Bernoulli **1**(3), 1995, 245–255

Joint distribution of maxima of concomitants of subsets of order statistics

S.N. JOSHI[1] and H.N. NAGARAJA[2]

[1]*Indian Statistical Institute, Bangalore Centre, 8th Mile Mysore Road, Post RVCE, Bangalore 560059, India*
[2]*Department of Statistics, Ohio State University, 141 Cockins Hall, 1958 Neil Avenue, Columbus OH 43210-1247, USA*

Let $(X_{i:n}, Y_{[i:n]})$, $1 \leq i \leq n$, denote the n pairs obtained by ordering a random sample of size n from an absolutely continuous bivariate population on the basis of X sample values. Here $Y_{[i:n]}$ is called the *concomitant* of the ith order statistic. For $1 \leq k \leq n$, let $V_1 = \max\{Y_{[n-k+1:n]}, \ldots, Y_{[n:n]}\}$, and $V_2 = \max\{Y_{[1:n]}, \ldots, Y_{[n-k:n]}\}$. In this paper, we discuss the finite-sample and asymptotic joint distribution of (V_1, V_2). The asymptotic results are obtained when $k = [np]$, $0<p<1$, and when k is held fixed, as $n \to \infty$. We apply our results to the bivariate normal population and indicate how they can be used to determine k such that V_1 is close to $Y_{n:n}$, the maximum of the values of Y in the sample.

Keywords: bivariate normal distribution; concomitants of order statistics; convergence in distribution; extreme values maximum

1. Introduction

Suppose we have a random sample of size n from an absolutely continuous bivariate population (X, Y). For $1 \leq i \leq n$, let $X_{i:n}$ and $Y_{i:n}$ denote the ith order statistics of the X and Y sample values, respectively. The Y value associated with $X_{i:n}$, denoted by $Y_{[i:n]}$, is called the *concomitant* of the ith order statistic or an induced order statistic. In view of their applications in selection procedures, functions of $Y_{[i:n]}$ have been extensively studied. For a recent review, see David (1991). Early work has been surveyed by Bhattacharya (1984).

Here we explore the joint distribution of V_1 and V_2, where $V_1 = \max\{Y_{[n-k+1:n]}, \ldots, Y_{[n:n]}\}$ and $V_2 = \max\{Y_{[1:n]}, \ldots, Y_{[n-k:n]}\}$. Recently, Feinberg and Huber (1994) have investigated some properties of V_1 in a study of cut-off rules under imperfect information. Assuming the sample is drawn from a bivariate normal distribution, they compared the value of $\mathrm{E}(V_1)$ with $\mathrm{E}(Y_{n:n})$ for selected values of n. This information was used to determine k, the number selected, that optimizes some cost function of interest. Motivated by these applications, Nagaraja and David (1994) have studied the finite-sample and the asymptotic properties of V_1. Our work extends their results to two dimensions by considering the joint cumulative distribution function (cdf) of V_1 and V_2.

Since $Y_{n:n} = \max\{V_1, V_2\}$, using our results, one may determine the smallest k such that $P(W_k \leq 1 - \epsilon) \leq \delta$ for a given n and prespecified small $\epsilon > 0$ and $\delta > 0$, where $W_k = (V_1/Y_{n:n})$. Since for $w < 1$, $P(W_k \leq w) = P((V_1/V_2) \leq w, V_2 > 0)$, the joint cdf of V_1 and V_2 is crucial for determining the cdf of W_k. When $k = [np]$, $0 < p < 1$, and the population is bivariate normal, Nagaraja and David (1994) noted that the limit distribution of V_1 is free of ρ, the correlation between X and Y, whenever $\rho > 0$. Hence, it seems more appropriate to determine k using W_k than by comparing the means of V_1 and $Y_{n:n}$. Thus, our results are applicable to the selection problem considered by Feinberg and Huber (1994).

In Section 2, we obtain an expression for the joint cdf of V_1 and V_2. Then, in Section 3, we investigate the limiting joint distribution of (V_1, V_2) in the quantile case where $k = [np]$, $0 < p < 1$, as $n \to \infty$. The limit distribution when k is held fixed (extreme case) is considered in Section 4. In both the cases we obtain simple sufficient conditions and suggest norming constants which ensure the convergence to a non-degenerate random vector. Finally, in Section 5 we apply our results to the bivariate standard normal population and discuss their implications on the choice of k that satisfies the constraint $P(W_k \le 1 - \epsilon) \le \delta$. As we march along this route, whenever we have a step that is similar to the one in Nagaraja and David (1994), we refer to it for details.

We now introduce some notation. For a random variable or vector T, F_T represents its cdf and f_T its probability density function (pdf). We write $a(x) \approx b(x)$ if the ratio tends to 1 as $x \to \infty$. We let $x_0 = F_X^{-1}(q)$, where $q = 1 - p$, and $y_0 = \sup\{y : F_Y(y) < 1\}$, the two special quantiles which appear in our analysis. The cdfs $F_1(y|x) = P(Y \le y|X > x)$, $F_2(y|x) = P(Y \le y|X < x)$ and $F_3(y|x) = P(Y \le y|X = x)$ represent three conditional cdfs associated with Y. The symbols ϕ and Φ represent the standard normal pdf and cdf, respectively.

2. Finite-sample joint cdf of V_1 and V_2

The joint cdf of V_1 and V_2 can be expressed in a compact form by conditioning on the value of $X_{n-k:n}$. From Theorem 2 of Kaufman and Reiss (1992) it follows that, conditioned on the event $X_{n-k:n} = x$, V_1 behaves like the sample maximum of a random sample of size k from $F_1(\cdot|x)$, and V_2 behaves like the maximum of another set of $n - k$ independent random variables where $(n - k - 1)$ of these have cdf $F_2(\cdot|x)$ and the remaining one has cdf $F_3(\cdot|x)$. Further, these two sets of random variables are (conditionally) independent. Thus, we obtain

$$F_{V_1 V_2}(v_1, v_2) \equiv P(V_1 \le v_1, V_2 \le v_2)$$
$$= \mathrm{E}\{h(v_1, v_2, X_{n-k:n})\}, \tag{2.1}$$

where

$$h(v_1, v_2, x) = \{F_1(v_1|x)\}^k F_2(v_2|x)^{n-k-1} F_3(v_2|x). \tag{2.2}$$

Note that the joint pdf of V_1 and V_2 is given by

$$f_{V_1,V_2}(v_1, v_2) = k \int \{(n-k-1)F_3(v_2|x) f_2(v_2|x) + f_3(v_2|x) F_2(v_2|x)\}$$
$$\times \{F_1(v_1|x)\}^{k-1} F_2(v_2|x)^{n-k-2} f_1(v_1|x) f_{X_{n-k:n}}(x)\,\mathrm{d}x,$$

for $1 \le k \le n - 1$. Since $Y_{n:n} = \max\{V_1, V_2\}$, the joint cdf of V_1 and $Y_{n:n}$ is

$$P(V_1 \le v, Y_{n:n} \le y) = F_{V_1 V_2}(v, y), \qquad v < y,$$
$$= \{F_Y(y)\}^n, \quad v \ge y,$$

where the joint cdf of V_1 and V_2 is given by (2.1).

The joint distribution of V_1 and V_2 can be used to obtain the cdf of $W_k = V_1/Y_{n:n}$. Since $W_k \geq 1$ whenever $V_2 < 0$, for $w < 1$,

$$\begin{aligned} F_{W_k}(w) &= P(V_1 \leq V_2 w, V_2 > 0) \\ &= \int_{v_2>0} \int_{v_1<v_2 w} f_{V_1,V_2}(v_1, v_2)\,\mathrm{d}v_1 \mathrm{d}v_2 \\ &= \int_{-\infty}^{\infty} \int_0^{\infty} \{(n-k-1)F_3(v_2|x)f_2(v_2|x) + f_3(v_2|x)F_2(v_2|x)\} \\ &\quad \times \{F_1(v_2 w|x)\}^k F_2(v_2|x)^{n-k-2} f_{X_{n-k:n}}(x)\mathrm{d}v_2\,\mathrm{d}x. \end{aligned} \tag{2.3}$$

3. Asymptotic joint distribution in the quantile case

Theorem 1 Let $k = [np]$, $0 < p < 1$. Assume (i) f_X is continuous at x_0; (ii) $f_X(x_0) > 0$; (iii) $f_3(y|x)$ is continuous at x_0 for all y; and (iv) there exist constants a_n, $b_n > 0$, c_n, $d_n > 0$, such that as $n \to \infty$,

$$\{F_1(a_n + b_n y|x_0)\}^n \to G_1(y) \tag{3.1a}$$

$$\{F_2(c_n + d_n y|x_0)\}^n \to G_2(y), \tag{3.1b}$$

for all real y, where G_1 and G_2 are non-degenerate cdfs. Further, assume that

$$P(X \text{ is between } x_0 \text{ and } x_0 + c/\sqrt{n}, Y > a_n + b_n y) = o(1/n), \tag{3.2a}$$

$$P(X \text{ is between } x_0 \text{ and } x_0 + c/\sqrt{n}, Y > c_n + d_n y) = o(1/n), \tag{3.2b}$$

for all fixed real c and all y. Then

$$F_{V_1,V_2}(a_n + b_n v_1, c_n + d_n v_2) \to \{G_1(v_1)\}^p \{G_2(v_2)\}^q, \tag{3.3}$$

for all real v_1 and v_2.

Proof
Define $Z_n = \{\sqrt{n} f_X(x_0)(X_{n-k:n} - x_0)/\sqrt{pq}\}$, and note that from (2.1) and (2.2) we have

$$F_{V_1,V_2}(a_n + b_n v_1, c_n + d_n v_2) = \mathrm{E}\{h(a_n + b_n v_1, c_n + d_n v_2, x_0 + c_0 Z_n/\sqrt{n})\},$$

where $c_0 = \sqrt{pq}/f_X(x_0)$. Recall that

$$\begin{aligned} &h(a_n + b_n v_1, c_n + d_n v_2, x_0 + c_0 z/\sqrt{n}) \\ &= \{F_1(a_n + b_n v_1|x_0 + c_0 z/\sqrt{n})\}^k \{F_2(c_n + d_n v_2|x_0 + c_0 z/\sqrt{n})\}^{n-k-1} F_3(c_n + d_n v_2|x_0 + c_0 z/\sqrt{n}). \end{aligned}$$

In view of the arguments presented in the proof of Result 2 in Nagaraja and David (1994), it is enough to show that, for all v_i in the support of G_i, $i = 1,2$, and for all z,

$$\{F_1(a_n + b_n v_1 | x_0 + c_0 z/\sqrt{n})\}^n \to G_1(v_1), \tag{3.4}$$

$$\{F_2(c_n + d_n v_2 | x_0 + c_0 z/\sqrt{n})\}^n \to G_2(v_2), \tag{3.5}$$

$$F_3(c_n + d_n v_2 | x_0 + c_0 z/\sqrt{n}) \to 1. \tag{3.6}$$

Condition (3.1a) is equivalent to the fact that $n\{1 - F_1(a_n + b_n v | x_0)\} \to -\log G_1(v)$ which can be expressed as $nP(X > x_0,\ Y > a_n + b_n v) \to -p \log G_1(v) = \eta$, say. Thus, in order to establish (3.4) we have to show that $nP(X > x_0 + c_0 z/\sqrt{n},\ Y > a_n + b_n v) \to \eta$, where $0 < \eta < \infty$. The difference between the two sequences is nothing but $nP(x_0 < X \le x_0 + c_0 z/\sqrt{n},\ \ Y > a_n + b_n v)$, which approaches 0 by the assumption made in (3.2a). Hence (3.4) holds. Using (3.1b) and (3.2b), we can show along similar lines that (3.5) holds.

To prove (3.6), note that $f_3(y | x_0 + c_0 z/\sqrt{n})$ is a pdf for every n and converges to a pdf $f_3(y|x_0)$. Then from a convergence theorem involving pdfs (see, for example, Rao 1973, p. 124) it follows that

$$\int_{-\infty}^{c_n + d_n v} |f_3(y | x_0 + c_0 z/\sqrt{n}) - f_3(y|x_0)| \mathrm{d}y \to 0. \tag{3.7}$$

Now, for any v with $0 < G_2(v) < 1$, $c_n + d_n v$ approaches the upper limit of the support of $F_2(y|x_0)$. By continuity, $F_2(y|x_0) = P(Y \le y | X \le x_0)$ and hence, the upper bound for the support of $F_2(y|x_0)$ is not less than the corresponding bound for $F_3(y|x_0)$. Thus, we can conclude that $\int_{-\infty}^{c_n + d_n v} f_3(y|x_0)\mathrm{d}y \to 1$. This, in view of (3.7), implies that (3.6) holds. □

We now examine the conditions we have imposed to establish Theorem 1.

Remark 1

Instead of assuming (3.1a) and (3.2a) to establish (3.4), we could have assumed the latter condition in the statement of the theorem. But (3.4) implies that (3.1a) and (3.2a) hold for v_1 in the support of G_1 (see Nagaraja and David 1994, p. 484).

Remark 2

Since f_X is assumed to be continuous at x_0, a sufficient condition for the continuity of $f_3(y|x)$ at x_0 is that the joint pdf is also continuous at x_0 as a function of x.

Remark 3

Condition (3.1a) is equivalent to saying that the sample maximum of a random sample from the cdf $F_1(\cdot|x_0)$ converges in distribution to a non-degenerate random variable with cdf G_1, where G_1 is one of the three extreme-value cdfs. If this is the case, we say that $F_1(\cdot|x_0)$ is in the domain of attraction of G_1 and we write $F_1(\cdot|x_0) \in D(G_1)$. Following, for example, Resnick (1987), we will denote the possible forms for G_1 by Φ_α, Ψ_α and Λ. Condition (3.1b) has a similar interpretation.

The domain of attraction entails a specific right tail behaviour for the cdf $F_1(\cdot|x_0)$. Quite often, its tail is messier to manipulate than that of F_Y. But sometimes the two cdfs may be tail-equivalent; that is, there exists a finite and positive β_1, possibly dependent on x_0, such that

$$\lim_{y \to y_0} \frac{1 - F_Y(y)}{1 - F_1(y|x_0)} = \beta_1. \tag{3.8}$$

Since $1 - F_Y(y) = p\{1 - F_1(y|x_0)\} + q\{1 - F_2(y|x_0)\}$, if $F_1(\cdot|x_0)$ and F_Y are tail-equivalent with $\beta_1 > p$, $F_2(\cdot|x_0)$ and F_Y are also tail-equivalent. So if (3.8) holds and $\beta_1 > p$, then

$$\beta_2 \equiv \lim_{y \to y_0} \frac{1 - F_Y(y)}{1 - F_2(y|x_0)} = \frac{q\beta_1}{\beta_1 - p}. \tag{3.9}$$

From Resnick (1987, Proposition 1.19) it follows that if $F_Y \in D(G)$ as $n \to \infty$ such that $\{F_Y(a_n + b_n y)\}^n \to G(y)$, and (3.8) holds, then (3.1a) holds with $G_1(v) = G(a + bv)$. Further,

$$\begin{aligned} a &= 0 \text{ and } b^{\alpha} = \beta_1, && \text{if } G = \Phi_{\alpha}, \\ a &= 0 \text{ and } b^{-\alpha} = \beta_1, && \text{if } G = \Psi_{\alpha}, \\ a &= \log \beta_1 \text{ and } b = 1, && \text{if } G = \Lambda. \end{aligned} \tag{3.10}$$

If $\beta_1 > p$, then (3.1b) also holds with $c_n = a_n$ and $d_n = b_n$ and $G_2(v) = G(c + dv)$. The constants c and d are determined using (3.10), by replacing a, b and β_1 by c, d and β_2 (from (3.9)), respectively.

When $f_{X,Y}(x, y) = 2\exp(-x - y)$, $0 < x < y < \infty$, $\beta_1 = \sqrt{p}$ (Nagaraja and David 1994, p. 491) and hence the above conclusions hold. For the bivariate standard normal population, $\beta_1 = p$, (Nagaraja and David 1994, pp. 486–487) and thus $F_2(\cdot|x_0)$ and F_Y are not tail-equivalent. However, $F_2(\cdot|x_0)$ is also in the domain of attraction of Λ, as we shall show in Section 5.

Remark 4

Now we examine (3.2a) and introduce conditions that validate it. First, define $\bar{F}(x, y) = P(X > x, Y > y)$ and let $\Delta_1 \bar{F}(x, y)$ denote its first partial derivative with respect to x. When $c > 0$, from the mean value theorem (under appropriate assumptions), $P(x_0 < X \leq x_0 + c/\sqrt{n}, Y > a_n + b_n y) = \bar{F}(x_0, a_n + b_n y) - \bar{F}(x_0 + c/\sqrt{n}, \ a_n + b_n y) = -\{c/\sqrt{n}\}\Delta_1 \bar{F}(x^*, a_n + b_n y)$, where x^* is between x_0 and $x_0 + c/\sqrt{n}$. Thus, if the conditions

$$\Delta_1 \bar{F}(x, a_n + b_n y) = o(1/\sqrt{n}), \tag{3.11a}$$

$$\Delta_1 \bar{F}(x, c_n + d_n y) = o(1/\sqrt{n}), \tag{3.11b}$$

hold uniformly in x in a neighbourhood of x_0, then (3.2a) and (3.2b) hold, respectively.

4. Asymptotic joint distribution in the extreme case

The premise now is that k is held fixed and $n \to \infty$.

Lemma 1 Let there exist constants $a_n', b_n' > 0$, c_n', $d_n' > 0$ such that for non-degenerate cdfs G_X and G_Y,

$$\{F_X(a_n' + b_n' x)\}^n \to G_X(x) \text{ and } \{F_Y(c_n' + d_n' y)\}^n \to G_Y(y), \tag{4.1}$$

for all real x and y. Further, assume

$$nP(X > a_n' + b_n' x, Y > c_n' + d_n' y) \to 0. \tag{4.2}$$

Then,

$$\{F_2(c'_n + d'_n y | a'_n + b'_n x)\}^n \to G_Y(y). \tag{4.3}$$

Proof
Condition (4.3) holds if, and only if, $n\{1 - F_2(c'_n + d'_n y | a'_n + b'_n x)\} \to -\log G_Y(y)$. Now,

$$n\{1 - F_2(c'_n + d'_n y | a'_n + b'_n x)\}$$
$$= nP(Y > c'_n + d'_n y)\{F_X(a'_n + b'_n x)\}^{-1}[1 - \{P(X > a'_n + b'_n x, Y > c'_n + d'_n y)/P(Y > c'_n + d'_n y)\}].$$

It follows from (4.1) that $nP(Y > c'_n + d'_n y) \to -\log G_Y(y)$ and $F_X(a'_n + b'_n x) \to 1$. From (4.2) we can conclude that the last factor on the right-hand side above tends to 1. That is, (4.3) holds. □

Now suppose F_X satisfies one of the three von Mises conditions (see, for example, Resnick, 1987, Propositions 1.15–1.17). Then, the pdf of $(X_{n-k:n} - a'_n)/b'_n$ converges to the pdf $g_{(k)}$ (see Lemma 1 of Nagaraja and David 1994), where

$$g_{(k)}(w) = \frac{\{-\log G_X(w)\}^k}{k!} g_X(w). \tag{4.4}$$

Theorem 2 Suppose F_X satisfies one of the von Mises conditions, and, for all x and y, assume (4.1) and (4.2) hold where the norming constants are such that $F_3(c'_n + d'_n y | a'_n + b'_n x) \to 1$. Further, suppose there exist constants a^*_n, $b^*_n > 0$, such that

$$F_1(a^*_n + b^*_n y | a'_n + b'_n x) \to H(x, y), \tag{4.5}$$

as $n \to \infty$. Then,

$$F_{V_1,V_2}(a^*_n + b^*_n v_1, c'_n + d'_n v_2) \to \{G_Y(v_2)\} \int \{H(x, v_1)\}^k g_{(k)}(x)\, \mathrm{d}x, \tag{4.6}$$

where $g_{(k)}$ is given by (4.4).

Proof
First observe from (2.1) that

$$F_{V_1,V_2}(a^*_n + b^*_n v_1, c'_n + d'_n v_2) = \int \{F_1(a^*_n + b^*_n v_1 | a'_n + b'_n u)\}^k F_2(c'_n + d'_n v_2 | a'_n + b'_n u)^{n-k-1}$$
$$\times F_3(c'_n + d'_n v_2 | a'_n + b'_n u)\{b'_n f_{X_{n-k:n}}(a'_n + b'_n u)\} \mathrm{d}u.$$

Next, use Lemma 1, and follow the proof of Result 1 in Nagaraja and David (1994). We omit the details. □

Remark 5
Note that (4.1) and (4.2) together imply that the marginal maxima are asymptotically independent (see, for example, Galambos 1987, p. 301). Under the conditions assumed above, asymptotically V_1 and V_2 are independent, just as in the quantile case. Here, while the limit distribution of V_1 is related to that of $X_{n:n}$, V_2 behaves like $Y_{n:n}$ asymptotically.

5. Bivariate normal population

We begin by deriving the asymptotic distribution.

Theorem 3 Let (X, Y) be a bivariate standard normal population with correlation coefficient ρ $(0 < \rho < 1)$. Then, for all real v_1, v_2,

$$P\left(\frac{V_1 - a_n}{b_n} \leq v_1, \frac{V_2 - c_n}{d_n} \leq v_2\right) \to \mathrm{e}^{-\{\exp(-v_1)+q\exp(-v_2)\}} \tag{5.1}$$

if $k = [np]$, $0 < p < 1$, and

$$P\left(\frac{V_1 - \rho a_n}{\theta} \leq v_1, \frac{V_2 - a_n}{b_n} \leq v_2\right) \to \{\Phi(v_1)\}^k \mathrm{e}^{-\exp(-v_2)}, \tag{5.2}$$

if k is held fixed as $n \to \infty$. Here $\theta = \sqrt{1-\rho^2}$, and the other norming constants may be chosen as

$$a_n = \sqrt{2\log n} - \frac{1}{2}\frac{\log(4\pi\log n)}{\sqrt{2\log n}}$$
$$b_n = 1/\sqrt{2\log n}, \tag{5.3}$$

$$c_n = \rho x_0 + \theta\left\{\sqrt{2\log n} - \frac{\log(4\pi\log n)}{\sqrt{2\log n}}\right\} - \theta\frac{(x_0^2/2) + \log(q\rho/\theta)}{\sqrt{2\log n}},$$
$$d_n = \theta b_n. \tag{5.4}$$

Proof

We verify the conditions assumed in Theorems 1 and 2 and identify the norming constants involved.

(a) *Quantile case*. It is well known that, when the a_n's and b_n's are given by (5.3), $\{F_Y(a_n + b_n y)\}^n \to \Lambda(y)$. Since $F_1(\cdot|x_0)$ and F_Y are tail-equivalent with p being the β_1 in (3.8) (Nagaraja and David 1994, p. 486), from (3.10) we may conclude that (3.1a) holds with $G_1(y) = \Lambda(y + \log p)$.

We now show that (3.1b) holds with $G_2 = \Lambda$ or equivalently $F_2 \in D(\Lambda)$, where F_2 stands for $F_2(\cdot|x_0)$. To prove this and other claims that follow, we make repeated use of L'Hôpital's rule and the following consequence:

$$\Phi(-y) = \{1 - \Phi(y)\} \approx \phi(y)/y \qquad \text{as } y \to \infty. \tag{5.5}$$

Note that the pdf of F_2 is given by

$$f_2(y) = q^{-1}\phi(y)\Phi(x_0k_0 + k_1y), \tag{5.6}$$

where

$$k_0 = 1/\theta, \qquad k_1 = -\rho/\theta < 0. \tag{5.7}$$

Hence, on using L'Hôpital's rule and (5.5), it can be shown that as $y \to \infty$, $f_2'(y)\{yf_2(y)\}^{-1} \to$

$-(1+k_1^2) = -1/\theta^2$ and

$$\frac{y(1-F_2(y))}{f_2(y)} \to \frac{1}{1+k_1^2} = \theta^2. \tag{5.8}$$

Thus, $f_2'(y)\{1-F_2(y)\}\{f_2(y)\}^{-1} \to -1$ as $y \to \infty$. Therefore, from Proposition 1.1(b) in Resnick (1987, p. 40) it follows that $F_2 \in D(\Lambda)$.

We now choose c_n and d_n using tail-equivalence ideas. From (5.5), (5.6) and (5.8) it can be shown that, for large y,

$$1 - F_2(y) \approx \theta^3\phi(y)\phi(x_0k_0 + k_1y)\{\rho q y^2\}^{-1} = 1 - F_0(y), \tag{5.9}$$

say. Thus the same set of norming constants works for both these cdfs. Now suppose r_n can be chosen to satisfy (see Resnick 1987, p. 40) $\log\{1-F_0(r_n)\} = -\log n$. Since $1/nf_0(r_n) = \{1-F_0(r_n)\}/f_0(r_n) \approx \theta^2/r_n$ (recall (5.8)), with $d_n = \theta^2/r_n$, $\{F_0(r_n + d_n y)\}^n \to \Lambda(y)$.

From (5.9) it follows that r_n satisfies

$$r_n^2 - 2\rho x_0 r_n + 4\theta^2 \log r_n + \alpha = 2\theta^2 \log n, \tag{5.10}$$

where $\alpha = x_0^2 - 2\theta^2\log(\theta^3/2\pi\rho q)$. Since the first term on the left is the dominating term, $r_n \approx \theta\sqrt{2\log n}$. Consequently, d_n can be chosen as $\{\theta/\sqrt{2\log n}\}$.

Now our goal is to determine other terms in r_n up to $o(1/t)$, where $t = \sqrt{2\log n}$. So, we define the function

$$h(t) = \theta t + \theta_1 + \theta_2(\log t)/t + \theta_3/t, \tag{5.11}$$

and determine the cofficients θ_i such that (5.10) holds as $t \to \infty$ when $h(t)$ replaces r_n. That is, we have to satisfy the constraint:

$$\{h(t)\}^2 - 2\rho x_0 h(t) + 4\theta^2 \log h(t) + \alpha - \theta^2 t^2 \to 0.$$

On using (5.11), the expression above can be written as $2\theta t(\theta_1 - \rho x_0) + 2\theta(\theta_2 + 2\theta)\log t + (\theta_1^2 + 2\theta\theta_3 - 2\rho x_0\theta_1 + 4\theta^2\log\theta + \alpha) + o(1)$. Since this must approach 0 as $t \to \infty$, we must have $\theta_1 = \rho x_0$, $\theta_2 = -2\theta$, and $\theta_3 = -\theta\{(x_0^2/2) + \log\theta_0\}$ where $\theta_0 = (2\pi q\rho/\theta)$. In other words, $h(t)$ in (5.11) must have the form

$$h(t) = \theta t + \rho x_0 - 2\theta(\log t)/t - \theta\{(x_0^2/2) + \log\theta_0\}/t, \tag{5.12}$$

with $\theta = \sqrt{1-\rho^2}$ and $\theta_0 = (2\pi q\rho/\theta)$.

We choose $c_n = h(\sqrt{2\log n})$ and show that $(c_n - r_n)/d_n \to 0$, where r_n satisfies (5.10). For this, call $r_n = c_n + \delta(t)$ with t being $\sqrt{2\log n}$. We have to show $t\delta(t) \to 0$ as $t \to \infty$, where $h(t)$ satisfies (5.12) and $r_n = h(t) + \delta(t)$ satisfies

$$\{h(t)+\delta(t)\}^2 - 2\rho x_0\{h(t)+\delta(t)\} + 4\theta^2\log h(t) + 4\theta^2\log[1+\delta(t)\{h(t)\}^{-1}] + \alpha - \theta^2t^2 \equiv 0. \tag{5.13}$$

On substituting for $h(t)$ from (5.12), and noting that $h(t) \approx \theta t$, it can be seen that the left-hand side of (5.13) is of the form $2\theta t\delta(t) + o(1) + o(t\delta(t))$ as $t \to \infty$. Since the right-hand side is zero, this means $t\delta(t) \to 0$. Hence we have shown that we can take $c_n = h(\sqrt{2\log n})$, where $h(t)$ is given by (5.12), and $d_n = \theta b_n$. This establishes (5.4).

We will now show that (3.11b) holds. Note that

$$\Delta_1\bar{F}(x,y) = -\int_y^\infty f(x,v)\mathrm{d}v = -\int_y^\infty \phi(x)\phi(k_0v+k_1x)\mathrm{d}v$$
$$= -\phi(x)\{1-\Phi(k_0y+k_1x)\},$$

where k_0 and k_1 are given in (5.7). Thus, it is enough to show that, uniformly in x in a neighbourhood of x_0, $n\{1-\Phi(k_0(c_n+d_ny)+k_1x)\}^2 \to 0$. As $d_n \to 0$, this is true if $n\{1-\Phi(k_0c_n+k_1x)\}^2 \to 0$. This holds if $\{1-\Phi(k_0c_n+k_1x)\}^2/\{1-F_2(c_n)\} \to 0$, since $n\{1-F_2(c_n)\} \to 1$. In view of (5.6) this indeed is the case, if $A_x(y) \to 0$ as $y \to \infty$, for some fixed $x < x_0$, where

$$A_x(y) = \frac{\{1-\Phi(k_0y+k_1x)\}^2}{\int_y^\infty \phi(v)\Phi(k_0x_0+k_1v)\mathrm{d}v} \approx -2k_0\frac{k_0x_0+k_1y}{k_0y+k_1x}\frac{\phi^2(k_0y+k_1x)}{\phi(y)\phi(k_0x_0+k_1y)}. \tag{5.14}$$

The coefficient of y^2 in the exponent arising from the ratio of the normal densities above is negative, which implies that $A_x(y) \to 0$. Thus we have shown that (3.11b) holds.

To establish (3.11a), we have to show that $n\{1-\Phi(k_0a_n+k_1x)\}^2$ approaches 0 uniformly in x as $n\to\infty$. Since $0<\theta<1$, $a_n > c_n$ for large n, and thus, $n\{1-\Phi(k_0a_n+k_1x)\}^2$ is ultimately bounded by $n\{1-\Phi(k_0c_n+k_1x)\}^2$. We have just shown that this bound approaches 0 uniformly in x.

In view of Remark 4, we have verified the validity of (3.2a) and (3.2b). Thus, (5.1) follows from Theorem 1.

(b) *Extreme case*. Now let us assume k is held fixed and $n\to\infty$. We will show that all the conditions needed for the validity of Theorem 2 are satisfied. First note that Φ satisfies the von Mises condition (equation (1.4) of Resnick 1987, p. 40), and that (4.1) holds with $G_X = G_Y = \Lambda$, $a'_n = c'_n = a_n$, and $b'_n = d'_n = b_n$ where convenient choices for a_n and b_n are given by (5.3). It is well known (see, for example, Reiss 1989, p. 237, or Resnick 1987, p. 297) that for the bivariate normal parent, (4.2) holds. Further, $F_3(c'_n+d'_ny|a'_n+b'_nx) \to 1$, since $P(Y\le c'_n+d'_ny|X=a'_n+b'_nx) = \Phi[\{a_n+b_ny-\rho(a_n+b_nx)\}/\theta]$. From Nagaraja and David (1994, p. 483) it follows that with $a^*_n = \rho a_n$ and $b^*_n = \theta$, $F_1(a^*_n+b^*_ny|a'_n+b'_nx) \to \Phi(y)$. Thus (4.5) holds with $H(x,y)=\Phi(y)$. Hence (5.2) follows from (4.6). □

We now discuss how to choose k. Suppose we are given n, and small positive ϵ and δ, and the goal is to choose the smallest k for which $P((V_1/Y_{n:n}) \le 1-\epsilon) = F_{W_k}(1-\epsilon) \le \delta$. For this purpose we may use the expression for F_{W_k} given in (2.3). For the bivariate standard normal cdf $\Phi_{X,Y}$, on substitution, we obtain, for $w<1$,

$$F_{W_k}(w) = \frac{n!}{k!(n-k-1)!}\int_{-\infty}^\infty\int_0^\infty \{(n-k-1)\Phi(k_0x+k_1v)\Phi(k_0v+k_1x)\phi(v) + \theta^{-1}\phi(k_0x+k_1v)\Phi_{X,Y}(x,v)\}\{\Phi(wv)-\Phi_{X,Y}(x,wv)\}^k\{\Phi_{X,Y}(x,v)\}^{n-k-2}\phi(x)\phi(v)\mathrm{d}v\mathrm{d}x, \tag{5.15}$$

where k_0 and k_1 are given in (5.7). One can use numerical integration to find k for which $F_{W_k}(1-\epsilon)$ falls just below δ.

For large n, we can use the asymptotic approximations in order to simplify the task of integration. Since for $w < 1$,

$$F_{W_k}(w) = P((V_1 - \rho a_n) \leq \{w(V_2 - a_n) + (w-\rho)a_n\}, V_2 > 0),$$

in view of (5.2), it can be approximated for small k by

$$\int_{-\infty}^{\infty} \Phi^k(\theta^{-1}\{wb_n v + (w-\rho)a_n\})\, \mathrm{e}^{-\exp(-v)}\, \mathrm{e}^{-v}\, \mathrm{d}v.$$

For a given $w = 1-\epsilon$, when this expression falls below δ, if the associated k is not too small (say, $k > 0.02n$), we suggest the approximation based on (5.1). In that case, note that

$$F_{W_k}(w) = P(\{(V_1 - a_n)/b_n\} \leq w\theta\{(V_2 - c_n)/d_n\} + \gamma_n, V_2 > 0),$$

where $\gamma_n = (wc_n - a_n)/b_n$. On using (5.1), we can approximate the above probability by

$$q\int_{-\infty}^{\infty} \mathrm{e}^{-\{\exp(-(w\theta v + \gamma_n)) + q\exp(-v)\}}\, \mathrm{e}^{-v}\, \mathrm{d}v,$$

where $q = (n-k)/n$, and choose k accordingly.

We can use the asymptotic results in another way. Let k^* be the value of k suggested by the asymptotic considerations. Then k^* can be used as the preliminary value in our search for the smallest k satisfying $F_{W_k}(1-\epsilon) \leq \delta$, where $F_{W_k}(w)$ is given by (5.15).

Remark 6

Theorem 3 assumed that ρ is positive. When $\rho = 0$, V_1 and V_2 are independent for all n. Further, (5.1) holds if c_n is replaced by a_n and (5.2) holds with θ being 1. When $\rho < 0$, with $Z = -Y$, (X, Z) will be bivariate standard normal with positive correlation $|\rho|$. Since $(X,Z) \stackrel{\mathrm{d}}{=} (-X, -Z)$, $V_1 = \max\{-Z_{[n-k+1:n]}, \ldots, -Z_{[n:n]}\} \stackrel{\mathrm{d}}{=} \max\{Z_{[1:n]}, \ldots, Z_{[k:n]}\} = V_1^*$, say. Similarly we observe that $V_2 \stackrel{\mathrm{d}}{=} \max\{Z_{[k+1:n]}, \ldots, Z_{n:n}\} = V_2^*$, say. In fact, $(V_1, V_2) \stackrel{\mathrm{d}}{=} (V_1^*, V_2^*)$. Consequently, we can modify Theorem 3 to conclude the following:

(i) If $k = [np]$, $0 < p < 1$,

$$P\left(\frac{V_1 - c_n^*}{d_n} \leq v_1, \frac{V_2 - a_n}{b_n} \leq v_2\right) \to \mathrm{e}^{-\{p\exp(-v_1) + \exp(-v_2)\}},$$

where c_n^* is obtained by replacing ρ by $|\rho|$ and q by p in the expression for c_n in (5.4).

(ii) If k is held fixed as $n \to \infty$, replace $(V_1 - \rho a_n)$ by $(V_1 - |\rho| a_n)$ on the left-hand side of (5.2).

Acknowledgements

We are indebted to the referees for their helpful comments. This work began when the first author was visiting the Department of Statistics, Ohio State University, during 1991–2.

References

Bhattacharya, P.K. (1984) Induced order statistics: theory and applications. In P.R. Krishnaiah and P.K. Sen (eds), *Handbook of Statistics*, Vol. **4**, pp. 383–403. Amsterdam: North-Holland.
David, H.A. (1991) Concomitants of order statistics: Review and recent developments. In F.M. Hoppe (ed.), *Multiple Comparisons, Selection, and Applications in Biometry*, pp. 507–518. New York: Marcel Dekker.
Feinberg, F.M. and Huber, J. (1994) A theory of cutoff severity under imperfect information. Working paper, Fuqua School of Business, Duke University (To appear in *Management Science*.)
Galambos, J. (1987) *The Asymptotic Theory of Extreme Order Statistics* (2nd edn). Malabar, FL: Krieger.
Kaufmann, E. and Reiss, R.-D. (1992) On conditional distributions of nearest neighbors. *J. Multivariate Anal.*, **42**, 67–76.
Nagaraja, H.N. and David, H.A. (1994) Distribution of the maximum of concomitants of selected order statistics. *Ann. Statist.*, **22**, 478–494.
Rao, C.R. (1973) *Linear Statistical Inference and Its Applications*. New York: Wiley.
Reiss, R.-D. (1989) *Approximate Distributions of Order Statistics*. New York: Springer-Verlag.
Resnick, S.I. (1987) *Extreme Values, Regular Variation, and Point Processes*. New York: Springer-Verlag.

Received May 1994 and revised October 1994