Ann. Inst. Statist. Math.
Vol. 54, No. 3, 667–680 (2002)

FISHER INFORMATION IN AN ORDER STATISTIC AND ITS CONCOMITANT

Z. A. ABO-ELENEEN* AND H. N. NAGARAJA

Department of Statistics, Ohio State University, Columbus OH 43210-1247, U.S.A.,
e-mail: zaher@stat.ohio-state.edu; hnn@stat.ohio-state.edu

(Received July 14, 2000; revised February 9, 2001)

Abstract. Let (X, Y) have an absolutely continuous distribution with parameter θ. We suggest regularity conditions on the parent distribution that permit the definition of Fisher information (FI) about θ in an X-order statistic and its Y-concomitant that are obtained from a random sample from (X, Y). We describe some general properties of the FI in such individual pairs. For the Farlie-Gumbel-Morgenstern parent with dependence parameter θ, we investigate the properties of this FI, and obtain the asymptotic relative efficiency of the maximum likelihood estimator of θ for Type II censored bivariate samples. Assuming (X, Y) is Gumbel bivariate exponential of second type, and θ is the mean of Y, we evaluate the FI in the Y-concomitant of an X-order statistic and compare it with the FI in a single Y-order statistic.

Key words and phrases: Concomitants of order statistics, Fisher information, Farlie-Gumbel-Morgenstern family, Gumbel Type II bivariate exponential distribution, Type II censoring, maximum likelihood estimator.

1. Introduction

Suppose we have a random sample of size n from a continuous distribution with cumulative distribution function (cdf) $F_1(x;\theta)$ and probability density function (pdf) $f_1(x;\theta)$, where θ is a real valued parameter and the sample is arranged in ascending order. The question about which part of the ordered sample has more information has been discussed by Tukey (1965) and Nagaraja (1994) in terms of linear sensitivity, a measure based on the first two moments of linear functions of order statistics. Mehrotra *et al.* (1979), Park (1996), and Zheng and Gastwirth (2000) have studied properties of the Fisher information (FI) measure in blocks and collections of order statistics. The FI plays a valuable role in statistical inference through the information (Cramer-Rao) inequality and its association with the asymptotic properties of the maximum likelihood estimators (MLE).

Now suppose (X, Y) is absolutely continuous with joint cdf $F(x, y; \theta)$ and pdf $f(x, y; \theta)$. For $1 \le r \le n$, let $X_{r:n}$ be the r-th X-order statistic and $Y_{[r:n]}$ be its concomitant obtained from a random sample of size n from f. David and Nagaraja (1998) provide a review of the area of the concomitants of order statistics.

The joint pdf of $(X_{r:n}, Y_{[r:n]})$ is given by

$$f_{r:n}(x, y; \theta) = cf(x, y; \theta)[F_1(x;\theta)]^{r-1}[1 - F_1(x;\theta)]^{n-r}, \tag{1.1}$$

*Now at Faculty of Computer Science, Zagazig University, Zagazig, Egypt.

where

$$c = \frac{n!}{(r-1)!(n-r)!}. \tag{1.2}$$

The marginal pdf of $Y_{[r:n]}$ is given by

$$f_{[r:n]}(y;\theta) = c\int_{-\infty}^{\infty} f(x,y;\theta)[F_1(x;\theta)]^{r-1}[1-F_1(x;\theta)]^{n-r}dx. \tag{1.3}$$

For a fixed n, we use the notation f_r instead of $f_{r:n}$ in (1.1) and we write $f_{[r]}$ for $f_{[r:n]}$ when n is fixed in (1.3).

We investigate the properties of $I_\theta(X_{r:n}, Y_{[r:n]})$ and $I_\theta(Y_{[r:n]})$, the FI contained in $(X_{r:n}, Y_{[r:n]})$ and $Y_{[r:n]}$, respectively. These play an important role in the asymptotic properties of MLE's based on censored bivariate samples. In Section 2 we show that the "standard" regularity conditions used to define $I_\theta(X,Y)$, the FI contained in (X,Y), are enough to define $I_\theta(X_{r:n}, Y_{[r:n]})$ and $I_\theta(Y_{[r:n]})$. Under the assumption that F_1 is free of θ, we obtain a simple useful recurrence relation satisfied by $I_\theta(X_{r:n}, Y_{[r:n]})$ and show that it is additive in r. This permits easy computation of the FI in an arbitrary collection of order statistics and their concomitants. In Section 3 we give an explicit expression for $I_\theta(X_{r:n}, Y_{[r:n]})$, when θ is the dependence parameter of the Farlie-Gumbel-Morgenstern (FGM) cdf

$$F(x,y) = F_1(x)F_2(y)[1+\theta(1-F_1(x))(1-F_2(y))], \tag{1.4}$$

where $-1 < \theta < 1$, and F_2 is the marginal cdf of Y. We prove the monotonicity property of FI as a function of r and evaluate $I_\theta(X_{r:n}, Y_{[r:n]})$ for selected θ, r, and n. We find the asymptotic variance and the asymptotic relative efficiency (ARE) of the MLE of θ under Type II censoring. In Section 4 we consider the Gumbel Type II bivariate exponential distribution (G_2BVE) where we take $\theta = E(Y)$ and give an explicit expression for $I_\theta(Y_{[r:n]})$. We compare it with $I_\theta(Y_{r:n})$ and discuss some implications.

2. Fisher information regularity conditions and properties

2.1 *Regularity conditions for* $I_\theta(X_{r:n}, Y_{[r:n]})$

The FI about the real parameter θ contained in X is defined by $I_\theta(X) = E(\frac{\partial \log f_1(X;\theta)}{\partial\theta})^2$ under the following regularity conditions (see, for example, Rao (1973), p. 329):

(1) The parameter space Ω is a real non-degenerate interval.

(2) The pdf of X w.r.t. the σ-finite measure μ, $f_1(x;\theta)$, is differentiable w.r.t. θ for all $\theta \in \Omega$.

(3) For every measurable set $C \subset S$, the sample space,

$$\frac{\partial}{\partial\theta}\int_C f_1(x;\theta)d\mu = \int_C \frac{\partial f_1(x;\theta)}{\partial\theta}d\mu,$$

assuming the existence of the right side integral. One commonly used condition which validates (3) is:

(3*) There exists an $H(x)$ integrable (w.r.t. μ) such that $|\frac{\partial f_1(x;\theta)}{\partial\theta}| \leq H(x)$ for all $\theta \in \Omega$.

Nagaraja (1983) showed that conditions (1), (2) and (3*) on $f_1(x,\theta)$ serve as regularity conditions for defining $I_\theta(X_{r:n})$. Now assume:

(A1) Ω is a real non-degenerate interval.

(A2) $f(x,y;\theta)$ is differentiable w.r.t. θ for all $\theta \in \Omega$.

(A3*) There exists an integrable $H(x,y)$ such that $|\frac{\partial f(x,y;\theta)}{\partial\theta}| \le H(x,y)$ for all θ.

Note that (A3*) validates the required assumption

(A3) For any measurable set $C \subset S$, the sample space,

$$\frac{\partial}{\partial\theta}\int_C f(x,y;\theta)d\mu = \int_C \frac{\partial f(x,y;\theta)}{\partial\theta}d\mu.$$

Conditions (A2) and (A3*) imply that $f_r(x,y;\theta)$, given by (1.1), is differentiable w.r.t. θ, for all $\theta \in \Omega$. Further, since $0 \le F_1(x;\theta) \le 1$, (1.1) implies that

$$\left|\frac{\partial f_r(x,y;\theta)}{\partial\theta}\right| \le c\left|\frac{\partial f(x,y;\theta)}{\partial\theta}\right| + c(n-1)f(x,y;\theta)\left|\frac{\partial F_1(x;\theta)}{\partial\theta}\right|. \tag{2.1}$$

Now since

$$\begin{aligned}\frac{\partial F_1(x;\theta)}{\partial\theta} &= \int_{-\infty}^{x}\int_{-\infty}^{\infty}\frac{\partial f(u,v;\theta)}{\partial\theta}dvdu \text{ from (A3)},\\ \left|\frac{\partial F_1(x;\theta)}{\partial\theta}\right| &\le \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} H(u,v)dudv \text{ from (A3*)}\\ &= c_1 \text{ say.}\end{aligned} \tag{2.2}$$

Then

$$\begin{aligned}\left|\frac{\partial f_r(x,y;\theta)}{\partial\theta}\right| &\le cH(x,y) + (n-1)cc_1 f(x,y;\theta)\\ &\le cH(x,y) + c_2 f(x,y;\theta),\end{aligned} \tag{2.3}$$

where $c_2 = (n-1)cc_1$ and $H(x,y)$ is integrable. We will now show that

$$\frac{\partial}{\partial\theta}\int_C f_r(x,y;\theta)dxdy = \int_C \frac{\partial f_r(x,y;\theta)}{\partial\theta}dxdy \tag{2.4}$$

for all interior points of Ω, and for all (two dimensional) measurable sets C. Let θ_0 be an arbitrary but fixed point in Ω. Then from the mean value theorem

$$\frac{f_r(x,y;\theta) - f_r(x,y;\theta_0)}{\theta-\theta_0} = \frac{\partial f_r(x,y;\theta^{(1)})}{\partial\theta}$$

for some $\theta^{(1)}$ between θ and θ_0. In view of (2.3), we have

$$\left|\frac{f_r(x,y;\theta) - f_r(x,y;\theta_0)}{\theta-\theta_0}\right| \le cH(x,y) + c_2 f(x,y;\theta^{(1)}). \tag{2.5}$$

But

$$f(x,y;\theta^{(1)}) = f(x,y;\theta_0) + (\theta^{(1)}-\theta_0)\frac{\partial f(x,y;\theta^{(2)})}{\partial\theta} \tag{2.6}$$

for some $\theta^{(2)}$ between θ_0 and $\theta^{(1)}$. So from (2.5) and (2.6) we have, for $|\theta - \theta_0| < 1$,

$$\begin{aligned}(2.7)\quad \left|\frac{f_r(x,y;\theta) - f_r(x,y;\theta_0)}{\theta - \theta_0}\right| &\le c_2\left\{f(x,y;\theta_0) + (\theta^{(1)} - \theta_0)\frac{\partial f(x,y;\theta^{(2)})}{\partial\theta}\right\} + cH(x,y)\\ &\le c_2\{f(x,y;\theta_0) + |\theta^{(1)} - \theta_0|H(x,y)\} + cH(x,y)\\ &\le c_2 f(x,y;\theta_0) + (c + c_2)H(x,y).\end{aligned}$$

Let $\{\theta_m, m \ge 1\}$ be a sequence converging to θ_0. Without loss of generality one can take $|\theta_m - \theta_0| < 1$ for all m. Then with $g_m(x,y) = \frac{f_r(x,y;\theta_m) - f_r(x,y;\theta_0)}{\theta_m - \theta_0}$, we have $|g_m(x,y)| \le G(x,y)$, where $G(x,y)$, the function on the right hand side of (2.7), is integrable. So at $\theta = \theta_0$,

$$\begin{aligned}(2.8)\quad \int_C \frac{\partial f_r(x,y;\theta)}{\partial\theta}dxdy &= \int_C \lim_{m\to\infty} g_m(x,y)dxdy\\ &= \lim_{m\to\infty}\int_C g_m(x,y)dxdy\\ &= \lim_{m\to\infty}\frac{\int_C f_r(x,y;\theta_m)dxdy - \int_C f_r(x,y;\theta_0)dxdy}{\theta_m - \theta_0}\\ &= \frac{\partial}{\partial\theta}\int_C f_r(x,y;\theta)dxdy.\end{aligned}$$

The second step above follows from the Lebesgue Dominated Convergence Theorem. Since θ_0 is arbitrary, (2.4) holds for all θ. Hence we have the following result.

THEOREM 2.1. *Conditions* (A1), (A2), *and* (A3*) *on* $f(x,y;\theta)$ *can serve as regularity conditions for defining* $I_\theta(X_{r:n}, Y_{[r:n]})$, *the FI in* $(X_{r:n}, Y_{[r:n]})$.

2.2 *Regularity conditions for* $I_\theta(Y_{[r:n]})$

Now we show that $I_\theta(Y_{[r:n]})$ exists if (A1), (A2) and (A3*) hold. Conditions (A2) and (A3) imply that $f_{[r]}(y;\theta)$, given by (1.3), is differentiable w.r.t. θ. Since $0 \le F_1(x;\theta) \le 1$,

$$\begin{aligned}\left|\frac{\partial f_{[r]}(y;\theta)}{\partial\theta}\right| &\le c\int_{-\infty}^{\infty}\left|\frac{\partial f(x,y;\theta)}{\partial\theta}\right|dx\\ &\quad + c(n-1)\int_{-\infty}^{\infty} f(x,y;\theta)\left|\frac{\partial F_1(x;\theta)}{\partial\theta}\right|dx\\ &\le c\int_{-\infty}^{\infty} H(x,y)dx + (n-1)cc_1\int_{-\infty}^{\infty} f(x,y;\theta)dx\end{aligned}$$

from (A3*) and (2.2). Thus

$$(2.9)\quad \left|\frac{\partial f_{[r]}(y;\theta)}{\partial\theta}\right| \le cH_1(y) + (n-1)cc_1 f_2(y;\theta),$$

where $H_1(y) = \int_{-\infty}^{\infty} H(x,y)dx$, is integrable. Note that (2.9) resembles (2.3) and we follow the approach used to prove (2.8) and conclude that

$$(2.10)\quad \frac{\partial}{\partial\theta}\int_C f_{[r]}(y;\theta)dy = \int_C \frac{\partial f_{[r]}(y;\theta)}{\partial\theta}dy$$

for all interior points $\theta \in \Omega$, for all measurable sets C. This leads to the following.

THEOREM 2.2. *Conditions* (A1), (A2), *and* (A3*) *on the bivariate pdf* $f(x, y; \theta)$ *can serve as regularity conditions for defining* $I_\theta(Y_{[r:n]})$, *for* $1 \le r \le n$.

Remark 1. In addition to conditions (A1)–(A3), with a regularity condition that permits the interchange of integration with respect to μ and second derivative of f_1 w.r.t. θ, $I_\theta(X) = -E\frac{\partial^2 \log f_1(X;\theta)}{\partial\theta^2}$. Under a similar additional condition on $f(x, y; \theta)$, parallel results hold for $I_\theta(X_{r:n}, Y_{[r:n]})$ and $I_\theta(Y_{[r:n]})$.

2.3 *Properties of* $I_\theta(X_{r:n}, Y_{[r:n]})$

THEOREM 2.3. *If* F_1 *is free of* θ, *the basic triangle relationship satisfied by the moments of order statistics* (*see Arnold et al.* (1992), *p.* 111) *holds for* $I_\theta(X_{r:n}, Y_{[r:n]})$; *that is*,

$$nI_\theta(X_{r:n-1}, Y_{[r:n-1]}) = (n-r)I_\theta(X_{r:n}, Y_{[r:n]}) + rI_\theta(X_{r+1:n}, Y_{[r+1:n]}). \tag{2.11}$$

PROOF. If F_1 is free of θ, from (1.1) it follows that

$$\frac{\partial \log f_{r:n}(x, y; \theta)}{\partial\theta} = \frac{\partial \log f(x, y; \theta)}{\partial\theta} = g(x, y), \quad \text{say}. \tag{2.12}$$

Further, (1.1) also implies that $nf_{r:n-1}(x, y; \theta) = (n-r)f_{r:n}(x, y; \theta) + rf_{r+1:n}(x, y; \theta)$. Consequently

$$nEh(X_{r:n-1}, Y_{[r:n-1]}) = (n-r)Eh(X_{r:n}, Y_{[r:n]}) + rEh(X_{r+1:n}, Y_{[r+1:n]})$$

for any h. With $h = g^2$ in (2.12) we obtain (2.11). □

THEOREM 2.4. *If* F_1 *is free of* θ, $I_\theta(X_{r:n}, Y_{[r:n]})$ *is additive*; *that is*,

$$I_\theta(X_{r_1:n}, \dots, X_{r_k:n}, Y_{[r_1:n]}, \dots, Y_{[r_k:n]}) = \sum_{i=1}^{k} I_\theta(X_{r_i:n}, Y_{[r_i:n]}).$$

PROOF. The joint pdf of the collection is a generalization of (1.1) and is given by

$$f_r(x_1, \dots, x_k, y_1, \dots, y_k; \theta) = \frac{n!}{(r_1-1)!(n-r_k)!}\{F_1(x_1)\}^{r_1-1}\{1-F_1(x_k)\}^{n-r_k} \times \prod_{i=2}^{k} \frac{\{F_1(x_i) - F_1(x_{i-1})\}^{r_i - r_{i-1} - 1}}{(r_i - r_{i-1} - 1)!} \prod_{i=1}^{k} f(x_i, y_i; \theta).$$

Hence,

$$\left\{\frac{\partial \log f_r}{\partial\theta}\right\}^2 = \left\{\sum_{i=1}^{k} \frac{\partial \log f(x_i, y_i; \theta)}{\partial\theta}\right\}^2.$$

We expand the right side before taking expectations, and note that the expected value of each of the cross product terms is zero. This yields the desired result. □

Remark 2. When F_1 is free of θ, Theorem 2.3 indicates that one needs to evaluate only n expectations in order to find $I_\theta(X_{r:m}, Y_{[r:m]})$ for all $1 \leq r \leq m \leq n$. Further, all the recurrence relations that hold for the moments of $X_{r:n}$ do apply for the FI in $(X_{r:n}, Y_{[r:n]})$. The additive property of the FI, proven in Theorem 2.4, makes the computation of the FI in an arbitrary collection of order statistics and their concomitants to be trivial once the FI in individual pairs are known.

3. Farlie-Gumbel-Morgenstern Family

3.1 *FI of the dependence parameter*

Hutchinson and Lai ((1990); Sec. 5.2) provide an excellent introduction to the FGM family whose cdf is given by (1.4), and discuss its properties and applications to a variety of situations. One of its useful property is that the dependence parameter θ and the marginal distributions can be modelled separately. For the FGM family with normal, exponential, and logistic marginals, the correlation coefficient is a scalar multiple of θ. With $F_1(x) = x$ and $F_2(y) = y$, one obtains the copula form of the FGM family. This copula is a special member of the family having copulas with quadratic sections (see Nelsen (1999), pp. 68–70) and for such distributions also, F_1 is free of the parameters determining the dependence structure. Another example is the bivariate normal distribuion where θ is the correlation coefficient. Thus Theorems 2.3 and 2.4 are also applicable to these distributions. Here we focus on the FGM family.

Recently Smith and Moffatt (1999) have investigated FI about θ in FGM type bivariate logistic models with some special sampling schemes. Scaria and Nair (1999) have discussed some distributional properties of concomitants of order statistics from the FGM family. The pdf for the cdf in (1.4) is

$$f(x, y; \theta) = f_1(x) f_2(y)[1 + \theta(1 - 2F_1(x))(1 - 2F_2(y))], \tag{3.1}$$

where $-1 < \theta < 1$, and f_i is the pdf of F_i, $i = 1, 2$. The copula form of the above density (see Nelsen (1999), Chapter 2) is

$$f(x, y; \theta) = [1 + \theta C(x, y)], \qquad 0 \leq x, y \leq 1, \tag{3.2}$$

where

$$C(x, y) = (1 - 2x)(1 - 2y). \tag{3.3}$$

Without loss of generality we use (3.2) to determine $I_\theta(X_{r:n}, Y_{[r:n]})$. Conditions (A1)-(A3) are satisfied; in fact, in (A3*) we can choose $H(x, y) = [1 - |C(x, y)|]^{-1}$. Now from (2.12) and (3.2)

$$\frac{\partial \log f_r(x, y; \theta)}{\partial \theta} = \frac{C(x, y)}{1 + \theta C(x, y)},$$

and

$$I_\theta(X_{r:n}, Y_{[r:n]}) = c \int_0^1 \int_0^1 \frac{[C(x, y)]^2}{[1 + \theta C(x, y)]} x^{r-1}(1 - x)^{n-r} dx dy. \tag{3.4}$$

With $u = (1 - 2x)$ and $v = (1 - 2y)$, we obtain

$$I_\theta(X_{r:n}, Y_{[r:n]}) = \frac{c}{2^{n+1}} \int_{-1}^1 u^2(1 + u)^{n-r}(1 - u)^{r-1} \left[\int_{-1}^1 \frac{v^2}{(1 + \theta uv)} dv \right] du. \tag{3.5}$$

The factor $(1+\theta uv)^{-1}$ above can be expanded as $\sum_{j=0}^{\infty}(-\theta)^j u^j v^j$. Since this representation is uniformly convergent for all $x, y \in (0,1)$, one can integrate term-by-term in (3.5). Upon doing so and noting that $\int_{-1}^{1} v^{j+2} dv = 0$ whenever j is odd, we obtain

$$I_\theta(X_{r:n}, Y_{[r:n]}) = \frac{c}{2^n} \sum_{j=0}^{\infty} \left[\frac{\theta^{2j}}{2j+3} \int_{-1}^{1} u^{2j+2}(1+u)^{n-r}(1-u)^{r-1} du \right]. \tag{3.6}$$

For $r = 1$ we get

$$I_\theta(X_{1:n}, Y_{[1:n]}) = \frac{n}{2^n} \sum_{j=0}^{\infty} \left[\frac{\theta^{2j}}{2j+3} \int_{-1}^{1} u^{2j+2}(1+u)^{n-1} du \right],$$

and on expanding $(1+u)^{n-1}$ binomially, we obtain

$$I_\theta(X_{1:n}, Y_{[1:n]}) = \frac{n}{2^{n-1}} \sum_{j=0}^{\infty} \frac{\theta^{2j}}{2j+3} \sum_{i=0}^{[\frac{n-1}{2}]} \binom{n-1}{2i} \frac{1}{2j+2i+3}. \tag{3.7}$$

With $n = 1$ in (3.7) we get the FI in a single pair to be

$$I_\theta(X, Y) = \sum_{j=0}^{\infty} \frac{\theta^{2j}}{(2j+3)^2}. \tag{3.8}$$

3.2 *Properties of* $I_\theta(X_{r:n}, Y_{[r:n]})$

From (3.6) it follows that $I_\theta(X_{r:n}, Y_{[r:n]}) = I_\theta(X_{n-r+1:n}, Y_{[n-r+1:n]})$. Now we show that for $1 \le r < (\frac{n}{2}+1)$, $I_\theta(X_{r:n}, Y_{[r:n]})$ decreases as r increases. For this, consider

$$a_r = c \int_{-1}^{1} u^k (1+u)^{n-r}(1-u)^{r-1} du, \tag{3.9}$$

where n and k are fixed, k is an even number, and c is given by (1.2). With $w = (1+u)/2$, a_r can be expressed as

$$a_r = c2^{n+k} \int_0^1 (w-0.5)^k w^{n-r}(1-w)^{r-1} dw = 2^{n+k} E(W_r - 0.5)^k, \tag{3.10}$$

where W_r is a Beta$(n-r+1, r)$ random variable. Note that since k is even, $E(W_r - 0.5)^k = E(W_{n-r+1} - 0.5)^k$. Thus

$$\begin{aligned} a_r &= 2^{n+k-1} c \int_0^1 (w-0.5)^k \{(1-w)^{n-r} w^{r-1} + w^{n-r}(1-w)^{r-1}\} dw \\ &= 2^{n+k} E(W_r^* - 0.5)^k, \end{aligned}$$

where the pdf of W_r^* is given by

$$g_r(w) = \frac{c}{2}\{w^{r-1}(1-w)^{n-r} + w^{n-r}(1-w)^{r-1}\}, \qquad 0 < w < 1. \tag{3.11}$$

Observe that $g_r(w) = g_r(1-w)$ and define $h(w) = g_r(w)/g_{r+1}(w)$.

LEMMA 3.1. *When $1 \le r < n/2$, $h(w)$ is increasing in w for $0.5 < w < 1$.*

PROOF. We have, from (3.11),

$$h(w) = \frac{r}{n-r}\left\{\frac{w^{r-1}(1-w)^{n-r} + w^{n-r}(1-w)^{r-1}}{w^r(1-w)^{n-r-1} + w^{n-r}(1-w)^r}\right\} = \frac{r}{n-r}h_1(t),$$

where $t = \frac{w}{1-w}$ and $h_1(t) = (\frac{1+t^{n-2r+1}}{t+t^{n-2r}})$. Since $h_1'(t) > 0$ for $t > 1$ whenever $r < \frac{n}{2}$, it increases for $t > 1$. Since t increases with w and $t > 1$ whenever $0.5 < w < 1$, we conclude that $h(w)$ increases for $w \in (0.5, 1)$. □

LEMMA 3.2. *Let $g_r(w)$ be as defined in (3.11). Then, for $x \in [0.5, 1]$,*

$$\int_{0.5}^{x} g_r(w)dw \le \int_{0.5}^{x} g_{r+1}(w)dw \tag{3.12}$$

whenever $r < \frac{n}{2}$ and equality holds in (3.12) only when $x = 1$.

PROOF. From Lemma 3.1 we know that $h(w) = g_r(w)/g_{r+1}(w)$ increases in $(0.5, 1)$. Consider

$$r(x) = \frac{\int_{0.5}^{x} g_r(w)dw}{\int_{0.5}^{x} g_{r+1}(w)dw}, \qquad 0.5 < x < 1. \tag{3.13}$$

Note that $r'(x) > 0$ if, and only if,

$$h(x) > \frac{\int_{0.5}^{x} g_r(w)dw}{\int_{0.5}^{x} g_{r+1}(w)dw} = \frac{\int_{0.5}^{x} h(w)g_{r+1}(w)dw}{\int_{0.5}^{x} g_{r+1}(w)dw}.$$

Since $h(w)$ strictly increases, the last term above is less than $h(x)$ for $0.5 < x < 1$. Consequently, $r(x)$ increases (strictly) in (0.5,1) and $r(1) = 1$ since $g_r(w)$ and $g_{r+1}(w)$ are both densities in (0,1) that are symmetric around 0.5. This establishes Lemma 3.2. □

To show that $a_r > a_{r+1}$ for a_r in (3.9) whenever $r < \frac{n}{2}$, we need to establish that $E(W_r^* - 0.5)^k > E(W_{r+1}^* - 0.5)^k$ where W_r^* has pdf $g_r(w)$ given by (3.11). Let $V_r = (W_r^* - 0.5)^k$. Then, we have to show that $E(V_r) > E(V_{r+1})$. This holds if $P(V_r \le v) < P(V_{r+1} \le v)$, $0 < v < 0.5^k$, since both V_r and V_{r+1} are positive random variables. Consider

$$\begin{aligned}
P(V_r \le v) &= P((W_r^* - 0.5)^k \le v) \\
&= P(0.5 - v^* \le W_r^* \le 0.5 + v^*), \qquad v^* = v^{1/k} \\
&= 2P(0.5 \le W_r^* \le x), \qquad x = v^* + 0.5 \\
&= 2\int_{0.5}^{x} g_r(w)dw \\
&< 2\int_{0.5}^{x} g_{r+1}(w)dw \text{ from Lemma 3.2} \\
&= P(V_{r+1} \le v).
\end{aligned}$$

Thus, we have shown that when $r < \frac{n}{2}, a_r > a_{r+1}$. Also, when $r = n/2$, $a_r = a_{r+1}$ Since (3.6) implies that

$$I_\theta(X_{r:n}, Y_{[r:n]}) = \frac{1}{2^n}\sum_{j=0}^{\infty}\frac{\theta^{2j}}{2j+3}a_r(j)$$

where $a_r(j) = a_r$ of (3.9) with $k = 2j+2$, we have established that $I_\theta(X_{r:n}, Y_{[r:n]})$ decreases as r increases whenever $r < (\frac{n}{2}+1)$. Using (3.10), the above equation can be expressed as

$$I_\theta(X_{r:n}, Y_{[r:n]}) = \sum_{i=1}^{\infty} \frac{\theta^{2(i-1)}}{2i+1} E(2U_{r:n}-1)^{2i} \tag{3.14}$$

where $U_{r:n}$ is a Beta$(r, n-r+1)$ random variable (or the r-th order statistic from a random sample of size n from the standard uniform distribution). Note that (3.14) implies that $I_\theta(X_{r:n}, Y_{[r:n]})$ increases as $|\theta|$ increases.

The above discussion leads to the following result.

THEOREM 3.1. *For the FGM family with dependence parameter θ, $I_\theta(X_{r:n}, Y_{[r:n]})$ is given by* (3.14), *and it has the following properties:* (a) $I_\theta(X_{r:n}, Y_{[r:n]}) = I_{-\theta}(X_{r:n}, Y_{[r:n]})$, (b) $I_\theta(X_{r:n}, Y_{[r:n]}) = I_\theta(X_{n-r+1:n}, Y_{[n-r+1:n]})$, (c) $I_\theta(X_{r:n}, Y_{[r:n]})$ *decreases as r increases for $1 \le r < (\frac{n}{2}+1)$, and* (d) $I_\theta(X_{r:n}, Y_{[r:n]})$ *increases as $\theta(>0)$ increases.*

Remark 3. From (3.14) it follows that $I_\theta(X_{r:n}, Y_{[r:n]}) < \sum_{i=1}^{\infty} \theta^{2(i-1)}(2i+1)^{-1} = \theta^{-2}\{2\theta^{-1}\log\{(1+\theta)/(1-\theta)\} - 1\}$ and hence the FI is bounded. For large n and $r = [np]$, $0 < p < 1$, $U_{r:n} \to p$ in probability, and since $(2U_{r:n}-1)^{2i}$ is bounded, $E(2U_{r:n}-1)^{2i} \to (2p-1)^{2i}$ (Lehmann (1999), p. 71). Hence

$$I_\theta(X_{r:n}, Y_{[r:n]}) \approx \sum_{i=1}^{\infty} \frac{\theta^{2(i-1)}}{2i+1}(2p-1)^{2i} = \frac{1}{\theta^2}\left\{\frac{1}{2\theta_p}\log\frac{1+\theta_p}{1-\theta_p} - 1\right\}$$

where $\theta_p = \theta(2p-1)$.

3.3 *Discussion*

Table 1 provides the values of $I_\theta(X_{r:n}, Y_{[r:n]})$ as a function of n, r $(\le (n+1)/2)$ and θ, for $n = 1(1)5(5)15$, and $\theta = 0.25, 0.5, 0.75, 0.99$. The entries were computed using (3.6) and MATHEMATICA. For $n = 1$, (3.8) was used. The infinite series was truncated after 11 terms and this provided adequate accuracy. The first row represents $I_\theta(X, Y)$, the FI in a single pair. Since the FI in a random sample of size n is $nI_\theta(X, Y)$, the table can also be used to compute the proportion of the sample FI contained in a single pair $I_\theta(X_{r:n}, Y_{[r:n]})$. For example, when $n = 10$, the FI in the extreme pair ranges from 21% to 26% of the total FI as θ ranges from 0.25 to 0.99. When $n = 15$ the FI in the extreme pair varies in the range of 16% to 21%. In contrast, the FI in the central pair is no more than 1% of what is available in the complete sample. One could use (3.7) and (2.11) to find $I_\theta(X_{r:n}, Y_{[r:n]})$ for all r and n, $1 \le r \le n$, in a recursive manner. But this leads to significant rounding off errors for n as small as 15.

Since the FI is additive in this case, Table 1 can readily be used to obtain the information contained in singly or multiply censored bivariate samples from the FGM distribution. One just adds up the FI in individual pairs that constitute the censored sample. For example, when $n = 10$, the FI in the Type II censored sample consisting of the bottom (or the top) two pairs ranges from 35% to 39% of the total FI as θ varies from 0.25 to 0.99. With the two extreme pairs together, it ranges from 42% to 52% of the total FI. We can also obtain an explicit expression for the FI in a Type II right censored bivariate sample using (3.14) and Theorem 2.4.

Table 1. Fisher information in $(X_{r:n}, Y_{[r:n]})$ about the dependence parameter θ in the FGM family.

		θ			
n	r	0.25	0.5	0.75	0.99
1	1	0.1137	0.1226	0.1437	0.2073
2	1	0.1137	0.1226	0.1437	0.2073
3	1	0.1367	0.1482	0.1758	0.2619
3	2	0.0678	0.0714	0.0794	0.0979
4	1	0.1596	0.1738	0.208	0.3166
4	2	0.0678	0.0714	0.0794	0.0979
5	1	0.1793	0.1959	0.2360	0.3922
5	2	0.0808	0.0856	0.0961	0.1229
5	3	0.0482	0.0502	0.0543	0.0626
10	1	0.2393	0.2641	0.3257	0.5366
10	2	0.1550	0.1668	0.1938	0.2658
10	3	0.0925	0.0976	0.1086	0.1329
10	4	0.0511	0.0532	0.0572	0.0651
10	5	0.0306	0.0314	0.0331	0.0359
15	1	0.2681	0.2978	0.3728	0.6408
15	2	0.2012	0.2187	0.2598	0.3773
15	3	0.1451	0.1550	0.1768	0.2289
15	4	0.0996	0.1049	0.1159	0.1389
15	5	0.0645	0.0672	0.0724	0.0823
15	6	0.0396	0.0408	0.0432	0.0472
15	7	0.0247	0.0252	0.0263	0.0280
15	8	0.0197	0.0201	0.0208	0.0218

THEOREM 3.2. *With* $(\boldsymbol{X}_{r,n}, \boldsymbol{Y}_{r,n}) = ((X_{i:n}, Y_{[i:n]}),\ 1 \le i \le r < n)$,

$$I_\theta(\boldsymbol{X}_{r,n}, \boldsymbol{Y}_{r,n}) = \frac{n}{2}\left\{ I_\theta(X,Y) + \sum_{i=1}^{\infty} \frac{\theta^{2(i-1)}}{(2i+1)^2} E(2U_{r:n-1}-1)^{2i+1} \right\}. \tag{3.15}$$

PROOF. From (3.14) and Theorem 2.4 we have $I_\theta(\boldsymbol{X}_{r,n}, \boldsymbol{Y}_{r,n}) = \sum_{i=1}^{\infty} \frac{\theta^{2(i-1)}}{2i+1} E(T_{i,r})$ where $T_{i,r} = \sum_{j=1}^{r}(2U_{j:n}-1)^{2i}$. Note that $E(T_{i,r}) = E(E(T_{i,r} \mid U_{r+1:n}))$, and given $U_{r+1:n} = u$, $U_{1:n} \cdots U_{r:n}$ behave like the order statistics of a random sample of size r from a uniform distribution over $(0,u)$. Hence

$$E(T_{i,r} \mid U_{r+1:n} = u) = rE((2U-1)^{2i} \mid U \le u) = \frac{r}{2u(2i+1)}\{1 + (2u-1)^{2i+1}\}.$$

Thus, one obtains

$$E(T_{i,r}) = \frac{n}{2(2i+1)}\{1 + E(2U_{r:n-1}-1)^{2i+1}\},$$

and

$$I_\theta(\boldsymbol{X}_{r,n}, \boldsymbol{Y}_{r,n}) = \frac{n}{2}\sum_{i=1}^{\infty} \frac{\theta^{2(i-1)}}{(2i+1)^2}\{1 + E(2U_{r:n-1}-1)^{2i+1}\}.$$

We now obtain (3.15) upon using (3.8). □

The above theorem can be used to obtain the asymptotic variance and ARE (Lehmann (1999), Chapter 7) of the MLE under Type II right censoring of X values and associated concomitants. Let $\hat{\theta}_r$ and $\hat{\theta}_n$ be the MLE of θ based the above censored sample, and the entire sample, respectively. When $r = [np]$, $0 < p < 1$, and n is large, $\text{Var}(\hat{\theta}_r) \approx [I_\theta(\boldsymbol{X}_{r,n}, \boldsymbol{Y}_{r,n})]^{-1}$ and $\text{Var}(\hat{\theta}_n) \approx [nI_\theta(X,Y)]^{-1}$. With $r = [np]$, $U_{r:n-1} \to p$ in probability and, as argued in Remark 3, $E(2U_{r:n-1} - 1)^{2i+1} \to (2p-1)^{2i+1}$. Hence from (3.15) and (3.8), it follows that

$$\frac{1}{n} I_\theta(\boldsymbol{X}_{r,n}, \boldsymbol{Y}_{r,n}) \to \frac{1}{2}\left\{ I_\theta(X,Y) + (2p-1)^3 \sum_{i=1}^{\infty} \frac{[\theta(2p-1)]^{2(i-1)}}{(2i+1)^2} \right\}$$
$$= \frac{1}{2}\{I_\theta(X,Y) + (2p-1)^3 I_{\theta_p}(X,Y)\},$$

where $\theta_p = \theta(2p-1)$. Thus

$$ARE(\hat{\theta}_r, \hat{\theta}_n) = \frac{I_\theta(X,Y) + (2p-1)^3 I_{\theta_p}(X,Y)}{2I_\theta(X,Y)}, \tag{3.16}$$

and the ARE can be computed for all θ and p, if $I_\theta(X,Y)$ is available for all $\theta > 0$. Hence Table 1 has another application. For example, when $\theta = 0.75$ and $p = 1/3$, $\theta_p = -0.25$, and upon using Table 1 entries, we obtain $ARE(\hat{\theta}_r, \hat{\theta}_n) = 0.49$. In other words, when $\theta = 0.75$, the MLE based on bottom 33% of the X's and their concomitants is as efficient as using a random sample consisting of 49% of the pairs.

4. Gumbel's type II bivariate exponential distribution

4.1 *FI in* $Y_{[r:n]}$ *about* $E(Y)$

A special distribution in the FGM family is the G_2BVE (Gumbel (1960)), where the marginal distributions are exponential and the joint cdf takes the form

$$F(x,y;\theta) = \left(1 - \exp\left\{\frac{-x}{\theta_1}\right\}\right)\left(1 - \exp\left\{\frac{-y}{\theta}\right\}\right)\left(1 + \alpha \exp\left\{-\frac{x}{\theta_1} - \frac{y}{\theta}\right\}\right), \tag{4.1}$$

where $x > 0$, $y > 0$, θ_1, $\theta > 0$, and $-1 < \alpha < 1$. Here, for convenience, we have labeled the mean of F_2, the parameter of interest, as θ, and the dependence parameter as α. The pdf of $Y_{[r:n]}$ (Balasubramanian and Beg (1997)) is

$$f_{[r]}(y;\theta) = \frac{1}{\theta}\exp\left\{\frac{-y}{\theta}\right\}\left(d_1 + d_2 \exp\left\{\frac{-y}{\theta}\right\}\right) \tag{4.2}$$

where

$$d_1 = d_1(r,n;\alpha) = 1 - \alpha + \frac{2\alpha r}{n+1} \quad \text{and} \quad d_2 = 2(1-d_1). \tag{4.3}$$

Since

$$\frac{\partial \log f_{[r]}(y;\theta)}{\partial \theta} = \frac{1}{\theta}\left\{ 2\frac{y}{\theta} - 1 - \frac{y}{\theta}\frac{d_1}{d_1 + d_2 \exp\left\{\frac{-y}{\theta}\right\}} \right\},$$

$$\text{(4.4)}\qquad \left(\frac{\partial \log f_{[r]}(y;\theta)}{\partial\theta}\right)^2 = \frac{1}{\theta^2}\left\{1-4w+4w^2-\frac{4w^2d_1}{d_1+d_2\exp\{-w\}} + \frac{2wd_1}{d_1+d_2\exp\{-w\}}+\frac{w^2{d_1}^2}{(d_1+d_2\exp\{-w\})^2}\right\}$$

where $W=Y/\theta$ has the pdf $f_W(w)=e^{-w}(d_1+d_2e^{-w})$. Upon taking expectation in (4.4), we obtain

$$\text{(4.5)}\qquad I_\theta(Y_{[r:n]})=\frac{1}{\theta^2}\left\{1-2d_1+{d_1}^2\int_0^\infty \frac{w^2\exp\{-w\}}{d_1+d_2\exp\{-w\}}dw\right\}$$

upon simplification, with d_1 and d_2 given by (4.3).

4.2 *Properties of* $I_\theta(Y_{[r:n]})$

Since $d_1(r,n;\alpha)=d_1(n-r+1,n;-\alpha)$, $I_\theta(Y_{[r:n]};\alpha)=I_\theta(Y_{[n-r+1:n]};-\alpha)$. We now fix $\alpha>0$ and examine (4.5) as a function of r and n. For this purpose we relabel d_1 as $t=1-\alpha+2\alpha r(n+1)^{-1}$ and note that $I_\theta(Y_{[r:n]})=\theta^{-2}g(t)$ with

$$\text{(4.6)}\qquad g(t)=\frac{1}{\theta^2}\left\{1-2t+t^2\int_0^\infty \frac{w^2\exp\{-w\}}{t+2(1-t)\exp\{-w\}}dw\right\},$$

and $0\le 1-\alpha\le t\le 1+\alpha\le 2$. We used MAPLE to evaluate $g(t)$ and as Fig. 1 shows, $g(t)$ decreases in (0, 0.383) and then increases. Further its maximum value is 1.80823, attained when $t=2$, and the minimum value is 0.82617. Thus, $I_\theta(Y_{[r:n]})$ increases with r if $1-\alpha>0.383$; otherwise it decreases for a while and then increases. The maximum FI is at $Y_{[n:n]}$. Further $I_\theta(Y_{[r:n]})$ does not exceed 1.81 and thus no single concomitant is more efficient than the mean of a random sample of size 2 from the exponential parent. In contrast, for $r=[np]$, $0<p<1$, and n large, $\theta^{-2}I_\theta(Y_{r:n})\approx n(1-p)p^{-1}[\log(1-p)]^2$; it becomes unbounded and peaks around 65% of the total FI in the complete Y-sample, when $p\approx 0.8$ (Arnold *et al.* (1992), p. 166).

Table 2 provides $I_\theta(Y_{[r:n]})$ and $I_\theta(Y_{r:n})$ values for $1\le r\le n$, $n=5,15$, $\alpha=0.25,0.5,0.75,0.99$, for $\theta=1$. The $I_\theta(Y_{r:n})$ values are evaluated using (7.3.8) of Arnold *et al.* (1992). As Table 2 and the above discussion show, $I_\theta(Y_{[r:n]};\alpha)$ is much smaller than $I_\theta(Y_{r:n})$.

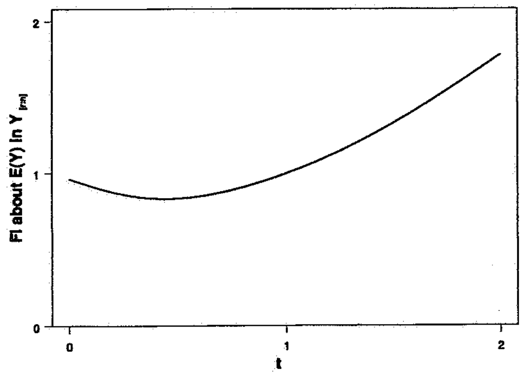

Fig. 1. $I_\theta(Y_{[r:n]})$ values for the G_2BVE parent with $\theta=1$; $t=1-\alpha+2\alpha r(n+1)^{-1}$.

Table 2. Fisher information in $Y_{[r:n]}$, $Y_{r:n}$, and (X, Y) for $\theta = 1$ where $\theta = E(Y)$ of the G_2BVE distribution with dependence parameter α.

		α				$I_\theta(Y_{r:n})$
n	r	0.25	0.5	0.75	0.99	
5	1	0.92515	0.86872	0.83414	0.82743	1.00000
5	2	0.96042	0.92515	0.89447	0.86965	1.80000
5	3	1.00000	1.00000	1.000000	1.00000	2.87500
5	4	1.04370	1.09139	1.14299	1.19618	3.56944
5	5	1.09139	1.19847	1.32127	1.45493	3.66204
15	1	0.90542	0.84423	0.82722	0.87892	1.00000
15	2	0.91704	0.85785	0.82796	0.83761	1.93330
15	3	0.92932	0.87467	0.83873	0.82617	2.98980
15	4	0.94223	0.89447	0.85785	0.83481	3.97247
15	5	0.95577	0.91704	0.88421	0.85877	4.94029
15	6	0.96992	0.94223	0.91704	0.89532	5.88606
15	7	0.98467	0.96992	0.95577	0.94276	6.79997
15	8	1.00000	1.00000	1.00000	1.00000	7.66830
15	9	1.01591	1.03240	1.04944	1.06634	8.47137
15	10	1.03240	1.06705	1.10392	1.14138	9.17998
15	11	1.04944	1.10392	1.16334	1.22501	9.74877
15	12	1.06705	1.14299	1.22768	1.31735	10.10290
15	13	1.08521	1.18423	1.29702	1.41889	10.10740
15	14	1.10392	1.22768	1.37155	1.53059	9.48492
15	15	1.12318	1.27334	1.45160	1.65442	7.51407
$I_\theta(X,Y)$		1.00612	1.02540	1.05954	1.11182	

How does $I_\theta(Y_{r:n})$ compare with the total information in the full bivariate sample, $nI_\theta(X, Y)$? Using the G_2BVE cdf in (4.1), we obtain

$$\frac{\partial \log f(x,y;\theta)}{\partial\theta}$$
$$= \frac{1}{\theta}\left\{-1+\frac{y}{\theta}+\frac{2\alpha y}{\theta}\frac{\exp\left\{\frac{-y}{\theta}\right\}\left(2\exp\left\{\frac{-x}{\theta_1}\right\}-1\right)}{\left[1+\alpha\left(2\exp\left\{\frac{-x}{\theta_1}\right\}-1\right)\left(2\exp\left\{\frac{-y}{\theta}\right\}-1\right)\right]}\right\}.$$

On letting $u = x/\theta_1$ and $v = y/\theta$, we get

$$(4.7)\; I_\theta(X,Y) = E\left(\frac{\partial \log f(X,Y;\theta)}{\partial\theta}\right)^2$$
$$= \frac{1}{\theta^2}\left\{1+4\alpha^2\int_0^\infty\int_0^\infty \frac{v^2\exp\{-3v\}(2\exp\{-u\}-1)^2\exp\{u\}}{[1+\alpha(2\exp\{-v\}-1)(2\exp\{-u\}-1)]}dudv\right\},$$

upon simplification. By expanding the denominator of the integrand above as a power series in α and carrying out term-by-term integration, we get

$$I_\theta(X,Y) = \frac{1}{\theta^2}\left\{1 + 4\sum_{j=0}^{\infty}\frac{\alpha^{2j+2}}{2j+3}\int_0^\infty v^2 \exp\{-3v\}(2\exp\{-v\}-1)dv\right\}. \tag{4.8}$$

This indicates $I_\theta(X,Y)$ is increasing with $|\alpha|$ whereas $I_\theta(Y_{r:n})$ is nonmonotonic. The last row of Table 2 contains $I_\theta(X,Y)$ values computed using (4.7) and these do not exceed 1.115 suggesting that the additional gain in efficiency while using the bivariate sample is no more than 12% when compared to the complete Y-sample. see below*

Acknowledgements

The authors thank the two referees for their helpful comments. The first author's research was supported by a training grant from the Egyptian government and the second author's research was supported in part by National Institutes of Health, USA, Grant # M01 RR00034.

References

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1992). *A First Course in Order Statistics*, Wiley, New York.

Balasubramanian, K. and Beg, M. I. (1997). Concomitants of order statistics in Morgenstern type bivariate exponential distribution, *J. Appl. Statist. Sci.*, **5**, 233–245.

David, H. A. and Nagaraja, H. N. (1998). Concomitants of order statistics, *Handbook of Statistics*, Vol. 16 (eds. N. Balakrishnan and C. R. Rao) 487–513, Elsevier, Amsterdam.

Gumbel, E. J. (1960). Bivariate exponential distributions, *J. Amer. Statist. Assoc.*, **55**, 698–707.

Hutchinson, T. P. and Lai, C. D. (1990). *Continuous Bivariate Distributions, Emphasising Applications*, Rumsby Scientific, Adelaide, Australia.

Lehmann, E. L. (1999). *Elements of Large-sample Theory*, Springer, New York.

Mehrotra, K. G., Johnson, R. A. and Bhattacharyya, G. K. (1979). Exact Fisher information for censored samples and the extended hazard rate functions, *Comm. Statist. Theory Methods*, **15**, 1493–1510.

Nagaraja, H. N. (1983). On the information contained in order statistics, Tech. Report, No. 278, Department of Statistics, The Ohio State University, Columbus, Ohio.

Nagaraja, H. N. (1994). Tukey's linear sensitivity and order statistics, *Ann. Inst. Statist. Math.*, **46**, 757–768.

Nelsen, R. B. (1999). *An Introduction to Copulas*, Lecture Notes in Statist., **139**, Springer, New York.

Park, S. (1996). Fisher information in order statistics, *J. Amer. Statist. Assoc.*, **91**, 385–390.

Rao, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed., Wiley, New York.

Scaria, J. and Nair, U. (1999). On concomitants of order statistics from Morgenstern family, *Biometrical J.*, **41**, 483–489.

Smith, M. D. and Moffatt, P. G. (1999). Fisher's information on the correlation coefficient in bivariate logistic models, *Australian & New Zealand Journal of Statistics*, **41**, 315–330.

Tukey, J. W. (1965). Which part of the sample contains the information?, *Proc. Nat. Acad. Sci., U.S.A.*, **53**, 127–134.

Zheng, G. and Gastwirth, J. L. (2000). Where is the Fisher information in an ordered sample?, *Statist. Sinica*, **10**, 1267–1280.

*It, along with the last column, also provides a comparison of $nI_\theta(X,Y)$ and $I_\theta(Y_{r:n})$ values.