JIRSS (2012)
Vol. 11, No. 1, pp 39-56

Prediction in a Trivariate Normal Distribution via Two Order Statistics

Alireza Arabpour, Mahsa Mahmood Molaiey, Ahad Jamalizadeh

Department of Statistics, Shahid Bahonar University of Kerman, Iran.

Abstract. In this paper, assuming that $(X, Y_1, Y_2)^T$ has a trivariate normal distribution, we derive the exact joint distribution of $\left(X, Y_{(1)}, Y_{(2)}\right)^T$, where $Y_{(1)}$ and $Y_{(2)}$ are order statistics arising from $(Y_1, Y_2)^T$. We show that this joint distribution is a mixture of truncated trivariate normal distributions and then use this mixture representation to derive the best (nonlinear) predictiors of X in terms of $\left(Y_{(1)}, Y_{(2)}\right)^T$. We also predict $Y_{(1)}$ in terms of $\left(X, Y_{(2)}\right)^T$, and $Y_{(2)}$ in terms of $\left(X, Y_{(1)}\right)^T$. Finally illustrate the usefulness of these results by using real-life data.

Keywords. Exchangeability, linear and nonlinear predictors, multivariate selection normal distribution, multivariate unified skew-normal distribution, order statistics, truncated trivariate normal distribution.

MSC: 62H05, 62H10, 62E15.

1 Introduction

In some situations, such as visual acuity, one is interested in studying relationships between an extreme measure or a function of extremes measures with one or more covariates. Let $(X, Y_1, Y_2)^T$ have a trivariate

Alireza Arabpour(✉)(arabpour@uk.ac.ir), Mahsa Mahmood Molaiey (m.molaiey@gmail.com), Ahad Jamalizadeh(a.jamalizadeh@uk.ac.ir)
Received: November 3, 2010; Accepted: October 11, 2011

normal distribution with mean vector $\boldsymbol{\mu}\in\mathbb{R}^3$ and positive definite covariance matrix $\boldsymbol{\Sigma}$, i.e. $(X,Y_1,Y_2)^T \sim N_3(\boldsymbol{\mu},\boldsymbol{\Sigma})$, We shall treat X as a covariate and write X as a covariate (such as age, weight, height, etc.), $\mathbf{Y}_{(2)} = \left(Y_{(1)},Y_{(2)}\right)^T$ for the vector of order statistics from $\mathbf{Y}=(Y_1,Y_2)^T$.

Olkin and Viana (1995) discussed the covariance structure of $\left(X,Y_{(1)},Y_{(2)}\right)^T$ in the case when $(X,Y_1,Y_2)^T$ has a trivariate normal distribution with Y_1 and Y_2 being exchangeable and $(X,Y_i)^T$, $i=1,2$, having a common correlation. They utilized this covariance structure to obtain the best linear predictors for X and $\left(Y_{(1)},Y_{(2)}\right)^T$. Viana (1998) considered the same exchangeable case and derived the best linear predictors for X and a linear combination of $Y_{(1)}$ and $Y_{(2)}$. Loperfido (2008) derived the exact joint distribution $\left(X,Y_{(2)}\right)^T$ as well as the conditional distribution of $X \mid Y_{(2)}$, under the same set-up. Jamalizadeh and Balakrishnan (2009) obtained the exact joint distribution of $\left(X,\mathbf{a}^T\mathbf{Y}_{(2)}\right)^T$, where $\mathbf{a}=(a_1,a_2)^T\in\mathbb{R}^2$ and $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ denoted the mean vector and covariance matrix. They showed that this joint distribution is a mixture of bivariate unified skew-normal distributions and used it to derive the best (non-linear) predictors of $\mathbf{a}^T\mathbf{Y}_{(2)}$ based on X as well as the predictors of X based on $\mathbf{a}^T\mathbf{Y}_{(2)}$. In this paper, we derive the exact joint distribution of $\left(X,Y_{(1)},Y_{(2)}\right)^T$, as mixture of truncated trivariate normal distributions and use it to derive the best predictors of X based on $\left(Y_{(1)},Y_{(2)}\right)^T$, as well as the predictors of $Y_{(1)}$ based on $\left(X,Y_{(2)}\right)^T$, and the predictors of $Y_{(2)}$ based on $\left(X,Y_{(1)}\right)^T$.

There are other possible fields of application besides visual acuity:

1. Reliability theory: Loperfido et al. (2007) obtained the distribution of order statistics arising from exchangeable random variables, as models for the failure time of parallel and series systems. However, they did not consider the use of covariates, whose values are often available. As an example, the failure time of a parallel system with two components (that is the maximum of failure times the components themselves) is definitely influenced by some properties of the components' material. The results in this paper my be useful in this setup, if both failure times and the main material' characteristic are modelled by a trivariate normal distribution.

2. Spatial statistics: As remarked by Loperfido and Guttorp (2008), monitoring networks are primarily aimed at finding large values of air pollution. Hence there is an interest in predicting the maximum level of air pollution on a given site, using the observed pollution level in a neighboring site, or the maximum value recorded in two neighboring

sites.

The rest of this paper is organized as follows. Section 2 presents a brief review of skew-normal distribution theory with special regard to the univariate and bivariate cases. We also, describe truncated univariate and trivariate normal distributions in this section. In Section 3, we derive the exact joint distribution of $\left(X, Y_{(1)}, Y_{(2)}\right)^T$ when $(X, Y_1, Y_2)^T$ has a mean vector and a covariance matrix. Finally, in Section 4, we illustrate our results using a real-life data.

2 Multivariate Selection Normal Distributions: Preliminaries

Let $\mathbf{U} \in \mathbb{R}^q$ and $\mathbf{V} \in \mathbb{R}^p$ be two random vectors and C be a measurable subset of $\mathbb{R}^q$. Arellano-Valle et al. (2006) defined a selection distribution as the conditional distribution of $\mathbf{V}$ given$\mathbf{U} \in C$. Specifically, a p-dimensional random vector $\mathbf{X}$ is said to have a multivariate selection distribution, denoted by $\mathbf{X} \sim SLCT_{p,q}$, with parameters depending on the characteristics of $\mathbf{U}$,$\mathbf{V}$, and C, if $\mathbf{X} \overset{d}{=} (\mathbf{V}|\mathbf{U} \in C)$. If $\mathbf{V}$ has a probability density function (pdf) $f_{\mathbf{V}}$, then $\mathbf{X}$ has a pdf $f_{SLCT_{p,q}}$ given by

$$f_{SLCT_{p,q}}(\mathbf{x}) = f_{\mathbf{V}}(\mathbf{x}) \frac{\Pr(\mathbf{U} \in C|\mathbf{V} = \mathbf{x})}{\Pr(\mathbf{U} \in C)}, \tag{1}$$

where $(\mathbf{U}, \mathbf{V})$ has a joint density function, $f_{\mathbf{U},\mathbf{V}}$. An alternative expression for the pdf of $\mathbf{X} \overset{d}{=} (\mathbf{V}|\mathbf{U} \in C)$ is given by

$$f_{SLCT_{p,q}}(\mathbf{x}) = \frac{\int_C f_{\mathbf{U},\mathbf{V}}(\mathbf{u}, \mathbf{x})\, d\mathbf{u}}{\int_C f_{\mathbf{U}}(\mathbf{u})\, d\mathbf{u}}.$$

One of the most important of selection distributions is when $\mathbf{U}$ and $\mathbf{V}$ are jointly normal distribution, i.e.

$$\begin{pmatrix} \mathbf{U} \\ \mathbf{V} \end{pmatrix} \sim N_{q+p}\left(\begin{pmatrix} \boldsymbol{\delta} \\ \boldsymbol{\xi} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Gamma} & \Delta^T \\ \Delta & \Omega \end{pmatrix}\right),$$

where $\boldsymbol{\delta} \in \mathbb{R}^q$, $\boldsymbol{\xi} \in \mathbb{R}^p$, $\Omega \in \mathbb{R}^{p\times p}$, $\Gamma \in \mathbb{R}^{q\times q}$ and $\Delta \in \mathbb{R}^{p\times q}$. In this case $\mathbf{X} \overset{d}{=} (\mathbf{V}|\mathbf{U} \in C)$ is said to have a multivariate selection normal distribution denoted by $\mathbf{X} \sim SLCT - N_{p,q}(\boldsymbol{\theta})$, where $\boldsymbol{\theta} = (\boldsymbol{\xi},\ \boldsymbol{\delta},\ \Omega,\ \Gamma,\ \Delta,\ C)$.

In this case the pdf of $\mathbf{X}$ can be easily seen to be [see the Arellano-Valle et al. (2006) for more details]

$$\phi_{SLCT-N_{p,q}}(\mathbf{x};\boldsymbol{\theta})=\frac{\phi_p(\mathbf{x};\boldsymbol{\xi},\Omega)\,\bar{\Phi}_q\left(C;\boldsymbol{\delta}+\Delta^T\Omega^{-1}(\mathbf{x}-\boldsymbol{\xi}),\Gamma-\Delta^T\Omega^{-1}\Delta\right)}{\bar{\Phi}_q(C;\boldsymbol{\delta},\Gamma)},\mathbf{x}\in\mathbb{R}^p, \tag{2}$$

where $\phi_p(\cdot;\boldsymbol{\xi},\Omega)$ denotes the pdf of $N_p(\boldsymbol{\xi},\Omega)$, and

$$\bar{\Phi}_q\left(C;\boldsymbol{\delta}+\Delta^T\Omega^{-1}(\mathbf{x}-\boldsymbol{\xi}),\boldsymbol{\Gamma}-\Delta^T\Omega^{-1}\Delta\right)$$

and $\bar{\Phi}_q(C;\boldsymbol{\delta},\Gamma)$ respectively denote $\Pr(\mathbf{Y}\in C)$ when

$$\mathbf{Y}\sim N_q\left(\boldsymbol{\delta}+\Delta^T\Omega^{-1}(\mathbf{x}-\boldsymbol{\xi}),\Gamma-\Delta^T\Omega^{-1}\Delta\right),$$

and $\mathbf{Y}\sim N_q(\boldsymbol{\delta};\Gamma)$.

The moment generating function is [see Arellano-Valle et al. (2006)]

$$M_{SLCT-N_{p,q}}(\mathbf{s};\boldsymbol{\theta})=\frac{\exp\left(\boldsymbol{\xi}^T\mathbf{s}+\frac{1}{2}\mathbf{s}^T\Omega\mathbf{s}\right)\bar{\Phi}_q(C;\boldsymbol{\delta}+\Delta^T\mathbf{s},\Gamma)}{\bar{\Phi}_q(C;\boldsymbol{\delta},\boldsymbol{\Gamma})},\quad s\in\mathbb{R}^p. \tag{3}$$

In the special case when $C=\{\mathbf{u}\in\mathbb{R}^q\,|\mathbf{u}>\mathbf{0}\}$ then the density function in (2) reduces to the density function of the multivariate unified skew-normal distribution presented by Arellano-Valle and Azzalini (2006),denoted by $\mathbf{X}\sim SUN_{p,q}(\boldsymbol{\theta})$, $\boldsymbol{\theta}=(\boldsymbol{\xi},\boldsymbol{\delta},\Omega,\Gamma,\Delta)$, as

$$\phi_{SUN_{p,q}}(\mathbf{x};\boldsymbol{\theta})=\frac{\phi_p(\mathbf{x};\boldsymbol{\xi},\Omega)\,\Phi_q\left(\boldsymbol{\delta}+\Delta^T\Omega^{-1}(\mathbf{x}-\boldsymbol{\xi});\Gamma-\Delta^T\Omega^{-1}\Delta\right)}{\Phi_q(\boldsymbol{\delta};\boldsymbol{\Gamma})},\quad \mathbf{x}\in\mathbb{R}^p\ ,$$

where $\Phi_q\left(\cdot;\Gamma-\Delta^T\Omega^{-1}\Delta\right)$ and $\Phi_q(\cdot,\Gamma)$ denote the cumulative distribution functions (cdf) of $N_q\left(\mathbf{0},\Gamma-\Delta^T\Omega^{-1}\Delta\right)$ and $N_q(\mathbf{0},\Gamma)$, respectively.

2.1 Univariate and Bivariate Unified Skew-Normal Distributions

In this section, we focus on two special cases of the multivariate unified skew-normal distribution with the density function given in (2). When $p=1$ and $q=1$ in (2), we obtain a univariate unified skew-normal random variable X with parameter $\boldsymbol{\theta} = \left(\xi,\, \delta, \omega,\, \gamma^2, \lambda\right)$, denoted by $X \sim SN\left(\boldsymbol{\theta}\right)$, with density function

$$\phi_{SN}\left(x; \boldsymbol{\theta}\right) = \frac{1}{\Phi\left(\frac{\delta}{\gamma}\right)}\ \phi(x; \xi, \omega)\ \Phi\left(\frac{\delta + \frac{\lambda}{\omega}\left(x - \xi\right)}{\sqrt{\gamma^2 - \frac{\lambda^2}{\omega}}}\right), \quad x \in \mathbb{R}, \tag{4}$$

where $\phi(\cdot\ ; \xi, \omega)$ is the pdf of $N\left(\xi, \omega\right)$ and $\Phi\left(\cdot\right)$ is the cdf of $N\left(0, 1\right)$ (standard normal distribution).

If $X \sim SN\left(\boldsymbol{\theta}\right)$, where $\boldsymbol{\theta} = \left(\xi,\, \delta, \omega,\, \gamma^2, \lambda\right)$, then the moment generating function, the first moment and the variance of X are, respectively,

$$\begin{aligned} M_{SN}\left(s; \boldsymbol{\theta}\right) &= \frac{1}{\Phi\left(\frac{\delta}{\gamma}\right)}\ \exp\left(\xi s + \frac{1}{2}\omega s^2\right)\ \Phi\left(\frac{\delta + \lambda s}{\gamma}\right), \\ E\left(X\right) &= \xi + \frac{\lambda}{\gamma}\, r\left(\frac{\delta}{\gamma}\right), \\ Var\left(X\right) &= \omega - \left(\frac{\lambda}{\gamma}\right)^2 r\left(\frac{\delta}{\gamma}\right)\left\{\frac{\delta}{\gamma} + r\left(\frac{\delta}{\gamma}\right)\right\}, \end{aligned}$$

where $r\left(t\right) = \frac{\phi(t)}{\Phi(t)}$, $t \in \Re$, is the reversed hazard rate function of the univariate standard normal distribution.

In the special case when $p = 2$ and $q = 1$ in (2), we obtain a bivariate unified skew-normal distribution. Specifically, if U and $\mathbf{V} = \left(V_1, V_2\right)^T$ are random variables of dimensions one and two, respectively, such that

$$\begin{pmatrix} U \\ \mathbf{V} \end{pmatrix} \sim N_3\left(\begin{pmatrix} \delta \\ \boldsymbol{\xi} \end{pmatrix}, \begin{pmatrix} \gamma^2 & \boldsymbol{\lambda}^T \\ \boldsymbol{\lambda} & \boldsymbol{\Omega} \end{pmatrix}\right),$$

where $\delta \in \Re, \boldsymbol{\xi} = \left(\xi_1, \xi_2\right)^T \in \Re^2,\ \gamma > 0,\ \boldsymbol{\lambda} = \left(\lambda_1,\, \lambda_2\right)^T \in \mathbb{R}^2$, and $\boldsymbol{\Omega} = \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix}$ is a two-dimensional positive definite covariance matrix. Then, a bivariate random vector $\mathbf{X} = \left(X_1, X_2\right)^T$ is said to have a bivariate unified skew-normal variable, with parameter $\boldsymbol{\theta} = \left(\boldsymbol{\xi},\, \delta, \boldsymbol{\Omega},\, \gamma^2, \boldsymbol{\lambda}\right)^T$, denoted by $BSN(\boldsymbol{\theta})$, if

$$\mathbf{X} \stackrel{d}{=} \mathbf{V} \mid U > 0. \tag{5}$$

From the general form in (2), the pdf of $\mathbf{X} = \left(X_1, X_2\right)^T$ becomes

$$\phi_{BSN}(\mathbf{x}; \boldsymbol{\theta}) = \frac{1}{\Phi\left(\frac{\delta}{\gamma}\right)}\ \phi_2(\mathbf{x}; \boldsymbol{\xi}, \boldsymbol{\Omega})\ \Phi\left(\frac{\delta + \boldsymbol{\lambda}^T\boldsymbol{\Omega}^{-1}\left(\mathbf{x} - \boldsymbol{\xi}\right)}{\sqrt{\gamma^2 - \boldsymbol{\lambda}^T\boldsymbol{\Omega}^{-1}\boldsymbol{\lambda}}}\right), \quad \mathbf{x} \in \mathbb{R}^2, \tag{6}$$

where $\phi_2(\cdot;\boldsymbol{\xi},\boldsymbol{\Omega})$ is the pdf of the bivariate normal distribution, $N_2(\boldsymbol{\xi},\boldsymbol{\Omega})$, and $\Phi(\cdot)$ is the univariate standard normal cdf, as before.

Lemma 2.1. *If* $\mathbf{X}=(X_1,X_2)^T \sim BSN(\boldsymbol{\theta})$, where $\boldsymbol{\theta}=(\boldsymbol{\xi},\delta,\boldsymbol{\Omega},\gamma^2,\boldsymbol{\lambda})$, then

(i)

$$X_1 \sim SN(\xi_1,\delta,\omega_{11},\gamma^2,\lambda_1). \tag{7}$$

(ii)

$$X_1|\,(X_2=x_2) \sim SN(\xi_{1.2},\delta_{1.2},\omega_{11.2},\gamma_{1.2}^2,\lambda_{1.2}), \tag{8}$$

where

$$\begin{aligned}
\xi_{1.2} &= \xi_1+\frac{\omega_{12}}{\omega_{22}}(x_2-\xi_2),\\
\delta_{1.2} &= \delta+\frac{\lambda_2}{\omega_{22}}(x_2-\xi_2),\\
\omega_{11.2} &= \omega_{11}-\frac{\omega_{12}^2}{\omega_{22}},\\
\gamma_{1.2}^2 &= \gamma^2-\frac{\lambda_2^2}{\omega_{22}},\ \lambda_{1.2}=\lambda_1-\frac{\omega_{12}}{\omega_{22}}\lambda_2.
\end{aligned}$$

(iii)

$$\begin{aligned}
E(X_1|X_2=x_2) &= \xi_{1.2}+\frac{\lambda_{1.2}}{\gamma_{1.2}}r\left(\frac{\delta_{1.2}}{\gamma_{1.2}}\right), \qquad (9)\\
Var(X_1|X_2=x_2) &= \omega_{11.2}-\left(\frac{\lambda_{1.2}}{\gamma_{1.2}}\right)^2 r\\
&\quad\times\left(\frac{\delta_{1.2}}{\gamma_{1.2}}\right)\left\{\frac{\delta_{1.2}}{\gamma_{1.2}}+r\left(\frac{\delta_{1.2}}{\gamma_{1.2}}\right)\right\};
\end{aligned}$$

here, $r(t)=\frac{\phi(t)}{\Phi(t)}$, $t\in\mathbb{R}$, is the reversed hazard rate function of the univariate standard normal distribution, as before.

2.2 Truncated Univariate and Trivariate Normal Distributions

Let $U\sim N(\delta,\gamma^2)$, then a univariate random variable X is said to have a truncated normal distribution on C, where C is a measurable subset of real line, denoted by $X\sim TN\left(\boldsymbol{\theta}=(\delta,\gamma^2,C)\right)$, if

$$X \stackrel{d}{=} U\,|(U\in C).$$

If $X \sim TN\left(\delta, \gamma^2, (-\infty, a)\right)$, $a \in \mathbb{R}$, then

$$\begin{aligned} E(X) &= \delta - \gamma \frac{\phi\left(\frac{a-\delta}{\gamma}\right)}{\Phi\left(\frac{a-\delta}{\gamma}\right)}, \\ Var(X) &= \gamma^2 \left(1 - \left(\frac{a-\delta}{\gamma}\right) \frac{\phi\left(\frac{a-\delta}{\gamma}\right)}{\Phi\left(\frac{a-\delta}{\gamma}\right)} - \left(\frac{\phi\left(\frac{a-\delta}{\gamma}\right)}{\Phi\left(\frac{a-\delta}{\gamma}\right)}\right)^2\right); \end{aligned}$$

Further, $X \sim TN\left(\delta, \gamma^2, (a, +\infty)\right)$, then

$$\begin{aligned} E(X) &= \delta + \gamma \frac{\phi\left(\frac{a-\delta}{\gamma}\right)}{1-\Phi\left(\frac{a-\delta}{\gamma}\right)}, \\ Var(X) &= \gamma^2 \left(1 + \left(\frac{a-\delta}{\gamma}\right) \frac{\phi\left(\frac{a-\delta}{\gamma}\right)}{1-\Phi\left(\frac{a-\delta}{\gamma}\right)} - \left(\frac{\phi\left(\frac{a-\delta}{\gamma}\right)}{1-\Phi\left(\frac{a-\delta}{\gamma}\right)}\right)^2\right). \end{aligned}$$

Let $\mathbf{U} = (U_1, U_2)^T$ and V be random variables of dimensions two and one, respectively, such that

$$\begin{pmatrix} \mathbf{U} \\ V \end{pmatrix} \sim N_3\left(\begin{pmatrix} \boldsymbol{\delta} \\ \xi \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Gamma} & \boldsymbol{\lambda} \\ \boldsymbol{\lambda}^T & \omega \end{pmatrix}\right),$$

where $\boldsymbol{\delta} = (\delta_1, \delta_2)^T \in \mathbb{R}^2$, $\xi \in \mathbb{R}$, $\Gamma = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}$ is a two-dimensional positive definite covariance matrix, $\omega > 0$ and $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)^T \in \mathbb{R}^2$. Then, a trivariate random vector $\mathbf{X} = (X_1, X_2, X_3)^T$ is said to have a truncated trivariate normal distribution, with parameter
$\boldsymbol{\theta} = \left(\boldsymbol{\psi} = \begin{pmatrix} \xi \\ \boldsymbol{\delta} \end{pmatrix}, \boldsymbol{\Psi} = \begin{pmatrix} \omega & \boldsymbol{\lambda}^T \\ \boldsymbol{\lambda} & \boldsymbol{\Gamma} \end{pmatrix}\right)$, denoted by $\mathbf{X} \sim TTN(\boldsymbol{\theta})$, if

$$\mathbf{X} \overset{d}{=} (V, U_1, U_2)^T \,|(U_1 < U_2) \tag{10}$$

The pdf of $\mathbf{X} = (X_1, X_2, X_3)^T$ is

$$\phi_{TTN}(\mathbf{x}; \boldsymbol{\theta}) = \begin{cases} \frac{\phi_3(\mathbf{x}; \boldsymbol{\psi}, \boldsymbol{\Psi})}{\Phi\left(\frac{\mathbf{c}^T \boldsymbol{\delta}}{\sqrt{\mathbf{c}^T \boldsymbol{\Gamma} \mathbf{c}}}\right)} & x_2 < x_3 \\ 0 & x_2 \geq x_3 \end{cases} \tag{11}$$

where $c=(-1,1)^T$.

Lemma 2.2. *If* $\mathbf{X}=(X_1,X_2,X_3)^T \sim TTN(\boldsymbol{\theta})$, *then*

(i)
$$X_1 \sim SN\left(\xi, c^T\boldsymbol{\delta}, \omega, c^T\boldsymbol{\Gamma}c, c^T\boldsymbol{\lambda}\right), \tag{12}$$
$$E\left(X_1\right)=\xi+\frac{c^T\boldsymbol{\lambda}}{\sqrt{\mathbf{c}^T\boldsymbol{\Gamma}\mathbf{c}}}r\left(\frac{c^T\boldsymbol{\delta}}{\sqrt{\mathbf{c}^T\boldsymbol{\Gamma}\mathbf{c}}}\right);$$

(ii)
$$X_2 \sim SN\left(\delta_1, c^T\boldsymbol{\delta}, \gamma_{11}, c^T\Gamma c, \gamma_{12}-\gamma_{11}\right), \tag{13}$$
$$E\left(X_2\right)=\delta_1+\frac{\gamma_{12}-\gamma_{11}}{\sqrt{\mathbf{c}^T\boldsymbol{\Gamma}\mathbf{c}}}r\left(\frac{c^T\boldsymbol{\delta}}{\sqrt{\mathbf{c}^T\boldsymbol{\Gamma}\mathbf{c}}}\right);$$

(iii)
$$X_3 \sim SN\left(\delta_2, c^T\boldsymbol{\delta}, \gamma_{22}, c^T\boldsymbol{\Gamma}c, \gamma_{22}-\gamma_{21}\right), \tag{14}$$
$$E\left(X_3\right)=\delta_2+\frac{\gamma_{22}-\gamma_{21}}{\sqrt{\mathbf{c}^T\boldsymbol{\Gamma}\mathbf{c}}}r\left(\frac{c^T\boldsymbol{\delta}}{\sqrt{\mathbf{c}^T\boldsymbol{\Gamma}\mathbf{c}}}\right);$$

(iv)
$$X_1 \mid \left(X_2=x_2, X_3=x_3\right) \sim N\left(\psi_{1.23}, \Psi_{11.23}\right), \tag{15}$$
$$E\left(X_1 \mid X_2=x_2, X_3=x_3\right)=\xi+\begin{pmatrix}\lambda_1 & \lambda_2\end{pmatrix}\boldsymbol{\Gamma}^{-1}\begin{pmatrix}x_2-\delta_1\\ x_3-\delta_2\end{pmatrix};$$

(v)
$$\left(X_1, X_3\right)^T \sim BSN\left(\begin{pmatrix}\xi\\ \delta_2\end{pmatrix}, \mathbf{c}^T\boldsymbol{\delta}, \begin{pmatrix}\omega & \lambda_2\\ \lambda_2 & \gamma_{22}\end{pmatrix}, \mathbf{c}^T\boldsymbol{\Gamma}\mathbf{c}, \begin{pmatrix}c^T\boldsymbol{\lambda}\\ \gamma_{22}-\gamma_{21}\end{pmatrix}\right),$$

(vi)
$$X_2 \mid \left(X_1=x_1, X_3=x_3\right) \sim TN\left(\psi_{2.13}, \Psi_{22.13}, (-\infty, x_3)\right), \tag{16}$$
$$E\left(X_2 \mid X_1=x_1, X_3=x_3\right)=\psi_{2.13}-\sqrt{\Psi_{22.13}}\frac{\phi\left(\frac{x_3-\psi_{2.13}}{\sqrt{\Psi_{22.13}}}\right)}{\Phi\left(\frac{x_3-\psi_{2.13}}{\sqrt{\Psi_{22.13}}}\right)};$$

(vii)
$$\left(X_1, X_2\right)^T \sim BSN\left(\begin{pmatrix}\xi\\ \delta_1\end{pmatrix}, \mathbf{c}^T\boldsymbol{\delta}, \begin{pmatrix}\omega & \lambda_1\\ \lambda_1 & \gamma_{11}\end{pmatrix}, \mathbf{c}^T\boldsymbol{\Gamma}\mathbf{c}, \begin{pmatrix}c^T\boldsymbol{\lambda}\\ \gamma_{12}-\gamma_{11}\end{pmatrix}\right),$$

(viii)

$$X_3 \mid (X_1 = x_1, X_2 = x_2) \sim TN\left(\psi_{3.12}, \Psi_{33.12}, (x_2, +\infty)\right), \quad (17)$$

$$E\left(X_3 \mid X_1 = x_1, X_2 = x_2\right) = \psi_{3.12} + \sqrt{\Psi_{33.12}} \frac{\phi\left(\frac{x_2 - \psi_{3.12}}{\sqrt{\Psi_{33.12}}}\right)}{1 - \Phi\left(\frac{x_2 - \psi_{3.12}}{\sqrt{\Psi_{33.12}}}\right)};$$

where

$$\psi_{1.23} = \xi + \begin{pmatrix} \lambda_1 & \lambda_2 \end{pmatrix} \Gamma^{-1} \begin{pmatrix} x_2 - \delta_1 \\ x_3 - \delta_2 \end{pmatrix},$$

$$\Psi_{11.23} = \omega - \begin{pmatrix} \lambda_1 & \lambda_2 \end{pmatrix} \Gamma^{-1} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix},$$

and

$$\psi_{2.13} = \delta_1 + \begin{pmatrix} \lambda_1 & \gamma_{21} \end{pmatrix} \begin{pmatrix} \omega & \lambda_2 \\ \lambda_2 & \gamma_{22} \end{pmatrix}^{-1} \begin{pmatrix} x_1 - \xi \\ x_3 - \delta_2 \end{pmatrix},$$

$$\Psi_{22.13} = \gamma_{11} - \begin{pmatrix} \lambda_1 & \gamma_{21} \end{pmatrix} \begin{pmatrix} \omega & \lambda_2 \\ \lambda_2 & \gamma_{22} \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1 \\ \gamma_{21} \end{pmatrix},$$

and

$$\psi_{3.12} = \delta_2 + \begin{pmatrix} \lambda_2 & \gamma_{12} \end{pmatrix} \begin{pmatrix} \omega & \lambda_1 \\ \lambda_1 & \gamma_{11} \end{pmatrix}^{-1} \begin{pmatrix} x_1 - \xi \\ x_2 - \delta_1 \end{pmatrix},$$

$$\Psi_{33.12} = \gamma_{22} - \begin{pmatrix} \lambda_2 & \gamma_{12} \end{pmatrix} \begin{pmatrix} \omega & \lambda_1 \\ \lambda_1 & \gamma_{11} \end{pmatrix}^{-1} \begin{pmatrix} \lambda_2 \\ \gamma_{12} \end{pmatrix}.$$

3 Joint Distribution of $\left(X, Y_{(1)}, Y_{(2)}\right)^T$

Let X and $\mathbf{Y} = (Y_1, Y_2)^T$ be two random variables such that

$$\begin{pmatrix} X \\ \mathbf{Y} \end{pmatrix} \sim N_3\left(\boldsymbol{\mu} = \begin{pmatrix} \mu_X \\ \boldsymbol{\mu}_{\mathbf{Y}} \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{XX} & \boldsymbol{\sigma}_{\mathbf{Y}X}^T \\ \boldsymbol{\sigma}_{\mathbf{Y}X} & \boldsymbol{\Sigma}_{\mathbf{YY}} \end{pmatrix}\right) \quad (18)$$

where $\boldsymbol{\mu}_{\mathbf{Y}} = (\mu_{Y_1}, \mu_{Y_2})^T \in \mathbb{R}^2$, and $\mu_X \in \mathbb{R}$, $\boldsymbol{\Sigma}_{\mathbf{YY}} = \begin{pmatrix} \sigma_{Y_1Y_1} & \sigma_{Y_1Y_2} \\ \sigma_{Y_2Y_1} & \sigma_{Y_2Y_2} \end{pmatrix}$ is a positive definite covariance matrix, $\sigma_{XX} > 0$ and $\boldsymbol{\sigma}_{\mathbf{Y}X} = (\sigma_{Y_1X}, \sigma_{Y_2X})^T \in [0, +\infty) \times [0, +\infty)$. Moreover, let $\mathbf{Y}_{(2)} = (Y_{(1)}, Y_{(2)})^T$ denote the vector of order statistics from $\mathbf{Y} = (Y_1, Y_2)^T$. In the following theorem, we show that the joint cdf of $(X, Y_{(1)}, Y_{(2)})^T$, is a mixture of two truncated trivariate normal distributions with the density function given in (11). For this purpose let $\mathbf{Y}^* = (Y_2, Y_1)^T$, $\boldsymbol{\mu}_{Y^*} = (\mu_{Y_2}, \mu_{Y_1})^T$, $\boldsymbol{\Sigma}_{\mathbf{Y}^*\mathbf{Y}^*} = \begin{pmatrix} \sigma_{Y_2Y_2} & \sigma_{Y_2Y_1} \\ \sigma_{Y_1Y_2} & \sigma_{Y_1Y_1} \end{pmatrix}$ and $\boldsymbol{\sigma}_{\mathbf{Y}^*X} = (\sigma_{Y_2X}, \sigma_{Y_1X})^T$.

Theorem 3.1. *The cdf of $(X, Y_{(1)}, Y_{(2)})^T$, is the mixture*

$$F_{\left(X,Y_{(1)},Y_{(2)}\right)}(\mathbf{t}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = p\Phi_{TTN}(\mathbf{t}; \boldsymbol{\theta}) + (1-p)\,\Phi_{TTN}(\mathbf{t}; \boldsymbol{\theta}^*); \qquad (19)$$

for $\mathbf{t} = (t_1, t_2, t_3) \in \mathbb{R}^3$, where $\Phi_{TTN}(\cdot;\theta)$ denotes the cdf of $TTN(\boldsymbol{\theta})$ in (11), and

$$p = \Phi\left(\frac{\mathbf{c}^T\mu_Y}{\sqrt{\mathbf{c}^T\boldsymbol{\Sigma}_{\mathbf{YY}}\mathbf{c}}}\right), \qquad \boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\Sigma}), \qquad \boldsymbol{\theta}^* = (\boldsymbol{\mu}^*, \boldsymbol{\Sigma}^*),$$

where

$$\boldsymbol{\mu}^* = \begin{pmatrix} \mu_X \\ \boldsymbol{\mu}_{Y^*} \end{pmatrix}, \boldsymbol{\Sigma}^* = \begin{pmatrix} \sigma_{XX} & \boldsymbol{\sigma}^T_{\mathbf{Y}^*X} \\ \boldsymbol{\sigma}_{\mathbf{Y}^*X} & \boldsymbol{\Sigma}_{\mathbf{Y}^*\mathbf{Y}^*} \end{pmatrix}.$$

with $c = (-1, 1)^T$.

Proof. For $\mathbf{t} = (t_1, t_2, t_3) \in \mathbb{R}^3$, we have

$$\begin{aligned} F_{\left(X,Y_{(1)},Y_{(2)}\right)}&(\mathbf{t}; \mu, \boldsymbol{\Sigma}) \qquad (20)\\ &= \Pr\left(X \le t_1, Y_{(1)} \le t_2, Y_{(2)} \le t_3\right) \\ &= \Pr(Y_1 < Y_2)\Pr(X \le t_1, Y_1 \le t_2, Y_2 \le t_3 \mid Y_1 < Y_2) \\ &\quad + \Pr(Y_2 \le Y_1)\Pr(X \le t_1, Y_2 \le t_2, Y_1 \le t_3 \mid Y_2 < Y_1) \end{aligned}$$

First, let us consider the first term on the RHS of (21). Now, $\Pr(Y_1 < Y_2) = p$, then

$$\Pr(X \le t_1, Y_1 \le t_2, Y_2 \le t_3 \mid Y_1 < Y_2) = \Phi_{TTN}(\mathbf{t}; \boldsymbol{\theta})$$

and consequently

$$\Pr(Y_1 < Y_2)\Pr(X \le t_1, Y_1 \le t_2, Y_2 \le t_3 \mid Y_1 < Y_2) = p\Phi_{TTN}(\mathbf{t}; \boldsymbol{\theta})$$

In a similar manner, we can show that

$$\Pr(Y_2 \le Y_1)\Pr(X \le t_1, Y_2 \le t_2, Y_1 \le t_3 \mid Y_2 < Y_1) = (1-p)F_{TTN}(\mathbf{t}; \boldsymbol{\theta}^*)$$

which completes the proof of the theorem. □

Corollary 3.1. *The joint density function of* $(X, Y_{(1)}, Y_{(2)})^T$ *is,*

$$f_{\left(X,Y_{(1)},Y_{(2)}\right)}\left(\mathbf{t};\boldsymbol{\mu},\boldsymbol{\Sigma}\right) = p\phi_{TTN}\left(\mathbf{t};\boldsymbol{\theta}\right) + (1-p)\,\phi_{TTN}\left(\mathbf{t};\boldsymbol{\theta}^*\right), \mathbf{t}\in\mathbb{R}^3$$

where $\phi_{TTN}(\cdot;\theta)$ *is as in (11).*

Corollary 3.2. *Assume (18) holds. Then the following are true:*

(i) The marginal cdf of $Y_{(2)}$ is given by

$$F_{Y_{(2)}}\left(t_3;\boldsymbol{\mu},\boldsymbol{\Sigma}\right) = p\Phi_{SN}\left(t_3;\boldsymbol{\theta}_3\right) + (1-p)\,\Phi_{SN}\left(t_3;\boldsymbol{\theta}_3^*\right), t_3\in\mathbb{R}$$

where $\Phi_{SN}(\cdot;\boldsymbol{\theta})$ is the cdf of $SN(\boldsymbol{\theta})$, and

$$\boldsymbol{\theta}_3 = \left(\mu_{Y_2}, \mathbf{c}^T\mu_Y, \sigma_{Y_2Y_2}, \mathbf{c}^T\boldsymbol{\Sigma}_{\mathbf{YY}}\mathbf{c}, \mathbf{c}^T\sigma_{Y_2\mathbf{Y}}\right),$$

$$\boldsymbol{\theta}_3^* = \left(\mu_{Y_2}, \mathbf{c}^T\mu_{Y^*}, \sigma_{Y_2Y_2}, \mathbf{c}^T\boldsymbol{\Sigma}_{\mathbf{YY}}\mathbf{c}, \mathbf{c}^T\sigma_{Y_2\mathbf{Y}^*}\right).$$

(ii) The conditional cdf of X given $\left(Y_{(1)}=t_2, Y_{(2)}=t_3\right)$ is given by

$$F_{X|\left(Y_{(1)}=t_2,Y_{(2)}=t_3\right)}\left(t_1;\boldsymbol{\mu},\boldsymbol{\Sigma}\right) = p\Phi\left(t_1;\boldsymbol{\theta}_{1.23}\right) + (1-p)\,\Phi\left(t_1;\boldsymbol{\theta}^*_{1.23}\right),$$

where

$$\begin{aligned}
\boldsymbol{\theta}_{1.23} &= \left(\mu_X + \sigma^T_{\mathbf{Y}X}\boldsymbol{\Sigma}^{-1}_{\mathbf{YY}}\begin{pmatrix} t_2-\mu_{Y_1} \\ t_3-\mu_{Y2}\end{pmatrix}, \sigma_{XX} - \sigma^T_{\mathbf{Y}X}\boldsymbol{\Sigma}^{-1}_{\mathbf{YY}}\sigma_{\mathbf{Y}X}\right),\\
\boldsymbol{\theta}^*_{1.23} &= \left(\mu_X + \sigma^T_{\mathbf{Y}^*X}\boldsymbol{\Sigma}^{-1}_{\mathbf{Y}^*\mathbf{Y}^*}\begin{pmatrix} t_2-\mu_{Y_2} \\ t_3-\mu_{Y_1}\end{pmatrix}, \sigma_{XX} - \sigma^T_{\mathbf{Y}^*X}\boldsymbol{\Sigma}^{-1}_{\mathbf{Y}^*\mathbf{Y}^*}\sigma_{\mathbf{Y}^*X}\right),
\end{aligned}$$

and the conditional mean of X given $\left(Y_{(1)}=t_2, Y_{(2)}=t_3\right)$ is given by

$$\begin{aligned}
&E\left(X \mid Y_{(1)}=t_2, Y_{(2)}=t_3\right)\\
&\quad= p\left(\mu_X + \sigma^T_{\mathbf{Y}X}\boldsymbol{\Sigma}^{-1}_{\mathbf{YY}}\begin{pmatrix} t_2-\mu_{Y_1} \\ t_3-\mu_{Y2}\end{pmatrix}\right)\\
&\quad\quad+ (1-p)\left(\mu_X + \sigma^T_{\mathbf{Y}^*X}\boldsymbol{\Sigma}^{-1}_{\mathbf{Y}^*\mathbf{Y}^*}\begin{pmatrix} t_2-\mu_{Y_2} \\ t_3-\mu_{Y_1}\end{pmatrix}\right).
\end{aligned}$$

(iii) The conditional cdf of $Y_{(1)}$ given $\left(X = t_1, Y_{(2)} = t_3\right)$ is given by

$$F_{Y_{(1)}|\left(X=t_1,Y_{(2)}=t_3\right)}\left(t_2;\boldsymbol{\mu},\boldsymbol{\Sigma}\right) = p\Phi_{TN}\left(t_2;\boldsymbol{\theta}_{2.13}\right) + (1-p)\,\Phi_{TN}\left(t_2;\boldsymbol{\theta}^*_{2.13}\right),$$

where $\Phi_{TN}\left(\cdot;\boldsymbol{\theta}\right)$ is the cdf of $TN\left(\boldsymbol{\theta}\right)$,

$$\begin{aligned}\boldsymbol{\theta}_{2.13} &= \left(\mu_{2.13}, \sigma_{22.13}, (-\infty, t_3)\right),\\ \boldsymbol{\theta}^*_{2.13} &= \left(\mu^*_{2.13}, \sigma^*_{22.13}, (-\infty, t_3)\right),\end{aligned}$$

and the conditional mean of $Y_{(1)}$ given $\left(X = t_1, Y_{(2)} = t_3\right)$ is given by

$$\begin{aligned}&E\left(Y_{(1)} \mid X = t_1, Y_{(2)} = t_3\right)\\ &\quad= p\left(\mu_{2.13} - \sqrt{\sigma_{22.13}}\frac{\phi\left(\frac{t_3-\mu_{2.13}}{\sqrt{\sigma_{22.13}}}\right)}{\Phi\left(\frac{t_3-\mu_{2.13}}{\sqrt{\sigma_{22.13}}}\right)}\right)\\ &\quad+(1-p)\left(\mu^*_{2.13} - \sqrt{\sigma^*_{22.13}}\frac{\phi\left(\frac{t_3-\mu^*_{2.13}}{\sqrt{\sigma^*_{22.13}}}\right)}{\Phi\left(\frac{t_3-\mu^*_{2.13}}{\sqrt{\sigma^*_{22.13}}}\right)}\right).\end{aligned}$$

(iv) The conditional cdf of $Y_{(2)}$ given $\left(X = t_1, Y_{(1)} = t_2\right)$ is given by

$$F_{Y_{(2)}|X=t_1,Y_{(1)}=t_2}\left(t_3;\boldsymbol{\mu},\boldsymbol{\Sigma}\right) = p\Phi_{TN}\left(t_3;\boldsymbol{\theta}_{3.12}\right) + (1-p)\,\Phi_{TN}\left(t_3;\boldsymbol{\theta}^*_{3.12}\right),$$

where

$$\begin{aligned}\boldsymbol{\theta}_{3.12} &= \left(\mu_{3.12}, \sigma_{33.12}, (t_2, \infty)\right),\\ \boldsymbol{\theta}^*_{3.12} &= \left(\mu^*_{3.12}, \sigma^*_{33.12}, (t_2, \infty)\right),\end{aligned}$$

and the conditional mean of $Y_{(2)}$ given $\left(X = t_1, Y_{(1)} = t_2\right)$ is given by

$$\begin{aligned}&E\left(Y_{(2)} \mid X = t_1, Y_{(1)} = t_2\right)\\ &\quad= p\left(\mu_{3.12} + \sqrt{\sigma_{33.12}}\frac{\phi\left(\frac{t_2-\mu_{3.12}}{\sqrt{\sigma_{33.12}}}\right)}{1-\Phi\left(\frac{t_2-\mu_{3.12}}{\sqrt{\sigma_{33.12}}}\right)}\right)\\ &\quad+(1-p)\left(\mu^*_{3.12} + \sqrt{\sigma^*_{33.12}}\frac{\phi\left(\frac{t_2-\mu^*_{3.12}}{\sqrt{\sigma^*_{33.12}}}\right)}{1-\Phi\left(\frac{t_2-\mu^*_{3.12}}{\sqrt{\sigma^*_{33.12}}}\right)}\right),\end{aligned}$$

where

$$\mu_{2.13} = \mu_{Y_1} + \begin{pmatrix} \sigma_{XY_1} & \sigma_{Y_2Y_1} \end{pmatrix} \begin{pmatrix} \sigma_{XX} & \sigma_{XY_2} \\ \sigma_{Y_2X} & \sigma_{Y_2Y_2} \end{pmatrix}^{-1} \begin{pmatrix} t_1 - \mu_X \\ t_3 - \mu_{Y_2} \end{pmatrix},$$

$$\mu^*_{2.13} = \mu_{Y_2} + \begin{pmatrix} \sigma_{XY2} & \sigma_{Y_1Y2} \end{pmatrix} \begin{pmatrix} \sigma_{XX} & \sigma_{XY_1} \\ \sigma_{Y_1X} & \sigma_{Y_1Y_1} \end{pmatrix}^{-1} \begin{pmatrix} t_1 - \mu_X \\ t_3 - \mu_{Y_1} \end{pmatrix},$$

$$\sigma_{22.13} = \sigma_{Y_1Y_1} - \begin{pmatrix} \sigma_{XY_1} & \sigma_{Y_2Y_1} \end{pmatrix} \begin{pmatrix} \sigma_{XX} & \sigma_{XY_2} \\ \sigma_{Y_2X} & \sigma_{Y_2Y_2} \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{XY_1} \\ \sigma_{Y_2Y_1} \end{pmatrix},$$

$$\sigma^*_{22.13} = \sigma_{Y_2Y_2} - \begin{pmatrix} \sigma_{XY_2} & \sigma_{Y_1Y_2} \end{pmatrix} \begin{pmatrix} \sigma_{XX} & \sigma_{XY_1} \\ \sigma_{Y_1X} & \sigma_{Y_1Y_1} \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{XY_2} \\ \sigma_{Y_1Y_2} \end{pmatrix},$$

and

$$\mu_{3.12} = \mu_{Y_2} + \begin{pmatrix} \sigma_{XY_2} & \sigma_{Y_1Y_2} \end{pmatrix} \begin{pmatrix} \sigma_{XX} & \sigma_{XY_1} \\ \sigma_{Y_1X} & \sigma_{Y_1Y_1} \end{pmatrix}^{-1} \begin{pmatrix} t_1 - \mu_X \\ t_2 - \mu_{Y_1} \end{pmatrix},$$

$$\mu^*_{3.12} = \mu_{Y_1} + \begin{pmatrix} \sigma_{XY_1} & \sigma_{Y_2Y_1} \end{pmatrix} \begin{pmatrix} \sigma_{XX} & \sigma_{XY_2} \\ \sigma_{Y_2X} & \sigma_{Y_2Y_2} \end{pmatrix}^{-1} \begin{pmatrix} t_1 - \mu_X \\ t_2 - \mu_{Y_2} \end{pmatrix},$$

$$\sigma_{33.12} = \sigma_{Y_2Y_2} - \begin{pmatrix} \sigma_{XY_2} & \sigma_{Y_1Y_2} \end{pmatrix} \begin{pmatrix} \sigma_{XX} & \sigma_{XY_1} \\ \sigma_{Y_1X} & \sigma_{Y_1Y_1} \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{XY_2} \\ \sigma_{Y_1Y_2} \end{pmatrix},$$

$$\sigma^*_{33.12} = \sigma_{Y_1Y_1} - \begin{pmatrix} \sigma_{XY_1} & \sigma_{Y_2Y_1} \end{pmatrix} \begin{pmatrix} \sigma_{XX} & \sigma_{XY_2} \\ \sigma_{Y_2X} & \sigma_{Y_2Y_2} \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{XY_1} \\ \sigma_{Y_2Y_1} \end{pmatrix}.$$

3.1 An Exchangeable Case

Here, we consider the special case when

$$\begin{pmatrix} X \\ Y_1 \\ Y_2 \end{pmatrix} \sim N_3\left(\begin{pmatrix} \mu_0 \\ \mu_1 \\ \mu_1 \end{pmatrix}, \begin{pmatrix} \sigma^2 & \eta\sigma\tau & \eta\sigma\tau \\ \eta\sigma\tau & \tau^2 & \rho\tau^2 \\ \eta\sigma\tau & \rho\tau^2 & \tau^2 \end{pmatrix}\right), \qquad (21)$$

where $\mu_0, \mu_1 \in \mathbb{R}, \tau > 0, |\rho| < 1, |\eta| < \sqrt{\frac{1+\rho}{2}}$.Olkin and Viana (1995) considered this case and derived the covariance matrix of $(X, Y_{(1)}, Y_{(2)})^T$, and then utilized it to obtain the best linear predictors of $Y_{(1)}$ given X, $Y_{(2)}$ given X, and X given $Y_{(1)}$ and $Y_{(2)}$. Under the same set-up, Loperfido (2008) derived the exact distribution of $X \mid Y_{(2)}$ and then an

expression for the conditional mean $E\left(X \mid Y_{(2)}\right)$.
From the mixture form in Theorem 1, we can easily obtain expressions for the conditional means $E\left(X \mid (Y_{(1)} = t_2, Y_{(2)} = t_3)\right)$, $E\left(Y_{(1)} \mid (X = t_1, Y_{(2)} = t_3)\right)$ and $E\left(Y_{(2)} \mid (X = t_1, Y_{(1)} = t_2)\right)$ as follows.

Corollary 3.3. *If $(X, Y_1, Y_2)^T$ is exchangeable as in (21), then*

(i) $E\left(X \mid Y_{(1)} = t_2, Y_{(2)} = t_3\right) = \mu_0 + \frac{\eta\sigma}{\tau(1+\rho)}\left(t_2 + t_3 - 2\mu_1\right),$

(ii) $E\left(Y_{(1)} \mid X = t_1, Y_{(2)} = t_3\right) = \mu_{2.13} - \sqrt{\sigma_{22.13}}\frac{\phi\left(\frac{t_3-\mu_{2.13}}{\sqrt{\sigma_{22.13}}}\right)}{\Phi\left(\frac{t_3-\mu_{2.13}}{\sqrt{\sigma_{22.13}}}\right)},$

(iii) $E\left(Y_{(2)} \mid X = t_1, Y_{(1)} = t_2\right) = \mu_{3.12} + \sqrt{\sigma_{33.12}}\frac{\phi\left(\frac{t_2-\mu_{3.12}}{\sqrt{\sigma_{33.12}}}\right)}{1-\Phi\left(\frac{t_2-\mu_{3.12}}{\sqrt{\sigma_{33.12}}}\right)},$

where

$$\mu_{2.13} = \mu_1 + \frac{\eta\tau\left(1-\rho\right)\left(t_1-\mu_0\right) - \sigma\left(\eta^2-\rho\right)\left(t_3-\mu_1\right)}{\sigma\left(1-\eta^2\right)},$$

$$\sigma_{22.13} = \frac{\tau^2}{1-\eta^2}\left(1-2\eta^2\left(1-\rho\right)-\rho^2\right),$$

and

$$\mu_{3.12} = \mu_1 + \frac{\eta\tau\left(1-\rho\right)\left(t_1-\mu_0\right) - \sigma\left(\eta^2-\rho\right)\left(t_2-\mu_1\right)}{\sigma\left(1-\eta^2\right)},$$

$$\sigma_{33.12} = \frac{\tau^2}{1-\eta^2}\left(1-2\eta^2\left(1-\rho\right)-\rho^2\right).$$

4 Illustration With Visual Acuity Data

Fishman et al. (1993) evaluated 43 patients with Best's vitelliform macular dystrophy for age X, visual acuity in left eye Y_1 and visual acuity in right eye Y_2. Olkin and Viana (1995) applied the model in (21) to these data and computed the maximum likelihood estimates (MLEs) of the parameters follows: the mean and standard deviation of age and the mean and standard deviation of vision in the eyes ** are respectively $\hat{\mu}_0 = 28.833$, $\hat{\mu}_1 = 0.424$, $\hat{\sigma} = 19.182$, $\hat{\tau} = 0.386$. Also, the correlation ρ between vision on the eyes was estimated as $\hat{\rho} = 0.496$ and the MLE of the correlation between age and either eye is $\hat{\eta} = 0.581$, then used the covariance structure of $(X, Y_{(1)}, Y_{(2)})^T$, along with the estimated first and second moments of $X, Y_{(1)}$ and $Y_{(2)}$, they compute the best linear predictor of $Y_{(1)}$ based on $X, Y_{(2)}$ as well as best linear predictor of $Y_{(2)}$ based on $X, Y_{(1)}$ as

$$\begin{aligned} E\left(Y_{(1)} \mid X = t_1, Y_{(2)} = t_3\right) &= -0.21 + 0.6351 t_3 + 0.0042 t_1, \\ E\left(Y_{(2)} \mid X = t_1, Y_{(1)} = t_2\right) &= 0.28 + 0.6351 t_2 + 0.0042 t_1. \end{aligned}$$

By Corollary 4, we can easily obtain the (estimated) conditional mean of X given $\left(Y_{(1)}, Y_{(2)}\right)$ as

$$E\left(X \mid Y_{(1)} = t_2, Y_{(2)} = t_3\right) = 12.47 + 19.3(t_2 + t_3),$$

which is exactly the same predictor as obtained by Olkin and Viana(1995). Further, under the exchangeable model in (21), the best (nonlinear) predictor $Y_{(1)}$ based on $\left(X, Y_{(2)}\right)$, and $Y_{(2)}$ based on $\left(X, Y_{(1)}\right)$, are

$$\begin{aligned} E\left(Y_{(1)} \mid X = t_1, Y_{(2)} = t_3\right) &= \mu_{2.13} - \sqrt{0.093}\frac{\phi\left(\frac{t_3 - \mu_{2.13}}{\sqrt{0.093}}\right)}{\Phi\left(\frac{t_3 - \mu_{2.13}}{\sqrt{0.093}}\right)}, \\ E\left(Y_{(2)} \mid X = t_1, Y_{(1)} = t_2\right) &= \mu_{3.12} + \sqrt{0.093}\frac{\phi\left(\frac{t_2 - \mu_{3.12}}{\sqrt{0.092}}\right)}{1 - \Phi\left(\frac{t_2 - \mu_{3.12}}{\sqrt{0.093}}\right)}, \end{aligned}$$

respectively, where

$$\begin{aligned} \mu_{2.13} &= \frac{0.113 t_1 + 3.04 t_3 + 0.84}{12.71}, \\ \mu_{3.12} &= \frac{0.113 t_1 + 3.04 t_2 + 0.84}{12.71}. \end{aligned}$$

If we consider the full model in (18) for these data then the estimated mean vector and covariance matrix of $(X, Y_1, Y_2)^T$ are [see Fishman et al. (1993)]

$$(\hat{\mu}_X, \hat{\mu}_1, \hat{\mu}_2)^T = (28.833, 0.412, 0.437)^T, \hat{\boldsymbol{\Sigma}} = \begin{pmatrix} 367.996 & 4.419 & 4.200 \\ 4.419 & 0.135 & 0.074 \\ 4.200 & 0.074 & 0.163 \end{pmatrix}.$$

Now, under the full model in (18), using the results in Corollary 2, the best predictor of X based on $\left(Y_{(1)}, Y_{(2)}\right)^T$, $Y_{(1)}$ based on $\left(X, Y_{(2)}\right)^T$ and $Y_{(2)}$ based on $\left(X, Y_{(1)}\right)^T$ can be easily obtained as

$$E\left(X \mid Y_{(1)} = t_2, Y_{(2)} = t_3\right) = 12.317 + 21.553 t_2 + 17.577 t_3,$$

$$\begin{aligned} E\left(Y_{(1)} \mid X = t_1, Y_{(2)} = t_3\right) \\ &= 0.57\left(\mu_{2.13} - \sqrt{0.077}\frac{\phi\left(\frac{t_3-\mu_{2.13}}{\sqrt{0.077}}\right)}{\Phi\left(\frac{t_3-\mu_{2.13}}{\sqrt{0.077}}\right)}\right) \\ &+0.43\left(\mu^*_{2.13} - \sqrt{0.109}\frac{\phi\left(\frac{t_3-\mu^*_{2.13}}{\sqrt{0.109}}\right)}{\Phi\left(\frac{t_3-\mu^*_{2.13}}{\sqrt{0.109}}\right)}\right), \end{aligned}$$

$$\begin{aligned} E\left(Y_{(2)} \mid X = t_1, Y_{(1)} = t_2\right) \\ &= 0.57\left(\mu_{3.12} + \sqrt{0.109}\frac{\phi\left(\frac{t_2-\mu_{3.12}}{\sqrt{0.109}}\right)}{1-\Phi\left(\frac{t_2-\mu_{3.12}}{\sqrt{0.109}}\right)}\right) \\ &+0.43\left(\mu^*_{3.12} + \sqrt{0.077}\frac{\phi\left(\frac{t_2-\mu^*_{3.12}}{\sqrt{0.077}}\right)}{1-\Phi\left(\frac{t_2-\mu^*_{3.12}}{\sqrt{0.077}}\right)}\right), \end{aligned}$$

where

$$\begin{aligned} \mu_{2.13} &= 0.044 + 0.010 t_1 + 0.204 t_3, \\ \mu^*_{2.13} &= 0.089 + 0.009 t_1 + 0.226 t_3, \end{aligned}$$

and

$$\begin{aligned} \mu_{3.12} &= 0.089 + 0.009 t_1 + 0.226 t_2, \\ \mu^*_{3.12} &= 0.044 + 0.010 t_1 + 0.204 t_2. \end{aligned}$$

5 Concluding Remarks

In this paper, we have derived the exact joint distribution of $(X, Y_{(1)}, Y_{(2)})^T$, when X is a covariate and $\mathbf{Y}_{(2)} = \left(Y_{(1)}, Y_{(2)}\right)^T$ is the vector of order statistics from $\mathbf{Y} = (Y_1, Y_2)^T$, in the case when $(X, Y_1, Y_2)^T \sim N_3(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. All the results established in this paper can be extended to the case when $(X, Y_1, Y_2)^T \sim EC_3(\boldsymbol{\mu}, \boldsymbol{\Sigma}, h^{(3)})$, a trivariate elliptical distribution with location parameter $\boldsymbol{\mu} \in \mathbb{R}^3$, positive definite dispersion matrix $\boldsymbol{\Sigma}_{3\times 3}$, and density generator $h^{(3)}$. Work is currently under progress on these issues, and we hope to report our findings in a future paper.

Acknowledgments

The authors thank Mahani Mathematical Research Center (Kerman, Iran) for supporting this research. The authors also sincerely thank the anonymous referees for their valuable comments and suggestions which led to a considerable improvement in the presentation.

References

Arellano-Valle, R. B. and Azzalini, A., (2006), On the unification of families of skew-normal distributions. Scand. J. Statist. **33**, 561-574.

Arellano-Valle, R. B., Branco, M. D., and Genton, M. G., (2006), A unified view on skewed distributions arising from selections. Canad. J. Statist., **34**, 1-21.

Fishman, G. A., Baca, W., Alexander, K. R., Derlacky, D. J., Glenn, A. M., and Viana, M. A. G. (1993), Visual acuity in patients with best vitelliform macular dystrophy, Ophthalmology. **100**, 1665-1670.

Jamalizadeh, A. and Balakrishnan, N. (2009), Prediction in a trivariate normal distribution via a linear combination of order statistics. Statist. Probab. Lett., **79**, 2289-2296.

Loperfido, N. (2008), Modeling maxima of longitudinal contralateral observations. Test, **17**, 370-380.

Loperfido N., Guttorp P. (2008), Network bias in air quality monitoring design. Environmetrics, **19**, 661-671.

Loperfido N., Navarro J., Ruiz J. M., and Sandoval C. J. (2007), Some Relationships Between Skew-Normal Distributions and Order Statistics from Exchangeable Normal Random Vectors. Comm. Statist.- Theor. Meth., **36**, 1719-1733.

Olkin, I., Viana, M. (1995), Correlation analysis of extreme observations from a multivariate normal distribution. J. Amer. Statist. Assoc., **90**, 1373 1379.

Viana, M. A. G. (1998), Linear combinations of ordered symmetric observations with applications to visual acuity, In: Handbook of Statistics. In: N. Balakrishnan, C. R. Rao, (Eds.), Order Statistics: Applications, vol. **17**. North-Holland, Amsterdam, 513-524.