Journal of Multivariate Analysis **83**, 208–231 (2002)
doi:10.1006/jmva.2001.2043

Distribution-Function-Based Bivariate Quantiles

L.-A. Chen

National Chiao Tung University, Hsinchu, Taiwan

and

A. H. Welsh

Australian National University, Canberra, Australia; and
University of Southampton, Southampton, United Kingdom

Received February 2, 2000; published online January 11, 2002

We introduce bivariate quantiles which are defined through the bivariate distribution function. This approach ensures that, unlike most multivariate medians or the multivariate M-quartiles, the bivariate quantiles satisfy an analogous property to that of the univariate quantiles in that they partition R^2 into sets with a specified probability content. The definition of bivariate quantiles leads naturally to the definition of quantities such as the bivariate median, bivariate extremes, the bivariate quantile curve, and the bivariate trimmed mean. We also develop asymptotic representations for the bivariate quantiles. © 2002 Elsevier Science (USA)

AMS 1991 subject classifications: 62G30; 62H05.

Key words and phrases: bivariate extreme; bivariate median; bivariate quantile; bivariate quantile curve; bivariate trimmed mean.

1. INTRODUCTION

Order statistics or quantiles are the basis for a variety of useful exploratory and robust procedures for univariate data. It is desirable to extend these procedures to multivariate data, but the lack of a natural ordering for multivariate data (Kendall, 1966; Bell and Haller, 1969) has hindered the definition of quantiles and hence the definition of procedures based on them in multivariate problems.

Much of the work in generalizing quantiles to multivariate distributions has concentrated on the particular case of the median or the extremes. Weber (1909) defined the multivariate ℓ_1 median by minimizing the multivariate version of the absolute residuals. More recently, Oja (1983) defined the multivariate simplex median by minimizing the sum of volumes of

0047-259X/02 $35.00

simplices with vertices on the observations, and Liu (1988, 1990) introduced the simplicial depth median maximizing an empirical simplicial depth function. An excellent review of this work is given by Small (1990). Extremes have been defined by Kudo (1957) as the observations with maximum Mahalanobis distance. The componentwise or marginal extreme has been studied by Sibuya (1960) and many other authors. This definition is quite reasonable for some applications but not for outlier detection because it does not in general identify a particular bivariate observation as the extreme from a sample (see Smith *et al.*, 1990). General multivariate quantiles (which of course include the multivariate median and extremes as special cases) are more difficult to define. The approach of taking a minimization problem whose solution is the univariate quantile, generalizing the minimization problem to the multivariate case, and then defining multivariate quantiles to be solutions of this minimization problem has been taken by Breckling and Chambers (1988) and Koltchinski (1997). Maller (1988) considered a fixed family of sets indexed by a univariate parameter (such as spheres) and implicitly defined αth multivariate quantiles to be the boundary of the largest member of the family (in terms of the index parameter) which has probability less than α. A related approach was developed by Einmahl and Mason (1992) who defined the multivariate αth quantile to be the smallest (based on a real-valued function) Borel set that has probability greater than or equal to α.

The αth quantile of a univariate distribution is a point that partitions the real line into two sets such that the probability of the set to the left of the quantile is approximately α and the probability of the set to the right of the quantile is approximately $1-\alpha$. Most of the multivariate medians and the multivariate M-quantiles do not satisfy this kind of probability cumulation condition because their definitions do not involve the cumulative probability distribution. Moreover, as noted by Chaudhuri (1996), most authors try to introduce descriptive statistics that generalize the concept of univariate quantiles to the multivariate setup without discussing what they are trying to estimate. That is, almost no attention is paid to the underlying population quantile. These issues, together with computational simplicity, motivate our definition of bivariate quantiles. Our approach is analogous to that used in the univariate case: We first specify the population quantile in terms of the underlying cumulative distribution and then construct estimators of the population quantiles simply by replacing the cumulative distributions by sample cumulative distributions. This definition leads naturally to the definition of quantities such as the bivariate median, bivariate extremes, the bivariate interquantile area, and the bivariate trimmed mean.

We define two different types of bivariate quantile points in Section 2. We present sample estimators of these bivariate quantile points and establish their large sample properties in Section 3. We introduce bivariate quantile curves in Section 4 and show how they can be used to define bivariate extremes, the bivariate interquantile range, and bivariate trimmed means. We apply the bivariate quantiles in Section 5 and briefly discuss extensions to higher dimensions in Section 6.

2. BIVARIATE QUANTILE POINTS

Our approach to the bivariate case is to define quantiles as points which satisfy natural generalisations of the probability cumulation condition. We begin by considering a natural, fixed direction in R^2 and then consider using the distribution of X to choose a particular direction.

2.1. *North–South Bivariate Quantile Points*

Suppose that we fix the direction for convenience from south to north. Then, each point $(a, b) \in R^2$ partitions R^2 into the sets $A_1 = \{(x_1, x_2)': x_2 \geqslant b\}$, $A_2 = \{(x_1, x_2)': x_1 \leqslant a, x_2 \leqslant b\}$, and $A_3 = \{(x_1, x_2)': x_1 \geqslant a, x_2 \leqslant b\}$. The point (a, b) can be thought of as a bivariate $(P(A_2), P(A_3))$th quantile point. It is convenient to express the formal definition in terms of the usual bivariate distribution function $F(x_1, x_2) = P(X_1 \leqslant x_1, X_2 \leqslant x_2)$ and the marginal distribution function F_2 of X_2. By analogy to the univariate quantile, we introduce the following definition.

DEFINITION 2.1. The (α_1, α_2)*th NS bivariate quantile point* is the vector $\xi(\alpha_1, \alpha_2) = (F_{12}^{-1}(\alpha_1, \alpha_2), F_2^{-1}(\alpha_1 + \alpha_2))'$ which satisfies

$$F_2^{-1}(\alpha_1 + \alpha_2) = \inf\{x_2: F_2(x_2) \geqslant \alpha_1 + \alpha_2\}$$

and

$$F_{12}^{-1}(\alpha_1, \alpha_2) = \inf\{x_1: F(x_1, F_2^{-1}(\alpha_1 + \alpha_2)) \geqslant \alpha_1\},$$

for $\alpha_1, \alpha_2 \geqslant 0$ and $\alpha_1 + \alpha_2 \leqslant 1$. The α*th NS bivariate quantile point* is defined as $\xi(\alpha) = \xi(\frac{1}{2}\alpha, \frac{1}{2}\alpha)$, $0 \leqslant \alpha \leqslant 1$, and we call the $\xi(\frac{1}{2})$ the *NS bivariate median point*.

The marginal quantiles arise as components of NS bivariate quantile points: The second component is the $\alpha = \alpha_1 + \alpha_2$ quantile of X_2 and when $\alpha_1 = 1 - \alpha_2 = \alpha$, the first component is $F_{12}^{-1}(\alpha, 1-\alpha) = F_1^{-1}(\alpha)$, the αth quantile of X_1 (and the second component is $F_2^{-1}(1)$). If X_1 and X_2 are independent, then $F_{12}^{-1}(\alpha_1, \alpha_2) = F_1^{-1}(\alpha_1/(\alpha_2 + \alpha_2))$.

EXAMPLE 1. Consider the random vector with the bivariate continuous uniform distribution on $(0, 1)\times(0, 1)$ which has probability density function

$$f(x_1, x_2) = \begin{cases} 1 & \text{if } \ 0 < x_1 < 1, 0 < x_2 < 1 \\ 0 & \text{otherwise.} \end{cases}$$

The (α_1, α_2)th and αth NS bivariate quantile points are

$$\xi(\alpha_1, \alpha_2) = \left(\frac{\alpha_1}{\alpha_1+\alpha_2}, \alpha_1+\alpha_2\right) \quad \text{and} \quad \xi(\alpha) = \left(\frac{1}{2}, \alpha\right).$$

The NS bivariate median point is $(\frac{1}{2}, \frac{1}{2})$.

2.2. *Bivariate Quantile Points*

One aspect of the definition of NS bivariate quantile points that is unsatisfactory is that the north–south direction, while very natural, is fixed and arbitrary. We therefore develop a definition of bivariate quantile points which allows the distribution of X to specify the appropriate direction. The resulting bivariate quantile has the additional advantage of satisfying an equivariance condition.

Suppose that $X = (X_1, X_2)'$ has location vector μ and positive definite spread matrix Σ. Since Σ is positive definite, there is an orthogonal matrix P such that $\Sigma = P\Lambda P'$, where Λ is the diagonal matrix of eigenvalues $\lambda_1 \leqslant \lambda_2$ of Σ. Let v_1 and v_2 denote the eigenvectors of Σ corresponding to λ_1 and λ_2, respectively. Set $\Sigma^{1/2} = P\Lambda^{1/2}$ so that $\Sigma = \Sigma^{1/2}\Sigma^{1/2'}$. Let the bivariate vectors σ_1' and σ_2' denote the rows of $\Sigma^{-1/2'}$. Then let

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \Sigma^{-1/2'}(X - \mu). \tag{2.1}$$

We denote the joint distribution function of Y_1 and Y_2 by G and the marginal distribution functions of Y_1 and Y_2 by G_1 and G_2, respectively.

DEFINITION 2.2. For $\alpha_1, \alpha_2 \geqslant 0$ and $\alpha_1+\alpha_2 \leqslant 1$, the bivariate vector $\eta(\alpha_1, \alpha_2)$ is an (α_1, α_2)*th bivariate quantile point* if

$$\eta(\alpha_1, \alpha_2) = \mu + \Sigma^{1/2}\xi^*(\alpha_1, \alpha_2),$$

where $\xi^*(\alpha_1, \alpha_2) = (G_{12}^{-1}(\alpha_1, \alpha_2), G_2^{-1}(\alpha_1+\alpha_2))'$ is the (α_1, α_2)th NS bivariate quantile point of $Y = \Sigma^{-1/2'}(X-\mu)$. We also write $\eta(\alpha) = \eta(\frac{1}{2}\alpha, \frac{1}{2}\alpha)$, $0 \leqslant \alpha \leqslant 1$, and call $\eta(\frac{1}{2})$ *the bivariate median point*.

At least for theoretical calculations, it is useful to note that, if Y_1 and Y_2 are independent, then $G_{12}^{-1}(\alpha_1, \alpha_2) = G_1^{-1}(\alpha_1/(\alpha_2+\alpha_2))$. In general,

$(G_{12}^{-1}(\alpha_1, \alpha_2), G_2^{-1}(\alpha_1+\alpha_2))'$ is the NS bivariate quantile point for the reweighted variable Y, and the bivariate quantile point is a back transformation of this NS bivariate quantile point to the scale of X. This means that the bivariate quantile point satisfies a rotated version of the probability cumulation condition.

The following theorem shows that the bivariate quantile points also satisfy a rotational equivariance property.

THEOREM 2.3. *Suppose that* $\Sigma^{1/2}(AX+b) = A\Sigma^{1/2}(X)$ *and* $\mu(AX+b) = A\mu(X)+b$. *Then the bivariate quantile satisfies*

$$\eta(\alpha_1, \alpha_2, AX+b) = A\eta(\alpha_1, \alpha_2, X)+b.$$

Proof. Notice that

$$\Sigma^{-1/2'}(AX+b)(AX+b-\mu(AX+b)) = \Sigma^{-1/2'}(X)(X-\mu(X))$$

so

$$\begin{pmatrix} G_{12}^{-1}(\alpha_1, \alpha_2, AX+b) \\ G_2^{-1}(\alpha_1+\alpha_2, AX+b) \end{pmatrix} = \begin{pmatrix} G_{12}^{-1}(\alpha_1, \alpha_2, X) \\ G_2^{-1}(\alpha_1+\alpha_2, X) \end{pmatrix}.$$

It follows immediately that

$$\begin{aligned} \eta(\alpha_1, \alpha_2, AX+b) &= \Sigma^{1/2}(AX+b)\begin{pmatrix} G_{12}^{-1}(\alpha_1, \alpha_2, AX+b) \\ G_2^{-1}(\alpha_1+\alpha_2, AX+b) \end{pmatrix} + \mu(AX+b) \\ &= A\Sigma^{1/2}(X)\begin{pmatrix} G_{12}^{-1}(\alpha_1, \alpha_2, X) \\ G_2^{-1}(\alpha_1+\alpha_2, X) \end{pmatrix} + A\mu(X)+b \\ &= A\eta(\alpha_1, \alpha_2, X)+b. \quad \blacksquare \end{aligned}$$

EXAMPLE 2. Consider the bivariate normal distribution

$$N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right).$$

Figure 1 shows the NS bivariate quantile points (on the curve from ns1 to ns2) and the bivariate quantile points (on the curve from bq1 to bq2) for $\rho = 0.8$.

Recall that for any bivariate point $x \in R^2$, the inner product $(x-\mu)'(q-\mu)$ is the projection of $x-\mu$ onto the vector $(q-\mu)$. We require the following lemma which describes the vector q^0 that maximizes the variance of projections of fixed length.

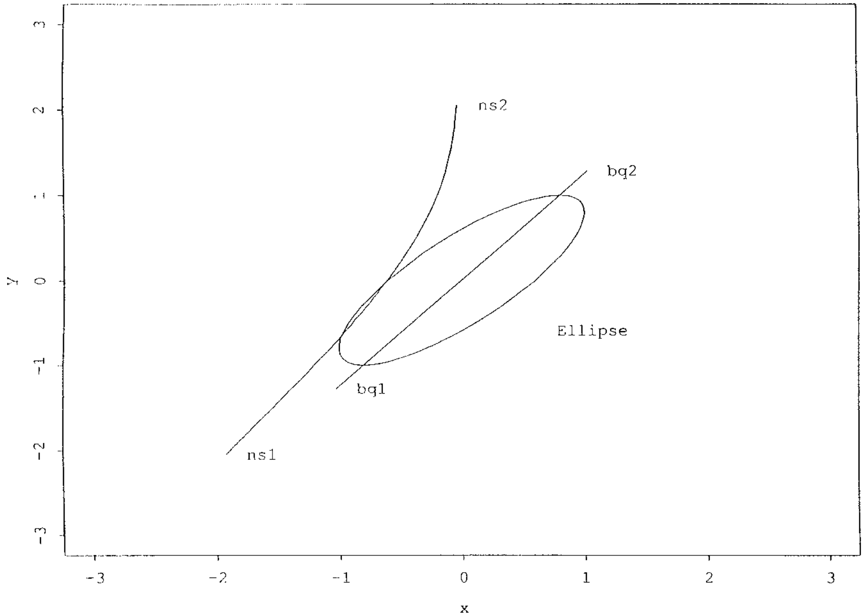

FIG. 1. The NS bivariate quantile points (on the curve from ns1 to ns2) and the bivariate quantile points (on the curve from bq1 to bq2) for the bivariate normal distribution with means zero, variances one, and correlation $\rho = 0.8$.

LEMMA 2.4. *Suppose that X has mean μ and covariance matrix Σ. Then $\mathrm{var}[(X-\mu)'(q-\mu)]$ is maximized among all bivariate vectors q satisfying $\|q-\mu\| = c$ for $c > 0$, by*

$$q^0 = cv_2 + \mu, \tag{2.2}$$

where v_2 is the eigenvector corresponding to the largest eigenvalue, say λ_2, of Σ. We can also write

$$q^0 = \Sigma^{1/2} \begin{pmatrix} 0 \\ \dfrac{c}{\sqrt{\lambda_2}} \end{pmatrix} + \mu. \tag{2.3}$$

Proof. Equation (2.2) follows from principal component analysis. Since $\Sigma^{1/2} = P\,\Lambda^{1/2}$, we have

$$\Sigma^{1/2} \begin{pmatrix} 0 \\ \dfrac{c}{\sqrt{\lambda_2}} \end{pmatrix} + \mu = P\,\Lambda^{1/2} \begin{pmatrix} 0 \\ \dfrac{c}{\sqrt{\lambda_2}} \end{pmatrix} + \mu = P \begin{pmatrix} 0 \\ c \end{pmatrix} + \mu = cv_2 + \mu$$

which implies (2.3). ∎

Lemma 2.4 establishes that if $q^0 = cv_2 + \mu$ satisfies the αth probability cumulation condition, then, $\operatorname{var}[(X-\mu)'(q^0-\mu)] \geqslant \operatorname{var}[(X-\mu)'(q-\mu)]$ for each q satisfying the αth probability cumulation condition and $\|q-\mu\| = \|q_0-\mu\|$.

The following theorem establishes conditions under which the bivariate quantiles lie on the principal component axis.

THEOREM 2.5. *If Y_1 and Y_2 are independent with symmetric distributions, then*

$$q^0 = \Sigma^{1/2}\begin{pmatrix} 0 \\ G_2^{-1}(\alpha) \end{pmatrix} + \mu = \sqrt{\lambda_2}\, G_2^{-1}(\alpha)\, v_2 + \mu. \tag{2.4}$$

That is, $q^0 = \eta(\alpha)$, the αth bivariate quantile.

Proof. Since the distribution of Y_1 is symmetric about zero, the result follows from the fact that $G_2(G_2^{-1}(\alpha)) = \alpha$ and $P(Y_1 \leqslant 0) = 1/2$. ∎

Some members of the elliptical family of distributions such as the bivariate normal distribution satisfy the conditions of Theorem 2.5.

2.3. *Relationships between NS and Bivariate Median Points*

The NS bivariate quantile points and the bivariate quantile points coincide when the random variables X_1 and X_2 are independent but not otherwise. Nonetheless, if the distribution of X is symmetric in the sense that $X-\mu$ and $-(X-\mu)$ have the same distribution, the bivariate median point can be expressed as the average of the NS bivariate median point and the SN bivariate median point. This may be useful for avoiding the potential loss of efficiency from having to estimate μ and Σ in order to estimate the bivariate median point.

THEOREM 2.6. *If X has a continuous and symmetric distribution, then the bivariate median*

$$\eta(\tfrac{1}{2}) = \tfrac{1}{2}\,(\xi(\tfrac{1}{2}) + \xi^*(\tfrac{1}{2})),$$

where $\xi^*(\frac{1}{2}) = (F_{12}^{*-1}(\frac{1}{4}, \frac{1}{4}), F_2^{*-1}(\frac{1}{2}))'$ satisfies

$$F_2^{*-1}(\tfrac{1}{2}) = \sup\{x_2 \colon P(X_2 \geqslant x_2) \geqslant \tfrac{1}{2}\}$$

and

$$F_{12}^{*-1}(\tfrac{1}{4}, \tfrac{1}{4}) = \sup\{x_1 \colon P(X_1 \geqslant x_1, X_2 \geqslant F_2^{*-1}(\tfrac{1}{2})) \geqslant \tfrac{1}{4}\}$$

with $F_{12}^*(x_1, x_2) = P(X_1 \geqslant x_1, X_2 \geqslant x_2)$ and $F_2^*(x_2) = P(X_2 \geqslant x_2)$.

Proof. Write $F_{12}^{*-1}(\alpha)=F_{12}^{*-1}(\alpha/2,\alpha/2)$. Clearly μ is the bivariate median. We see from the continuity and symmetry of the distribution that $F_2^{*-1}(\frac{1}{2})=\sup\{x_2: P(X_2\geqslant x_2)\geqslant\frac{1}{2}\}=\inf\{x_2: P(X_2\leqslant x_2)\geqslant\frac{1}{2}\}=\mu_2$. Again, by symmetry, $F_{12}^{-1}(\frac{1}{2})$ satisfies $P(X_1\leqslant F_{12}^{-1}(\frac{1}{2}), X_2\leqslant\mu_2)=\frac{1}{4}$ and $F_{12}^{*-1}(\frac{1}{2})$ satisfies $P(X_1\geqslant F_{12}^{*-1}(\frac{1}{2}), X_2\geqslant\mu_2)=\frac{1}{4}$. It follows from the symmetry property that $\xi(\frac{1}{2})$ and $\xi^*(\frac{1}{2})$ are equidistant from μ_1 and this implies that $\frac{1}{2}(F_{12}^{-1}(\frac{1}{2})+F_{12}^{*-1}(\frac{1}{2}))=\mu_1$. ∎

3. SAMPLE BIVARIATE QUANTILE POINTS

Let $X_i=(X_{1i}, X_{2i})'$ be a random sample from the distribution with distribution function F and marginal distribution functions F_1 and F_2. Let the density functions of F and F_2 be f and f_2, respectively. We assume the set of assumptions listed in the Appendix throughout the rest of this paper.

3.1. *Sample NS Bivariate Quantile Points*

The empirical marginal distribution function of X_2 and the empirical left joint distribution function of X_1 and X_2 are $\hat{F}_2(x_2)=n^{-1}\sum_{i=1}^n I(X_{2i}\leqslant x_2)$ and $\hat{F}(x_1, x_2)=n^{-1}\sum_{i=1}^n I(X_{1i}\leqslant x_1, X_{2i}\leqslant x_2)$, respectively.

DEFINITION 3.1. The *sample* (α_1,α_2)*th NS bivariate quantile point* $\hat{\xi}(\alpha_1,\alpha_2)$, the *sample* α*th NS bivariate quantile point* $\hat{\xi}(\alpha)$, and the *sample NS bivariate median point* $\hat{\xi}(\frac{1}{2})$ are defined as in Definition 2.1 with F_2 and F replaced by $\hat{F}_2$ and $\hat{F}$, respectively.

The sample (α_1,α_2) NS bivariate quantile has breakdown point $\min\{\alpha_1+\alpha_2, 1-(\alpha_1+\alpha_2)\}$. This implies that the breakdown point of the sample NS bivariate median is 0.5. For comparison, the breakdown points for Weber's (1909) ℓ_1 median is 0.5, for Oja's simplex median is 0, and for the half space median is 1/3 (see Small (1990)).

To obtain the large sample properties of $\hat{\xi}(\alpha_1,\alpha_2)$, let $\delta_1(\alpha_1,\alpha_2)=\int_{-\infty}^{F_2^{-1}(\alpha_1+\alpha_2)} f(F_{12}^{-1}(\alpha_1,\alpha_2), x_2)\,dx_2$ and $\delta_2(\alpha_1,\alpha_2)=\int_{-\infty}^{F_{12}^{-1}(\alpha_1,\alpha_2)} f_{x_1\mid x_2}(x_1\mid F_2^{-1}(\alpha_1+\alpha_2))\,dx_1$, where $f_{x_1\mid x_2}$ is the conditional probability density function of X_1 given $X_2=x_2$ and $F_{12}^{-1}(\alpha_1,\alpha_2)$ satisfies $F(F_{12}^{-1}(\alpha_1,\alpha_2), F_2^{-1}(\alpha_1+\alpha_2))=\alpha_1$. Our main result is the following theorem.

THEOREM 3.2. *The components of the sample* (α_1,α_2)*th NS bivariate quantile* $\hat{\xi}(\alpha_1,\alpha_2)$ *satisfy*

$$
\begin{aligned}
n^{1/2}(\hat{F}_{12}^{-1}&(\alpha_1, \alpha_2) - F_{12}^{-1}(\alpha_1, \alpha_2)) \\
&= \delta_1(\alpha_1, \alpha_2)^{-1} \Bigg[\left\{ \frac{\alpha_1}{\alpha_1+\alpha_2} - \delta_2(\alpha_1, \alpha_2) \right\} \\
&\quad \times n^{-1/2} \sum_{i=1}^{n} \{\alpha_1+\alpha_2 - I(X_{2i} \leqslant F_2^{-1}(\alpha_1+\alpha_2))\} \\
&\quad + n^{-1/2} \sum_{i=1}^{n} \left\{ \left(\frac{\alpha_1}{\alpha_1+\alpha_2} - I(X_{1i} \leqslant F_{12}^{-1}(\alpha_1, \alpha_2)) \right) \right. \\
&\quad \left. \times I(X_{2i} \leqslant F_2^{-1}(\alpha_1+\alpha_2)) \right\} \Bigg] + o_p(1),
\end{aligned}
$$

and

$$
\begin{aligned}
n^{1/2}(\hat{F}_2^{-1}&(\alpha_1+\alpha_2) - F_2^{-1}(\alpha_1+\alpha_2)) \\
&= f_2^{-1}(F_2^{-1}(\alpha_1+\alpha_2))\, n^{-1/2} \sum_{i=1}^{n} \{\alpha_1+\alpha_2 - I(X_{2i} \leqslant F_2^{-1}(\alpha_1+\alpha_2))\} + o_p(1).
\end{aligned}
$$

The proof is given in the Appendix.

COROLLARY 3.3. *The asymptotic distribution of the centered and normalized* (α_1, α_2)*th NS bivariate quantile* $n^{1/2}(\hat{\xi}(\alpha_1, \alpha_2) - \xi(\alpha_1, \alpha_2)) = n^{1/2}(\hat{F}_{12}^{-1}(\alpha_1, \alpha_2) - F_{12}^{-1}(\alpha_1, \alpha_2), \hat{F}_2^{-1}(\alpha_1+\alpha_2) - F_2^{-1}(\alpha_1+\alpha_2))'$ *is bivariate normal with mean vector zero and covariance matrix* $\tilde{\Sigma} = (\sigma_{ij}, i, j = 1, 2)$*, where*

$$
\sigma_{11} = \delta_1(\alpha_1, \alpha_2)^{-2} \left[\left(\frac{\alpha_1}{\alpha_1+\alpha_2} - \delta_2(\alpha_1, \alpha_2) \right)^2 (\alpha_1+\alpha_2)(1-(\alpha_1+\alpha_2)) + \frac{\alpha_1\alpha_2}{\alpha_1+\alpha_2} \right]
$$

$$
\begin{aligned}
\sigma_{12} &= \delta_1(\alpha_1, \alpha_2)^{-1} f_2(F_2^{-1}(\alpha_1+\alpha_2))^{-1} \left(\frac{\alpha_1}{\alpha_1+\alpha_2} - \delta_2(\alpha_1, \alpha_2) \right) \\
&\quad \times (\alpha_1+\alpha_2)(1-(\alpha_1+\alpha_2))
\end{aligned}
$$

and

$$
\sigma_{22} = f_2(F_2^{-1}(\alpha_1+\alpha_2))^{-2} (\alpha_1+\alpha_2)(1-(\alpha_1+\alpha_2)).
$$

3.2. *Sample Bivariate Quantile Points*

Put $Y_i = (Y_{1i}, Y_{2i})' = S^{-1/2'}(X_i - \hat{\mu})$, where S and $\hat{\mu}$ represent estimators of Σ and μ, respectively. The corresponding empirical distribution functions of $Y = (Y_1, Y_2)'$ and Y_2 are $\hat{G}(y_1, y_2) = \frac{1}{n} \sum_{i=1}^{n} I(Y_{1i} \leqslant y_1, Y_{2i} \leqslant y_2)$ and $\hat{G}_2(y_2) = n^{-1} \sum_{i=1}^{n} I(Y_{2i} \leqslant y_2)$, respectively.

DEFINITION 3.4. The *sample* (α_1, α_2)*th bivariate quantile point* and the *sample bivariate median point* $\hat{\eta}(\frac{1}{2})$ are defined as in Definition 2.2 with μ, Σ, G_2, and G replaced by $\hat{\mu}$, S, $\hat{G}_2$, and $\hat{G}$, respectively.

Arguments similar to those used to prove Theorem 2.4 show that the sample bivariate quantile point is equivariant provided the estimators S and $\hat{\mu}$ are equivariant.

We assume that g and g_2 are continuous, positive, and finite and that g_2, g_2', g, $\partial g/\partial y_1$, and $\partial g/\partial y_2$ are bounded functions.

THEOREM 3.5. *Let* $(Y_{1i}^*, Y_{2i}^*)' = \Sigma^{-1/2'}(X_i - \mu)$. *Then the sample bivariate quantile point satisfies*

$$n^{1/2}(\hat{\eta}(\alpha_1, \alpha_2) - \eta(\alpha_1, \alpha_2)) = n^{1/2}(S^{1/2} - \Sigma^{1/2}) \begin{pmatrix} G_{12}^{-1}(\alpha_1, \alpha_2) \\ G_2^{-1}(\alpha_1 + \alpha_2) \end{pmatrix}$$

$$+ \Sigma^{1/2} n^{1/2} \left(\begin{pmatrix} \hat{G}_{12}^{-1}(\alpha_1, \alpha_2) \\ \hat{G}_2^{-1}(\alpha_1 + \alpha_2) \end{pmatrix} - \begin{pmatrix} G_{12}^{-1}(\alpha_1, \alpha_2) \\ G_2^{-1}(\alpha_1 + \alpha_2) \end{pmatrix} \right) + n^{1/2}(\hat{\mu} - \mu) + o_p(1),$$

where

$$n^{1/2}(\hat{G}_2^{-1}(\alpha_1 + \alpha_2) - G_2^{-1}(\alpha_1 + \alpha_2))$$

$$= g_2(G_2^{-1}(\alpha_1 + \alpha_2))^{-1} \, n^{-1/2} \sum_{i=1}^{n} \{\alpha_1 + \alpha_2 - I(Y_{2i}^* \leqslant G_2^{-1}(\alpha_1 + \alpha_2))\}$$

$$- \sigma_2' n^{1/2}(\hat{\mu} - \mu) + \tilde{E}_2' n^{1/2}(s_2 - \sigma_2) + o_p(1)$$

and

$$n^{1/2}(\hat{G}_{12}^{-1}(\alpha_1, \alpha_2) - G_{12}^{-1}(\alpha_1, \alpha_2))$$

$$= (p_{21}\, g_1(G_{12}^{-1}(\alpha_1, \alpha_2)))^{-1} \Bigg[n^{-1/2} \sum_{i=1}^{n} \left\{ \frac{\alpha_1}{\alpha_1 + \alpha_2} - I(Y_{1i}^* \leqslant G_{12}^{-1}(\alpha_1, \alpha_2)) \right\}$$

$$\times I(Y_{2i}^* \leqslant G_2^{-1}(\alpha_1 + \alpha_2)) + \left[\frac{\alpha_1}{\alpha_1 + \alpha_2} - p_{12} \right] n^{-1/2}$$

$$\times \sum_{i=1}^{n} \{\alpha_1 + \alpha_2 - I(Y_{2i}^* \leqslant G_2^{-1}(\alpha_1 + \alpha_2))\}$$

$$- g_1(G_{12}^{-1}(\alpha_1, \alpha_2)) \, p_{21} \sigma_1' n^{1/2}(\hat{\mu} - \mu) + g_1(G_{12}^{-1}(\alpha_1, \alpha_2)) \, \tilde{E}_{21}' n^{1/2}(s_1 - \sigma_1)$$

$$+ g_2(G_2^{-1}(\alpha_1 + \alpha_2))(\tilde{E}_{12}' - p_{12}\tilde{E}_2') \, n^{1/2}(s_2 - \sigma_2) \Bigg] + o_p(1)$$

with s_1' and s_2' the rows of $S^{-1/2'}$,

$$\tilde{E}_2 = E(X-\mu \mid Y_2^* = G_2^{-1}(\alpha_1+\alpha_2)),$$
$$p_{21} = P(Y_2^* \leqslant G_2^{-1}(\alpha_1+\alpha_2) \mid Y_1^* = G_{12}^{-1}(\alpha_1, \alpha_2)),$$
$$p_{12} = P(Y_1^* \leqslant G_{12}^{-1}(\alpha_1, \alpha_2) \mid Y_2^* = G_2^{-1}(\alpha_1+\alpha_2)),$$
$$\tilde{E}_{12} = E[(X-\mu)\, I(Y_1^* \leqslant G_{12}^{-1}(\alpha_1, \alpha_2) \mid Y_2^* = G_2^{-1}(\alpha_1+\alpha_2)],$$
$$\tilde{E}_{21} = E[(X-\mu)\, I(Y_2^* \leqslant G_2^{-1}(\alpha_1+\alpha_2) \mid Y_1^* = G_{12}^{-1}(\alpha_1, \alpha_2)].$$

3.3. *An Estimator of the Bivariate Median Point under Symmetry*

We showed in Section 2.3 that $\eta(\frac{1}{2}) = \frac{1}{2}(\xi(\frac{1}{2})+\xi^*(\frac{1}{2}))$ under symmetry. We now explore the properties of the estimator of $\eta(\frac{1}{2})$ constructed from $\frac{1}{2}(\xi(\frac{1}{2})+\xi^*(\frac{1}{2}))$.

Let $\hat{F}_2^*(x_2) = n^{-1}\sum_{i=1}^n I(X_{2i} \geqslant x_2)$ and $\hat{F}_{12}^*(x_1, x_2) = n^{-1}\sum_{i=1}^n I(X_{1i} \geqslant x_1, X_{2i} \geqslant x_2)$.

THEOREM 3.6. *Let $\hat{\xi}_m = \frac{1}{2}(\hat{\xi}^*(\frac{1}{2})+\hat{\xi}(\frac{1}{2}))$ where $\hat{\xi}^*(\frac{1}{2})$ is defined in Theorem 2.6 with F_2^* and F_{12}^* replaced by $\hat{F}_2^*$ and $\hat{F}_{12}^*$, respectively. Then, if the distribution of X is symmetric, $n^{1/2}(\hat{\xi}_m-\mu)$ has the same asymptotic distribution as that of $n^{1/2}(\hat{F}_{12}^{-1}(1/2)-F_{12}^{-1}(1/2), \hat{F}_2^{-1}(1/2)-F_2^{-1}(1/2))'$ in Corollary 3.3 with $\alpha_1 = \alpha_2 = \frac{1}{4}$.*

The proof of Theorem 3.6 is analogous to that of Theorem 3.5 and so is omitted.

Thus when the bivariate distribution is symmetric, we can obtain a consistent estimator of the bivariate median without having to estimate μ and Σ.

If we use $\hat{\eta}(\frac{1}{2})$ to estimate $\eta(\frac{1}{2})$, we need to estimate at least Σ and perhaps also μ, and the estimates of these quantities affect the efficiency of $\hat{\eta}(\frac{1}{2})$. On

TABLE I

Efficiencies of the Sample Mean $\bar{X}$, Sample Bivariate Median $\hat{\eta}$, and $\hat{\xi}_m$ for $\rho = 0.2$

Estimate	$\sigma=1$	$\sigma=3$	$\sigma=5$	$\sigma=10$	$\sigma=15$
$\delta=0.1$					
$\bar{X}$	1.000	0.968	0.538	0.171	0.083
$\hat{\eta}$	0.657	0.973	0.962	0.876	0.832
$\hat{\xi}_m$	0.679	1.000	1.000	1.000	1.000
$\delta=0.2$					
$\bar{X}$	1.000	0.807	0.389	0.110	0.048
$\hat{\eta}$	0.675	0.971	0.950	0.946	0.929
$\hat{\xi}_m$	0.671	1.000	1.000	1.000	1.000

TABLE II

Efficiencies of the Sample Mean $\bar{X}$, Sample Bivariate Median $\hat{\eta}$, and $\hat{\xi}_m$ for $\rho = 0.8$

Estimate	$\sigma = 1$	$\sigma = 3$	$\sigma = 5$	$\sigma = 10$	$\sigma = 15$
$\delta = 0.1$					
$\bar{X}$	1.000	0.872	0.512	0.151	0.074
$\hat{\eta}$	0.698	0.933	0.955	0.842	0.853
$\hat{\xi}_m$	0.725	1.000	1.000	1.000	1.000
$\delta = 0.2$					
$\bar{X}$	1.000	0.707	0.334	0.097	0.043
$\hat{\eta}_y$	0.636	1.000	0.985	0.972	0.969
$\hat{\eta}_m$	0.679	0.977	1.000	1.000	1.000

the other hand, $\hat{\xi}_m$ does not require estimates of either Σ or μ so it may be more efficient than $\hat{\eta}(\frac{1}{2})$. To explore this possibility, we computed the asymptotic variances of the three estimators $\bar{X}$, $\hat{\eta}(\frac{1}{2})$ (using the sample mean and variance to estimate μ and Σ, respectively), and $\hat{\xi}_m$ under the bivariate mixture distribution

$$(1-\delta)\, N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}\right) + \delta N\left(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}, \begin{pmatrix}\sigma^2 & 0\\ 0 & \sigma^2\end{pmatrix}\right)$$

and compared their efficiencies. We present the results in Tables I and II in terms of the ratio of the minimum asymptotic variance of the three estimators to the asymptotic variance of each estimator so that the efficiency is always less than one. That is, the most efficient estimator has efficiency equal to one.

Not surprisingly, for small σ, the sample mean $\bar{X}$ is the most efficient estimator. As σ increases, $\hat{\xi}_m$ becomes the most efficient estimator. While $\hat{\xi}_m$ is mostly more efficient than $\hat{\eta}(\frac{1}{2})$, the improvement in efficiency through using $\hat{\xi}_m$ is quite small.

4. QUANTILE CURVES AND OTHER DERIVED QUANTITIES

The analogue of the real interval $[F^{-1}(\alpha_1), F^{-1}(\alpha_2)]$ in two dimensions is a set $J(\alpha)$ whose boundaries can be called bivariate quantile curves. By analogy to the univariate case, it is most useful to think of $J(\alpha)$ as a set bounded by two quantile curves. Thus, while we have thought of bivariate quantiles as points in R^2, for many purposes, it is more natural to think of a bivariate quantile as a curve in R^2.

Once we have defined an appropriate set $J(\alpha)$ or equivalently appropriate quantile curves, we can then define the extremes to be the extreme quantile curves (the boundaries of the extreme set $J(0, 0, \frac{1}{2}, \frac{1}{2})$), we can generalize the interquantile range to the interquantile area which is the area $A(\alpha) = \int_{J(\alpha)} dx$ of $J(\alpha)$, and we can define the trimmed mean to be the mean over the set $J(\alpha)$, namely $\mu(\alpha) = (\int_{J(\alpha)} dF(x))^{-1} \int_{J(\alpha)} x\, dF(x)$. These derived quantities are equivariant or not according to whether the quantile curves are equivariant or not, so it is of particular interest to construct equivariant quantile curves.

One simple analogy to the univariate quantile interval is the bivariate quantile parallelogram.

DEFINITION 4.1. The $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)'$th *bivariate quantile parallelogram* is

$$P(\alpha) = \left\{ \Sigma^{1/2} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \mu : G_{12}^{-1}(\alpha_1, \alpha_2) \leqslant y_1 \leqslant G_{12}^{-1}(\alpha_3, \alpha_4), \right.$$

$$\left. G_2^{-1}(\alpha_1 + \alpha_2) \leqslant y_2 \leqslant G_2^{-1}(\alpha_3 + \alpha_4) \right\}$$

for $\alpha_1 \leqslant \alpha_3$ and $\alpha_2 \leqslant \alpha_4$.

Just as we can think of the quantile interval as the intersection of the two intervals $[F^{-1}(\alpha_1), \infty)$ and $(-\infty, F^{-1}(\alpha_2)]$, with the finite boundaries defining the quantiles, we can think of the quantile parallelogram as the intersection of the two sets with finite boundaries

$$C(\alpha_1, \alpha_2) = \text{boundary} \left\{ \Sigma^{1/2} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \mu : y_1 \leqslant G_{12}^{-1}(\alpha_1, \alpha_2),\ y_2 \leqslant G_2^{-1}(\alpha_1 + \alpha_2) \right\}$$

if $\alpha_1, \alpha_2 \geqslant \frac{1}{2}$ and

$$C(\alpha_1, \alpha_2) = \text{boundary} \left\{ \Sigma^{1/2} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \mu : y_1 \geqslant G_{12}^{-1}(\alpha_1, \alpha_2),\ y_2 \geqslant G_2^{-1}(\alpha_1 + \alpha_2) \right\}$$

if $\alpha_1, \alpha_2 \leqslant \frac{1}{2}$. (In most practical applications, it will be sensible to have both arguments equal and hence on the same side of $\frac{1}{2}$.) The curves $C(\alpha_1, \alpha_2)$ are potential quantile curves which we call the quantile parallelogram curves.

DEFINITION 4.2. The αth *quantile parallelogram curve* is given by

$$C(\alpha) = \text{boundary} \left\{ \Sigma^{1/2} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \mu : y_1 \leqslant G_{12}^{-1}\left(1 - \alpha, \frac{1}{2}\alpha\right),\ y_2 \leqslant G_2^{-1}\left(1 - \frac{1}{2}\alpha\right) \right\}$$

if $\alpha \geqslant \frac{1}{2}$ and

$$C(\alpha) = \text{boundary}\left\{\Sigma^{1/2}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \mu : y_1 \geqslant G_{12}^{-1}\left(\frac{1}{2}\alpha,\, 1-\alpha\right),\, y_2 \geqslant G_2^{-1}\left(\frac{1}{2}\alpha\right)\right\}$$

if $\alpha \leqslant \frac{1}{2}$.

Example 2 (continued). Consider again the bivariate normal distribution of Example 2 with $\rho = 0.8$. Quantile parallelograms for this distribution are shown in Fig. 2 with $\alpha = 0.025$, 0.05, 0.1, and 0.2 (denoted by p1, p2, p3, and p4, respectively).

A different approach is to consider defining a bivariate quantile point for each possible rotation of the coordinate system and then rotate the resulting curve back into the original coordinate system. Thus, if we let

$$\begin{pmatrix} Z_1(\theta) \\ Z_2(\theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix},$$

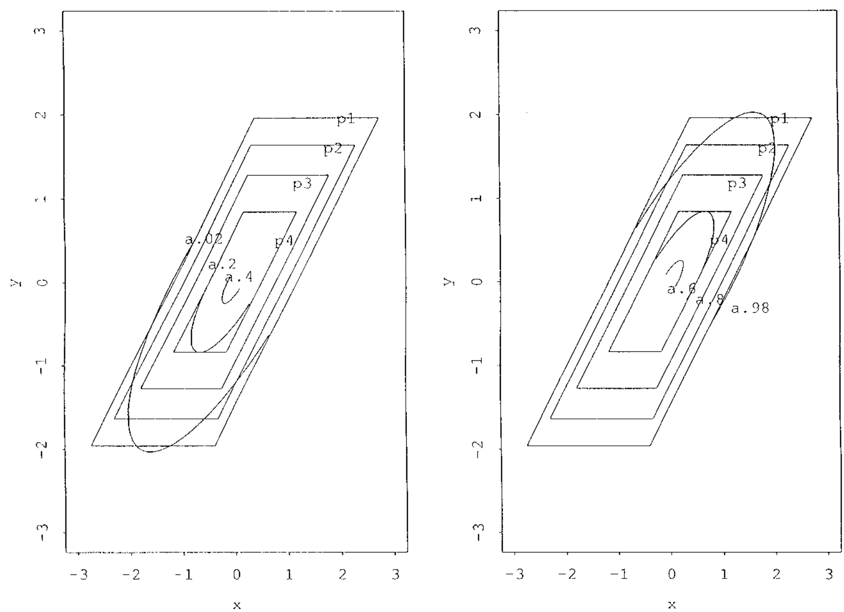

FIG. 2. The quantile curves for $\alpha = 0.02$, 0.2, and 0.4 (denoted qc.02, qc.2, and qc.4 in (a)) and the quantile curves for $\alpha = 0.6$, 0.8, and 0.98 (denoted qc.6, qc.8 and qc.98 in (b)). (a) Quantile parallelograms and quantile curves ($0 \leqslant \theta \leqslant \pi$). (b) Quantile parallelograms and quantile curves ($\pi \leqslant \theta < 2\pi$).

we can define the (α_1, α_2)th NS bivariate quantile point $\xi_\theta(\alpha_1, \alpha_2)$ for $(Z_1(\theta), Z_2(\theta))'$. The (α_1, α_2)th quantile curve is then

$$\Sigma^{1/2}\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \xi_\theta(\alpha_1, \alpha_2) + \mu \tag{4.1}$$

viewed as a function of θ for fixed α_1 and α_2. (We can define the (α_1, α_2)th NS bivariate quantile curve by replacing Y by X and omitting the final renormalization by $\Sigma^{1/2}$ and μ.)

If we consider all possible rotations (all possible values of θ), the quantile curves are closed. (For the bivariate normal distribution, the curves are ellipses in R^2.) These quantiles reduce to a point in the upper tail and one in the lower tail for the univariate case. This is not really what we think of as a univariate quantile. An approach which reduces in the univariate case to the univariate quantile is to consider only half the set of possible rotations corresponding to the intersection of the closed curve with the half-plane $x_2 > x_1$ if $\alpha_1 + \alpha_2 > \frac{1}{2}$ and with the other half-plane otherwise. In this case, the quantile curve is defined in the half-plane on either side of the line $x_2 = -x_1$. To partition the space R^2, we can extend the curve linearly along the boundary line $x_2 = -x_1$.

DEFINITION 4.3. The *bivariate quantile curve* is the intersection of the curve defined by (4.1) with the set $\{(x_1, x_2)': x_2 \geqslant -x_1\}$ if $\alpha > \frac{1}{2}$ and $\{(x_1, x_2)': x_2 \leqslant -x_1\}$ if $\alpha < \frac{1}{2}$.

EXAMPLE 2 (continued). Consider again the bivariate normal distribution of Example 2 with $\rho = 0.8$. Figure 2a shows the quantile curves for $\alpha = 0.02$, 0.2, and 0.4 (denoted qc.02, qc.2, and qc.4) and Fig. 2b shows the quantile curves for $\alpha = 0.6$, 0.8, and 0.98 (denoted qc.6, qc.8, and qc.98).

We can apply the quantile curve approach using the marginal quantile instead of a bivariate quantile point. Of course, the points on the curve then no longer satisfy bivariate probability cumulation conditions. We could also consider using the boundaries of the sets defined by Einmahl and Mason (1992) to define quantile curves but our approach is computationally simpler.

Estimators of the quantile curves and the derived quantities are easily constructed simply by replacing the unknown population quantities by their empirical analogues.

5. EXAMPLES

In this section, we examine the quantiles of two real data sets. In the first example we display the two proposed bivariate quantiles and illustrate the partitions of the data implied by them. In the second example, we illustrate the use of NS bivariate quantile points as a basis for statistical inference.

Reaven and Miller (1979) measured several variables to compare normal patients and diabetics. Among the variables, the three variables of major interest were X_1, glucose intolerance; X_2, insulin response to oral glucose; and X_3, insulin resistance. For our bivariate quantile analysis, we consider the variables X_1 and X_2. Figures 3a and 3b show, respectively, the NS bivariate quantile points $\hat{\xi}(\alpha)$, for $\alpha = 0.1, 0.2, \ldots, 0.9$ (labeled $\times$ and denoted ns1, ..., ns9), and the bivariate quantile points $\hat{\eta}(\alpha)$ (labeled $\triangle$ and denoted bq1, ..., bq9). We have also included the observations (labeled $\cdot$) and the partitions of the data implied by the NS bivariate quantile point and the bivariate quantile point at $\alpha = 0.7$ (denoted ns.7 and bq.7).

The sales price of rural land depends on many variables, including the closeness of a parcel to transportation facilities. Maddala (1988) gives a sample of size 67 of data on sales prices (per acre) of rural land near Sarasota, Florida, and some other variables. For a bivariate variables

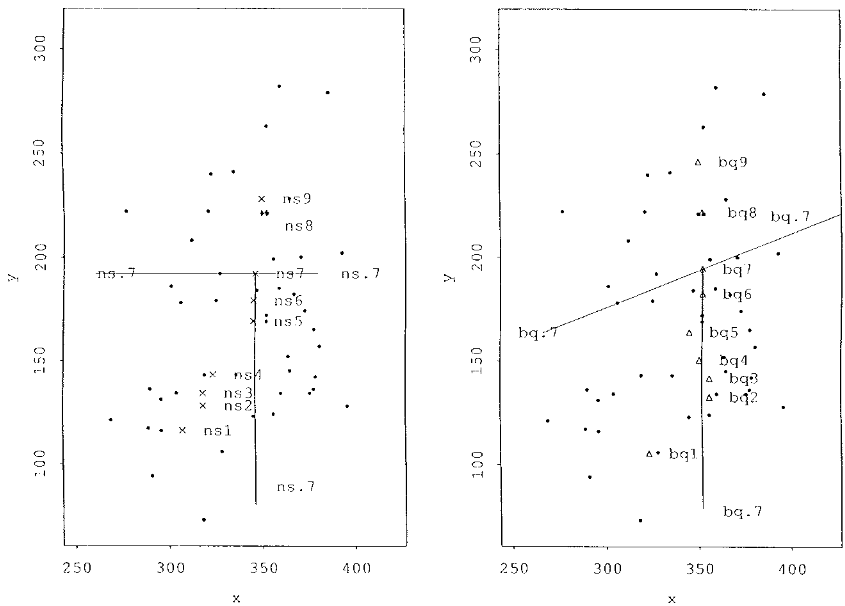

FIG. 3. The NS bivariate quantile points $\hat{\xi}(\alpha)$, for $\alpha = 0.1, 0.2, \ldots, 0.9$ (labeled $\times$ and denoted ns1, ..., ns9 in (a)) and the bivariate quantile points $\hat{\eta}(\alpha)$ (labeled $\triangle$ and denoted bq1, ..., bq9 in (b)). We also show the observations (labeled $\cdot$) and the partitions of the data implied by the NS bivariate quantile point and the bivariate quantile point at $\alpha = 0.7$ (denoted ns.7 and bq.7, respectively). (a) Sample NS bivariate quantile points for normal patients. (b) Sample bivariate quantile points for normal patients.

TABLE III

Sample NS Bivariate Quantiles

α	$\hat{\xi}(\alpha)$	α	$\hat{\xi}(\alpha)$	α	$\hat{\xi}(\alpha)$
0.1	$\begin{pmatrix} 3000 \\ 2 \end{pmatrix}$	0.4	$\begin{pmatrix} 3977 \\ 3.2 \end{pmatrix}$	0.7	$\begin{pmatrix} 5172 \\ 5.4 \end{pmatrix}$
0.2	$\begin{pmatrix} 4234 \\ 2.5 \end{pmatrix}$	0.5	$\begin{pmatrix} 4821 \\ 4 \end{pmatrix}$	0.8	$\begin{pmatrix} 5000 \\ 6.4 \end{pmatrix}$
0.3	$\begin{pmatrix} 4764 \\ 2.9 \end{pmatrix}$	0.6	$\begin{pmatrix} 5172 \\ 4.9 \end{pmatrix}$	0.9	$\begin{pmatrix} 4835.5 \\ 12.4 \end{pmatrix}$

analysis, we consider the variables sale price (X_1) and distance from the parcel to the I-75 freeway (X_2). To see if these two variables are related, we could fit a simple linear regression model with X_1 as the dependent variable and X_2 as the independent variable and then test the significance of the slope parameter, not to be greater than zero. We can explore the hypothesis informally by examining the NS bivariate quantiles

$$\hat{\xi}(\alpha) = \begin{pmatrix} \hat{F}_{12}^{-1}(\alpha/2, \alpha/2) \\ \hat{F}_2^{-1}(\alpha) \end{pmatrix}.$$

Since $\hat{F}_{12}^{-1}(\alpha/2, \alpha/2)$ represents the median of the observations $(x_{1i}, x_{2i})'$ subject to $X_{2i} \leqslant \hat{F}_2^{-1}(\alpha)$, there will be a positive relationship between X_1 and X_2 if $\hat{F}_{12}^{-1}(\alpha/2, \alpha/2)$ decreases as α increases. We display the NS bivariate quantiles for $\alpha = 0.1, 0.2, ..., 0.9$ in Table III. Clearly, $\hat{F}_{12}^{-1}(\alpha/2, \alpha/2)$ does not decrease in α, providing evidence against the hypothesis. Note that this exploration does not require the assumption of a linear relationship between X_1 and X_2.

6. HIGHER DIMENSIONS

The bivariate quantile can be extended to higher dimensional observations. Suppose that the random vector $X = (X_1, X_2, ..., X_p)'$ has a location vector μ and positive spread matrix Σ. Again, let the p-vectors $\sigma'_1, \sigma'_2, ..., \sigma'_p$ denote the rows of $\Sigma^{-1/2'}$. We denote the distribution function of the random variables $X_j, ..., X_p$ by $F_{j...p}(x_j, ..., x_p)$. When $j = 1$ we also write $F = F_{1...p}$. The $(\alpha_1, \alpha_2, ..., \alpha_p)$*th NS multivariate quantile point* is the vector $\xi(\alpha_1, \alpha_2, ..., \alpha_p) = (F_{12..p}^{-1}(\alpha_1, \alpha_2, ..., \alpha_p), F_{2..p}^{-1}(\alpha_1 + \alpha_2, \alpha_3, ..., \alpha_p), ..., F_{p-1p}^{-1}(\alpha_1 + \cdots + \alpha_{p-1}, \alpha_p), F_p^{-1}(\alpha_1 + \alpha_2 + \cdots + \alpha_p))'$ which satisfies

$$F_p^{-1}(\alpha_1+\alpha_2+\cdots+\alpha_p)=\inf\{x_p: F_p(x_p)\geqslant \alpha_1+\alpha_2+\cdots+\alpha_p\},$$

$$F_{p-1p}^{-1}(\alpha_1+\cdots+\alpha_{p-1},\alpha_p)=\inf\{x_{p-1}: F_{p-1p}(x_{p-1},F_p^{-1}(\alpha_1+\alpha_2+\cdots+\alpha_p))\geqslant \alpha_1+\cdots+\alpha_{p-1}\},$$

$$\vdots$$

$$F_{23..p}^{-1}(\alpha_1+\alpha_2,\alpha_3,...,\alpha_p)=\inf\{x_2: F_{23..p}(x_2,F_{3..p}^{-1}(\alpha_1+\alpha_2+\alpha_3,\alpha_4,...,\alpha_p),...,F_p^{-1}(\alpha_1+\alpha_2+\cdots+\alpha_p))\geqslant \alpha_1+\alpha_2\},$$

and

$$F_{12..p}^{-1}(\alpha_1,\alpha_2,...,\alpha_p)=\inf\{x_1: F(x_1,F_{2...p}^{-1}(\alpha_1+\alpha_2,\alpha_3,...,\alpha_p),...,F_p^{-1}(\alpha_1+\alpha_2+\cdots+\alpha_p))\geqslant \alpha_1\},$$

for $\alpha_1,\alpha_2,...,\alpha_p\geqslant 0$ and $\alpha_1+\alpha_2+\cdots+\alpha_p\leqslant 1$.

For $\alpha_1,\alpha_2,...,\alpha_p\geqslant 0$ and $\alpha_1+\alpha_2+\cdots+\alpha_p\leqslant 1$, the vector $\eta(\alpha_1,\alpha_2,...,\alpha_p)$ is an $(\alpha_1,\alpha_2,...,\alpha_p)$*th multivariate quantile point* if

$$\eta(\alpha_1,\alpha_2,...,\alpha_p)=\mu+\Sigma^{1/2}\xi^*(\alpha_1,\alpha_2,...,\alpha_p),$$

where $\xi^*(\alpha_1,\alpha_2,...,\alpha_p)=(G_{12..p}^{-1}(\alpha_1,\alpha_2,...,\alpha_p),G_{23..p}^{-1}(\alpha_1+\alpha_2,\alpha_3,...,\alpha_p),...,G_{p-1p}^{-1}(\alpha_1+\cdots+\alpha_{p-1},\alpha_p),G_p^{-1}(\alpha_1+\alpha_2+\cdots+\alpha_p))'$ is the $(\alpha_1,\alpha_2,...,\alpha_p)$th *NS multivariate quantile point* of $Y=\Sigma^{-1/2'}(X-\mu)$.

APPENDIX

Let α_1 and $\alpha_2\geqslant 0$ and $\alpha_1+\alpha_2\leqslant 1$. The following conditions are assumed to be true for random vector X, $\hat{\mu}$, and S:

(a1) The probability density function f_2 and the conditional probability density function $f_{x_1|x_2}$, with $x_2=F_2^{-1}(\alpha_1+\alpha_2)$, and their derivatives are both bounded and bounded away from 0 in neighborhoods of $F_2^{-1}(\alpha)$ and $F_{12}^{-1}(\alpha_1,\alpha_2)$, respectively.

(a2) There exists $t>0$ such that the probability density function of $(X-\mu)'(\sigma_j+\delta)$ is uniformly bounded in a neighborhood of H, with $H=G_2^{-1}(\alpha)$ if $j=2$ and $H=G_{12}^{-1}(\alpha_1,\alpha_2)$ if $j=1$, for $\|\delta\|\leqslant t$ and the probability density function of $(X-\mu)'(\sigma_j+\delta)(X-\mu)'u((X-\mu)'\sigma_j)$ is uniformly bounded away from zero for $\|u\|=1$ and $\|\delta\|\leqslant t$, for $j=1,2$.

(a3) $E(((X-\mu)'\sigma_j)^2\|(X-\mu)\|)<\infty$ for $j=1,2$.

(a4) $n^{1/2}(\hat{\mu}-\mu)=O_p(1)$ and $n^{1/2}(S-\Sigma)=O_p(1)$.

Proof of Theorem 3.2. The representation of $\hat{F}_2^{-1}(\alpha_1, \alpha_2)$ can be seen in Ruppert and Carroll (1980). The first component of the sample NS bivariate quantile point $\hat{F}_{12}^{-1}(\alpha_1, \alpha_2)$ can be formulated as a solution of the problem

$$\min_a \sum_{i=1}^n (X_{1i} - a)\left(\frac{\alpha_1}{\alpha_1 + \alpha_2} - I(X_{1i} \leqslant a)\right) I(X_{2i} \leqslant \hat{F}_2^{-1}(\alpha_1, \alpha_2)).$$

Let

$$S(t_1, t_2) = n^{-1/2} \sum_{i=1}^n \left\{\frac{\alpha_1}{\alpha_1 + \alpha_2} - I(X_{1i} \leqslant F_{12}^{-1}(\alpha_1, \alpha_2) + n^{-1/2} t_1)\right\} \times I(X_{2i} \leqslant F_2^{-1}(\alpha_1 + \alpha_2) + n^{-1/2} t_2).$$

Then we need to show that

$$\sup_{\|t_1\| \leqslant k, \|t_2\| \leqslant k'} \left| S(t_1, t_2) - S(0, 0) + \left[\frac{\alpha_1}{\alpha_1 + \alpha_2} f_2(F_2^{-1}(\alpha_1 + \alpha_2))\, t_2 - \delta_1(\alpha_1, \alpha_2)\, t_1 - f_2(F_2^{-1}(\alpha_1 + \alpha_2))\, \delta_2(\alpha_1, \alpha_2)\, t_2\right]\right| = o_p(1). \quad (7.1)$$

Now, (7.1) is bounded above by the two terms

$$\frac{\alpha_1}{\alpha_1 + \alpha_2} \sup_{\|t_2\| \leqslant k} \left| n^{-1/2} \sum_{i=1}^n \{I(X_{2i} \leqslant F_2^{-1}(\alpha_1 + \alpha_2) + n^{-1/2} t_2) - I(X_{2i} \leqslant F_2^{-1}(\alpha_1 + \alpha_2))\} - f_2(F_2^{-1}(\alpha_1 + \alpha_2))\, t_2 \right|,$$

which is $o_p(1)$ by the properties of univariate quantiles, and

$$\sup_{\|t_1\| \leqslant k, \|t_2\| \leqslant k'} |\tilde{S}(t_1, t_2) - \tilde{S}(0, 0) - \{\delta_1(\alpha_1, \alpha_2)\, t_1 + f_2(F_2^{-1}(\alpha_1 + \alpha_2))\, \delta_2(\alpha_1, \alpha_2)\, t_2\}|, \quad (7.2)$$

where

$$\tilde{S}(t_1, t_2) = n^{-1/2} \sum_{i=1}^n I(X_{1i} \leqslant F_{12}^{-1}(\alpha_1, \alpha_2) + n^{-1/2} t_1, X_{2i} \leqslant F_2^{-1}(\alpha_1 + \alpha_2) + n^{-1/2} t_2).$$

The result (7.1) will follow if we can show that the term in (7.2) is $o_p(1)$.

We have

$$E\,|I\{X_{1i} \leqslant F_{12}^{-1}(\alpha_1, \alpha_2) + n^{-1/2} t_1^1,\, X_{2i} \leqslant F_2^{-1}(\alpha_1 + \alpha_2) + n^{-1/2} t_2^1\}$$
$$- I\{X_{1i} \leqslant F_{12}^{-1}(\alpha_1, \alpha_2) + n^{-1/2} t_1^2,\, X_{2i} \leqslant F_2^{-1}(\alpha_1 + \alpha_2) + n^{-1/2} t_2^2\}|$$
$$\leqslant E\,|I\{X_{1i} \leqslant F_{12}^{-1}(\alpha_1, \alpha_2) + n^{-1/2} t_1^1\} - I\{X_{1i} \leqslant F_{12}^{-1}(\alpha_1, \alpha_2) + n^{-1/2} t_1^2\}|$$
$$+ E\,|I\{X_{2i} \leqslant F_2^{-1}(\alpha_1 + \alpha_2) + n^{-1/2} t_2^1\} - I\{X_{2i} \leqslant F_2^{-1}(\alpha_1 + \alpha_2) + n^{-1/2} t_2^2\}|$$
$$\leqslant M(\|t_1^2 - t_1^1\| + \|t_2^2 - t_2^1\|)$$

and, similarly,

$$E \sup_{\|t_1 - t_1^1\| \leqslant k,\, \|t_2 - t_2^1\| \leqslant k'} |I\{X_{1i} \leqslant F_{12}^{-1}(\alpha_1, \alpha_2) + n^{-1/2} t_1,\, X_{2i} \leqslant F_2^{-1}(\alpha_1 + \alpha_2) + n^{-1/2} t_2\}$$
$$- I\{X_{1i} \leqslant F_{12}^{-1}(\alpha_1, \alpha_2) + n^{-1/2} t_1^1,\, X_{2i} \leqslant F_2^{-1}(\alpha_1 + \alpha_2) + n^{-1/2} t_2^1\}|$$
$$\leqslant C\{P(|X_{1i} - F_{12}^{-1}(\alpha_1, \alpha_2)| \leqslant k + k') + P(|X_{2i} - F_2^{-1}(\alpha_1 + \alpha_2)| \leqslant k + k')\}$$
$$\leqslant C'(k + k').$$

We can apply Lemma 3.2 of Bai and He (1998) to show that (7.2) is $o_p(1)$ provided

$$\sup_{\|t_1\| \leqslant k,\, \|t_2\| \leqslant k'} \|E(\tilde{S}(t_1, t_2) - \tilde{S}(0, 0)) - \{\delta_1(\alpha_1, \alpha_2)\, t_1 + f_2(F_2^{-1}(\alpha_1 + \alpha_2))\, \delta_2(\alpha_1, \alpha_2)\, t_2\}\| = o_p(1). \tag{7.3}$$

Using the techniques of Jurečková (1984), we can establish both (7.3) and also that $n^{1/2}(\hat{F}_{12}(\alpha_1, \alpha_2) - F_{12}^{-1}(\alpha_1, \alpha_2)) = O_p(1)$. Using the fact that $n^{1/2}(\hat{F}_2^{-1}(\alpha_1 + \alpha_2) - F_2^{-1}(\alpha_1 + \alpha_2)) = O_p(1)$, the theorem then follows.

Proof of Theorem 4.2. The second component of the sample bivariate quantile point $\hat{G}_2^{-1}(\alpha)$ can be formulated as a solution of the minimization problem

$$\min_a \sum_{i=1}^{n} (Y_{2i} - a)(\alpha - I(Y_{2i} \leqslant a)).$$

By letting

$$\Phi_1(t_1, t_2, t_3) = n^{-1/2} \sum_{i=1}^{n} [\alpha - I\{(\sigma_2 + n^{-1/2} t_3)'\,(X_i - \mu) \leqslant G_2^{-1}(\alpha) + n^{-1/2}((\sigma_2 + n^{-1/2} t_3)'\, t_1 + n^{-1/2} t_2\}],$$

the representation for $\hat{G}_2^{-1}(\alpha)$ follows from:

(a) $\Phi_1(T_1, T_2, T_3) - \Phi_1(0,0,0) - g_2(G_2^{-1}(\alpha))[\sigma_2' T_1 + T_2 - T_3' \tilde{E}_{c_2}(\alpha)] = o_p(1)$ for any sequences T_j with $T_j = O_p(1)$.

(b) $n^{1/2}(\hat{G}_2^{-1}(\alpha) - G_2^{-1}(\alpha)) = O_p(1)$, and

(c) $n^{-1/2} \sum_{i=1}^n (\alpha - I(Y_{2i} \leqslant \hat{G}_2^{-1}(\alpha)) = \Phi_1(n^{1/2}(\hat{\mu} - \mu), n^{1/2}(\hat{G}_2^{-1}(\alpha) - G_2^{-1}(\alpha)), n^{1/2}(s_2 - \sigma_2)) = o_p(1)$.

The proof of the above statements can be derived by similar arguments to those used in the proof in Chen *et al.* (1999) and is therefore omitted. The proof for $\hat{G}_{12}^{-1}(\alpha_1, \alpha_2)$ which follows also establishes a more general case.

The first component of the sample bivariate quantile, $\hat{G}_{12}^{-1}(\alpha_1, \alpha_2)$, is a solution of the minimization problem

$$\min_a \sum_{i=1}^n (Y_{1i} - a)\left(\frac{\alpha_1}{\alpha_1 + \alpha_2} - I(Y_{1i} \leqslant a)\right) I(Y_{2i} \leqslant \hat{G}_2^{-1}(\alpha_1 + \alpha_2)).$$

Let

$$\begin{aligned}\tilde{\Phi}_1(t_j) = n^{-1/2} \sum_{i=1}^n \Bigg[&\frac{\alpha_1}{\alpha_1 + \alpha_2} - I\{(\sigma_1 + n^{-1/2} t_4)' (X_i - \mu) \leqslant G_{12}^{-1}(\alpha_1, \alpha_2) \\ &+ n^{-1/2}((\sigma_1 + n^{-1/2} t_4)' t_1 + t_3)) \Bigg] I((\sigma_2 + n^{-1/2} t_5)' (X_i - \mu) \\ &\leqslant G_2^{-1}(\alpha_1 + \alpha_2) + n^{-1/2}((\sigma_2 + n^{-1/2} t_5)' t_1 + t_2))\end{aligned}$$

and $\tilde{\Phi}(t_j) = \tilde{\Phi}_1(t_j) - \tilde{\Phi}(0)$. Put

$$\begin{aligned}\tilde{M}(t_j) = &\frac{\alpha_1}{\alpha_1 + \alpha_2} g_2(G_2^{-1}(\alpha_1 + \alpha_2))[-t_5' \tilde{E}_2 + \sigma_2' t_1 + t_2] \\ &- g_1(G_{12}^{-1}(\alpha_1, \alpha_2))[p_{21}(\sigma_1' t_1 + t_3) - t_4' \tilde{E}_{21}] \\ &- g_2(G_2^{-1}(\alpha_1 + \alpha_2))[-t_5' \tilde{E}_{12} + p_{12}(\sigma_2' t_1 + t_2)].\end{aligned}$$

Then, we want to show that

$$\max_{t_j \geqslant k_j} |\tilde{\Phi}(t_j) - \tilde{M}(t_j)| = o_p(1). \tag{7.4}$$

Let us denote

$$\begin{aligned}B(t_j) = n^{-1/2} \sum_{i=1}^n \, [&I\{(\sigma_1 + n^{-1/2} t_4)' (X_i - \mu) \\ &\leqslant G_{12}^{-1}(\alpha_1, \alpha_2) + n^{-1/2}((\sigma_1 + n^{-1/2} t_4)' t_1 + t_3), (\sigma_2 + n^{-1/2} t_5)' (X_i - \mu) \\ &\leqslant G_2^{-1}(\alpha_1 + \alpha_2) + n^{-1/2}((\sigma_2 + n^{-1/2} t_5)' t_1 + t_2)\} \\ &- I\{\sigma_1'(X_i - \mu) \leqslant G_{12}^{-1}(\alpha_1, \alpha_2), \sigma_2'(X_i - \mu) \leqslant G_2^{-1}(\alpha_1 + \alpha_2)\}].\end{aligned}$$

From the result in the first part of this proof, (7.4) is shown from the fact that

$$\sup_{\|t_j\| \leqslant k_j} |B(t_j) - h(t_j)| = o_p(1), \tag{7.5}$$

where

$$h(t_j) = g_1(G_{12}^{-1}(\alpha_1, \alpha_2))[-t_4'\tilde{E}_{21} + p_{21}(\sigma_1' t_1 + t_3)] \\ + g_2(G_2^{-1}(\alpha_1 + \alpha_2))[-t_5'\tilde{E}_{12} + p_{12}(\sigma_2' t_1 + t_2)],$$

which is implied from the following (see Chen *et al.* (2001) for analogous proofs),

$$n^{-1} \sum_{i=1}^{n} E(\tilde{h}_i(t_j^1) - \tilde{h}_i(t_j^2))^2 \leqslant n^{-1/2} M \sum_{j=1}^{5} \|t_j^2 - t_j^1\|, \tag{7.6}$$

with writing $B(t_j) = n^{-1/2} \sum_{i=1}^{n} \tilde{h}_i(t_j)$ and, fixing t_j^0,

$$E \sup_{\|t_j^1 - t_j^0\| \leqslant k} |\tilde{h}_i(t_j^0 - t_j^1)| \leqslant n^{-1/2} Mk \qquad \text{for some} \quad M > 0. \tag{7.7}$$

From (7.6), (7.7), and Lemma 3.2 of Bai and He (1999), we have

$$\sup_{\|t_j^0\| \leqslant k} |\tilde{\Phi}(t_j^0) - E\tilde{\Phi}(t_j^0)| = o_p(1).$$

To complete the proof, we still need to show that

$$\sup_{\|t_j^0\| \leqslant k} |E\tilde{\Phi}(t_j^0) - \tilde{M}(t_j^0)| = o(1). \tag{7.8}$$

Write

$$\begin{aligned} E\tilde{\Phi}(t_j^0) = & \frac{\alpha_1}{\alpha_1 + \alpha_2} n^{1/2} EI((\sigma_2 + n^{-1/2} t_5^0)' (X - \mu) \leqslant G_2^{-1}(\alpha_1 + \alpha_2) \\ & + n^{-1/2}((\sigma_2 + n^{-1/2} t_5^0)' t_1^0 + t_2^0)) - n^{1/2} EI((\sigma_1 + n^{-1/2} t_4^0)' (X - \mu) \\ & \leqslant G_{12}^{-1}(\alpha_1, \alpha_2) + n^{-1/2}((\sigma_1 + n^{-1/2} t_4^0)' t_1^0 + t_3^0)), (\sigma_2 + n^{-1/2} t_5^0)' (X - \mu) \\ & \leqslant G_2^{-1}(\alpha_1 + \alpha_2) + n^{-1/2}((\sigma_2 + n^{-1/2} t_5^0)' t_1^0 + t_2^0)) \\ = & H_1 + H_2. \end{aligned}$$

We denote variables $Y_1^* = Y_1 - G_{12}^{-1}(\alpha_1, \alpha_2)$ and $Y_2^* = Y_2 - G_2^{-1}(\alpha_1 + \alpha_2)$. Let $Z_1 = t_4^{0'}(X - \mu)$, $Z_2 = t_5^{0'}(X - \mu)$, $\delta_1 = n^{-1/2}((\sigma_1 + n^{-1/2} t_4^0)' t_1^0 + t_3^0)$, and $\delta_2 = n^{-1/2}((\sigma_2 + n^{-1/2} t_5^0)' t_1^0 + t_2^0)$. Then we can see that the two terms in the following

$$|H_2+g_1(G_{12}^{-1}(\alpha_1+\alpha_2))[-t_4^{0'}\tilde{E}_{21}+p_{21}(\sigma_1' t_1^0+t_3^0)]$$
$$+g_2(G_2^{-1}(\alpha_1,\alpha_2))[-t_5^{0'}\tilde{E}_{12}+p_{12}(\sigma_2' t_1^0+t_2^0)]|$$
$$\left|H_1-\frac{\alpha_1}{\alpha_1+\alpha_2}\,g_2(G_2^{-1}(\alpha_1+\alpha_2))[-t_5^{0'}\tilde{E}_2+\sigma_2' t_1^0+t_2^0]\right|\leqslant Mn^{-1/2}$$

are all bounded by $Mn^{-1/2}$ which establishes (7.8).

Finally, as in Ruppert and Carroll (1980),

$$n^{-1/2}\sum_{i=1}^{n}\left\{\frac{\alpha_1}{\alpha_1+\alpha_2}-I(Y_{1i}\leqslant \hat{G}_{12}^{-1}(\alpha_1,\alpha_2))\right\}I\{Y_{2i}\leqslant \hat{G}_2^{-1}(\alpha_1+\alpha_2)\}=o_p(1)$$

and, by an analogous argument to that given in Jurečková (1977, Lemma 5.2), for $\varepsilon>0$ there exists $k>0$, $\ell>0$, and N such that

$$P\left(\inf_{\|t_3\|\geqslant k} n^{-1/2}\left|\sum_{i=1}^{n}\left(\frac{\alpha_1}{\alpha_1+\alpha_2}-I(Y_{1i}\leqslant G_{12}^{-1}(\alpha_1+\alpha_2)+t_3)\right)\right.\right.$$
$$\left.\left. I(Y_{2i}\leqslant \hat{G}_2^{-1}(\alpha_1+\alpha_2))\right|<\ell\right)<\varepsilon$$

for $n\geqslant N$. These two results together establish that

$$n^{1/2}(\hat{G}_{12}^{-1}(\alpha_1,\alpha_2)-G_{12}^{-1}(\alpha_1,\alpha_2))=O_p(1)$$

and the representation for $\hat{G}_{12}^{-1}(\alpha_1+\alpha_2)$ follows.

REFERENCES

D. F. Andrews and A. M. Herzberg, "Dta," Springer-Verlag, New York, 1985.

Z.-D. Bai and X. He, Asymptotic distributions of the maximal depth estimators for regression and multivariate location, *Ann. Statist.* **27** (1999), 1616–1637.

V. Barnett, The ordering of multivariate data (with discussion), *J. Roy. Statist. Soc. A* **139** (1976), 318–354.

V. Barnett and T. Lewis, "Outliers in Statistical Data," Wiley, New York, 1994.

J. Breckling and R. Chambers, M-quantiles, *Biometrika* **75** (1988), 761–771.

P. Chaudhuri, On a geometric notion of quantiles for multivariate data, *J. Amer. Statist. Assoc.* **91** (1996), 862–872.

L.-A. Chen, A. H. Welsh, and W. Chan, Linear winsorized means for the linear regression model, *Statist. Sinica* **11** (2001), 147–172.

J. H. J. Einhmahl and D. Mason, Generalized quantile process, *Ann. Statist.* **20** (1992), 1062–1078.

J. Jureckova, Asymptotic relation of M-estimates and R-estimates in linear regression model, *Ann. Math. Statist.* **5** (1977), 464–472.

J. Jureckova, Regression quantiles and trimmed least squares estimator under general design, *Kybernetika* **20** (1984), 345–357.

V. I. Koltchinskii, M-estimation, convexity and quantiles, *Ann. Statist.* **25** (1997), 435–477.

A. Kudo, The extreme value in a multivariate normal sample, *Mem. Fac. Sci. Kyushu Univ. (A)* (1957), 143–156.

R. Y. Liu, On a notion of simplicial depth, *Proc. Nat. Acad. Sci. USA* **18** (1988), 1732–1734.

R. Y. Liu, On a notion of data depth based on random simplices, *Ann. Statist.* **18** (1990), 405–414.

G. S. Maddala, "Introduction to Econometrics," Macmillan, New York, 1988.

R. A. Maller, Normality of trimmed means in higher dimensions, *Ann. Probab.* **16** (1988), 1608–1622.

H. Oja, Descriptive statistics for multivariate distributions, *Statist. Probab. Lett.* **1** (1983), 327–332.

G. M. Reaven and R. G. Miller, An attempt to define the nature of chemical diabets using a multidimensional analysis, *Diabetologia* **16** (1979), 17–24.

D. Ruppert and R. J. Carroll, Trimmed least squares estimation in the linear model, *J. Amer. Statist. Assoc.* **75** (1980), 828–838.

M. Sibuya, Bivariate extreme statistics, *Ann. Inst. Statist. Math., Tokyo* **11** (1960), 195–210.

C. G. Small, A survey of multidimensional medians, *Internat. Statist. Rev.* **58** (1990), 273–277.

R. L. Smith, J. A. Tawn, and H. K. Yuen, Statistics of multivariate extremes, *Internat. Statist. Rev.* **58** (1990), 47–58.

A. Weber, "Über den Standort der Industrien," Tubingen; "Alfred Weber's Theory of Location of Industries," University of Chicago Press, Chicago, 1909.