. Equating $\mathbb{E}(X)$ to $\bar{x}$ and $\mathbb{E}(Y)$ to $\bar{y}$ directly leads to the estimates

$$\hat{\theta}_1 = \bar{x}/(1+\bar{x}) \tag{11a}$$

$$\hat{\theta}_2 = \bar{y}/(1+\bar{y}) \tag{11b}$$

for the two marginal parameters. However, as we have seen in Section 4, the mixed moment $\mathbb{E}(XY)$ does not have a closed-form expression, so the standard version of the method of moments is not viable for deriving the remaining parameter θ_3.

A possible modification that tackles this technical issue considers the s.f. $S(x,y)$. We have that $S(1,1) = P(X \geq 1, Y \geq 1) = \theta_1\theta_2\theta_3$; then, we can equate it to the corresponding sample quantity, $\hat{s}_{11} = \sum_{i=1}^n \mathbb{1}_{\{x_i\geq 1, y_i\geq 1\}}/n$, thus deriving an estimate for θ_3 as $\hat{\theta}_3^{MMP} = \hat{s}_{11}/(\hat{\theta}_1\hat{\theta}_2)$. This method can be seen as an hybrid method blending the method of moments, used for estimating the two marginal parameters, and the method of proportion, which is particular suitable for discrete models (see Khan et al. 1989), and is used here for estimating the dependence parameter; this is the reason for the abbreviation MMP (Method of Moments and Proportion).

Another possible modification of the method of moments could exploit the characterization of the bivariate geometric model in terms of bivariate failure rates and mean residual lifetimes recalled in Section 2. Since we have:

$$\lambda_1(x,1) = 1-\theta_1\theta_3 \quad \forall x \in \mathbb{Z}_0^+$$
$$\lambda_2(1,y) = 1-\theta_2\theta_3 \quad \forall y \in \mathbb{Z}_0^+$$

and, by particularizing Eq.(5),

$$\lambda_1(0,1)\cdot\mu_1(0,0) = 1 = \lambda_2(1,0)\cdot\mu_2(0,0)$$

where $\mu_1(0,0) = \mathbb{E}(X|X>0, Y>0)$, $\mu_2(0,0) = \mathbb{E}(Y|X>0, Y>0)$; then, plugging in the moment estimates of θ_1 and θ_2, and computing the sample conditional moments $\bar{x}^+ := \sum_{i=1}^n x_i \cdot \mathbb{1}_{\{x_i>0,y_i>0\}} / \sum_{i=1}^n \mathbb{1}_{\{x_i>0,y_i>0\}}$ and $\bar{y}^+ := \sum_{i=1}^n y_i \cdot \mathbb{1}_{\{x_i>0,y_i>0\}} / \sum_{i=1}^n \mathbb{1}_{\{x_i>0,y_i>0\}}$, we get:

$$(1-\hat{\theta}_1\theta_3)\bar{x}^+ = 1 \tag{12a}$$
$$(1-\hat{\theta}_2\theta_3)\bar{y}^+ = 1 \tag{12b}$$

from which one can derive an estimate of θ_3 as

$$\hat{\theta}_3^{MM1} = \frac{\bar{x}^+ - 1}{\bar{x}^+\hat{\theta}_1} \tag{13}$$

or

$$\hat{\theta}_3^{MM2} = \frac{\bar{y}^+ - 1}{\bar{y}^+\hat{\theta}_2}, \tag{14}$$

$MM1$ and $MM2$ standing for "method of moments 1" (based on the bivariate failure rate and mean residual life of the first component of the bivariate random vector) and "method of moments 2" (based on the bivariate failure rate and mean residual life of the second component).

Alternatively, instead of choosing between the two estimators $MM1$ and $MM2$, one can seek for an estimator derived as a solution of the two equations (12a) and (12b) "put together". Such an estimator is expected to have a better performance in terms of bias and standard error than either estimator (13) or (14). Summing up the left members of Eq.(12a) and (12b) and equating it to the sum of the right members of the same equations, we can derive an estimator of θ_3 as

$$\hat{\theta}_3^{MM3} = \frac{\bar{x}^+ + \bar{y}^+ - 2}{\hat{\theta}_1\bar{x}^+ + \hat{\theta}_2\bar{y}^+}. \tag{15}$$

As a further possible modification, on can sum up the two equations (12a) and (12b) after rewriting them as $\bar{x}^+ = (1-\hat{\theta}_1\theta_3)^{-1}$ and $\bar{y}^+ = (1-\hat{\theta}_2\theta_3)^{-1}$:

$$\bar{x}^+ + \bar{y}^+ = \frac{1}{1-\hat{\theta}_1\theta_3} + \frac{1}{1-\hat{\theta}_2\theta_3},$$

from which one derives a quadratic equation in θ_3 with two real roots; the smallest is given by

$$\hat{\theta}_3^{MM4} = \frac{(\hat{\theta}_1 + \hat{\theta}_2) \cdot (\bar{s}^+ - 1) - \sqrt{\bar{s}^+ \cdot (\bar{s}^+ - 2) \cdot (\hat{\theta}_1 - \hat{\theta}_2)^2 + (\hat{\theta}_1 + \hat{\theta}_2)^2}}{2 \cdot \bar{s}^+ \cdot \hat{\theta}_1 \cdot \hat{\theta}_2} \tag{16}$$

with $\bar{s}^+ = \bar{x}^+ + \bar{y}^+$.

Finally, another possible modification simply arises from equating the left members of Equations (12a) and (12b) and then solving for θ_3:

$$\hat{\theta}_3^{MM5} = \frac{\bar{y}^+ - \bar{x}^+}{\hat{\theta}_2 \bar{y}^+ - \hat{\theta}_1 \bar{x}^+}.$$

This estimator looks less reliable in practice, since it can yield negative values. However, all the methods presented in this subsection for estimating the dependence parameter θ_3 suffer from the disadvantage of possibly producing inconsistent estimates, i.e., values outside the natural parameter space of θ_3.

6.3 Least-squares method

For the model at study, we consider the expression (4) of the joint s.f. $S(x, y)$ for a single bivariate observation (x_i, y_i) and take its natural logarithm:

$$\ln S(x_i, y_i) = x_i \ln \theta_1 + y_i \ln \theta_2 + x_i y_i \ln \theta_3.$$

The equation above can be seen as a linear regression model:

$$\mathbf{s} = \mathbf{X}\boldsymbol{\beta} \tag{17}$$

by setting $\mathbf{s} = (s_1 \dots s_n)'$, with $s_i = \ln S(x_i, y_i)$ as the dependent variable, $\mathbf{X} = \begin{pmatrix} x_1 & y_1 & x_1 y_1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & x_n y_n \end{pmatrix}$ as the matrix of covariates, $\boldsymbol{\beta} = (\ln \theta_1, \ln \theta_2, \ln \theta_3)'$ as the vector or regression coefficients. Obviously, $S(x_i, y_i)$, depending on the unknown parameters θ_1, θ_2, θ_3, has to be somehow estimated; a natural choice is the "standard" empirical joint s.f., already introduced in Section 6.2:

$$\hat{S}(x_i, y_i) = \sum_{j=1}^{n} \mathbb{1}_{\{x_j \geq x_i, y_j \geq y_i\}} / n;$$

note that since $0 < \hat{S}(x_i, y_i) \leq 1$, for any observed (x_i, y_i), $\hat{s}_i = \ln S(x_i, y_i)$ is always defined. An alternative sample estimator of S is given by

$$\hat{S}^*(x_i, y_i) = \sum_{j=1}^{n} \mathbb{1}_{\{x_j \geq x_i, y_j \geq y_i\}} / (n - 1);$$

for which it holds $0 < \hat{S}^*(x_i, y_i) < 1$. Then, in (17), we can substitute $\hat{\mathbf{s}} = (\hat{s}_1 \dots \hat{s}_n)'$ to $\mathbf{s}$ and the solution of the linear model is given by the standard expression: $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\hat{\mathbf{s}}$, by which the LS parameter estimates for the bivariate geometric model can be derived as $\hat{\boldsymbol{\theta}}^{LS} = \exp(\hat{\boldsymbol{\beta}})$.

Although not much can be said about the statistical properties of the LS estimators, nevertheless they possess the advantage, with respect to the ML estimators, of being computable analytically. Unfortunately, similarly to the method of moments, the values of the estimates are not guaranteed to lie in the corresponding parameter space, so some samples can yield inconsistent estimates.

7 MONTE CARLO STUDY

In this section, a simulation study is described that is carried out for the bivariate geometric model of Eq.(3) $N = 10,000$ times, for several parameter combinations $(\theta_1, \theta_2, \theta_3)$, with sample sizes $n = 50, 100, 200$. Even though the simulation study cannot be exhaustive, nevertheless we considered a selection of all the consistent combinations of values of the dependence parameter θ_3 and the two marginal parameters θ_1 and θ_2. Indeed, the aim of this study is not to assess and rank the "performances" of the parameter estimators for each possible parameter combination (which is obviously impracticable due also to the particular form of the parameter space), but to give a rough idea of their "placing" and how it depends on the parameter combination itself.

For each simulated setting and for all the estimators described in Section 6, except MM5, the study calculates the following measures:

- Average value (AV) of the simulated estimates:

$$AV(\hat{\theta}_j) = \bar{\hat{\theta}}_j := \frac{1}{N}\sum_{i=1}^{N}\hat{\theta}_{ij},$$

  where $\hat{\theta}_{ij}$ is the value on the i-th sample of one of the estimators $\hat{\theta}_j$ of θ_j, $i = 1, \dots, N$, $j = 1, 2, 3$.
- Standard Error (SE) of the simulated estimates:

$$SE(\hat{\theta}_j) := \frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_{ij} - \bar{\hat{\theta}}_j\right)^2.$$

The Monte Carlo study is carried out in the `R` programming environment using *ad hoc* developed code, which is available at the following link: https://bit.ly/2ETSP8B. An `R` package, `bivgeom`, including routines for computing the joint s.f, c.d.f., p.m.f., bivariate failure rates, and correlation, and for implementing random generation and sample estimation, is freely available on the CRAN repository (`https://CRAN.R-project.org/package=bivgeom`).

We mainly focus on the dependence parameter θ_3; the results about the two marginal parameters θ_1 and θ_2 are less relevant, since the maximum likelihood method and the different versions of the method of moments (all sharing the same expression for the two estimates, see equations (11a) and (11b)) lead sample by sample to very similar values for the estimates and overall for the two synthesis measures. Tables 2, 3, and 4 report the AV and SE for all the estimators for the examined parameter combinations, for sample sizes $n = 50$, $n = 100$, and $n = 200$, respectively.

Table 2 Average value (AV) and standard error (SE) for the estimators of θ_3, $n = 50$

θ_1	θ_2	θ_3	ML		MMP		MM1		MM2		MM3		MM4		LS	
			AV	SE	AV	SE	AV	SE	AV	SE	AV	SE	AV	SE	AV	SE
0.3	0.3	0.1	0.087	0.128	0.106	0.157	0.058	0.298	0.048	0.261	0.069	0.253	0.069	0.252	0.292	0.226
0.3	0.7	0.1	0.070	0.076	0.105	0.100	0.060	0.285	0.060	0.182	0.073	0.186	0.072	0.180	0.178	0.126
0.5	0.5	0.1	0.075	0.075	0.104	0.093	0.062	0.216	0.055	0.207	0.073	0.186	0.073	0.185	0.166	0.124
0.3	0.3	0.3	0.270	0.209	0.317	0.267	0.184	0.474	0.185	0.472	0.229	0.403	0.228	0.401	0.469	0.289
0.3	0.7	0.3	0.275	0.115	0.311	0.170	0.221	0.434	0.232	0.262	0.257	0.244	0.255	0.238	0.380	0.179
0.5	0.5	0.3	0.282	0.103	0.310	0.154	0.240	0.305	0.231	0.300	0.266	0.231	0.266	0.231	0.377	0.166
0.3	0.3	0.6	0.551	0.255	0.625	0.357	0.434	0.604	0.445	0.609	0.511	0.481	0.510	0.478	0.763	0.348
0.3	0.7	0.6	0.579	0.116	0.620	0.214	0.519	0.466	0.544	0.234	0.567	0.198	0.564	0.195	0.686	0.168
0.5	0.5	0.6	0.583	0.108	0.620	0.200	0.545	0.288	0.544	0.286	0.576	0.188	0.575	0.188	0.695	0.163
0.6	0.6	0.6	0.494	0.144	0.615	0.170	0.569	0.200	0.565	0.200	0.588	0.124	0.587	0.124	0.681	0.118
0.3	0.3	0.9	0.803	0.221	0.930	0.398	0.728	0.634	0.746	0.638	0.814	0.475	0.812	0.472	1.057	0.348
0.3	0.7	0.9	0.877	0.092	0.922	0.217	0.829	0.418	0.867	0.154	0.867	0.154	0.875	0.133	0.959	0.114
0.5	0.5	0.9	0.879	0.091	0.924	0.210	0.857	0.225	0.862	0.222	0.882	0.143	0.881	0.143	0.974	0.111
0.6	0.6	0.9	0.890	0.057	0.918	0.172	0.880	0.136	0.880	0.134	0.892	0.082	0.892	0.082	0.953	0.069
0.7	0.7	0.9	0.893	0.036	0.914	0.144	0.891	0.083	0.889	0.083	0.897	0.047	0.897	0.047	0.939	0.043

Table 3 Average value (AV) and standard error (SE) for the estimators of θ_3, $n = 100$

θ_1	θ_2	θ_3	ML		MMP		MM1		MM2		MM3		MM4		LS	
			AV	SE	AV	SE	AV	SE	AV	SE	AV	SE	AV	SE	AV	SE
0.3	0.3	0.1	0.094	0.097	0.103	0.108	0.058	0.284	0.060	0.286	0.074	0.253	0.074	0.252	0.193	0.178
0.3	0.7	0.1	0.090	0.072	0.101	0.069	0.068	0.273	0.071	0.174	0.081	0.172	0.080	0.167	0.137	0.094
0.5	0.5	0.1	0.095	0.076	0.101	0.063	0.071	0.199	0.068	0.197	0.082	0.166	0.082	0.165	0.134	0.098
0.3	0.3	0.3	0.285	0.149	0.306	0.180	0.217	0.443	0.219	0.444	0.219	0.444	0.254	0.360	0.391	0.220
0.3	0.7	0.3	0.290	0.077	0.306	0.117	0.258	0.347	0.263	0.205	0.279	0.180	0.277	0.178	0.357	0.126
0.5	0.5	0.3	0.292	0.071	0.304	0.109	0.267	0.235	0.266	0.231	0.266	0.231	0.266	0.231	0.355	0.119
0.3	0.3	0.6	0.579	0.179	0.613	0.241	0.510	0.486	0.509	0.488	0.555	0.359	0.555	0.359	0.715	0.254
0.3	0.7	0.6	0.590	0.080	0.610	0.150	0.563	0.329	0.573	0.158	0.585	0.132	0.583	0.131	0.666	0.117
0.5	0.5	0.6	0.592	0.076	0.610	0.144	0.575	0.199	0.572	0.199	0.589	0.128	0.589	0.128	0.671	0.116
0.6	0.6	0.6	0.552	0.107	0.607	0.120	0.585	0.139	0.585	0.138	0.595	0.085	0.595	0.085	0.659	0.085
0.3	0.3	0.9	0.850	0.153	0.917	0.271	0.816	0.462	0.818	0.459	0.859	0.328	0.859	0.328	1.025	0.236
0.3	0.7	0.9	0.890	0.063	0.911	0.152	0.873	0.286	0.885	0.098	0.891	0.088	0.889	0.087	0.951	0.080
0.5	0.5	0.9	0.890	0.065	0.911	0.149	0.881	0.151	0.881	0.151	0.892	0.097	0.892	0.097	0.959	0.081
0.6	0.6	0.9	0.895	0.039	0.909	0.121	0.890	0.092	0.891	0.091	0.897	0.057	0.897	0.057	0.942	0.052
0.7	0.7	0.9	0.897	0.022	0.907	0.101	0.895	0.057	0.895	0.057	0.898	0.033	0.898	0.033	0.930	0.033

Table 4 Average value (AV) and standard error (SE) for the estimators of θ_3, $n = 200$

θ_1	θ_2	θ_3	ML		MMP		MM1		MM2		MM3		MM4		LS	
			AV	SE	AV	SE	AV	SE	AV	SE	AV	SE	AV	SE	AV	SE
0.3	0.3	0.1	0.098	0.071	0.101	0.075	0.064	0.266	0.065	0.269	0.078	0.232	0.078	0.232	0.143	0.131
0.3	0.7	0.1	0.102	0.077	0.101	0.049	0.076	0.229	0.076	0.144	0.083	0.135	0.083	0.133	0.121	0.065
0.5	0.5	0.1	0.103	0.069	0.101	0.045	0.081	0.167	0.080	0.168	0.089	0.130	0.089	0.130	0.122	0.071
0.3	0.3	0.3	0.293	0.106	0.304	0.126	0.248	0.361	0.246	0.355	0.271	0.271	0.271	0.271	0.360	0.157
0.3	0.7	0.3	0.296	0.053	0.303	0.080	0.280	0.250	0.281	0.147	0.289	0.123	0.289	0.123	0.340	0.086
0.5	0.5	0.3	0.297	0.051	0.304	0.077	0.282	0.167	0.282	0.170	0.291	0.116	0.291	0.116	0.340	0.082
0.3	0.3	0.6	0.590	0.125	0.608	0.166	0.557	0.351	0.553	0.350	0.579	0.245	0.579	0.245	0.690	0.177
0.3	0.7	0.6	0.595	0.055	0.605	0.103	0.580	0.232	0.587	0.109	0.592	0.091	0.592	0.091	0.648	0.082
0.5	0.5	0.6	0.596	0.053	0.605	0.100	0.588	0.138	0.588	0.139	0.595	0.090	0.595	0.090	0.652	0.082
0.6	0.6	0.6	0.588	0.057	0.603	0.084	0.593	0.096	0.592	0.097	0.592	0.097	0.598	0.060	0.641	0.060
0.3	0.3	0.9	0.875	0.109	0.908	0.183	0.858	0.318	0.860	0.316	0.880	0.219	0.880	0.219	0.999	0.162
0.3	0.7	0.9	0.895	0.043	0.906	0.106	0.885	0.193	0.893	0.067	0.895	0.061	0.894	0.060	0.941	0.057
0.5	0.5	0.9	0.895	0.045	0.905	0.104	0.890	0.104	0.892	0.104	0.896	0.068	0.896	0.068	0.947	0.059
0.6	0.6	0.9	0.897	0.027	0.904	0.086	0.895	0.063	0.895	0.064	0.898	0.040	0.898	0.040	0.932	0.038
0.7	0.7	0.9	0.899	0.016	0.903	0.072	0.898	0.039	0.897	0.039	0.899	0.023	0.899	0.023	0.922	0.024

From these results, we try to outline some relevant points. The MLE of θ_3, as one could expect, represent overall, for the vast majority of scenarios, the best choice (with the smallest absolute bias and standard error). The superiority of MLE is guaranteed also for the smallest value of sample size ($n = 50$). The MLE of θ_3 is negatively biased (i.e., the value of its AV is smaller than the true value of θ_3) over most of the examined scenarios. Exceptions to this behavior are represented by two scenarios with $\theta_3 = 0.1$ and $n = 200$.

The second best performer is the MMP estimator: its absolute bias is almost always even smaller than that of MLE, although its standard error is at the same time (far) larger. The best scenarios for MMP are those characterized by $\theta_3 = 0.1$. Like MLE, for a fixed value of θ_3, the variability of MMP estimator is strongly affected by the combination (θ_1, θ_2); larger values of the two marginal parameters ensure smaller dispersion.

As to the four versions of the method of moments, their performances are overall worse than those of ML and MMP estimators. However, as foreseen in Section 6, MM3 and MM4 present better features than MM1 and MM2. Their biases are always negative (a characteristic they share with MLE) and their magnitude, as well as the magnitude of the SE, tends to decrease with increasing values of θ_1 and θ_2, for a fixed value of θ_3.

Despite its analytical expression, the LS estimator does not represent a valid alternative to MLE. It is always largely and positively biased and apart from few exceptions is systematically more variable than MLE. This behavior is particularly apparent for small sample size.

A graphical comparison of the empirical distributions of the set of estimators is provided in Figure 3, which refers to the scenario with $\theta_1 = \theta_2 = 0.5$, $\theta_3 = 0.3$, and $n = 100$. The figure, reporting the boxplots of the Monte Carlo distributions of the seven estimators investigated, clearly underlines how in this case for every sample the MLE, MMP, and LS methods provide a consistent estimator for θ_3 (i.e., a value falling in its parameter space, which is the unit interval), whereas the four variants of the moment method often yield inconsistent estimates, i.e. values larger than 1. The graph also highlights the smaller variability of the MLE compared to the other methods.

All these elements strongly support the use of MLE for the estimation of the dependence parameter θ_3; the other methods can be used, if possible, to provide a “starting value” for the optimization routine of the ML method.

8 APPLICATION TO REAL DATA

In this section, a numerical example is provided to illustrate the application of the bivariate geometric distribution. The data, considered in Mitchell and Paulson (1981), consist of the number of aborts by 109 aircrafts in two consecutive semesters (number of aborts in the first period $= x$, in the second period $= y$), see Table 5. Summary statistics for the dataset are provided below:

$$\bar{x} = 0.624, \quad \bar{y} = 0.725, \quad \mathrm{var}(x) = 1.024, \quad \mathrm{var}(y) = 1.062, \quad \hat{\rho}_{xy} = -0.161.$$

Fig. 3 Boxplot of the empirical distribution of the estimators of θ_3, with $\theta_3 = 0.3$, $\theta_1 = \theta_2 = 0.5$, $n = 100$.

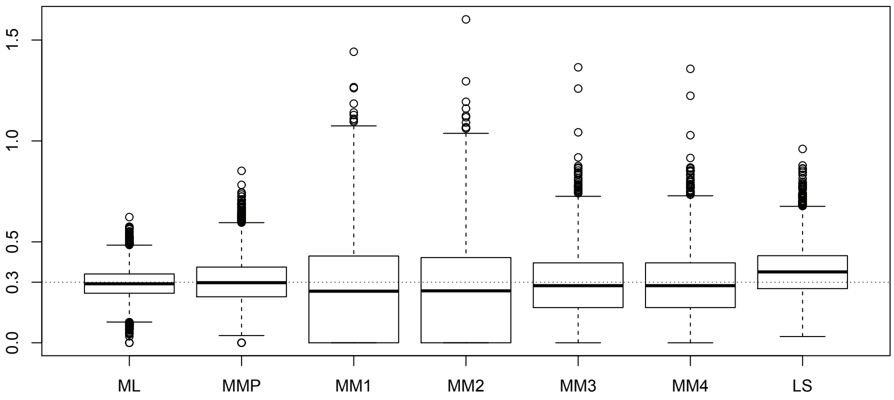


Table 5 Bivariate distribution of the data taken from Mitchell and Paulson (1981): number of flight aborts by 109 aircrafts in two consecutive periods

x, y	0	1	2	3	4	
0	34	20	4	6	4	68
1	17	7	0	0	0	24
2	6	4	1	0	0	11
3	0	4	0	0	0	4
4	0	0	0	0	0	0
5	2	0	0	0	0	2
	59	35	5	6	4	109

Looking at the empirical marginal distributions of x and y, one can assume they come from two geometric rvs. Through a proper univariate goodness-of-fit test one can check this hypothesis. By resorting to Kolmogorov-Smirnov test, see Bracquemond et al. (2002), we accept the null hypotheses that x and y come from two geometric distributions (bootstrapped p-values 0.7989 and 0.2984, respectively). Then, we can further assume and check whether the bivariate sample (x_i, y_i), $i = 1, \ldots, n$, comes from the bivariate geometric model proposed by Roy (1993). Fitting it to these data, being $\bar{x}^+ = 1.8125$, $\bar{y}^+ = 1.0625$, and $\bar{s}^+ = 2.875$, and estimating its parameters through each of the methods presented in Section 6 leads to the results reported in Table 6.

The estimates obtained through the LS method are quite different from those obtained through the ML method and the MMP; the dependence parameter θ_3 is closer to 1, and the marginal parameters θ_1 and θ_2 are a bit larger. Focusing on θ_3, the values provided by MM1 and MM2 are very different from each other and from the estimates derived through ML, MMP, and LS. MM3 and MM4 provide values much closer to the MLE. Note that the MLE of ρ_{xy}, obtained by plugging in Equation 7 the MLEs of θ_1, θ_2, and θ_3, takes the value -0.160, which is almost the same as $\hat{\rho}_{xy}$.

Table 6 Parameter estimates for the bivariate geometric model applied to Mitchell and Paulson (1981) data.

method	θ_1	θ_2	θ_3
ML	0.3857	0.4200	0.8033
MMP	0.3842	0.4202	0.9092
MM1	0.3842	0.4202	1.1668
MM2	0.3842	0.4202	0.1400
MM3	0.3842	0.4202	0.7657
MM4	0.3842	0.4202	0.7561
LS	0.4296	0.4341	0.9231

Table 7 Theoretical distribution of the data taken from Mitchell and Paulson (1981) obtained by fitting the bivariate geometric model. Cell borders highlight a possible aggregation of cells into 9 groupings with frequencies' sum greater than 5, where the chi-squared goodness-of-fit statistic is computed.

x, y	0	1	2	3	4
0	35.37	17.15	7.98	3.61	2.85
1	16.04	6.19	2.30	0.83	0.46
2	6.93	2.14	0.64	0.18	0.07
3	2.90	0.72	0.17	0.04	0.01
4	1.19	0.24	0.05	0.01	0.00
5	0.80	0.11	0.02	0.00	0.00

The value of the log-likelihood function computed at the MLEs is $\ell_{\max} = -244.0191$ and the corresponding value of the Akaike Information Criterion ($\text{AIC} = 2k - 2\ell_{\max}$, with $k = 3$ being the number of parameters) is 494.0382. The bivariate geometric model introduced in Hawkes (1972), see also Eq.(2), provides the MLEs $\hat{p}_{11} = 0.1355411$, $\hat{p}_{10} = 0.2475996$, and $\hat{p}_{01} = 0.2856990$, a value of the log-likelihood function equal to -245.0645, and an AIC equal to 496.129, which denotes a worse fit to the data. The bivariate negative binomial distribution proposed in Mitchell and Paulson (1981) shows an AIC equal to 498.54, thus indicating a worse fit than Roy's model. The bivariate geometric model introduced in Phatak and Sreehari (1981) and fully described in Krishna and Pundir (2009) provides even a worse fit: the maximum value of the log-likelihood function is equal to -259.5122 and the corresponding AIC is 523.0244. This could have been expected, since the latter model allows only for positive correlations, whereas the data show negative dependence.

In order to obtain an absolute measure of fit of the proposed bivariate model, we can resort to the standard chi-squared goodness-of-fit test. We compute the theoretical absolute joint frequencies, by using the p.m.f. in (3) with the MLEs of the parameters θ_1, θ_2 and θ_3; they are displayed in Table 7. Then we aggregate cells in order to obtain for each grouping an aggregate frequency larger than 5; we calculate the chi-squared statistic as $\chi^2 = \sum_{g=1}^{G} (\hat{n}_g - n_g)^2/\hat{n}_g$, where n_g is the observed count for grouping g, $\hat{n}_g$ is its theoretical analogue, G is the number of groupings (in this case $G = 9$: 6 single cells plus 3 actual groupings). Under the null hypothesis that the bivariate sample comes from the proposed distribution, χ^2 is approximately distributed as a chi-squared random variable with $9 - 3 - 1 = 5$ degrees of freedom. The empirical value of χ^2 is 5.151; its p-value is 0.398 and, being far larger than zero, it denotes a satisfactory fit of the model to the data.

9 CONCLUSIONS

Random simulation, attainable correlations, point estimation, and reliability concepts have been described for a bivariate geometric model characterized by locally constant failure rates. The model can be seen as the very bivariate counterpart of the univariate geometric distribution, characterized by a constant failure rate, and thus could play an important role among bivariate discrete distributions.

The marginal and conditional distributions and probability mass functions have accessible analytical expressions, which allows straightforward pseudo-random simulation. A Monte Carlo simulation study has empirically proved that despite the presence of non-trivial parameter constraints and the lack of a closed-form solution, the performance of the maximum likelihood method for the dependence parameter is overall superior than those of alternative moment or least-squares methods, which are developed by exploiting the characterizing features of the model. An application to a well-known dataset, which has been previously analysed in the literature by several other authors, illustrates the feasibility of the proposed estimation techniques and shows that the model can fit real dependent count data even better than extant alternatives, making a significant contribution to the discussion on count data modelling.

However, some model limitations have also to be pointed out. First, the model has been proved to yield negative (to the limit, null) dependence only, thus its use should be bounded to bivariate count data exhibiting negative sample correlation. Second, the dependence parameter vary in a range that depends on the values of the two marginal parameters; so, as mentioned before, maximum-likelihood estimation is subject to non-trivial constraints and this makes the study of the properties and asymptotic distribution of the corresponding estimators quite challenging; it can be the object of further work.

Acknowledgements I would like to thank the Editor-in-Chief, the Guest Editor, and the anonymous referees for their valuable comments on an earlier draft of this article. I acknowledge the financial support to the present research by the University of Milan (Piano di Sostegno alla Ricerca 2015/2017-Linea 2A).

References

Barnett, V. (1980). Some bivariate uniform distributions. *Communications in Statistics - Theory and Methods* 9(4), 453-461.

Basu, A.P., & Dhar, S.K. (1995) Bivariate geometric distribution, *Journal of Applied Statistical Science* 2, 33-44.

Bracquemond, C, Cretois, E., & Gaudoin, O. (2002). A comparative study of goodness-of-fit tests for the geometric distribution and application to discrete time reliability. Laboratoire Jean Kuntzmann, Applied Mathematics and Computer Science, Technical Report.

Dhar, S.K. (1998). Data analysis with discrete analog of Freund's model, *Journal of Applied Statistical Science* 7, 169-183.

Fiondella, L., Zeephongsekul, P. (2016). Trivariate Bernoulli distribution with application to software fault tolerance, *Annals of Operations Research* 244(1), 241-255.

Freund, J.E. (1961) A bivariate extension of the exponential distribution, *Journal of the American Statistical Association* 56(296), 971-977.

Galambos, J., Kotz, S. (1978) *Characterizations of probability distributions*, Springer Verlag, Berlin/New York.

Gumbel, E.J. (1960). Bivariate exponential distributions. *Journal of the American Statistical Association* 55(292), 698-707.

Hawkes, A.G. (1972). A bivariate exponential distribution with applications to reliability. *Journal of the Royal Statistical Society Series B* 34(1), 129-131.

Huber, M., Maric, N. (2014). Minimum correlation for any bivariate Geometric distribution. *Alea* 11(1): 459-470.

Jovanović, M. (2017) Estimation of $P\{X < Y\}$ for Geometric-exponential Model Based on Complete and Censored Samples, *Communications in Statistics - Simulation and Computation* 46(4), 3050-3066.

Khan, M.S.A., Khalique, A., & Abouammoh, A.M. (1989). On Estimating Parameters in a Discrete Weibull Distribution. *IEEE Transactions on Reliability* 38(3), 348-350

Krishna, H., & Pundir, P.S. (2009) A bivariate geometric distribution with applications to reliability, *Communications in Statistics - Theory and Methods* 38(7), 1079-1093.

Maiti, S.S. (1995). Estimation of $P(X \leq Y)$ in the geometric case, *Journal of Indian Statistical Association* 33, 87-91.

Mari, D.D., & Kotz, S. (2001) *Correlation and dependence*. World Scientific, Singapore.

Marshall, A.W., & Olkin, I. (1967) A multivariate exponential distribution, *Journal of the American Statistical Association* 62(317), 30-44.

Mitchell, C.R., & Paulson, A.S. (1981) A new bivariate negative binomial distribution, *Naval Research Logistic Quarterly* 28(3), 359-374.

Nelsen, R.B. (1999) *An Introduction to Copulas*. Springer, New York.

Paulson, A.S., & Uppuluri, V.R.R. (1972). A characterization of the geometric distribution and a bivariate geometric distribution, *Sankhyā Series A* 34(3), 297-300.

Phatak, A.G., & Sreehari, M. (1981). Some characterizations of a bivariate geometric distribution, *Journal of Indian Statistical Association* 19, 141-146.

R Core Team, R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria. URL `https://www.R-project.org/` (2018).

Roy, D. (1993) Reliability measures in the discrete bivariate set-up and related characterization results for a bivariate geometric distribution, *Journal of Multivariate Analysis* 46(2), 362-373.

Sklar, A. (1959) Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de l'Université de Paris*, 8, 229-231.

Sun. K., & Basu, A.P. (1995). A characterization of a bivariate geometric distribution. *Statistics and Probability Letters* 23(4), 307-311.

Xekalaki, E. (1983) Hazard functions and life distributions in discrete time. *Communications in Statistics - Theory and Methods* 12(21), 2503-2509