

RESEARCH

Open Access



A note on inconsistent families of discrete multivariate distributions

Sugata Ghosh¹, Subhajit Dutta¹ and Marc G. Genton^{2*}

*Correspondence:

marc.genton@kaust.edu.sa
²CEMSE Division, King Abdullah
University of Science and
Technology, Thuwal 23955-6900,
Saudi Arabia
Full list of author information is
available at the end of the article

Abstract

We construct a d -dimensional discrete multivariate distribution for which *any proper subset* of its components belongs to a specific family of distributions. However, the joint d -dimensional distribution fails to belong to that family and in other words, it is 'inconsistent' with the distribution of these subsets. We also address preservation of this 'inconsistency' property for the symmetric Binomial distribution, and some discrete distributions arising from the multivariate discrete normal distribution.

Keywords: Binomial distribution, Discrete normal and skew-normal, Multivariate distributions, Sign function, Symmetric distributions

AMS Subject Classification: 62E10; 62H05

1 Introduction

If a multivariate distribution is parametrically specified, often the lower-dimensional marginals follow the same distribution of an appropriate dimension. For example, all the lower-dimensional distributions of a multivariate Gaussian distribution are Gaussian. The converse is however not necessarily true. For example, Dutta and Genton (2014) gave a construction of a non-Gaussian multivariate distribution with all lower-dimensional Gaussians and considered some generalizations of this for a certain class of elliptical and skew-elliptical distributions. In the discrete case, all the lower-dimensional marginals of a bivariate Binomial distribution (Bairamov and Gultekin 2010) follow the Binomial distribution. Conversely, given a set of marginal distributions, different dependence structures can give rise to different joint distributions. In a more general setting, Hoeffding (1940) and Fréchet (1951) independently obtained a characterization of the class of bivariate distributions with given univariate marginals. Characterization problems for discrete distributions using conditional distributions and their expectations have been investigated by several authors; see, e.g., Dahiya and Korwar (1977), Ruiz and Navarro (1995), and Nguyen et al. (1996). The paper by Conway (1979) also discussed some additional properties and derived appropriate relationships for such systems, while a method of constructing multivariate distributions with specified univariate marginals and a given correlation matrix was studied by Cuadras (1992).

The main aim of this paper is the converse question for the discrete multivariate setup. More specifically, we construct a discrete multivariate distribution (see, e.g., Johnson et al. (1997)), all of whose lower-dimensional marginals follow the same symmetric distribution but the joint distribution does not conform to that pattern and has a different distribution.

We say that the joint distribution is ‘*inconsistent*’ with the distribution of its marginals. The main idea of this construction is based on a transformation using the sign function applied to a class of discrete symmetric random variables. The sign function plays a key role in yielding a distribution with ‘restricted support’.

The structure of the paper is as follows. In Section 2, we first motivate our construction for the bivariate set-up with a simple case starting from a symmetric Binomial distribution. Generalization of this result to symmetric discrete distributions with a specific support in higher dimensions is explored in Section 3 where our main result is stated. We also address preservation of this ‘inconsistency’ property for the symmetric Binomial distribution in Section 4, a class of symmetric discrete distributions constructed from symmetric continuous distributions in Section 5, and the multivariate discrete skew-normal distribution in Section 6. Proofs of all the theorems are provided in the Appendix.

2 A motivating bivariate example

Consider a random variable U that follows a Binomial distribution with parameters $n = 2$ and $p = \frac{1}{2}$. The support of this distribution is the set $\mathcal{N}_2 = \{0, 1, 2\}$, and it has the following probability mass function (pmf):

$$q_0(u) = \binom{2}{u} \frac{1}{4} \text{ if } u \in \mathcal{N}_2.$$

Let U_1 and U_2 be two independent copies of U . The support of the joint distribution of U_1 and U_2 is $\mathcal{N}_2^2 = \{0, 1, 2\} \times \{0, 1, 2\}$. Thus, the joint pmf of U_1 and U_2 is

$$q_0(u_1, u_2) = \binom{2}{u_1} \binom{2}{u_2} \frac{1}{16} \text{ if } (u_1, u_2) \in \mathcal{N}_2^2.$$

Define $X = U - 1$. The support of the distribution of X is the set $\mathcal{Z}_1 = \{-1, 0, 1\}$, and it has the pmf:

$$q(x) = \binom{2}{x+1} \frac{1}{4} \text{ if } x \in \mathcal{Z}_1.$$

Clearly, this is a version of the Binomial distribution which is symmetric about 0. Let X_1 and X_2 be two independent copies of X . The support of the joint distribution $(X_1, X_2)^T$ is $\mathcal{Z}_1^2 = \{-1, 0, 1\} \times \{-1, 0, 1\}$. Thus, the joint pmf of X_1 and X_2 is

$$q(x_1, x_2) = \binom{2}{x_1+1} \binom{2}{x_2+1} \frac{1}{16} \text{ if } (x_1, x_2) \in \mathcal{Z}_1^2. \tag{1}$$

Recall the univariate sign function, namely,

$$S(x) = \begin{cases} -1, & \text{if } x < 0, \\ 0, & \text{if } x = 0, \\ 1, & \text{if } x > 0. \end{cases}$$

We use a modified version of this sign function. Consider the following transformation:

$$X_1^* = X_1 S_{2,1} \text{ and } X_2^* = X_2 S_{1,2},$$

where $S_{1,2}$ and $S_{2,1}$ are defined as follows:

$$S_{1,2} = \begin{cases} -1, & \text{if } X_1 < 0, \text{ or } (X_1 = 0 \text{ and } X_2 \neq 0), \\ 0, & \text{if } (X_1 = 0 \text{ and } X_2 = 0), \\ 1, & \text{if } X_1 > 0, \end{cases}$$

and

$$S_{2,1} = \begin{cases} -1, & \text{if } X_2 < 0, \text{ or } (X_2 = 0 \text{ and } X_1 \neq 0), \\ 0, & \text{if } (X_2 = 0 \text{ and } X_1 = 0), \\ 1, & \text{if } X_2 > 0. \end{cases}$$

We now study the distribution of X_1^* .

$$\begin{aligned} P(X_1^* = -1) &= P(X_1 S_{2,1} = -1) \\ &= P(X_1 = 1, S_{2,1} = -1) + P(X_1 = -1, S_{2,1} = 1) \\ &= P(X_1 = 1, X_2 < 0) + P(X_1 = 1, X_2 = 0, X_1 \neq 0) + P(X_1 = -1, X_2 > 0) \\ &= P(X_1 = 1, X_2 = -1) + P(X_1 = 1, X_2 = 0) + P(X_1 = -1, X_2 = 1) \\ &= \frac{1}{16} + \frac{1}{8} + \frac{1}{16} = \frac{1}{4}, \end{aligned}$$

and

$$\begin{aligned} P(X_1^* = 0) &= P(X_1 S_{2,1} = 0) \\ &= P(X_1 = 0, S_{2,1} \neq 0) + P(X_1 \neq 0, S_{2,1} = 0) + P(X_1 = 0, S_{2,1} = 0) \\ &= \{P(X_1 = 0, X_2 < 0) + P(X_1 = 0, X_2 > 0)\} + 0 + P(X_1 = 0, X_2 = 0) \\ &= P(X_1 = 0, X_2 = -1) + P(X_1 = 0, X_2 = 1) + P(X_1 = 0, X_2 = 0) \\ &= \frac{1}{8} + \frac{1}{8} + \frac{1}{4} = \frac{1}{2}. \end{aligned}$$

This now implies that $P(X_1^* = 1) = 1/4$. Hence, the distribution of X_1^* is the same as that of the distribution of X . Similarly, one can show that X_2^* is identically distributed with X .

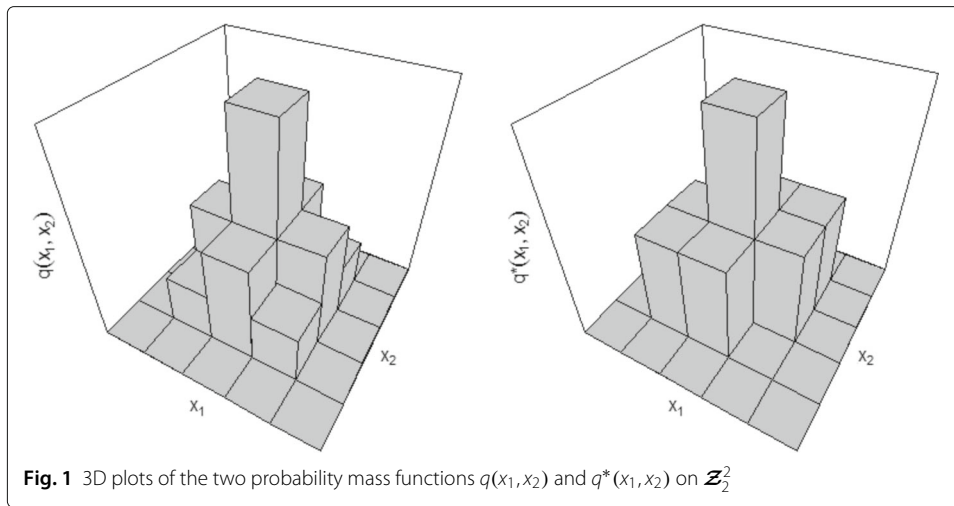
By our construction, the support of the joint distribution is on a restricted set, i.e.,

$$X_1^* \cdot X_2^* = X_1 S_{2,1} \cdot X_2 S_{1,2} = X_1 S_{1,2} \cdot X_2 S_{2,1}$$

which is non-negative with probability one. The joint distribution of $(X_1^*, X_2^*)^T$ is as follows:

$$\begin{aligned} P(X_1^* = -1, X_2^* = -1) &= P(X_1 = -1, X_2 = 1) + P(X_1 = 1, X_2 = -1) = \frac{1}{16} + \frac{1}{16} = \frac{1}{8}, \\ P(X_1^* = -1, X_2^* = 0) &= P(X_1 = 1, X_2 = 0) = \frac{1}{8}, \\ P(X_1^* = 0, X_2^* = -1) &= P(X_1 = 0, X_2 = 1) = \frac{1}{8}, \\ P(X_1^* = 0, X_2^* = 0) &= P(X_1 = 0, X_2 = 0) = \frac{1}{4}, \\ P(X_1^* = 0, X_2^* = 1) &= P(X_1 = 0, X_2 = -1) = \frac{1}{8}, \\ P(X_1^* = 1, X_2^* = 0) &= P(X_1 = -1, X_2 = 0) = \frac{1}{8}, \text{ and} \\ P(X_1^* = 1, X_2^* = 1) &= P(X_1 = -1, X_2 = -1) + P(X_1 = 1, X_2 = 1) = \frac{1}{16} + \frac{1}{16} = \frac{1}{8}. \end{aligned}$$

Let $\mathbb{I}_A(x)$ be the indicator function, which is defined as



$$\mathbb{I}_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases}$$

It is easy to check that the joint pmf of $(X_1^*, X_2^*)^T$ can be re-written in a concise form as follows:

$$q^*(x_1, x_2) = q(x_1, x_2)\mathbb{I}_{\{0\}}(x_1x_2) + 2q(x_1, x_2)\mathbb{I}_{(0,\infty)}(x_1x_2) \text{ for } (x_1, x_2) \in \mathcal{Z}_1^2,$$

where the expression of $q(x_1, x_2)$ is defined in (1). The joint distribution of $(X_1^*, X_2^*)^T$ is asymmetric, and clearly does not belong to the family of distributions containing the joint distribution of $(X_1, X_2)^T$ as the support of $(X_1^*, X_2^*)^T$ is restricted to two quadrants only, although the univariate distributions of X_1^* and X_2^* are the same as that of the distribution of X . The plot in Fig. 1 demonstrates the basic idea of shifting of probability mass to only two quadrants (see the right panel of Fig. 1), and clearly shows the difference between the pmfs $q(x_1, x_2)$ and $q^*(x_1, x_2)$.

This motivates us to construct a d -dimensional random vector such that the distribution of any proper subset of its components belongs to a certain family of distributions, but the joint distribution, with all the d components taken together, fails to conform to that family of distributions.

3 General symmetric discrete multivariate distributions

We now consider a general discrete distribution symmetric about the point 0 with support on a finite or countably infinite set \mathbb{S} , given by the pmf

$$q(x) = \sum_{j \in \mathbb{S}} \delta(x - j)m_j, \tag{2}$$

where $0 \leq m_j \leq 1$ and $m_j = m_{-j}$ for all $j \in \mathbb{S}$ with $\sum_{j \in \mathbb{S}} m_j = 1$.

Let X_1, \dots, X_d be independent copies of $X \sim q(\cdot)$. Here, the function $\delta(\cdot)$ is defined as follows:

$$\delta(x) = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0. \end{cases}$$

We have $I_{\{0\}}(\mathbf{x}) = \delta(\mathbf{x})$. The joint pmf of $(X_1, \dots, X_d)^T$ is given by

$$q(\mathbf{x}) = \prod_{i=1}^d q(x_i) = \prod_{i=1}^d \left\{ \sum_{j \in \mathbb{S}} \delta(x_i - j) m_j \right\},$$

where $\mathbf{x} = (x_1, \dots, x_d)^T \in \mathbb{S}^d$.

Again, consider the transformation $(X_1, \dots, X_d) \rightarrow (X_1^*, \dots, X_d^*)$ such that

$$X_1^* = X_1 S_{2,1}, \dots, X_{i-1}^* = X_{i-1} S_{i,i-1}, \dots, X_{d-1}^* = X_{d-1} S_{d,d-1}, X_d^* = X_d S_{1,d}, \tag{3}$$

where the modified sign function is defined as

$$S_{i,i-1} = \begin{cases} -1, & \text{if } X_i < 0 \text{ or } (X_i = 0 \text{ and } X_{i-1} \neq 0), \\ 0, & \text{if } (X_i = 0 \text{ and } X_{i-1} = 0), \\ 1, & \text{if } X_i > 0, \end{cases}$$

for $i = 2, \dots, d$, and

$$S_{1,d} = \begin{cases} -1, & \text{if } X_1 < 0 \text{ or } (X_1 = 0 \text{ and } X_d \neq 0), \\ 0, & \text{if } (X_1 = 0 \text{ and } X_d = 0), \\ 1, & \text{if } X_1 > 0. \end{cases}$$

Now, $\prod_{i=1}^d X_i^* = X_1 S_{2,1} \cdot X_2 S_{3,2} \cdots X_d S_{1,d} = X_1 S_{1,d} \cdot X_2 S_{2,1} \cdots X_d S_{d,d-1}$. Fix $i \in \{2, \dots, d\}$. Again, $X_i < 0 \Rightarrow S_{i,i-1} = -1$, hence $X_i S_{i,i-1} > 0$, while $X_i > 0 \Rightarrow S_{i,i-1} = 1$, hence $X_i S_{i,i-1} > 0$. Also, $X_i = 0 \Rightarrow X_i S_{i,i-1} = 0$. To summarize, $X_i S_{i,i-1} \geq 0$ for any $i = 2, \dots, d$. Similarly, $X_1 S_{1,d} \geq 0$. Hence, we obtain $\prod_{i=1}^d X_i^* \geq 0$.

Theorems 1 and 2 below state the joint distribution of $(X_1^*, \dots, X_d^*)^T$ as well as the distribution of the lower-dimensional vectors.

Theorem 1 *The joint pmf of $(X_1^*, \dots, X_d^*)^T$ is*

$$q^*(\mathbf{x}) = q(\mathbf{x}) \mathbb{I}_{\{0\}} \left(\prod_{i=1}^d x_i \right) + 2q(\mathbf{x}) \mathbb{I}_{(0,\infty)} \left(\prod_{i=1}^d x_i \right) \text{ for } \mathbf{x} \in \mathbb{S}^d.$$

Theorem 2 *Any sub-vector $(X_{k_1}^*, \dots, X_{k_{d'}}^*)^T$ of $(X_1^*, \dots, X_d^*)^T$ with $d' < d$ is component-wise independent, and the joint distribution of $(X_{k_1}^*, \dots, X_{k_{d'}}^*)^T$ is the same as that of $(X_{k_1}, \dots, X_{k_{d'}})^T$.*

In particular, this result holds true if $\mathbb{S} = \mathbb{Z} = \{\dots, -1, 0, 1, \dots\}$ or any proper subset of \mathbb{Z} which is symmetric about 0, (i.e., $x \in \mathbb{S} \Leftrightarrow -x \in \mathbb{S}$).

4 The symmetric binomial distribution

We started our investigation in Section 2 using a discrete distribution symmetric about 0 with support on the set \mathcal{N}_2 , and extended it to the case of a general symmetric discrete distribution symmetric about 0, supported on a finite or countably infinite subset \mathbb{S} of \mathbb{R} . However, it is interesting to see whether such ‘inconsistency’ results continue to hold for other symmetric distributions for which the point of symmetry is not necessarily 0.

Suppose that U follows a symmetric distribution with a finite or countably infinite support and point of symmetry u_0 . Define $X = U - u_0$. Then, X follows a symmetric distribution about 0 with a finite or countably infinite support (as mentioned in (2)).

Assume $U \sim \text{Binomial}(n, 1/2)$, and consider u_0 to be $n/2$. We now have versions of Theorems 1 and 2 for symmetric Binomial distributions.

5 Discrete symmetric distributions

Given any continuous distribution with distribution function (df) F on \mathbb{R} , we can define its ‘discrete analogue’ (see, e.g., Alzaatreh et al. 2012) with pmf $q_d(\cdot)$ as follows:

$$q_d(x) = F(x + 1) - F(x), \quad x \in \mathbb{Z},$$

with $q_d(x) \geq 0$ and $\sum_{x \in \mathbb{Z}} q_d(x) = 1$. If $U \sim F(\cdot)$, then this implies that $[U] \sim q_d(\cdot)$. Here, $[u]$ denotes the largest integer less than or equal to u .

Assume F to be symmetric about the point 0, i.e., $F(x) + F(-x) = 1$ for any $x \in \mathbb{R}$. Now, note that $q_d(0) = q_d(-1)$ and the point of symmetry of $[U]$ is clearly $-\frac{1}{2}$. Define $X = [U] + \frac{1}{2}$. Then, X is a discrete variate with support on a countably infinite set \mathbb{S} which is symmetric about the point 0, say, $X \sim q_{dS}(\cdot)$. In particular, if we take $F(x) = \Phi(x)$, where $\Phi(\cdot)$ is the df of the standard normal distribution, then we obtain the discrete normal distribution (dN) (Roy 2003). The pmf of X (say, $q_{dN}(\cdot)$) simplifies to be $\Phi(x + 1/2) - \Phi(x - 1/2)$ with $x \in \mathbb{S}$. Using this discretization idea, Chakraborty and Chakravarty (2016) have proposed a discrete logistic distribution starting from the continuous two-parameter logistic distribution. Now, we can construct versions of Theorems 1 and 2 for such symmetric discrete probability distributions.

6 The multivariate discrete normal and related distributions

In the multivariate setting (say, \mathbb{R}^d), let $U_d \sim N_d(\mathbf{0}, \mathbf{I}_d)$ be the d -dimensional normal distribution with \mathbf{I}_d as the $d \times d$ identity matrix. We can define an analogue of the discrete normal (dN) distribution on \mathbb{S}^d as follows:

$$\mathbf{X}_d = [U_d] + \frac{1}{2} \mathbf{1}_d = \left([U_1] + \frac{1}{2}, \dots, [U_d] + \frac{1}{2} \right)^T. \tag{4}$$

The joint pmf of \mathbf{X}_d is the product $\prod_{i=1}^d q_{dN}(x_i)$.

We now apply the transformation defined in Eq. (6) on the vector \mathbf{X}_d stated in Eq. (4) to obtain \mathbf{X}_d^* . Define

$$\mathbf{Y}_d^* = A \mathbf{1}_d + B \mathbf{X}_d^*,$$

where $A, B \in \mathbb{S}$ are discrete random variables independent of \mathbf{X}_d^* . The support of the random vector \mathbf{Y}_d^* is the set \mathbb{S}_*^d , where \mathbb{S}_* is a countable set which depends on the joint distribution of A and B . Define $\mathbf{Y}_{d-1,d}^*$ based on \mathbf{Y}_d^* by selecting a subset of size $d - 1$ from the set $\{1, \dots, d\}$.

Theorem 3 *The distributions of all the $(d - 1)$ -dimensional random vectors $\mathbf{Y}_{d-1,d}^*$ belong to the same family of distributions, but that of \mathbf{Y}_d^* does not.*

We now state consequences of Theorem 3 in some popular and important classes of symmetric and asymmetric distributions derived from the multivariate discrete normal distribution:

- (Discrete Scale Mixture) Consider $A = 0$ (a degenerate random variable) and $B = W$, where W is a non-negative, discrete random variable independent of \mathbf{X}_d^* .
- (Discrete Skew-Normal) Consider a discrete normal variate $X_{d+1} \sim q_{dN}(\cdot)$ independent of \mathbf{X}_d^* . Define $A = \delta|X_{d+1}|$ and $B = \sqrt{1 - \delta^2}$ (a degenerate random variable) with $-1 < \delta < 1$ (also see pp.128-129 of Azzalini (2014)).

Using Theorem 3, we now obtain this ‘inconsistency result’ for a class of discrete skew-normal distributions. One can also extend this results to a class of skew-elliptical distributions starting from the discrete scale mixture. More general results using the idea of modulation of symmetric discrete distributions proposed recently by Azzalini and Regoli (2014) still remain open.

Appendix: Proofs and Mathematical Details

Proof 1 (Proof of Theorem 1) We break the proof into two parts, namely, Case I and Case II.

Case I: $\prod_{i=1}^d X_i^* = 0$. This fact now implies that at least one of the X_i^* 's is zero. Without loss of generality, let $X_1^* = 0$ and consider the event $(X_1^* = 0, X_2^* = x_2, \dots, X_d^* = x_d)$. Under the assumption that $X_1^* = 0$, we now establish the following facts:

(F1) $X_i^* = 0 \Rightarrow X_i = 0$ for any $i = 1, \dots, d$. To show this, we note that if $X_1^* = 0$, then $X_1 S_{2,1} = 0$. Now, either $X_1 = 0$ or $S_{2,1} = 0$. If $X_1 = 0$ then we are done. If $S_{2,1} = 0$, then $X_2 = X_1 = 0$, in particular $X_1 = 0$. More generally, $X_i^* = 0 \Rightarrow X_i = 0$ for any $i = 1, \dots, d$.

(F2) If $X_1^* = 0$, then $X_d = -x_d$. To show this, suppose that $X_d^* \neq 0$, then we obtain $X_d S_{1,d} \neq 0$ and in particular $X_d \neq 0$. We have assumed that $X_1^* = 0$, now using (F1) we know that $X_1 = 0$. Combining the facts that $X_1 = 0$ and $X_d \neq 0$, we have by definition $S_{1,d} = -1$. Again, $X_d^* = x_d \Rightarrow X_d S_{1,d} = x_d$ and thus we obtain $X_d = -x_d$. On the other hand, suppose that $X_d^* = 0$. In this case $x_d = 0$. By (F1), $X_d^* = 0 \Rightarrow X_d = 0$. Thus we trivially have $X_d = -x_d$ (both sides being equal to 0). Hence, in all the cases, we have $X_d = -x_d$.

(F3) We have $X_i = x_i S_{i+1,i}$ for $i = 2, \dots, d - 1$. To show this, we start with any $i \in \{2, \dots, d - 1\}$. Now $X_i^* = x_i$, which implies $X_i S_{i+1,i} = x_i$. First take $x_i \neq 0$. Then $X_i \neq 0$ and $S_{i+1,i} \neq 0$.

Since $X_i \neq 0$, $S_{i+1,i}$ simplifies to the following function:

$$S_{i+1,i} = \begin{cases} -1, & \text{if } X_{i+1} \leq 0, \\ 1, & \text{if } X_{i+1} > 0. \end{cases}$$

Thus, $S_{i+1,i}$ only takes the values -1 or 1 , which implies that $S_{i+1,i}^2 = 1$. We can multiply $S_{i+1,i}$ on both sides of the equation $X_i S_{i+1,i} = x_i$ to obtain $X_i = x_i S_{i+1,i}$.

On the other hand, suppose $x_i = 0$. Using (F1), we obtain $X_i = 0$ which trivially satisfies $X_i = x_i S_{i+1,i}$ (as both sides equal zero). This is true for all $i = 2, \dots, d - 1$.

(F4) We have $P(X_i = x_i S_{i+1,i}) = P(X_i = x_i)$. To show this, we first consider the case when $x_i \neq 0$. Then $S_{i+1,i} = \pm 1$. Hence $P(X_i = x_i S_{i+1,i}) = P(X_i = \pm x_i)$, which is $P(X_i = x_i)$ by virtue of symmetry. The claim follows trivially when $x_i = 0$.

Now, we consider the probability

$$\begin{aligned}
 &P(X_1^* = 0, X_2^* = x_2, \dots, X_d^* = x_d) \\
 &= P(X_1 = 0, X_2 = x_2 S_{3,2}, \dots, X_{d-1} = x_{d-1} S_{d,d-1}, X_d = -x_d) \\
 &= P(X_1 = 0)P(X_2 = x_2 S_{3,2}) \cdots P(X_{d-1} = x_{d-1} S_{d,d-1})P(X_d = -x_d) \text{ [Using independence]} \\
 &= P(X_1 = 0)P(X_2 = x_2) \cdots P(X_{d-1} = x_{d-1})P(X_d = x_d) \text{ [Using (F4)]} \\
 &= P(X_1 = 0, X_2 = x_2, \dots, X_{d-1} = x_{d-1}, X_d = x_d)
 \end{aligned}$$

In this case, the joint pmf of $(X_1^*, \dots, X_d^*)^T$ denoted by $q^*(\mathbf{x})$ is given as

$$q^*(\mathbf{x}) = q(\mathbf{x}).$$

This completes the first part.

Case II: Assume $\prod_{i=1}^d X_i^* > 0$. This implies that none of the X_i^* 's are zero. Since $X_i^* \neq 0$ implies that $X_i \neq 0$ for $i = 1, \dots, d$, we can ignore the value of the sign function $S(\cdot)$ at 0 and it simplifies to the following function:

$$S(x) = \begin{cases} -1, & \text{if } x < 0, \\ 1, & \text{if } x > 0. \end{cases}$$

Thus, in this case we always have $S(x) = \pm 1$. Also, note that here the modified sign function $S_{i,i-1}$ simplifies to $S(X_i)$, and hence $X_i^* = X_i S(X_{i+1})$, $i = 1, \dots, d - 1$ and $X_d^* = X_d S(X_1)$.

We now consider enumerating the joint probability $P(X_1^* = x_1, \dots, X_d^* = x_d)$ with the restriction that $\prod_{i=1}^d x_i > 0$.

Lemma 1 *The event $(X_1^* = x_1, \dots, X_d^* = x_d)$ is equivalent to the occurrence of either of the following two events:*

- (i) $(X_1 = x_1, X_2 = x_2 S(x_2), X_3 = x_3 S(x_2 x_3), \dots, X_d = x_d S(\prod_{j=2}^d x_j))$, or
- (ii) $(X_1 = -x_1, X_2 = -x_2 S(x_2), X_3 = -x_3 S(x_2 x_3), \dots, X_d = -x_d S(\prod_{j=2}^d x_j))$.

Proof 2 (Proof of Lemma 1) Note that $|X_k| = |X_k^*|$ for all $k = 1, \dots, d$. Also, in this case, the value of $S(X_2)$ can either be 1, or -1 . We now consider two separate cases.

Case (i): Assume $S(X_2) = 1$, and we get $X_1 = x_1$. Recall that $X_2^* = X_2 S(X_3) = |X_2| S(X_2) S(X_3)$, hence we get $X_2^* = |X_2| S(X_3)$ and $X_2 = X_2^* / S(X_3)$. Now, the following holds:

$$X_2^* = x_2 \Rightarrow |X_2| S(X_3) = x_2 \Rightarrow |x_2| S(X_3) = x_2 \Rightarrow S(X_3) = \frac{x_2}{|x_2|} = S(x_2).$$

Then $X_2^* = X_2 S(X_3) \Rightarrow X_2 = \frac{X_2^*}{S(X_3)} = X_2^* S(X_3) = x_2 S(x_2)$. Here, we use the fact that $S(u) = 1/S(u)$.

Again, $X_3^* = X_3 S(X_4) = |X_3| S(X_3) S(X_4)$. Using the expression of $S(X_3)$ derived above, we obtain

$$X_3^* = x_3 \Rightarrow |X_3| S(X_3) S(X_4) = x_3 \Rightarrow |x_3| S(x_2) S(X_4) = x_3.$$

Therefore,

$$S(X_4) = \frac{x_3}{|x_3| S(x_2)} = \frac{S(x_3)}{S(x_2)} = S(x_3) S(x_2) = S(x_2 x_3).$$

Note that $S(xy) = S(x)S(y)$. Then

$$X_3^* = X_3 S(X_4) \Rightarrow X_3 = \frac{X_3^*}{S(X_4)} = X_3^* S(X_4) = x_3 S(x_2 x_3).$$

Further, $X_4^* = X_4 S(X_5) = |X_4| S(X_4) S(X_5)$ and using the expression of $S(X_4)$ obtained above we have, $X_4^* = x_4 \Rightarrow |X_4| S(X_4) S(X_5) = x_4 \Rightarrow |x_4| S(x_2 x_3) S(X_5)$. This implies that

$$S(X_5) = \frac{x_4}{|x_4| S(x_2 x_3)} = \frac{S(x_4)}{S(x_2 x_3)} = S(x_2 x_3 x_4).$$

Therefore, $X_4^* = X_4 S(X_5) \Rightarrow X_4 = \frac{X_4^*}{S(X_5)} = \frac{x_4}{S(x_2 x_3 x_4)} = x_4 S(x_2 x_3 x_4)$. Proceeding in a similar fashion, we obtain $X_d = x_d S(x_2 \cdots x_d)$.

Case (ii): The proof follows by taking $S(X_2) = -1$, and repeating the line of arguments stated above for **Case (i)**.

Now, we can enumerate the quantity $P(X_1^* = x_1, X_2^* = x_2, \dots, X_d^* = x_d)$ as follows

$$\begin{aligned} &P(X_1^* = x_1, X_2^* = x_2, \dots, X_d^* = x_d) \\ &= P(X_1 S(X_2) = x_1, X_2 S(X_3) = x_2, \dots, X_d S(X_1) = x_d) \\ &= P\left(X_1 = -x_1, X_2 = -x_2 S(x_2), \dots, X_i = -x_i S\left(\prod_{j=1}^i x_j\right), \dots, X_d = -x_d S\left(\prod_{j=1}^d x_j\right)\right) \\ &\quad + P\left(X_1 = x_1, X_2 = x_2 S(x_2), \dots, X_i = x_i S\left(\prod_{j=1}^i x_j\right), \dots, X_d = x_d S\left(\prod_{j=1}^d x_j\right)\right) \\ &= P(X_1 = -x_1) P(X_2 = -x_2 S(x_2)) \cdots P\left(X_i = -x_i S\left(\prod_{j=1}^i x_j\right)\right) \cdots P\left(X_d = -x_d S\left(\prod_{j=1}^d x_j\right)\right) \\ &\quad + P(X_1 = x_1) P(X_2 = x_2 S(x_2)) \cdots P\left(X_i = x_i S\left(\prod_{j=1}^i x_j\right)\right) \cdots P\left(X_d = x_d S\left(\prod_{j=1}^d x_j\right)\right) \end{aligned}$$

[Using independence]

$$= P(X_1 = x_1) \cdots P(X_i = x_i) \cdots P(X_d = x_d) + P(X_1 = x_1) \cdots P(X_i = x_i) \cdots P(X_d = x_d)$$

[Using symmetry about 0, and the fact that $S(\cdot) = \pm 1$]

$$= 2P(X_1 = x_1, \dots, X_d = x_d).$$

Hence in this case, the joint pmf of $(X_1^*, \dots, X_d^*)^T$ is given by

$$q^*(\mathbf{x}) = 2q(\mathbf{x}).$$

This completes the proof of the second part.

Combining these two cases, we may write the joint pmf of $(X_1^*, \dots, X_d^*)^T$ as follows:

$$q^*(\mathbf{x}) = q(\mathbf{x}) \mathbb{I}_{\{0\}}\left(\prod_{i=1}^d x_i\right) + 2q(\mathbf{x}) \mathbb{I}_{(0, \infty)}\left(\prod_{i=1}^d x_i\right). \tag{5}$$

This completes the proof. □

Proof 3 (Proof of Theorem 2) First we consider the univariate distributions, namely, when $d = 1$ and compute the probability $P(X_t^* = x_t)$, denoted by $q^*(x_t)$ for a fixed $t \in \{1, \dots, d\}$. Now, suppose that $x_t = 0$. Note that $X_t^* = X_t S_{t+1,t} = 0 \Leftrightarrow X_t = 0$ since $S_{t+1,t} = 0$ also requires X_t to be zero. We trivially have the reverse, i.e., $X_t = 0 \Rightarrow X_t^* = 0$. So, we have $X_t^* = 0 \iff X_t = 0$ and hence $P(X_t^* = 0) = P(X_t = 0)$. Thus, we obtain $q^*(x_t) = q(x_t)$ for any $t = 1, \dots, d$.

On the other hand, if $x_t \neq 0$, then $S_{t+1,t}$ simplifies to the following function

$$S_{t+1,t} = \begin{cases} -1, & \text{if } X_{t+1} \leq 0, \\ 1, & \text{if } X_{t+1} > 0, \end{cases}$$

and we get

$$\begin{aligned} P(X_t^* = x_t) &= P(X_t S_{t+1,t} = x_t) \\ &= P(X_t = x_t, S_{t+1,t} = 1) + P(X_t = -x_t, S_{t+1,t} = -1) \\ &= P(X_t = x_t, X_{t+1} > 0) + P(X_t = -x_t, X_{t+1} \leq 0) \\ &= P(X_t = x_t)P(X_{t+1} > 0) + P(X_t = -x_t)P(X_{t+1} \leq 0) \\ &= P(X_t = x_t)\{P(X_{t+1} > 0) + P(X_{t+1} \leq 0)\} \\ &= P(X_t = x_t). \end{aligned}$$

Thus, in all these cases, we obtain

$$q^*(x_t) = q(x_t) \text{ for } t = 1, \dots, d. \tag{6}$$

Define the following random vectors:

$$\mathbf{X}_{(-1)} = (X_2, \dots, X_d)^T, \mathbf{X}_{(-t)} = (X_1, \dots, X_{t-1}, X_{t+1}, \dots, X_d)^T, \mathbf{X}_{(-d)} = (X_1, \dots, X_{d-1})^T \text{ and}$$

$$\mathbf{X}_{(-1)}^* = (X_2^*, \dots, X_d^*)^T, \mathbf{X}_{(-t)}^* = (X_1^*, \dots, X_{t-1}^*, X_{t+1}^*, \dots, X_d^*)^T, \mathbf{X}_{(-d)}^* = (X_1^*, \dots, X_{d-1}^*)^T,$$

$$\text{and the sets } I_{(-1)} = \{2, \dots, d\}, I_{(-t)} = \{1, \dots, t-1, t+1, \dots, d\}, I_{(-d)} = \{1, \dots, d-1\}.$$

Let $\mathbf{x}_{(-t)} = (x_1, \dots, x_{t-1}, x_{t+1}, \dots, x_d)^T$, and $q(\mathbf{x}_{(-t)}) = P(\mathbf{X}_{(-t)} = \mathbf{x}_{(-t)})$ for $t = 1, \dots, d$. We want to compute the probability $P(\mathbf{X}_{(-t)}^* = \mathbf{x}_{(-t)})$, and denote it by $q^*(\mathbf{x}_{(-t)})$. Further, we denote the joint probability $P(\mathbf{X}_{(-t)} = \mathbf{x}_{(-t)}, X_t = x_t)$ by $q(\mathbf{x}_{(-t)}, x_t)$, which is nothing but $P(X_1 = x_1, \dots, X_d = x_d)$, and the joint probability $P(\mathbf{X}_{(-t)}^* = \mathbf{x}_{(-t)}, X_t^* = x_t)$ by $q^*(\mathbf{x}_{(-t)}, x_t)$, which is nothing but $P(X_1^* = x_1, \dots, X_d^* = x_d)$. We now consider three separate cases.

Case I: $\prod_{i \in I_{(-t)}} x_i = 0$

Note that $\prod_{i \in I_{(-t)}} x_i = 0 \Rightarrow \prod_{i=1}^d x_i = 0$. Hence, we get $q^*(\mathbf{x}) = q(\mathbf{x}) = q(\mathbf{x}_{(-t)})q(x_t)$. This now gives the following:

$$q^*(\mathbf{x}_{(-t)}) = \sum_{x_t \in \mathbb{S}} q^*(\mathbf{x}) = q(\mathbf{x}_{(-t)}) \sum_{x_t \in \mathbb{S}} q(x_t) = q(\mathbf{x}_{(-t)}) = \prod_{i \in I_{(-t)}} q(x_i) = \prod_{i \in I_{(-t)}} q^*(x_i) \text{ [Using (6)].}$$

Case II: $\prod_{i \in I_{(-t)}} x_i > 0$

Consider three separate sub-cases

(i) If $x_t = 0$, then $\prod_{i=1}^d x_i = 0$, and $q^*(\mathbf{x}) = q(\mathbf{x}) = q(\mathbf{x}_{(-t)})q(0)$. So, $q^*(\mathbf{x}_{(-t)}, 0) = q(\mathbf{x}_{(-t)})q(0)$.

(ii) If $x_t > 0$, then $\prod_{i=1}^d x_i > 0$. Thus, $q^*(\mathbf{x}) = 2q(\mathbf{x}) = 2q(\mathbf{x}_{(-t)})q(x_t)$. Hence, $q^*(\mathbf{x}_{(-t)}, x_t) = 2q(\mathbf{x}_{(-t)}) \sum_{t: x_t > 0} q(x_t)$.

(iii) If $x_t < 0$, then $\prod_{i=1}^d x_i < 0$, which cannot happen by the construction of X_i^* 's. Hence, $q^*(\mathbf{x}_{(-t)}, x_t) = 0$.

Considering these three sub-cases, we have

$$\begin{aligned}
 q^*(\mathbf{x}_{(-t)}) &= \sum_{t:x_t < 0} q^*(\mathbf{x}_{(-t)}, x_t) + q^*(\mathbf{x}_{(-t)}, 0) + \sum_{t:x_t > 0} q^*(\mathbf{x}_{(-t)}, x_t) \\
 &= 0 + q(\mathbf{x}_{(-t)})q(0) + 2q(\mathbf{x}_{(-t)}) \sum_{t:x_t > 0} q(x_t) \\
 &= q(\mathbf{x}_{(-t)}) \sum_{x_t \in \mathbb{S}} q(x_t) \text{ [Using symmetry about 0]} \\
 &= q(\mathbf{x}_{(-t)}) = \prod_{i \in I_{(-t)}} q(x_i) = \prod_{i \in I_{(-t)}} q^*(x_i) \text{ [Using (6)].}
 \end{aligned}$$

Case III: $\prod_{i \in I_{(-t)}} x_i < 0$

Again, consider three separate sub-cases

(i) If $x_t = 0$, then $\prod_{i=1}^d x_i = 0$. Hence, $q^*(\mathbf{x}) = q(\mathbf{x}) = q(\mathbf{x}_{(-t)})q(0)$, i.e., $q^*(\mathbf{x}_{(-t)}, 0) = q(\mathbf{x}_{(-t)})q(0)$.

(ii) If $x_t < 0$, then $\prod_{i=1}^d x_i > 0$. Hence, $q^*(\mathbf{x}) = 2q(\mathbf{x}) = 2q(\mathbf{x}_{(-t)})q(x_t)$, i.e., $q^*(\mathbf{x}_{(-t)}, x_t) = 2q(\mathbf{x}_{(-t)}) \sum_{t:x_t < 0} q(x_t)$.

(iii) If $x_t > 0$, then $\prod_{i=1}^d x_i < 0$. Hence, $q^*(\mathbf{x}_{(-t)}, x_t) = 0$.

Combining these three sub-cases, we obtain

$$\begin{aligned}
 q^*(\mathbf{x}_{(-t)}) &= \sum_{x_t < 0} q^*(\mathbf{x}_{(-t)}, x_t) + q^*(\mathbf{x}_{(-t)}, 0) + \sum_{x_t > 0} q^*(\mathbf{x}_{(-t)}, x_t) \\
 &= 2q(\mathbf{x}_{(-t)}) \sum_{t:x_t < 0} q(x_t) + q(\mathbf{x}_{(-t)})q(0) + 0 \\
 &= q(\mathbf{x}_{(-t)}) \sum_{x_t \in \mathbb{S}} q(x_t) \text{ [Using symmetry about 0]} \\
 &= q(\mathbf{x}_{(-t)}) = \prod_{i \in I_{(-t)}} q(x_i) = \prod_{i \in I_{(-t)}} q^*(x_i) \text{ [Using (6)].}
 \end{aligned}$$

In all these cases, the components of a sub-vector $(X_{k_1}^*, \dots, X_{k_{d'}}^*)^T$ of $(X_1^*, \dots, X_d^*)^T$ are independent of each other. Again, if we take a sub-vector with size less than $d - 1$, then the components will also be independent as all the sub-vectors of size $d - 1$ are component-wise independent. Thus, we can say that any sub-vector $(X_{k_1}^*, \dots, X_{k_{d'}}^*)^T$ of $(X_1^*, \dots, X_d^*)^T$ with $d' < d$ is component-wise independent. Now,

$$\begin{aligned}
 &P(X_{k_1}^* = x_1, X_{k_2}^* = x_2, \dots, X_{k_{d'}}^* = x_{d'}) \\
 &= P(X_{k_1}^* = x_1)P(X_{k_2}^* = x_2) \cdots P(X_{k_{d'}}^* = x_{d'}) \text{ [Using independence]} \\
 &= P(X_{k_1} = x_1)P(X_{k_2} = x_2) \cdots P(X_{k_{d'}} = x_{d'}) \\
 &\quad \text{[Since the marginal distributions of } X_i^* \text{ and } X_i \text{ are identical for } i = k_1, \dots, k_{d'}\text{]} \\
 &= P(X_{k_1} = x_1, X_{k_2} = x_2, \dots, X_{k_{d'}} = x_{d'}) \text{ [Again using independence].}
 \end{aligned}$$

Hence, the joint distribution of $(X_{k_1}^*, \dots, X_{k_{d'}}^*)^T$ is same as that of $(X_{k_1}, \dots, X_{k_{d'}})^T$. \square

Proof 4 (Proof of Theorem 3) Let us denote the joint distribution of (A, B) by the pmf $G(a, b)$ with $a, b \in \mathbb{S}$. Recall that we have assumed A, B to be independent of \mathbf{X}_d^* . So, the conditional distribution of \mathbf{Y}_d^* given $(A = a, B = b)$ is same as its unconditional distribution. We will use this fact throughout the proof of this theorem.

The joint distribution of \mathbf{Y}_d^* is given by

$$\begin{aligned}
 r^*(y_d) &= \sum_{a,b \in \mathbb{S}} P(Y_d^* = y_d | a, b) G(a, b) \\
 &= \sum_{a,b \in \mathbb{S}} P\left(X_d^* = \frac{y_d - a \mathbf{1}_d}{b}\right) G(a, b) \\
 &= \sum_{a,b \in \mathbb{S}} q^*\left(\frac{y_d - a \mathbf{1}_d}{b}\right) G(a, b) \\
 &= \sum_{a,b \in \mathbb{S}} \left\{ q\left(\frac{y_d - a \mathbf{1}_d}{b}\right) \mathbb{I}\left(\prod_{i=1}^d \frac{y_i - a}{b} = 0\right) + 2q\left(\frac{y_d - a \mathbf{1}_d}{b}\right) \mathbb{I}\left(\prod_{i=1}^d \frac{y_i - a}{b} > 0\right) \right\} G(a, b).
 \end{aligned}$$

Recall the expression of $q^*(\mathbf{x})$ from Eq. (5).

Define $\mathbf{Y}_{(-1)}^* = (Y_2^*, Y_3^*, \dots, Y_d^*)^T$ with $I_{(-1)} = \{2, 3, \dots, d\}$; $\mathbf{Y}_{(-t)}^* = (Y_1^*, \dots, Y_{t-1}^*, Y_{t+1}^*, \dots, Y_d^*)^T$ with $I_{(-t)} = \{1, \dots, t-1, t+1, \dots, d\}$, $t = 2, \dots, (d-1)$; and $\mathbf{Y}_{(-d)}^* = (Y_1^*, Y_2^*, \dots, Y_{d-1}^*)^T$ with $I_{(-d)} = \{1, 2, \dots, d\}$. Also, define $\mathbf{y}_{(-t)} = (y_1, \dots, y_{t-1}, y_{t+1}, \dots, y_d)$. Consider

$$r^*(\mathbf{y}_{(-t)}) = \sum_{k \in S_*} \sum_{a,b \in \mathbb{S}} P(Y_1^* = y_1, \dots, Y_t^* = k, \dots, Y_d^* = y_d | a, b) G(a, b).$$

Now, for any fixed $a, b \in \mathbb{S}$, we have

$$P(Y_1^* = y_1, \dots, Y_t^* = k, \dots, Y_d^* = y_d) = P\left(X_1^* = \frac{y_1 - a}{b}, \dots, X_t^* = \frac{k - a}{b}, \dots, X_d^* = \frac{y_d - a}{b}\right).$$

Then,

$$\begin{aligned}
 r^*(\mathbf{y}_{(-t)}) &= \sum_{a,b \in \mathbb{S}} \sum_{k \in S_*} P\left(X_1^* = \frac{y_1 - a}{b}, \dots, X_t^* = \frac{k - a}{b}, \dots, X_d^* = \frac{y_d - a}{b}\right) G(a, b) \\
 &= \sum_{a,b \in \mathbb{S}} P\left(X_1^* = \frac{y_1 - a}{b}, \dots, X_{t-1}^* = \frac{y_{t-1} - a}{b}, X_{t+1}^* = \frac{y_{t+1} - a}{b}, \dots, X_d^* = \frac{y_d - a}{b}\right) G(a, b) \\
 &= \sum_{a,b \in \mathbb{S}} q^*\left(\frac{\mathbf{y}_{(-t)} - a \mathbf{1}_{d-1}}{b}\right) G(a, b) \quad [\text{Recall the expression of } q^*(\mathbf{x}_{(-t)})] \\
 &= \sum_{a,b \in \mathbb{S}} q\left(\frac{\mathbf{y}_{(-t)} - a \mathbf{1}_{d-1}}{b}\right) G(a, b) = \sum_{a,b \in \mathbb{S}} \left\{ \prod_{i \in I_{(-t)}} q\left(\frac{y_i - a}{b}\right) \right\} G(a, b).
 \end{aligned}$$

Thus, the random vectors $\mathbf{Y}_{(-t)}^*$ for $t = 2, \dots, d-1$ possess the same joint distribution. Similarly, one may argue that the random vectors $\mathbf{Y}_{(-1)}^*$ and $\mathbf{Y}_{(-d)}^*$ also follow the same joint distribution as well. Further, we can argue that any sub-vector $(Y_{k_1}^*, Y_{k_2}^*, \dots, Y_{k_{d'}}^*)^T$ of $(Y_1^*, Y_2^*, \dots, Y_d^*)^T$, where $d' < d$, has the same joint distribution. \square

Acknowledgements

We thank the Reviewers and Editors for comments that improved the paper.

Funding

This research was supported by the King Abdullah University of Science and Technology (KAUST).

Availability of data and materials

Not applicable.

Authors' contributions

SG, SD, MG contributed equally to the research. All authors read and approved the final manuscript.

Competing interests

No competing interests.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Publisher's Note

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

Author details

¹Department of Mathematics and Statistics, Indian Institute of Technology, Kanpur 208016, India. ²CEMSE Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia.

Received: 13 February 2017 Accepted: 31 May 2017

Published online: 05 July 2017

References

- Alzaatreh, A, Lee, C, Famoye, F: On the discrete analogues of continuous distributions. *Stat. Methodol.* **9**, 589–603 (2012)
- Azzalini, A: With the collaboration of A. Capitanio. *The Skew-Normal and Related Families*. Cambridge University Press, UK (2014)
- Azzalini, A, Regoli, G: Modulation of symmetry for discrete variables and some extensions. *Stat* **3**, 56–67 (2014)
- Bairamov, I, Gultekin, OE: Discrete distributions connected with the bivariate Binomial. *Hacetatepe J. Math. Stat.* **39**, 109–120 (2010)
- Chakraborty, S, Chakravarty, D: A new discrete probability distribution with integer support on $(-\infty, \infty)$. *Commun. Stat. Theory Methods*. **45**, 492–505 (2016)
- Conway, DA: Multivariate distributions with specified marginals (1979). Available at <https://statistics.stanford.edu/sites/default/files/OLK%20NSF%20145.pdf>. Accessed 10 June 2017
- Cuadras, CM: Probability distributions with given multivariate marginals and given dependence structure. *J. Multivar. Anal.* **42**, 51–66 (1992)
- Dahiya, R, Korwar, R: On characterizing some bivariate discrete distributions by linear regression. *Sankhyā: Indian J. Stat. Series A*. **39**, 124–129 (1977)
- Dutta, S, Genton, MG: A non-Gaussian multivariate distribution with all lower-dimensional Gaussians and related families. *J. Multivar. Anal.* **132**, 82–93 (2014)
- Fréchet, M: Sur les tableaux de corrélation dont les marges sont données. *Ann. Univ. de Lyon Sect. A, Series 3*. **14**, 53–77 (1951)
- Hoeffding, W: Massababinvariate korrelations-theorie. *Schriften Math. Inst. Univ. Berlin*. **5**, 181–233 (1940)
- Johnson, NL, Kotz, S, Balakrishnan, N: *Discrete Multivariate Distributions*. Wiley, New York (1997)
- Nguyen, TT, Gupta, AK, Wang, Y: A characterization of certain discrete exponential families. *Ann. Inst. Stat. Math.* **48**, 573–576 (1996)
- Roy, D: The discrete normal distribution. *Commun. Stat. Theory Methods*. **32**, 1871–1883 (2003)
- Ruiz, JM, Navarro, J: Characterization of discrete distributions using expected values. *Stat Papers*. **36**, 237–252 (1995)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com