ON THE ORDERING OF RANDOM VECTORS IN A NORM SENSE

I.G.Bairamov and O.L.Gebizlioglu

Ankara University, Faculty of Science
Department of Statistics, 06100
Tandogan, Ankara, Turkey

Abstract

Let $R^m, m \geq 1$, be the real Euclidean space. Suppose $X_1, X_2, .., .X_n \in R^m$ are independent identically distributed (i.i.d.) random variables ($m > 1$ random vectors) (r.v.'s) with distribution function (d.f.) F. Denote by $\|.\|$ the norm defined in R^m. It is clear that $\|X_1\|, \|X_2\|, ..., \|X_n\|$ are i.i.d. r.v. with d.f. $P\{\|X_i\| \leq x\} \equiv F^*(x)$, $x \in R$. If F is assumed to be continuous, the probability of any two or more of these r.v. assuming equal magnitudes is zero. Therefore, there exists a unique ordered arrangement within the r.v. $\|X_i\|$, $i = 1, 2, ..., n$. We say that X_1 precedes X_2 (or that X_1 is less than X_2 in a norm sense) if $\|X_1\| \leq \|X_2\|$ and denote $X_1 \prec X_2$. Suppose $X^{(1)}$ denotes the smallest of the set $X_1, X_2, .., .X_n$; $X^{(2)}$ denotes the second smallest, etc. ; and $X^{(n)}$ denotes the largest in a norm sense. In this paper we have investigated the distributional properties of r.v.'s $X^{(1)}$, $X^{(2)}$, ..., $X^{(n)}$ and it's applications in statistical inference.

AMS 1991 Subject Classifications. 62E20,62G30; 62E15.
Key words and phrases. Random vectors, distribution function, probability density function, order statistics, norm ordered statistics, empirical distribution function.

1. INTRODUCTION.

Let $(\Omega, \Im, P)$ be a probability space, where Ω is a non-empty set of points ω, $\Im$ is a $\sigma-$field of subsets of Ω and P is a probability measure defined on the $\{\Omega, \Im\}$. Let us consider the real Euclidean space R^m. $\|.\|$ denotes the norm defined in R^m, $\Re^m$ denotes the Borel $\sigma-$algebra of subsets of R^m. Let $\xi(\omega)$, $\omega \in \Omega$, be the r.v. mapping Ω into R^m, so $\xi^{-1}(B) \in \Im$ for any $B \in \Re^m$. Let us denote by Ξ the class of continuous r.v.'s. If ω is fixed, then $\xi = \xi(\omega)$ is a point of R^m and we denote by $\|\xi(\omega)\|$ the norm of point $\xi \in R^m$. It is clear $\|\xi(\omega)\|$, $\omega \in \Omega$ is a random variable. In fact $\{\omega : \|\xi(\omega)\| \leq x\} = \{\omega : \xi(\omega) \in S(0, x)\} \in \Im$, where $S(0, x) = \{y \in R^m : \|y\| \leq x\} \in \Re^m$. Note that if ξ is an element of some class of r.v.'s, say Ξ , then we use also the notation $\xi \in \Xi$.

Definition 1.1. Two r.v.'s ξ_1 and ξ_2 are said to be equal in a norm sense if $\|\xi_1(\omega)\| = \|\xi_2(\omega)\|$ for each $\omega \in \Omega$. To denote "equal in a norm sense " we use the notation $\xi_1 \overset{n}{=} \xi_2$, where $\overset{n}{=}$ is read "has the same norm as". Note that $\overset{n}{=}$ is an equivalence relation.

Suppose Ξ_1 is the sub class of Ξ such that the elements of Ξ_1 have the same continuous (but not necessarily absolutely continuous) distribution, say F. Then for any $X_1(\omega) \in \Xi_1$ and $X_2(\omega) \in \Xi_1$, we have $P\{\omega : \|X_1(\omega)\| = \|X_2(\omega)\|\} = 0$. That is any two or more elements of Ξ_1 are different in a norm sense with probability 1.

Definition 1.2. Let ξ_1 and ξ_2 be the r.v.'s defined in $\{\Omega, \Im, P\}$. ξ_1 is said to be less than ξ_2 in a norm sense if

$$\|\xi_1(\omega)\| \leq \|\xi_2(\omega)\| \text{ for each } \omega \in \Omega \ .$$

To denote "less in a norm sense " we use the notation $\xi_1 \prec \xi_2$. By Definition 1.2, the following equalities are satisfied:

1. $\xi \prec \xi$ for each r.v. ξ;
2. If $\xi_1 \prec \xi_2$ and $\xi_2 \prec \xi_3$, then $\xi_1 \prec \xi_3$;
3. If $\xi_1 \prec \xi_2$ and $\xi_2 \prec \xi_1$, then $\xi_1 \overset{n}{=} \xi_2$.

Let Ω_1 be some subset of Ω. Note that if $\|\xi_1(\omega)\| = \|\xi_2(\omega)\|$ for each ω in Ω_1, then ξ_1 and ξ_2 are said to be equal in a norm sense on Ω_1. Analogously, if $\|\xi_1(\omega)\| \leq \|\xi_2(\omega)\|$ for each ω on Ω_1 we say "ξ_1 less than ξ_2 in a norm sense on Ω_1 " and write "$\xi_1 \prec \xi_2$ on Ω_1".

Definition 1.3. The r.v.'s $\xi_1$and ξ_2 are said to be comparable in a norm sense on Ω_1, if the following equalities are satisfied:

1. $P\left\{\omega \in \Omega_1 : \xi_1(\omega) \overset{n}{=} \xi_2(\omega)\right\} = 0$ 2. $\xi_1 \prec \xi_2$ or $\xi_2 \prec \xi_1$ on Ω_1

Now, suppose $X_1 = X_1(\omega)$, $X_2 = X_2(\omega), ..., X_n = X_n(\omega)$, $\omega \in \Omega$, be i.i.d. r.v.'s mapping Ω into R^m, with continuous d.f. F, i.e. $X_i \in \Xi_1, i = 1, 2, ..., n$. For any $r \in \{1, 2, ..., n\}$ Ω may be presented as follows:

$$\Omega = A \cup A_r^1 \cup A_r^2 \cup ... \cup A_r^n \qquad ,$$

where $A_r^k = \{\omega : \|X_1(\omega)\|, \|X_2(\omega)\|, ..., \|X_n(\omega)\|$ are different and $\|X_k(\omega)\|$ is the r th smallest among $\|X_1(\omega)\|, \|X_2(\omega)\|, ..., \|X_n(\omega)\|\}$, $r = 1, 2.., n$. Let us define the event

$$A = \{\omega : \text{two or more of r.v.'s } \|X_1(\omega)\|, \|X_2(\omega)\|, ..., \|X_n(\omega)\| \text{ have equal magnitudes}\} \, . \tag{1.1}$$

It is clear that $P(A) = 0$ and this fact will be used in what follows.

Define the r.v.'s $X^{(r)}$ for any $r = 1, 2, ..., n$ as follows:

$$X^{(r)} \equiv X^{(r)}(\omega) = X_k \quad \text{if} \quad \omega \in A_r^k, \quad k = 1, 2, ..., n.$$

If $\omega \in A$ and A has, for example, the following structure

$$A = \{\omega : \|X_{i_1}\| = \|X_{i_2}\| = \|X_{i_3}\| < \|X_{i_4}\| < \|X_{i_5}\| = \|X_{i_6}\| < \|X_{i_7}\| < ... < \|X_{i_n}\|\},$$

then $X^{(1)} = X_{i_1},\ X^{(2)} = X_{i_2},\ X^{(3)} = X_{i_3},\ X^{(4)} = X_{i_4},\ X^{(5)} = X_{i_5},\ X^{(6)} = X_{i_6}, ..., X^{(n)} = X_{i_n}$. It is clear that the r.v.'s $X^{(1)}, X^{(2)}, ..., X^{(n)}$ comparable in a norm sense in Ω and it satisfies

$$X^{(1)} \prec X^{(2)} \prec ... \prec X^{(n)}.$$

This paper aims to present an investigation of the distributional properties of random variables (vectors) $X^{(1)}$, $X^{(2)}, ..., X^{(n)}$ and their applications in statistical inference. We will call $X^{(1)}$, $X^{(2)}, ...,$ $X^{(n)}$ a "norm ordered statistics". Such statistics is closely related to order statistics if $n = 1$ but not the same with them at all. Record values can be considered in the sense of norms. So a theoretical investigation and illustration of applications of norm ordered statistics at the conjunction of order statistics and record values will be a fruitful attempt. The details on the theory of order statistics and record values can be found in the works of David (1970), Wilks (1962), Galambos (1987) and Ahsanullah (1995), among others. A general insight on stochastic ordering and its application areas is lucidly expressed by Shaked and Shanthikumar(1994). The ordering of multivariate data is discussed and four-fold classification of sub-ordering principles is expressed by Barnett (1976).

There are many situations in practice for which we need to investigate the distributional properties of a random vector whose elements are magnitudes of distance related characteristics of an event. For instance, in a two dimensional space, bombing on and around a target point has destructive effects on the point itself depending on its distance from the site of explosion in conjunction with some other factors. Similarly, multidimensional epidemiological processes can be analyzed in terms of the norm ordered statistics for the spread of disease analysis. It is possible to extend the examples to other fields and application areas.

Here we have presented some major theoretical results of our work. The distribution of the above mentioned random vectors and the empirical structural function are discussed.

2. THE DISTRIBUTIONS OF NORM ORDERED STATISTICS

Let us consider the case $m = 2$, for simplicity. Suppose F has the probability density function f, i.e. for $X_i = (X_i^1, X_i^2) \in R^2,\ i = 1, 2, ..., n$, one has

$$F(x, y) = P\left\{X_i^1 \le x, X_i^2 \le y\right\} = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u, v)dudv.$$

Let us denote $h(x, y) = P\{\|X_1\| \le \|\bar{x}\|\}$, where $\bar{x} = (x, y) \in R^2$. We call the function $h(x, y)$ the structural function of the sample $X_1, X_2, ..., X_n$. Consider

$X^{(1)} \prec X^{(2)} \prec ... \prec X^{(n)}$. It is clear that the r.v.'s $X^{(1)}, X^{(2)}, ..., X^{(n)}$ are not independent. By using $P(A) = 0$ (see (1.1)), for any $B \in \Re^2$ one can write

$$P\left\{X^{(n)} \in B\right\} = \sum_{k=1}^{n} P\left\{X_k \in B, \|X_k\| \geq \|X_i\|,\ i = 1,2,...,n; i \neq k\right\}$$

$$= \sum_{k=1}^{n} \int\int P\left\{X_k \in B, \|X_k\| \geq \|X_i\|,\ i = 1,2,...,n; i \neq k / X_k^1 = x, X_k^2 = y\right\} dF(x,y)$$

$$= \sum_{k=1}^{n} \int\int_{B} P\left\{X_k \in B, \|X_i\| \leq \|\bar{x}\|,\ i = 1,2,...,n;\ i \neq k\right\} dF(x,y)$$

$$= \sum_{k=1}^{n} \int\int_{B} \left[P\left\{\|X_1\| \leq \|\bar{x}\|\right\}\right]^{n-1} dF(x,y) = n \int\int_{B} \left[h(x,y)\right]^{n-1} dF(x,y).$$

Hence the probability density function of $X^{(n)}$ is

$$f_n(x,y) = n\left[h(x,y)\right]^{n-1} f(x,y),$$

where by $f_r(x,y)$ we denote the probability density function of r.v. $X^{(r)}$, $r = 1,2,...,n$. Similarly one can write

$$f_1(x,y) = n\left[1 - h(x,y)\right]^{n-1} f(x,y).$$

In general one has for $B \in \Re^2$

$$P\left\{X^{(r)} \in B\right\} = \sum_{k=1}^{n} P\left\{X_k \in B, \|X_k\| \text{ is the } r \text{ th smallest among } \|X_1\|, \|X_2\|, ..., \|X_n\|\right\}$$

$$= \sum_{k=1}^{n} \binom{n-1}{r-1} P\left\{X_k \in B, \|X_1\| \leq \|X_k\|, \|X_2\| \leq \|X_k\|, ..., \|X_{r-1}\| \leq \|X_k\|, \|X_{r+1}\| > \|X_k\|, ...,\right.$$

$$\left.\|X_n\| > \|X_k\|\right\} = \sum_{k=1}^{n} \binom{n-1}{r-1} \iint_{B} \left[P\left\{\|X_1\| \leq \|\bar{x}\|\right\}\right]^{r-1} \left[1 - P\left\{\|X_1\| \leq \|\bar{x}\|\right\}\right]^{n-r} dF(x,y) =$$

$$= n\binom{n-1}{r-1} \iint_{B} \left[h(x,y)\right]^{r-1} \left[1 - h(x,y)\right]^{n-r} dF(x,y).$$

Therefore the probability density function of $X^{(r)}, 1 \leq r \leq n$ is

$$f_r(x,y) = n\binom{n-1}{r-1} \left[h(x,y)\right]^{r-1} \left[1 - h(x,y)\right]^{n-r} f(x,y). \tag{1}$$

From (2.1) one can write

$$\int\int \left[h(x,y)\right]^{r-1} \left[1 - h(x,y)\right]^{n-r} dF(x,y) = \left(n\binom{n-1}{r-1}\right)^{-1}. \tag{2}$$

From (2.2) it follows for any $n \geq 1$

$$\int\int [1-h(x,y)]^n \, dF(x,y) = \frac{1}{n+1},$$

$$\int\int [h(x,y)]^n \, dF(x,y) = \frac{1}{n+1} \; .$$

It is clear that in the case of $m = k$ (2.1), has the following form:

$$f_r(x_1,x_2,...,x_k) = n\binom{n-1}{r-1} [h(x_1,x_2,...,x_k)]^{r-1} [1-h(x_1,x_2,...,x_k)]^{n-r} f(x_1,x_2,...,x_k).$$

Let us consider some examples:

Example 2.1. Let $X_1, X_2, ..., X_n$ be i.i.d. r.v.'s and $X_1 = (X_1^1, X_1^2, ..., X_1^k)$, where $X_1^1, X_1^2, ..., X_1^k$ are i.i.d. normal distributed random variables with $EX_1^1 = 0$, $var(X_1^1) = 1$. The probability density function of X_1 is

$$f(x_1,x_2,...,x_k) = \frac{1}{(2\pi)^{\frac{k}{2}}} \exp\left\{-\frac{x_1^2+x_2^2+...+x_k^2}{2}\right\}.$$

Suppose for any $\bar{x} = (x_1,x_2,...,x_k) \in R^k$ the norm is defined as $\|\bar{x}\|^2 = x_1^2+x_2^2+...+x_k^2$. Then $(X_1^1)^2+(X_1^2)^2+...+(X_1^k)^2$ is distributed as a random variable with χ^2 distribution of degree k :

$$P\left\{\|X_1\|^2 \leq x\right\} = G_{\frac{1}{2},\frac{k}{2}}(x),$$

where $G_{\alpha,\beta}(x)$ denotes the d.f of gamma distribution with the parameters (α,β). Then

$$h(x_1,x_2,...,x_k) = G_{\frac{1}{2},\frac{k}{2}}(x_1^2+x_2^2+...+x_k^2).$$

Example 2.2. Let $(X_1^1, X_1^2, ..., X_1^k)$ has a probability density function

$$f(x_1,x_2,...,x_k) = \begin{cases} \lambda^k \exp\{-\lambda(x_1+x_2+...+x_k\}, & \text{if } x_1 \geq 0, x_2 \geq 0, ..., x_k \geq 0 \\ 0, & \text{otherwise} . \end{cases}$$

Suppose for any $\bar{x} = (x_1,x_2,...,x_k) \in R^k$ the norm is denoted by $\|\bar{x}\| = \sum_{i=1}^k |x_i|$ .One can write for $X_1 = (X_1^1, X_1^2, ..., X_1^k)$

$$h(x_1,x_2,...,x_k) = P\left\{\left|X_1^1\right|+\left|X_1^2\right|+...+\left|X_1^k\right| \leq |x_1|+|x_2|+...+|x_k|\right\} =$$
$$P\left\{X_1^1+X_1^2+...+X_1^k \leq |x_1|+|x_2|+...+|x_k|\right\} = G_{k,\lambda}(|x_1|+|x_2|+...+|x_k|),$$

since $X_1^1, X_1^2, ..., X_1^k$ are i.i.d. r.v.'s with d.f. $F(u) = 1-e^{-\lambda u}$, $u \geq 0$.

Example 2.3. Let $\bar{x} = (x_1, x_2, ..., x_k) \in R^k$ and $\|\bar{x}\| = \max(x_1, x_2, ..., x_k) \equiv x_{(k)}$. Consider $X_1 = (X_1^1, X_1^2, ..., X_1^k)$ with d.f. $P\left\{X_1^1 \le x_1, X_1^2 \le x_2, ..., X_1^k \le x_k\right\} = F(x_1, x_2, ..., x_k)$. It is clear that

$$h(x_1, x_2, ..., x_k) = P\left\{\|X_1\| \le \|\bar{x}\|\right\} = P\left\{\max(X_1^1, X_1^2, ..., X_1^k) \le \max(x_1, x_2, ..., x_k)\right\} = P\left\{X_1^1 \le x_{(k)}, X_1^2 \le x_{(k)}, ..., X_1^k \le x_{(k)}\right\} = F(x_{(k)}, x_{(k)}, ..., x_{(k)}).$$

As an important class of two dimensional distributions let us consider the class of exponential distributions for $k = 2$. This class may be defined as

$$F(x, y) = 1 - \exp(-x) - \exp(-y) + G(x, y),$$

where $G(x, y) = P\left\{X_1^1 \ge x, X_1^2 \ge y\right\}$. For example selecting $G(x, y) = \exp(-x - y - \lambda \max(x, y))$, $\lambda > 0$, we have the two dimensional exponential distribution of Marshal-Olkin. In reliability theory this distribution expresses the distribution of life time of two dependent elements (see Barlow and Proschan, 1975). It is clear that in this case $h(x, y) = 1 - 2\exp\left\{-\max(x, y)\right\} + \exp\left\{-(2 + \lambda)\max(x, y)\right\}$.

2.1. The marginal distributions

Let $X_1, X_2, ..., X_n \in R^2$ be i.i.d. r.v.'s and $X_i = (X_i^1, X_i^2)$, $i = 1, 2, ..., n$. Consider $X^{(1)} \prec X^{(2)} \prec ... \prec X^{(n)}$, $X^{(i)} = (X_1^{(i)}, X_2^{(i)})$, $i = 1, 2, ..., n$. It is easy to see that the marginal density function $f_1^{(r)}$ of $X_1^{(r)}$ and $f_2^{(r)}$ of $X_2^{(r)}$,$1 \le r \le n$, are expressed as follows:

$$f_1^{(r)}(x) = \int_{-\infty}^{\infty} f_r(x, y)dy = n\binom{n-1}{r-1} \int_{-\infty}^{\infty} \left[h(x, y)\right]^{r-1} \left[1 - h(x, y)\right]^{n-r} f(x, y)dy,$$
$$f_2^{(r)}(x) = \int_{-\infty}^{\infty} f_r(x, y)dx = n\binom{n-1}{r-1} \int_{-\infty}^{\infty} \left[h(x, y)\right]^{r-1} \left[1 - h(x, y)\right]^{n-r} f(x, y)dx.$$

For $r = n$, we have $f_1^{(n)}(x) = n \int_{-\infty}^{\infty} \left[h(x, y)\right]^{n-1} f(x, y)dy$. Denote by $f_{1|2}^{(n)}(x \mid y)$ the conditional probability density function of $X_1^{(n)}$ with respect to $\left\{X_2^{(n)} = y\right\}$. Then we have

$$f_{1|2}^{(n)}(x \mid y) = \frac{\left[h(x, y)\right]^{n-1} f(x, y)}{\int_{-\infty}^{\infty} \left[h(x, y)\right]^{n-1} f(x, y)dx}.$$

3. THE JOINT DISTRIBUTIONS OF TWO OR MORE NORM ORDERED STATISTICS

Let $1 \le r < s \le n$, $X^{(r)} \prec X^{(s)}.$ Consider $B_1, B_2 \in \Re^2$, such that $\|\bar{x}_1\| \le \|\bar{x}_2\|$ for any $\bar{x}_1 \in B_1$, $\bar{x}_2 \in B_2$, $\bar{x}_k = (x_k, y_k)$, $k = 1, 2$. Suppose F has the probability density function of f.

Theorem 3.1. Under the above mentioned assumptions it is true that

$$P\left\{X^{(r)}\in B_1, X^{(s)}\in B_2\right\} = n(n-1)\binom{n-2}{r-1}\binom{n-r-1}{s-r-1}\times$$

$$\times\underset{B_1\times B_2}{\int\int\int\int}\left[h(x_1,y_1)\right]^{r-1}\left[h(x_2,y_2)-h(x_1,y_1)\right]^{s-r-1}\left[1-h(x_2,y_2)\right]^{n-s}dF(x_1,y_1)dF(x_2,y_2)$$

and the joint probability density function of $X^{(r)}$ and $X^{(s)}$ is

$$f_{rs}(x_1,y_1,x_2,y_2)=\begin{cases} n(n-1)\binom{n-2}{r-1}\binom{n-r-1}{s-r-1}\left[h(x_1,y_1)\right]^{r-1}\left[h(x_2,y_2)-h(x_1,y_1)\right]^{s-r-1}\times \\ \times\left[1-h(x_2,y_2)\right]^{n-s}f(x_1,y_1)f(x_2,y_2)\,, \qquad if \quad \|\bar{x}_1\|\le\|\bar{x}_2\| \\ 0\;, \qquad elswhere\;. \end{cases} \quad (3)$$

Proof. By using $P(A)=0$ for $1\le r<s\le n$,see (1.1), one has

$$P\left\{X^{(r)}\in B_1, X^{(s)}\in B_2\right\}=\sum_{k\neq j}P\left\{X^{(r)}\in B_1, X^{(s)}\in B_2, X_k \text{ is } r \text{ th smallest in a norm sense,}\right.$$

$$\left. X_j \text{ is } s \text{ th smallest in a norm sense}\right\} = \binom{n-2}{r-1}\binom{n-2-(r-1)}{s-r-1}\times$$

$$\times\sum_{k\neq j}P\left\{X_k\in B_1, X_j\in B_2, \|X_1\|<\|X_k\|,...,\|X_{r-1}\|<\|X_k\|, \|X_k\|<\|X_{r+1}\|<\|X_j\|,...,\right.$$

$$\left.\|X_k\|<\|X_{s-1}\|<\|X_j\|, \|X_j\|<\|X_{s+1}\|,...,\|X_j\|<\|X_n\|\right\}=\binom{n-2}{r-1}\binom{n-2-(r-1)}{s-r-1}\times$$

$$\times\sum_{k\neq j}\int P\left\{X_k\in B_1, X_j\in B_2, \|X_1\|<\|X_k\|,...,\|X_{r-1}\|<\|X_k\|, \|X_k\|<\|X_{r+1}\|<\|X_j\|,...,\right.$$

$$\left.\|X_k\|<\|X_{s-1}\|<\|X_j\|, \|X_j\|<\|X_{s+1}\|,...,\|X_j\|<\|X_n\| / X_k^1=x_1, X_k^2=y_1, X_j^1=x_2, X_j^2=y_2\right\}\times$$

$$dF(x_1,y_1)dF(x_2,y_2)=\binom{n-2}{r-1}\binom{n-2-(r-1)}{s-r-1}\sum_{k\neq j}\underset{B_1\times B_2}{\int\int\int\int}\left[P\left\{\|X_1\|<\|\bar{x}_1\|\right\}\right]^{r-1}\times$$

$$\left[P\left\{\|X_1\|<\|\bar{x}_2\|\right\}-P\left\{\|X_1\|<\|\bar{x}_1\|\right\}\right]^{s-r-1}\left[1-P\left\{\|X_1\|<\|\bar{x}_2\|\right\}\right]^{s-r}dF(x_1,y_1)dF(x_2,y_2)$$

$$=\binom{n-2}{r-1}\binom{n-2-(r-1)}{s-r-1}\sum_{k\neq j}\underset{B_1\times B_2}{\int\int\int\int}\left[h(x_1,y_1)\right]^{r-1}\left[h(x_2,y_2)-h(x_1,y_1)\right]^{s-r-1}\times$$

$$\times\left[1-h(x_2,y_2)\right]^{n-s}dF(x_1,y_1)dF(x_2,y_2)\,.$$

$$=n(n-1)\binom{n-2}{r-1}\binom{n-2-(r-1)}{s-r-1}\underset{B_1\times B_2}{\int\int\int\int}\left[h(x_1,y_1)\right]^{r-1}\left[h(x_2,y_2)-h(x_1,y_1)\right]^{s-r-1}\times$$

$$\times\left[1-h(x_2,y_2)\right]^{n-s}dF(x_1,y_1)dF(x_2,y_2)\,. \qquad (Q.E.D.)$$

Corollary 3.1. *Let* $1 \leq r_1 < r_2 < ... < r_k \leq n$*. The joint probability density function of* $X^{(r_1)}, X^{(r_2)}, ..., X^{(r_k)}$ *is*

$$f_{r_1,r_2,...,r_k}(x_1, y_1, x_2, y_2, ..., x_k, y_k) = \frac{n!}{(r_1-1)!(r_2-r_1-1)!...(n-r_k)!} \left[h(x_1, y_1)\right]^{r_1-1} \times$$

$$\left[h(x_2, y_2) - h(x_1, y_1)\right]^{r_2-r_1-1} ... \left[h(x_k, y_k) - h(x_{k-1}, y_{k-1})\right]^{r_k-r_{k-1}-1} \left[1 - h(x_k, y_k)\right]^{n-r_k} \times$$

$$\times f(x_1, y_1)f(x_2, y_2)...f(x_k, y_k),$$

$$\text{if} \quad \|\bar{x}_1\| < \|\bar{x}_2\| < ... < \|\bar{x}_k\|$$

and

$$f_{r_1,r_2,...,r_k}(x_1, y_1, x_2, y_2, ..., x_k, y_k) = 0, \qquad \text{elsewhere.}$$

The joint probability density function of $X^{(1)}, X^{(2)}, ..., X^{(n)}$ *is*

$$f_{1,2,...,n}(x_1, y_1, x_2, y_2, ..., x_n, y_n) = \begin{cases} n!f(x_1, y_1)f(x_2, y_2)...f(x_n, y_n) & \text{if} \quad \|\bar{x}_1\| < \|\bar{x}_2\| < ... < \|\bar{x}_k\| \\ 0, & \text{elsewhere}. \end{cases}$$

4. THE DISTRIBUTION OF SAMPLE RANGE

Let us consider $X^{(s)} = (X_1^{(s)}, X_2^{(s)})$ and $X^{(r)} = (X_1^{(r)}, X_2^{(r)})$, where $X_i^{(l)}, i = 1, 2$ denotes the i th coordinate of random vector $X^{(l)} \in R^2, l = 1, 2, ..., n$. Denote $R_{rs} = X^{(s)} - X^{(r)}$, $1 \leq r < s \leq n$.

Theorem 4.1. *The probability density function of* R_{rs} *is*

$$f_{R_{rs}}(x, y) = \frac{n!}{(r-1)!(s-r-1)!(n-s)!} \iint\limits_{\|\bar{t}\| \leq \|\bar{t}+\bar{x}\|} \left[h(t_1, t_2)\right]^{r-1} \left[h(t_1 + x, t_2 + y) - h(t_1, t_2)\right]^{s-r-1} \times$$

$$\left[1 - h(t_1 + x, t_2 + y)\right]^{n-s} f(t_1, t_2)f(t_1 + x, t_2 + y)dt_1dt_2. \qquad (4)$$

Proof. Let $B = \{(u, v) : -\infty < u \leq x, -\infty < v \leq y\}$, $c_{rs} = \frac{n!}{(r-1)!(s-r-1)!(n-s)!}$,$\bar{v} = (v_1, v_2)$, $\bar{u} = (u_1, u_2)$, $\bar{z} = (z_1, z_2)$. By using (3.1) one can write

$$P\left\{X^{(s)} - X^{(r)} \in B\right\} = P\left\{X_1^{(s)} - X_1^{(r)} \leq x, X_2^{(s)} - X_2^{(r)} \leq y\right\}$$

$$= \iiiint\limits_{\substack{u_1 - v_1 \leq x \\ u_2 - v_2 \leq y \\ \|\bar{v}\| \leq \|\bar{u}\|}} f_{rs}(v_1, v_2, u_1, u_2)du_1du_2dv_1dv_2$$

$$= c_{rs} \iiiint\limits_{\substack{u_1 - v_1 \le x \\ u_2 - v_2 \le y \\ \|\bar{v}\| \le \|\bar{u}\|}} h(v_1, v_2)^{r-1} \left[h(u_1, u_2) - h(v_1, v_2)\right]^{s-r-1} \times .$$

$$\times \left[1 - h(u_1, u_2)\right]^{n-s} dF(u_1, u_2) dF(v_1, v_2). \qquad (5)$$

After changing the variables $u_1 - v_1 = z_1$, $v_1 = t_1$ $u_2 - v_2 = z_2$, $v_2 = t_2$, from the (4.2) one obtains

$$P\left\{X^{(s)} - X^{(r)} \in B\right\} = c_{rs} \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} \iint\limits_{\|\bar{t}\| \le \|\bar{t} + \bar{z}\|} \left[h(t_1, t_2)\right]^{r-1} \left[h(t_1 + z_1, t_2 + z_2) - h(t_1, t_2)\right]^{s-r-1} \times$$

$$\left[1 - h(t_1 + z_1, t_2 + z_2)\right]^{n-s} f(t_1, t_2) f(t_1 + z_1, t_2 + z_2) dt_1 dt_2 dz_1 dz_2. \qquad (Q.E.D.)$$

5. THE EMPIRICAL STRUCTURAL FUNCTION

If it is difficult or impossible to obtain admissible expression for $h(x_1, x_2, ..., x_k)$ one needs to estimate the structural function empirically. Let $B \in \Re^2$. Denote by $\nu_n(B)$ the number of observations $X_1, X_2, ..., X_n$ falling in B. One can write

$$\nu_n(B) = \sum_{i=1}^{n} I_{X_i}(B), \qquad (6)$$

where $I_x(B) = 1$ if $x \in B$ and $I_x(B) = 0$ if $x \notin B$. Let us denote the points of R^2 as $\bar{u} = (u_1, u_2)$, $\bar{x} = (x_1, x_2)$ etc. We call $F_n^*(t_1, t_2) = \frac{\nu_n(B_{\bar{t}})}{n}$, the empirical structural function, where $B_{\bar{t}} \equiv \left\{(y_1, y_2) : \|\bar{y}\| \le \|\bar{t}\|\right\}$. One can write

$$F_n^*(t_1, t_2) = \frac{1}{n} \sum_{i=1}^{n} I_{X_i}(B_{\bar{t}}) = F_n(\|\bar{t}\|), \qquad (7)$$

where $F_n(u)$ denotes the empirical d.f. of the sample $\|X_1\|, \|X_2\|, ..., \|X_n\|$. Hence by using the representation (5.1) and the equality (5.2) the existing results for empirical d.f. (see Borovkov 1984, Gaensler and Stute 1987 etc.) can be formulated for $F_n^*(t_1, t_2)$. Note that $EF_n^*(t_1, t_2) = h(t_1, t_2)$ and $var(F_n^*(t_1, t_2))$ $= h(t_1, t_2)(1 - h(t_1, t_2))$.

Theorem 5.1. For k=1,2,...,n the following relation is true:

$$P\left\{nF_n^*(t_1, t_2) = k\right\} = \binom{n}{k} \left[h(t_1, t_2)\right]^k \left[1 - h(t_1, t_2)\right]^{n-k}.$$

Theorem 5.2. It is true that

$$\sup_{\bar{t}\in R^2} |F_n^*(t_1,t_2) - h(t_1,t_2)| \to 0 \text{ , almost surely as } n\to\infty.$$

6. THE RECORDS IN A NORM SENSE

Let $X_i = (X_i^1, X_i^2, ..., X_i^m) \in R^m \;\; i = 1, 2, ...$ be a sequence of i.i.d. r.v.'s with d.f. $F(x_1, x_2, ..., x_m) = P\left\{X_1^1 \le x_1, X_1^2 \le x_2, ..., X_1^m \le x_m\right\}$ and probability density function $f(x_1, x_2, ..., x_m)$. Let $Y_n = \max(\min)(\|X_1\|, \|X_2\|, ..., \|X_n\|)$. We say that X_i is upper(lower) record value in a norm sense (or norm-record value) of this sequence if $Y_i > (<) Y_{i-1}, i > 1$. By definition, X_1 is an upper as wel as a lower record value. Let us define the following random variables

$$V(1) = 1, V(r) = \min\left\{i : i > V(r-1), \left\|X_{V(i)}\right\| > \left\|X_{V(r-1)}\right\|\right\}, r > 1.$$

We will call the random variables $V(1), V(2), ..., V(r), ...$ upper norm- record times. Accordingly the r.v.'s $X_{V(1)}, X_{V(2)}, ..., X_{V(r)}, ...$ will be called norm - records.

Example 6.1. In analyzing the growth of Olympic records in weighlifting, the sum of two-hand snatch and two hand clean and jark is also considered as a record. It is clear that in this case we use the norm $\|\bar{x}\| = |x_1| + |x_2|$ defined in R^2.

The proof of the following lemma is easy and hence omitted.

Lemma 6.1. 1) For any $r = 1, 2, ...$ the random variable $V(r)$ has the same distribution as that of the record times of arbitrary continuous i.i.d. r.v.'s.

2) The sequence of random variables $V(n), n \ge 2$ is a Markov chain with the transition probabilities

$$P\left\{V(n) = k / V(n-1) = j\right\} = \begin{cases} \frac{j}{k(k-1)}, & \text{if} \quad k > j \ge n-1 \ge 2 \\ 0, & \text{otherwise .} \end{cases}$$

From Lemma 6.1 and using the known results of theory of records (for more details one can look of works of Ahsanullah(1995), Nevzorov (1988), Nagaraja (1988) etc.) we have

$$P\left\{V(2) = k\right\} = \frac{1}{k(k-1)}, k \ge 2,$$

$$P\left\{V(3) = k\right\} = \sum_{j=2}^{\infty} P\left\{V(3) = k / V(2) = j\right\} P\left\{V(2) = j\right\} = \frac{1}{k(k-1)} \sum_{j=1}^{k-2} \frac{1}{j}, \text{ etc.}$$

One can easily obtain the probability density function of $X_{V(r)}, r = 1, 2, ...$ expressed by $h(x_1, ..., x_m)$. For simplicity consider the case of $m = 2$. Let $B(\bar{t}) = \left\{(u,v) : -\infty < u \le t_1, -\infty < v \le t_2\right\}$. Elements of R^2 will be denoted as above by $\bar{x}$, $\bar{y}$ and so forth, or more explicitly as $\bar{x} = (x_1, x_2)$, $\bar{y} = (y_1, y_2)$, and so on. Let $0 < F(x_1, x_2) < 1$, $(x_1, x_2) \in R^2$. Then $0 < h(x_1, x_2) < 1, (x_1, x_2) \in R^2$. One can write

$$P\left\{X_{V(2)} \in B(\bar{t})\right\} = \sum_{j=2}^{\infty} P\left\{X_{V(2)} \in B(\bar{t}), V(2) = j\right\} = \sum_{j=2}^{\infty} \int\int\int\int P\{\ X_j \in B(\bar{t}), \|X_1\| > \|X_2\|,$$

$$\|X_1\| > \|X_3\|, ..., \|X_1\| > \|X_{j-1}\|, \|X_1\| < \|X_j\| / X_1^1 = x_1, X_1^2 = x_2, X_i^1 = y_1, X_i^2 = y_2\ \} dF(x_1, x_2) \times$$

$$\times dF(y_1, y_2) = \sum_{j=2}^{\infty} \iint_{B(\bar{t})} \iint_{\|\bar{x}\| < \|\bar{y}\|} P\{\|X_2\| \leq \|\bar{x}\|, ..., \|X_{j-1}\| \leq \|\bar{x}\|\} dF(x_1, x_2) dF(y_1, y_2)$$

$$= \sum_{j=2}^{\infty} \iint_{B(\bar{t})} \iint_{\|\bar{x}\| < \|\bar{y}\|} [h(x_1, x_2)]^{j-2} dF(x_1, x_2) dF(y_1, y_2). \tag{8}$$

It easy to see that if $\bar{x} \in \{(x_1, x_2) : \|\bar{x}\| < \|\bar{y}\|\}$ it satisfies

$$h(x_1, x_2) = P\{\|X_1\| < \|\bar{x}\|\} \leq P\{\|Y_1\| < \|\bar{y}\|\} \leq h(y_1, y_2).$$

For any $(y_1, y_2) \in R^2$, the series $\sum_{j=2}^{\infty} [h(y_1, y_2)]^{j-2}$ is convergent, therefore $\sum_{j=2}^{\infty} [h(x_1, x_2)]^{j-2}$ converges uniformly, due to Weierstrass theorem (see Rudin,1959, Theorem 7.10, p.134). That is to say that the series in (6.1) may be integrated term by term. Hence from (6.1) it follows

$$P\left\{X_{V(2)} \in B(\bar{t})\right\} = \iint_{B(\bar{t})} \iint_{\|\bar{x}\| < \|\bar{y}\|} \sum_{j=2}^{\infty} [h(x_1, x_2)]^{j-2} dF(x_1, x_2) dF(y_1, y_2)$$

$$= \iint_{B(\bar{t})} \left[\iint_{\|\bar{x}\| < \|\bar{y}\|} \frac{1}{1 - h(x_1, x_2)} dF(x_1, x_2) \right] dF(y_1, y_2).$$

Hence the probability density function of $X_{V(2)}$ is

$$f_{X_{V(2)}}(t_1, t_2) = \iint_{\|\bar{x}\| < \|\bar{y}\|} \frac{1}{1 - h(x_1, x_2)} dF(x_1, x_2).$$

Similarly one can obtain the probability density function of $X_{V(3)}$

$$f_{X_{V(3)}}(t_1, t_2) = \iiiint_{\|\bar{x}\| \leq \|\bar{y}\| \leq \|\bar{t}\|} \frac{f(x_1, x_2)}{1 - h(x_1, x_2)} \frac{f(y_1, y_2)}{1 - h(y_1, y_2)} dx_1 dx_2 dy_1 dy_2 \text{ etc.}$$

ACKNOLEDGMENT

The authors thank the Referee and an Editor for their valuable suggestions which resulted in the improvement of the presentation of this paper.

REFERENCES

1. Ahsanullah M. (1995) *Record Statistics.* Nova Science Publishers. Inc.Commack,NY.
2. Barlow R., Proschan F. (1975) *Statistical Theory of Reliability and Life Testing. Probability Models.* Holt, Rinehart and Winston, New-York.
3. Barnett V. (1976) The Ordering of Multivariate Data. *J.Royal.Statist.Soc.*A,139,Part 3, 318-343.
4. Borovkov A.A. (1984) *Mathematical Statistics.*Nauka, Moscow, (in Russian).
5. David H.A. (1970) *Order Statistics.* John Wiley & Sons.New York-London -Sydney - Toronto.
6. Gaensler P.,Stute W. (1987) *Seminar on Empirical Processes.* Birkhauser Verlag. Basel.Boston.
7. Galambos J. (1987) *The Asymptotic Theory of Extreme Order Statistics.*(2nd ed). Klieger, Malabar (FL).
8. Nagaraja H.N. (1988) Record Values and Related Statistics. A Review.*Commun.in Statist. Theor. Meth.* 17, 2223-2238.
9. Nevzorov V.B. (1988) Records. *Theory. Probab.Appl.* 32,201-228.
10. Rudin W. (1959). *Principles of Mathematical Analysis.* Mc GRAW-HILL Kogakusha, LTD. Tokyo, Düsseldorf, London,Singapure, Sidney.
11. Shaked M., Shanthikumar J.G. (1994) *Stochastic Orders and Their Applications.* Academic Press. Boston.
12. Wilks S.S. (1962) *Mathematical Statistics.*(2nd ed.) New York-London. John Wiley & Sons.