A. Barbiero*

A bivariate count model with discrete Weibull margins

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

Abstract

Multivariate discrete data arise in many fields (statistical quality control, epidemiology, failure and reliability analysis, etc) and modeling such data is a relevant task. Here we consider the construction of a bivariate model with discrete Weibull margins, based on Farlie-Gumbel-Morgenstern copula, analyze its properties especially in terms of attainable correlation, and propose several methods for the point estimation of its parameters. Two of them are the standard one-step and two-step maximum likelihood procedures; the other two are based on an approximate method of moments and on the method of proportion, which represent intuitive alternatives for estimating the dependence parameter. A Monte Carlo simulation study is presented, comprising more than one hundred artificial settings, which empirically assesses the performance of the different estimation techniques in terms of statistical properties and computational cost. For illustrative purposes, the model and related inferential procedures are fitted and applied to two datasets taken from the literature, concerning failure data, presenting either positive or negative correlation between the two observed variables. The applications show that the proposed bivariate discrete Weibull distribution can model correlated counts even better than existing and well-established joint distributions.

Keywords: Farlie-Gumbel-Morgenstern copula, method of moments, Monte Carlo simulation, Pearson's correlation, two-step maximum likelihood

1 Introduction

In many scientific fields, researchers are concerned with multivariate random variables (r.v.s). Although quantities measured on a continuous scale are more frequent, nevertheless multivariate count data often arise in several contexts (statistical quality control, epidemiology, failure and reliability analysis, etc). Such data are frequently modelled through the multivariate normal distribution or some multivariate exponential distribution, which however, being continuous r.v.s, fit the data hardly adequately; or through a multivariate Poisson model, which would require the data to have marginal means almost equal to the marginal variances. There are several forms of alternative multivariate discrete distributions defined and studied in the literature: for an extensive account of these distributions, we refer the reader to the early books by [13, 19], and references therein. Various methods have been proposed for constructing alternative multivariate r.v.s, see the excellent review in [31]. Methods and issues related to the construction of bivariate discrete distributions, which are disseminated in the literature, have been reviewed in [21]; for more recent proposals, see e.g. [16, 22]. Whereas the construction of multivariate distributions based on the definition of their joint probability mass or density function poses some difficulties and often results in practical limitations, for example, in the range of possible pairwise correlations; the specification via the marginal distributions and a copula function that provides the dependence structure, is much more straightforward. The main advantage provided by using copula-based models is indeed that the selection of an appropriate model for the dependence among the margins, represented by the copula, can proceed independently from the choice of the marginal distributions, resulting in a great flexibility. Copulas (see [25]) nowadays are widely used for modeling correlated continuous data, whereas the literature on copulas used for discrete data is more limited. Recently, however, some papers (see, e.g., [26, 11]) discussed the modeling of multivariate count data based on copulas and highlight possible pitfalls and challenges [27].

In this work we consider the discrete Weibull distribution as an alternative to the Poisson r.v. for modelling count data. It is a two-parameter distribution derived as a discrete counterpart of the popular continuous Weibull model and can reveal itself to be more flexible than Poisson. It was employed by Englehardt and Li [7] in a correlated random multiplicative growth model for microbial counts in water; more recently, Englehardt [8] showed that the discrete Weibull distribution can model products of autocorrelated causes, generated via copula. In [3], a procedure for generating correlated discrete Weibull variables linked via a Gaussian copula was described, with an assigned correlation matrix. In [5], a bivariate model with discrete Weibull margins and Farlie-Gumbel-Morgenstern (FGM) copula was suggested. The FGM copula is a dependence structure particularly interesting since i) it allows for either positive or negative dependence ii) it has an analytically closed-form and simple expression iii) it is easy to simulate iv) its unique parameter has a straightforward relationship with the linear correlation coefficient. In this work, we reclaim this model, through an in-depth examination of its properties, with special attention at conditional distributions, correlation range, and estimation issues.

The rest of the paper is structured as follows: in the next section, the discrete Weibull distribution is introduced and its features briefly described. Section 3 presents and discusses the FGM copula and introduces the new bivariate discrete Weibull model, with a special focus on conditional distributions and Pearson's correlation. In Section 4, procedures for parameter estimation of the model are suggested. Section 5 outlines a Monte Carlo simulation study that assesses the statistical properties of the different estimation techniques of Section 4. Section 6 considers two real datasets that are fitted by the bivariate discrete Weibull model. The final section concludes the paper with some remarks and future research perspectives.

2 The type I discrete Weibull distribution

The discrete Weibull distribution was introduced by Nakagawa and Osaki [24] as a discrete counterpart of the continuous Weibull distribution and is usually referred to as "type I discrete Weibull distribution". Two other models built as discrete analogues were proposed later by Stein and Dattero [33] (type II discrete Weibull) and Padgett and Spurrier [28] (type III discrete Weibull). For the model proposed by Nakagawa and Osaki [24], the cumulative distribution function (c.d.f.) is

$$F(x; q, \beta) = 1 - q^{(x+1)^\beta}; \quad x \in \mathbb{N}_0, \tag{1}$$

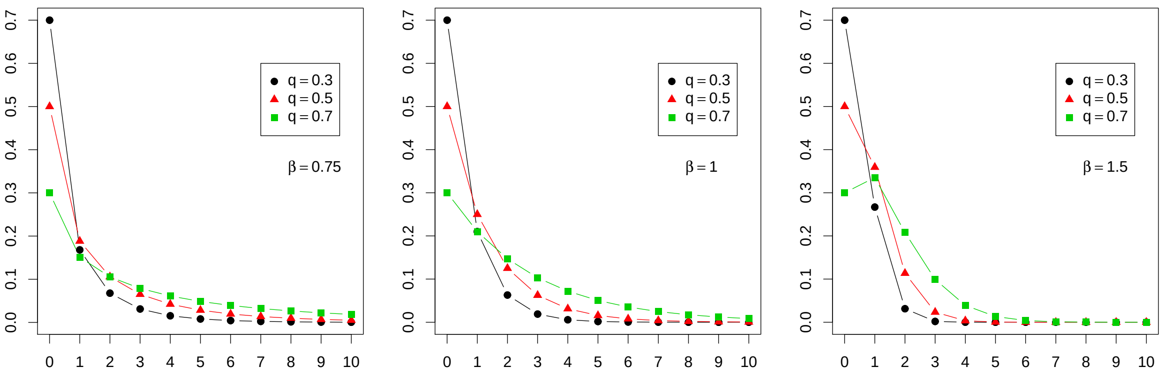
with $0 < q < 1$ and $\beta > 0$, and consequently the expression of its probability mass function is:

$$p(x; q, \beta) = q^{x^\beta} - q^{(x+1)^\beta} \quad x \in \mathbb{N}_0. \tag{2}$$

This distribution, differently from the two alternative ones, retains the expression of the c.d.f. of the continuous Weibull model. Hereafter, we will denote it simply as "discrete Weibull" distribution or model.

Note that the first parameter q in (2) has an easy and nice interpretation: since $P(X = 0) = 1 - q$, it represents the probability of a positive value. As to the second parameter β, it does not possess an equally nice interpretation. However, it has been shown that the hazard function, or failure rate – for discrete r.v.s defined as $r(x) = p(x)/(1-F(x))$ – is an increasing function if $\beta > 1$, a decreasing function if $\beta < 1$, a constant function if $\beta = 1$; note that in the latter case the discrete Weibull distribution reduces to the geometric distribution with success probability $1 - q$. Figure 1 displays the p.m.f. (limited to the values $x = 0, 1, \dots, 10$) for several combinations of q and β. Here the role of β, for a fixed value of q, is a bit clearer: larger values of β lead to less dispersed distributions, with most of the probability mass concentrated on the first integer values; conversely, smaller values of β lead to more dispersed distributions, with a thicker right tail.

Figure 1: Probability mass function of the discrete Weibull distribution for several values' combinations of its parameters q and β



The discrete Weibull distribution can be used for modelling the number of shocks, cycles, runs a component or structure can overcome before failing, or the number of periods (e.g., days, weeks, etc.) it successfully works until failure (discrete lifetimes). More generally, it can virtually model any type of count data. In fact, contrary to the Poisson r.v., which cannot adequately model counts whose variance exceeds the mean, the discrete Weibull r.v. can model both underdispersed and overdispersed data (see Englehardt and Li [7], Barbiero [3]). This distribution can also handle count data presenting an excess of zeros, which arise in many physical situations (see Englehardt and Li [7], Englehardt [8]). With regard to the issues related to point and interval estimation of its parameters, one can refer to Khan and Khalique and Abouammoh [12], Kulasekera [15], Barbiero [4]. The discrete Weibull model is implemented in the `R` environment [30] through the package `DiscreteWeibull` [2], which comprises several functions computing the p.m.f., the c.d.f., the quantile function, the first and second moments, and implementing the pseudo-random generation and sample estimation.

3 Modelling correlated discrete Weibull r.v.s via the FGM copula

In this section, we formally introduce the bivariate discrete Weibull distribution based on the FGM copula, by defining its c.d.f. and then deriving the expression of its p.m.f., conditional distributions, and Pearson's correlation.

3.1 The model

The FGM copula [10] possesses an easy analytical expression. For the dimension 2, it is given by

$$C(u_1, u_2; \theta) = u_1 u_2 [1 + \theta(1 - u_1)(1 - u_2)], \quad (u_1, u_2) \in [0, 1]^2 \tag{3}$$

with $-1 \leq \theta \leq +1$, and it can be seen as a "perturbation" of the independence copula $\Pi(u_1, u_2) = u_1 u_2$ via its parameter θ. The bivariate FGM copula allows for a moderate level of linear correlation; it can be shown that Pearson's rho between the two uniform components is $\rho(U_1, U_2) = \theta/3$. This is quite interesting since not all the bivariate copulas can deal with both negative and positive dependence: for example, the well-known Gumbel and Clayton copulas only allow positive dependence. Given the constraint on θ, the linear correlation between the two FGM copula margins is bounded in $[-1/3, 1/3]$.

Let consider two r.v.s X_1 and X_2 with c.d.f.s F_1 and F_2, respectively. A bivariate random vector (X_1, X_2), having marginal distributions F_1 and F_2 and FGM copula C with parameter θ as in Eq. (3), can be constructed defining the bivariate joint c.d.f.

$$F(x_1, x_2; \theta) = C(F_1(x_1), F_2(x_2)) = F_1(x_1)F_2(x_2)\left[1 + \theta(1 - F_1(x_1))(1 - F_2(x_2))\right], \tag{4}$$

with (x_1, x_2) belonging to the support of (X_1, X_2).

If X_1 and X_2 are discrete r.v.s with p.m.f. $p_1(x_1)$ and $p_2(x_2)$, respectively, recalling that for a bivariate r.v. defined on non-negative integers we have

$$p(x_1, x_2) = F(x_1, x_2) - F(x_1 - 1, x_2) - F(x_1, x_2 - 1) + F(x_1 - 1, x_2 - 1),$$

the bivariate p.m.f. of (X_1, X_2) can be recovered from (4) as

$$p(x_1, x_2; \theta) = p_1(x_1)p_2(x_2)\left\{1 + \theta\left[1 + p_1(x_1) - 2F_1(x_1)\right]\left[1 + p_2(x_2) - 2F_2(x_2)\right]\right\}. \tag{5}$$

If F_1 and F_2 are the c.d.f.s of a discrete Weibull r.v., $F_i(x; q_i, \beta_i) = 1 - q_i^{(x+1)^{\beta_i}}$, $i = 1, 2$, then a FGM copula-based bivariate discrete Weibull r.v. is generated. The bivariate joint c.d.f. (4) takes the form

$$F(x_1, x_2; q_1, \beta_1, q_2, \beta_2, \theta) = \left[1 - q_1^{(x_1+1)^{\beta_1}}\right]\left[1 - q_2^{(x_2+1)^{\beta_2}}\right]\left[1 + \theta q_1^{(x_1+1)^{\beta_1}} q_2^{(x_2+1)^{\beta_2}}\right],$$

with $(x_1, x_2) \in \mathbb{N}_0^2$, and the bivariate p.m.f. (5) is then given by

$$p(x_1, x_2; q_1, \beta_1, q_2, \beta_2, \theta) = \left(q_1^{x_1^{\beta_1}} - q_1^{(x_1+1)^{\beta_1}}\right)\left(q_2^{x_2^{\beta_2}} - q_2^{(x_2+1)^{\beta_2}}\right) \cdot \left[1 + \theta\left(q_1^{x_1^{\beta_1}} + q_1^{(x_1+1)^{\beta_1}} - 1\right)\left(q_2^{x_2^{\beta_2}} + q_2^{(x_2+1)^{\beta_2}} - 1\right)\right]. \tag{6}$$

We must remark that due to the discrete nature of the two margins, the parameter θ in (6) is now allowed to span a wider range than that corresponding to the FGM copula in (3). Following [6, 29], we have in fact that

$$-1 \leq \theta \leq \min\{1/q_1, 1/q_2\}. \tag{7}$$

For more details on the FGM copula, and its extension, combining discrete margins, see [29].

3.2 Conditional distributions and simulation

For two discrete r.v.s X_1 and X_2 linked by the FGM copula, recalling Eqs. (2) and (5), the conditional distribution of X_2 given $X_1 = x_1$ is

$$p_{2|1}(x_2|x_1) = \frac{p(x_1, x_2)}{p(x_2)} = p_2(x_2)\left\{1 + \theta\left[1 + p_1(x_1) - 2F_1(x_1)\right]\left[1 + p_2(x_2) - 2F_2(x_2)\right]\right\}. \tag{8}$$

For the bivariate discrete Weibull distribution, (8) reduces to:

$$p_{2|1}(x_2|x_1) = \left(q_2^{x_2^{\beta_2}} - q_2^{(x_2+1)^{\beta_2}}\right)\left[1 + \theta\left(q_1^{x_1^{\beta_1}} + q_1^{(x_1+1)^{\beta_1}} - 1\right)\left(q_2^{x_2^{\beta_2}} + q_2^{(x_2+1)^{\beta_2}} - 1\right)\right]. \tag{9}$$

As an example, let us consider two discrete Weibull r.v.s with marginal parameters $q_1 = q_2 = 0.9$ and $\beta_1 = \beta_2 = 1.2$, linked by a FGM copula with parameter $\theta = \pm 1/2$; the corresponding conditional distributions of X_2 given $X_1 = 1$ are plotted in Figure 2, along with the marginal distribution $p_2(x_2)$. We note that for $\theta = 1/2$, the first five integer values of the conditional distribution have each a larger probability than in the marginal distribution ($\theta = 0$), whereas the values greater than 4 have smaller probabilities. The opposite holds for $\theta = -1/2$.

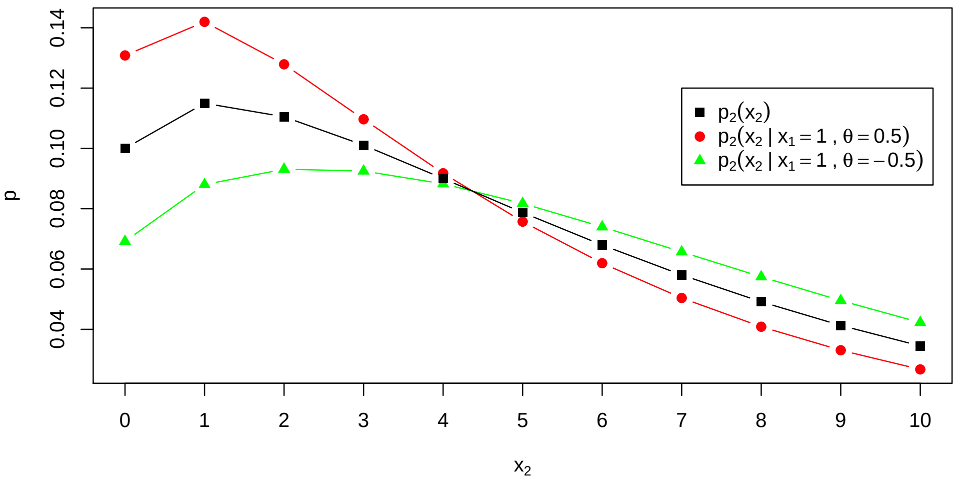
By using (9), the conditional expected value of X_2 given $X_1 = x_1$ can be derived as follows:

$$\begin{aligned}
\mathbb{E}(X_2|X_1 = x_1) = \mathbb{E}(X_2) + \theta \cdot \left[q_1^{(x_1+1)^{\beta_1}} + q_1^{x_1^{\beta_1}} - 1\right]&\left\{\sum_{x_2=0}^{\infty} x_2 \left[q_2^{x_2^{\beta_2}} - q_2^{(x_2+1)^{\beta_2}}\right] + \right.\\
&\left. - \sum_{x_2=0}^{\infty} x_2 \left[q_2^{x_2^{\beta_2}} - q_2^{(x_2+1)^{\beta_2}}\right]\left[2 - q_2^{x_2^{\beta_2}} - q_2^{(x_2+1)^{\beta_2}}\right]\right\}\\
&= \mathbb{E}(X_2) + \theta \cdot \left[q_1^{(x_1+1)^{\beta_1}} + q_1^{x_1^{\beta_1}} - 1\right]\left\{\mathbb{E}(X_2) - 2\mathbb{E}(X_2) + \sum_{x_2=0}^{\infty} x_2 \left[q_2^{2x_2^{\beta_2}} - q_2^{2(x_2+1)^{\beta_2}}\right]\right\}\\
&= \mathbb{E}(X_2) + \theta \cdot \left[q_1^{(x_1+1)^{\beta_1}} + q_1^{x_1^{\beta_1}} - 1\right]\left[\mathbb{E}(Y_2) - \mathbb{E}(X_2)\right],
\end{aligned}$$

where Y_2 is a discrete Weibull r.v. with parameters q_2^2 and β_2, which implies that $\mathbb{E}(Y_2) - \mathbb{E}(X_2) < 0$ for any (q_2, β_2). Values of $\mathbb{E}(X_2|X_1 = x_1)$ can be computed numerically. For the example above, when $\theta = 0.5$, we have that $\mathbb{E}(X_2|X_1 = 1) = 4.721$, whereas $\mathbb{E}(X_2) = 5.641$.

Given the symmetric nature of the FGM copula, analogous results hold for the distribution of X_1 given $X_2 = x_2$.

Figure 2: Conditional distributions $p_{2|1}(x_2|x_1 = 1)$ for bivariate discrete Weibull r.v. with FGM copula with $\theta = \pm 0.5$ compared to the marginal distribution $p_2(x_2)$



The simulation of this bivariate discrete distribution is straightforward: one can resort to the general algorithm for the FGM copula sketched in [19, p.185] and inverse transform sampling, when $-1 \leq \theta \leq 1$. If $\theta > 1$, an appropriate modification has to be made, but the rationale remains the same: first, sample x_1 from the marginal p.m.f. p_1, and then sample x_2 from the conditional p.m.f. $p_{2|1}$. The steps of the algorithm, which works for any consistent value of θ, are the following:

1. Simulate two independent uniform r.v.s in $(0, 1)$, $V_1 \sim \text{Unif}(0, 1)$ and $V_2 \sim \text{Unif}(0, 1)$;
2. Set $U_1 = V_1$ and let $X_1 = F_1^{-1}(U_1)$;
3. Set $U_2 = 2V_2/(a + b)$, where $a = 1 + \theta(1 - p_1(X_1) - 2F_1(X_1 - 1))$ and $b = \left[a^2 - 4(a - 1)V_2\right]^{1/2}$. Let $X_2 = F_2^{-1}(U_2)$;

where the generalized inverse of the discrete Weibull cdf (1) is $F_i^{-1}(u) = \left\lceil \left(\frac{\log(1-u)}{\log q_i}\right)^{1/\beta_i} \right\rceil - 1$. Pseudo-random simulation can then be easily implemented, by employing the `R` package `DiscreteWeibull` [2], which provides, among all, random generation of the univariate discrete Weibull distribution.

3.3 Pearson's correlation

As anticipated, Pearson's correlation between the two uniform components of the FGM copula with parameter θ is given by $\theta/3$; one may be interested in determining Pearson's correlation between two non-uniform (discrete) r.v.s linked by the FGM copula. We know, this is one of its drawbacks, that Pearson's correlation does not depend only on the copula linking the two distributions, but also on the margins themselves; so, in general, it differs from $\theta/3$. Reference [32] showed that if the margins are absolutely continuous, Pearson's correlation cannot exceed $1/3$.

Let us consider two discrete r.v.s X_1 and X_2 with c.d.f.s F_1 and F_2, respectively, defined on $\mathbb{N}_0$ or in any subset thereof. Then, recall the formula for Pearson's ρ:

$$\rho_{x_1x_2} = \frac{\mathbb{E}(X_1X_2) - \mathbb{E}(X_1)\mathbb{E}(X_2)}{\sqrt{\mathrm{Var}(X_1)\mathrm{Var}(X_2)}}$$

where $\mathbb{E}(X_1X_2) = \sum_i \sum_j ijp(i,j)$, with $p(i,j)$ given by Eq. (5), $\mathbb{E}(X_1) = \sum_i ip_1(i)$, and $\mathrm{Var}(X_1) = \sum_i (i - \mathbb{E}(X_1))^2 p_1(i)$, with p_1 being the marginal p.m.f. of X_1, and similarly for X_2; the sum over i (j) extends over the entire support of X_1 (X_2). Then Pearson's correlation coefficient between the discrete margins X_1 and X_2 of the model (5) can be easily computed, since the numerator of the expression for the correlation coefficient is

$$\begin{aligned}\mathbb{E}(X_1X_2) - \mathbb{E}(X_1)\mathbb{E}(X_2) &= \theta \sum\sum ijp_1(i)p_2(j)[1 + p_1(i) - 2F_1(i)][1 + p_2(j) - 2F_2(j)] \\ &= \theta \sum_i ip_1(i)[1 + p_1(i) - 2F_1(i)] \sum_j jp_2(j)[1 + p_2(j) - 2F_2(j)]\end{aligned}$$

so that for assigned margins, the ratio between the correlation coefficient ρ of the two r.v.s and the copula parameter θ is a constant (i.e. independent from θ). In other terms, the relationship between ρ and θ is linear, with null intercept and an angular coefficient depending on the two margins.

Table 1 reports the minimum and maximum attainable values of Pearson's correlation coefficient when we consider two discrete Weibull r.v.s linked through the FGM copula, for several marginal parameters' combinations. They are obtained letting θ take the minimum and maximum value, respectively, of its range (7). These values are computed numerically and are affected by the approximation error since the support of the discrete Weibull r.v. is infinite and there are no closed-form analytical expression for its moments. The only exception is when $\beta_1 = \beta_2 = 1$, i.e., when the two margins are geometrically distributed; in this case, the minimum attainable linear correlation can be analytically calculated as $\rho_{\min} = -\frac{\sqrt{q_1}}{1+q_1} \cdot \frac{\sqrt{q_2}}{1+q_2}$ (which can never be smaller than $-1/4$); whereas the maximum attainable linear correlation is $\rho_{\max} = \frac{\sqrt{q_1}}{1+q_1} \cdot \frac{\sqrt{q_2}}{1+q_2} \cdot \min\{1/q_1, 1/q_2\}$. In fact, in this case, it can be easily shown that $\mathbb{E}(X_1) = q_1/(1-q_1)$, $\mathrm{V}(X_1) = q_1/(1-q_1)^2$, and $\mathbb{E}(X_1X_2) - \mathbb{E}(X_1)\mathbb{E}(X_2) = \theta \frac{q_1q_2}{(1-q_1)(1+q_1)(1-q_2)(1+q_2)}$.

For example, if we consider two discrete Weibull r.v.s with parameters $q_1 = 0.7$, $\beta_1 = 0.8$, and $q_2 = 0.9$ and $\beta_2 = 1.2$, the minimum correlation coefficient attainable through the FGM copula-based bivariate model is about -0.238, the maximum correlation coefficient is 0.264. Note that relying on Table 1, the maximum value of ρ can exceed the value 1/3 (this fact occurs for "small" values of q_i and "large" values of β_i), differently from what happens with bivariate continuous distribution with the FGM dependence structure.

4 Estimation

Much less straightforward than simulation, is the estimation of parameters of the proposed bivariate discrete Weibull model. We consider a bivariate random sample of size n, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ (with $\boldsymbol{x}_i = (x_{1i}, x_{2i})$, $i = 1, \ldots, n$) of non-negative integers, which we assume has been drawn by the bivariate discrete Weibull r.v. We illustrate several possible

Table 1: Minimum and maximum Pearson correlation ρ for the bivariate discrete Weibull r.v. of (6), for several combinations of the marginal parameters

q		0.5			0.7			0.9		
	β	0.8	1	1.2	0.8	1	1.2	0.8	1	1.2
0.5	0.8	-0.191	-0.206	-0.215	-0.199	-0.215	-0.225	-0.201	-0.218	-0.229
		0.383	0.413	0.429	0.284	0.308	0.321	0.223	0.243	0.254
	1		-0.222	-0.231	-0.214	-0.232	-0.242	-0.217	-0.235	-0.247
			0.444	0.462	0.306	0.331	0.346	0.241	0.262	0.274
	1.2			-0.240	-0.223	-0.241	-0.252	-0.225	-0.245	-0.256
				0.481	0.318	0.345	0.360	0.250	0.272	0.285
0.7	0.8				-0.206	-0.224	-0.233	-0.209	-0.227	-0.238
					0.295	0.319	0.333	0.232	0.252	0.264
	1					-0.242	-0.253	-0.226	-0.246	-0.257
						0.346	0.361	0.251	0.273	0.286
	1.2						-0.264	-0.236	-0.257	-0.269
							0.377	0.262	0.285	0.299
0.9	0.8							-0.211	-0.229	-0.240
								0.235	0.255	0.267
	1								-0.249	-0.261
									0.277	0.290
	1.2									-0.274
										0.304

methods for estimating the parameters q_1, β_1, q_2, β_2, and θ. We will start from the most standard method (maximum likelihood), which should reveal the most efficient but also the most computationally demanding, and then move to easier, but also presumably less efficient, techniques. Estimation of this bivariate discrete model inherits all the drawbacks related to the estimation of copulas for count data, see [11].

4.1 Full maximum likelihood

The log-likelihood function has the usual expression:

$$\log L(q_1, \beta_1, q_2, \beta_2, \theta; \boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = \sum_{i=1}^{n} \log p(x_{1i}, x_{2i}; \boldsymbol{\eta}), \tag{10}$$

with $\boldsymbol{\eta}$ denoting the parameter vector $(q_1, \beta_1, q_2, \beta_2, \theta)$ and $p(x_{1i}, x_{2i}; \boldsymbol{\eta})$ from (2). The "full maximum likelihood method" simply consists in maximizing the log-likelihood function of Eq. (10) with respect to all the five parameters simultaneously, incorporating the constraints on the parameters' values:

$$\hat{\boldsymbol{\eta}} = \arg\max_{\boldsymbol{\eta}} \log L(\boldsymbol{\eta}; \boldsymbol{x}_1, \dots, \boldsymbol{x}_n).$$

This task can be performed numerically; under the `R` environment, it can be implemented using appropriate functions (e.g. `optim`). Deriving the asymptotic variances of the full maximum likelihood estimators is an arduous task, because of the complex form of the p.m.f. in (2) and the unavailability of closed-form expressions for its expected value and variance.

4.2 Two-step maximum likelihood

The method, largely used in many other situations for copula estimation, see [17], consists in computing first the maximum likelihood estimates of the two discrete Weibull margins independently:

$$(\hat{q}_j, \hat{\beta}_j) = \arg\max_{q_j,\beta_j} \log L_j(q_j, \beta_j; x_{j1}, \ldots, x_{jn}) = \sum_{i=1}^{n} \log p_j(x_{ji}; q_j, \beta_j) \quad j = 1, 2 \tag{11}$$

(one can use the function `estdweibull` in the `DiscreteWeibull` package) and then maximize the log-likelihood of Eq. (10) with respect to θ only, with the four marginal components of the $\boldsymbol{\eta}$ parameter vector substituted by their corresponding MLEs of Eq. (11). Note that for univariate samples consisting of only 0s and 1s the maximum-likelihood procedure is inapplicable [4].

4.3 Marginal maximum likelihood plus approximate method of moments

This method estimates the marginal parameters and the copula parameter separately. The former are computed as (independent) maximum likelihood estimates for the two discrete Weibull distributions (as for the two-step maximum likelihood method, Eq. (11)). The latter is computed resorting to the known relationship existing between the θ parameter of the FGM copula and Spearman's rank correlation (holding when X_1 and X_2 are continuous r.v.s, see e.g. [20, p.213]:

$$\rho_S(X_1, X_2) = \frac{\theta}{3}. \tag{12}$$

If one compute a sample estimate of ρ_S (e.g. the sample rank correlation), $\hat{\rho}_S$, then an estimate of θ can be computed as

$$\hat{\theta}_M = 3\hat{\rho}_S. \tag{13}$$

Of course, while this method may produce reliable estimates when X_1 and X_2 are continuous r.v.s, nothing can be said when they are discrete: in this case, Eq. (12) does not hold exactly.

Alternatively, one can consider Kendall's tau $\tau(X_1, X_2)$: for the FGM copula linking two continuous r.v.s, the following relationship holds [20, p.213]:

$$\tau(X_1, X_2) = \frac{2\theta}{9}. \tag{14}$$

If one compute a sample estimate of τ, $\hat{\tau}$, then an estimate of θ can be computed as $\hat{\theta}'_M = 9\hat{\tau}/2$.

Note that although the estimate provided for θ by Eq. (12) (or Eq. (14)) is very easy to compute, it may be "not feasible", i.e., may fall outside the valid range for θ.

4.4 Method of proportion

The original method of proportion we propose here is itself a two-step method. First, we compute the estimates of the marginal parameters of the two r.v.s X_1 and X_2 independently, by resorting to the univariate method of proportion (see [12, 15, 4]). Thus we obtain the estimates

$$\hat{q}_{i,P} = 1 - \hat{p}_{0i}, i = 1, 2, \tag{15}$$

where $\hat{p}_{0i}$ is the sample rate of zeros of X_i;

$$\hat{\beta}_{i,P} = \log\left[\log(\hat{q}_{i,P} - \hat{p}_{1i})/\log(\hat{q}_{i,P})\right]/\log(2),$$

where $\hat{p}_{1i}$ is the sample rate of ones of X_i. Then the copula parameter θ can be estimated by observing that

$$F(0,0) = P(X_1 = 0, X_2 = 0) = (1 - q_1)(1 - q_2)[1 + \theta q_1 q_2]$$

and then, substituting to q_1 and q_2 the corresponding estimates in (15), we derive an easy estimate for the copula parameter θ, by taking into account the proportion of $(0,0)$ pairs in the sample, denoted by $\hat{p}_{00}$:

$$\hat{\theta}_P = \left(\frac{\hat{p}_{00}}{(1 - \hat{q}_{1,P})(1 - \hat{q}_{2,P})} - 1\right)/(\hat{q}_{1,P}\hat{q}_{2,P}). \tag{16}$$

Thus, according to this method, all the five parameters can be computed through analytical expressions. However, some caution is needed. First, as known (see, for example, [4]), some samples may be not able to provide (univariate) estimates for the q and β marginal parameters: this happens if either (univariate) sample does not contain at least a 0 and at least a 1. Secondly, even when this first issue is overcome, the estimate of θ of Eq. (16) that can be then calculated may be "not feasible", i.e., may fall outside the valid range for θ.

Estimating θ by equating the probability of a double zero to the sample proportion of a double zero ($\hat{p}_{00}$) does not represent the only possible choice: ideally, we can use any joint cumulative probability $F(x_1, x_2)$, $(x_1, x_2) \in \mathbb{N}_0^2$ and equate it to the corresponding sample cumulative frequency $\hat{F}(x_1, x_2)$. The choice of $(0,0)$ is just the one providing the simplest expression for $\hat{\theta}_P$ and also the most natural, since the marginal parameters q_i and β_i are estimated on the basis of the sample proportions of zeros for the two empirical marginal distributions. As an example, equating the joint $(1,1)$ probability to the corresponding sample rate, we would get

$$\hat{\theta}_P^* = \left[\frac{\hat{F}_{11}}{(1 - \hat{q}_{1,P}^{2^{\hat{\beta}_{1,P}}})(1 - \hat{q}_{2,P}^{2^{\hat{\beta}_{2,P}}})} - 1\right]/(\hat{q}_{1,P}^{2^{\hat{\beta}_{1,P}}} \cdot \hat{q}_{2,P}^{2^{\hat{\beta}_{2,P}}})$$

5 A Monte Carlo study

In order to study the behaviour of the parameters' estimators in terms of biasedness and variability, one would like to obtain their expectations and variances. Finding these means and variances is almost impossible, especially for methods involving maximum likelihood, which provide estimates in a numerical form only. Therefore, we study numerically the expressions for expected values and variances with a Monte Carlo simulation study. We consider several representative combinations of the five parameters characterizing the bivariate discrete Weibull distribution; in particular, for the sake of simplicity, we consider all the "reasonable" and "consistent" combinations arising from the following choice of parameters: $\theta = -0.9, -0.6; -0.3; 0; +0.3, +0.6, +0.9, +1.2$ (thus leading to negative, null or positive dependence between the margins); $q_1, q_2 = 0.5; 0.7; 0.9$; $\beta_1, \beta_2 = 0.8; 1.2$. "Reasonable" here means that given the symmetrical nature of the problem, if the combination corresponding to the parameter vector $\boldsymbol{\eta} = (q_1, \beta_1, q_2, \beta_2, \theta)$ is considered, the combination associated to the "symmetrical" parameter vector $\boldsymbol{\eta}^* = (q_2, \beta_2, q_1, \beta_1, \theta)$ will be skipped. "Consistent" means that the value of θ should satisfy the constraint of Eq. (7); this leads to exclude the combinations where $\theta = 1.2$ and either q_1 or q_2 is equal to 0.9. This means that $21 \times 7 + 10 = 157$ combinations for $\boldsymbol{\eta}$ are eventually considered. For each of these artificial settings, the Monte Carlo simulation study is performed by generating $1,000$ samples of size $n = 100$, using the algorithm given in Section 3. For each sample, the four types of estimators of q_1, q_2, β_1, β_2 and θ, described in Section 4, are calculated. Since the actual evaluation of $\hat{\tau}$ for large n is time consuming (in comparison with $\hat{\rho}_S$), among the two moment method's estimators of Section 4.3, only the estimator (13) is considered. Then, over the $1,000$ simulations, the means, $\bar{\hat{\theta}}$, and standard deviations, $\mathrm{sd}(\hat{\theta})$, of these estimates are obtained.

Here we report only the results related to the estimates of the dependence parameter θ, which are expected to be the most meaningful, since the copula parameter is intuitively the most difficult to be estimated. We will discuss the performance of the estimators comparatively, underlying how the parameter values (of θ itself and the other marginal parameters) affect it. Results are reported in Tables 2–9.

Table 2: Simulation results: Monte Carlo expected value and standard deviation of the four estimators for the θ parameter presented in Section 4, with $\theta = 1.2$ and $n = 100$. Abbreviations: P=method of proportion, M=two-step maximum likelihood plus moment method; TSML=two-step maximum likelihood; ML=full maximum likelihood

		$q = 0.5, \beta = 0.8$		$q = 0.5, \beta = 1.2$		$q = 0.7, \beta = 0.8$		$q = 0.7, \beta = 1.2$	
		$\hat{\hat{\theta}}$	sd($\hat{\theta}$)	$\hat{\hat{\theta}}$	sd($\hat{\theta}$)	$\hat{\hat{\theta}}$	sd($\hat{\theta}$)	$\hat{\hat{\theta}}$	sd($\hat{\theta}$)
P		1.218	0.395						
M	$q = 0.5,$	1.034	0.281						
TSML	$\beta = 0.8$	1.195	0.314						
ML		1.202	0.318						
P		1.222	0.394	1.222	0.394				
M	$q = 0.5,$	1.025	0.278	1.016	0.275				
TSML	$\beta = 1.2$	1.198	0.316	1.201	0.318				
ML		1.205	0.319	1.208	0.322				
P		1.222	0.429	1.222	0.429	1.209	0.504		
M	$q = 0.7,$	1.095	0.269	1.085	0.269	1.153	0.259		
TSML	$\beta = 0.8$	1.172	0.241	1.172	0.243	1.172	0.214		
ML		1.179	0.244	1.178	0.245	1.179	0.217		
P		1.223	0.429	1.223	0.429	1.210	0.503	1.210	0.503
M	$q = 0.7,$	1.087	0.266	1.078	0.266	1.143	0.256	1.133	0.256
TSML	$\beta = 1.2$	1.174	0.239	1.174	0.241	1.172	0.212	1.172	0.212
ML		1.181	0.241	1.180	0.243	1.181	0.214	1.180	0.214

Table 3: Simulation results: Monte Carlo expected value and standard deviation of the four estimators for the θ parameter presented in Section 4, with $\theta = 0.9$ and $n = 100$. Abbreviations: P=method of proportion, M=two-step maximum likelihood plus moment method; TSML=two-step maximum likelihood; ML=full maximum likelihood

		$q = 0.5, \beta = 0.8$		$q = 0.5, \beta = 1.2$		$q = 0.7, \beta = 0.8$		$q = 0.7, \beta = 1.2$		$q = 0.9, \beta = 0.8$		$q = 0.9, \beta = 1.2$	
		$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$
P		0.916	0.404										
M	$q = 0.5,$	0.776	0.293										
TSML	$\beta = 0.8$	0.898	0.336										
ML		0.903	0.338										
P		0.917	0.402	0.917	0.402								
M	$q = 0.5,$	0.768	0.293	0.762	0.293								
TSML	$\beta = 1.2$	0.899	0.339	0.901	0.344								
ML		0.903	0.341	0.905	0.346								
P		0.912	0.443	0.911	0.443	0.908	0.501						
M	$q = 0.7,$	0.821	0.285	0.812	0.285	0.866	0.282						
TSML	$\beta = 0.8$	0.896	0.300	0.895	0.303	0.893	0.276						
ML		0.901	0.302	0.899	0.305	0.898	0.279						
P		0.912	0.440	0.912	0.440	0.913	0.502	0.913	0.502				
M	$q = 0.7,$	0.815	0.282	0.807	0.282	0.860	0.281	0.854	0.281				
TSML	$\beta = 1.2$	0.898	0.300	0.897	0.301	0.894	0.276	0.895	0.278				
ML		0.903	0.302	0.902	0.304	0.900	0.279	0.901	0.281				
P		0.931	0.641	0.931	0.641	0.869	0.839	0.869	0.839	0.921	1.459		
M	$q = 0.9,$	0.833	0.286	0.827	0.285	0.878	0.281	0.871	0.280	0.892	0.280		
TSML	$\beta = 0.8$	0.863	0.250	0.865	0.252	0.870	0.236	0.870	0.237	0.871	0.229		
ML		0.866	0.250	0.868	0.252	0.873	0.237	0.874	0.238	0.873	0.231		
P		0.931	0.642	0.931	0.642	0.869	0.840	0.869	0.840	0.919	1.460	0.919	1.460
M	$q = 0.9,$	0.830	0.285	0.825	0.283	0.876	0.281	0.869	0.280	0.888	0.281	0.888	0.280
TSML	$\beta = 1.2$	0.861	0.250	0.864	0.250	0.869	0.237	0.868	0.237	0.869	0.229	0.869	0.230
ML		0.865	0.250	0.868	0.251	0.872	0.237	0.872	0.238	0.872	0.230	0.873	0.230

Table 4: Simulation results: Monte Carlo expected value and standard deviation of the four estimators for the θ parameter presented in Section 4, with $\theta = 0.6$ and $n = 100$. Abbreviations: P=method of proportion, M=two-step maximum likelihood plus moment method; TSML=two-step maximum likelihood; ML=full maximum likelihood

		$q=0.5, \beta=0.8$		$q=0.5, \beta=1.2$		$q=0.7, \beta=0.8$		$q=0.7, \beta=1.2$		$q=0.9, \beta=0.8$		$q=0.9, \beta=1.2$	
		$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$
P		0.615	0.412										
M	$q=0.5,$	0.519	0.306										
TSML	$\beta=0.8$	0.600	0.352										
ML		0.602	0.354										
P		0.611	0.410	0.611	0.410								
M	$q=0.5,$	0.511	0.304	0.507	0.302								
TSML	$\beta=1.2$	0.597	0.354	0.600	0.356								
ML		0.600	0.356	0.602	0.358								
P		0.605	0.455	0.605	0.455	0.607	0.505						
M	$q=0.7,$	0.549	0.301	0.543	0.301	0.579	0.295						
TSML	$\beta=0.8$	0.600	0.330	0.601	0.334	0.598	0.301						
ML		0.603	0.332	0.604	0.336	0.601	0.304						
P		0.602	0.452	0.602	0.452	0.607	0.502	0.607	0.502				
M	$q=0.7,$	0.543	0.298	0.537	0.299	0.574	0.294	0.571	0.290				
TSML	$\beta=1.2$	0.601	0.329	0.600	0.334	0.598	0.303	0.601	0.302				
ML		0.604	0.332	0.603	0.336	0.602	0.305	0.604	0.305				
P		0.630	0.676	0.630	0.676	0.576	0.810	0.576	0.810	0.613	1.400		
M	$q=0.9,$	0.558	0.304	0.554	0.303	0.587	0.296	0.582	0.294	0.597	0.298		
TSML	$\beta=0.8$	0.597	0.317	0.600	0.319	0.596	0.291	0.596	0.291	0.596	0.285		
ML		0.600	0.319	0.602	0.321	0.598	0.292	0.598	0.292	0.598	0.287		
P		0.629	0.676	0.629	0.676	0.577	0.809	0.577	0.809	0.616	1.399	0.616	1.319
M	$q=0.9,$	0.556	0.303	0.551	0.303	0.585	0.296	0.581	0.293	0.595	0.298	0.592	0.298
TSML	$\beta=1.2$	0.596	0.317	0.597	0.320	0.595	0.292	0.595	0.291	0.595	0.287	0.594	0.289
ML		0.599	0.318	0.599	0.321	0.597	0.293	0.598	0.292	0.598	0.289	0.597	0.290

Table 5: Simulation results: Monte Carlo expected value and standard deviation of the four estimators for the θ parameter presented in Section 4, with $\theta = 0.3$ and $n = 100$. Abbreviations: P=method of proportion, M=two-step maximum likelihood plus moment method; TSML=two-step maximum likelihood; ML=full maximum likelihood

		$q = 0.5, \beta = 0.8$		$q = 0.5, \beta = 1.2$		$q = 0.7, \beta = 0.8$		$q = 0.7, \beta = 1.2$		$q = 0.9, \beta = 0.8$		$q = 0.9, \beta = 1.2$	
		$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$	$\hat{\hat{\theta}}$	$\mathrm{sd}(\hat{\theta})$
P		0.313	0.415										
M	$q = 0.5,$	0.262	0.309										
TSML	$\beta = 0.8$	0.301	0.353										
ML		0.303	0.355										
P		0.313	0.415	0.313	0.415								
M	$q = 0.5,$	0.259	0.308	0.255	0.305								
TSML	$\beta = 1.2$	0.302	0.357	0.300	0.358								
ML		0.303	0.358	0.301	0.359								
P		0.308	0.457	0.308	0.457	0.303	0.502						
M	$q = 0.7,$	0.277	0.307	0.273	0.305	0.293	0.306						
TSML	$\beta = 0.8$	0.302	0.334	0.301	0.336	0.301	0.315						
ML		0.303	0.336	0.302	0.337	0.302	0.316						
P		0.308	0.457	0.308	0.457	0.303	0.502	0.303	0.502				
M	$q = 0.7,$	0.275	0.306	0.271	0.304	0.290	0.305	0.286	0.305				
TSML	$\beta = 1.2$	0.303	0.336	0.302	0.338	0.302	0.317	0.300	0.320				
ML		0.304	0.338	0.303	0.339	0.303	0.318	0.301	0.321				
P		0.316	0.692	0.317	0.690	0.279	0.815	0.279	0.815	0.258	1.286		
M	$q = 0.9,$	0.281	0.307	0.277	0.304	0.296	0.307	0.292	0.307	0.300	0.306		
TSML	$\beta = 0.8$	0.303	0.330	0.303	0.330	0.301	0.310	0.300	0.314	0.299	0.303		
ML		0.305	0.331	0.304	0.331	0.303	0.312	0.301	0.315	0.300	0.305		
P		0.316	0.692	0.316	0.692	0.279	0.815	0.279	0.815	0.258	1.286	0.258	1.286
M	$q = 0.9,$	0.279	0.307	0.276	0.305	0.294	0.306	0.290	0.307	0.298	0.306	0.298	0.307
TSML	$\beta = 1.2$	0.302	0.331	0.302	0.332	0.300	0.311	0.298	0.314	0.297	0.304	0.298	0.305
ML		0.303	0.332	0.303	0.334	0.301	0.313	0.300	0.316	0.299	0.305	0.300	0.306

Table 6: Simulation results: Monte Carlo expected value and standard deviation of the four estimators for the θ parameter presented in Section 4, with $\theta = 0$ and $n = 100$. Abbreviations: P=method of proportion, M=two-step maximum likelihood plus moment method; TSML=two-step maximum likelihood; ML=full maximum likelihood

		$q=0.5, \beta=0.8$		$q=0.5, \beta=1.2$		$q=0.7, \beta=0.8$		$q=0.7, \beta=1.2$		$q=0.9, \beta=0.8$		$q=0.9, \beta=1.2$	
		$\bar{\hat{\theta}}$	sd($\hat{\theta}$)	$\bar{\hat{\theta}}$	sd($\hat{\theta}$)	$\bar{\hat{\theta}}$	sd($\hat{\theta}$)	$\bar{\hat{\theta}}$	sd($\hat{\theta}$)	$\bar{\hat{\theta}}$	sd($\hat{\theta}$)	$\bar{\hat{\theta}}$	sd($\hat{\theta}$)
P		0.006	0.412										
M	$q=0.5,$	0.002	0.308										
TSML	$\beta=0.8$	0.002	0.357										
ML		0.002	0.358										
P		0.006	0.412	0.006	0.412								
M	$q=0.5,$	0.002	0.306	0.001	0.306								
TSML	$\beta=1.2$	0.001	0.359	0.000	0.363								
ML		0.001	0.361	0.000	0.364								
P		-0.003	0.462	-0.003	0.462	0.003	0.497						
M	$q=0.7,$	0.002	0.311	0.001	0.309	0.004	0.307						
TSML	$\beta=0.8$	0.003	0.342	0.001	0.345	0.004	0.319						
ML		0.003	0.344	0.001	0.346	0.004	0.320						
P		-0.003	0.462	-0.003	0.462	0.003	0.497	0.003	0.497				
M	$q=0.7,$	0.003	0.311	0.001	0.309	0.005	0.307	0.004	0.307				
TSML	$\beta=1.2$	0.004	0.346	0.002	0.348	0.004	0.322	0.004	0.325				
ML		0.004	0.348	0.002	0.350	0.004	0.323	0.004	0.326				
P		-0.016	0.718	-0.014	0.715	-0.029	0.786	-0.029	0.786	-0.021	1.169		
M	$q=0.9,$	0.001	0.311	0.000	0.308	0.004	0.309	0.003	0.309	0.004	0.308		
TSML	$\beta=0.8$	0.003	0.336	0.002	0.336	0.004	0.315	0.003	0.318	0.003	0.308		
ML		0.003	0.337	0.002	0.338	0.003	0.316	0.003	0.319	0.003	0.310		
P		-0.016	0.718	-0.016	0.718	-0.029	0.786	-0.029	0.786	-0.021	1.169	-0.021	1.169
M	$q=0.9,$	0.000	0.310	-0.001	0.308	0.002	0.308	0.002	0.308	0.003	0.308	0.004	0.308
TSML	$\beta=1.2$	0.001	0.336	0.000	0.338	0.002	0.315	0.002	0.318	0.002	0.309	0.002	0.310
ML		0.002	0.338	-0.001	0.340	0.002	0.317	0.002	0.319	0.002	0.311	0.002	0.311

Table 7: Simulation results: Monte Carlo expected value and standard deviation of the four estimators for the θ parameter presented in Section 4, with $\theta = -0.3$ and $n = 100$. Abbreviations: P=method of proportion, M=two-step maximum likelihood plus moment method; TSML=two-step maximum likelihood; ML=full maximum likelihood

		$q = 0.5, \beta = 0.8$		$q = 0.5, \beta = 1.2$		$q = 0.7, \beta = 0.8$		$q = 0.7, \beta = 1.2$		$q = 0.9, \beta = 0.8$		$q = 0.9, \beta = 1.2$	
		$\bar{\hat{\theta}}$	sd($\hat{\theta}$)	$\bar{\hat{\theta}}$	sd($\hat{\theta}$)	$\bar{\hat{\theta}}$	sd($\hat{\theta}$)	$\bar{\hat{\theta}}$	sd($\hat{\theta}$)	$\bar{\hat{\theta}}$	sd($\hat{\theta}$)	$\bar{\hat{\theta}}$	sd($\hat{\theta}$)
P		-0.293	0.405										
M	$q = 0.5,$	-0.254	0.298										
TSML	$\beta = 0.8$	-0.293	0.341										
ML		-0.294	0.343										
P		-0.293	0.405	-0.293	0.405								
M	$q = 0.5,$	-0.252	0.297	-0.249	0.299								
TSML	$\beta = 1.2$	-0.293	0.343	-0.292	0.348								
ML		-0.294	0.344	-0.293	0.349								
P		-0.304	0.454	-0.304	0.454	-0.301	0.483						
M	$q = 0.7,$	-0.269	0.302	-0.266	0.303	-0.281	0.303						
TSML	$\beta = 0.8$	-0.292	0.327	-0.292	0.332	-0.291	0.312						
ML		-0.294	0.328	-0.293	0.333	-0.292	0.313						
P		-0.304	0.454	-0.304	0.454	-0.301	0.483	-0.301	0.483				
M	$q = 0.7,$	-0.266	0.302	-0.263	0.303	-0.279	0.303	-0.278	0.303				
TSML	$\beta = 1.2$	-0.292	0.332	-0.292	0.337	-0.291	0.315	-0.292	0.318				
ML		-0.294	0.333	-0.293	0.338	-0.292	0.317	-0.293	0.319				
P		-0.334	0.699	-0.334	0.699	-0.310	0.723	-0.310	0.723	-0.315	1.055		
M	$q = 0.9,$	-0.274	0.303	-0.272	0.304	-0.286	0.304	-0.286	0.304	-0.292	0.304		
TSML	$\beta = 0.8$	-0.293	0.322	-0.293	0.326	-0.292	0.307	-0.293	0.309	-0.293	0.301		
ML		-0.294	0.323	-0.294	0.328	-0.293	0.308	-0.294	0.310	-0.295	0.303		
P		-0.334	0.699	-0.334	0.699	-0.310	0.723	-0.310	0.723	-0.315	1.055	-0.315	1.055
M	$q = 0.9,$	-0.275	0.302	-0.272	0.303	-0.287	0.304	-0.286	0.304	-0.293	0.304	-0.292	0.304
TSML	$\beta = 1.2$	-0.295	0.323	-0.295	0.327	-0.293	0.308	-0.294	0.310	-0.295	0.302	-0.295	0.303
ML		-0.297	0.324	-0.296	0.328	-0.294	0.309	-0.295	0.311	-0.296	0.304	-0.297	0.304

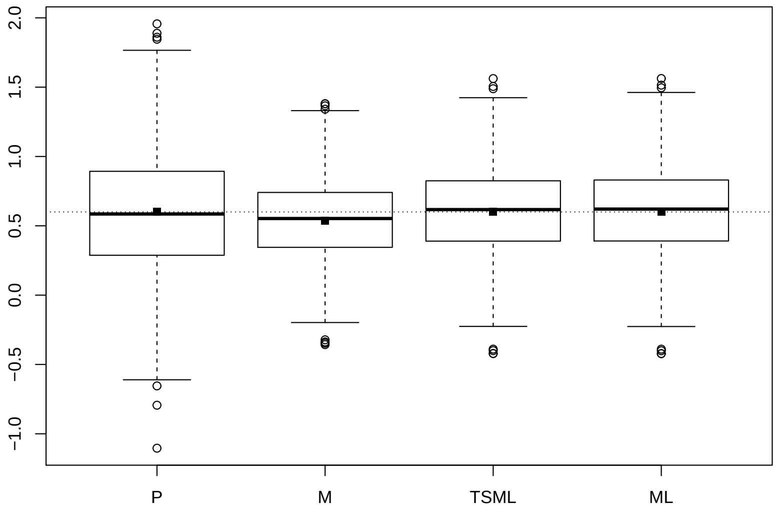
Table 8: Simulation results: Monte Carlo expected value and standard deviation of the four estimators for the θ parameter presented in Section 4, with $\theta = -0.6$ and $n = 100$. Abbreviations: P=method of proportion, M=two-step maximum likelihood plus moment method; TSML=two-step maximum likelihood; ML=full maximum likelihood

		$q = 0.5, \beta = 0.8$		$q = 0.5, \beta = 1.2$		$q = 0.7, \beta = 0.8$		$q = 0.7, \beta = 1.2$		$q = 0.9, \beta = 0.8$		$q = 0.9, \beta = 1.2$	
		$\hat{\hat{\theta}}$	sd($\hat{\theta}$)	$\hat{\hat{\theta}}$	sd($\hat{\theta}$)	$\hat{\hat{\theta}}$	sd($\hat{\theta}$)	$\hat{\hat{\theta}}$	sd($\hat{\theta}$)	$\hat{\hat{\theta}}$	sd($\hat{\theta}$)	$\hat{\hat{\theta}}$	sd($\hat{\theta}$)
P		-0.595	0.409										
M	$q = 0.5,$	-0.511	0.292										
TSML	$\beta = 0.8$	-0.575	0.304										
ML		-0.577	0.305										
P		-0.595	0.409	-0.595	0.409								
M	$q = 0.5,$	-0.506	0.292	-0.500	0.291								
TSML	$\beta = 1.2$	-0.574	0.306	-0.572	0.309								
ML		-0.576	0.307	-0.574	0.309								
P		-0.606	0.451	-0.606	0.451	-0.600	0.457						
M	$q = 0.7,$	-0.540	0.295	-0.533	0.293	-0.571	0.294						
TSML	$\beta = 0.8$	-0.576	0.292	-0.574	0.294	-0.578	0.276						
ML		-0.578	0.292	-0.577	0.295	-0.580	0.277						
P		-0.606	0.451	-0.606	0.451	-0.600	0.457	-0.600	0.457				
M	$q = 0.7,$	-0.535	0.295	-0.529	0.293	-0.566	0.294	-0.560	0.296				
TSML	$\beta = 1.2$	-0.576	0.296	-0.574	0.298	-0.578	0.280	-0.577	0.285				
ML		-0.578	0.297	-0.576	0.299	-0.580	0.280	-0.579	0.285				
P		-0.630	0.696	-0.630	0.696	-0.595	0.644	-0.595	0.644	-0.596	0.909		
M	$q = 0.9,$	-0.549	0.296	-0.543	0.295	-0.579	0.294	-0.574	0.296	-0.589	0.293		
TSML	$\beta = 0.8$	-0.577	0.289	-0.576	0.291	-0.579	0.274	-0.579	0.279	-0.582	0.270		
ML		-0.579	0.290	-0.578	0.292	-0.582	0.274	-0.581	0.280	-0.585	0.272		
P		-0.630	0.696	-0.630	0.696	-0.595	0.644	-0.595	0.644	-0.596	0.909	-0.596	0.909
M	$q = 0.9,$	-0.549	0.295	-0.543	0.294	-0.579	0.293	-0.574	0.296	-0.589	0.293	-0.588	0.292
TSML	$\beta = 1.2$	-0.579	0.289	-0.577	0.291	-0.581	0.274	-0.580	0.279	-0.583	0.271	-0.583	0.272
ML		-0.581	0.290	-0.579	0.291	-0.583	0.275	-0.582	0.280	-0.586	0.272	-0.586	0.273

Table 9: Simulation results: Monte Carlo expected value and standard deviation of the four estimators for the θ parameter presented in Section 4, with $\theta = -0.9$ and $n = 100$. Abbreviations: P=method of proportion, M=two-step maximum likelihood plus moment method; TSML=two-step maximum likelihood; ML=full maximum likelihood

		$q = 0.5, \beta = 0.8$		$q = 0.5, \beta = 1.2$		$q = 0.7, \beta = 0.8$		$q = 0.7, \beta = 1.2$		$q = 0.9, \beta = 0.8$		$q = 0.9, \beta = 1.2$	
		$\hat{\hat{\theta}}$	sd($\hat{\theta}$)	$\hat{\hat{\theta}}$	sd($\hat{\theta}$)	$\hat{\hat{\theta}}$	sd($\hat{\theta}$)	$\hat{\hat{\theta}}$	sd($\hat{\theta}$)	$\hat{\hat{\theta}}$	sd($\hat{\theta}$)	$\hat{\hat{\theta}}$	sd($\hat{\theta}$)
P		-0.892	0.404										
M	$q = 0.5,$	-0.769	0.275										
TSML	$\beta = 0.8$	-0.817	0.222										
ML		-0.819	0.222										
P		-0.892	0.404	-0.892	0.404								
M	$q = 0.5,$	-0.761	0.278	-0.753	0.282								
TSML	$\beta = 1.2$	-0.814	0.226	-0.813	0.232								
ML		-0.816	0.225	-0.814	0.232								
P		-0.906	0.449	-0.906	0.449	-0.907	0.430						
M	$q = 0.7,$	-0.813	0.279	-0.803	0.282	-0.859	0.277						
TSML	$\beta = 0.8$	-0.822	0.214	-0.819	0.219	-0.829	0.200						
ML		-0.824	0.213	-0.821	0.218	-0.832	0.199						
P		-0.906	0.449	-0.906	0.449	-0.907	0.430	-0.907	0.430				
M	$q = 0.7,$	-0.806	0.279	-0.796	0.282	-0.852	0.278	-0.846	0.277				
TSML	$\beta = 1.2$	-0.820	0.216	-0.817	0.221	-0.828	0.202	-0.827	0.201				
ML		-0.823	0.216	-0.819	0.221	-0.830	0.202	-0.829	0.200				
P		-0.931	0.669	-0.931	0.669	-0.904	0.576	-0.904	0.576	-0.917	0.647		
M	$q = 0.9,$	-0.826	0.280	-0.817	0.283	-0.872	0.277	-0.865	0.276	-0.885	0.273		
TSML	$\beta = 0.8$	-0.823	0.213	-0.821	0.217	-0.829	0.198	-0.829	0.198	-0.832	0.192		
ML		-0.826	0.212	-0.824	0.217	-0.832	0.198	-0.832	0.197	-0.835	0.192		
P		-0.931	0.669	-0.931	0.669	-0.904	0.576	-0.904	0.576	-0.917	0.647	-0.917	0.647
M	$q = 0.9,$	-0.826	0.279	-0.816	0.282	-0.871	0.276	-0.865	0.276	-0.884	0.273	-0.882	0.273
TSML	$\beta = 1.2$	-0.824	0.212	-0.821	0.217	-0.830	0.198	-0.830	0.198	-0.833	0.193	-0.833	0.193
ML		-0.826	0.212	-0.824	0.216	-0.833	0.198	-0.832	0.197	-0.836	0.192	-0.836	0.193

Figure 3: Boxplots displaying the MC empirical distribution of the θ sample estimators for $q_1 = 0.7$, $q_2 = 0.5$, $\beta_1 = \beta_2 = 1.2$, $\theta = 0.6$, $n = 100$. The horizontal line indicates $\theta = 0.6$. For each boxplot, a black filled square indicates the Monte Carlo average of the corresponding estimator distribution. Abbreviations: P=method of proportion, M=two-step maximum likelihood plus moment method; TSML=two-step maximum likelihood; ML=full maximum likelihood



It is apparent that the method of proportion is the least efficient amongst the four estimators of θ. Indeed, it is sometimes the least biased, but has always by far the largest standard deviation, thus producing the largest root mean square error. This was somewhat expected, since as in the univariate case (see [4]), this method discards most of the information contained in the sample, being based only on the rates of 0s and 1s in the data. Moreover, the percentage of samples leading to non-feasible estimates (not reported here for each single scenario) is often non-negligible. Note that the estimates of θ derived through the method of proportion (see Eq. (16)) depends on the two marginal frequencies of 0s and on the bivariate sample frequency of double 0s. But the corresponding probabilities do not depend on the β parameters, so the Monte Carlo mean of $\hat{\theta}$ is a function of θ, q_1 and q_2 only, and is independent from the two β's. This is easily confirmed by the simulation results which also show that even the standard deviation of $\hat{\theta}$ does not depend on β's, but only on q's, θ itself and n. In particular, the standard deviation, for a fixed θ, increases when the q's increase; for fixed values of the marginal parameters, the standard deviation does not possess a clear trend with respect to θ.

As for the estimator obtained through the method of moment, we can note that the bias is usually negative for $\theta > 0$ and positive for $\theta < 0$, and, for a fixed value of θ, in absolute value it tends to diminish when q increases and tends to 1, and - fixed q - when β is smaller; fixed q's and β's, it tends to increase as the absolute value of θ increases. Thus, for a fixed θ, the smallest value of the bias is obtained for the scenario $q_1 = q_2 = 0.9, \beta_1 = \beta_2 = 0.8$. The standard deviation, for a fixed θ, tends to slightly decrease by increasing q, whereas it is quite stable with β; it tends to diminish as the absolute value of θ increases, for a fixed combination of q's and β's; this means that "extreme" values of θ (1.2 and -0.9, in our study) are those yielding the smallest variability, for a given marginal setting.

The estimators obtained through the full maximum likelihood and two-step maximum likelihood methods have a very similar behaviour. In terms of bias, the former would be preferable to the latter, except when $\theta = 0.3$. In terms of variability, the latter is better than the former for most scenarios. The bias magnitude is negligible for $\theta = -0.3, 0, 0.3, 0.6$; it becomes not entirely insignificant for the remaining values of θ, and in particular for $\theta = -0.9$. For a fixed θ, the biases increase in absolute value when the q's get close to the smallest value here considered (0.5) and is practically independent from β_i. Looking for example Table 9, with $q_1 = q_2 = 0.5$ and $\beta_1 = \beta_2 = 0.8$, the Monte Carlo mean of the TSML and ML estimators of θ are -0.817 and -0.819, whereas the "true" value of θ is -0.9. As for their standard deviations, they do not depend on θ heavily, and for a fixed θ they decrease with q, and once q is fixed, they slightly increase with β's. The corresponding root mean square error (which is not reported here for the sake of brevity) almost always turns out to be smaller - even if to a barely detectable extent - for the full maximum likelihood estimator. The "price" to pay is the greater computational cost. Although the full maximization of the log-likelihood function can be computationally demanding (however, in the order few seconds), however it did not raise particular convergence problems, under any of the examined scenarios. The TSML and ML estimators of θ are overall almost always better than the estimator derived through the method of moments. For most scenarios they have a far smaller absolute bias (even if their variance may be slightly larger); for a few other scenarios, they have a slightly larger absolute bias, which is however compensated by a far smaller variance; so that, in terms of root mean square error the maximum likelihood estimators are preferable to the method of moments. The only exception is when θ is zero (i.e., the margins are independent): in this case, the method of moment can be considered the best performer. See Table 6: for $q_1 = 0.9$, $\beta_1 = 1.2$, $q_2 = 0.5$, $\beta_2 = 0.8$, we have that the method of moment's estimator of θ is practically unbiased, with standard deviation 0.310, whereas the TSML and ML estimators have bias 0.001 and 0.002, and standard deviation 0.336 and 0.338, respectively.

Figure 3 displays the Monte Carlo empirical distribution of the four estimators of θ for another of the examined settings. From this boxplot, it is apparent that the estimator derived by the method of proportion presents the largest variability, although it is almost unbiased, and is characterized by the presence of several "outliers", which can even fall outside the admissible range for θ. The method of moments yields a less variable estimator which however presents a non-negligible bias. The one-step and two-step maximum likelihood methods produce estimators with a very similar distribution, with negligible bias and a variability a bit larger than that of the moment estimator.

6 Application to real data

In this section, we fit the bivariate discrete Weibull model to two real datasets taken from the literature and derive the parameter estimates through each of the methods described in Section 4. The computations have been carried out by developing appropriate code in the `R` environment and by using the available `DiscreteWeibull` package, and in particular the `estdweibull` function, with the option `method=''ML''` (for the two-step maximum likelihood method and the marginal maximum likelihood plus approximate method of moments) or `method=''P''` (for the proportion method). The two datasets are characterized by a positive (the first) and negative (the second) correlation coefficient between the two observed count variables.

6.1 Bivariate data with positive dependence

In this section, we apply the proposed bivariate discrete Weibull distribution to model a dataset studied in Kocherlakota and Kocherlakota [13], taken from Arbous and Kerrich [1] and reported in Table 10. The authors studied accidents among 122 experienced shunters. Here, x_1 represents the number of accidents in 6-year period from 1937-1942, and x_2 represents the number of accidents in 5-year period from 1943-1947.

Table 10: Dataset from Kocherlakota and Kocherlakota [13]

x_1, x_2	0	1	2	3	4	5	6	7	
0	21	13	4	2	0	0	0	0	40
1	18	14	5	1	0	0	0	1	39
2	8	10	4	3	1	0	0	0	26
3	2	1	2	2	1	0	0	0	8
4	1	4	1	0	0	0	0	0	6
5	0	1	0	1	0	0	0	0	2
6	0	0	0	1	0	0	0	0	1
	50	43	16	10	2	0	0	1	122

Summary statistics for the dataset are provided below:

$$\bar{x}_1 = 1.270 \quad \bar{x}_2 = 0.984 \quad \text{var}(x_1) = 1.640 \quad \text{var}(x_2) = 1.311 \quad \hat{\rho} = 0.283.$$

Note that both x_1 and x_2 present overdispersion (the variance is greater than the mean) and then Poisson margins would not fit the univariate distributions adequately. Note also that the Pearson's correlation is (moderately) positive.

If we fit the bivariate discrete Weibull model, the parameter estimates, derived through each of the methods presented in Section 4, are those displayed in Tables 11 and 12. Note that the four methods provide similar estimates for each marginal parameter; more importantly, the estimates of the two marginal β parameters are all greater than 1 (the two univariate distribution are estimated to present both an increasing failure rate). The estimates of the copula parameter θ derived through the full and two-step maximum likelihood methods are very close to +1; the estimates by the moment method and, more apparently, the method of proportion are smaller.

To test the null hypothesis H_0: "The data come from the bivariate discrete Weibull model" against the alternative H_1: "The data do not come from the bivariate discrete

Table 11: Parameter estimates for Arbous and Kerrich [1] data

method,estimate	q_1	β_1	q_2	β_2	θ
P	0.672	1.392	0.590	1.446	0.708
M	0.671	1.402	0.578	1.311	0.859
TSML	0.671	1.402	0.578	1.311	0.957
ML	0.678	1.414	0.585	1.319	0.961

Table 12: MLEs, standard errors and p-values for Arbous and Kerrich [1] data

Parameter	Estimate	Std. error	p-value
q_1	0.678	0.040	≈ 0
β_1	1.414	0.120	≈ 0
q_2	0.585	0.043	≈ 0
β_2	1.319	0.117	≈ 0
θ	0.961	0.277	0.0005

Weibull model", we employ the asymptotic chi-squared goodness-of-fit test, based on the statistic

$$\chi^2 = \sum_i \sum_j \frac{(n_{ij} - n_{ij}^*)^2}{n_{ij}^*}, \tag{17}$$

where n_{ij} is the observed frequency of $(i, j) \in \mathbb{N}_0^2$, n_{ij}^* its theoretical analogue. Under H_0, χ^2 is asymptotically distributed as a chi squared r.v. with $r - e - 1$, where r is the number of cells, e the number of estimated parameters.

After properly pooling the cells in the theoretical frequency table built by using the MLEs of the parameters (in order to get aggregated frequencies greater than 5; see Table 13), we compute the chi-square statistic, with $12 - 5 - 1 = 6$ degrees of freedom, and we get $\chi^2 = 5.49$, with a corresponding p-value equal to 0.483. Hence, it is reasonable to claim that the fit is good.

Additionally, we computed the value of the likelihood-ratio-test statistic, given by

$$\lambda = 2 \sum_i \sum_j n_{ij} \log \left(\frac{n_{ij}}{n \hat{p}_{ij}} \right) \tag{18}$$

which - based on the same cells pooling as before - equals 6.02, with a corresponding p-value of 0.421.

The χ^2 goodness-of-fit statistic for the bivariate geometric distribution in [14], although computed over a different cell grouping, returns a p-value of 0.08481, thus indicating a worse fit.

Table 13: Expected joint frequency distribution and cell aggregation for the computation of the chi-squared statistic for the first dataset.

x_1, x_2	0	1	2	3	4	5	6	7
0	22,53	11,42	3,70	1,14	0,36	0,11	0,03	0,01
1	16,65	12,64	6,20	2,55	0,93	0,31	0,09	0,04
2	7,20	8,25	4,97	2,22	0,84	0,28	0,09	0,03
3	2,73	4,20	2,77	1,28	0,48	0,16	0,05	0,02
4	1,00	1,81	1,24	0,58	0,22	0,07	0,02	0,01
5	0,35	0,69	0,48	0,22	0,09	0,03	0,01	0,00
6	0,17	0,34	0,24	0,11	0,04	0,01	0,00	0,00

Table 14: Bivariate distribution of the data taken from [23]: number of flight aborts by 109 aircrafts in two consecutive periods

$x_1 \setminus x_2$	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

Table 15: Parameter estimates for Mitchell and Paulson [23] data

method,estimate	q_1	β_1	q_2	β_2	θ
P	0.376	0.926	0.459	1.348	-0.442
M	0.379	0.977	0.450	1.120	-0.401
TSML	0.379	0.977	0.450	1.120	-0.635
ML	0.371	0.965	0.459	1.133	-0.655

6.2 Bivariate data with negative dependence

The data, considered in Mitchell and Paulson [23] and reported in Table 14, consist of the number of aborts by 109 aircrafts in two (first = x_1, second = x_2) consecutive 6 months of a 1-year period. Summary statistics for the dataset are provided below:

$$\bar{x}_1 = 0.624 \quad \bar{x}_2 = 0.725 \quad \text{var}(x_1) = 1.024 \quad \text{var}(x_2) = 1.062 \quad \hat{\rho} = -0.161$$

Note that both x_1 and x_2 present overdispersion and then, as for the dataset discussed in the previous section, Poisson margins would not fit the univariate distributions adequately. Note also that Pearson's correlation is (moderately) negative. Mitchell and Paulson [23] used a new bivariate negative binomial distribution derived by convoluting a bivariate geometric distribution to model the data.

If we fit the bivariate discrete Weibull model, we obtain the parameter estimates, derived through each of the methods presented in Section 4, displayed in Table 15. Note that all the β_1 estimates are smaller than 1, indicating a decreasing failure rate for x_1, and all the β_2 estimates are greater than 1, conversely indicating an increasing failure rate for x_2. The method of proportion provides estimates for the two β parameters that are slightly different from those derived through the other methods. On the contrary, the values of the q_1 estimates (and q_2 as well) are very close to each other. As to the estimate of the dependence parameter θ, the full maximum likelihood and two-step maximum likelihood methods return quite similar negative values (-0.655 and -0.635); the proportion method and the two-step method based on Spearman's correlation return two similar negative values (-0.442 and -0.401) which are considerably smaller – in absolute value – than the other two.

The maximum value of the loglikelihood function, computed at the MLEs, is $\ell_{\max} = -243.966$; the corresponding value of the Akaike Information Criterion is $AIC = 497.932$. Focusing on the (full) MLEs, we have that all the parameters are significant except for the

Table 16: MLEs, standard errors and p-values for Mitchell and Paulson [23] data

Parameter	Estimate	Std. error	p-value
q_1	0.371	0.046	≈ 0
β_1	0.965	0.118	≈ 0
q_2	0.459	0.047	≈ 0
β_2	1.133	0.121	≈ 0
θ	-0.655	0.405	0.106

copula parameter, whose p-value is 10.6% (see Table 16). We note that also in [9], where a bivariate generalized Poisson distribution was fitted to the same dataset, the association parameter of the model was the only one that turned out to be non-significant. In terms of AIC, the bivariate discrete Weibull model has a better fit than the bivariate Poisson, the bivariate negative binomial, and the bivariate generalized Poisson distributions [9, 34].

If we want to provide an absolute measure of goodness-of-fit, as we did in the previous section, we can easily derive the theoretical joint p.m.f. and the expected joint frequency table (based on the MLEs). To test the null hypothesis H_0: "The data come from the bivariate discrete Weibull model" against the alternative H_1: "The data do not come from the bivariate discrete Weibull model", we employ again the asymptotic chi-squared goodness-of-fit test, based on the statistic of Eq. (17). Since $n = 109$, this statistic can be usefully employed for assessing whether the bivariate discrete Weibull model fits the data adequately. If we proceed and collapse adjacent cells, ensuring that every expected frequency is greater than 5 (in order to apply the asymptotic results satisfactorily), we can choose the aggregations of Table 17 and the chi-squared statistic takes the value 7.13, whose corresponding p-value (the degrees of freedom are $9 - 5 - 1 = 3$) is 0.068, which means that at the 5% level we accept the null hypothesis that the data come from a bivariate discrete Weibull model. The value of the likelihood-ratio-test statistic Eq. (18), computed on the same cell grouping, is 7.32, with a corresponding p-value of 0.062. Again, at a 5% significance level, the hypothesis H_0 would be accepted.

Table 17: Expected joint frequency distribution and cell aggregation for the computation of the chi-squared statistic for the second dataset.

$x_1 \setminus x_2$	0	1	2	3	≥ 4
0	32.97	20.71	9.26	3.62	2.00
1	15.33	6.08	2.15	0.76	0.40
2	6.38	2.14	0.66	0.22	0.11
3	2.58	0.81	0.23	0.07	0.04
4	1.04	0.32	0.09	0.03	0.01
≥ 5	0.70	0.21	0.06	0.02	0.01

7 Conclusions

A bivariate count model, with discrete Weibull margins and the FGM copula has been proposed. It allows for moderate levels of positive and negative linear correlation, whose relationship with the dependence parameter θ and the marginal parameters has been examined. Its simulation is straightforward and easily implemented in the R statistical environment. Sample estimation of its parameters can be performed via standard methods or ad hoc procedures (moment method and method of proportion) which are less efficient but lead to analytical expressions for $\hat{\theta}$. The relative efficiency of the proposed estimators are also investigated through a Monte Carlo study, which reveals potential practical pitfalls of each procedure. Applications to real datasets have shown that this new distribution can fit count data even better than existing models (e.g., bivariate geometric, negative binomial or Poisson models). Further research will address generalization of this distribution taking into account extensions of the FGM copula able to enlarge the range of attainable values for Pearson's correlation.

References

[1] A.G. Arbous and J.E. Kerrich, Accident statistics and the concept of accident proneness, Biometrics 7 (1951), 340-432

[2] A. Barbiero, DiscreteWeibull: Discrete Weibull Distributions (Type 1 and 3). R package version 1.1. `http://CRAN.R-project.org/package=DiscreteWeibull` (2015).

[3] A. Barbiero, Simulating correlated discrete Weibull variables: a proposal and an implementation in the `R` environment. AIP Conference Proceedings 1702, 190017 (2015); doi: 10.1063/1.4938984

[4] A. Barbiero, A comparison of methods for estimating parameters of the type I discrete Weibull distribution, Stat Interface 9(2) (2016), 203-212.

[5] A. Barbiero, Discrete Weibull variables linked by Farlie-Gumbel-Morgenstern copula. AIP Conference Proceedings 1863, 240006 (2017); doi: 10.1063/1.4992408

[6] Cambanis, S. (1977). Some properties and generalizations of multivariate Eyraud-Gumbel-Morgenstern distributions. J Multivariate Anal 7(4), 551-559.

[7] J.D. Englehardt and R.C. Li, The discrete Weibull distribution: An alternative for correlated counts with confirmation for microbial counts in water, Risk Anal 31 (2011), 370–381.

[8] J.D. Englehardt, Distributions of Autocorrelated First-Order Kinetic Outcomes: Illness Severity, PLoS ONE 10(6). doi:10.1371/journal.pone.0129042 (2015).

[9] F. Famoye, A new bivariate generalized Poisson distribution, Statistica Neerlandica 64(1) (2011) 112-124.

[10] D.J.G. Farlie, The performance of some correlation coefficient for a general bivariate distribution, Biometrika **47** (1960), 307–323.

[11] C. Genest and J. Nešlehová, A primer on copulas for count data, Astin Bulletin 37(02) (2007), 475-515.

[12] M.S.A. Khan, A. Khalique, and A.M. Abouammoh, On estimating parameters in a discrete Weibull distribution, IEEE Trans Reliability 38(3) (1989), 348-350.

[13] S. Kocherlakota and K. Kocherlakota, Bivariate Discrete Distributions, Marcel Dekker Inc., New York, 1992.

[14] H. Krishna and P.S. Pundir, A Bivariate Geometric Distribution with Applications to Reliability, Comm Stat - Theory Methods 38(7) (2009), 1079-1093.

[15] K.B. Kulasekera, Approximate MLE's of the parameters of a discrete Weibull distribution with type I censored data, Microel Reliability 34(7) (1984), 1185–1188.

[16] X. Jiang, J. Chu, and S. Nadarajah, New classes of discrete bivariate distributions with application to football data, Comm Stat - Theory Methods 46(16) (2017), 8069-8085

[17] H. Joe, Asymptotic efficiency of the two-stage estimation method for copula-based models, J Multivar Anal 94(2) (2005) 401–419.

[18] M.E. Johnson (1987). Multivariate statistical simulation. John Wiley and Sons, New York, 1987.

[19] N.L. Johnson, S. Kotz, and N. Balakrishnan, Discrete multivariate distributions. John Wiley and Sons, New York, 1997.

[20] H. Joe, Dependence Modelling with Copulas, CRC Press, Boca Raton, 2015.

[21] C.D. Lai (2006) Constructions of discrete bivariate distributions. In: Balakrishnan N, Castillo E, Sarabia Alegria JM (eds) Advances in Distribution Theory, Order Statistics, and Inference, pp 29-58.

[22] H. Lee, J.H. Cha, and G. Pulcini "Modeling Discrete Bivariate Data with Applications to Failure and Count Data", Qual Reliab Eng Int 33(7) (2017), 1455–1473

[23] C.R. Mitchell and A.S. Paulson, A new bivariate negative binomial distribution, Naval Res Log Quarterly 28 (1981), 359-374 .

[24] T. Nakagawa and S. Osaki, The discrete Weibull distribution, IEEE Trans Reliability 24(5) (1975), 300–301.

[25] R.B. Nelsen, An Introduction to Copulas, Springer, New York, 1999.

[26] A. K. Nikoloulopoulos, D. Karlis, Modeling multivariate count data using copulas, Comm Stat - Simulation Computation 39(1) (2010), 172–187.

[27] A.K. Nikoloulopoulos, "Copula-Based Models for Multivariate Discrete Response Data," Copulae in Mathematical and Quantitative Finance, Lecture Notes in Statistics, in P. Jaworski et al. (eds.), Springer, New York, 2013, pp. 231–249.

[28] W.J. Padgett and J.D. Spurrier, Discrete failure models, IEEE Trans Reliability 34(3) (1985), 253–256.

[29] V.E. Piperigou, Discrete distributions in the extended FGM family: The pgf approach, J Stat Plan Inference 139(11) (2009), 3891-3899.

[30] R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, http://www.R-project.org/ (2016)

[31] J.M. Sarabia and E. Gòmez-Déniz, Construction of multivariate distributions: A review of some recent results, SORT 32(1) (2008), 3–36.

[32] W. Schucany, W.C. Parr, and J.E. Boyer, Correlation structure in Farlie-Gumbel-Morgenstern distributions, Biometrika 65(3) (1978), 650–653.

[33] W.E. Stein and R. Dattero, A new discrete Weibull distribution, IEEE Trans Reliability 33 (1984) 196–197.

[34] H. Zamani, P. Faroughi, and N. Ismail, Bivariate Poisson-Lindley distribution with application, J Math Stat, 11(1) (2015), 1–6.