

ON THE LIMIT BEHAVIOUR OF FINITE-SUPPORT BIVARIATE DISCRETE PROBABILITY DISTRIBUTIONS UNDER ITERATED PARTIAL SUMMATIONS

L. LEŠŠOVÁ AND J. MAČUTEK

ABSTRACT. One type of bivariate partial-sums discrete probability distributions is defined. It is shown that in analogy to the univariate case, there is one-to-one relation between the summations and bivariate discrete distributions, namely for each partial summation, there is one and only one distribution which remains unchanged under the summation. The question of the existence of a limit distribution for iterated partial summations is solved for finite-support bivariate distributions which satisfy conditions under which the power method (known from matrix theory) can be used. Examples of both a converging sequence of distributions with its limit and an oscillating sequence which does not converge are presented.

1. INTRODUCTION

Let $\{P_x^{(1)}\}_{x=0}^\infty$ and $\{P_x^*\}_{x=0}^\infty$ be probability mass functions of two univariate discrete probability distributions defined on nonnegative integers. The distribution $\{P_x^{(1)}\}_{x=0}^\infty$ (the descendant distribution) is a partial-sums distribution created from $\{P_x^*\}_{x=0}^\infty$ (the parent distribution) if

$$(1) \quad P_x^{(1)} = c_1 \sum_{j=x}^{\infty} g(j) P_j^*, \quad x = 0, 1, 2, \dots,$$

where c_1 is a normalization constant and $g(j)$ a real function (we note that also another type of the partial summations is mentioned in [7], but it is not the subject of this study). Several examples of partial summations – for different choices of $g(j)$ – are mentioned in the comprehensive monograph by Johnson et al. [2]. An extensive survey of pairs of parents and descendants was provided in [9]. According to [7], for each function $g(j)$, there is one and only one distribution which remains unchanged under the summation (i.e., if function $g(x)$ is fixed, the distributions $\{P_x^{(1)}\}_{x=0}^\infty$ and $\{P_x^*\}_{x=0}^\infty$ are identical). More detailed analyses (e.g.

Received July 7, 2019; revised December 13, 2019.

2010 *Mathematics Subject Classification.* Primary 60E05.

Key words and phrases. Discrete probability distributions; partial-sums distributions; convergence.

relations between probability generating functions of the parent and descendant distributions) can also be found in [7].

Partial summations from (1) can be applied iteratively. Take $\{P_x^{(1)}\}_{x=0}^\infty$, i.e., the descendant distribution from (1), as the parent, with function $g(j)$ remaining unaltered. We obtain the descendant of the second generation

$$P_x^{(2)} = c_2 \sum_{j=x}^{\infty} g(j) P_j^{(1)}, \quad x = 0, 1, 2, \dots,$$

and, repeatedly applying the partial summation, for any $k \in \mathbb{N}$, the descendant of the k -th generation

$$(2) \quad P_x^{(k)} = c_k \sum_{j=x}^{\infty} g(j) P_j^{(k-1)}, \quad x = 0, 1, 2, \dots,$$

c_2, c_k being normalization constants.

The question whether the sequence of the descendant distributions from (2) has a limit for $k \rightarrow \infty$ and a constant function $g(j)$ was investigated in [8]. In this case, the answer is positive for a wide class of parent distributions, with the limit distribution being geometric (we note that the geometric distribution is the only distribution which is invariant with respect to the partial summations with a constant function $g(j)$, see [10] and [7]). A solution – albeit not a general one – of the problem under condition that the parent distribution has a finite support was presented in [4].

2. BIVARIATE PARTIAL-SUMS DISTRIBUTIONS

Research on partial-sums distributions has been so far almost exclusively dedicated to univariate distributions (see [11] and references therein). The only note on the bivariate (and r -variate) partial-sums distributions can be found in Kotz and Johnson [5], who more or less restrict themselves to a suggestion to study not only univariate, but also multivariate cases.

Univariate partial-sums distributions from (1) can be naturally generalized to two dimensions as follows. Let $\{P_{x,y}^*\}_{x,y=0}^\infty$ and $\{P_{x,y}^{(1)}\}_{x,y=0}^\infty$ be bivariate discrete distributions defined on nonnegative integers and let $g(x, y)$ be a real function. Then $\{P_{x,y}^{(1)}\}_{x,y=0}^\infty$ is the descendant of the parent $\{P_{x,y}^*\}_{x,y=0}^\infty$ if

$$(3) \quad P_{x,y}^{(1)} = c_1 \sum_{i=x}^{\infty} \sum_{j=y}^{\infty} g(i, j) P_{i,j}^*.$$

Relations between properties of the parent and descendant distributions from (3), such as, e.g., moments or probability generating functions, can be easily obtained using the same mathematical apparatus as for univariate partial-sums distributions (see [7]).

In analogy to univariate partial-sums distributions (see [7]), for each bivariate discrete distribution there is one and only one function $g(x, y)$ (and consequently,

one and only one partial summation) which leaves the distribution unchanged (i.e., the parent and the descendant distributions are identical).

Lemma 2.1. *A bivariate discrete distribution $\{P_{x,y}\}_{x,y=0}^{\infty}$ with $P_{x,y} \neq 0$, for all x, y is invariant under partial summation (3), i.e.,*

$$P_{x,y} = \sum_{i=x}^{\infty} \sum_{j=y}^{\infty} g(i,j) P_{i,j},$$

if and only if

$$g(x,y) = 1 - \frac{P_{x+1,y} + P_{x,y+1} - P_{x+1,y+1}}{P_{x,y}},$$

for $x, y = 0, 1, 2, \dots$

Proof. Suppose that the parent and the descendant distributions are identical. Using (3), we have

$$\begin{aligned} P_{x,y} &= \sum_{i=x}^{\infty} \sum_{j=y}^{\infty} g(i,j) P_{i,j} \\ (4) \quad &= \sum_{i=x+1}^{\infty} \sum_{j=y+1}^{\infty} g(i,j) P_{i,j} + \sum_{i=x}^{\infty} g(i,y) P_{i,y} + \sum_{j=y}^{\infty} g(x,j) P_{x,j} \\ &\quad - g(x,y) P_{x,y} \\ &= P_{x+1,y+1} + \sum_{i=x}^{\infty} g(i,y) P_{i,y} + \sum_{j=y}^{\infty} g(x,j) P_{x,j} - g(x,y) P_{x,y}, \end{aligned}$$

similarly for $P_{x+1,y}$ and $P_{x,y+1}$, we obtain

$$(5) \quad P_{x+1,y} = \sum_{i=x+1}^{\infty} \sum_{j=y}^{\infty} g(i,j) P_{i,j} = P_{x+1,y+1} + \sum_{i=x}^{\infty} g(i,y) P_{i,y} - g(x,y) P_{x,y}$$

and

$$(6) \quad P_{x,y+1} = \sum_{i=x}^{\infty} \sum_{j=y+1}^{\infty} g(i,j) P_{i,j} = P_{x+1,y+1} + \sum_{j=y}^{\infty} g(x,j) P_{x,j} - g(x,y) P_{x,y}.$$

From (5) and (6), we have

$$(7) \quad \sum_{i=x}^{\infty} g(i,y) P_{i,y} = P_{x+1,y} - P_{x+1,y+1} + g(x,y) P_{x,y}$$

and

$$(8) \quad \sum_{j=y}^{\infty} g(x,j) P_{x,j} = P_{x,y+1} - P_{x+1,y+1} + g(x,y) P_{x,y}.$$

Substituting (7) and (8) into (4), we obtain

$$P_{x,y} - P_{x+1,y} - P_{x,y+1} = -P_{x+1,y+1} + g(x,y) P_{x,y},$$

and consequently,

$$g(x, y) = 1 - f(x, y) = 1 - \frac{P_{x+1,y} + P_{x,y+1} - P_{x+1,y+1}}{P_{x,y}},$$

which completes the proof. \square

The lemma applies analogously also to bivariate distributions on a finite support of size $m \times n$ (one formally considers probabilities $\{P_{x,y}^*\}_{x,y}$ to be zeroes if $x > m$ or $y > n$).

Also the bivariate partial summations from (3) can be applied iteratively. The descendant of the k -th generation is obtained analogously to the univariate case (2), i.e., the k -th descendant is

$$(9) \quad P_{x,y}^{(k)} = c_k \sum_{i=x}^{\infty} \sum_{j=y}^{\infty} g(i, j) P_{i,j}^{(k-1)}.$$

In the next section, we extend the result from [4] on the limit behaviour of some univariate discrete distributions under the iterated partial summations to bivariate distributions. We show that if the parent distribution has a finite support of size $m \times n$, the power method, which is a computational approach to finding matrix eigenvalues and eigenvectors, can in some cases be used to find the limit distribution for $k \rightarrow \infty$ in (9).

3. POWER METHOD AND ITS APPLICATION

The power method (see, e.g., [1]) was suggested as a computational tool which under certain conditions enables to find an approximation of square matrix eigenvalues. The method can be applied if

1. the matrix is diagonalizable (i.e., it has linearly independent eigenvectors, or equivalently, it is similar to a diagonal matrix), and
2. it has a unique dominant eigenvalue (denote the eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$; there exists k such that $|\lambda_k| > |\lambda_i| \ \forall i \neq k$); the eigenvector corresponding to the dominant eigenvalue is the dominant eigenvector.

If a matrix A satisfies the abovementioned conditions, then there exists a non-zero vector x_0 such that the sequence $\{A^k x_0\}_{k=1}^{\infty}$ converges to a multiple of the dominant eigenvector. The iterations may start from any vector x_0 which 1) is not orthogonal to the vector space associated with the dominant eigenvalue and 2) is not such a linear combination of other eigenvectors which does not contain the dominant eigenvector.

While the application of the power method is straightforward for univariate iterated partial summations (see [4]), a bivariate distribution requires an additional step, namely, a vectorization of the probability matrix. The vectorization is, however, a standard operation in matrix theory (see, e.g., [1]). Denote \mathbb{P}^* the parent

distribution from (3), i.e.,

$$\mathbb{P}^* = \begin{pmatrix} P_{0,0}^* & P_{0,1}^* & \cdots & P_{0,n-1}^* \\ P_{1,0}^* & P_{1,1}^* & \cdots & P_{1,n-1}^* \\ \vdots & \vdots & \ddots & \vdots \\ P_{m-1,0}^* & P_{m-1,1}^* & \cdots & P_{m-1,n-1}^* \end{pmatrix}.$$

The vectorization of \mathbb{P}^* yields a vector of probabilities

$$(10) \quad v(\mathbb{P}^*) = (P_{0,0}^*, \dots, P_{m-1,0}^*, P_{0,1}^*, \dots, P_{m-1,1}^*, \dots, P_{0,n-1}^*, \dots, P_{m-1,n-1}^*)^T.$$

Now we construct a square matrix \tilde{G} with $m \times n$ rows and $m \times n$ columns from the values of the function $g(i, j)$ from (3) as

$$\tilde{G} = \begin{pmatrix} g(0,0) & g(1,0) & \cdots & g(m-1,0) & g(0,1) & g(1,1) & \cdots & g(m-1,n-1) \\ 0 & g(1,0) & \cdots & g(m-1,0) & 0 & g(1,1) & \cdots & g(m-1,n-1) \\ 0 & 0 & \cdots & g(m-1,0) & 0 & 0 & \cdots & g(m-1,n-1) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & g(m-1,0) & 0 & \cdots & 0 & g(m-1,n-1) \\ 0 & 0 & \cdots & 0 & g(0,1) & g(1,1) & \cdots & g(m-1,n-1) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & g(m-1,n-1) \end{pmatrix}.$$

Denote D the diagonal matrix consisting of the elements from the main diagonal of matrix \tilde{G} , i.e.,

$$D = \text{diag}(g(0,0), g(1,0), \dots, g(m-1,n-1)),$$

and A the upper triangular matrix of ones with dimensions $m \times m$, i.e.,

$$A = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}_{m \times m}.$$

Then it holds

$$\tilde{G} = \begin{pmatrix} A & A & A & \cdots & A \\ 0 & A & A & \cdots & A \\ 0 & 0 & A & \cdots & A \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & A \end{pmatrix} D.$$

Matrix \tilde{G} is an upper triangular matrix with dimensions $nm \times nm$, each of its columns contains only one particular value of $g(i, j)$ (which occurs several times

in its column). Its main diagonal consists of all values of $g(i, j)$, each of them occurring just once.

The notation established above allows us to write

$$v(\mathbb{P}^{(1)}) = \frac{\tilde{G}v(\mathbb{P}^*)}{\|\tilde{G}v(\mathbb{P}^*)\|_1}$$

and the k -th descendant of \mathbb{P}^* , defined by (9), can be expressed after applying (10), in its vector form as

$$(\mathbb{P}^{(k)}) = \frac{\tilde{G}v(\mathbb{P}^{(k-1)})}{\|\tilde{G}v(\mathbb{P}^{(k-1)})\|_1} = \frac{\tilde{G}^k v(\mathbb{P}^*)}{\|\tilde{G}^k v(\mathbb{P}^*)\|_1}.$$

The assumptions under which the power method converges are satisfied if the parent distribution has a finite support, if the matrix \tilde{G} is diagonalizable and has a unique dominant eigenvalue (as the eigenvalues of the upper triangular matrix are the elements of its main diagonal, the dominant eigenvalue is unique if and only if the greatest absolute value of $g(i, j)$ is unique), and if iterations start from a suitable starting vector \mathbb{P}^* (i.e., from a vector which satisfies conditions from the beginning of this section). Then the sequence of vectors

$$\frac{\tilde{G}v(\mathbb{P}^*)}{\|\tilde{G}v(\mathbb{P}^*)\|_2}, \frac{\tilde{G}^2 v(\mathbb{P}^*)}{\|\tilde{G}^2 v(\mathbb{P}^*)\|_2}, \dots, \frac{\tilde{G}^k v(\mathbb{P}^*)}{\|\tilde{G}^k v(\mathbb{P}^*)\|_2}, \dots$$

converges to the unit dominant eigenvector of matrix \tilde{G} . In other words, if its assumptions are satisfied, the power method ensures that limit

$$\mathbb{P}^{(\infty)} = \lim_{k \rightarrow \infty} \mathbb{P}^{(k)}$$

exists. We obtain the limit distribution by multiplying the dominant unit eigenvector of matrix \tilde{G} by a normalization constant, i.e., the vectorized form of the limit distribution is

$$v(\mathbb{P}^{(\infty)}) = \lim_{k \rightarrow \infty} \frac{v(\mathbb{P}^{(k)})}{\|v(\mathbb{P}^{(k)})\|_1} = \lim_{k \rightarrow \infty} \frac{\tilde{G}^k v(\mathbb{P}^*)}{\|(\tilde{G}^k v(\mathbb{P}^*))\|_1}.$$

4. EXAMPLES

4.1. Limit distribution

Let $N_1, N_2, N_3 \in \mathbb{N}$, $k \in \{1, 2, \dots, N_3\}$ and $N = N_1 + N_2 + N_3$. Vector $\binom{X}{Y}$ has the bivariate inverse hypergeometric distribution (see [3]) if

$$P_{x,y} = \frac{N_3 - k + 1}{N - (x + y + k - 1)} \frac{\binom{N_1}{x} \binom{N_2}{y} \binom{N_3}{k-1}}{\binom{N}{x+y+k-1}}$$

for $x = 0, 1, 2, \dots, N_1$, $y = 0, 1, 2, \dots, N_2$.

We derive such a function $g(i, j)$ which leaves the bivariate inverse hypergeometric distribution unchanged. We choose parameter values $N_1 = N_2 = 2$, $N_3 = 5$, $k = 2$, i.e.,

$$(11) \quad \mathbb{P} = \begin{pmatrix} \frac{5}{18} & \frac{10}{63} & \frac{5}{126} \\ \frac{10}{63} & \frac{10}{63} & \frac{4}{63} \\ \frac{5}{126} & \frac{4}{63} & \frac{5}{126} \end{pmatrix}.$$

According to Lemma 2.1,

$$(12) \quad \begin{aligned} g(i, j) &= 1 - \frac{P_{i+1,j} + P_{i,j+1} - P_{i+1,j+1}}{P_{i,j}} \\ &= \frac{(N_1 - i)(N_2 - j)(i + j + k)(i + j + k + 1)}{(i + 1)(j + 1)(N - k - i - j)(N - k - i - j - 1)} \\ &\quad - \frac{(i + j + k)(N_1 - i)}{(i + 1)(N - i - j - k)} - \frac{(i + j + k)(N_2 - j)}{(j + 1)(N - i - j - k)} + 1, \end{aligned}$$

where $i, j = 0, 1, 2$. For the distribution under consideration, we obtain

$$\tilde{G} = \begin{pmatrix} \frac{3}{7} & \frac{3}{20} & -\frac{3}{5} & \frac{3}{20} & \frac{9}{20} & \frac{3}{8} & -\frac{3}{5} & \frac{3}{8} & 1 \\ 0 & \frac{3}{20} & -\frac{3}{5} & 0 & \frac{9}{20} & \frac{3}{8} & 0 & \frac{3}{8} & 1 \\ 0 & 0 & -\frac{3}{5} & 0 & 0 & \frac{3}{8} & 0 & 0 & 1 \\ 0 & 0 & 0 & \frac{3}{20} & \frac{9}{20} & \frac{3}{8} & -\frac{3}{5} & \frac{3}{8} & 1 \\ 0 & 0 & 0 & 0 & \frac{9}{20} & \frac{3}{8} & 0 & \frac{3}{8} & 1 \\ 0 & 0 & 0 & 0 & 0 & \frac{3}{8} & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{3}{5} & \frac{3}{8} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{3}{8} & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Matrix \tilde{G} is diagonalizable and has a unique dominant eigenvalue, hence conditions under which the power method can be applied are satisfied in this case. We note that the normalized dominant eigenvector of \tilde{G} from this example is $v(\mathbb{P})$, i.e., the vectorized matrix of the bivariate inverse hypergeometric distribution from (11).

If we start from any suitable probability vector, i.e., from one which 1) is not orthogonal to the vector space associated with the dominant eigenvalue and 2) is not such a linear combination of other eigenvectors which does not contain the dominant eigenvector, the iterated partial summations (9) converge to the unit eigenvector corresponding to the dominant eigenvalue of the matrix \tilde{G} . In this case, the dominant eigenvalue is 1, hence the normalized multiple of its corresponding

eigenvector is the limit distribution

$$\mathbb{P}^{(\infty)} = \begin{pmatrix} \frac{5}{18} & \frac{10}{63} & \frac{5}{126} \\ \frac{10}{63} & \frac{10}{63} & \frac{4}{63} \\ \frac{5}{126} & \frac{4}{63} & \frac{5}{126} \end{pmatrix}.$$

We remind that (almost – with the exception of vectors orthogonal to the space of the dominant eigenvalue and of distributions containing zero terms) regardless of the parent distribution, the limit distribution $\mathbb{P}^{(\infty)}$ is the bivariate hypergeometric distribution with the parameters $N_1 = N_2 = 2$, $N_3 = 5$, $k = 2$ which is invariant under the partial summation with $g(i, j)$ from (12). The bivariate discrete uniform distribution with constant $P_{x,y} = \frac{1}{9}$ is one of many parent distributions which converges under iterated partial summations given by (3), with $g(i, j)$ from (12) to the bivariate inverse hypergeometric distribution. The limit distribution plays, in a way, the role of an attractor for this partial summation.

4.2. Oscillation

There are also sequences of descendant distributions which do not converge. Let the parent be

$$\mathbb{P}^* = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & 0 \end{pmatrix}$$

and let the matrix \tilde{G} be

$$\tilde{G} = \begin{pmatrix} -1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

which means that the function $g(i, j)$ from (3) is $g(0, 0) = -1$, $g(0, 1) = 1$, $g(1, 0) = 1$ and $g(1, 1) = 0$. We remind that the matrix does not have a unique dominant eigenvalue, which means that the power method cannot be applied. After the first partial summation, we obtain

$$\mathbb{P}^{(1)} = \begin{pmatrix} 0 & \frac{1}{2} \\ \frac{1}{2} & 0 \end{pmatrix},$$

and after the second summation,

$$\mathbb{P}^{(2)} = \begin{pmatrix} \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & 0 \end{pmatrix},$$

i.e., the distribution identical to the parent \mathbb{P}^* .

Some examples and a discussion on oscillating sequences of univariate partial-sums discrete probability distributions can be found in [6].

Acknowledgment. The research was supported by grant VEGA 2/0054/18 (L. Leššová, J. Mačutek) and by the Comenius University grant UK/334/2019 (L. Leššová).

REFERENCES

1. Golub H. G. and Van Loan Ch. F., *Matrix Computations*, The Johns Hopkins University Press, Baltimore/London, 1996.
2. Johnson N. L., Kemp A. W. and Kotz S., *Univariate Discrete Distributions*, Wiley, Hoboken (NJ), 2005.
3. Johnson N. L., Kotz S. and Balakrishnan N., *Discrete Multivariate Distributions*, Wiley, Hoboken (NJ), 1997.
4. Koščová M., Harman R. and Mačutek J., *Iterated partial summations applied to finite-support discrete distributions*, Math. Slovaca **70** (2020), 489–496.
5. Kotz S. and Johnson N. L., *A note on renewal (partial sums) distributions for discrete variables*, Statist. Probab. Lett. **12** (1991), 229–231.
6. Leššová L., *Oscillating sequences of partial-sums discrete probability distributions*, in: Proceedings from the 21st European Young Statisticians Meeting (B. Milošević, M. Obradović, eds.), Faculty of Mathematics, Belgrade, 2019, 41–45.
7. Mačutek J., *On two types of partial summations*, Tatra Mt. Math. Publ. **26** (2003), 403–410.
8. Mačutek J., *A limit property of the geometric distribution*, Theory Probab. Appl. **50**(2) (2006), 316–319.
9. Wimmer G. and Altmann G., *On the generalization of the STER distribution applied to generalized hypergeometric parents*, Acta Univ. Palack. Olomuc. Fac. Rerum Natur. Math. **39**(1) (2000), 215–247.
10. Wimmer G. and Kalas J., *A characterization of the geometric distribution*, Tatra Mt. Math. Publ. **17** (1999), 325–329.
11. Wimmer G. and Mačutek J., *New integrated view at partial-sums distributions*, Tatra Mt. Math. Publ. **51** (2012), 183–190.

L. Leššová, Department of Applied Mathematics and Statistics, Faculty of Mathematics, Physics and Informatics, Comenius University in Bratislava, Mlynská dolina, 842 48 Bratislava, Slovakia,
e-mail: livia.lessova@fmph.uniba.sk

J. Mačutek, Mathematical Institute Slovak Academy of Sciences, Štefánikova 49, 814 73 Bratislava, Slovakia; Department of Mathematics, Faculty of Natural Sciences, Constantine the Philosopher University in Nitra, Trieda Andreja Hlinku 1, 949 74 Nitra, Slovakia,
e-mail: jmacutek@yahoo.com