Published in final edited form as:

J Appl Stat. 2015 ; 42(12): 2734–2753. doi:10.1080/02664763.2015.1049130.

Measures of biomarker dependence using a copula-based multivariate epsilon-skew-normal family of distributions

Alan D. Hutson*, Gregory E. Wilding, Terry L. Mashtare, and Albert Vexler

Department of Biostatistics, University at Buffalo, 706 Kimball Tower, 3435 Main St., Buffalo, NY 14214-3000, USA

Abstract

In this note we develop a new multivariate copula model based on epsilon-skew-normal marginal densities for the purpose of examining biomarker dependency structures. We illustrate the flexibility and utility of this model via a variety of graphical tools and a data analysis example pertaining to salivary biomarker. The multivariate normal model is a sub-model of the multivariate epsilon-skew-normal distribution.

Keywords

association; conditional expectation; conditional correlation; conditional dependence

1. Introduction

In early stage clinical phase I and II trials, as well as pre-clinical investigations, it is oftentimes of importance to develop measures of dependence for biomarkers between each other as well as relevant measures of efficacy used in the phase clinical trials III setting, both cross-sectionally and longitudinally. It was Prentice [12] who was one of the first statisticians to put a spotlight on the importance of examining the relative dependence of surrogate endpoints, as opposed to association, which is a critical distinction in that we put dependence at a higher degree of ‘relatedness’ than association, for example, $F(x, y)$ $F(x)F(y)$ versus $E(XY)$ $E(X)E(Y)$. This is of particular importance when examining surrogate endpoints. The dependence measures that we outline in this note can be both conditional on other risk factors such as stage or grade of disease or can be examined unconditionally. We can also handle other issues such as censoring or truncation. The copula models that we will develop as part of this note have the relative flexibility to tackle a wide range of problems, including the surrogate endpoint analyses for several different disease settings.

Unlike analyses based on more straightforward models of association and correlation, for example, calculating the Pearson correlation coefficient or a common odds ratio, the purpose of this note is to develop novel copula-based statistical models of both dependence and

*Corresponding author. ahutson@buffalo.edu.

Disclosure statement

No potential conflict of interest was reported by the authors.

association specific to the biomarker setting. The rationale for developing dependence measures based on copula models is to examine the dependence and interdependence of a set of biomarkers used in clinical trials relative to their merit as surrogate markers of clinical efficacy in early stage investigations and examining the optimal set of biomarkers and disease progression as surrogate measures of treatment efficacy in the phase III setting with the goal of designing phase III trials with shorter follow-up times, that is, utilizing surrogate endpoints which measure treatment efficacy in a shorter period of time as compared to overall survival. In addition to dependence measures, the copula model allows for several different novel approaches to examining various measures of dependence and association, for example, the correlation of two or more variables conditional on other factors, which may or may not be censored.

Biomarker studies themselves represent a substantial portion of the clinical and biostatistical literature. Studies based on biomarkers have been known to be more susceptible to problems, for example, controlling statistical Type I errors due to incorrect dependence structure specifications and multiple comparisons. Furthermore, inappropriate or misleading statistical evaluations have been known to cause biased results, lack of generalizability relative to true efficacy and often contradictory findings relative to the different phases of clinical study [7,8]. Many of these issues found in clinical practice relative to biomarker studies and surrogate endpoints are related to the use of association measures versus dependence measures. We take a copula model-based approach to examining these issues.

Copula models have been a relatively recent development in the statistical and econometrics literature and have been extremely popular modeling tools in financial, business and the social sciences relative to modeling various dependence structures [10,14]. However, they have seen very little usage in biostatistical applications even though the types of problems faced in these seemingly disparate fields have similar dependence structures. Our goal is to take advantage of the power and flexibility of the copula modeling approach in order to develop new and novel copula models and testing methods for use in the clinical trials setting in order to examine both the dependence and association structures found in multivariate biomarker and outcomes data. The ultimate goal is to utilize clinical trial information in an efficient manner in order to improve the phase I, II and III clinical trials processes via defining new endpoints and generating cost-effective and accurate decision rules. In addition, this will provide a new and powerful statistical approach towards examining various sets of multivariate biomarker and correlative study data.

In terms of background a distribution function C on the unit cube $[0, 1]^p$ in $\mathbf{R}^m$ with uniform marginal distributions is called a copula. A classic result due to Sklar [13] relates an arbitrary distribution function F on $\mathbf{R}^m$ to a copula function C via the marginal distribution functions $F_1, F_2, \dots, F_m$ of F using the following theorem:

THEOREM 1 ([13])

Suppose that F is a distribution function on $\mathbf{R}^m$ with one-dimensional marginal distribution functions $F_1, F_2, \dots, F_m$. Then, there is a copula C such that

$$F(x_1, x_2, \dots, x_m) = C(F(x_1), F(x_2), \dots, F(x_m)). \quad (1)$$

If F is continuous, then the copula C satisfying Equation (1) is unique and is given by

$$C(u_1, u_2, \dots, u_m) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_m^{-1}(u_m)) \quad (2)$$

for $u = (u_1, u_2, \dots, u_m) \in (0, 1)^m$, where $F_i^{-1}(u) = \inf x : F_i(x) \geq u, i = 1, \dots, m$.

Conversely, if C is a copula on $(0, 1)^m$ and $F_1, \dots, F_m$ are distribution functions on $\mathbf{R}$, then the function F defined by Equation (1) is a distribution function on $\mathbf{R}^m$ with one-dimensional marginal distributions $F_1, \dots, F_m$.

In this note we utilize the approach of Sklar [13] to have a specific form relative to an m -dimensional multivariate normal form combined with a flexible family of epsilon-skew-normal distributions [4,9]. In Section 2 we provide background on the epsilon-skew-normal distribution, which form our marginal models within the copula framework. In Section 3 we describe a two-stage maximum likelihood approach to estimation and provide some large sample theoretical results. In Section 4 we provide a data analysis example based on salivary biomarker pre-clinical data. We conclude with some remarks about future work.

2. Multivariate epsilon-skew-normal families of distributions

Suppose now that Φ_{Σ} is an m -dimensional multivariate normal distribution function with $m \times 1$ mean vector $\mu = \mathbf{0}$ and $m \times m$ variance covariance matrix Σ with elements equal to 1 for $i = j$ and ρ_{ij} for $\rho_{ij} = \rho_{ji}$, where $i = 1, 2, \dots, m$ and $j = i + 1, i + 2, \dots, m$ with marginal distribution functions $F_1, F_2, \dots, F_m$. Then, there is a copula C such that

$$F(x_1, x_2, \dots, x_m) = C(F_1(x_1), F_2(x_2), \dots, F_m(x_m)). \quad (3)$$

If F is continuous, then the copula C satisfying Equation (3) is unique and is defined to be

$$C(u_1, u_2, \dots, u_m) = \Phi_{\Sigma}(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \dots, \Phi^{-1}(u_m)), \quad (4)$$

where $\Phi^{-1}(\cdot)$ denotes the quantile function of a standard normal random variate. In this framework we still have the result that given $F_1, \dots, F_m$ are distribution functions on $\mathbf{R}$, then the function F defined by Equation (3) is a distribution function on $\mathbf{R}^m$ with one-dimensional marginal distributions $F_1, \dots, F_m$. The result follows straightforward based on standard multivariate normal theory and given standard change-of-variable methods. This approach allows us to choose from a variety of distribution for the F_i 's and makes for a straightforward estimation problem. Our approach focuses on the epsilon-skew-normal family of distributions.

Relative to our choice of F_i , we briefly outline the epsilon-skew-normal distribution [9] in terms of its incorporation in our new copula model. There are several other skewed distributions that may be considered [1]. What we term the 'base model' or the standardized model for the epsilon-skew-normal distribution $ESN(\mu, \sigma, \varepsilon)$ is defined to be a unimodal distribution with the mode at μ and probability mass $(1 - \varepsilon)/2$ below the mode. The

probability density function (p.d.f.), the distribution function (d.f.), and quantile function (q.f.) of its canonical form $ESN(0, 1, \varepsilon)$ are, respectively

$$f_0(x) = \begin{cases} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2(1-\varepsilon)^2}\right) & \text{if } x < 0, \\ \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2(1+\varepsilon)^2}\right) & \text{if } x \geq 0, \end{cases} \quad (5)$$

$$F_0(x) = \begin{cases} (1-\varepsilon)\Phi\left(\frac{x}{1-\varepsilon}\right) & \text{if } x < 0, \\ -\varepsilon + (1+\varepsilon)\Phi\left(\frac{x}{1+\varepsilon}\right) & \text{if } x \geq 0, \end{cases} \quad (6)$$

$$\text{and } Q_0(u) = F_0^{-1}(u) = \begin{cases} (1-\varepsilon)\Phi^{-1}\left(\frac{u}{1-\varepsilon}\right) & \text{if } 0 < u < \frac{(1-\varepsilon)}{2}, \\ (1+\varepsilon)\Phi^{-1}\left(\frac{u+\varepsilon}{1+\varepsilon}\right) & \text{if } \frac{(1-\varepsilon)}{2} \leq u < 1, \end{cases} \quad (7)$$

where $-1 < \varepsilon < 1$, and $\Phi(x)$ denotes the standard normal d.f.

The standard epsilon-skew-normal distribution, $ESN(0, 1, \varepsilon)$, is a mixture of two half-normal distributions and reduces to the standard normal distribution when $\varepsilon = 0$. In other words, $ESN(0, 1, \varepsilon)$ may be viewed as generated by $(1-U)(1-\varepsilon)|N_1| - U(1+\varepsilon)|N_2|$, where U, N_1, N_2 are independent, $P(U=1) = (1+\varepsilon)/2 = 1 - P(U=0)$, and N_1, N_2 are independent standard normal variates. The distribution is skewed right for values of $\varepsilon > 0$ and skewed left for values of $\varepsilon < 0$. The limiting cases of Equation (5) as $\varepsilon \rightarrow \pm 1$ are the well-known half-normal distributions.

The p.d.f. $f_0(\cdot)$ in Equation (5) has derivatives of arbitrary orders. It is differentiable once at the mode. The p.d.f. of $ESN(\mu, \sigma, \varepsilon)$ is given as $f_0((x-\mu)/\sigma)/\sigma$, where $f_0(\cdot)$ is given by Equation (5). Its quantile function, $Q(u) = \mu + \sigma Q_0(u)$, can be used to generate samples from the $ESN(\mu, \sigma, \varepsilon)$ population. Note the relationship $Q((1-\varepsilon)/2) = \mu$.

Mode, mean, variance and median

The mode of $ESN(\mu, \sigma, \varepsilon)$ distribution is μ . It is easy to verify that its mean and variance [9] are

$$E(X) = \mu + \frac{4\sigma\varepsilon}{\sqrt{2\pi}}, \quad (8)$$

the variance is

$$\text{var}(X) = \frac{\sigma^2}{\pi} \left[(3\pi - 8)\varepsilon^2 + \pi \right], \quad (9)$$

respectively, and that its median is

$$Q\left(\frac{1}{2}\right) = \begin{cases} \mu + \sigma(1-\varepsilon)\Phi^{-1}\left(\frac{1}{2(1-\varepsilon)}\right) & \text{if } \varepsilon < 0, \\ \mu + \sigma(1+\varepsilon)\Phi^{-1}\left(\frac{1/2+\varepsilon}{1+\varepsilon}\right) & \text{if } \varepsilon \geq 0. \end{cases} \quad (10)$$

For detailed theoretical results pertaining to this model, see Mudholkar and Hutson [9].

Then, the ESN family of distributions corresponding to marginal distribution function F_i , $i = 1, 2, \dots, m$, is given by

$$u_i = F_i(x_i) = F_0\left(\frac{(x_i - \mu_i)}{\sigma_i}\right), \quad (11)$$

where F_0 is defined in Equation (6). Furthermore, the p.d.f. corresponding to the i th marginal, $i = 1, 2, \dots, m$, is given by

$$f_i(x_i) = \frac{dF_i}{dx_i} = \frac{1}{\sigma_i} f_0\left(\frac{(x_i - \mu_i)}{\sigma_i}\right), \quad (12)$$

where $f_{0,i}$ is defined in Equation (5). Some example shapes for the ESN family as a function of ε are provided in Figures 1–5.

Comment

Note that the multivariate normal distribution is a special case given all ε_i 's are equal to 0.

Comment

It is reasonable to consider different skewed families of distributions such as the skewed- t distribution [5] corresponding to each marginal distribution within the overall modeling framework. However, our focus in this note will be the ESN distribution. Other alternatives for model choice are the so-called multivariate skew-elliptical distributions [2].

3. Estimation

Even though estimation for the multivariate ESN family of distributions via maximum likelihood techniques is a complex numerical methods problem, it can be simplified by breaking the process into two components, marginal estimation and joint estimation. As before, let $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_n$ be i.i.d. distributed as an m -dimensional multivariate ESN family of distribution $F(x_1, x_2, \dots, x_m) = C(F_1(x_1), F_2(x_2), \dots, F_m(x_m))$.

The two-step maximum likelihood estimation process is as follows:

Step 1

Estimate $\mu_i, \sigma_i, \varepsilon_i$, $i = 2, 3, \dots, m$, via maximizing the marginal likelihood per each i [9]. For integer $k = k(x_{(1)}, \dots, x_{(n)})$, $0 \leq k \leq n$, the log-likelihood per each margin $l(\mu_i, \sigma_i^2, \varepsilon_i)$ can be expressed as

$$l(\mu_i, \sigma_i^2, \varepsilon_i) = \begin{cases} -\frac{n}{2} \log 2\pi \sigma_i^2 - \frac{1}{8\sigma_i^2} \left[\sum_{i=1}^n (x_{(i)} - \mu_i)^2 \right] & \text{if } k=0, n, \\ -\frac{n}{2} \log 2\pi \sigma_i^2 - \frac{1}{2\sigma_i^2} \left[\sum_{i=1}^k \frac{(x_{(i)} - \mu_i)^2}{(1 - \varepsilon_i)^2} + \sum_{i=k+1}^n \frac{(x_{(i)} - \mu_i)^2}{(1 + \varepsilon_i)^2} \right] & \text{if } 1 \leq k < n. \end{cases} \quad (13)$$

As noted in Mudholkar and Hutson [9], if $k = 0$ or $k = n$ the maximum likelihood estimate $(\hat{\mu}, \hat{\sigma}^2, \hat{\varepsilon})$ is given by

$$(\hat{\mu}, \hat{\sigma}^2, \hat{\varepsilon}) = \begin{cases} (x_{(1)}, s_0^2, -1) & \text{if } k=0, \\ (x_{(n)}, s_n^2, 1) & \text{if } k=n, \end{cases} \quad (14)$$

where $s_0^2 = \sum_{i=2}^n (x_{(i)} - x_{(1)})^2 / (4n)$ and $s_n^2 = \sum_{i=1}^{n-1} (x_{(i)} - x_{(n)})^2 / (4n)$.

Theorem 4.7 [9] of Mudholkar and Hutson shows that as $n \rightarrow \infty$, the vector of maximum likelihood estimators $\sqrt{n}(\hat{\mu}_{ml} - \mu, \hat{\sigma}_{ml}^2 - \sigma^2, \hat{\varepsilon}_{ml} - \varepsilon)$ is a centered trivariate normal distribution with variance–covariance matrix

$$\Sigma_{ml} = \begin{pmatrix} I^{\mu\mu} = \frac{3\pi(1-\varepsilon^2)\sigma^2}{3\pi-8} & I^{\mu\sigma^2} = 0 & I^{\mu\varepsilon} = -\frac{2\sqrt{2\pi}(1-\varepsilon^2)\sigma}{3\pi-8} \\ I^{\sigma^2\sigma^2} = 2\sigma^4 & I^{\sigma^2\varepsilon} = 0 & I^{\varepsilon\varepsilon} = \frac{\pi(1-\varepsilon^2)}{3\pi-8} \end{pmatrix}. \quad (15)$$

Note that the original paper of Mudholkar and Hutson [9] had a typo where the ‘–’ sign was mistakenly left off of $I^{\mu\varepsilon}$ term in Equation (15).

Step 2

Once the vectors of parameters $\theta_i = (\mu_i, \sigma_i, \varepsilon_i)$ have been estimated as $\hat{\theta}_i = (\hat{\mu}_i, \hat{\sigma}_i, \hat{\varepsilon}_i)$, $i = 2, 3, \dots, m$, then utilize the overall likelihood equations to estimate the parameters ρ_{kl} , $k = 1, 2, \dots, m$ and $l = k + 1, k + 2, \dots, m$, and given as

$$L(x|\theta_i = \hat{\theta}_i) = \prod_{i=1}^n f(x_{i,1}, x_{i,2}, \dots, x_{i,m}) \\ = \frac{\prod_{i=1}^n \phi \sum (\Phi^{-1}(F_1(x_{i,1})), \Phi^{-1}(F_2(x_{i,2})), \dots, \Phi^{-1}(F_m(x_{i,m}))) \prod_{j=1}^m f_j(x_{i,j})}{\prod_{j=1}^m \phi(\Phi^{-1}(F_j(x_{i,j})))}, \quad (16)$$

where $f_j(x_j)$ is given in Equation (12), $F_j(x_j)$ is given in Equation (11). It is then clear that the log-likelihood factors are divided into three components, where the component dependent on the ρ_{kl} , $k = 1, 2, \dots, m$ and $l = 1, 2, \dots, m$, is given as

$$\frac{\log \rho_{kl} \propto -\frac{n}{2} \log(1 - \rho_{kl}^2)}{\sum_{i=1}^n \Phi^{-1}(F_k(x_{i,k}))^2 - 2\rho_{kl} \sum_{i=1}^n \Phi^{-1}(F_k(x_{i,k}))\Phi^{-1}(F_l(x_{i,l})) + \sum_{i=1}^n \Phi^{-1}(F_l(x_{i,l}))^2} \quad (17)$$

such that the following score equation for each ρ_{kl} is given as

$$\frac{n\rho_{kl}}{1-\rho_{kl}^2} - \frac{\rho \left[\sum_{i=1}^n \Phi^{-1}(F_k(x_{i,k}))^2 - 2\rho_{kl} \sum_{i=1}^n \Phi^{-1}(F_k(x_{i,k}))\Phi^{-1}(F_l(x_{i,l})) + \sum_{i=1}^n \Phi^{-1}(F_l(x_{i,l}))^2 \right]}{(1-\rho_{kl}^2)^2} + \frac{\sum_{i=1}^n \Phi^{-1}(F_k(x_{i,k}))\Phi^{-1}(F_l(x_{i,l}))}{1-\rho_{kl}^2} = 0, \quad (18)$$

which involves finding the real root of a cubic equation. This is similar in nature to the maximum likelihood estimation in the bivariate normal case with known means being 0 and variances equal to 1 [6]. One way to facilitate the numerical maximization is to use the moment-based starting point

$$\hat{\rho}_{kl,0} = \frac{\sum_{i=1}^n \left(\Phi^{-1}(F_k(x_{i,k})) - \overline{\Phi^{-1}(F_k)} \right) \left(\Phi^{-1}(F_l(x_{i,l})) - \overline{\Phi^{-1}(F_l)} \right)}{\sqrt{\sum_{i=1}^n \left(\Phi^{-1}(F_k(x_{i,k})) - \overline{\Phi^{-1}(F_k)} \right)^2} \sqrt{\sum_{i=1}^n \left(\Phi^{-1}(F_l(x_{i,l})) - \overline{\Phi^{-1}(F_l)} \right)^2}}, \quad (19)$$

where $\overline{\Phi^{-1}(F_j)} = \sum_{i=1}^n \Phi^{-1}(F_j(x_{i,j})) / n$, $j = 1, 2, \dots, m$, with $\theta_i = (\mu_i, \sigma_i, \varepsilon_i)$ estimated by $\hat{\theta}_i = (\hat{\mu}_i, \hat{\sigma}_i, \hat{\varepsilon}_i)$ from Step 1, $i = 1, 2, \dots, m$. Alternatively, the explicit roots to the solution to the cubic may be developed similar to a special case of the mean being zero and variances set equal to 1 as given by Fosdick and Raftery [3] and are not presented here. The total sample information for each estimator $\hat{\rho}_{kl}$ is given in the appendix. In practice, $\hat{\rho}_{kl,0}$ will be quite close to the true maximum likelihood estimate.

4. Application example

In order to demonstrate some of the potential application of the multivariate ESN family of distributions, we use a data set from a study of candidate biomarkers associated with alveolar bone loss [11]. The data presented in Table 1 represent the first 10 observations of a data set from a sample of size $n = 96$ (data available upon request). The columns represent the log-transformed values of a subset of cytokines that were sampled as part of this research.

Table 2 provides the estimated parameter values for θ_i , σ_i and ε_i along with the corresponding approximate 95% confidence intervals for $i = 1, 2, \dots, 6$. In addition, the estimated expected value and variance estimated from Equations (8) and (9) along with their corresponding 95% approximate confidence intervals are provided. These intervals were estimated via the delta-method approach. The estimated means and variances from the ESN marginal models are very close to their sample counterparts $\bar{x}$ and s^2 (not shown). The calculations were carried forth using SAS PROC NLMIXED (SAS version 9.3, SAS, Cary, NC). The estimated densities based on the ESN model for each measure are plotted in Figures 6–11. An examination of the parameter estimates along with visual inspection of the estimated densities indicates that most of the marginal densities are far from normal, for example, the case $\varepsilon_i = 0$. Only the confidence interval for ε_3 corresponding to log(IL1-b) contains the value of 0. Throughout this example, the indices 1–6 correspond to the measures 1 = log (HGF), 2 = log (IFN), 3 = log (IL-1b), 4 = log (IL-4), 5 = log (IL-6) and 6 = log (IL-8), respectively.

Next, we estimated the dependence measure ρ_{ij} , $i = 1, 2, \dots, 6, j = i + 1, \dots, 6$. The estimates along with the corresponding 95% confidence intervals are contained in Table 3. We see that there appears to be moderate to strong dependence under the ESN multivariate model with the exception of the (2,3) pairing. For illustration purposes, we can plot the joint probability contour plots for any given pair of variables. We chose the (i, j) pairs (1,2), (2,4) and (2,6) in order to illustrate the wide variety of non-elliptical shapes these densities may take. See Figures 12, 13 and 14, which imply that the joint probability densities are far from bivariate normal or even more general symmetric-type elliptical shapes as indicated via the corresponding parameter estimates. In addition, in Figures 15, 16 and 17 for the pairs (1,2), (2,4) and (2,6) we plotted the estimated conditional expectation $E(X_2|X_1)$, $E(X_4|X_2)$ and $E(X_6|X_2)$ for a select set of points, respectively. As you can see from Figure 17, there appears to be strong nonlinear dependence between the two variables, which the ESN multivariate family captures nicely, whereas the Pearson correlation between these two variables is $r = 0.06$. For the other two relationships, we appear to capture a linear dependence given the model assumptions, where the Pearson correlation values for the pairs (1,2) and (2,4) were $r = 0.40$ and $r = 0.95$, respectively. Figure 18 illustrates an estimate of $\text{Var}(X_2|X_1)$. It is clear that even though X_2 and X_1 appear to be linearly related the variance is an increasing function of X_1 . Plots of this type are useful in examining homoscedastic error assumption.

It should be clear from this straightforward example that the multivariate normal model would be a poor choice as a model for this particular data structure as evidenced by the estimates for the ε_i 's and that many of the joint bivariate densities are clearly non-elliptical. There are several other layers of association and dependence, which can be generated from this model. As an example, we calculated the condition correlation $\text{Corr}(X_4, X_5|X_6)$ in order to exam if the linear association between log (IL-4) and log (IL-6) changes conditional on values for log (IL-8). We see from Figure 19 that the linear dependence does appear to be stable over most of the values for X_6 with a slight increase towards across increasing values. This is a highly numerical calculation relying on conditional means, variances and covariances, and hence this contributes to some of the roughness of the plot. More precise estimates would require additional computational time. The actual Pearson correlation between X_4 and X_5 is $r = 0.77$, which implies from Figure 19 that there is a moderate strengthening of the overall association between these two variables conditional on a third factor. This type of model provides an alternative method for calculating conditional correlations to what is currently in the literature [14] and can be extended to conditioning on more than one variable. Similarly, Figures 20 and 21 illustrate how we can examine graphically dependence structures changing between two variables conditional on a third variable. These types of 'moving' contour plots may be quite useful if the extra dimension for which we condition on is time.

5. Future work

In terms of future work based on our newly developed model, we expect to develop an exact test about $H_0 : \rho_{ij} = 0$ for all $i = 1, 2, \dots, 6, j = i + 1, \dots, 6$, as well as any sub-tests that may be of interest. We will examine more complex conditional expectation expressions and consider how to incorporate censoring and truncation into the estimation process. We will

also consider developing confidence bands for the estimates of quantities such as $E(X_i|X_j)$ as well consider longitudinal type data structures.

Acknowledgments

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. We wish to thank the reviewers for their helpful suggestions, which led to a much improved presentation.

References

1. Azzalini, A. The Skew-Normal and Related Families. Cambridge University Press; New York: 2003.
2. Branco MD, Dey DK. A general class of multivariate skew-elliptical distributions. *J Multivar Anal.* 2001; 79:99–113.
3. Fosdick BK, Raftery AE. Estimating the correlation in bivariate normal data with known variances and small sample sizes. *Amer Stat.* 2012; 66:34–41. [PubMed: 23378667]
4. Hutson AD. Utilizing the flexibility of the epsilon-skew-normal distribution for common regression problems. *J Appl Stat.* 2004; 31:673–683.
5. Jones MC, Faddy MJ. A skew extension of the t-distribution, with applications. *J R Stat Soc Ser B.* 2003; 65:159–174.
6. Kendall, SM.; Stuart, A. The Advanced Theory of Statistics. 4th. Vol. 2. Charles Griffin & Company Limited; London: 1979.
7. Mann U, Srivastava RG, Srivastava S. Recent advances in biomarkers for cancer diagnosis and treatment. *Drug Discov Today.* 2005; 10:965–976. [PubMed: 16023055]
8. Mayeux R. Biomarkers: Potential uses and limitations. *NeuroRx.* 2004; 1:182–188. [PubMed: 15717018]
9. Mudholkar GS, Hutson AD. The epsilon-skew-normal distribution for analyzing near-normal data. *J Stat Plann Infer.* 2000; 83:291–309.
10. Nelson, RB. An introduction to copulas. 2nd. Springer; New York: 2010.
11. Ng PYB, Donley M, Hausmann E, Hutson AD, Rossomando EF, Scannapieco FA. Candidate salivary biomarkers associated with alveolar bone loss: cross-sectional and in vitro studies. *Immunol Med Microbiol.* 2007; 49:252–260.
12. Prentice RL. Surrogate endpoints in clinical trials: Definition and operational criteria. *Stat Med.* 1989; 8:431–440. [PubMed: 2727467]
13. Sklar A. Fonctions de répartition à n dimensions et leurs marges. *Publ Inst Statist Univ Paris.* 1959; 8:229–231.
14. Wilding GE, Cai X, Hutson AD, Yu Z. A linear model-based test for the heterogeneity of conditional correlations. *J Appl Stat.* 2011; 38:2355–2366.

Figure 1.
Some typical ESN probability density functions.

Figure 2.
Example multivariate ESN family for parameter values $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$, $\varepsilon_1 = \varepsilon_2 = 0$ and $\rho_0 = 0.5$.

Figure 3.
Example multivariate ESN family for parameter values $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$, $\varepsilon_1 = \varepsilon_2 = 0.5$ and $\rho_0 = 0.5$.

Figure 4. Example multivariate ESN family for parameter values $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$, $\varepsilon_1 = 0.5$ and $\varepsilon_2 = -0.5$ and $\rho_0 = 0.5$.

Figure 5.
Example multivariate ESN family for parameter values $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2 = 1$, $\varepsilon_1 = 0.9$ and $\varepsilon_2 = -0.25$ and $\rho_0 = 0.5$.

Figure 6.
Estimated ESN family-based density for $\log(\text{HGF})$.

Figure 7.
Estimated ESN family-based density for $\log(\text{IFN})$.

Figure 8.
Estimated ESN family-based density for $\log(\text{IL-1b})$.

Figure 9.
Estimated ESN family-based density for $\log(\text{IL-4})$.

Figure 10.
Estimated ESN family-based density for $\log(\text{IL-6})$.

Figure 11.
Estimated ESN family-based density for $\log(\text{IL-8})$.

Figure 12.
Estimated ESN family-based contour plot of log (HGF) on the x -axis versus log (IFN) on the y -axis.

Figure 13.
Estimated ESN family-based contour plot of log (IFN) on the x -axis versus log (IL-4) on the y -axis.

Figure 14. Estimated ESN family-based contour plot of log (IFN) on the x -axis versus log (IL-8) on the y -axis.

Figure 15.
Estimated conditional expectation of log (IFN) on the y-axis given values of log (HGF) on the x-axis.

Figure 16.
Estimated conditional expectation of log (IL-4) on the y-axis given values of log (IFN) on the x-axis.

Figure 17.
Estimated conditional expectation of log (IL-8) on the y-axis given values of log (IFN) on the x-axis.

Figure 18.
Estimated conditional variance of log (IFN) on the y-axis given values of log (HGF) on the x-axis.

Figure 19. Estimated conditional correlation of log (IL-4) and log (IL-6) on the y-axis given values of log (IL-8) on the x-axis.

Figure 20.
Estimated ESN family-based conditional contour plot of log (IFN) on the x -axis versus log (IL-4) on the y -axis given $\log(\text{IL-8}) = 0.5$

Figure 21.
Estimated ESN family-based conditional contour plot of log (IFN) on the x -axis versus log (IL-4) on the y -axis given $\log(\text{IL-8}) = 2$.

Table 1

Example data: salivary biomarker data (first 10 observations out of a dataset of sample size $n = 96$).

$\log(\text{HGF})$	$\log(\text{IFN})$	$\log(\text{IL-1b})$	$\log(\text{IL-4})$	$\log(\text{IL-6})$	$\log(\text{IL-8})$
0.88	1.62	2.93	2.29	2.23	1.1
0.49	0.28	0.32	0.63	0.38	0.32
0.8	0.8	0.84	1.11	0.97	1.02
0.82	0.77	0.82	1.09	0.93	1
0.79	0.77	0.84	1.08	0.98	0.97
0.87	0.88	1.38	1.18	1.09	1.86
0.94	0.91	1.69	1.24	1.28	2.24
1.18	1.11	2.11	1.44	1.53	2.65
1.52	1.48	2.57	2.05	1.91	3.12
1.96	1.87	3.02	2.73	2.36	3.55
⋮	⋮	⋮	⋮	⋮	⋮

Table 2

Marginal parameter estimates for cytokine data and corresponding large sample 95% confidence intervals.

i	$\hat{\mu}_i$	$\hat{\sigma}_i$	$\hat{\epsilon}_i$	$E(\widehat{X}_j)$	$\widehat{\text{Var}}(\widehat{X}_j)$
1	1.48(1.29,1.67)	0.31(0.22,0.40)	0.33(0.14,0.52)	1.78(1.66,1.89)	0.33(0.23,0.42)
2	0.68(0.61,0.76)	0.20(0.15,0.26)	0.73(0.63,0.83)	1.21(1.11,1.31)	0.25(0.18,0.33)
3	2.70(2.49,2.73)	0.43(0.31,0.55)	-0.10(-.28, 0.07)	2.59(2.45,2.73)	0.43(0.31,0.56)
4	0.94(0.86,1.01)	0.25(0.18,0.32)	0.80(0.72,0.88)	1.58(1.46,1.69)	0.32(0.23,0.42)
5	0.96(0.88,1.04)	0.22(0.16,0.29)	0.67(0.57,0.78)	1.47(1.37,1.58)	0.27(0.19, 0.35)
6	3.78(3.57,3.99)	0.41(0.29,0.52)	-0.53(-0.70, -0.35)	3.23(3.10,3.38)	0.46(0.32,0.60)

Table 3
Estimates for ρ_{ij} for cytokine data and corresponding large sample 95% confidence intervals.

$\hat{\rho}_{ij}$	1	2	3	4	5	6
1	1	0.30(0.18,0.42)	0.36(0.25,0.47)	0.39(0.27,0.50)	0.44(0.32,0.56)	0.84(0.80,0.89)
2		1	-0.01(-0.09, 0.07)	0.88(0.85,0.92)	0.88(0.84,0.92)	0.39(0.22,0.55)
3			1	0.48(0.35,0.61)	0.62(0.52,0.72)	0.83(0.79,0.88)
4				1	0.92 (0.90,0.94)	0.37 (0.23,0.51)
5					1	0.46(0.33,0.60)
6						1