

MULTIVARIATE DISPERSION ORDER AND THE NOTION OF COPULA APPLIED TO THE MULTIVARIATE t -DISTRIBUTION

J. P. ARIAS-NICOLÁS

*Department of Mathematics
University of Extremadura
Extremadura, Spain
E-mail: jparias@unex.es*

J. M. FERNÁNDEZ-PONCE AND P. LUQUE-CALVO

*Department of Statistics
University of Sevilla
41013 Sevilla, Spain
E-mail: ferpon@us.es; calvo@us.es*

A. SUÁREZ-LLORENS

*Department of Statistics
University of Cádiz
11002 Cádiz, Spain
E-mail: alfonso.suarez@uca.es*

We study the concept of multivariate dispersion order, defined as the existence of an expansion function that maps a random vector to another one, for multivariate distributions with the same dependence structure. As a particular case, we can order the multivariate t -distribution family in dispersion sense. Finally, we use these results in the problem of detection and characterization of influential observations in regression analysis. This problem can often be used to compare two multivariate t -distributions.

1. INTRODUCTION

Stochastic orderings arise in statistical decision theory in the comparison of experiments and estimation problems (see [17]). In particular, dispersion has been used

to characterize the variability for distributions and it has been extensively studied (see [5,10,13,15,16], among others).

For univariate and multivariate distributions, the concept of dispersion is fundamental; a statistical research is unthinkable for a phenomenon without variability. Unfortunately, there is not a unique definition of dispersion and this problem is much more complicated for distributions on \mathbb{R}^n . For univariate distributions, Lewis and Thompson [13] introduced a concept of variability through the definition of the dispersion order (LT sense). Let F and G be two distribution functions; we say that F is less dispersive than G , denoted $F <_{\text{Disp}} G$, if any pair of quantiles of G are at least as widely separated as the corresponding quantiles of F . Let u be a real value in $(0,1)$; we use the definition of univariate quantile as follows:

$$Q_X(u) = F_X^-(u) = \inf\{x : F_X(x) \geq u\}.$$

Many useful characterizations of the dispersion order can be found in the literature. An excellent handbook is that by Shaked and Shanthikumar [17]. One of the most interesting characterizations of this univariate order is given in [16]. Let F and G be two strictly increasing and absolutely continuous distribution functions; then $X <_{\text{Disp}} Y$ or $F <_{\text{Disp}} G$, if and only if there exists a function $\Phi : S_F \rightarrow S_G$ (where S_F and S_G are the support of F and G , respectively) such that $Y =_{\text{st}} \Phi(X)$ and $\Phi'(x) \geq 1$ for all x in S_F . Note that under the last condition, the function Φ is an expansion function; that is, the function Φ verifies $\Phi(x) - \Phi(x') \geq x - x'$, for all $x > x'$. Hence, the dispersion ordering in the LT sense is based on the existence of an expansion function that depends on the corresponding distribution functions. Furthermore, in this case, $\Phi(x) = Q_Y(F_X(x))$ for all x in S_F .

An extension of the univariate dispersion order to the multivariate case was given by Giovagnoli and Wynn [9]. A function $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is called an *expansion* if

$$\|\Phi(\mathbf{x}) - \Phi(\mathbf{x}')\|_2 \geq \|\mathbf{x} - \mathbf{x}'\|_2 \quad \text{for all } \mathbf{x} \text{ and } \mathbf{x}' \text{ in } \mathbb{R}^n.$$

Let \mathbf{X} and \mathbf{Y} be two n -dimensional random vectors. Suppose that

$$\mathbf{Y} =_{\text{st}} \Phi(\mathbf{X}) \quad \text{for some expansion function } \Phi.$$

Then we say that \mathbf{X} is less than \mathbf{Y} in the strong multivariate dispersion order (denoted by $\mathbf{X} \leq_{\text{SD}} \mathbf{Y}$).

Roughly speaking, the strong multivariate dispersive order is based on the existence of an expansion function that maps stochastically a random vector to another one. Obviously, the ordering in the \leq_{SD} sense is intuitively reasonable and it satisfies many desirable properties. For instance, the strong dispersion ordering implies that $\|\mathbf{X} - \mathbf{X}'\|_2 <_{\text{st}} \|\mathbf{Y} - \mathbf{Y}'\|_2$, where $\|\cdot\|_2$ corresponds to the Euclidean norm and \mathbf{X}' and \mathbf{Y}' are two independent values for \mathbf{X} and \mathbf{Y} , respectively. It also implies that $\text{Tr}(\Sigma_X) \leq \text{Tr}(\Sigma_Y)$, where Σ_X and Σ_Y are the covariance matrices for \mathbf{X} and \mathbf{Y} , respectively, see Giovagnoli and Wynn [9]. As a consequence of these properties, the $<_{\text{SD}}$ multivariate order has a clear interpretation in dispersion terms.

In the multivariate case, there exist several transformations that map one multi-dimensional random variable to another one. For this reason, it seems intuitive to define a multivariate dispersion order based on a particular transformation. Note that in the univariate case, the function Φ depends on the corresponding distribution functions, so it has a unique expression. These considerations led Fernández-Ponce and Suárez-Llorens [7] to define a concept of multivariate dispersion order based on the existence of a particular expansion function and, in addition, has a particular interpretation as multivariate quantiles more widely separated.

From this point forward, we assume that X_1, \dots, X_n have an absolutely continuous joint distribution and the corresponding conditioned variables are also absolutely continuous with density functions strictly positive. Let $\mathbf{u} = (u_1, \dots, u_n)$ be a vector in $[0, 1]^n$. The multivariate \mathbf{u} -quantile for \mathbf{X} , denoted as $\hat{\mathbf{x}}(\mathbf{u})$, is defined as follows:

$$\begin{aligned} \hat{x}_1(u_1) &= Q_{X_1}(u_1), \\ \hat{x}_2(u_1, u_2) &= Q_{X_2|X_1=\hat{x}_1(u_1)}(u_2), \\ &\vdots \\ \hat{x}_n(u_1, \dots, u_n) &= Q_{X_n|\cap_{j=1}^{n-1} X_j=\hat{x}_j(u_j)}(u_n). \end{aligned}$$

This last construction is widely used in simulation theory, and it is named the standard construction. Obviously, the standard construction depends on the choice of the ordering of the marginal distributions. The notion of the standard construction can be interpreted as a quantile function on \mathbb{R}^n . Fernández-Ponce and Suárez-Llorens [7] provide the notion of corrected orthant associated to the standard construction that interprets the accumulated probability in all orthants.

The definition of the multivariate \mathbf{u} -quantile for \mathbf{X} led us to define the multivariate \mathbf{x} -rate vector, denoted $\hat{\mathbf{x}}^*(\mathbf{x})$, as

$$\begin{aligned} \hat{x}_1^*(x_1) &= F_{X_1}(x_1), \\ \hat{x}_2^*(x_1, x_2) &= F_{X_2|X_1=x_1}(x_2) \\ &\vdots \\ \hat{x}_n^*(x_1, \dots, x_n) &= F_{X_n|\cap_{j=1}^{n-1} X_j=x_j}(x_n). \end{aligned}$$

Under the notion of the standard construction and the interpretation as multivariate quantiles, Fernández-Ponce and Suárez-Llorens [7] defined the multivariate dispersive order as a generalization of the univariate dispersive order in the LT sense.

DEFINITION 1: Let \mathbf{X} and \mathbf{Y} be two random vectors in \mathbb{R}^n . We say that \mathbf{X} is less than \mathbf{Y} in dispersion sense, denoted as $\mathbf{X} <_{\text{Disp}} \mathbf{Y}$, if

$$\|\hat{\mathbf{x}}(\mathbf{v}) - \hat{\mathbf{x}}(\mathbf{u})\|_2 \leq \|\hat{\mathbf{y}}(\mathbf{v}) - \hat{\mathbf{y}}(\mathbf{u})\|_2,$$

for all \mathbf{u} and \mathbf{v} in $(0, 1)^n$.

Note that this new ordering depends on the chosen permutation (see [7]). Definition 1 defines a multivariate dispersion ordering as quantiles that are more widely separated. Theorem 1 characterizes the multivariate dispersion order by means of a particular expansion function.

THEOREM 1: *Let \mathbf{X} and \mathbf{Y} be two random vectors in \mathbb{R}^n with their distribution functions satisfying the regularity conditions. Then it holds that $\mathbf{X} <_{\text{Disp}} \mathbf{Y}$ if and only if there exists a function Φ such that*

$$\begin{aligned} \Phi(\mathbf{X}) &=_{\text{st}} \mathbf{Y} \quad \text{with } [\Phi(\mathbf{x})]_i = \Phi_i(x_1, \dots, x_i), \\ I_n &<_L J_\Phi(\mathbf{x})^t J_\Phi(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}^n, \end{aligned} \tag{1}$$

where $J_\Phi(\mathbf{x})$ is the Jacobian matrix of Φ .

$$\frac{\partial \Phi_i(x_1, \dots, x_i)}{\partial x_i} \geq 0.$$

Moreover, in this case,

$$\Phi_i(x_1, \dots, x_i) = (\hat{y}_i \circ \hat{x}_i)(x_1, \dots, x_i).$$

Note that the symbol $<_L$ represents the well-known Loewner ordering of matrices, where $A <_L B$ if and only if the matrix $B - A$ is nonnegative definite. Condition (1) implies that the function Φ is an expansion function; that is, it holds that

$$\|\Phi(\mathbf{y}) - \Phi(\mathbf{x})\|_2 \geq \|\mathbf{y} - \mathbf{x}\|_2$$

for all \mathbf{x}, \mathbf{y} in \mathbb{R}^n (see [9]).

To summarize, the multivariate dispersion order can be simplified just by checking whether the multivariate function $\Phi = (\Phi_1, \dots, \Phi_n)$, defined as

$$\begin{aligned} \Phi_i(x_1, \dots, x_i) &= (\hat{y}_i \circ \hat{x}_i)(x_1, \dots, x_i) \\ &= Q_{Y_i | \bigcap_{j=1}^{i-1} Y_j = \Phi_j(x_1, \dots, x_j)}(F_{X_i | \bigcap_{j=1}^{i-1} X_j = x_j}(x_i)) \end{aligned} \tag{2}$$

for $i = 1, \dots, n$, is an expansion function.

It is apparent that this ordering is a generalization of the dispersive ordering in the LT sense. The multivariate dispersive ordering is characterized through a particular expansion function, so it obviously implies the strong multivariate dispersive ordering. Then we will always consider the multivariate concept of dispersion in the Giovagnoli and Wynn sense. However, when we study the particular case defined by function (2), we will call the strong multivariate dispersion order as the multivariate dispersion order, denoted $<_{\text{Disp}}$.

The organization in this article is as follows. In Section 2, we will show the multivariate dispersion order between two multivariate random variables, with the same copula, is characterized by the univariate dispersion order for the corresponding marginal distributions. In Section 3, we will use the results from Section 2 to order the multivariate t -Student family in a dispersion sense, $<_{\text{Disp}}$ or $<_{\text{SD}}$, accord-

ing to properties of the precision matrix and the degrees of freedom. Finally, in Section 4, we will apply this ordering in the problem of the detection and characterization of influential observations in regression analysis. This problem can be often used to compare two multivariate t -distributions.

2. MULTIVARIATE DISPERSION ORDER UNDER THE NOTION OF DISTRIBUTION FUNCTIONS WITH THE SAME COPULA

Now, we characterize the multivariate dispersion order, $<_{\text{Disp}}$, for random variables with the same dependence structure in the sense that they have the same copula. A copula C is a cumulative distribution function with uniform margins on $[0, 1]$. Furthermore, it is shown that if H is an n -dimensional distribution function with margins F_1, \dots, F_n , then there exists an n -copula C such that for all \mathbf{x} in \mathbb{R}^n , it holds that

$$H(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)).$$

Moreover, if F_1, \dots, F_n are continuous, then C is unique (for more details about copulas, see Nelsen [14]). It follows that if $\mathbf{X} = (X_1, \dots, X_n)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$ are two n -dimensional random variables, then they have the same copula if and only if

$$(X_1, \dots, X_n) =_{\text{st}} (Q_{X_1}[F_{Y_1}(Y_1)], \dots, Q_{X_n}[F_{Y_n}(Y_n)]).$$

Within this setting, we can formulate the following theorem.

THEOREM 2: *Let $\mathbf{X} = (X_1, \dots, X_n)$ and $\mathbf{Y} = (Y_1, \dots, Y_n)$ be two n -dimensional random vectors such that they have the same copula, denoted as C . Then $\mathbf{X} <_{\text{Disp}} \mathbf{Y}$ if and only if $X_i <_{\text{Disp}} Y_i$ for all $i = 1, \dots, n$.*

PROOF: In light of Theorem 1, it is only necessary to prove that the component Φ_i of the function Φ has the following expression:

$$\Phi_i(x_1, \dots, x_i) = Q_{Y_i}(F_{X_i}(x_i)), \tag{3}$$

for $i = 1, \dots, n$. In other words, the component Φ_i of the function Φ , which maps the random vector \mathbf{X} to \mathbf{Y} , only depends on the i th marginal variable. Note that if (3) holds, the Jacobian matrix of Φ is a diagonal matrix in which the diagonal elements are the functions that map the univariate marginal distribution of \mathbf{X} to the corresponding one of \mathbf{Y} . Hence, the condition (1) in Theorem 1 is apparent.

The proof of (3) is by mathematical induction. For $n = 1$, it is trivial. Let us assume that it is true for $i = 1, \dots, j - 1$; then we need to show it for $i = j$. By induction hypothesis, it holds that

$$\Phi_j(x_1, \dots, x_j) = Q_{Y_j}(\bigcap_{i=1}^{j-1} Y_i = Q_{Y_i}(F_{X_i}(x_i))) (F_{X_j}(\bigcap_{i=1}^{j-1} X_i = x_i)(x_j)).$$

Furthermore, in light of the equality 2.9.1 in Nelsen [14], it is easy to show that

$$P \left[X_j \leq Q_{X_j}(v) \mid \bigcap_{i=1}^{j-1} X_i = Q_{X_i}(u_i) \right] = \frac{\partial C(u_1, \dots, u_{j-1}, v)}{\partial u_1, \dots, \partial u_{j-1}}, \tag{4}$$

where C is the copula of the distribution function \mathbf{F} . By assumption, \mathbf{F} and \mathbf{G} have the same copula; hence, using (4), it holds that

$$F_{X_j}(Q_{X_j|\cap_{i=1}^{j-1} X_i=Q_{X_i}(u_i)}(p)) = G_{Y_j}(Q_{Y_j|\cap_{i=1}^{j-1} Y_i=Q_{Y_i}(u_i)}(p)), \tag{5}$$

for all $0 \leq p \leq 1$. Now, if we take $u_i = F_{X_i}(x_i)$ in (5), we obtain that

$$Q_{Y_j|\cap_{i=1}^{j-1} Y_i=Q_{Y_i}(F_{X_i}(x_i))}(p) = Q_{Y_j}(F_{X_j}(Q_{X_j|\cap_{i=1}^{j-1} X_i=x_i}(p))). \tag{6}$$

Note that if we now consider

$$p = F_{X_j|\cap_{i=1}^{j-1} X_i=x_i}(x_j)$$

in (6), it is easy to verify that

$$\begin{aligned} \Phi_j(x_1, \dots, x_j) &= Q_{Y_j|\cap_{i=1}^{j-1} Y_i=Q_{Y_i}(F_{X_i}(x_i))}(F_{X_j|\cap_{i=1}^{j-1} X_i=x_i}(x_j)) \\ &= Q_{Y_j}(F_{X_j}(Q_{X_j|\cap_{i=1}^{j-1} X_i=x_i}(F_{X_j|\cap_{i=1}^{j-1} X_i=x_i}(x_j)))) \\ &= Q_{Y_j}(F_{X_j}(x_j)). \end{aligned}$$

Therefore, the proof is concluded. ■

COROLLARY 1: *Let (X_1, \dots, X_n) be a multivariate random vector. Let us consider h_i a univariate strictly increasing expansion function for $i = 1, \dots, n$. Then it holds that*

$$(X_1, \dots, X_n) <_{\text{Disp}} (h_1(X_1), \dots, h_n(X_n)).$$

PROOF: The corollary follows from the facts that any copula is invariant under monotone increasing transformations (see [14]) and the univariate dispersion order is invariant under strictly increasing expansion functions (see [17]). ■

Note that this corollary provides many possible comparisons. In particular, if we take a real number $a > 1$, then it is easy to show that $\mathbf{X} <_{\text{Disp}} a\mathbf{X}$.

Example 1: Let $\mathbf{X} \rightsquigarrow N_n(\mu_1, \Sigma_1)$ and $\mathbf{Y} \rightsquigarrow N_n(\mu_2, \Sigma_2)$ be two multivariate normal distributions. If $\rho_{ij}^{\mathbf{X}} = \rho_{ij}^{\mathbf{Y}}$ for all i and j and $\sigma_i^{\mathbf{X}} < \sigma_i^{\mathbf{Y}}$, then $\mathbf{X} \leq_{\text{Disp}} \mathbf{Y}$. It is well known that under the last assumption, \mathbf{X} and \mathbf{Y} have the same copula. Thus, the proof is apparent using Theorem 2.

3. DISPERSION PROPERTIES OF THE MULTIVARIATE STUDENT DISTRIBUTION

In this section, we apply some results obtained in the last section to the particular t -distribution family. For this purpose, we use the corresponding definition of the t -distribution from Bernardo and Smith [2, pp. 139–140]. A continuous random vector \mathbf{X} has a multivariate t -distribution or a multivariate Student distribution of dimension k , with parameters $\mu = (\mu_1, \dots, \mu_k)$, Σ , and n , where $\mu \in \mathbb{R}^k$, Σ is

a symmetric positive-definite $k \times k$ matrix and $n > 0$ if its probability density function, denoted $St_k(\mathbf{x}|\mu, \Sigma, n)$, is

$$St_k(\mathbf{x}|\mu, \Sigma, n) = c \left[1 + \frac{1}{n} (\mathbf{x} - \mu)^t \Sigma (\mathbf{x} - \mu) \right]^{-(n+k/2)}, \quad \mathbf{x} \in \mathbb{R}^k,$$

where

$$c = \frac{\Gamma((n+k)/2)}{\Gamma(n/2)(n\pi)^{k/2}} |\Sigma|^{1/2}.$$

Although not exactly equal to the inverse of the covariance matrix, the parameter Σ is often referred to as the precision matrix of the distribution or, equivalently, the inverse matrix of the dispersion matrix. In the general case, $E[\mathbf{X}] = \mu$ and $\text{Var}(\mathbf{X}) = \Sigma^{-1}[n/(n-2)]$.

Before introducing the results in this section, we need the definition of a univariate partial order strongly connected with the univariate dispersive ordering. We consider the tail ordering defined by Lawrence [12]. Let X and Y be two univariate random variables symmetric about zero; then we say that X is less in the tail order sense, denoted $X <_r Y$, if the ratio $Q_Y(u)/Q_X(u)$ is nondecreasing (nonincreasing) for $u \in (\frac{1}{2}, 1)$ ($u \in (0, \frac{1}{2})$). In the following theorem, we will use the definition of the tail ordering to order the univariate t -Student family.

THEOREM 3: *Let $St_1(0, 1, m)$ and $St_1(0, 1, n)$ be two univariate t -distributions. Then if $n < m$, it holds that $St_1(0, 1, m) <_{\text{Disp}} St_1(0, 1, n)$.*

PROOF: To simplify, denote t_n as the univariate t -distribution with n degrees of freedom. Capéraà [3] showed that if $n \leq m$, then $t_m <_r t_n$. In addition, Doksum [5] showed that for univariate absolutely continuous distributions with $F(0) = G(0) = 0$ such that $f(0) \geq g(0) > 0$ and $Q_Y(u)/Q_X(u)$ nondecreasing for all $u \in (0, 1)$, it holds that $F <_{\text{Disp}} G$.

Under the last discussion, we consider the random variable $|t_n|$ with the density function given by

$$f_{|t_n|}(t) = 2f_{t_n}(t) \quad \text{if } t > 0 \text{ and } 0 \text{ otherwise.}$$

A straightforward computation shows that the distribution function of $|t_n|$ is $F_{|t_n|}(x) = 2F_{t_n}(x) - 1$ for $x \geq 0$. Hence, $Q_{|t_n|}(u) = Q_{t_n}((u+1)/2)$ for all u in the interval $(0, 1)$. Therefore, it is apparent, using the work of Capéraà [3], that if $n \leq m$, then $Q_{|t_m|}(u)/Q_{|t_n|}(u)$ is nondecreasing for all $u \in (0, 1)$. Since $F_{|t_m|}(0) = F_{|t_n|}(0) = 0$ and, of course $f_{|t_m|}(0) > f_{|t_n|}(0)$, we obtain, using the result in [5] that $|t_m| <_{\text{Disp}} |t_n|$. It is easy to check, using properties of symmetry, that $|t_m| <_{\text{Disp}} |t_n|$ implies that $t_m <_{\text{Disp}} t_n$. ■

Note that although in the literature the degrees of freedom of a t -distribution are always associated with the dispersion and the lack of knowledge of the experiment, the previous result provides that the t -distribution family is ordered in a really

strict dispersion order. To study in depth the implications of the univariate dispersion order, see Shaked and Shanthikumar [17].

We have ordered two univariate t -distributions pertaining to the degrees of freedom. If we consider the more general class when the precision is different, the following corollary holds.

COROLLARY 2: *Let $St_1(0, \sigma_1, m)$ and $St_1(0, \sigma_2, n)$ be two univariate t -distributions that satisfy that for $n \leq m$ and $\sigma_2 \leq \sigma_1$, $St_1(0, \sigma_1, m) <_{Disp} St_1(0, \sigma_2, n)$.*

PROOF: The proof is apparent. ■

Note that the precision matrix is related to the variance through the expression $Var(St_1(0, \sigma_1, m)) = \sigma_1^{-1}(m/m - 2)$.

From this point forward, we will consider two multivariate t -distributions. We generalize the results obtained in Theorem 3 and Corollary 2 to the multivariate case.

THEOREM 4: *Let $\mathbf{Y}_n \sim St_k(0, \Sigma, n)$ and $\mathbf{Y}_m \sim St_k(0, \Sigma, m)$ be two multivariate t -distributions with the same precision matrix and different degrees of freedom. Then if $n \leq m$, it holds that $St_k(0, \Sigma, m) <_{Disp} St_k(0, \Sigma, n)$.*

Although, of course, Theorem 4 is more general than Theorem 3, we first need the results of this one to prove the multivariate case.

PROOF: Let $\mathbf{Y}_m = (Y_{m,1}, \dots, Y_{m,k})$ and $\mathbf{Y}_n = (Y_{n,1}, \dots, Y_{n,k})$ be the corresponding multivariate t -distributions. First, we need to prove that two multivariate t -distributions with the same precision matrix have the same copula. For this purpose, we define the random vector \mathbf{X} such as

$$\mathbf{X} = (X_1, \dots, X_k) = [Q_{Y_{m,1}}(F_{Y_{n,1}}(Y_{n,1})), \dots, Q_{Y_{m,k}}(F_{Y_{n,k}}(Y_{n,k}))].$$

From Exercise 2.15 in Nelsen [14], it holds that the multivariate distributions \mathbf{X} and \mathbf{Y}_n have the same copula. We only have to prove that $\mathbf{X} =_{st} \mathbf{Y}_m$.

First, from the well-known result that the function $Q_{Y_{m,i}}(F_{Y_{n,i}})(\cdot)$ for $i = 1, \dots, k$, maps the univariate random variable $Y_{n,i}$ to $Y_{m,i}$, we know that all marginal distributions of \mathbf{X} are generalized t -distributions. Hence, using Theorem 1 from Arellano-Valle and Bolfarine [1], we know that \mathbf{X} is a multivariate t -distribution with parameters $\mathbf{X} \sim St_k(0, \Sigma, m)$. From the fact that \mathbf{X} and \mathbf{Y}_m have the same degrees of freedom, it is only necessary to show that $Var(\mathbf{X}) = Var(\mathbf{Y}_m)$. Of course, it holds that $Var(X_i) = Var(Y_{m,i})$ for all $i = 1, \dots, n$. Hence, both matrices have the same diagonal elements.

Since \mathbf{X} and \mathbf{Y}_n have the same copula, it is easy to show, using Theorem 5.1.3 in Nelsen [14], that

$$\tau_{X_i, X_j} = \tau_{Y_{n,i}, Y_{n,j}},$$

where $\tau_{X,Y}$ is the population version of Kendall’s tau for X and Y . Therefore, using Theorem 3.3 from Frahm, Junker, and Szimayer [8], it holds that they also have the same Pearson’s coefficient, ρ , so

$$\frac{\text{cov}(X_i, X_j)}{\sigma_{X_i} \sigma_{X_j}} = \frac{\text{cov}(Y_{n,i}, Y_{n,j})}{\sigma_{Y_{n,i}} \sigma_{Y_{n,j}}}.$$

Using the fact that \mathbf{Y}_n and \mathbf{Y}_m have the same precision matrix, if we denote $\text{Var}(\mathbf{Y}_n) = (\sigma_{ij,n})$ and $\text{Var}(\mathbf{Y}_m) = (\sigma_{ij,m})$, it holds that

$$\frac{\sigma_{ij,m}}{\sigma_{ij,n}} = \left[\frac{m(n-2)}{n(m-2)} \right]^{1/2},$$

for all $i, j = 1, \dots, k$. Hence, it is apparent that $\text{cov}(X_i, X_j) = \text{cov}(Y_{m,i}, Y_{m,j})$ and, of course, $\tilde{\Sigma} = \Sigma$, which implies that $\mathbf{Y}_m =_{\text{st}} \mathbf{X}$.

We have already shown that \mathbf{Y}_m and \mathbf{Y}_n have the same copula. Hence, using first Corollary 2 and then Theorem 2, we have that $\mathbf{Y}_m \prec_{\text{Disp}} \mathbf{Y}_n$. ■

At this point, we focus our attention on ordering two multivariate t -distributions with different precision matrices.

THEOREM 5: *Let $\mathbf{Y}_1 \sim \text{St}_k(0, \Sigma_1, n)$ and $\mathbf{Y}_2 \sim \text{St}_k(0, \Sigma_2, n)$ be two multivariate t -distributions with different precision matrices and the same degrees of freedom. Then the following conditions are equivalent:*

1. $\mathbf{Y}_2 =_{\text{st}} k(\mathbf{Y}_1)$, where k is one-to-one, linear, and expansion.
2. There is an orthogonal matrix Γ such that $\Sigma_2^{-1} \geq_L \Gamma \Sigma_1^{-1} \Gamma^t$.
3. $\lambda(\Sigma_2^{-1}) \geq \lambda(\Sigma_1^{-1})$ (where $\lambda(\cdot)$ is the vector of ordered eigenvalues and \leq refers to the usual entrywise ordering).

PROOF: It is analogous to Theorem 4 in Giovagnoli and Wynn [9]. ■

Note that if both distributions have the same degrees of freedom, it is easy to find several linear transformations that map one t -distribution to the other one. Obviously, this not the case when they have different degrees of freedom; the possible transformations are clearly not linear.

It is necessary to take in account that Theorem 5 does not provide a characterization of the multivariate dispersion order. As we mentioned earlier, the \prec_{Disp} ordering is associated to a particular transformation given by the function in (2). Just looking at Example 4.1 in [7] for the multivariate Normal distributions, it is easy to show the linear expression of Φ when we are interested in comparing two multivariate t -distributions. This expression depends on the Cholesky decomposition of the matrices Σ_2^{-1} and Σ_1^{-1} .

However, in the conditions of Theorem 5, it holds that $\text{St}_k(0, \Sigma_1, n) \leq_{\text{SD}} \text{St}_k(0, \Sigma_2, n)$. We emphasize that \leq_{SD} is a weaker multivariate dispersion order than the \prec_{Disp} ordering. Hence, at least there exists an expansion function that maps one t -distribution to the other one. As we will show in Section 4, we only need to know about the \leq_{SD} order in order to define a new measure of Bayesian influence.

Corollary 3 is needed to compare t -distributions when both degrees of freedom and precision matrices are different.

COROLLARY 3: Let $\mathbf{Y}_1 \sim \text{St}_k(0, \boldsymbol{\Sigma}_1, m)$ and $\mathbf{Y}_2 \sim \text{St}_k(0, \boldsymbol{\Sigma}_2, n)$ be two multivariate t -distributions with different precision matrices and degrees of freedom. Then if it holds that $\lambda(\boldsymbol{\Sigma}_2^{-1}) \geq \lambda(\boldsymbol{\Sigma}_1^{-1})$ and $n < m$, then $\text{St}_k(0, \boldsymbol{\Sigma}_1, m) \leq_{\text{SD}} \text{St}_k(0, \boldsymbol{\Sigma}_2, n)$.

PROOF: The result follows straightforward from the fact that

$$\begin{aligned} \text{St}_k(0, \boldsymbol{\Sigma}_1, m) &\underbrace{<_{\text{Disp}}}_{\text{using Theorem 4}} \text{St}_k(0, \boldsymbol{\Sigma}_1, n) \\ &\underbrace{\leq_{\text{SD}}}_{\text{Using Theorem 5}} \text{St}_k(0, \boldsymbol{\Sigma}_2, n). \end{aligned}$$

From the well-known result that a composition of two expansion functions is also an expansion function, it easily holds that $\text{St}_k(0, \boldsymbol{\Sigma}_1, m) \leq_{\text{SD}} \text{St}_k(0, \boldsymbol{\Sigma}_2, n)$. ■

Note that Corollary 3 provides a sufficient condition for the \leq_{SD} order. It is easy to show that under this condition, it also holds that

$$\lambda(\text{Var}(\text{St}_k(0, \boldsymbol{\Sigma}_2, n))) \geq \lambda(\text{Var}(\text{St}_k(0, \boldsymbol{\Sigma}_1, m))).$$

However, this last implication cannot be considered a sufficient condition. The reason is that the variance matrix is defined through both the precision matrix and degrees of freedom.

4. APPLICATION TO THE INFLUENTIAL OBSERVATIONS IN REGRESSION ANALYSIS

4.1. The Model

Johnson and Geisser [11] proposed a method of assessing the influence of specified subsets of the data when the goal is to predict future observations using predictive densities. For this purpose, they considered the following model: $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, where $\boldsymbol{\varepsilon}$ is an $N \times 1$ random vector distributed as $\text{MN}_n(\mathbf{0}, \boldsymbol{\theta}\mathbf{I})$ (N -dimensional multivariate normal (MN)) with mean vector $\mathbf{0}$ and covariance matrix $\boldsymbol{\theta}\mathbf{I}$, $\boldsymbol{\theta}$ scalar, $\boldsymbol{\beta}$ is the $p \times 1$ vector of regression coefficients, \mathbf{X} is an $N \times p$ matrix of fixed “independent” variables, and \mathbf{Y} is the $N \times 1$ vector of responses on the “dependent” variable. Although they noted that a more general model is possible, they assumed the prior density for $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ to be $g(\boldsymbol{\beta}, \boldsymbol{\theta}) \propto \boldsymbol{\theta}^{-1}$, which presumes that little prior information is available relative to that information inherent in the data. Assume the case when a particular subset of size k has been deleted; we denote this by (i) , and the subset itself is indicated by i . Then the general linear model can then be expressed as

$$\mathbf{y}' = (\mathbf{y}'_i, \mathbf{y}'_{(i)}) = \boldsymbol{\beta}'(\mathbf{X}'_i, \mathbf{X}'_{(i)}) + (\boldsymbol{\varepsilon}'_i, \boldsymbol{\varepsilon}'_{(i)}).$$

Thus, the predictive densities based on full and subset deleted datasets, when $\boldsymbol{\theta}$ is unknown, are two multivariate t -distributions with parameters

$$\text{St}_N(\hat{\mathbf{y}}, (s^2(\mathbf{I} + \mathbf{H}))^{-1}, N - p) \quad \text{and} \quad \text{St}_N(\hat{\mathbf{y}}^{(i)}, (s_{(i)}^2(\mathbf{I} + \mathbf{H}^{(i)}))^{-1}, N - k - p),$$

where

$$\begin{aligned} \mathbf{S} &= \mathbf{X}'\mathbf{X}, & \mathbf{H} &= \mathbf{X}\mathbf{S}^{-1}\mathbf{X}', & \mathbf{H}^{(i)} &= \mathbf{X}\mathbf{S}_{(i)}^{-1}\mathbf{X}', & \hat{\mathbf{y}} &= \mathbf{X}\hat{\boldsymbol{\beta}}, \\ \mathbf{r} &= \mathbf{y} - \hat{\mathbf{y}}, & \hat{\mathbf{y}}^{(i)} &= \mathbf{X}\hat{\boldsymbol{\beta}}_{(i)}, & \mathbf{a}^2 &= \mathbf{r}'\mathbf{r}, & s^2 &= \mathbf{a}^2/N - p, \end{aligned}$$

and let $\mathbf{S}_{(i)}$, $\mathbf{a}_{(i)}^2$, and $s_{(i)}^2$ be similarly defined.

4.2. The Problem

In this case, the problem of detecting influential observations is based on comparing two multivariate t -distributions. If we only study the comparison in terms of variability, it seems logical that if we delete a subset of data, then the obtained predictive density will be expected to be more dispersive than the one based on full data. In other words, it would be expected that $f(\cdot) <_{SD} f_{(i)}(\cdot)$. This may be interpreted as the added variability, due to deletion of data subset i . However, it is not true that every subset of data with a fixed size k has the same influence. First, using Corollary 3, we checked that $f(\cdot) <_{SD} f_{(i)}(\cdot)$ for all $\binom{N}{k}$ subsets, $k = 1, 2, 3$. We do not consider it necessary to present these comparisons in this article. After these comparisons, and clearly inspired in Corollary 3, we can define a dispersion Bayesian influence in terms of variability (DBIV) measure to the i subset as

$$Q_i^2 = \|\lambda(s_{(i)}^2(\mathbf{I} + \mathbf{H}^{(i)})) - \lambda(s^2(\mathbf{I} + \mathbf{H}))\|_2^2$$

and we will order the subsets from least to most influential according to the magnitude of Q_i^2 . Note that under the assumptions in Corollary 3, if we have the inequality $\lambda(s_{(i)}^2(\mathbf{I} + \mathbf{H}^{(i)})) \geq \lambda(s^2(\mathbf{I} + \mathbf{H}))$, then it holds that $f(\cdot) <_{SD} f_{(i)}(\cdot)$.

4.3. The Dataset and Conclusions

We consider data from the 1975 Florida Area Cumulus Experiment (FACE) previously discussed in great detail by Cook and Weisberg [4]. This experiment was conducted to determine the merits of using silver iodide to produce rainfall increases and to isolate factors contributing to treatment until additivity. There were 24 observations on 11 variables, including the response rainfall, an indicator variable determining whether clouds were seeded or not seeded, and 8 other variables, including interaction terms that were determined to be related to rainfall.

Initially, we will analyze the full dataset and then delete case 2 (observation 2) and reanalyze in its absence. For an initial analysis, a computer program was written to compute relevant statistics for all $\binom{N}{k}$ subsets, $k = 1, 2, 3$ and $N = 23, 24$ using the Maple 6 package.

A summary of these results for the full dataset, as well as for the full dataset with case 2 deleted, is given in Table 1. It is clear from Table 1 that observations 2, 18, 1, 15, 6, and 3 are most influential in the dispersion sense when $k = 1$. Moreover, case 2 is clearly an outlier and it significantly affects predictive inferences based in this dataset from the dispersion point of view. Also, it is interesting to note that when case 2 is deleted, the most influential observations in dispersion sense are

TABLE 1. Top Six Most Influential Subsets $k = 1, 2, 3$

	Full Dataset		Observation 2 Deleted	
	Observation	Q_i^2	Observation	Q_i^2
$k = 1$	2	52,226.08	18	156.47
	18	289.36	1	71.50
	1	137.74	15	62.43
	15	70.92	5	54.14
	6	67.93	6	37.22
	3	56.50	23	37.17
$k = 2$	(2,9)	139,060.84	(1,18)	964.98
	(2,23)	106,460.70	(1,15)	755.63
	(2,5)	92,471.40	(9,18)	630.90
	(2,19)	88,912.05	(13,18)	480.58
	(2,16)	78,617	(7,18)	421.42
	(2,17)	77,522	(8,18)	368.62
$k = 3$	(2,9,23)	300,931.65	(3,5,11)	3,899.09
	(2,9,16)	284,113.70	(9,13,18)	3,573.22
	(2,9,18)	253,606.55	(1,15,18)	2,565.88
	(2,5,9)	245,137.72	(7,9,18)	2,330.72
	(2,9,19)	242,907.97	(1,16,18)	2,090.76
	(2,3,9)	228,659.05	(1,9,15)	2,016.76

18, 1, 15, 5, 6, and 23. Although the order has been changed, we can observe that the differences among the influence measures are not significant. When $k = 2$ or 3, it is clear from Table 1 that case 2 will be included in most influential subsets. However, we choose to delete case 2 and perform a more careful analysis. Inspection of Table 1 reveals that observations 18 and 1 are not only the most individual influence observations but also the most influential pair. Also, we can observe that case 18 appears jointly with case 1 in the third most influential triple. Also, we note that the pair (1,18) appears in two influential triples. Consequently, the influence for case (1,18) must be taken into account in a dispersion sense.

Surprisingly, the most influential triple is not composed of the three most influential cases, but the third most influential triple coincides with cases that are most influential individually and in pairs.

With this analysis, we present a way to study the influence of specified subsets of the data in a dispersion sense complementary to the analysis of Johnson and Geisser [10]. This is not obviously an alternative method for studying the divergence between $f(\cdot)$ and $f_{(i)}(\cdot)$. The reason is why the study of the dispersion is, of course, location independent and the location component given by the means of the t -distributions is specially important to study the lack of fit of the model, as John-

son and Geisser [11] showed. However, this study provides an interesting point of view and we would like to emphasize that it was not necessary for the approximation of substituting the multivariate t -distribution for a scaled multivariate Normal density.

Acknowledgment

We wish to thank the referees whose careful reading and marking in the first version of this manuscript led to a number of corrections.

References

1. Arellano-Valle, R.B. & Bolfarine, H. (1995). On some characterizations of the t -distribution. *Statistics and Probability Letters* 25: 79–85.
2. Bernardo, J.M. & Smith, A.F.M. (1994). *Bayesian theory*. New York: Wiley.
3. Capéraà, P. (1988). Tail ordering and asymptotic efficiency of rank test. *The Annals of Statistics* 16(1): 470–478.
4. Cook, R.D. & Weisberg, S. (1980). Finding influential cases in regression: A review. *Technometrics* 22: 495–508.
5. Deshpandé, J.V. & Kochar, S.C. (1983). Dispersive ordering is the same as tail ordering. *Advances in Applied Probability* 15: 686–687.
6. Doksum, K. (1969). Starshaped transformations and the power of rank tests. *Annals of Mathematical Statistics* 40: 1167–1176.
7. Fernández-Ponce, J.M. & Suárez-Llorens, A. (2003). A multivariate dispersion ordering based on quantiles more widely separated. *Journal of Multivariate Analysis* 85: 40–53.
8. Frahm, G., Junker, M., & Szimayer, A. (2003). Elliptical copulas: Applicability and limitations. *Statistics and Probability Letters* 63: 275–286.
9. Giovagnoli, A. & Wynn, H.P. (1995). Multivariate dispersion orderings. *Statistics and Probability Letters* 22: 325–332.
10. Hickey, R.J. (1986). Concepts of dispersion in distributions: A comparative note. *Journal of Applied Probability* 23: 924–929.
11. Johnson, W. & Geisser, S. (1983). A predictive view of the detection and characterization of influential observations in regression analysis. *Journal of the American Statistical Association* 78(381): 137–144.
12. Lawrence, M.J. (1975). Inequalities of s -ordered distributions. *The Annals of Statistics* 3: 413–428.
13. Lewis, T. & Thompson, J.W. (1981). Dispersive distributions and the connection between dispersivity and strong unimodality. *Journal of Applied Probability* 18: 76–90.
14. Nelsen, R.B. (1999). *An introduction to copulas*. New York: Springer-Verlag.
15. Rojo, J. & He, G.Z. (1991). New properties and characterizations of the dispersive ordering. *Statistics and Probability Letters* 11: 365–372.
16. Shaked, M. (1982). Dispersive ordering of distributions. *Journal of Applied Probability* 19: 310–320.
17. Shaked, M. & Shanthikumar, J.G. (1994). *Stochastic orders and their applications*. New York: Academic Press.