

On the correlation structures of multivariate skew-normal distribution

ENE KÄÄRIK, MEELIS KÄÄRIK, AND INGER-HELEN MAADIK

ABSTRACT. Skew-normal distribution is an extension of the normal distribution where the symmetry of the normal distribution is distorted with an extra parameter. A multivariate skew-normal distribution has been parametrized differently to stress different aspects and constructions behind the distribution. There are several possible parametrizations available to define the skew-normal distribution. The current most common parametrization is through Ω and α , as an alternative, parametrization through Ω and δ can be used if straightforward relation to marginal distributions is of interest. The main problem with $\{\Omega, \delta\}$ -parametrization is that the vector δ cannot be chosen independently of Ω . This motivated us to investigate what are the possibilities of choosing δ under different correlation structures of Ω . We also show how the assumptions on structure of δ and Ω affect the asymmetry parameter α and correlation matrix \mathbf{R} of corresponding skew-normal random variable.

1. Background

We consider multivariate correlated data in broader sense including repeated measurements over time (longitudinal data) or space. To handle the dependence between measurements, it is essential to construct a full probability model that integrates the marginal distributions and the correlation coherently. For the multivariate normal distribution, such approach has been extensively investigated (see, for example, [23]). There are also many comprehensively examined analyses using simple correlation structure (uniform correlation and/or serial correlation).

However, in many cases, the assumption of normality might be unrealistic, especially for skewed data. Data may be skewed naturally (consider,

Received December 3, 2015.

2010 *Mathematics Subject Classification.* 60E15, 62H05, 60E05.

Key words and phrases. Multivariate skew-normal distribution, compound symmetry correlation structure, autoregressive correlation structure.

<http://dx.doi.org/10.12697/ACUTM.2016.20.07>

e.g., the environmental pollution data in [11]) or because of conditioning while gathering the data (for example medical data of patients with certain diagnosis [13]). Conditioning causes a situation where (although the overall underlying distribution is normal) the gathered data actually follows a skew-normal distribution.

The history of skew-normal distribution is not very long, but the development has been rapid and intensive. The first systematic representation of univariate skew-normal distribution is given by Azzalini [5]. The multivariate skew-normal distribution is introduced by Azzalini and Dalla Valle [10] and structured by Azzalini and Capitanio [7]. Extensions to the class are proposed by several authors [2, 8, 6, 16, 12]. Multivariate skew-normal distribution also has several different equivalent parametrizations, each of which offers advantages depending on a particular problem.

To fully understand the mechanics of skewing, it is important to study the underlying dependence structures. A related widely studied example is the data of Australian athletes collected by Australian Institute of Sport (AIS-data). Various skewed models have been applied by different authors, main focus has been on the bivariate skew-normal distribution [10, 4, 3, 16]. Similar problems in more general setup are addressed in [7], where a 5-dimensional skew-normal model is applied to a diabetes data. Arellano-Valle et al. developed a skew-normal mixed model for fitting longitudinal cholesterol levels data in Framingham heart study [1]. An extensive list of applications is given in [15, 6, 9].

Despite the growing number of publications, there remain several things to be explored. In this paper we examine the behaviour of the correlation matrices of multivariate skew-normal random variables in different specific situations starting from the more simple correlation structures like compound symmetry and autoregressive structure.

2. Framework

The multivariate skew-normal distribution has received considerable attention over the last years. Nevertheless, even the “classical” multivariate skew-normal distribution, introduced by Azzalini and Dalla Valle in 1996, can be parametrized in many different ways starting from the initial $\{\Psi, \lambda\}$ -parametrization to the currently prevalent $\{\Omega, \alpha\}$ -parametrization (see, e.g., [7], or, for a comprehensive analysis, [9]). Yet, because of the construction of α , the $\{\Omega, \alpha\}$ -parametrization does not seem to be the most appropriate choice if the connection to the parameters of corresponding marginal distributions is of importance (see, e.g., [17, 19]). We have shown that the $\{\Omega, \delta\}$ -parametrization is a reasonable choice, especially in the situation where direct connection with marginal parameters is of importance [20].

We use $\{\mathbf{\Omega}, \boldsymbol{\delta}\}$ -parametrization and analyze properties of the corresponding skew-normal random variable in different situations. We will reveal how the choice of correlation structure in $\mathbf{\Omega}$ affects possible choices for $\boldsymbol{\delta}$ and how the vector $\boldsymbol{\alpha}$ is expressed in these special cases (recall that the vector $\boldsymbol{\alpha}$ is determined by the vector $\boldsymbol{\delta}$ and the concentration matrix $\mathbf{\Omega}^{-1}$). The correlation matrix \mathbf{R} and the region of possible correlation coefficients of the corresponding skew-normal random variable are studied as well.

Let us first recall the definition of a multivariate skew-normal random variable through $\{\mathbf{\Omega}, \boldsymbol{\delta}\}$ -parametrization.

Definition 1. Let $\mathbf{\Omega}$ be a positive definite $k \times k$ correlation matrix and let $\boldsymbol{\delta}$ be a k -dimensional vector such that $\boldsymbol{\delta}^T \mathbf{\Omega}^{-1} \boldsymbol{\delta} < 1$. We say that a random variable $\mathbf{Z} = (Z_1, \dots, Z_k)^T$ has the k -variate skew-normal distribution with skewness parameter $\boldsymbol{\delta}$, and write $\mathbf{Z} \sim SN(\mathbf{\Omega}, \boldsymbol{\delta})$, if its probability density function is given by

$$f(\mathbf{z}; \mathbf{\Omega}, \boldsymbol{\delta}) = 2\phi_k(\mathbf{z}; \mathbf{\Omega})\Phi\left(\frac{\boldsymbol{\delta}^T \mathbf{\Omega}^{-1} \mathbf{z}}{\sqrt{1 - \boldsymbol{\delta}^T \mathbf{\Omega}^{-1} \boldsymbol{\delta}}}\right), \quad \mathbf{z} \in \mathbb{R}^k, \quad (1)$$

where ϕ_k is the probability density function of the k -variate normal distribution with standard normal marginals and correlation matrix $\mathbf{\Omega}$, and Φ is the univariate standard normal distribution function.

We also recall the following conditional representation of skew-normal variable (see, e.g., [10], [14] or [20] for $\{\mathbf{\Omega}, \boldsymbol{\delta}\}$ -parametrization). If we have a k -dimensional random variable $\mathbf{X} \sim N(\mathbf{0}, \mathbf{\Omega})$, a one-dimensional random variable $X_0 \sim N(0, 1)$ and a vector $\boldsymbol{\delta} = (\delta_1, \dots, \delta_k)^T$ such that

$$\boldsymbol{\delta}^T \mathbf{\Omega}^{-1} \boldsymbol{\delta} < 1 \quad (2)$$

and $\boldsymbol{\delta}$ is the vector of correlations between \mathbf{X} and X_0 , then a random variable $\mathbf{Z} = \mathbf{X} | X_0 > 0$ has the skew-normal distribution $\mathbf{Z} \sim SN(\mathbf{\Omega}, \boldsymbol{\delta})$.

Considering the interpretation of $\boldsymbol{\delta}$ as a vector of correlations it is useful to introduce the matrices

$$\mathbf{\Omega}^* = \begin{pmatrix} 1 & \boldsymbol{\delta}^T \\ \boldsymbol{\delta} & \mathbf{\Omega} \end{pmatrix}, \quad \mathbf{\Omega}_* = \begin{pmatrix} \mathbf{\Omega} & \boldsymbol{\delta} \\ \boldsymbol{\delta}^T & 1 \end{pmatrix},$$

where $\mathbf{\Omega}^*$ is the correlation matrix for $(X_0, \mathbf{X})^T$ and $\mathbf{\Omega}_*$ is the correlation matrix for $(\mathbf{X}, X_0)^T$.

In different applications, these matrices have important role and meaning. For example, in repeated measurement studies, the former has intuitive meaning if we consider conditioning by baseline and the latter in imputation of dropouts.

In consequence, we have four correlation matrices related to a skew-normal random variable

- $\mathbf{\Omega} = (\omega_{ij})$ – the (scale) parameter matrix of the skew-normal distribution (see, e.g., (1)),
- $\mathbf{\Omega}^*, \mathbf{\Omega}_*$ – denoted above,
- $\mathbf{R} = (r_{ij})$ – the correlation matrix of a skew-normal random variable \mathbf{Z} ,

and it is now quite interesting to analyze the nature of restrictions on elements of these matrices.

The general formula for computing the correlation matrix $\mathbf{R} = (r_{ij})$ of a skew-normal random variable $\mathbf{Z} = (Z_1, \dots, Z_k)$ is given by Azzalini and Dalla Valle [10] as follows:

$$r_{ij} = \text{corr}(Z_i, Z_j) = \frac{\omega_{ij} - \frac{2}{\pi}\delta_i\delta_j}{\sqrt{1 - \frac{2}{\pi}\delta_i^2}\sqrt{1 - \frac{2}{\pi}\delta_j^2}}, \quad i, j = 1, \dots, k. \quad (3)$$

An important special case is when all marginals have the same distribution $SN(\delta)$, that means all the components of $\boldsymbol{\delta}$ are identical, i.e., $\boldsymbol{\delta} = (\delta, \dots, \delta)^T$. Then the last formula simplifies to

$$r_{ij} = \frac{\omega_{ij} - \frac{2}{\pi}\delta^2}{1 - \frac{2}{\pi}\delta^2}, \quad i, j = 1, \dots, k. \quad (4)$$

We are interested in structures of correlation matrices $\mathbf{\Omega}$, $\mathbf{\Omega}^*$, $\mathbf{\Omega}_*$, especially when they have the same structure, and we would like to examine how their structure affects the structure of the correlation matrix \mathbf{R} and the parameter vectors $\boldsymbol{\delta}$ and $\boldsymbol{\alpha}$ in different situations. A natural starting point is to use certain simple correlation structures depending on one parameter only. Thus, we consider the following correlation structures of $\mathbf{\Omega}$:

- (1) the exchangeable correlation structure or the *compound symmetry (CS)* when the correlations between all variables are equal, $\omega_{ij} = \omega$, $i, j = 1, \dots, k, i \neq j$;
- (2) the correlation structure, where the distance between variables in space or time determines the correlation (consider, e.g., autoregressive structure in repeated measurements or spatial processes). We focus on the simplest structure with this property: $\omega_{ij} = \omega^{|j-i|}$, $i, j = 1, \dots, k, i \neq j$, and call this structure the *autoregressive correlation structure (AR)*.

Of course, there are various alternative structures for correlations available. We begin our analysis with the *CS* and *AR* correlation structures because of the distinctive reasons: they are the simplest most commonly used approaches and require the estimation of one parameter ω only, no matter how big the number of variables is.

3. Formulation of the problem

Let us formulate the problem setup in more detail. Let $\mathbf{Z} = (Z_1, \dots, Z_k)^T$ have the k -variate skew-normal distribution, $\mathbf{Z} \sim SN(\mathbf{\Omega}, \boldsymbol{\delta})$, as specified in (1). Our aim is to investigate the effect of special cases of $\mathbf{\Omega}$ to the corresponding skew-normal distribution. In the following we are going to study the aforementioned two special cases of $\mathbf{\Omega}$ (*CS*, *AR*), starting with $\mathbf{\Omega} = \mathbf{I}$ as the simplest subcase of either structure.

Two main special cases for vector $\boldsymbol{\delta}$ are considered:

- (a) the case when all the components of parameter vector $\boldsymbol{\delta}$ are identical, $\boldsymbol{\delta} = (\delta, \dots, \delta)^T$, i.e., all marginals have the same distribution (i.e., $SN(\delta)$);
- (b) the situation when the correlation matrix $\mathbf{\Omega}^*$ (or $\mathbf{\Omega}_*$) has also some simple structure.

We search for answers to the following questions.

- (1) How can the parameter vector $\boldsymbol{\alpha}$ of the corresponding $\{\mathbf{\Omega}, \boldsymbol{\alpha}\}$ -parametrization be calculated?
- (2) What is the region of valid values for ω_{ij} (given $\boldsymbol{\delta}$)?
- (3) How can the correlation coefficients r_{ij} (defined in (3)) be calculated?
- (4) What is the region of possible correlation coefficients r_{ij} ?

4. The case when $\mathbf{\Omega} = \mathbf{I}$ and $\boldsymbol{\delta} = (\delta, \dots, \delta)^T$

Let us start with the simplest possible structure. Let $\mathbf{\Omega} = \mathbf{I}$ and $\boldsymbol{\delta} = (\delta, \dots, \delta)^T$, i.e., the skew-normal distribution has i.i.d. marginals and the corresponding normal distribution has independent marginals.

Let us recall the dual relation between the parameter vectors $\boldsymbol{\delta}$ and $\boldsymbol{\alpha}$ (see, e.g., [7]):

$$\boldsymbol{\alpha} = \frac{\mathbf{\Omega}^{-1}\boldsymbol{\delta}}{\sqrt{1 - \boldsymbol{\delta}^T \mathbf{\Omega}^{-1} \boldsymbol{\delta}}} \quad (5)$$

and

$$\boldsymbol{\delta} = \frac{\mathbf{\Omega}\boldsymbol{\alpha}}{\sqrt{1 + \boldsymbol{\alpha}^T \mathbf{\Omega}\boldsymbol{\alpha}}}. \quad (6)$$

Taking into account our current assumptions, formulas (5) and (6) simplify to

$$\boldsymbol{\alpha} = \frac{\boldsymbol{\delta}}{\sqrt{1 - k\delta^2}}, \quad \boldsymbol{\delta} = \frac{\boldsymbol{\alpha}}{\sqrt{1 + k\alpha^2}}. \quad (7)$$

In other words, the vector $\boldsymbol{\alpha}$ is in the form $(\alpha, \dots, \alpha)^T$, where $\alpha = \frac{\delta}{\sqrt{1 - k\delta^2}}$.

The region of valid values for δ (with $\boldsymbol{\delta} = (\delta, \dots, \delta)^T$) also follows from (7):

$$\delta \in \left(-\frac{1}{\sqrt{k}}; \frac{1}{\sqrt{k}} \right). \quad (8)$$

Also, $\mathbf{\Omega} = \mathbf{I}$ means that $\omega_{ij} = 0, i \neq j$, which together with (4) implies that $r_{ij} = -\frac{2\delta^2}{\pi - 2\delta^2}$, where $r_{ij} = \text{corr}(Z_i, Z_j)$ are the correlation coefficients corresponding to the skew-normal random variable \mathbf{Z} . The relationship between correlation coefficients $r_{ij} = r$ and δ for some values of k is shown in Figure 1.

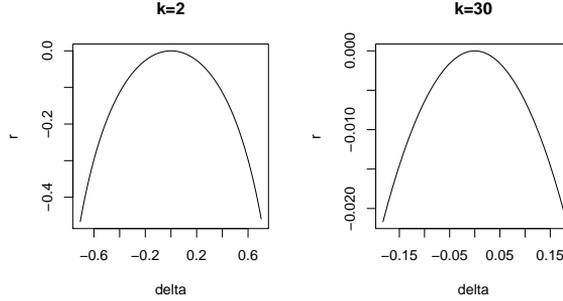


FIGURE 1. Relationship between r and δ when $\mathbf{\Omega} = \mathbf{I}$.

We conclude that the correlation matrix of the random variable \mathbf{Z} has the CS structure and, by (8), the region of possible values of correlation coefficients $r_{ij} = r$ is given by

$$r \in \left(-\frac{2}{k\pi - 2}; 0 \right]. \quad (9)$$

Remark 1. If the components of $\boldsymbol{\delta}$ are identical, $\boldsymbol{\delta} = (\delta, \dots, \delta)^T$, then the components of $\boldsymbol{\lambda}$ are also identical, $\boldsymbol{\lambda} = (\lambda, \dots, \lambda)^T$, and the following relations between δ and λ hold (see, for example, [10]):

$$\delta = \frac{\lambda}{\sqrt{1 + \lambda^2}}; \quad \lambda = \frac{\delta}{\sqrt{1 - \delta^2}}.$$

Now, given $\mathbf{\Omega} = \mathbf{I}$, the region of valid values for δ (8) also specifies the following region for valid values of λ :

$$\lambda \in \left(-\frac{1}{\sqrt{k-1}}; \frac{1}{\sqrt{k-1}} \right).$$

5. The case when $\mathbf{\Omega}$ has the CS structure

Let us begin with a motivating example.

Example 1 (Customer satisfaction data). The customer satisfaction index (CSI) is an economic indicator that measures the satisfaction of consumers. This is found by a customer satisfaction survey, which consists a questionnaire where the respondents (customers) are requested to give scores.

Now, for example, in case of benchmarking, we are interested in the behaviour of the customers whose score for specific questions is bigger (or smaller) than the overall mean or certain threshold. This implies that we have skewed data. Also, the whole customer satisfaction survey has usually several blocks of similar questions and the questions in blocks are correlated. The natural assumption is that the ordering of questions within a block does not matter, so we can assume the *CS* correlation structure.

5.1. The case when $\mathbf{\Omega}$ has the *CS* structure and $\boldsymbol{\delta} = (\delta, \dots, \delta)^T$. Let us assume that the correlation matrix $\mathbf{\Omega} = (\omega_{ij})$ has the *CS* correlation structure, i.e., $\omega_{ij} = \omega$, for $i, j = 1, \dots, k, i \neq j$. Consider the case of identical marginals, i.e., the vector $\boldsymbol{\delta}$ has the form $\boldsymbol{\delta} = (\delta, \dots, \delta)^T$. Formula (4) now simplifies to

$$r = \frac{\pi\omega - 2\delta^2}{\pi - 2\delta^2}. \quad (10)$$

In other words, the correlation matrix \mathbf{R} also has the *CS* structure with parameter r . Now, from (10) it is easy to see that if δ is fixed, then r is a linear function of ω .

To calculate $\boldsymbol{\alpha}$, let us first denote the elements of $\mathbf{\Omega}^{-1}$ by $\omega_{ij}^{(-1)}$, i.e., $\mathbf{\Omega}^{-1} = (\omega_{ij}^{(-1)})$. Then the matrix $\mathbf{\Omega}^{-1}$ has the following structure (see [21]):

$$\omega_{ii}^{(-1)} = \frac{1 + (k-2)\omega}{(1 + (k-1)\omega)(1-\omega)}; \quad \omega_{ij}^{(-1)} = -\frac{\omega}{(1 + (k-1)\omega)(1-\omega)}, \quad i \neq j.$$

Let us now apply formula (5) for $\boldsymbol{\alpha}$. The numerator $\mathbf{\Omega}^{-1}\boldsymbol{\delta}$ is a k -dimensional vector with equal components

$$\omega_{ii}^{(-1)}\delta + (k-1)\omega_{ij}^{(-1)}\delta = \frac{\delta}{1 + (k-1)\omega}$$

and for the the denominator we get

$$1 - \boldsymbol{\delta}^T \mathbf{\Omega}^{-1} \boldsymbol{\delta} = \frac{1 + (k-1)\omega - k\delta^2}{1 + (k-1)\omega}.$$

In summary, the parameter vector $\boldsymbol{\alpha}$ also has identical components, $\boldsymbol{\alpha} = (\alpha, \dots, \alpha)^T$ with

$$\alpha = \frac{\delta}{\sqrt{1 + (k-1)\omega} \sqrt{1 + (k-1)\omega - k\delta^2}}. \quad (11)$$

Let us look what the restriction (2) implies on the correlation matrix $\mathbf{\Omega}$ and on the correlation matrix \mathbf{R} . Considering the positive definiteness of $\mathbf{\Omega}$ and restriction (2) we get the following requirement for δ and ω :

$$0 < \frac{k\delta^2}{1 + (k-1)\omega} < 1. \quad (12)$$

To find the region of valid values for ω , we start from (12). Since the numerator is always nonnegative, the denominator must be positive, which implies that the region of valid values for ω is given by

$$\omega \in \left(\frac{k\delta^2 - 1}{k - 1}; 1 \right). \quad (13)$$

This region is illustrated in Figure 2.

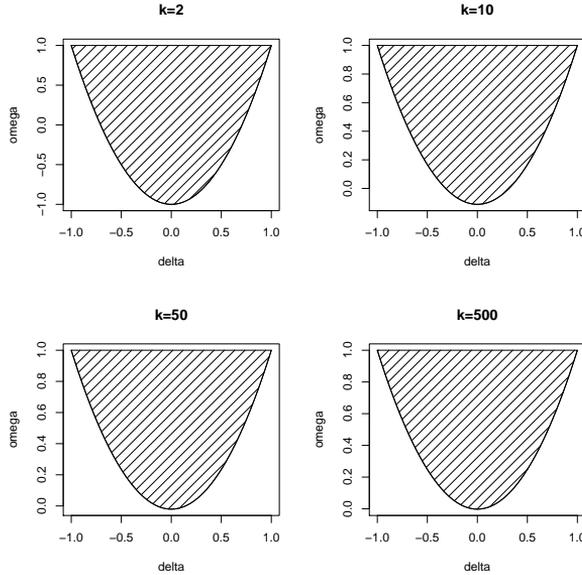


FIGURE 2. The region of valid values for ω when $\mathbf{\Omega}$ has the CS structure and $\boldsymbol{\delta} = (\delta, \dots, \delta)^T$.

Remark 2. It is easy to see that in case $\delta = 0$, that is, if we have normal distribution, the region of valid values for ω is $\omega \in \left(-\frac{1}{k-1}; 1 \right)$, which is the whole region where $\mathbf{\Omega}$ is positive definite, as expected.

Now, we have found the region (13) of valid elements ω of the correlation matrix $\mathbf{\Omega}$. This region also helps us to find the limits for the corresponding correlation coefficients r . Recall that the correlation matrix \mathbf{R} has the CS structure with the correlation coefficient r defined as in (10), and observe that

$$\omega \downarrow \frac{k\delta^2 - 1}{k - 1} \quad \text{and} \quad \omega \uparrow 1.$$

Using formula (10), straightforward calculations lead to the following region for possible values for r :

$$r \in \left(\frac{\delta^2(\pi k - 2k + 2) - \pi}{(k - 1)(\pi - 2\delta^2)}, 1 \right). \quad (14)$$

Remark 3. We have proved that if $\mathbf{\Omega}$ has the CS structure and $\boldsymbol{\delta} = (\delta, \dots, \delta)^T$, then \mathbf{R} also has the CS structure. It is not hard to prove that if \mathbf{R} has the CS structure and $\boldsymbol{\delta} = (\delta, \dots, \delta)^T$, then $\mathbf{\Omega}$ also has the CS structure. Nevertheless, it is also easy to show that the CS structure of \mathbf{R} is not sufficient to ensure that $\mathbf{\Omega}$ has the CS structure or $\boldsymbol{\delta} = (\delta, \dots, \delta)^T$.

Such situation is described in the following simple example.

Example 2. Let us have a vector $\boldsymbol{\delta}$ such that its elements are not identical: $\boldsymbol{\delta} = (0, \frac{1}{2}, 0)^T$. Let us also have $\mathbf{\Omega}$ that does not have the CS structure:

$$\mathbf{\Omega} = \begin{pmatrix} 1 & 1 - \frac{1}{2\pi} & \sqrt{1 - \frac{1}{2\pi}} \\ 1 - \frac{1}{2\pi} & 1 & 1 - \frac{1}{2\pi} \\ \sqrt{1 - \frac{1}{2\pi}} & 1 - \frac{1}{2\pi} & 1 \end{pmatrix}.$$

But it turns out that the matrix \mathbf{R} has the CS structure (apply formula (3)):

$$r_{12} = \sqrt{1 - \frac{1}{2\pi}}, \quad r_{13} = \sqrt{1 - \frac{1}{2\pi}} \quad \text{and} \quad r_{23} = r_{12} = \sqrt{1 - \frac{1}{2\pi}}.$$

Thus, the CS structure of \mathbf{R} does not ensure that the vector $\boldsymbol{\delta}$ has equal components nor that the matrix $\mathbf{\Omega}$ has the CS structure.

5.2. The case when $\mathbf{\Omega}$ and $\mathbf{\Omega}^*$ (and $\mathbf{\Omega}_*$) have the CS structures.

Let us assume that the correlation matrix $\mathbf{\Omega}$ has the CS structure as in the previous section. It is obvious that $\mathbf{\Omega}^*$ and $\mathbf{\Omega}_*$ also have the CS structures if $\boldsymbol{\delta} = (\omega, \dots, \omega)^T$. So, we have an evidential subcase of the previous case.

Let us see, how this simplification affects the quantities of interest. First, since $\boldsymbol{\delta} = (\omega, \dots, \omega)^T$, equality (11) reduces to

$$\alpha = \frac{\omega}{\sqrt{1 + (k-1)\omega} \sqrt{1 + (k-1)\omega - k\omega^2}}.$$

Also, the restriction (12) with $\boldsymbol{\delta} = \omega$ takes the form

$$0 < \frac{k\omega^2}{1 + (k-1)\omega} < 1, \quad (15)$$

which implies that if $\mathbf{\Omega}$ has the CS structure and $\boldsymbol{\delta} = (\omega, \dots, \omega)^T$, the region of valid values for ω is

$$\omega \in \left(-\frac{1}{k}; 1 \right). \quad (16)$$

The assumption that $\boldsymbol{\delta} = \omega$ also simplifies formula (10) as follows:

$$r = \frac{\pi\omega - 2\omega^2}{\pi - 2\omega^2} = \frac{\omega(\pi - 2\omega)}{\pi - 2\omega^2}. \quad (17)$$

This relationship is shown in Figure 3.

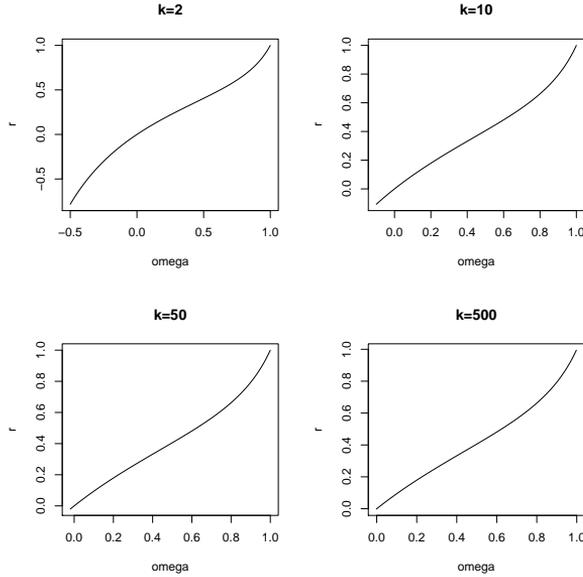


FIGURE 3. The relationship between the correlation coefficients r and ω for some values of k when Ω has the *CS* structure and $\delta = (\omega, \dots, \omega)^T$.

To find the region for possible values for r , consider the limit situations for $\omega \in \left(-\frac{1}{k}; 1\right)$. If $\omega \downarrow -\frac{1}{k}$, then $r \downarrow -\frac{\pi k + 2}{\pi k^2 - 2}$, and if $\omega \uparrow 1$, then $r \uparrow 1$. Therefore, the region for possible correlation coefficients r is

$$r \in \left(-\frac{\pi k + 2}{\pi k^2 - 2}; 1\right). \quad (18)$$

6. The case when Ω has the *AR* structure

Let us again begin with a motivating example.

Example 3 (PW170 test)). Let us consider the test for estimating aerobic fitness of athletes, which measures the physical working capacity (PWC) at a heart rate of 170 beats per minute (test called PWC170). Fifteen athletes (Estonian cross-country skiing team) performed six consecutive workloads on a bicycle ergometer and the average heart rate was recorded before test and at every step. So we had repeated measurements at seven time points. Now, if the model for the whole dataset was of interest, the multivariate normal fit was well justified (see, e.g., [18]). On the other hand, if the purpose is to analyze only some selection of data, for example, athletes whose heart rate before test was higher (or lower) than the average, the joint distribution is no longer multivariate normal but multivariate skew-normal. Analyzing the dependence between measurements showed that the correlations decreased

monotonically, so the natural choice is to use the autoregressive correlation structure.

6.1. The case when Ω has the AR structure and $\delta = (\delta, \dots, \delta)^T$. Let us now assume that the correlation matrix Ω has autoregressive structure (AR), i.e.,

$$\Omega = (\omega_{ij}) = (\omega^{|i-j|}). \quad (19)$$

The inverse Ω^{-1} of the correlation matrix is a three-diagonal matrix with well-known properties. The structure of matrix Ω^{-1} is the following (see, e.g., [22]):

$$\Omega^{-1} = \frac{1}{1-\omega^2} \begin{pmatrix} 1 & -\omega & 0 & \dots & 0 & 0 \\ -\omega & (1+\omega^2) & -\omega & \dots & 0 & 0 \\ 0 & -\omega & (1+\omega^2) & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & (1+\omega^2) & -\omega \\ 0 & 0 & 0 & \dots & -\omega & 1 \end{pmatrix}.$$

We are interested in the same questions as in previous sections. To find an expression for calculation of α , we will observe how the current assumptions (Ω has the AR structure and δ has the form $\delta = (\delta, \dots, \delta)^T$) influence formula (5). Let us study the numerator and denominator of (5) separately. For the numerator we derive

$$\Omega^{-1}\delta = \frac{\delta}{1+\omega}(1, 1-\omega, \dots, 1-\omega, 1)^T, \quad (20)$$

and the denominator of (5) is

$$1 - \delta^T \Omega^{-1} \delta = \frac{1 + \omega - (k(1 - \omega) + 2\omega)\delta^2}{1 + \omega}. \quad (21)$$

Now, the form of $\alpha = (\alpha_1, \dots, \alpha_k)$ follows from formulas (20) and (21):

$$\alpha_1 = \alpha_k = \frac{\delta}{\sqrt{1+\omega}\sqrt{1+\omega - (k(1-\omega) + 2\omega)\delta^2}}$$

and $\alpha_2 = \dots = \alpha_{k-1} = \alpha_k(1-\omega) = \alpha_1(1-\omega)$. Applying the derivation (21) to restriction (2), the region of valid values for ω is specified by

$$\omega > \frac{k\delta^2 - 1}{(k-2)\delta^2 + 1}. \quad (22)$$

The region corresponding to this inequality is shown in Figure 4.

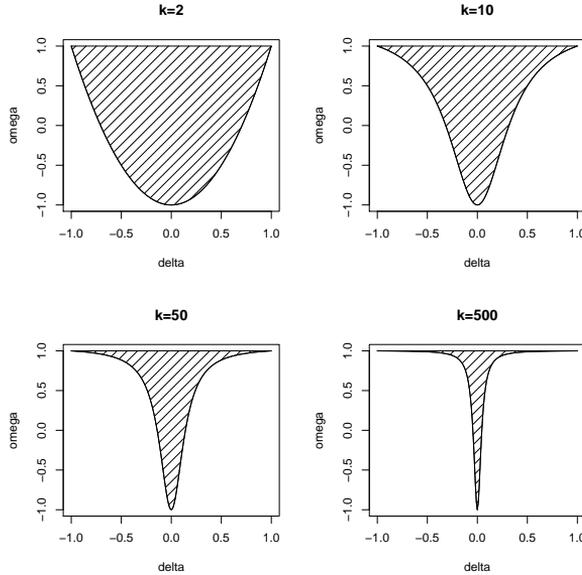


FIGURE 4. The region of valid values for ω (given δ and k) when $\mathbf{\Omega}$ has the CS structure and $\boldsymbol{\delta} = (\omega, \dots, \omega)^T$.

One can also note that the border curve for valid values of ω (specified by equality in (22)) is a convex function for $k = 2$. For $k > 2$, there is a flex point at $\pm \frac{1}{\sqrt{3(k-2)}}$, the curve is convex in $\left(-\frac{1}{\sqrt{3(k-2)}}, \frac{1}{\sqrt{3(k-2)}}\right)$ and concave outside.

Let us now focus on the correlation coefficients r_{ij} . Formula (3) together with $\omega_{ij} = \omega^{|i-j|}$ and $\boldsymbol{\delta} = (\delta, \dots, \delta)^T$ implies that

$$r_{ij} = \frac{\pi\omega^{|i-j|} - 2\delta^2}{\pi - 2\delta^2}. \quad (23)$$

The remaining task is to find the limits of r_{ij} depending on ω . For upper limit we consider $\omega \uparrow 1$, which, by (23), results in $r_{ij} \uparrow 1$. By formula (23) we can also see that the infimum value of the r_{ij} depends on whether $|i-j|$ is odd or even. If $|i-j|$ is even, then r_{ij} tends to its infimum if $\omega \rightarrow 0$, thus

$$r_{ij} \in \left(-\frac{2\delta^2}{\pi - 2\delta^2}; 1\right).$$

Because of (22), $\omega = 0$ yields that $|\delta| < 1/\sqrt{k}$. For $|\delta| \in \left[1/\sqrt{k}, 1\right)$, we have $\omega > 0$, the infimum of the r_{ij} is found in the process $\omega \downarrow \frac{k\delta^2 - 1}{(k-2)\delta^2 + 1}$, and by (23) we have

$$r_{ij} \in \left(\frac{\pi \left(\frac{k\delta^2 - 1}{(k-2)\delta^2 + 1} \right)^{|i-j|} - 2\delta^2}{\pi - 2\delta^2}; 1 \right). \quad (24)$$

Exactly the same reasoning can be used if $|i - j|$ is odd, thus the corresponding region of valid values for r_{ij} is also specified by (24).

Remark 4. From a practical perspective, the situation where $\omega > 0$ is obviously more common (consider, e.g., repeated measurements where the correlation decays in time). If $\omega > 0$, then the region of valid values for r_{ij} becomes

$$r_{ij} \in \left(-\frac{2\delta^2}{\pi - 2\delta^2}; 1 \right)$$

for both even and odd values of $|i - j|$, with $|\delta| < 1/\sqrt{k}$, because of (22). This means that even if $\omega > 0$, the r_{ij} can also have negative values.

6.2. The case when Ω and Ω_* have the AR structures. Let us assume that Ω and Ω_* have the AR structures. Note that if Ω_* has the AR structure, then the vector δ is specified as follows: $\delta = (\omega^k, \dots, \omega)^T$.

Let us now derive a formula for corresponding α . We again start from the general relation (5) and calculate the numerator and denominator of given expression separately. For the numerator we get

$$\Omega^{-1}\delta = (0, \dots, 0, \omega)^T,$$

and the denominator therefore reduces to $1 - \delta^T \Omega^{-1} \delta = 1 - \omega^2$. Also, as $\delta^T \Omega^{-1} \delta > 0$ if $|\omega| < 1$, the region of valid values of ω is $\omega \in (-1; 1)$.

Taking into account the forms of the numerator and the denominator of (5) under current assumptions, the components of $\alpha = (\alpha_1, \dots, \alpha_k)^T$ are the following: $\alpha_1 = \dots = \alpha_{k-1} = 0$ and $\alpha_k = \frac{\omega}{\sqrt{1 - \omega^2}}$. For the correlation coefficients we can see that the general formula (3) transforms to

$$r_{ij} = \frac{\omega^{|i-j|} - \frac{2}{\pi}\omega^{2k+2-i-j}}{\sqrt{1 - \frac{2}{\pi}\omega^{2(k+1-i)}}\sqrt{1 - \frac{2}{\pi}\omega^{2(k+1-j)}}}. \quad (25)$$

The behaviour of r_{ij} for different choices of ω and k is shown in Figure 5.

Remark 5. Let us now examine how the choice of ω affects the correlation coefficients r_{ij} . In the process $\omega \uparrow 1$ we have $r_{ij} \uparrow 1$. In the process $\omega \rightarrow 0$ we have $r_{ij} \rightarrow 0$ (the case of independent marginals). In the process $\omega \downarrow -1$, the value of r_{ij} depends on indices i and j : if $|i - j|$ is even, then $r_{ij} \uparrow 1$, and if $|i - j|$ is odd, then $r_{ij} \downarrow -1$.

The dependence of the regions for valid values of r_{ij} from the parity of $|i - j|$ is also illustrated in Figure 5. One can also see from this figure that bigger values of i and j produce smoother curve (especially if the difference

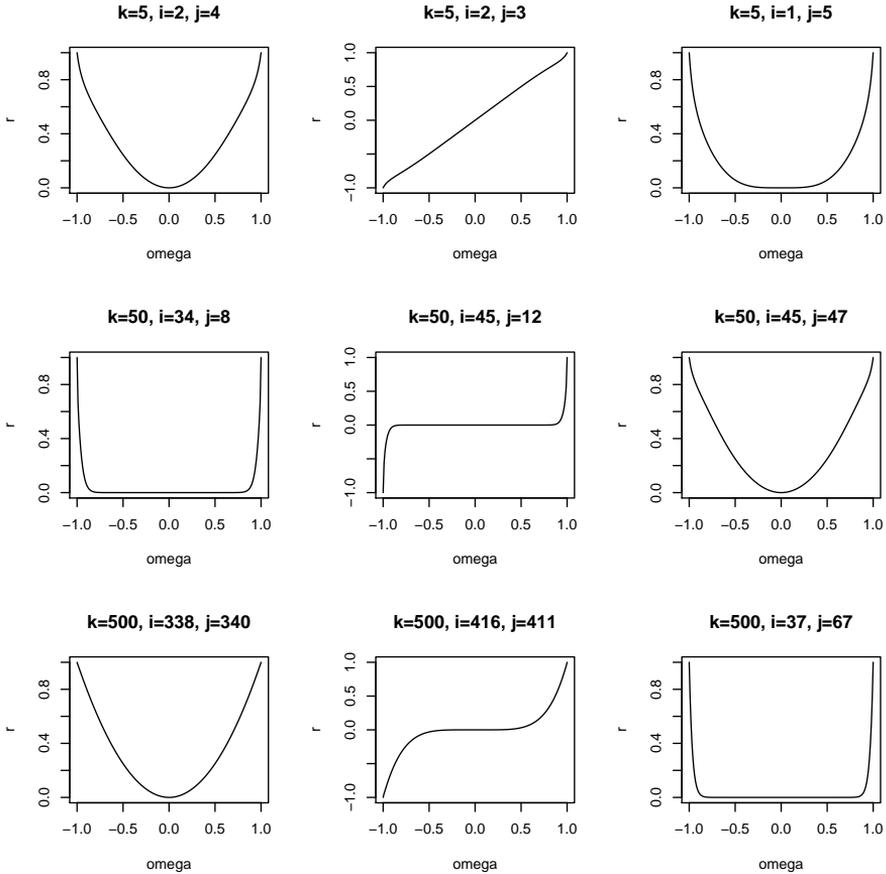


FIGURE 5. The relationship between the correlation coefficients r and ω when $\mathbf{\Omega}$ and $\mathbf{\Omega}_*$ have the AR structures.

of i and j is small) and big difference between i and j results in $r_{ij} = 0$ except for the extreme cases when $|\omega|$ is close to 1.

Although the correlation matrix \mathbf{R} does not retain the AR -structure, it is worth noting that there still exists a certain structure.

Proposition 1. *Let \mathbf{Z} follow a k -variate skew-normal distribution, $\mathbf{Z} \sim SN(\mathbf{\Omega}, \delta)$, such that $\mathbf{\Omega}$ and $\mathbf{\Omega}_*$ have the AR structures. Then the correlation matrix $\mathbf{R} = (r_{ij})$ has the following property:*

$$r_{ij} = I_{\{i < j\}} \prod_{m=i}^{j-1} r_{m,m+1} + I_{\{j < i\}} \prod_{m=j}^{i-1} r_{m,m+1}, \quad i \neq j,$$

where $I_{\{i < j\}}$ and $I_{\{j < i\}}$ are indicator functions.

The *proof* is straightforward, as the correlation coefficients r_{ij} and $r_{m,m+1}$ are on the form (25).

In other words, the elements on the first off-diagonal of the correlation matrix \mathbf{R} describe the whole structure of \mathbf{R} .

Remark 6. Similar results for the relations between ω and r_{ij} and the regions of their values can be proved for the case when $\mathbf{\Omega}$ and $\mathbf{\Omega}^*$ have the *AR* structures. As the calculations are similar, they are omitted. The obtained formulas are available in the tables in the summary section. It can also be shown that under these assumptions the correlation matrix \mathbf{R} has exactly the same structure as in Proposition 1.

7. Summary

The main aim of the paper was to investigate the effect of special cases of $\mathbf{\Omega}$ and δ to the corresponding skew-normal distribution.

We studied the two special cases of $\mathbf{\Omega}$ (*CS*, *AR*), starting with $\mathbf{\Omega} = \mathbf{I}$ as the simplest subcase of either structure and find simple formulas for the parameter vector α , for the correlation coefficients r_{ij} and for the valid regions of ω_{ij} and r_{ij} .

The following interesting findings can be pointed out:

- when $\mathbf{\Omega} = \mathbf{I}$ and δ has identical components, the only valid values for correlation coefficients in r_{ij} are non-positive and the region of possible values for r_{ij} diminishes with the increase of k (see formula (9));
- when $\mathbf{\Omega}$ has the *CS* structure and δ has identical components, the correlation matrix \mathbf{R} has the *CS* structure as well, but the converse implication is not true;
- when $\mathbf{\Omega}^*$ and $\mathbf{\Omega}_*$ have the *AR* structures, it does not imply that \mathbf{R} has the *AR* structure, yet there is some special structure (see Proposition 1).

More detailed results are summarized in the following tables.

Assumptions	
$\delta = (\delta, \dots, \delta)^T$	$\delta = (\omega, \dots, \omega)^T$
Results	
$\alpha = (\alpha, \dots, \alpha)^T$	$\alpha = (\alpha, \dots, \alpha)^T$
$\alpha = \frac{\delta}{\sqrt{1+(k-1)\omega}\sqrt{1+(k-1)\omega-k\delta^2}}$	$\alpha = \frac{\omega}{\sqrt{1+(k-1)\omega}\sqrt{1+(k-1)\omega-k\omega^2}}$
$r_{ij} = r$	$r_{ij} = r$
$r = \frac{\pi\omega-2\delta^2}{\pi-2\delta^2}$	$r = \frac{\omega(\pi-2\omega)}{\pi-2\omega^2}$
$r \in \left(\frac{\delta^2(\pi k-2k+2)-\pi}{(k-1)(\pi-2\delta^2)}; 1 \right)$	$r \in \left(-\frac{\pi k+2}{\pi k^2-2}; 1 \right)$
$\omega \in \left(\frac{k\delta^2-1}{k-1}; 1 \right)$	$\omega \in \left(-\frac{1}{k}; 1 \right)$

TABLE 1. The case when $\mathbf{\Omega}$ (or also $\mathbf{\Omega}^*$ and $\mathbf{\Omega}_*$) has the *CS* structure.

Assumptions	
$\boldsymbol{\delta} = (\delta, \dots, \delta)^T$	
Results	
$\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_k)^T$	
$\alpha_1 = \alpha_k = \frac{\delta}{\sqrt{1+\omega}\sqrt{1+\omega-(k(1-\omega)+2\omega)\delta^2}}$	
$\alpha_2 = \dots = \alpha_{k-1} = \alpha_k(1-\omega) = \alpha_1(1-\omega)$	
$r_{ij} = \frac{\pi\omega^{ i-j -2\delta^2}}{\pi-2\delta^2}$	
(a) $r_{ij} \in \left(\frac{\pi\left(\frac{k\delta^2-1}{(k-2)\delta^2+1}\right)^{ i-j -2\delta^2}}{\pi-2\delta^2}; 1 \right)$, if $ i-j $ is odd or $ \delta \in \left[\sqrt{\frac{1}{k}}, 1\right)$	
(b) $r_{ij} \in \left(-\frac{2\delta^2}{\pi-2\delta^2}; 1\right)$, if $ i-j $ is even and $ \delta < \sqrt{\frac{1}{k}}$ (or $\omega > 0$)	
$\omega \in \left(\frac{k\delta^2-1}{(k-2)\delta^2+1}; 1\right)$	

TABLE 2. The case when $\boldsymbol{\Omega}$ has the AR structure.

Assumptions	
$\boldsymbol{\delta} = (\omega^k, \dots, \omega)^T$	$\boldsymbol{\delta} = (\omega, \dots, \omega^k)^T$
Results	
$\boldsymbol{\alpha} = \left(0, \dots, 0, \frac{\omega}{\sqrt{1-\omega^2}}\right)^T$	$\boldsymbol{\alpha} = \left(\frac{\omega}{\sqrt{1-\omega^2}}, 0, \dots, 0\right)^T$
$r_{ij} = \frac{\omega^{ i-j -\frac{2}{\pi}\omega^{2k+2-i-j}}}{\sqrt{1-\frac{2}{\pi}\omega^{2(k+1-i)}}\sqrt{1-\frac{2}{\pi}\omega^{2(k+1-j)}}}$	$r_{ij} = \frac{\omega^{ i-j -\frac{2}{\pi}\omega^{i+j}}}{\sqrt{1-\frac{2}{\pi}\omega^{2i}}\sqrt{1-\frac{2}{\pi}\omega^{2j}}}$
$r_{ij} \in (-1; 1)$, if $ i-j $ is odd	$r_{ij} \in (-1; 1)$, if $ i-j $ is odd
$r_{ij} \in (0; 1)$, if $ i-j $ is even or $\omega > 0$	$r_{ij} \in (0; 1)$, if $ i-j $ is even or $\omega > 0$
$\omega \in (-1; 1)$	$\omega \in (-1; 1)$

TABLE 3. The cases when $\boldsymbol{\Omega}_*$ or $\boldsymbol{\Omega}^*$ has the AR structure.

Acknowledgements

The research was supported by institutional research funding IUT34-5 of the Estonian Ministry of Education and Research. The authors thank editors and anonymous referees for their constructive and valuable comments and suggestions.

References

- [1] R. B. Arellano-Valle, H. Bolfarine, and V. H. Lachos, *Skew-normal linear mixed models*, *J. Data Sci.* **3** (2005), 415–438.
- [2] R. B. Arellano-Valle, G. Del Pino, and E. San Martin, *Definition and probabilistic properties of skew distributions*, *Stat. Prob. Lett.* **58** (2002), 111–121.
- [3] R. B. Arellano-Valle, H. W. Gomez, and F. A. Quintana, *A new class of skew-normal distributions*, *Commun. Stat. Theory Methods* **33** (2004), 1465–1480.
- [4] B. C. Arnold and R. J. Beaver, *Skewed multivariate models related to hidden truncation and/or selective reporting*, *Sociedad de Estadística e Investigación Operativa Test* **11**(1) (2002), 7–54.
- [5] A. Azzalini, *A class of distributions which includes the normal ones*, *Scand. J. Stat.* **12** (1985), 171–178.
- [6] A. Azzalini, *The skew-normal distribution and related multivariate families (with discussion)*, *Scand. J. Stat.* **32** (2005), 159–188.
- [7] A. Azzalini and A. Capitanio, *Statistical applications of the multivariate skew normal distribution*, *J. R. Stat. Soc. Ser. B Stat. Methodol.* **61**(3) (1999), 579–602.
- [8] A. Azzalini and A. Capitanio, *Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t distribution*, *J. R. Stat. Soc. Ser. B Stat. Methodol.* **65**(2) (2003), 367–389.
- [9] A. Azzalini and A. Capitanio, *The Skew-Normal and Related Families*, Cambridge University Press, New York, 2015.
- [10] A. Azzalini and A. Dalla Valle, *The multivariate skew-normal distribution*, *Biometrika* **83** (4) (1996), 715–726.
- [11] S. Bartoletti and N. Loperfido, *Modelling air pollution data by the skew-normal distribution*, *Stoch. Env. Res. Risk A.* **24** (2010), 513–517.
- [12] L. M. Castro, E. San Martin, and R. B. Arellano-Valle, *A note on the parameterization of multivariate skewed-normal distributions*, *Braz. J. Prob. Stat.* **27** (2013), 110–115.
- [13] C. Crocetta and N. Loperfido, *Maximum likelihood estimation of correlation between maximal oxygen consumption and the 6-min walk test in patients with chronic heart failure*, *J. Appl. Stat.* **36**(10) (2009), 1101–1108.
- [14] O. Dunajeva, T. Kollo, and I. Traat, *Bias correction for the shape parameter of the skew normal distribution*, *Tatra Mt. Math. Publ.* **26** (2003), 281–289.
- [15] M. G. Genton (Editor), *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*, Chapman&Hall, Boca Raton, FL, 2004.
- [16] M. G. Genton and N. M. R. Loperfido, *Generalized skew-elliptical distributions and their quadratic forms*, *Ann. Inst. Stat. Math.* **57**(2) (2005), 389–401.
- [17] P. Ghosh, M. D. Branco, and H. Chakraborty, *Bivariate random effect model using skew-normal distribution with application to HIV-RNA*, *Stat. Med.* **26**(6) (2007), 1255–1267.
- [18] E. Käärrik, *Imputation algorithm using copulas*, *Adv. Methodol. Statist.* **3**(1) (2006), 109–120.
- [19] M. Käärrik, A. Selart, and E. Käärrik, *The use of copulas to model conditional expectation for multivariate data*, in: *ISI Proc. 58th World Statistical Congress*, 2011, Dublin, pp. 5533–5538.
- [20] M. Käärrik, A. Selart, and E. Käärrik, *On parametrization of multivariate skew-normal distribution*, *Comm. Stat. Theory Methods* **44**(9) (2015), 1869–1885.
- [21] C. R. Rao, *Linear Statistical Inference and Its Applications*, Wiley, New York, 2002.
- [22] A. Raveh, *On the use of the inverse of the correlation matrix in multivariate data analysis*, *Amer. Statist.* **39**(1) (1985), 39–42.

- [23] P. X. K. Song, *Correlated Data Analysis: Modeling, Analytics, and Applications*, Springer, New York, 2007.

INSTITUTE OF MATHEMATICAL STATISTICS, UNIVERSITY OF TARTU, J. LIIVI 2, 50409
TARTU, ESTONIA

E-mail address: ene.kaarik@ut.ee

E-mail address: meelis.kaarik@ut.ee

E-mail address: ingerhelenmaadik@gmail.com