Methodol Comput Appl Probab (2026) 28:41
https://doi.org/10.1007/s11009-026-10263-0

RESEARCH

New Discrete Bivariate Distributions Generated from Discrete Time Point Process Model

Hyunju Lee[1] · **Ji Hwan Cha**[2]

Received: 8 February 2026 / Revised: 22 March 2026 / Accepted: 26 March 2026 /
Published online: 25 April 2026
© The Author(s) 2026

Abstract
In this paper, we develop a new class of discrete bivariate distributions based on a discrete time nonhomogeneous Poisson process shock model. The proposed class of distributions is general and flexible in the sense that numerous parametric families of discrete bivariate distributions can be generated from it. Furthermore, the flexibility of the developed model also exists in modeling the proportion of tied observations. We discuss the properties of the developed class of bivariate distributions including the dependence structure and the ageing properties. Furthermore, specific discrete bivariate distributions are generated and they are applied to analyze bivariate reliability failure data and bivariate environmental ordinal data to demonstrate the utility and the flexibility of it.

Keywords Discrete bivariate distribution · Discrete time shock model · Discrete time nonhomogeneous Poisson process · Dependence structure · Bivariate lack of memory property

1 Introduction

Traditionally, the concepts and probabilistic models in reliability and survival analysis have substantially contributed to the development of bivariate or multivariate distributions and they have taken crucial roles in generating new statistical models. Especially, various shock models have been developed in the literature (Ozkut (2023), Goyal et al. (2022, 2023)). Traditionally, shock models have contributed to the development of important statistical distributions. For example, Marshall and Olkin (1967) obtained a class of multivariate exponential distributions based on a fatal shock model. After the work of Marshall and Olkin

✉ Ji Hwan Cha
jhcha@ewha.ac.kr

Hyunju Lee
hyunjlee@hufs.ac.kr

[1] Department of Statistics, Hankuk University of Foreign Studies, Yongin 17035, Republic of Korea

[2] Department of Statistics, Ewha Womans University, Seoul 120-750, Republic of Korea

(1967), other shock models also contributed to the development of various bivariate or multivariate distributions. For example, Lu (1989) suggested the Weibull extension of the Marshall and Olkin model. See also Arnold (1975), Raftery (1984), O'cinneide and Raftery (1989) and Wang and Li (2005) for some other applications of shock models in the development of bivariate or multivariate distributions. Recently, Lee and Cha (2024) has developed a new general class of bivariate distributions based on a Poisson shock model.

Note that, in most of the foregoing works, dependent distributions have been extensively studied in the continuous time scale assuming that the external shock process follows the homogeneous Poisson process (HPP) or nonhomogeneous Poisson process (NHPP). However, in practice, discrete type of data can also be frequently encountered. For example, in the reliability and survival analysis, field failures are often collected and reported daily or weekly. Systems often operate in cycles, and the experimenter observes the number of cycles completed before failure. In the literature, several discrete bivariate models have been proposed and studied (see, e.g., Li et al. (2024) and Bermúdez and Karlis (2022), and see also Johnson et al. (1997) for an encyclopedic review). In this paper, we develop a new general class of discrete bivariate distributions based on a new shock model in the discrete time (DT) scale. In the literature, some DT shock models have been developed and studied (see, e.g., Eryilmaz (2013), Eryilmaz and Kan (2021)). In this paper, the DT shock model employed for the modeling is defined based in the DT NHPP, recently developed in Cha and Limnos (2021) based on the concept of DT stochastic intensity. The discrete bivariate distribution model developed in this paper can be, in a sense, regarded as a discrete analogue of the model developed in Lee and Cha (2024). However, the theoretical foundations needed for the development are completely different, as deep understanding of new advanced concepts and properties recently developed for DT scale point process theory are prerequisite for the modeling.

The discrete bivariate distribution model is developed by modeling the discrete lifetimes of two components in a system operated in DT scale. We assume that the external shocks are the only cause of component failures and there are two types of shock processes: the first type of shock process causes sequential failures of the components, whereas the second type of shock process causes the simultaneous failures of the components. The shock processes are assumed to follow the DT NHPPs, respectively. Under these assumptions, we construct a class of discrete bivariate distributions. The developed class of discrete bivariate distributions is very general and flexible in the sense that, by specifying the parameters of the external shock processes, numerous discrete bivariate distributions can be generated. Furthermore, as there are two independent shock processes which control and govern simultaneous and non-simultaneous failures, respectively, the flexibility of the developed model also exists in modeling the proportion of tied observations. This will be illustrated in the real data analysis part, where, in the first data set, the tied observations are very sparse, whereas, in the second data set, the proportion of tied observation is very high.

This paper is organized as follows. In Sect. 2, some preliminary concepts related to the discrete random variable are introduced. Furthermore, the formal definition of the DT NHPP is suggested. In Sect. 3, a general class of discrete bivariate distributions is developed based on the shock model. The joint probability mass function, the joint survival function and the marginal distributions for the class are derived. In Sect. 4, the dependence structure and the ageing properties of the class are discussed. In Sect. 5, by specifying the parameters of the external shock processes, some useful specific parametric families of distributions are gen-

erated. In Sect. 6, the parametric families of distributions generated in Sect. 5 are applied to two real data sets from reliability and environmental sciences to demonstrate the utility and the flexibility of the proposed model.

2 Discrete Time Nonhomogeneous Poisson Process

In this section, we briefly review the concept and properties of the DT NHPP (Cha and Limnios 2021). In reliability area, a repair type defines the corresponding point process (see Cha 2014). For instance, the perfect repair process defines the renewal process and the minimal repair process defines the NHPP. Due to this sense, a new repair type, which is called the DT minimal repair, can define the DT NHPP. For this, let T be the discrete lifetime of the item which can assume non-negative integers 0,1,2,.... Define $S(n)$ and $f(n)$, $n = 0,1,2,\ldots,$ as the survival function (sf) and the probability mass function (pmf) of T, respectively. Then the DT failure rate, which is denoted by $r(n)$, is defined by (see, e.g., Salvia and Bollinger (1982), Limnios and Oprisan (2001) and Barbu and Limnios (2008)).

$$r(n) = \frac{f(n)}{S(n-1)} = P(T = n | T > n-1), n = 0,1,2,\ldots \tag{1}$$

where $S(-1) \equiv 1$ because we are assuming that the discrete random variables can assume the value '0' and hence $S(0) < 1$ when $f(0) > 0$. From (1), it is straightforward to obtain that

$$S(n) = P(T > n) = \prod_{i=0}^{n-1} (1 - r(i))$$

$$f(n) \equiv P(T = n) = \prod_{i=0}^{n-1} (1 - r(i)) r(n), n = 0,1,2,\ldots \tag{2}$$

Analogously to the continuous time minimal repair, the discrete time minimal repair is defined as "the repair which does not change the failure rate of the item". This means that if an item has failed and has been instantaneously minimally repaired at time n, its failure rate is the same as that just prior to the failure and should be given by $r(n)$. Based on this type of DT minimal repair, in Cha and Limnios (2021), the corresponding process of failure occurrences is defined as DT NHPP. For the formal definition, a new concept of discrete time stochastic intensity needs to be defined in a similar manner to the continuous time scale. In Cha and Limnios (2021), in developing the DT NHPP, it was assumed that an event cannot occur at time $n = 0$, and thus $N(0) = 0$. However, in this paper, we are modelling the joint distribution of two dependent discrete random variables which can assume "0". Thus, in defining the DT NHPP, some necessary modifications should be made and we assume that an event can also occur at time $n = 0$, assuming that $N(-1) = 0$.

In the DT point process, it is said that a point process is orderly (or simple) if $\{N(n), n = 0,1,2,\ldots\}$ satisfies $P(N(n) - N(n-1) > 1) = 0$, for all $n = 0,1,2,....$ We also define $\mathrm{H}_n \equiv \{N(k), 0 \le k \le n\}$, $n = 0,1,2,...,$ as the history (natural filtration) of the process.

Definition 1. (Stochastic Intensity in Discrete Time) For a DT orderly (simple) point process $N \equiv \{N(n), n = 0, 1, 2, \dots\}$, the stochastic intensity r_n is defined as

$$r_n \equiv P(N(n) - N(n-1) = 1 | \, \mathrm{H}_{n-1}) = E[N(n) - N(n-1) \mid \mathrm{H}_{n-1}], n \geq 0$$

where, when $n = 0$, $\mathrm{H}_{n-1} \equiv \{N(-1) = 0\}$.

Based on the "minimal repair process reasoning", the stochastic intensity for the discrete time minimal repair process should be given by

$$r_n \equiv P(N(n) - N(n-1) = 1|\mathrm{H}_{n-1}) = E[N(n) - N(n-1)|\mathrm{H}_{n-1}] = r(n), n = 0,1,2,...,$$

which is deterministic and does not depend on H_{n-1}. Accordingly, this yields the following definition of DT NHPP in Cha and Limnios (2021).

Definition 2. (DT NHPP, Cha and Limnios (2021)) An orderly (simple) counting process $\{N(n), n = 0, 1, 2, \dots\}$ is called the DT NHPP with intensity function $r(n)$ if

$$r_n = P(N(n) - N(n-1) = 1|\mathrm{H}_{n-1}) = r(n), n = 0,1,2,\dots.$$

From Definition 2, we can see that the DT NHPP possesses the independent increments property. If the stationary increment property is additionally assumed, then it is said to be the DT homogeneous Poisson process (DP HPP) with intensity function $r(n) = r$.

We now provide some basic properties of the DT NHPP. The following proposition presents the distribution of number of events in an arbitrary interval and, its expectation and variance. For the description, we define a set of combinations, $C(m, k)$, as follows:

$C(m, k)$, $1 \leq k \leq m$, is the set of all combinations of size k from m elements $\{1, 2, \dots, m\}$

where $C(m, 0) \equiv \varnothing$. For example, $C(3, 2) = \{\{1, 2\}, \{1, 3\}, \{2, 3\}\}$.

Proposition 1. (Cha and Limnios (2021)) The distribution of $N(n_2) - N(n_1), n_2 > n_1$, and its expectation and variance are given by.

$$P(N(n_2) - N(n_1) = k)$$

$$= \sum_{\forall \{i_1, i_2, \dots, i_k\} \in C(n_2 - n_1, k)} \left[\prod_{i \in \{i_1, i_2, \dots, i_k\}} r(n_1 + i) \cdot \prod_{i \in \{(1,2,\dots,n_2 - n_1) - \{i_1, i_2, \dots, i_k\}\}} \overline{r}(n_1 + i) \right], \; for \; k = 0, 1, 2, \dots, n_2 - n_1$$

$$E[N(n_2) - N(n_1)] = \sum_{n=n_1+1}^{n_2} r(n)$$

$$Var[N(n_2) - N(n_1)] = \sum_{n=n_1+1}^{n_2} r(n)\overline{r}(n)$$

Respectively, where $\overline{r}(n) = 1 - r(n), n = 0,1,2,\dots$

As a special case, when $r(n) = r$ (constant), for all n, and when we assume that an event can also occur at time $n = 0$, we have

$$P(N(n) = k) = P(N(n) - N(-1) = k) = \binom{n+1}{k} r^k (1-r)^{n+1-k}, k = 0, 1, 2, ..., n+1$$

3 Construction of Discrete Bivariate Distribution

We assume that a system consists of two components and they are subject to two external DT NHPP shock processes $\{N(n), n = 0, 1, 2, \dots\}$ and $\{M(n), n = 0, 1, 2, \dots\}$. Under the DT NHPP $\{N(n), n = 0, 1, 2, \dots\}$ (with intensity function $r(n)$) of external shocks, it is assumed that the first shock destroys the first component with probability θ and it destroys the second component with probability $(1-\theta)$; and the second shock destroys the remaining component. In addition, under the DT-NHPP $\{M(n), n = 0, 1, 2, \dots\}$ (with intensity function $v(n)$) of external shocks, a shock destroys the two components simultaneously. We further assume that $\{N(n), n = 0, 1, 2, \dots\}$ and $\{M(n), n = 0, 1, 2, \dots\}$ are independent. Under this model, let us now denote by X_i the lifetime of component i, $i = 1,2$, and by T_i the time from 0 to the occurrence of i th shock, $i = 1,2$, in the DT NHPP $\{N(n), n = 0, 1, 2, \dots\}$. We first derive the joint probability mass function of (X_1, X_2). For the sake of convenience, throughout this paper, we define $I = 1$ if the first shock destroys the first component and $I = 0$ otherwise under the shock process $\{N(n), n = 0, 1, 2, \dots\}$.

Theorem 1. The joint probability mass function of (X_1, X_2) is given by.

$$P(X_1 = n_1, X_2 = n_2) = \begin{cases} \theta \frac{f(n_1)}{S(n_1)} \{f(n_2) S^*(n_2 - 1) + S(n_2) f^*(n_2)\}, n_1 < n_2 \\ (1-\theta) \frac{f(n_2)}{S(n_2)} \{f(n_1) S^*(n_1 - 1) + S(n_1) f^*(n_1)\}, n_1 > n_2 \\ f^*(n) S(n-1), n_1 = n_2 = n \end{cases}$$

where $f(n)(f^*(n))$ and $S(n)(S^*(n))$ are the pmf and sf constructed by Eq. (2) using the intensity functions $r(n)(v(n))$ as the corresponding failure rates, respectively.

Proof. Define S_1 as the time from 0 to the occurrence of the first shock in the DT NHPP $\{M(n), n = 0,1, 2, ...\}$. For the case $n_1 < n_2$,

$$P(X_1 = n_1, X_2 = n_2) = P(I = 1, T_1 = n_1, T_2 = n_2, M(n_2 - 1) = 0)$$

$$+P(I = 1, T_1 = n_1, T_2 > n_2, S_1 = n_2)$$

$$= \theta \prod_{i=0}^{n_1-1} (1 - r(i)) r(n_1) \cdot \prod_{i=n_1+1}^{n_2-1} (1 - r(i)) r(n_2) \cdot \prod_{i=0}^{n_2-1} (1 - v(i))$$

$$+\theta \prod_{i=0}^{n_1-1} (1 - r(i)) r(n_1) \cdot \prod_{i=n_1+1}^{n_2} (1 - r(i)) \cdot \prod_{i=0}^{n_2-1} (1 - v(i)) v(n_2)$$

$$= \theta \left\{ \frac{f(n_1) f(n_2) S^*(n_2 - 1)}{S(n_1)} + \frac{f(n_1) S(n_2) f^*(n_2)}{S(n_1)} \right\}, n_1 < n_2$$

Symmetrically, in the case of $I = 0$, the joint probability mass function of (X_1, X_2) is given by

$$P\left(X_1 = n_1, X_2 = n_2\right)$$

$$= (1-\theta)\left\{\frac{f\left(n_1\right) f\left(n_2\right) S^*\left(n_1 - 1\right)}{S\left(n_2\right)} + \frac{f(n_2)S(n_1)f^*(n_1)}{S(n_2)}\right\}, n_1 > n_2$$

For the case of $n_1 = n_2 = n$, note that the only way to have a common failures at time n is that there should be the first shock from $\{M\left(n\right), n = 0, 1, \dots\}$ at time n and there should be no shock from $\{N\left(n\right), n = 0, 1, \dots\}$ until time $n-1$. Therefore, the joint probability mass function is given by

$$P(X_1 = n, X_2 = n)$$

$$= \prod_{i=0}^{n-1}(1 - v(i))v(n) \cdot \prod_{i=0}^{n-1}(1 - r(i)) = f^*(n)S(n-1)$$

Finally, combining all the cases, we have that

$$P(X_1 = n_1, X_2 = n_2) = \begin{cases} \theta\left\{\frac{f(n_1)f(n_2)S^*(n_2-1)}{S(n_1)} + \frac{f(n_1)S(n_2)f^*(n_2)}{S(n_1)}\right\}, n_1 < n_2 \\ (1-\theta)\left\{\frac{f(n_1)f(n_2)S^*(n_1-1)}{S(n_2)} + \frac{f(n_2)S(n_1)f^*(n_1)}{S(n_2)}\right\}, n_1 > n_2 \\ f^*(n)S(n-1), n_1 = n_2 = n \end{cases}$$

It is well known that the discrete failure rate is bounded; that is, $r\left(n\right) \leq 1$, which in turn, the discrete failure rate does not satisfy some properties of the failure rate in continuous case. This can be regarded as a critical restriction, especially from a modeling point of view. To overcome such restriction, the alternative discrete failure rate has been introduced in the literature (see Roy and Gupta (1992); Dewan and Sudheesh (2009)). The alternative discrete failure rate $r^*\left(n\right)$ is defined as follows:

$$r^*\left(n\right) = -ln\left(\frac{S\left(n\right)}{S\left(n-1\right)}\right), n = 0,1,2,\dots.$$

Note that $r\left(n\right)$ and $r^*\left(n\right)$ are related by the following relationship:

$$r^*(n) = -ln[1 - r(n)] \; or \; r(n) = 1 - exp(-r^*(n))$$

Thus, the sf $S\left(n\right)$ can be also expressed in terms of $r^*\left(n\right)$ as follows:

$$S\left(n\right) = exp\left\{-\sum_{i=0}^{n} r^*\left(i\right)\right\}, n = 0,1,2,\dots$$

Observe that $0 < r^*(n) < \infty$ and $r^*(n)$ does not have the restriction that $r(n)$ has. Note that the alternative discrete failure rate does not have a clear meaning of probability, but it can be regarded as a suitable transformation of $r(n)$. In addition, we can see that $r(n)$ and $r^*(n)$ have the same monotonicity properties.

The following result is obtained by replacing $r(n)$ by $r^*(n)$ in Theorem 1. This will play a crucial role in our further discussions in generating various specific families of distributions (see Sect. 5).

Corollary 1. From Theorem 1, the joint probability mass function of (X_1, X_2) can be expressed by.

$$P(X_1 = n_1, X_2 = n_2) = \theta[exp\{r^*(n_1)\} - 1][exp\{r^*(n_2) + v^*(n_2)\} - 1]$$

$$\times exp\left\{-\left(\sum_{i=0}^{n_2} r^*(i) + v^*(i)\right)\right\}, n_1 < n_2$$

$$P(X_1 = n_1, X_2 = n_2) = (1-\theta)[exp\{r^*(n_2)\} - 1][exp\{r^*(n_1) + v^*(n_1)\} - 1]$$

$$\times exp\left\{-\left(\sum_{i=0}^{n_1} r^*(i) + v^*(i)\right)\right\}, n_1 > n_2$$

$$P(X_1 = n_1, X_2 = n_2) = [1 - exp\{-v^*(n)\}]exp\left\{-\left(\sum_{i=0}^{n-1} r^*(i) + v^*(i)\right)\right\}, n_1 = n_2 = n$$

where $r^*(n)$ and $v^*(n)$ are the corresponding alternative discrete failure rates of $r(n)$ and $v(n)$, respectively.

Now we derive the bivariate joint survival function of (X_1, X_2).

Theorem 2. The bivariate joint survival function of (X_1, X_2) is given by

$$P(X_1 > n_1, X_2 > n_2)$$

$$= \begin{cases} \prod_{i=0}^{n_2}(1-v(i))\prod_{i=0}^{n_2}(1-r(i))\left(1+\theta\sum_{k=n_1+1}^{n_2}\frac{r(k)}{1-r(k)}\right), n_1 < n_2 \\ \prod_{i=0}^{n_1}(1-v(i))\prod_{i=0}^{n_1}(1-r(i))\left(1+(1-\theta)\sum_{k=n_2+1}^{n_1}\frac{r(k)}{1-r(k)}\right), n_1 > n_2 \\ \prod_{i=0}^{n}(1-v(i))\prod_{i=0}^{n}(1-r(i)), n_1 = n_2 = n \end{cases}$$

Proof. As before, define S_1 as the time from 0 to the occurrence of the first shock in the DT NHPP $\{M(n), n = 0,1,2,...\}$. For $n_1 < n_2$, it follows that.

$$P(X_1 > n_1, X_2 > n_2)$$

$$= \theta P(T_1 > n_1, T_2 > n_2, S_1 > n_2) + (1-\theta)P(T_2 > n_1, T_1 > n_2, S_1 > n_2)$$

$$= \theta P(T_1 > n_1, T_2 > n_2, S_1 > n_2) + (1-\theta)P(T_1 > n_2, S_1 > n_2)$$

$$= \theta P(N(n_1) = 0, N(n_2) \le 1, M(n_2) = 0) + (1-\theta)P(N(n_2) = 0, M(n_2) = 0)$$

$$= \theta P(N(n_2) = 0)P(M(n_2) = 0) + \theta P(N(n_1) = 0, N(n_2) - N(n_1) = 1)P(M(n_2) = 0) + (1-\theta)P(N(n_2) = 0)P(M(n_2) = 0)$$

$$= P\left(N\left(n_2\right)=0\right) P\left(M\left(n_2\right)=0\right)+\theta P\left(N\left(n_1\right)=0, N\left(n_2\right)-N\left(n_1\right)=1\right) P\left(M\left(n_2\right)=0\right)$$

$$= \left[\prod_{i=0}^{n_2}(1-r(i))+\theta \prod_{i=0}^{n_1}(1-r(i))\left\{\sum_{k=1}^{n_2-n_1} r(n_1+k) \cdot \prod_{j \in\{1,2, \ldots, n_2-n_1\}-\{k\}}(1-r(n_1+j))\right\}\right] \cdot \prod_{i=0}^{n_2}(1-v(i))$$

$$= \left[\prod_{i=0}^{n_2}(1-r(i))+\theta \prod_{i=0}^{n_1}(1-r(i))\left\{\prod_{i=1}^{n_2-n_1}(1-r(n_1+i))\right\}\left\{\sum_{k=1}^{n_2-n_1} \frac{r(n_1+k)}{1-r(n_1+k)}\right\}\right] \cdot \prod_{i=0}^{n_2}(1-v(i))$$

$$= \left[\prod_{i=0}^{n_2}(1-r(i))+\theta \prod_{i=0}^{n_1}(1-r(i)) \prod_{i=n_1+1}^{n_2}(1-r(i)) \sum_{k=n_1+1}^{n_2} \frac{r(k)}{1-r(k)}\right] \cdot \prod_{i=0}^{n_2}(1-v(i))$$

$$= \prod_{i=0}^{n_2}(1-v(i)) \prod_{i=0}^{n_2}(1-r(i))\left(1+\theta \sum_{k=n_1+1}^{n_2} \frac{r(k)}{1-r(k)}\right).$$

For $n_1 > n_2$, we have that

$$P\left(X_1>n_1, X_2>n_2\right)$$

$$= \theta P\left(T_1>n_1, T_2>n_2, S_1>n_1\right)+(1-\theta) P\left(T_2>n_1, T_1>n_2, S_1>n_1\right)$$

$$= \theta P\left(T_1>n_1, S_1>n_1\right)+(1-\theta) P\left(T_2>n_1, T_1>n_2, S_1>n_1\right)$$

$$= \theta P\left(N\left(n_1\right)=0, S_1>n_1\right)+(1-\theta) P\left(N\left(n_2\right)=0, N\left(n_1\right) \leq 1, S_1>n_1\right)$$

$$= \theta P\left(N\left(n_1\right)=0\right) P\left(S_1>n_1\right)+(1-\theta) P\left(N\left(n_1\right)=0\right) P\left(S_1>n_1\right) + (1-\theta) P\left(N\left(n_2\right)=0, N\left(n_1\right)-N\left(n_2\right)=1\right) P\left(S_1>n_1\right)$$

$$= \prod_{i=0}^{n_1}(1-v(i)) \prod_{i=0}^{n_1}(1-r(i))+(1-\theta) \prod_{i=0}^{n_1}(1-v(i))\left[\prod_{i=0}^{n_2}(1-r(i))\right] \cdot\left[\sum_{k=1}^{n_1-n_2} r(n_2+k)\left\{\prod_{j \in\{1,2, \ldots, n_1-n_2\}-\{k\}}(1-r(n_2+j))\right\}\right]$$

$$= \prod_{i=0}^{n_1}(1-v(i)) \prod_{i=0}^{n_1}(1-r(i))\left(1+(1-\theta) \sum_{k=n_2+1}^{n_1} \frac{r(k)}{1-r(k)}\right).$$

For $n_1 = n_2 = n$, we obviously have

$$P\left(X_1>n, X_2>n\right)=P\left(N\left(n\right)=0, M\left(n\right)=0\right)=\prod_{i=0}^{n}\left(1-v\left(i\right)\right) \prod_{i=0}^{n}\left(1-r\left(i\right)\right)$$

From Theorem 2, the marginal survival functions of X_1 and X_2 can be obtained as follows.

Corollary 2. The marginal survival functions of X_1 and X_2 are given by.

$$P\left(X_1 \ > \ n_1\right) \ = \ \prod_{i=0}^{n_1}\left(1-v\left(i\right)\right) \ \prod_{i=0}^{n_1}\left(1-r\left(i\right)\right) \ \left(1+\ (1-\theta) \ \sum_{k=0}^{n_2} \ \frac{r\left(k\right)}{1-r\left(k\right)}\right), \ n_1=0,1,2, \ldots .$$

$$P\left(X_2 > n_2\right) = \prod_{i=0}^{n_2}\left(1-v\left(i\right)\right)\prod_{i=0}^{n_2}\left(1-r\left(i\right)\right)\left(1+\theta\sum_{k=0}^{n_2}\frac{r\left(k\right)}{1-r\left(k\right)}\right),\ n_2 = 0,1,2,....$$

4 Statistical Properties of the Proposed Class

In this section, we now discuss some statistical properties of the developed class of the discrete bivariate distributions. To this end, we first introduce a concept of conditional dependence, conditional positive quadrant dependence, which was originally defined by Brady and Singpurwalla (1990).

Definition 3. The random variables if $(Z_1,\ Z_2)$ are said to be conditionally positive quadrant dependent with respect to random variable W, if

$$P\left(Z_1 >_{Z_1},\ Z_2 >_{Z_2} |W=w\right) \geq P\left(Z_1 >_{Z_1} |W=w\right) P\left(Z_2 >_{Z_2} |W=w\right),$$

for all $z_1, z_2,\ w$.

Theorem 3. The random variables (X_1, X_2) are conditionally positive quadrant dependent with respect to random variable I, that is,

$$P\left(X_1 > n_1, X_2 > n_2 | I=i\right) \geq P\left(X_1 > n_1 | I=i\right) P\left(X_2 > n_2 | I=i\right)$$

for all n_1, n_2, and $i = 0,1$.

Proof.

For $n_1 \leq n_2$,

$$P\left(X_1 > n_1,\ X_2 > n_2 \mid I=1\right)$$

$$= P\left(T_1 > n_1,\ T_2 > n_2\ ,\ S_1 > n_2\right)$$

$$= P\left(N\left(n_1\right) = 0,\ N\left(n_2\right) \leq 1,\ M\left(n_2\right)=0\right)$$

$$= P\left(N\left(n_1\right) = 0,\ N\left(n_2\right) - N\left(n_1\right) \leq 1,\ M\left(n_2\right)=0\ \right)$$

$$= P\left(N\left(n_1\right) = 0\right) P\left(N\left(n_2\right) - N\left(n_1\right) \leq 1\right) P\left(M\left(n_2\right) = 0\right)$$

On the other hand,

$$P\left(X_1 > n_1 | I=1\right) = P\left(T_1 > n_1, S_1 > n_1\right) = P\left(N\left(n_1\right)=0\right) P\left(M\left(n_1\right)=0\right)$$

$$P\left(X_2 > n_2 | I=1\right) = P\left(T_2 > n_2, S_1 > n_2\right) = P\left(N\left(n_2\right) \leq 1\right) P\left(M\left(n_2\right)=0\right)$$

As, $N\left(n_2\right) = \left[N\left(n_2\right) - N\left(n_1\right)\right] + N\left(n_1\right)$, we have $\left[N\left(n_2\right) - N\left(n_1\right)\right] \leq_{st} N\left(n_2\right)$, where $''\leq_{st}''$ stands for the usual stochastic order (see Shaked and Shanthikumar (2007), Navarro (2016, 2018), Badía and Sangüesa (2014)), which entails

$$P\left(N\ \left(n_2\right)\ -\ N\ \left(n_1\right)\ \leq 1\right)\ \geq P\left(N\left(n_2\right)\ \leq 1\right).$$

$P\left(X_1\ >\ n_1,\ X_2\ >n_2\ \mid I=1\right)$
$=P\left(N\left(n_1\right)\ =0\right)P\left(N\left(n_2\right)\ -N\left(n_1\right)\ \leq 1\right)P\left(M\left(n_2\right)\ =0\right)$
$\geq P\left(N\left(n_1\right)\ =0\right)P\left(N\left(n_2\right)\ \leq 1\right)P\left(M\left(n_2\right)\ =0\right)$
$\geq P\left(N\left(n_1\right)\ =0\right)P\left(M\left(n_1\right)\ =0\right)P\left(N\left(n_2\right)\ \leq 1\right)P\left(M\left(n_2\right)\ =0\right)=P\left(X_1>n_1|I=1\right)P\left(X_2>n_2|I=1\right)$
Thus, it follows that
For $n_1 > n_2$, obviously, it holds that

$$P\left(X_1>n_1,X_2>n_2|I=1\right)$$

$$=P\left(T_1>n_1,T_2>n_2,S_1>n_1\right)$$

$$=P\left(T_1>n_1\right)P\left(S_1>n_1\right)$$

$$\geq P\left(T_1>n_1\right)P\left(S_1>n_1\right)P\left(T_2>n_2\right)P\left(S_1>n_2\right)=P\left(X_1>n_1|I=1\right)P\left(X_2>n_2|I=1\right)$$

The other case with $I = 0$ can be shown symmetrically.

Now, we will show that a well-defined subclass of the proposed class satisfies the discrete-type bivariate lack of memory property (BLMP). In the case of continuous random variables (X, Y), the joint distribution of (X, Y) is said to have the bivariate lack of memory property (BLMP) iff (see Block and Basu (1974)).

$$P\left(X>t+s_1,Y>t+s_2|X>t,Y>t\right)=P\left(X>s_1,Y>s_2\right),\ for\ all\ t,s_1,s_2\geq 0 \qquad (3)$$

Interpreting (3), if both of the two items are alive at t, then the joint distribution of their remaining lifetimes is the original joint distribution. In Block and Basu (1974), it has been shown that Freund's bivariate exponential distribution (see Freund (1961)) is absolutely continuous and that it possesses the BLMP. It has been known that Freund's bivariate exponential distribution (see Freund (1961)) is 'one of the few absolutely continuous bivariate continuous distributions' which possess the BLMP.

Obviously, for defining the BLMP of discrete-type distribution, the aforementioned definition can be applied in the same way (except for the fact that t, s_1, and s_2 are nonnegative integers). The following theorem asserts that all the members belonging to well-defined subclasses of the developed class possess discrete-type BLMP.

Theorem 4. Suppose that the intensity functions $r\left(n\right)$ and $v\left(n\right)$ are given by constants r and v, respectively, that is, $r\left(n\right)=r$ and $v\left(n\right)=v$, for $n = 1,2,\ldots$ and $r\left(0\right)=\nu\left(0\right)=0$. Then the joint probability distribution of (X_1, X_2) possesses the discrete BLMP.

Proof. For $n_1 < n_2$ and, $r\left(n\right)=r$ and $v\left(n\right)=v$, for $n = 1,2,\ldots$ and $r\left(0\right)=\nu\left(0\right)=0$, from Theorem 2,

$$P(X_1>n_1,X_2>n_2)=\prod_{i=1}^{n_2}(1-v)\prod_{i=1}^{n_2}(1-r)\left(1+\theta\sum_{l=n_1+1}^{n_2}\frac{r}{1-r}\right)$$

and

$$P(X_1 > k, X_2 > k) = \prod_{i=1}^{k} (1 - v(i)) \prod_{i=1}^{k} (1 - r(i))$$

Then,

$$\begin{aligned}
&P(X_1 > k + n_1, X_2 > k + n_2 | X_1 > k, X_2 > k) \\
&= P(X_1 > k + n_1, X_2 > k + n_2) / P(X_1 > k, X_2 > k) \\
&= \frac{\prod_{i=1}^{k+n_2} (1 - v) \prod_{i=1}^{k+n_2} (1 - r) \left(1 + \theta \sum_{l=k+n_1+1}^{k+n_2} \frac{r}{1-r}\right)}{\prod_{i=1}^{k}(1 - r) \prod_{i=1}^{k}(1 - v)} \\
&= \prod_{i=1}^{n_2}(1 - v) \prod_{i=1}^{n_2}(1 - r) \left(1 + \theta \sum_{k=n_1+1}^{n_2} \frac{r}{1 - r}\right) \\
&= P(X_1 > n_1, X_2 > n_2)
\end{aligned}$$

The case for $n_1 > n_2$ can also be shown symmetrically. The case with $n_1 = n_2$ can be shown easily.

5 Specific Families of Distributions

In this section, we obtain some specific families of discrete bivariate distribution using the result in Theorem 1. The class of bivariate distributions proposed in this paper is very general, in the sense that, by specifying $r(n)$ and $v(n)$, numerous specific families of discrete bivariate distributions can be generated. For the first four specific models, $r(n)$ is specified as the failure rate functions of the well-known typical discrete distributions: geometric, Poisson, and negative binomial distributions. To be simple, for all the considered models, we assume that the function of $v(n)$ is a constant which corresponds to the failure rate of the geometric distribution.

5.1 Model 1 Bivariate Geometric Distribution

We will obtain a family of discrete bivariate distributions based on the geometric distribution. Suppose.

$$f(n) = p(1 - p)^n, 0 < p < 1, \ for\ n = 0, 1, 2, \ldots$$

$$S(n) = (1 - p)^{n+1}, \ r(n) = p, r^*(n) = -ln(1 - p), n = 0, 1, 2, \ldots$$

and

$$f^*(n) = q(1-q)^n, 0 < q < 1, \ for\ n = 0, 1, 2, \ldots$$

$$S^*(n) = (1-q)^{n+1}, v(n) = q, v^*(n) = -ln(1-q), n = 0,1,2,...$$

Then, by Theorem 1, the joint distribution is given by.

$$P\left(X_1 = n_1, X_2 = n_2\right) = \theta p \left(1-p\right)^{n_2-1} \left(1-q\right)^{n_2} \left\{1-pq\right\}, \ for\ n_1 < n_2$$

$$P\left(X_1 = n_1, X_2 = n_2\right) = (1-\theta) p^2 \left(1-p\right)^{n_1-1} \left(1-q\right)^{n_1} \left\{1-pq\right\}, \ n_1 > n_2$$

$$P\left(X_1 = n, X_2 = n\right) = q \left(1-q\right)^n \left(1-p\right)^n, \ n = 0, 1, 2, \ldots$$

The proposed distribution admits a natural interpretation through a discrete-time random experiment based on sequential Bernoulli trials. Define two independent Bernoulli trials (denoted by B1 and B2) with success probabilities p and q, respectively, where the trial number starts from 0. In trial B1, the experiment continues until the second success is observed. If the first success in B1 occurs before that of B2, then the trial number until the first success becomes X_1 and the minimum of the trial number until the second success in B1 and that until the first success in B2 becomes X_2 with probability θ. In the same situation, the definitions of X_1 and X_2 are reversed with probability $1-\theta$. If the first success in B2 occurs before that of B1, then the trial number until the first success in B2 becomes X_1 and X_2 (in this case, $X_1 = X_2$).

5.2 Model 2 Bivariate Poisson Distribution

Let us consider the following Poisson distribution as the baseline distribution for $r\left(n\right)$, that is,

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

$$S(n) = \sum_{j=n+1}^{\infty} \frac{\lambda^j e^{-\lambda}}{j!} = \frac{\gamma(n+1,\lambda)}{n!}, \quad n = 0,1,2,...$$

where $\gamma\left(m, \lambda\right)$ denotes the lower incomplete gamma function:

$$\gamma(m,\lambda) \equiv \int_0^{\lambda} t^{m-1}e^{-t}dt = (m-1)!\sum_{n=m}^{\infty}\frac{\lambda^n e^{-\lambda}}{n!} = (m-1)!\left(1-\sum_{n=0}^{m-1}\frac{\lambda^n e^{-\lambda}}{n!}\right)$$

and then the failure rate functions are given by

$$r(n) = \frac{\lambda^n e^{-\lambda}}{n\,\gamma(n,\lambda)}, \quad r^*(n) = -\ln(1-r(n)), \quad n = 0,1,2,...$$

We also specify $v\,(n) = q$:

$$f^*(n) = q(1-q)^n, 0 < q < 1,\ for\ n = 0,1,2,\ldots$$

$$S^*(n) = (1-q)^{n+1}, v(n) = q, v^*(n) = -ln(1-q), n = 0,1,2,...$$

Then, from Theorem 1, we have.

$$P\,(X_1 = n_1, X_2 = n_2) = \theta\frac{\lambda^{n_1}e^{-\lambda}\,(1-q)^{n_2}}{n_2!\cdot\gamma\,(n_1+1,\lambda)}\left\{\lambda^{n_2}e^{-\lambda} + q\gamma\,(n_2+1,\lambda)\right\},\ for\ n_1 < n_2$$

$$P\,(X_1 = n_1, X_2 = n_2) = (1-\theta)\,\frac{\lambda^{n_2}e^{-\lambda}\,(1-q)^{n_1}}{n_1!\cdot\gamma\,(n_2+1,\lambda)}\left\{\lambda^{n_1}e^{-\lambda} + q\gamma\,(n_1+1,\lambda)\right\},\ n_1 > n_2$$

$$P\,(X_1 = n, X_2 = n) = \frac{q\,(1-q)^n\,\gamma\,(n,\lambda)}{(n-1)!},\ for\ n = 0,1,2,\ldots$$

5.3 Model 3 Bivariate Negative Binomial Distribution

Let us now consider the following negative binomial distribution:

$$f\,(n) = \left(\begin{array}{c} r+n-1 \\ n \end{array}\right)\beta^n(1-\beta)^r, r > 0, 0 < \beta < 1, n = 0,1,2,\ldots$$

$$S\,(n) = \sum_{j=n+1}^{\infty}\left(\begin{array}{c} r+j-1 \\ j \end{array}\right)\beta^j(1-\beta)^r = l_\beta\,(n+1,r)\,, n = 0,1,2,\ldots$$

where $l_\beta\,(a,b)$ denotes the regularized incomplete beta function:

Then observe that

$$r(n)=\frac{\binom{r+n-1}{n}\beta^n(1-\beta)^r}{l_\beta(n,r)},\quad r^*(n)=\begin{cases}-lnS(0)=-ln(1-(1-\beta)^r), & n=0\\ -ln\frac{S(n)}{S(n-1)}=-ln\frac{l_\beta(n+1,r)}{l_\beta(n,r)}, & n=1,2,...\end{cases}$$

We also specify $v(n)=q$:

$$f^*(n)=q(1-q)^n, 0<q<1,\ for\ n=0,1,2,\ldots$$

$$S^*(n)=(1-q)^{n+1}, v(n)=q, v^*(n)=-ln(1-q), n=0,1,2,...$$

Then, from Theorem 1, the joint distribution is given by.

$$P(X_1=n_1,X_2=n_2)=\theta\frac{\binom{r+n_1-1}{n_1}\beta^{n_1}(1-\beta)^r(1-q)^{n_2}}{l_\beta(n_1,r)}\left\{\binom{r+n_2-1}{n_2}\beta^{n_2}(1-\beta)^r+q\cdot l_\beta(n_2+1,r)\right\}, \text{for } n_1<n_2$$

$$P(X_1=n_1,X_2=n_2)=(1-\theta)\frac{\binom{r+n_2-1}{n_2}\beta^{n_2}(1-\beta)^r(1-q)^{n_1}}{l_\beta(n_2+1,r)}\left\{\binom{r+n_1-1}{n_1}\beta^{n_1}(1-\beta)^r+q\cdot l_\beta(n_1+1,r)\right\}, \text{for } n_1>n_2$$

$$P(X_1=n,X_2=n)=q(1-q)^n l_\beta(n,r),\ \text{for } n=0,1,2,\ldots$$

Note that Models 1–3 are obtained based on some specific parametric discrete distributions using the result in Theorem 1. However, relying on Corollary 1, we can generate additional families of distributions just by determining a proper functional forms of $r^*(n)$ ($v^*(n)$). The following models (Model 4–5) are generated in this way and they demonstrate the flexibility of the class proposed in this paper.

5.4 Model 4

Assume that the alternative failure rates are given by.

$$r^*(n)=\alpha n+\beta,\ and\ v^*(n)=\rho n+\nu, n=0,1,2,\ldots$$

Then, from Corollary 1, we have that

$$P(X_1=n_1,X_2=n_2)=\theta[exp(\alpha n_1+\beta)-1][exp((\alpha+\rho)n_2+\beta+v)-1]$$

$$\times exp\left\{-\sum_{i=0}^{n_2}((\alpha+\rho)i+\beta+v)\right\}, n_1<n_2$$

$$P(X_1=n_1,X_2=n_2)=(1-\theta)[exp(\alpha n_2+\beta)-1][exp((\alpha+\rho)n_1+\beta+v)-1]$$

$$\times exp\left\{-\sum_{i=0}^{n_1}((\alpha+\rho)i+\beta+v)\right\}, n_1>n_2$$

$$P\left(X_1 = n_1, X_2 = n_2\right) = \left[1 - exp\left\{-\left(\rho n + v\right)\right\}\right] exp\left\{-\sum_{i=0}^{n-1}\left(\left(\alpha + \rho\right) i + \beta + v\right)\right\}, n_1 = n_2 = n$$

5.5 Model 5

Let us consider that the alternative failure rates are given by

$$r^*\left(n\right) = \alpha n\left(n+1\right) + \beta \; and \; v^*\left(n\right) = \rho n + \nu, n = 0, 1, 2, \ldots$$

Then, from Corollary 1, the joint distribution is presented by

$$P(X_1 = n_1,\, X_2 = n_2) = \theta\left[exp\{\alpha n_1(n_1+1) + \beta\} - 1\right]\left[exp\{\alpha n_2(n_2+1) + \rho n_2 + \beta + v\} - 1\right] \times exp\left\{-\sum_{i=0}^{n_2}(\alpha\, i\,(i+1) + \rho\, i + \beta + v)\right\}, \quad n_1 < n_2$$

$$P(X_1 = n_1,\, X_2 = n_2) = (1-\theta)\left[exp\{\alpha n_2(n_2+1) + \beta\} - 1\right]\left[exp\{\alpha n_1(n_1+1) + \rho n_1 + \beta + v\} - 1\right] \times exp\left\{-\sum_{i=0}^{n_1}(\alpha\, i\,(i+1) + \rho\, i + \beta + v)\right\}, \quad n_1 > n_2$$

$$P(X_1 = n_1, X_2 = n_2) = [1 - \exp\{-(\rho n + v)\}]\exp\left\{-\sum_{i=0}^{n-1}(\alpha i(i+1) + \rho i + \beta + v)\right\},\; n_1 = n_2 = n$$

As mentioned before, the flexibility of the developed model also exists in modeling the proportion of tied observations. Using the simplest Model 1, it will now be illustrated how the proportion of tied observations can be modeled by appropriately specifying the corresponding parameters of the model. As the DT NHPP shock processes $\{N\left(n\right), n = 0, 1, 2, \ldots\}$ (with intensity function $r\left(n\right)$) governs the non-simultaneous failures of the components and the DT-NHPP $\{M\left(n\right), n = 0, 1, 2, \ldots\}$ (with intensity function $v\left(n\right)$) governs the simultaneous failures, in Model 1, by increasing q and decreasing p, the proportion of the tied observation can be increased (also, by decreasing q and increasing p, it can be deceased). In Fig. 1, the log joint pmfs of Model 1 for different values of q and p are given (with $\theta = 0.5$), where the flexibility of the developed model in modeling the proportion of tied observations is well illustrated.

6 Application to a Real Data Set

6.1 Motor Data

We now analyze a bivariate failure time data set reported in Reliability Edge (2002), summarized in Table 1. The data were obtained from a life test conducted on 18 identical parallel systems, each consisting of two components: Motor A (X_1) and Motor B (X_2), with failure times recorded in days. Unlike the data set to be analyzed in the next subsection, where the observed outcomes were restricted to a finite range, the failure times in this study take values on an unbounded support, that is, $X_1, X_2 \in \{0,1,2,\ldots\}$, allowing in principle arbitrarily large lifetimes. This makes the present data particularly suitable for assessing the performance of the proposed models in a fully non-truncated setting. The system operates in a redundant (parallel) configuration, meaning that it remains functional as long as at least one motor continues to operate. This configuration induces dependence between component

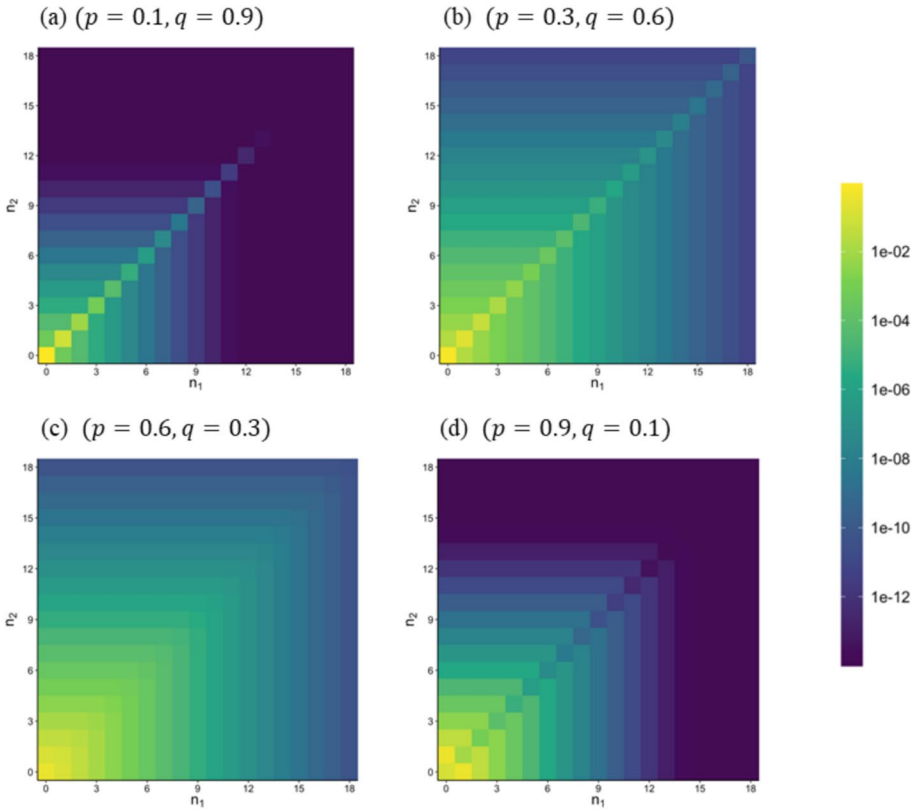


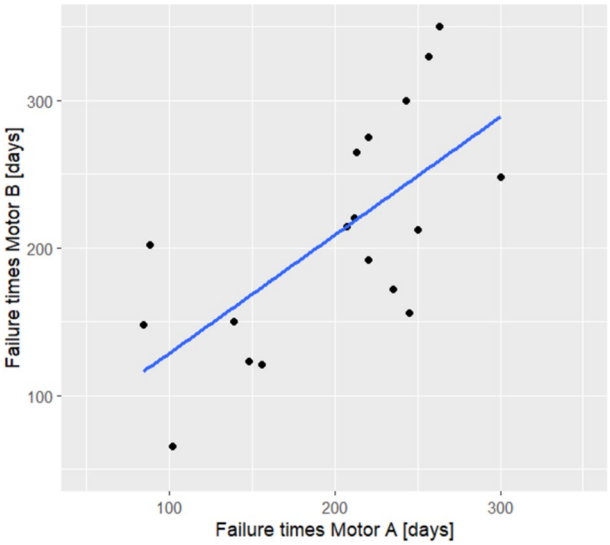
Fig. 1 The log joint pmfs of Model 1 for different values of q and p with $\theta = 0.5$ ((**a**) $q = 0.9$, $p = 0.1$; (**b**) $q = 0.6$,$p = 0.3$; (**c**) $q = 0.3$,$p = 0.6$; (**d**) $q = 0.1$,$p = 0.9$)

Table 1 Failure times (in days) of motors A and B

System	Motor A (X_1)	Motor B (X_2)	System	Motor A (X_1)	Motor B (X_2)
1	102	65	10	207	214
2	84	148	11	250	212
3	88	202	12	212	220
4	156	121	13	213	265
5	148	123	14	220	275
6	139	150	15	243	300
7	245	156	16	300	248
8	235	172	17	257	330
9	220	192	18	263	350

lifetimes, and an exploratory analysis of the paired failure times reveals a strong positive association, with a sample correlation coefficient of $\rho = 0.669$ (The scatter plot is presented in Fig. 2). This motivates the use of flexible bivariate lifetime models capable of capturing dependence over an unbounded domain.

Fig. 2 Scatter plot



To model this dependence structure, we fitted the proposed bivariate models (Models 1–5) developed in Sect. 4. For benchmarking purposes, we also considered the bivariate Poisson distribution (BPHD) model introduced by Holgate (1964) (see also Johnson et al. 1997). The joint probability mass function of the BPHD model is given by

$$P\left(X_1 = x_1, X_2 = x_2\right)$$

$$= exp[-(a+b-d)]\sum_{u=0}^{min\{x_1,x_2\}} \frac{(a-d)^{x_1-u}(b-d)^{x_2-u}d^u}{(x_1-u)!\,(x_2-u)!u!}, x_1 = 0,1,2,...,x_2 = 0,1,2,...$$

and it is noteworthy that, under this model, the marginal means of X_1 and X_2 are equal to their corresponding variances, reflecting the equidispersion property of the Poisson distribution. Parameter estimation for all models was carried out using maximum likelihood estimation (MLE), with numerical optimization used to obtain the estimates.

Table 2 summarizes the maximum likelihood (ML) estimates of the model parameters, together with the corresponding estimated log-likelihood $\widehat{l}$ and the values of the Akaike Information Criterion (AIC; Akaike 1974) and the Bayesian Information Criterion (BIC). As shown in Table 2, Model 3 outperforms the other competing models, including the BPHD model, as it attains the smallest AIC and BIC values.

Note that the data set has no tied observation and, due to it, the intensity $v(n)$ of the DT-NHPP $\{M\left(n\right), n = 0,1,2,\ldots\}$ which governs the simultaneous failures was estimated to be $v(n) \approx 0$ ($q \approx 0$ in Models 1–3 and $\rho \approx 0, \nu \approx 0$ for Models 4–5). Furthermore, for Models 4–5, the estimated values for the other parameters are very small. This is due to the fact that, as the observations are quite large, the speed of increments in $r^*(n)$ should also be very slow.

Note that this data set was previously analyzed using four models proposed by Lee et al. (2017) (see Table III in their paper). Their results indicated that the model providing the best fit among those considered was Model 3 in Lee et al. (2017), corresponding to the bivariate negative binomial distribution, with $\widehat{l} = -200.33$, AIC=412.66, and BIC=418.00. Table 2 demonstrates that Model 3 in the present study outperforms all models considered in Lee

Table 2 Estimation results

Model	parameter	estimate	$\widehat{l}$	AIC	BIC
Model 1	p	0.0089	−218.937	443.874	446.545
	q	7.608×10^{-11}			
	θ	0.5556			
Model 2	λ	203.273	−504.3826	1014.765	1017.4363
	q	0.0012			
	θ	0.5557			
Model 3	r	10	−201.612	411.224	413.782
	β	0.9448			
	q	1.510×10^{-11}			
	θ	0.5554			
Model 4	α	0.00007	−203.606	419.172	414.785
	β	2.487×10^{-9}			
	ρ	6.092×10^{-7}			
	ν	2.374×10^{-10}			
	θ	0.5556			
Model 5	α	0.00007	−203.548	417.096	421.5479
	β	7.440×10^{-10}			
	ρ	2.916×10^{-7}			
	ν	1.322×10^{-9}			
	θ	0.5554			
BPHD (Holgate 1964)	α	199.000	−513.958	1033.916	1036.587
	b	207.944			
	d	81.097			

et al. (2017) by attaining smaller AIC and BIC values. Notably, although the estimated log-likelihood of our Model 3 is comparable to that of Lee et al. (2017)'s Model 3, the proposed model achieves the best fit as it has smaller number of parameters, leading to a lower AIC and indicating superior model parsimony and statistical efficiency. In addition, Table 3 sum-

Table 3 Empirical and maximum likelihood estimates under the analyzed models

	$\widehat{E}[X_1]$	$\widehat{E}[X_2]$	$\widehat{V}[X_1]$	$\widehat{V}[X_2]$	$\widehat{\rho}$
Empirical estimate	199.000	207.944	4113.294	5855.585	0.669
MLE under Model 1	158.045	170.289	19,011.988	20,190.767	0.406
MLE under Model 2	184.755	185.907	3051.113	3103.918	0.963
MLE under Model 3	194.979	200.915	3978.466	4020.575	0.393
MLE under BPHD	199.000	207.944	199.000	207.944	0.399

marizes the empirical and model-based estimates of the marginal means, variances, and correlation coefficient for the motor data. Among the competing models, Model 3 provides the closest overall agreement with the empirical second-order characteristics, particularly in reproducing the marginal variances and dependence structure, supporting its selection as the most suitable model for this dataset.

Further diagnostic evidence for the adequacy of Model 3 is provided in Figs. 3 and 4. Figure 3 compares the observed failure time pairs (X_1, X_2) with simulated pairs generated from Model 3 using the maximum likelihood estimates. The simulated observations reproduce the overall dependence structure and dispersion pattern observed in the data, including the positive association between the lifetimes of Motors A and B. The observed data points are well embedded within the cloud of simulated values, indicating that Model 3 captures the joint behavior of the two failure times reasonably well. Figure 4 presents a comparison between the empirical cumulative distribution functions (ECDFs) of X_1 and X_2 and the corresponding marginal CDFs implied by Model 3. Although the failure times are discrete,

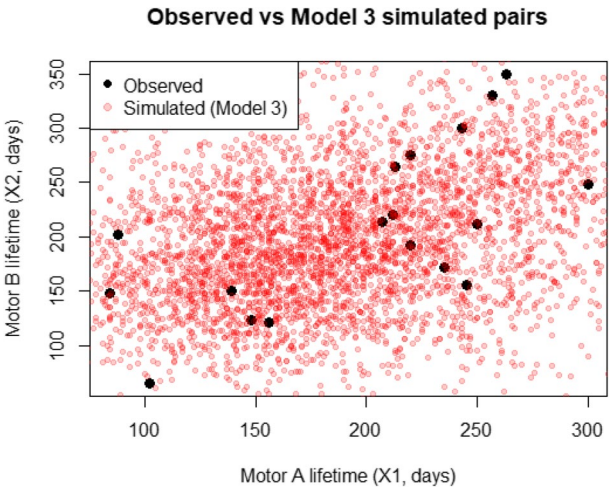


Fig. 3 Scatter plot of observed failure times and simulated pairs generated from Model 3

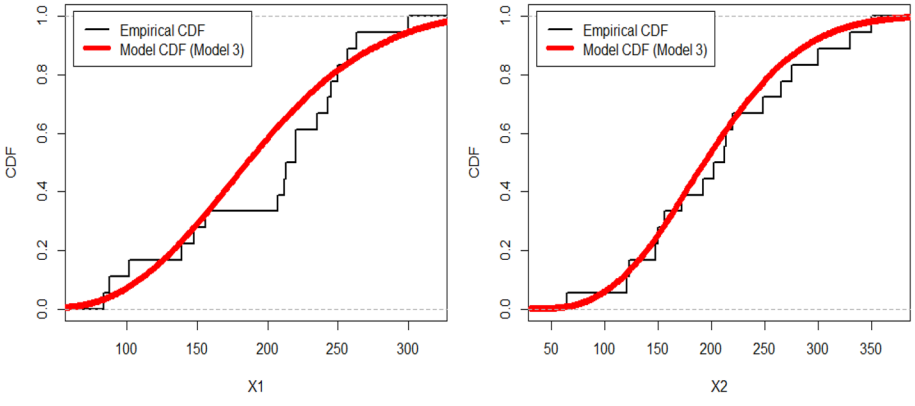


Fig. 4 Comparison of empirical marginal CDFs and Model 3-based marginal CDFs for Motor A (X_1) and Motor B (X_2)

the model-based CDFs appear smooth as they are estimated using a large Monte Carlo sample. The close agreement between the empirical and model-based CDFs across the entire range of the data suggests that Model 3 provides an adequate fit to the marginal distributions of both motor lifetimes. Minor deviations at the extremes can be attributed to the limited sample size, but no systematic lack of fit is observed.

Taken together, these diagnostic plots support the conclusions drawn from the information criteria, confirming that Model 3 not only achieves superior AIC and BIC values but also offers a reliable and interpretable representation of both the joint and marginal characteristics of the failure time data.

6.2 Air Pollution Data

The second dataset to be analyzed in this study is a bivariate ordinal data set from environmental science. The data set consists of hourly air pollution measurements collected in Seoul, South Korea, and publicly available through the Kaggle repository (Air pollution in Seoul, sourced from the Seoul Metropolitan Government Open Data Portal). The dataset covers the period from January 2017 to December 2019 and includes observations from 25 monitoring stations across Seoul. For each station and hour, the dataset provides concentration levels of major air pollutants, including PM10 (particulate matter $\leq 10\mu m$) and PM2.5 (particulate matter $\leq 2.5\mu m$), along gaseous pollutants such as SO_2, NO_2, CO, and O_3. Given that PM10 and PM2.5 are widely used as key indicators of air quality and public health risk, they were selected as the primary variables of interest.

Hourly air pollution measurements were first aggregated into daily mean concentrations to reduce short-term noise and improve data stability. Because pollutant levels exhibit temporal persistence, a weekly random subsampling procedure was applied, whereby one day per station per week was randomly selected. This process reduces serial autocorrelation while preserving seasonal variability, yielding a temporally thinned dataset consisting of 3,925 discrete bivariate observations, which is suitable for subsequent statistical modelling.

Daily PM10 and PM2.5 concentrations were discretized into four ordinal air-quality categories in accordance with widely adopted environmental guidelines (Good, Moderate, Unhealthy, Very Unhealthy). Table 4 presents a summary of the joint distribution of category levels for the two air pollution measurements (PM10 and PM2.5). These categories were encoded as integer-valued discrete variables (0–3), yielding a bivariate ordinal discrete outcome pair (X_1,X_2), where X_1 and X_2 represent the PM10 and PM2.5 category levels, respectively. The sample correlation between X_1 and X_2 was found to be moderately strong ($\rho = 0.66$), indicating meaningful cross-dependence between particulate matter categories and supporting the use of a discrete bivariate modeling framework.

Table 4 Joint frequency distribution of PM10 and PM2.5 category levels in Seoul (2017–2019)

PM10	PM2.5				Total
	Good	Moderate	Unhealthy	Very Unhealthy	
Good	1036	364	42	2	1444
Moderate	216	1446	497	12	2171
Unhealthy	10	50	148	68	276
Very Unhealthy	1	6	13	14	34
Total	1263	1866	700	96	3925

To examine the temporal structure of the processed dataset, autocorrelation functions (ACF) were applied to both continuous pollutant concentrations and their discretized counterparts. The analysis indicates that long-range temporal dependence is negligible, with only weak short-lag autocorrelation remaining. This suggests that the combined procedure of daily aggregation and weekly random subsampling substantially reduced temporal dependence while preserving meaningful seasonal variation in pollutant levels. Because pollutant category levels take nonnegative integer values and exhibit moderate cross-dependence between PM10 and PM2.5, the resulting dataset provides a natural and empirically meaningful foundation for discrete bivariate modeling. The final dataset consists of weekly subsampled discrete bivariate observations (X_1,X_2) and is well suited for evaluating probabilistic models of dependent count-like outcomes.

6.3 Model Truncation

Because the observed data take values only in the finite range $\{0,1,2,3\}$, all candidate models introduced in Sect. 4 were fitted under a truncated likelihood framework, in which the joint probability distribution is conditioned on the restricted support of the data. Specifically, for each model, the joint pmf $P(X_1 = n_1, X_2 = n_2)$ was renormalized over the truncated sample space so that the total probability over $(n_1, n_2) \in \{0,1,2,3\}^2$ sums to one. This truncation approach was applied consistently across all model classes, including Model 1 (geometric-based), Model 2 (Poisson-based), Model 3 (negative binomial-based), and Models 4–5 (alternative failure rate specifications). Specifically, if we denote by $P(X_1 = n_1, X_2 = n_2|\boldsymbol{\theta})$ the original joint pmf under a given model with parameter vector $\boldsymbol{\theta}$, the truncated likelihood is constructed using

$$P_T\left(X_1 = n_1, X_2 = n_2|\boldsymbol{\theta}\right) = \frac{P(X_1 = n_1, X_2 = n_2|\boldsymbol{\theta})}{\sum_{i=0}^{3}\sum_{j=0}^{3} P(X_1 = i, X_2 = j|\boldsymbol{\theta})}, n_1, n_2 \in \{0,1,2,3\}$$

This formulation ensures valid probability mass functions on the restricted domain and allows for fair and coherent comparison across models.

This data set has been analyzed by applying the truncated bivariate models based on Models 1–5 introduced in Sect. 4, where the likelihood was conditioned on the observed finite support of the data. For comparative purposes, we also considered the truncated version of BPHD model to analyze the same data set. For all models considered in this study, maximum likelihood estimation (MLE) was employed to estimate the model parameters, and the estimates were obtained numerically.

6.4 Estimation Results

Table 5 summarizes the ML estimates of the parameters for each model, together with the corresponding estimated log-likelihood $\widehat{l}$ and the AIC and BIC values. As shown in Table 5, Model 4 provides the best fit among all competing models, including the BPHD model, as it attains the smallest AIC and BIC values. This indicates that Model 4 achieves the most favorable balance between goodness of fit and model complexity. In addition, Table 6 reports the empirical and model-based estimates of the marginal means, variances, and correlation coefficient. The results show that all competing models reproduce the first-order

Table 5 Estimation results

Model	parameter	estimate	$\widehat{l}$	AIC	BIC
Model 1	p	0.3402	−7877.199	15,760.40	15,779.22
	q	0.3668			
	θ	0.7689			
Model 2	λ	1.2598	−7561.294	15,128.59	15,147.41
	q	0.4004			
	θ	0.7689			
Model 3	r	86.7654	−7565.061	15,138.12	15,163.22
	β	0.0144			
	q	0.3998			
	θ	0.7689			
Model 4	α	0.9524	−6957.129	13,924.26	13,955.63
	β	0.2491			
	ρ	0.6373			
	ν	0. 3121			
	θ	0.7689			
Model 5	α	0.4252	−6972.572	13,955.14	13,986.52
	β	0.2526			
	ρ	0.6351			
	ν	0.3127			
	θ	0.7689			
BPHD (Holgate 1964)	α	0.7666	−7248.250	14,502.50	14,521.33
	b	0.9666			
	d	0.6563			

Table 6 Empirical and maximum likelihood estimates under the analyzed models

	$\widehat{E}[X_1]$	$\widehat{E}[X_2]$	$\widehat{V}[X_1]$	$\widehat{V}[X_2]$	$\widehat{\rho}$
Empirical estimate	0.720	0.905	0.394	0.589	0.656
MLE under Model 1	0.634	0.880	0.741	0.947	0.587
MLE under Model 2	0.630	0.860	0.647	0.889	0.640
MLE under Model 4	0.719	0.908	0.435	0.538	0.632
MLE under BPHD	0.720	0.905	0.647	0.785	0.735

moments reasonably well, while noticeable differences arise in variance and dependence estimates, with Model 4 providing closer agreement with the empirical correlation structure.

Further insight into model performance is provided by Fig. 5, which compares the observed joint frequency distribution with the expected frequencies under Models 1, 2, 4, and the BPHD model. These models are presented because they represent either the strongest competitors or substantively different modeling structures, whereas the remaining models produced nearly indistinguishable fits. The figure demonstrates that Model 4 most closely reproduces the observed joint frequency pattern, particularly in capturing the distribution of outcomes along the diagonal $X_1 = X_2$, where a substantial proportion (67.3%) of observations is concentrated. Note that, compared with analysis of the data set in sub-Sect. 6.1, the intensity $v(n)$ of the DT-NHPP $\{M(n), n = 0,1,2,\dots\}$ which governs the simultaneous failures was estimated to be quite large to model the large proportion of tied observations. Thus, compared with the BPHD model, Model 4 exhibits a clear advantage in

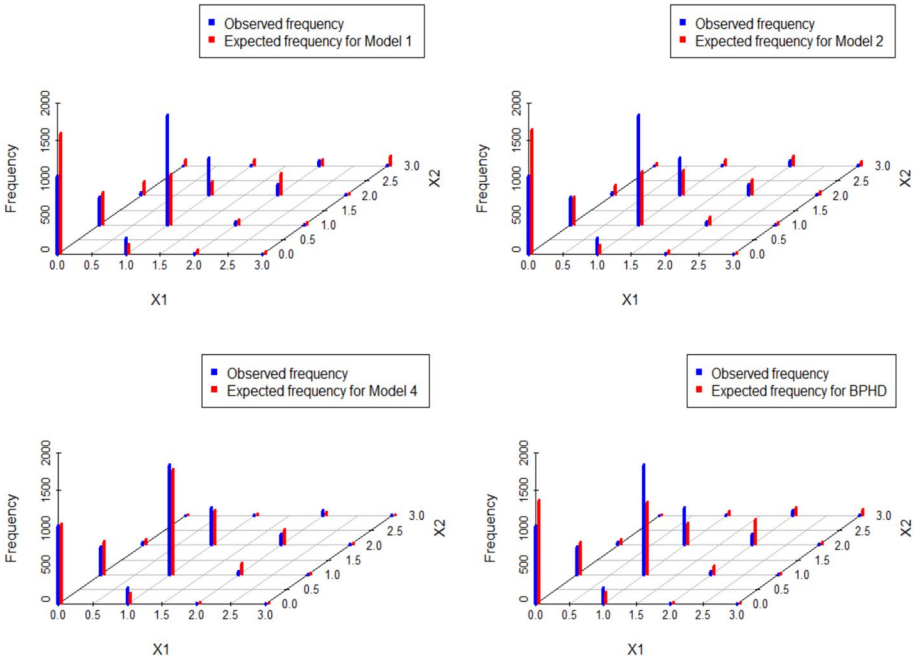


Fig. 5 The observed and estimated joint frequency for Model 1, 2, 4 and BPHD

explaining the high frequency of equal-score events, a prominent characteristic of the present dataset. While the BPHD model tends to underestimate or overestimate these diagonal probabilities, Model 4 provides a markedly more accurate representation of these outcomes. This improvement is visually evident in Fig. 5, where the expected frequencies under Model 4 align more closely with the observed bars across both diagonal and off-diagonal cells.

Overall, the combined evidence from Tables 5,6 and Fig. 5 indicates that Model 4 offers a superior and more data-adaptive description of the joint distribution, making it particularly well suited for datasets characterized by a high incidence of matched outcomes.

Acknowledgements The authors greatly thank the reviewer for helpful and insightful comments and suggestions.

Author Contributions Hyunju Lee and Ji Hwan Cha derived the main theoretical results and wrote the main manuscript text. Hyunju Lee performed the numerical analysis. All authors reviewed the manuscript.

Funding Open Access funding enabled and organized by Ewha Womans University. The work of the first author was supported by Hankuk University of Foreign Studies Research Fund of 2025 and the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. RS-2023–00240817). The work of the second author was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No. 2023R1A2C1003238).

Data Availability The motor failure dataset supporting the findings of this study is fully described and reported within the main text of this article. The air pollution dataset analyzed in this study is publicly available through the Kaggle repository under the title “Air Pollution in Seoul”, which compiles measurements from the Seoul Metropolitan Government Open Data Portal: https://www.kaggle.com/datasets/bappekim/air-pollution-in-seoul, https://www.kaggle.com/datasets/bappekim/air-pollution-in-seoul.

Declarations

Competing Interests Ji Hwan Cha is a member of the editorial board for Methodology and Computing in Applied Probability.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

References

Akaike H (1974) A new look at the statistical model identification. IEEE Trans Automat Control 19:716–723
Arnold BC (1975) Multivariate exponential distributions based on hierarchical successive damage. J Appl Probab 12:142–147
Badía FG, Sangüesa C (2014) Log-concavity for Bernstein-type operators using stochastic orders. J Math Anal Appl 413:953–962
Barbu VS, Limnios N (2008) Semi-Markov chains and hidden Markov models toward applications: Their use in reliability and DNA analysis. Springer, New York
Bermúdez L, Karlis D (2022) Copula-based bivariate finite mixture regression models with an application for insurance claim count data. TEST 31:1082–1099
Block HW, Basu AP (1974) A continuous bivariate exponential extension. J Am Stat Assoc 69:1031–1037
Brady B, Singpurwalla ND (1990) Stochastically monotone dependence. In: Block HW, Sampson AR, Savits TH, eds. Lecture Notes-Monograph Series. Institute of Mathematical Statistics. 93–102
Cha JH (2014) Characterization of the generalized Polya process and its applications. Adv Appl Probab 46:1148–1171
Cha JH, Limnios N (2021) Discrete time minimal repair process and its reliability applications under random environments. Stoch Models https://doi.org/10.1080/15326349.2021.2018336
Dewan I, Sudheesh KK (2009) On proportional (reversed) hazard model for discrete data. ISI Delhi Centre Preprint No. isid/ms/2009/11, Indian Statistical Institute, New Delhi, India
Eryilmaz SN (2013) On the lifetime behavior of a discrete time shock model. J Comput Appl Math 237(1):384–388
Eryilmaz SN, Kan C (2021) Reliability assessment for discrete time shock models via phase-type distributions. Appl Stoch Model Bus Ind 37(3):513–524
Freund JE (1961) A bivariate extension of the exponential distribution. J Am Stat Assoc 56:971–977
Goyal D, Hazra NK, Finkelstein M (2022) On the general δ-shock model. TEST 31:994–1029
Goyal D, Hazra NK, Finkelstein M (2023) A general class of shock models with dependent inter-arrival times. TEST 32:1079–1105
Holgate B (1964) Estimation for the bivariate Poisson distribution. Biometrika 51:241–245
Johnson NL, Kotz S, Balakrishnan N (1997) Discrete multivariate distributions. Wiley, New York
Lee H, Cha JH (2024) New continuous bivariate distributions generated from shock models. Statistics 58:437–449
Lee H, Cha JH, Pulcini G (2017) Modeling discrete bivariate data with applications to failure and count data. Qual Reliab Eng Int 33:1455–1473
Li XJ, Li S, Tian GL, Shi J (2024) Modeling paired binary data by a new bivariate Bernoulli model with flexible beta kernel correlation. TEST 33:1180–1224
Limnios N, Oprişan G (2001) Semi–Markov processes and reliability. Birkhäuser, Boston
Lu JC (1989) Weibull extensions of the Freund and Marshall-Olkin bivariate exponential models. IEEE Trans Reliab 38:615–619
Marshall AW, Olkin I (1967) A multivariate exponential distribution. J Am Stat Assoc 62:30–44
Navarro J (2016) Stochastic comparisons of generalized mixtures and coherent systems. TEST 25:150–169

Navarro J (2018) Stochastic comparisons of coherent systems. Metrika 81:465–482
O'cinneide CA, Raftery AE (1989) A continuous multivariate exponential distribution that is multivariate phase type. Stat Probab Lett 7:323–325
Ozkut M (2023) Reliability and optimal replacement policy for a generalized mixed shock model. TEST 32:1038–1054
Raftery AE (1984) A continuous multivariate exponential distribution. Commun Stat 13:947–965
ReliaSoft's R&D Staff (2002) Using QALT models to analyze system configurations with load sharing. Reliability Edge 3:1–4
Roy D, Gupta R (1992) Classification of discrete lives. Microelectron Reliab 32:1459–1473
Salvia AA, Bollinger RC (1982) On discrete hazard functions. IEEE Trans Reliab R-31:458–459
Shaked M, Shanthikumar JG (2007) Stochastic orders. Springer, New York
Wang J, Li Y (2005) Dependency measures under bivariate homogeneous shock models. Stat 39:73–80

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.