Bounds are derived on the rate of convergence of the joint distribution of order statistics to the corresponding multivariate normal distribution.

Let \( X_{i_1} \leq X_{i_2} \leq \ldots \leq X_{i_p} \) be the ordered observations corresponding to a sample from a population with distribution density \( g(x) \). The sample quantiles \( X_{h_1}, X_{h_2}, \ldots, X_{h_n} \), where \( n = \lfloor nh \rfloor + 1 \) and \( 0 < h_1 < h_2 < \ldots < h_n < h \), are known to be asymptotically normal (see, e.g., [1]). The aim of this article is to derive optimal bounds on the rate of convergence of the joint distribution of the order statistics to the multivariate normal law.

Let \( 0 = k_0 < k_1 < \ldots < k_p < k_{p+1} = nh \) be integer numbers (note that \( h_i - h_{i-1} \) are not necessarily nonzero). Define \( d_i(\ell=i,\ldots,p) \) as follows: 
\[
\lambda_i = \int_{-\infty}^{d_i} g(t) dt,
\]
where for simplicity, in order to avoid ambiguity in the definition of \( d_i \), we assume that the density \( g(x) \) is nonzero on its interval \((-\infty < x < \infty)\). We also assume that \( g(x) \) is differentiable on \([a,b]\). Let \( M = \sup_{x \geq 0} |g'(x)| \) and \( g_i = g(d_i) \).

We introduce the following notation:
\[
\Phi_{\theta, \Sigma}(Y) - \text{the distribution function (d.f.) of a normal vector with mean } \theta \text{ and variance-covariance matrix } \Sigma;
\]
\[
\Theta_{\theta, \Sigma}(A) - \text{the probability that the corresponding normal vector belongs to the set } A;
\]
\[
\Lambda = p \times p \text{ matrix with the elements } \Lambda_{ii} = \frac{\lambda_i}{g_i}; \quad \Sigma_i = \sqrt{h_i} \left( X_{h_i} - d_i \right) \quad (i = 1, \ldots, p).
\]

We have the following theorem.

**Theorem 1.**

\[
\sup_{y \in [a,b]} \left| P \left[ z_1 < y_1, z_2 < y_2, \ldots, z_p < y_p \right] - \Phi_{\Theta, \Sigma}(Y) \right| \leq C(p) \left[ \frac{1}{\sqrt{h_i} k_i} + \frac{1}{\sqrt{h_i} k_i} + \ldots + \frac{1}{\sqrt{h_i} k_i} + \frac{1}{\sqrt{h_i} k_i} + \frac{1}{\sqrt{h_i} k_i} + \ldots + \frac{1}{\sqrt{h_i} k_i} \right].
\]

(1)

We first prove the theorem for uniform order statistics, i.e., for the case

\[
g(x) = \begin{cases} 
1, & \text{if } 0 \leq x < 1 \\
0, & \text{if } x \in [0,1].
\end{cases}
\]
The proof will be based on the following lemmas.

**LEMMA 1.** The distribution of the vector \((X_{i_1}, \ldots, X_{i_n})\) of uniform order statistics coincides with the distribution of the vector 

\[
\left( \frac{s_1}{s_1 + s_2 + \cdots + s_n}, \frac{s_1 + s_2}{s_1 + s_2 + \cdots + s_n}, \ldots, \frac{s_1 + s_2 + \cdots + s_n}{s_1 + s_2 + \cdots + s_n} \right),
\]

where \(s_1, s_2, \ldots\) are independent random variables with the same d.f. \(P(s < x) = \max \{0, 1 - e^{-x}\}\).

The proof of Lemma 1 is given in [2].

**LEMMA 2.** Let \(\hat{s}_1, \hat{s}_2, \ldots\) be independent identically distributed random variables with \(E\hat{s}_i = 0, \sigma^2\hat{s}_i = 1, \beta_i = E|\hat{s}_i|^p < \infty (i = 1, 2, \ldots)\). Let

\[
S_k = \sqrt{k} \sum_{j=k+1}^{k+i} \hat{s}_j, \quad (i = 1, 2, \ldots, p+1), \quad S = (s_1, \ldots, s_{p+1})^T.
\]

Then

\[
\sup_{\mathcal{U}} |P\{S \in \mathcal{U} \} - \Phi_{p,v}(d)| \leq C\left(\frac{1}{\sqrt{k}} + \frac{1}{\sqrt{k-1}} + \frac{1}{\sqrt{k-2}} + \cdots + \frac{1}{\sqrt{k-p-1}}\right),
\]

where \(\mathcal{U}\) is the collection of convex sets in \(\mathbb{R}^{p+1}\), the constant \(c\) depends only on \(\beta_i\) and \(p\), and

\[
V = \begin{pmatrix}
V_1 & 0 & \cdots & 0 \\
0 & V_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & V_{p+1}
\end{pmatrix}
\]

with the elements

\[
V_i = -\frac{k_i - k_{i-1}}{k_{i+1}} \quad (i = 1, \ldots, p+1).
\]

The proposition of Lemma 2 follows directly from Theorem 2 in [3].

**LEMMA 3.** Let \(\Lambda\) be a positive-definite, and \(C\Lambda\) a nonnegative-definite matrix and let for some \(\varepsilon > 0\) and all \(X \in \mathbb{R}^p\)

\[
X^\top \Lambda X - X^\top C X \leq \varepsilon (X^\top \Lambda X).
\]

Then

\[
0 \leq X^\top \Lambda^{-1} X - X^\top C^{-1} X \leq \varepsilon (X^\top \Lambda^{-1} X) \quad \text{(for any } X \in \mathbb{R}^p)\]

and

\[
(t + \varepsilon)^p |\Lambda| \geq |C| \cdot |\Lambda|.
\]

The lemma follows directly from Lemmas 5 and 6 in [3].

**LEMMA 4.** Let \(\nu_1, \nu_2, \ldots, \nu_p\) be the eigenvalues of the symmetrical matrix \(\Lambda\). Then

\[
\nu_p = \inf_X \frac{\Lambda X}{X^\top X} \quad \text{and} \quad \nu_1 = \sup_X \frac{\Lambda X}{X^\top X}.
\]

2276