SORT 33 (2) July-December 2009, 213-232

How much Fisher information is contained in record values and their concomitants in the presence of inter-record times?

Morteza Amini and Jafar Ahmadi*

Abstract

It is shown that, although the distribution of inter-record time does not depend on the parent distribution, Fisher information increases when inter-record times are included. The general results concern different classes of bivariate distributions and propose a comparison study of the Fisher information. This study is done in situations in which the univariate counterpart of the underlying bivariate family belongs to a general continuous parametric family and its well-known subclasses such as location-scale and shape families, exponential family and proportional (reversed) hazard model. We derived some explicit formulas for the additional information of record time given records and their concomitants (bivariate records) for some classes of bivariate distributions. Some common distributions are considered as examples for illustrations and are classified according to this criterion. A simulation study and a real data example from bivariate normal distribution are considered to study the relative efficiencies of estimator based on bivariate record values and inter-record times with respect to the corresponding estimator based on iid sample of the same size and bivariate records only.

MSC: 62G30; 62B10; 62G32; 62B05.

Keywords: Bivariate family, hazard rate function, reversed hazard rate, location and scale families, proportional (reversed) hazard model, shape family.

1 Introduction

Let $\{(X_i,Y_i), i \geq 1\}$ be a sequence of bivariate random variables from a continuous distribution with the real valued parameter θ. Let $\{R_n, n \geq 1\}$ be the sequence of record values in the sequence of X's. Then the Y-variable associated with the X-value which is

**Address for correspondence:* Department of Statistics, School of Mathematical Sciences, Ferdowsi University of Mashhad, P.O. Box 91775-1159, Mashhad, Iran. E-mail addresses: mort.amini@gmail.com (Morteza Amini), ahmadi-j@um.ac.ir (J. Ahmadi).

Received: June 2009
Accepted: October 2009

quantified as the nth record is called the concomitant of the nth record and is denoted by $R_{[n]}$. The most important use of concomitants of record values arises in experiments in which a specified characteristic's measurements of an individual are made sequentially and only values that exceed or fall below the current extreme value are recorded. So the only observations are bivariate record values, i.e., records and their concomitants. Such situations often occur in industrial stress, life time experiments, sporting matches, weather data recording and some other experimental fields.

Under certain regularity conditions, the Fisher information about the real parameter θ contained in a random variable X with density $f(x;\theta)$ is defined by (see, for example, Lehmann, 1989, p. 115), $I_X(\theta) = E\left(\frac{\partial \log f(X;\theta)}{\partial\theta}\right)^2 = -E\left(\frac{\partial^2 \log f(X;\theta)}{\partial\theta^2}\right)$. The Fisher information plays an important role in statistical estimation and inference through the information (Cramér-Rao) inequality and its association with the asymptotic properties, especially the asymptotic variance of the maximum likelihood estimators. It can also be used to compute the variance of the estimator whose variance is equal to Cramér-Rao lower bound, i.e., $\delta(X)$, $\text{Var}(\delta(X)) = (\frac{\partial}{\partial\theta}\text{E}\delta(X))^2/I_X(\theta)$. Abo-Eleneen and Nagaraja (2002) investigated some properties of Fisher information in an order statistic and its concomitant. Recently, Nagaraja and Abo-Eleneen (2008) considered bivariate censored samples and evaluated the Fisher information contained in a collection of order statistics and their concomitans.

Several authors have considered the amount of Fisher information in record data and have discussed its applications in inference [see, for example, Ahmadi and Arghami (2001, 2003), Hofmann and Nagaraja (2003), Balakrishnan and Stepanov (2005) and references therein]. However, the treatment of Fisher information contained in the bivariate record values is very limited. The question "How much information is contained in records and their concomitants about a specified parameter?" was addressed by Amini and Ahmadi (2007, 2008).

The time at which a record appears is called *record time*. There is no information, in record times themselves, about the sampling distribution, since for a continuous sampling distribution F, the joint distribution of record times does not depend on F (see, Arnold *et al.*, 1998, Section 2.5). Nevertheless, there is crucial information about F in the joint distribution of record times and record values. Actually, in the process of obtaining the bivariate record values, one usually observes the record times. So, it is worthwhile to use them, since they provide meaningful additional information. Ahmadi and Arghami (2003) and Hofmann (2004) presented some comparison results of Fisher information in univariate record values and record times with the Fisher information contained in the same number of random univariate observations. The aim of this paper is to investigate the amount of Fisher information in bivariate record values in the presence of inter-record times in some well-known bivariate classes of distributions. We have especially focused on the increment of Fisher information by considering inter-record times. We also study some estimation results based on bivariate record values and inter-record times.

The rest of paper is organized as follows. Section 2 contains some preliminaries and introduction to some classes of univariate and bivariate distributions. In Section 3, we establish some general results to compare the amount of the Fisher information contained in a set of the first n bivariate record values and inter-record times with a bivariate random sample of same size from the parent distribution. For each result, we give some examples for illustration. In Section 4, a simulation study and a real data example from bivariate normal distribution are also presented.

2 Preliminaries

Let $\{(X_i,Y_i), i \geq 1\}$ be a sequence of iid bivariate random variables with an absolutely continuous cumulative distribution function (cdf) $F_{X,Y}(x,y;\theta)$, where θ is a real valued parameter. The marginal probability density function (pdf) and cdf of X are denoted by $f_X(x;\theta)$ and $F_X(x;\theta)$, respectively. Furthermore, $h_X(x;\theta) = f_X(x;\theta)/\bar{F}_X(x;\theta)$ and $\tilde{h}_X(x;\theta) = f_X(x;\theta)/F_X(x;\theta)$ are the hazard rate and the reversed hazard rate functions of X, respectively, where $\bar{F}_X(x;\theta) = 1 - F_X(x;\theta)$.

The sequence of bivariate record values is defined as $(R_n, R_{[n]}) = (X_{T_n}, Y_{T_n})$, $n \geq 1$, where $T_1 = 1$ with probability one and for $n \geq 2$, $T_n = \min\{j : j > T_{n-1},\ X_j > X_{T_{n-1}}\}$.

An analogous definition deals with lower records and their concomitants. In this paper, we assume that the data available for study are records (upper or lower), inter-record times and their concomitants. Such data may be rewritten as $(R_1,\Delta_1,R_{[1]}),(R_2,\Delta_2,R_{[2]}), \ldots, (R_n,\Delta_n,R_{[n]})$, where $\Delta_i = T_{i+1} - T_i - 1$, $i = 1,2,\ldots,n-1$, $\Delta_n = 0$ are the number of trials needed to obtain new records. Let us denote

$$\mathbf{R}_n = (R_1,\ldots,R_n),\ \mathbf{\Delta}_n = (\Delta_1,\ldots,\Delta_n), \mathbf{C}_n = (R_{[1]},\ldots,R_{[n]}).$$

Suppose the observed data is $(r_1,\delta_1,s_1),\ldots,(r_n,\delta_n,s_n)$, then the joint pdf of the first n upper records and inter-record times is (see Arnold *et al.*, 1998, p. 169)

$$f_{(\mathbf{R}_n,\mathbf{\Delta}_n)}(\mathbf{r}_n, \boldsymbol{\delta}_n;\theta) = \prod_{i=1}^{n} f_X(r_i;\theta)\{F_X(r_i;\theta)\}^{\delta_i}. \tag{1}$$

Using (1) the joint pdf of records, inter-record times and their concomitants is given by

$$f_{(\mathbf{R}_n,\mathbf{\Delta}_n,\mathbf{C}_n)}(\mathbf{r}_n, \boldsymbol{\delta}_n, \mathbf{s}_n;\theta) = \prod_{i=1}^{n} f_{X,Y}(r_i,s_i;\theta)\{F_X(r_i;\theta)\}^{\delta_i}. \tag{2}$$

So, the conditional probability mass function of $\mathbf{\Delta}_n$ given $(\mathbf{R}_n,\mathbf{C}_n)$ is given by

$$f_{(\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n)}(\boldsymbol{\delta}_n|\mathbf{r}_n,\mathbf{s}_n;\theta) = \prod_{i=1}^{n-1}[F_X(r_i;\theta)]^{\delta_i}\bar{F}_X(r_i;\theta). \tag{3}$$

In order to perform a comparison study, first let us consider some classes of univariate and bivariate distributions as follows:

$$\mathscr{F} = \{f_{X,Y} : f_{Y|X} \text{ is free of } \theta\},$$

$$\mathscr{B} = \{f_{X,Y} : f_{X,Y}(x,y;\theta) = a(\theta)b(x,y)\exp\{c(\theta)d(x,y)\}, a(\theta) > 0, b(x,y) > 0\},$$

$$\mathscr{K} = \{f_{X,Y} : f_{Y|X}(y|x) \text{ is in the form of } f_{X,Y} \text{ in } \mathscr{B}\},$$

$$\mathscr{C}_1 = \{F_X : \bar{F}_X(x;\theta) = (\bar{G}(x))^{\alpha(\theta)}\},$$

$$\mathscr{C}_2 = \{F_X : F_X(x;\theta) = (H(x))^{\beta(\theta)}\},$$

$$\mathscr{D}_i = \{f_{X,Y} \in \mathscr{F} : F_X \in \mathscr{C}_i\},\ i = 1,2,$$

$$\mathscr{E}_i = \{f_{X,Y} \in \mathscr{K} : F_X \in \mathscr{C}_i, \text{ with } c(\theta) = \alpha(\theta)I_1(i) + \beta(\theta)I_2(i)\},\ i = 1,2,$$

$$\mathscr{G} = \{f_{X,Y} \in \mathscr{F} : f_X \in \mathscr{E}\},$$

$$\mathscr{H} = \{f_{X,Y} \in \mathscr{K} : f_X \in \mathscr{E}\},$$

$$\mathscr{L}_{\mathscr{B}} = \{f_{X,Y} \in \mathscr{B} :\ F_X(x;\theta) = F_0(x-\theta),\ \theta \in \mathbb{R} \text{ or } F_X(x;\theta) = F_1(\theta x),\ \theta > 0\},$$

$$\mathscr{S}_{\mathscr{B}} = \{f_{X,Y} \in \mathscr{B} :\ F_X(x;\theta) = F_1(x^{\theta}),\ \theta > 0,\ x > 0\},$$

$$\mathscr{L}_{\mathscr{K}} = \{f_{X,Y} \in \mathscr{K} :\ F_X(x;\theta) = F_0(x-\theta),\ \theta \in \mathbb{R} \text{ or } F_X(x;\theta) = F_1(\theta x),\ \theta > 0\}$$

and

$$\mathscr{S}_{\mathscr{K}} = \{f_{X,Y} \in \mathscr{K} :\ F_X(x;\theta) = F_1(x^{\theta}),\ \theta > 0,\ x > 0\},$$

where $\alpha(\theta)$ and $\beta(\theta)$ are real positive functions, $G(x)$ and $H(x)$ are arbitrary continuous cdf's, free of θ, $\bar{G}(x) = 1 - G(x)$, $\mathscr{E}$ in $\mathscr{G}$ and $\mathscr{H}$ stands for the well-known exponential family, F_0 and F_1 are arbitrary cdf's, free of θ ($F_i(t) = F_X(t;i),\ i = 0,1$) and $\bar{F}_i(x) = 1 - F_i(x),\ i = 0,1$. Let $h_i(x)$ and $\tilde{h}_i(x),\ i = 0,1$ stand for the standard hazard rate and the reversed hazard rate functions of a random variable with pdf f_i and cdf $F_i,\ i = 0,1$, respectively.

Indeed, $\mathscr{C}_1$ and $\mathscr{C}_2$ stand for two well-known families of distributions in life-time experiments literature, the proportional hazard model and proportional reversed hazard model, respectively (see for example Lawless, 2003). Classes $\mathscr{B}$, $\mathscr{D}_1$ and $\mathscr{D}_2$ include several well-known distributions (see Amini and Ahmadi, 2008). We should emphasize that, although in the two classes $\mathscr{D}_1$ and $\mathscr{D}_2$, $f_{Y|X}$ is free of θ, f_Y may depend on it. In fact by considering a single (X,Y), one would find X a sufficient statistic for θ. Since $\mathscr{C}_1$ and $\mathscr{C}_2$ are both subsets of $\mathscr{E}$, $\mathscr{D}_1$ and $\mathscr{D}_2$ are both subsets of $\mathscr{G}$.

It is clear that $\mathscr{L}_{\mathscr{B}} \subset \mathscr{B}$, $\mathscr{S}_{\mathscr{B}} \subset \mathscr{B}$ and $\mathscr{D}_i \subset \mathscr{G} \subset \mathscr{H} \subset \mathscr{B}, i = 1,2$. Note that in the functional form of $\mathscr{B}$, one may let $d(x,y,\eta) = 0$ and $a(\theta,\eta) = 1$ to obtain a form of $f_{Y|X}(y|x)$ that is free of θ.

We shall note that, although we have used bivariate upper records and times to obtain the results of this paper, corresponding results for bivariate lower records are derived and classified in Table 8.

The hazard rate function and the reversed hazard rate functions are important characteristics for the analysis of reliability data. A random variable X is said to be Increasing Hazard Rate (Reversed Hazard Rate), Decreasing Hazard Rate (Decreasing Reversed Hazard Rate) or Constant Hazard Rate (Constant Reversed Hazard Rate), and is denoted by IHR (IRHR), DHR (DRHR) or CHR (CRHR), if its hazard rate (reversed hazard rate) function is increasing, decreasing or constant, respectively.

3 Main results

Since reparameterizing $\theta = z(\gamma)$, for a differentiable $z(.)$, transforms the Fisher information of any data to $(\frac{\partial}{\partial\gamma} z(\gamma))^2 I_X(z(\gamma))$ (see Lehmann, 1989), we may assume throughout that $c(\theta) = \theta$. To prove the main results of this paper, we need the following lemma. The proof is easy and hence is omitted.

Lemma 1 *The pdf $f_{X,Y}(x,y;\theta)$ belongs to $\mathscr{B}$ with natural parameter θ ($c(\theta) = \theta$) if and only if $\frac{\partial^2}{\partial\theta^2} \log f_{X,Y}(x,y;\theta)$ does not depend on x and y.*

Note: Obviously, we have

$$I_{\mathbf{R}_n,\mathbf{\Delta}_n,\mathbf{C}_n}(\theta) = I_{\mathbf{R}_n,\mathbf{C}_n}(\theta) + I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta), \tag{4}$$

where $I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta) = I_{\mathbf{\Delta}_n|\mathbf{R}_n}(\theta)$ is indeed $\mathrm{E}_{\mathbf{R}_n}(I_{\mathbf{\Delta}_n|\mathbf{R}_n}(\theta))$. Hereafter, we will use the notation $I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta)$ instead of $\mathrm{E}_{\mathbf{R}_n}(I_{\mathbf{\Delta}_n|\mathbf{R}_n}(\theta))$.

Proposition 1 *Let $\{(X_i,Y_i), i \geq 1\}$ be a sequence of iid bivariate random variables with pdf $f_{X,Y}(x,y;\theta)$, then $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) \geq I_{\mathbf{R}_n,\mathbf{C}_n}(\theta)$, with equality while F_X is free of θ and the increment of Fisher information by considering inter-record times is given by*

$$I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta) = -\sum_{i=1}^{n-1} \mathrm{E}\left[\frac{F_X(R_i;\theta)}{\bar{F}_X(R_i;\theta)} \frac{\partial^2}{\partial\theta^2} \log F_X(R_i;\theta) + \frac{\partial^2}{\partial\theta^2} \log \bar{F}_X(R_i;\theta)\right].$$

So $I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta) = 0$ when F_X is free of θ.

Proof From (4), we conclude that $I_{\mathbf{R}_n,\mathbf{C}_n,\boldsymbol{\Delta}_n}(\theta) \geq I_{\mathbf{R}_n,\mathbf{C}_n}(\theta)$. Using (3) and the fact that $\mathrm{E}(\delta_i|R_i) = F_X(R_i;\theta)/\bar{F}_X(R_i;\theta)$ along with the definition of Fisher information, the proof is complete. ■

The univariate case of Proposition 1 is obtained by Hofmann (2004).

***Example 1** **(Farlie-Gumbel-Morgenstern family of distributions)** Let*

$$f_{X,Y}(x,y;\theta) = f_X(x)f_Y(y)[1+\theta(1-2F_X(x))(1-2F_Y(y))], -1 < \theta < 1.$$

Amini and Ahmadi (2007) showed that for this family $I_{\mathbf{R}_n,\mathbf{C}_n}(\theta) > nI_{(X,Y)}(\theta)$. However, since F_X is free of θ, Proposition 1 yields that $I_{\boldsymbol{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta) = 0$. So $I_{\mathbf{R}_n,\mathbf{C}_n,\boldsymbol{\Delta}_n}(\theta) > nI_{(X,Y)}(\theta)$.

Theorem 1 *Suppose $f_{X,Y}(x,y;\theta)$ belongs to $\mathscr{K}$ and let $l(x;\theta) = \dfrac{\partial^2}{\partial\theta^2}\log f_X(x;\theta)$. Then*

(i) if $l(x;\theta)$ is decreasing in x and $F_X(x;\theta)$ is strictly log-concave or log-linear in θ, then $I_{\mathbf{R}_n,\mathbf{C}_n,\boldsymbol{\Delta}_n}(\theta) > nI_{(X,Y)}(\theta)$;

(ii) if $l(x;\theta)$ is increasing in x and $F_X(x;\theta)$ is strictly log-convex or log-linear in θ, then $I_{\mathbf{R}_n,\mathbf{C}_n,\boldsymbol{\Delta}_n}(\theta) < nI_{(X,Y)}(\theta)$.

Proof

(i) Equation (2) yields

$$\frac{\partial^2}{\partial\theta^2}\log f_{(\mathbf{R}_n,\boldsymbol{\Delta}_n,\mathbf{C}_n)}(\mathbf{r}_n,\boldsymbol{\delta}_n,\mathbf{s}_n;\theta) = \sum_{i=1}^{n}\frac{\partial^2}{\partial\theta^2}\log f_{X,Y}(r_i,s_i;\theta) + \sum_{i=1}^{n-1}\delta_i\frac{\partial^2}{\partial\theta^2}\log F_X(r_i;\theta). \tag{5}$$

The second term of the right hand side of (5) is non-positive by assumption. On the other hand

$$\frac{\partial^2}{\partial\theta^2}\log f_{X,Y}(r_i,s_i;\theta) = \frac{\partial^2}{\partial\theta^2}\log f_X(r_i;\theta) + \frac{\partial^2}{\partial\theta^2}\log f_{Y|X}(s_i|r_i;\theta). \tag{6}$$

By assumptions and in the view of Lemma 1 the second term on the right hand side of (6) does not depend on r_i and s_i. Noting that record values are stochastically ordered, i.e., $R_i <_{st} R_{i+1}$, we have $\mathrm{E}(l(R_i;\theta)) > \mathrm{E}(l(R_{i+1};\theta))$ for each i, since $l(x;\theta)$ is decreasing in x. Thus the proof is complete using the definition of Fisher information.

(ii) The proof is similar to that of part (i). ■

Remark 1 *One can easily see that*

$$\frac{\partial^2}{\partial\theta^2}\log F_X(x;\theta) = \frac{L(x;\theta)}{F_X^2(x;\theta)},$$

where $L(x;\theta) = F_X(x;\theta)\partial^2/\partial\theta^2 F_X(x;\theta) - (\partial/\partial\theta F_X(x;\theta))^2$*. So* $F_X(x;\theta)$ *is strictly log-concave, log-linear or strictly log-convex if and only if* $L(x;\theta)$ *is negative, zero or positive. This approach is used in the next illustrative examples.*

Remark 2 *For the case of lower records, their concomitants and inter-record times the result of the Theorem 1 holds by considering* $\bar{F}_X$ *instead of* F_X *and replacing increasing by decreasing and vice versa.*

Example 2 ***Bivariate normal with a known correlation*** r ***and*** $\mu_{\mathbf{X}} = r^{-\mathbf{1}}\mu_{\mathbf{Y}} = \sigma_{\mathbf{X}} = \sigma_{\mathbf{Y}} = \theta$**.** *This family does not belong to class* $\mathscr{B}$*. However, the distribution of* Y *given* $X = x$ *is normal with mean* rx *and variance* $\theta^2(1-r^2)$ *which is a member of* $\mathscr{B}$*. So, this family is a member of* $\mathscr{K}$*. Taking* $\alpha = \theta^{-1}$*,* $l(x;\alpha) = -\alpha^{-4}/2 - \alpha^{-3}x/4$ *which is decreasing in* x*. Also*

$$L(x;\alpha) = \frac{1}{2\pi}e^{-(1/2)(\alpha x-1)^2}\left[(\alpha x - 1)\int_{\alpha x-1}^{\infty} e^{-(1/2)u^2}du - e^{-(1/2)(\alpha x-1)^2}\right],$$

it can be shown that the expression in the bracket on the right hand side of the above equation is negative (see Ahmadi and Arghami, 2001). Hence $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) > nI_{(X,Y)}(\theta)$ *by Theorem 1.*

Theorem 2 *Let* $f_{X,Y}(x,y;\theta)$ *belong to* $\mathscr{B}$*. Then* $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta)$ *is less than, equal to or greater than* $nI_{(X,Y)}(\theta)$ *if* $F_X(x;\theta)$ *is strictly log-convex, log-linear and strictly log-concave, respectively in* θ*.*

Proof By Lemma 1, the first term on the right hand side of (5) does not depend on r_i's and s_i's, and it's expected value equals $-nI_{(X,Y)}(\theta)$. This completes the proof. ∎

Some illustrative examples of Theorem 2 are bivariate normal with a known correlation r, $\mu_{\mathbf{X}} = \theta$ $\mu_{\mathbf{Y}} = \mathbf{0}$ and $\sigma_{\mathbf{X}} = \sigma_{\mathbf{Y}} = \mathbf{1}$, Arnold and Strauss's bivariate exponential (Arnold and Strauss, 1988 [See also Amini and Ahmadi, 2008]), Mardia's bivariate Pareto distribution with the joint pdf

$$f_{X,Y}(x,y;\theta) = \theta(\theta+1)(1+x+y)^{-(\theta+2)},\ x,\ y,\ \theta > 0,$$

McKay's bivariate gamma distribution and Bilateral bivariate Pareto distribution. The results of these examples are summarized in Table 8 and the last two are presented below.

Example 3 McKay's bivariate gamma distribution *(McKay, 1934). Suppose (X,Y) has the joint pdf*

$$f_{X,Y}(x,y;\theta) = \frac{\theta^{a+b}}{\Gamma(a)\Gamma(b)} x^{a-1}(y-x)^{b-1}e^{-\theta y},\ y > x > 0,\ \theta > 0, \tag{7}$$

where a and b are known positive real numbers and $\Gamma(.)$ is the well-known gamma function.

This family is a member of $\mathscr{B}$ and the marginal distribution of X is gamma with parameters a and θ. Hence $L(x;\theta) = \frac{\theta^{2a-2}x^a}{\Gamma(a)^2}e^{-\theta x}\left\{(a-1-\theta x)\int_0^x y^{a-1}\exp(-\theta y)\,dy - x^a e^{-\theta x}\right\}$, which is negative (see Ahmadi and Arghami, 2003). Therefore, applying Theorem 2, $\boldsymbol{I}_{\mathbf{R}_n,\mathbf{C}_n,\boldsymbol{\Delta}_n}(\theta) > n\boldsymbol{I}_{(X,Y)}(\theta)$.

Example 4 Bilateral bivariate Pareto distribution. *This family has the joint pdf (for example, see De Groot, 1970)*

$$f_{X,Y}(x,y;\theta) = \theta(\theta+1)(a-b)^{\theta}(y-x)^{-(\theta+2)},\ x < b < a < y,\ \theta > 1, \tag{8}$$

where the two quantities a and b are known positive real numbers.

This is again a member of $\mathscr{B}$, and the marginal pdf of X is given by $f_X(x;\theta) = \frac{\theta(a-b)^{\theta}}{(a-x)^{\theta+1}}$. We obtain $L(x;\theta) = 0$. Hence applying Theorem 2, $\boldsymbol{I}_{\mathbf{R}_n,\mathbf{C}_n,\boldsymbol{\Delta}_n}(\theta) = n\boldsymbol{I}_{(X,Y)}(\theta)$.

Theorem 3 ***(Location or scale marginal in $\mathscr{B}$)*** *Let $f_{X,Y}(x,y;\theta,\eta)$ belong to $\mathscr{L}_{\mathscr{B}}$, then:*

(i) $I_{\mathbf{R}_n,\mathbf{C}_n,\boldsymbol{\Delta}_n}(\theta)$ is less than, equal to or greater than $nI_{(X,Y)}(\theta)$ if X, is IRHR, CRHR or DRHR, respectively;

(ii) the increment of Fisher information by considering inter-record times is equal to

$$I_{\boldsymbol{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta) = \sum_{i=1}^{n-1} E\left[\frac{h_0^2(R_i-\theta)}{F_0(R_i-\theta)}\right], \tag{9}$$

for a location marginal and equals

$$I_{\boldsymbol{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta) = \sum_{i=1}^{n-1} E\left[\frac{R_i^2 h_1^2(\theta R_i)}{F_1(\theta R_i)}\right], \tag{10}$$

for the scale marginal.

Proof

(i) One can easily show that for both location and scale families, $\frac{\partial^2}{\partial\theta^2}\log F_X(x;\theta)$ and $\frac{\partial^2}{\partial x^2}\log F_X(x;\theta)$ have the same sign, that is, convexity, linearity and concavity of

$\log F_X(x;\theta)$, in x is similar to that in θ. On the other hand, $F_X(x;\theta)$ is strictly log-convex, log-linear or strictly log-concave in x if and only if the reversed hazard rate function, $\tilde{h}_X(x;\theta)$, is increasing, constant or decreasing in x, respectively. So the results of part(i) follow from Theorem 2 and Remark 2.

(ii) Use Proposition 1. Note that for location and scale families, $\frac{\partial^2}{\partial\theta^2}\log F_X(x;\theta)$ is equal to $\frac{\partial^2}{\partial x^2}\log F_X(x;\theta)$ and $x^2\frac{\partial^2}{\partial x^2}\log F_X(x;\theta)$, respectively. Also $\frac{\partial^2}{\partial x^2}\log F_X(x;\theta)$ equals $\frac{\partial}{\partial x}\log \tilde{h}_X(x;\theta)$ and $\frac{\partial^2}{\partial x^2}\log \bar{F}_X(x;\theta)$ equals $\frac{\partial}{\partial x}\log h_X(x;\theta)$. So

$$I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta) = \sum_{i=1}^{n-1} \mathrm{E}\left(h_0'(R_i-\theta) - \frac{\tilde{h}_0'(R_i-\theta)F_0(R_i-\theta)}{\bar{F}_0(R_i-\theta)}\right)$$
$$= \sum_{i=1}^{n-1} E\left[\frac{h_0^2(R_i-\theta)}{F_0(R_i-\theta)}\right],$$

for a location marginal and

$$I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta) = \sum_{i=1}^{n-1} E\left[R_i^2\left(h_1'(\theta R_i) - \frac{\tilde{h}_1'(\theta R_i)F_1(\theta R_i)}{\bar{F}_1(\theta R_i)}\right)\right]$$
$$= \sum_{i=1}^{n-1} E\left[\frac{R_i^2 h_1^2(\theta R_i)}{F_1(\theta R_i)}\right],$$

for the scale marginal. ■

Example 5 Bivariate normal with known correlation *r, $\mu_{\mathbf{X}} = \theta$ $\mu_{\mathbf{Y}} = \mathbf{0}$ **and** $\sigma_{\mathbf{X}} = \sigma_{\mathbf{Y}} = \mathbf{1}$. The considered bivariate normal family belongs to $\mathscr{L}_{\mathscr{B}}$, and the normal distribution is DRHR. Hence Theorem 3 also yields that $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) > nI_{(X,Y)}(\theta)$. Table 1 shows the values of $I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta)$ for $n = 2,3,5,7,10$ of the normal distribution.*

Example 6 *(Continuation of Example 3) This family belongs to $\mathscr{L}_{\mathscr{B}}$ and the distribution of X is DRHR. Therefore, applying Theorem 3, $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) > nI_{(X,Y)}(\theta)$. Table 2 shows the values of $\theta^2 I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta)$ for different values of n and a. As can be seen, these values increase as the shape parameter a increases.*

Example 7 Bivariate gamma exponential distribution (i). *Suppose that*

$$f_{X,Y}(x,y;\theta) = \theta dx \exp\{-(\theta x + dxy)\} \quad x>0,\ y>0,\ \theta>0, \tag{11}$$

where d is a known positive real number. This family belongs to $\mathscr{L}_{\mathscr{B}}$, and the exponential distribution is DRHR. Therefore Theorem 3 yields that $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) > nI_{(X,Y)}(\theta)$. The values of $\theta^2 I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta)$ in Table 2 with $a = 1$ are the corresponding Fisher information for the exponential distribution.

Table 1: *The values of* $I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(0)$ *for* $n = 2,3,5,7,10$ *from standard normal distribution.*

n	2	3	5	7	10
$I_{\mathbf{\Delta}_n\|\mathbf{R}_n,\mathbf{C}_n}(0)$	1.6718	4.7961	15.7557	33.5634	73.9717

Corollary 1 ***(Shape marginal in*** $\mathscr{B}$***)*** *Let* $f_{X,Y}(x,y;\theta)$ *belong to* $\mathscr{S}_{\mathscr{B}}$*. Then* $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta)$ *is less than, equal to or greater than* $nI_{(X,Y)}(\theta)$*, if* $k(x) = x\tilde{h}_1(x)$ *is increasing, constant or decreasing in x.*

Proof As in the proof of Theorem 2, and since we have in shape family $F_X(x;\theta) = F_1(x^\theta)$, taking $\gamma = x^\theta$ it follows that

$$\frac{\partial^2}{\partial\theta^2}\log F_1(x^\theta) = x^\theta(\log x)^2[\frac{\partial}{\partial\gamma}\gamma\tilde{h}_1(\gamma)].$$

Since $x > 0$, this gives us the result. ■

Example 8 Sub-class of $\mathscr{H}$ ***with power distribution marginal.*** *In order to illustrate the result of Corollary 1, a sub-class of* $\mathscr{S}$ *with* $F_X(x) = x^\theta$, $x > 0$, $\theta > 0$ *is concerned. Hence,* $f_{Y|X}(y|x)$ *must have the functional form of* $\mathscr{B}$*. So this class is also a sub-class of* $\mathscr{H}$ *with power distribution marginal. For power distribution,* $k(x) = x\tilde{h}_1(x) = 1$, $x > 0$, *which is constant. So* $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) = nI_{(X,Y)}(\theta)$ *by Corollary 1. An example of such bivariate distributions can be*

$$f_{X,Y}(x,y;\theta) = \theta^2 x^{-1}\exp\{\theta(\log x + x - y)\},\ 0 < x < y,\ \theta > 0.$$

Corollary 2 ***(Location or scale marginal in*** $\mathscr{K}$***)*** *Let* $f_{X,Y}(x,y;\theta)$ *belong to* $\mathscr{L}_{\mathscr{K}}$*. Then* $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta)$ *is less (greater) than* $nI_{(X,Y)}(\theta)$ *if X is IRHR (DRHR), or CRHR and* $l(x;\theta)$ *is increasing (decreasing) in x.*

Proof The proof is similar to that of Theorems 1 and 3. ■

Table 2: *The values of* $\theta^2 I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta)$ *for* $n = 3(2)7, 10$ *and* $a = 0.5, 1, 2$ *of gamma distribution.*

		a	
n	0.5	1	2
3	5.2036	8.8980	15.4526
5	27.0356	41.6880	66.4591
7	78.9683	114.1098	172.0214
10	245.0912	332.3383	473.1286

Example 9 *(Continuation of Example 2) The considered bivariate normal family belongs to $\mathscr{L}_{\mathscr{K}}$ with respect to parameter α, $l(x;\theta)$ is decreasing in x and the normal distribution is DRHR. Hence $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) > nI_{(X,Y)}(\theta)$ by Corollary 2.*

Corollary 3 ***(Shape marginal in $\mathscr{K}$)*** *Let $f_{X,Y}(x,y;\theta,\eta)$ belong to $\mathscr{S}_{\mathscr{K}}$. Then $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta)$ is less (greater) than $nI_{(X,Y)}(\theta)$, if $k(x) = x\tilde{h}_1(x)$ is increasing (decreasing) or constant and $l(x;\theta)$ is increasing (decreasing) in x.*

Proof The proof is similar to Theorem 1 and Corollary 1. ∎

Corollary 4 *Let $\{(X_i,Y_i),\ i \geq 1\}$ be distributed as the family*

$$\{f_{X,Y}(x,y;\theta) \in \mathscr{B}; F_X \text{ is free of } \theta\},$$

then $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) = nI_{(X,Y)}(\theta)$.

Proof It is deduced from Theorem 2, since $L(x;\theta) = 0$. ∎

Example 10 Bivariate gamma exponential distribution (ii). *Consider the joint pdf*

$$f_{X,Y}(x,y;\theta) = \frac{a^b\theta}{\Gamma(b)} x^b \exp\{-(ax+\theta xy)\}\ \ x>0,\ y>0,\ \theta>0, \tag{12}$$

where a and b are known positive real numbers. This family is a member of $\mathscr{B}$ and F_X is free of θ. Therefore by Corollary 4, $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) = nI_{(X,Y)}(\theta)$.

Remark 3 *For the case of lower records, their concomitants and inter-record times the results of Theorem 3 and Corollaries 1, 2 and 3 are reversed. For example in Corollary 2 $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta)$ is less than (greater than) $nI_{(X,Y)}(\theta)$, if X is DHR (IHR) or CHR and $l(x;\theta)$ is decreasing (increasing) in x. Note that in this case, we consider the standard hazard rate function in location and scale families, i.e., $h_i(x)$, $i = 0,1$.*

Theorem 4 *Let $f_{X,Y}(x,y;\theta)$ belong to $\mathscr{E}_1$, then:*

(i) $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) > nI_{(X,Y)}(\theta)$;

(ii) the increment of Fisher information by considering inter-record times is equal to

$$I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta) = \left(\frac{\alpha'(\theta)}{\alpha(\theta)}\right)^2 \sum_{i=1}^{n-1} i(i+1)\xi(i+2), \tag{13}$$

where $\xi(.)$ is the Riemann Zeta function.

Table 3: *The values of* $(\alpha(\theta)/\alpha'(\theta))^2 I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\alpha)$ *for* $n = 2(3)14$ *in* (13).

n	2	5	8	11	14
$(\alpha(\theta)/\alpha'(\theta))^2 I_{\mathbf{\Delta}_n\vert\mathbf{R}_n,\mathbf{C}_n}(\alpha)$	2.404	41.688	170.222	442.365	912.397

Proof

(i) The class $\mathscr{E}_1$ is a subclass of $\mathscr{B}$. Assuming $\alpha(\theta) = \alpha$, we have

$$L(x;\alpha) = -(\log \bar{G}(x))^2 \bar{G}(x)^{\alpha},$$

which is clearly negative. Hence, the result follows from Theorem 2.

(ii) Using Proposition 1, we have

$$\begin{aligned} I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\alpha) &= \sum_{i=1}^{n-1} \mathrm{E}\left(\frac{(\log \bar{G}(R_i))^2}{1-\bar{G}(R_i)^{\alpha}} \right) \\ &= \alpha^{-2} \sum_{i=1}^{n-1} \mathrm{E}\left(\frac{(\log \bar{F}(R_i))^2}{1-\bar{F}(R_i)} \right) \\ &= \alpha^{-2} \sum_{i=1}^{n-1} \frac{1}{(i-1)!} \int_0^1 \frac{(-\log v)^{i+1}}{1-v}\, dv. \end{aligned}$$

Expanding the term $1/(1-v)$, we get

$$\begin{aligned} I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\alpha) &= \alpha^{-2} \sum_{i=1}^{n-1} \frac{1}{(i-1)!} \sum_{j=1}^{\infty} \int_0^1 v^{j-1}(-\log v)^{i+1}\, dv \\ &= \alpha^{-2} \sum_{i=1}^{n-1} i(i+1)\xi(i+2). \end{aligned}$$

■

Table 3 shows the values of $(\alpha(\theta)/\alpha'(\theta))^2 I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\alpha)$ for $n = 2(3)14$ in class $\mathscr{E}_1$.

Example 11 *(Continuation of Example 7) The distribution of X is exponential with parameter θ, which belongs to $\mathscr{C}_1$. Also the conditional distribution of Y given $X = x$ is free of θ. Hence, this family is a member of $\mathscr{E}_1$. Therefore Corollary 4 yields that* $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) > nI_{(X,Y)}(\theta)$.

Theorem 5 *Let $f_{X,Y}(x,y;\theta,\eta)$ belong to $\mathscr{E}_2$, then:*

(i) $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) = nI_{(X,Y)}(\theta)$;

(ii) the increment of Fisher information by considering inter-record times is equal to

$$I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\theta) = \left(\frac{\beta'(\theta)}{\beta(\theta)} \right)^2 \sum_{i=1}^{n-1} \varphi(i), \tag{14}$$

where

$$\varphi(i) = \sum_{r=1}^{\infty}\sum_{s=1}^{\infty}\frac{1}{rs}\left[\frac{1}{(r+s-1)^i} - \frac{1}{(r+s)^i}\right]. \tag{15}$$

Proof

(i) The class $\mathscr{E}_2$ is a subclass of $\mathscr{B}$ with $c(\theta) = \beta(\theta)$. We may assume without loss of generality that $\beta(\theta) = \beta$. The result follows from Theorem 2, since $L(x;\beta) = 0$.

(ii) Using Proposition 1, we have

$$\begin{aligned} I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\beta) &= \sum_{i=1}^{n-1}\mathrm{E}\left(\frac{H(R_i)^{\beta}(\log H(R_i))^2}{(1-H(R_i)^{\beta})^2}\right) \\ &= \beta^{-2}\sum_{i=1}^{n-1}\mathrm{E}\left(\frac{F(R_i)(\log F(R_i))^2}{(1-\bar{F}(R_i))^2}\right) \\ &= \beta^{-2}\sum_{i=1}^{n-1}\frac{1}{(i-1)!}\int_0^1\frac{v}{(1-v)^2}(\log v)^2(-\log(1-v))^{i-1}\,dv. \end{aligned}$$

Expanding $\log(v)$ we have

$$\begin{aligned} I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\beta) &= \beta^{-2}\sum_{i=1}^{n-1}\frac{1}{(i-1)!}\sum_{r=1}^{\infty}\sum_{s=1}^{\infty}\frac{1}{rs}\int_0^1 v(1-v)^{r+s-2}(-\log(1-v))^{i-1}\,dv \\ &= \beta^{-2}\sum_{i=1}^{n-1}\sum_{r=1}^{\infty}\sum_{s=1}^{\infty}\frac{1}{rs}\left[\frac{1}{(r+s-1)^i} - \frac{1}{(r+s)^i}\right]. \end{aligned}$$

■

Table 4 shows the values of $\varphi(i)$ for $i = 1(1)7$, which are calculated to 4 decimal places using the R package. These values tend very quickly to 1 as i increases, such that they are approximately equal to one, for $i \geq 7$. Hence using these values is a proper approach to calculate $I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\beta)$. Table 5 shows these values for $n = 2(3)14$ in class $\mathscr{E}_2$.

Table 4: *The values of* $\varphi(i)$ *in* (15) *for* $i = 1(1)7$.

i	1	2	3	4	5	6	7
$\varphi(i)$	0.8857	0.9772	0.9943	0.9984	0.9995	0.9999	1.0000

Table 5: *The values of* $(\beta(\theta)/\beta'(\theta))^2 I_{\mathbf{\Delta}_n|\mathbf{R}_n,\mathbf{C}_n}(\beta)$ *in* (14) *for* $n = 2(3)14$.

n	2	5	8	11	14
$(\beta(\theta)/\beta'(\theta))^2 I_{\mathbf{\Delta}_n\|\mathbf{R}_n,\mathbf{C}_n}(\beta)$	0.8857	3.8608	6.8602	9.8602	12.8602

Example 12 *(Continuation of Example 4). We have* $F_X(x;\theta) = \left(\frac{a-b}{a-x}\right)^{\theta}$*. Hence* F_X *belongs to* $\mathscr{C}_2$ *and therefore* $f_{X,Y}(x,y;\theta) \in \mathscr{E}_2$*. Thus, using Theorem 5,* $I_{\mathbf{R}_n,\mathbf{C}_n,\boldsymbol{\Delta}_n}(\theta) = nI_{(X,Y)}(\theta)$*.*

Theorem 6 *Let* $f_{X,Y}(x,y;\theta)$ *belong to* $\mathscr{F}$ *or* $\mathscr{K}$*. Then* $I_{\mathbf{R}_n,\mathbf{C}_n,\boldsymbol{\Delta}_n}(\theta)$ *is less than, equal to or greater than* $nI_{(X,Y)}(\theta)$ *if and only if* $I_{\mathbf{R}_n,\boldsymbol{\Delta}_n}(\theta)$ *is less than, equal to or greater than* $nI_X(\theta)$*, respectively.*

Proof From equations (1) and (2)

$$I_{\mathbf{R}_n,\mathbf{C}_n,\boldsymbol{\Delta}_n}(\theta) = I_{\mathbf{R}_n,\boldsymbol{\Delta}_n}(\theta) - \mathrm{E}\left[\sum_{i=1}^{n} \frac{\partial^2}{\partial\theta^2} \log f_{Y|X}(R_{[i]}|R_i;\theta)\right].$$

The expectation above is equal to zero and $n\mathrm{E}\left[\frac{\partial^2}{\partial\theta^2} \log f_{Y|X}(Y|X;\theta)\right]$ in $\mathscr{F}$ and $\mathscr{K}$, respectively. Hence, in both classes

$$I_{\mathbf{R}_n,\mathbf{C}_n,\boldsymbol{\Delta}_n}(\theta) - nI_{(X,Y)}(\theta) = I_{\mathbf{R}_n,\boldsymbol{\Delta}_n}(\theta) - nI_X(\theta). \qquad \blacksquare$$

A result similar to Theorem 6 holds for lower records.

4 Estimation

To illustrate the applications of comparison study of Fisher information, discussed in previous section, we present a simulation study and a real data example.

4.1 A simulation study

In order to compare the performance of estimators based on bivariate records and inter-record times with corresponding estimators based on other types of data, consider a bivariate normal distribution. For simplicity, we may assume that the only unknown parameter in this model is $\theta = \mathrm{E}(X)$, i.e.,

$$f_{X,Y}(x,y;\theta) = \frac{1}{2\pi\sqrt{1-r^2}} \exp\left\{ \frac{\left[(x-\theta)^2 + y^2 - 2r(x-\theta)y\right]}{-2(1-r^2)} \right\}, \tag{16}$$

$$x,\ y \in \mathbb{R},\ \theta \in \mathbb{R}.$$

The likelihood equation for deriving the MLE of θ based on bivariate record values and inter-record times ($\hat{\theta}_{RCT}$, if exists) is as follows:

$$\sum_{i=1}^{n} R_i - n\theta - r\sum_{i=1}^{n} R_{[i]} - (1-r^2)\sum_{i=1}^{n} \delta_i \tilde{h}_0 (R_i - \theta) = 0. \tag{17}$$

In this case, $\hat{\theta}_{RCT}$ has no explicit form and the values of this estimator have to be derived by numerical methods.

Now, the following criteria are interesting:

(a) Relative efficiency (RE) of estimator based on bivariate record values and inter-record times with respect to estimator based on bivariate record values only.

(b) RE of estimator based on bivariate record values and inter-record times with respect to estimator based on an independent bivariate random sample of the same size.

For deriving the RE of case **(a)**, we may consider the likelihood equation for deriving the MLE of θ based on bivariate record values only ($\hat{\theta}_{RC}$, if exists) as follows

$$\sum_{i=1}^{n} R_i - n\theta - r\sum_{i=1}^{n} R_{[i]} - (1-r^2)\sum_{i=1}^{n-1} h_0 (R_i - \theta) = 0. \tag{18}$$

Again, the values of $\hat{\theta}_{RC}$ have to be derived by numerical methods. For deriving the RE of case **(b)**, note that the MLE of θ based on an iid sample of size n from this bivariate family equals

$$\hat{\theta}_{IID} = n^{-1}[\sum_{i=1}^{n} X_i - r\sum_{i=1}^{n} Y_i],$$

which is an unbiased estimator of θ with a variance equal to $(1-r^2)/n$.

Table 6: ***(a)*** *RE*$(\hat{\theta}_{RCT}, \hat{\theta}_{RC})$ *in bivariate normal distribution for different values of r and n.*

	r								
n	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
3	3.052	2.799	2.633	2.431	2.207	1.909	1.494	1.175	1.035
5	6.583	6.088	5.517	4.756	3.707	2.975	2.062	1.372	1.044
7	12.291	11.325	9.506	7.469	5.735	4.137	2.777	1.624	1.057
10	25.743	22.500	18.092	13.622	9.527	6.489	3.821	2.060	2.052

Table 7: ***(b)*** *RE*$(\hat{\theta}_{RCT}, \hat{\theta}_{IID})$ *in bivariate normal distribution for different values of r and n.*

	r								
n	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
3	1.680	1.630	1.594	1.596	1.550	1.625	1.727	1.918	2.767
5	2.568	2.557	2.425	2.420	2.307	2.216	2.172	2.244	2.810
7	3.716	3.605	3.563	3.446	3.240	3.132	2.813	2.694	2.884
10	5.784	5.865	5.573	5.382	4.950	4.439	3.872	3.448	3.142

Tables 6 and 7 show the simulated values of RE of $\hat{\theta}_{RCT}$ based on the first n bivariate upper records and inter-record times relative to $\hat{\theta}_{RC}$ and $\hat{\theta}_{IID}$, respectively, which are derived using 100,000 iterations generated by the R package. The minimum number of iterations is used to derive the root of equations (17) and (18) to 3 decimal places. Also the default method of finding the roots of equations in the R package is considered. As one can see in Figure 1, MSE($\hat{\theta}_{RCT}$) decreases as n or r increases. The simulated values showed that MSE($\hat{\theta}_{RCT}$) has similar values for positive and negative values of r. Also, since θ is a location parameter, the values of MSE($\hat{\theta}_{RCT}$) does not depend on θ. The values of RE($\hat{\theta}_{RCT}, \hat{\theta}_{RC}$) and RE($\hat{\theta}_{RCT}, \hat{\theta}_{IID}$) increase as n increases. The values of Table 7 seem to have a minimum point when r increases and the value of r for which RE($\hat{\theta}_{RCT}, \hat{\theta}_{IID}$) is minimum, tends to 1 by increasing n.

These values show that, in this example, the estimator of $\theta = \mathrm{E}(X)$ based on bivariate record values and inter-record times is more efficient than the corresponding estimator based on bivariate record values only and the estimator based on an iid bivariate sample of the same size. The result of Fisher information comparison for the parameter $\theta = \mathrm{E}(X)$ in this model and the fact that considering inter-record times causes an increment of Fisher information, uphold these estimation results.

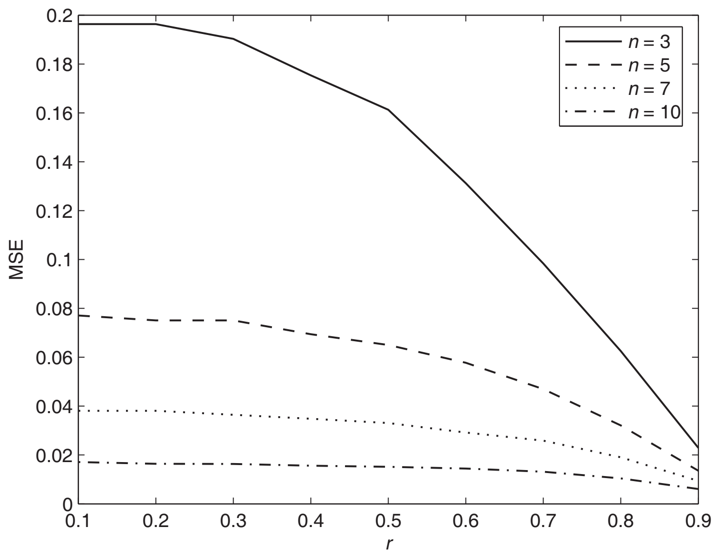


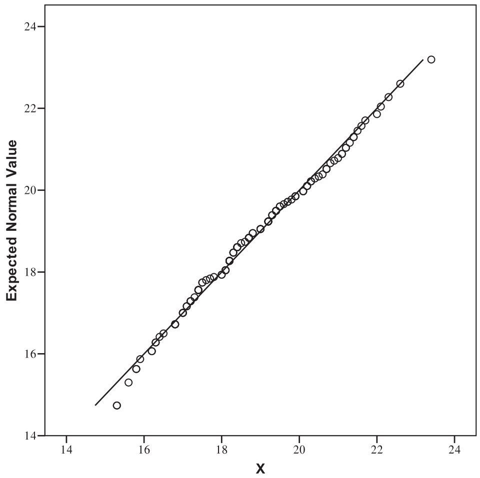
Figure 1: *MSE($\hat{\theta}_{RCT}$) in bivariate normal distribution for different values of r and n.*

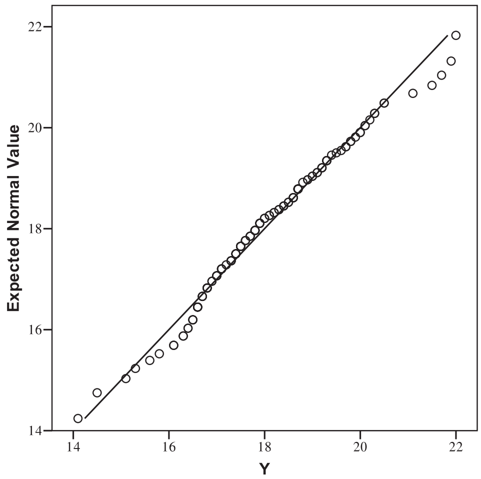
4.2 A real data example

As a real data example, we have considered 130 observations of temperatures at Neuenburg, Switzerland, on July (X) and August (Y), during 1864-1993 (Arnold *et al.*, 1998, p. 278). For these data, bivariate record values and inter-record times are given as follows:

Year	1864	1865	1869	1870	1881	1904	1911	1928	1983
i	1	2	3	4	5	6	7	8	9
Records (July), R_i	19.0	20.1	21.0	21.4	21.7	22.0	22.1	22.6	23.4
Concomitants (August), $R_{[i]}$	17.3	16.7	17.5	16.1	18.5	19.5	21.7	20.1	19.6
Inter-record times, Δ_i	0	3	0	10	22	6	16	54	0

In order to check the normality of the marginal distributions of X and Y, the corresponding Q-Q plots are drawn as follows.





The values of the Mardia test statistics (Mardia, 1974) are obtained as $V_1^* = 8.36 \times 10^{-141}$ and $V_2^* = -0.289$. Since the null hypothesis is rejected for large values of V_1^* and $|V_2^*|$, this indicates that the bivariate normal model provides a good fit to the above data.

Maximum likelihood estimates of the parameters, based on bivariate record values and also based on bivariate record values and inter-record times, are obtained by solving likelihood equations of bivariate normal distribution numerically as follows:

Parameter (θ)	μ_1	μ_2	σ_1^2	σ_2^2	ρ
Bivariate records	20.35	17.36	0.89	2.67	0.60
Bivariate records and times	20.12	17.21	1.32	2.82	0.63
Complete sample ($n = 130$)	18.79	18.04	2.89	2.15	0.31
$I_{\mathbf{\Delta}_n\mid\mathbf{R}_n,\mathbf{C}_n}(\hat{\theta})$	58.64	0	168.06	0	0

The complete sample estimators and the values of $I_{\mathbf{\Delta}_n\mid\mathbf{R}_n,\mathbf{C}_n}(\theta)$ (estimated values if unknown) are also given. As we can see, larger values of $I_{\mathbf{\Delta}_n\mid\mathbf{R}_n,\mathbf{C}_n}(\theta)$ cause a larger difference of the estimate based on bivariate records and complete sample estimates, with respect to the corresponding difference of the estimate based on bivariate records and times.

5 Concluding remarks

In this paper, we have considered the problem of studying Fisher information in bivariate records in the present of inter-record times. Although, there is no information in record times themselves about the sampling distribution, the joint distribution of records and record times depends on it. We have seen that they provide significant additional information (see Table 8). For various cases an explicit formula for the increment of the Fisher information in the presence of inter-record times have obtained. Some general results have established to compare the amount of Fisher information in bivariate records and inter-record times with a random sample. Several classes of common univariate and bivariate families of distributions have been taken into account and some examples have been given in each cases to explain the results. The results of Section 4 show that the estimator on the basis of bivariate record values including inter-record times is more efficient than the corresponding estimator based on iid sample of the same size and the estimator based on bivariate records only. These results agree with the facts that $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) > nI_{(X,Y)}(\theta)$ and $I_{\mathbf{R}_n,\mathbf{C}_n,\mathbf{\Delta}_n}(\theta) > I_{\mathbf{R}_n,\mathbf{C}_n}(\theta)$ (when F_X depends on θ) for bivariate normal distribution.

Table 8: *Classification of some bivariate distributions based on information properties, by considering their marginal properties.*

Bivariate distribution	URC	URCT	LRC	LRCT	UR	URT	LR	LRT
Bivariate Normal, $\theta = \mathrm{E}(X)$ or $\mathrm{Var}(X)$ $\theta = \mathrm{E}(Y)$ or $\mathrm{Var}(Y)$	 $<$ $=$	 $>$ $=$	 $<$ $=$	 $>$ $=$	 $<$ $=$	 $>$ $=$	 $<$ $=$	 $>$ $=$
McKay's Biv. Gamma (7) $0 < a < 1$ $a = 1$ $a > 1$	 $>$ $=$ $<$	 $>$ $>$ $>$	 $<$ $<$ $<$	 $<$ $=$ $>$	 $>$ $=$ $<$	 $>$ $>$ $>$	 $<$ $<$ $<$	 $<$ $=$ $>$
Biv. Gamma exponential (11) (12)	$=$ $=$	$>$ $=$	$<$ $=$	$=$ $=$	$=$ $=$	$>$ $=$	$<$ $=$	$=$ $=$
Bilateral Biv. Pareto (8)	$<$	$=$	$=$	$>$	$<$	$=$	$=$	$>$
Mardia Biv. Pareto	$=$	$>$	$<$	$=$	$=$	$>$	$<$	$=$
Arnold and Strauss's Bivariate Exponential [7]	 $>$	 $>$	 $<$	 $<$	 $<$	 $>$	 $<$	 $>$
Class $\mathscr{E}_1$	$=$	$>$	$<$	$=$	$=$	$>$	$<$	$=$
Class $\mathscr{E}_2$	$<$	$=$	$=$	$>$	$<$	$=$	$=$	$>$

Finally, some common bivariate distribution are classified in Table 8, according to the introduced criteria. The abbreviations URC (LRC), URCT (LRCT), UR (LR) and URT (LRT) are considered for upper (lower) records with their concomitants, upper (lower) records with their concomitants and inter-record times, upper (lower) records and upper (lower) records and inter-record times, respectively. The symbols "$>$", "$=$" and "$<$" mean that the Fisher information contained in the first n of the aforementioned statistics about θ is greater than, equal to and less than Fisher information contained in a random sample of size n from the parent bivariate distribution (or its X-marginal distribution for record statistics without concomitants). The results of the columns URC, LRC, UR and LR are given by Amini and Ahmadi (2008). The columns URT and LRT are the results of Theorem 6. From Table 8 we observe that there is a marked increase in the Fisher information by including inter-record times.

Acknowledgement

The authors would like to thank two referees for their constructive comments and suggestions which led to a considerable improved on an earlier version of this paper. Partial support from "Ordered and Spatial Data Center of Excellence of Ferdowsi University of Mashhad" is acknowledged.

References

Abo-Eleneen, Z. A. and Nagaraja, H. N. (2002). Fisher information in an order statistic and its concomitant. *Annals of the Institute of Statistical Mathematics*, 54, 667-680.

Ahmadi, J. and Arghami, N. R. (2001). On the Fisher information in record values. *Metrika*, 53, 195-206.

Ahmadi, J. and Arghami, N. R. (2003). Comparing the Fisher information in record values and i.i.d. observations. *Statistics*, 37, 435-441.

Amini, M. and Ahmadi, J. (2007). Fisher information in record values and their concomitants about the dependence and correlation parameters. *Statistics & Probability Letters*, 77, 964-972.

Amini, M. and Ahmadi, J. (2008). Comparing Fisher information in record values and their concomitants with random observations. *Statistics*, 42, 393-405.

Arnold, B. C., Balakrishnan, N. and Nagaraja, H. N. (1998). *Records*. John Wiley & Sons, New York.

Arnold, B. C. and Strauss, D. (1988). Bivariate distributions with exponential conditionals. *Journal of the American Statistical Association*, 83, 552-527.

Balakrishnan, N. and Stepanov, A. (2005). On the Fisher information in record data. *Statistics & Probability Letters*, 76, 537-545.

De Groot, M. H. (1970). *Optimal Statistical Decisions*. McGraw-Hill, New York.

Hofmann, G. (2004). Comparing the Fisher information in record data and random observations. *Statistical Papers*, 45, 517-528.

Hofmann, G. and Nagaraja, H. N. (2003). Fisher information in record data. *Metrika*, 57, 177-193.

Lawless, J. L. (2003). *Statistical Models and Methods for Lifetime Data*. Second Edition. John Wiley, New York.

Lehmann, E. L. (1989). *Theory of Point Estimation*. John Wiley, New York.

Mardia, K. V. (1974). Applications of some measures of multivariate skewness and kurtosis for testing normality and robustness studies. *Sankhyā*, 58, 105-121.

McKay, A. T. (1934). Sampling from batches. *Journal of the Royal Statistical Society*, Supplement 1, 207-216.

Nagaraja, H. N. and Abo-Eleneen, Z. A. (2008). Fisher information in order statistics and their concomitants in bivariate censored samples. *Metrika*, 67, 327-347.