

Applications of Mathematics

María Dolores Esteban; Domingo Morales; Leandro Pardo; María Luisa Menéndez
Order statistics and (r, s) -entropy measures

Applications of Mathematics, Vol. 39 (1994), No. 5, 321–337

Persistent URL: <http://dml.cz/dmlcz/134262>

Terms of use:

© Institute of Mathematics AS CR, 1994

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://dml.cz>

ORDER STATISTICS AND (r, s) -ENTROPY MEASURES

M. D. ESTEBAN, D. MORALES, L. PARDO, M. L. MENÉNDEZ, Madrid

(Received April 10, 1992)

Summary. K. M. Wong and S. Chen [9] analyzed the Shannon entropy of a sequence of random variables under order restrictions. Using (r, s) -entropies, I. J. Taneja [8], these results are generalized. Upper and lower bounds to the entropy reduction when the sequence is ordered and conditions under which they are achieved are derived. Theorems are presented showing the difference between the average entropy of the individual order statistics and the entropy of a member of the original independent identically distributed (i.i.d.) population. Finally, the entropies of the individual order statistics are studied when the probability density function (p.d.f.) of the original i.i.d. sequence is symmetric about its mean.

Keywords: Unified (r, s) -entropy measure, order statistics, Shannon entropy, logistic distribution.

AMS classification: 62B10, 62G30, 94A15.

1. INTRODUCTION

Statisticians have been studying the properties of order statistics for some time and have applied them to solve nonparametric inference problems like tolerance intervals for distribution, coverages, confidence interval estimates for quantiles and so on. Recently, applications of order statistics in diverse areas have been found such as in engineering, signal processing, speech processing, image coding, image and picture processing, echo removal and image coding. A partial list, with very readable references, can be found in K. M. Wong and S. Chen [9].

This work was partially supported by the Dirección General de Investigación Científica y Técnica (DGICYT) under the contracts PB91-0387 and PB91-0155.

Here we examine some interesting properties of order sequences and order statistics using the (r, s) -entropy studied by Taneja [8]. A numerical example with the logistic distribution is given.

Suppose we have a set of random variables X_1, X_2, \dots, X_N , with joint probability density function (p.d.f.) $f(x)$ at a point $\mathbf{x} = (x_1, \dots, x_N)$. If we write $\exp_2(a)$ to denote 2^a , then the (r, s) -entropy of the sequence $\mathbf{X} = (X_1, \dots, X_N)$ is defined as

$$\mathcal{E}_r^s(\mathbf{X}) = \begin{cases} H_r^s(\mathbf{X}) &= (2^{1-s} - 1)^{-1} \left[\left(\int_{\mathbb{R}^N} f(x)^r dx \right)^{\frac{s-1}{r-1}} - 1 \right], & r \neq 1, s \neq 1, \\ H_1^s(\mathbf{X}) &= (2^{1-s} - 1)^{-1} \left\{ \exp_2 \left((s-1) \int_{\mathbb{R}^N} f(\mathbf{x}) \log_2 f(\mathbf{x}) dx \right) - 1 \right\}, & r = 1, s \neq 1, \\ H_r^1(\mathbf{X}) &= (1-r)^{-1} \log_2 \int_{\mathbb{R}^N} f(\mathbf{x})^r dx, & r \neq 1, s = 1, \\ H(\mathbf{X}) &= - \int_{\mathbb{R}^N} f(\mathbf{x}) \log_2 f(\mathbf{x}) dx, & r = 1, s = 1 \end{cases}$$

for all $r \in (0, \infty)$ and any $s \in (-\infty, \infty)$, provided the integrals exist.

This measure includes as particular and / or limiting cases the measures studied by Shannon [6], Renyi [5], Havrda and Charvat [4], Arimoto [1] and Sharma and Mittal [7].

For every set of random variables X_1, X_2, \dots, X_N , the following limit relations hold

$$H_1^s(\mathbf{X}) = \lim_{r \rightarrow 1} H_r^s(\mathbf{X}), \quad H_r^1(\mathbf{X}) = \lim_{s \rightarrow 1} H_r^s(\mathbf{X})$$

and

$$H(\mathbf{X}) = \lim_{s \rightarrow 1} H_1^s(\mathbf{X}) = \lim_{r \rightarrow 1} H_r^1(\mathbf{X}).$$

We arrange the set of random variables X_1, X_2, \dots, X_N , in ascending order of magnitude so that

$$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(N)}$$

where the subscript (n) denotes the index of the variable after ordering. For convenience of notation, we denote the set after ordering by Y_1, Y_2, \dots, Y_N so that $Y_n \equiv X_{(n)}$. Then Y_n is called the n th-order statistic ($n = 1, \dots, N$).

Let $f_i(x)$ be the p.d.f. of X_i and let $F_i(x)$ be its cumulative distribution function (c.d.f.). If X_1, \dots, X_N are independent, then the joint p.d.f. of the order statistics $\mathbf{Y} = (Y_1, \dots, Y_N)$ at $\mathbf{y} = (y_1, \dots, y_N)$ is given by

$$f(\mathbf{y}) = \begin{vmatrix} + & & & + \\ f_1(y_1) & f_2(y_1) & \dots & f_N(y_1) \\ f_1(y_2) & f_2(y_2) & \dots & f_N(y_2) \\ \vdots & \vdots & & \vdots \\ f_1(y_N) & f_2(y_N) & \dots & f_N(y_N) \\ - & & & - \end{vmatrix}$$

for $y_1 \leq y_2 \leq \dots \leq y_N$ and $f(y) = 0$ otherwise, where $\dagger \dagger$ is the permanent which is defined like the determinant, except that all signs are positive. The p.d.f. of Y_k at y is given by

$$f_k^*(y) = \frac{1}{(N-k)!(k-1)!} \begin{array}{cccc} \dagger & F_1(y) & F_2(y) & \dots & F_N(y) & \dagger \\ & \cdot & \cdot & \dots & \cdot & \\ & F_1(y) & F_2(y) & \dots & F_N(y) & \\ & f_1(y) & f_2(y) & \dots & f_N(y) & \\ 1 - F_1(y) & 1 - F_2(y) & \dots & 1 - F_N(y) & & \\ \cdot & \cdot & \dots & \cdot & & \\ 1 - F_1(y) & 1 - F_2(y) & \dots & 1 - F_N(y) & & \end{array} \left. \begin{array}{l} \left. \begin{array}{l} \dots \\ \dots \\ \dots \end{array} \right\} k-1 \text{ rows} \\ \left. \begin{array}{l} \dots \\ \dots \\ \dots \end{array} \right\} N-k \text{ rows} \end{array} \right.$$

The extension of the derivation of the joint p.d.f. of $\mathbf{Y} = (Y_1, \dots, Y_N)$ to the case where (X_1, \dots, X_N) are dependent, having a joint p.d.f. $f(x_1, \dots, x_N)$, is given by

$$f^*(y) = \sum_{j=1}^{N!} f(\mathbf{y}_j),$$

where \mathbf{y}_j is the j th permutation of the elements in the vector $\mathbf{y}_1 = (y_1, \dots, y_N)$ and $y_1 \leq y_2 \leq \dots \leq y_N$.

In the following sections we examine some interesting properties of the (r, s) -entropy of the ordered sequence $\mathbf{Y} = (Y_1, \dots, Y_N)$ and of the individual order statistics.

2. (r, s) -ENTROPY MEASURE OF AN ORDERED SEQUENCE

In this section a result is presented showing the amount of (r, s) -entropy reduction when the sequence is ordered. Upper and lower bounds to the (r, s) -entropy measure reduction and conditions under which they are achieved are derived.

Theorem 1. *The (r, s) -entropy measure of the ordered sequence $\mathbf{Y} = (Y_1, \dots, Y_N)$ is given by*

$$\mathcal{E}_r^s(\mathbf{Y}) = \begin{cases} (2^{1-s} - 1)^{-1} \left[\left(\int_{\mathbf{R}^N} f(\mathbf{y}_1)^r \frac{(\sum_{j=1}^{N!} f(\mathbf{y}_j))^r}{\sum_{j=1}^{N!} f^r(\mathbf{y}_j)} d\mathbf{y}_1 \right)^{\frac{s-1}{r-1}} - 1 \right], & r \neq 1, s \neq 1, \\ (2^{1-s} - 1)^{-1} \left[\exp_2 \left\{ (s-1) \int_{\mathbf{R}^N} f(\mathbf{y}_1) \log_2 \left(\sum_{j=1}^{N!} f(\mathbf{y}_j) \right) d\mathbf{y}_1 \right\} - 1 \right], & r = 1, s \neq 1, \\ (1-r)^{-1} \log_2 \left[\int_{\mathbf{R}^N} f^r(\mathbf{y}_1) \frac{(\sum_{j=1}^{N!} f(\mathbf{y}_j))^r}{\sum_{j=1}^{N!} f^r(\mathbf{y}_j)} d\mathbf{y}_1 \right], & r \neq 1, s = 1, \\ - \int_{\mathbf{R}^N} f(\mathbf{y}_1) \log_2 \left(\sum_{j=1}^{N!} f(\mathbf{y}_j) \right) d\mathbf{y}_1, & r = 1, s = 1 \end{cases}$$

for all $r \in (0, \infty)$ and any $s \in (-\infty, \infty)$, provided the integrals exist.

Proof. The region of integration of the joint p.d.f. of the ordered sequence \mathbf{Y} is governed by the condition $-\infty < y_1 \leq y_2 \leq \dots \leq y_N < \infty$. To incorporate this condition into the expression for $f(\mathbf{y})$, we use the unit step function

$$U(y) = \begin{cases} 1, & y \geq 0 \\ 0, & y < 0 \end{cases}$$

and introduce the notation

$$U(\mathbf{y}_1) = U(y_N - y_{N-1}) \cdot U(y_{N-1} - y_{N-2}) \cdot \dots \cdot U(y_2 - y_1).$$

As $U^r(\mathbf{y}_1) = U(\mathbf{y}_1)$, the unified (r, s) -entropy of the ordered sequence can be rewritten as

$$\begin{aligned} H_r^s(\mathbf{Y}) &= (2^{1-s} - 1)^{-1} \left\{ \left[\int_{\mathbf{R}^N} \left(U(\mathbf{y}_1) \sum_{j=1}^{N!} f(\mathbf{y}_j) \right)^r d\mathbf{y}_1 \right]^{\frac{s-1}{r-1}} - 1 \right\} \\ &= (2^{1-s} - 1)^{-1} \left\{ \left[\int_{\mathbf{R}^N} U(\mathbf{y}_1) \frac{(\sum_{j=1}^{N!} f(\mathbf{y}_j))^r \sum_{k=1}^{N!} f^r(\mathbf{y}_k)}{\sum_{j=1}^{N!} f^r(\mathbf{y}_j)} d\mathbf{y}_1 \right]^{\frac{s-1}{r-1}} - 1 \right\}. \end{aligned}$$

Interchanging the order of integration and summation, we obtain

$$H_r^s(\mathbf{Y}) = (2^{1-s} - 1)^{-1} \left\{ \left[\sum_{k=1}^{N!} \int_{\mathbf{R}^N} f^r(\mathbf{y}_k) U(\mathbf{y}_1) \left[\frac{(\sum_{j=1}^{N!} f(\mathbf{y}_j))^r}{\sum_{j=1}^{N!} f^r(\mathbf{y}_j)} \right] d\mathbf{y}_1 \right]^{\frac{s-1}{r-1}} - 1 \right\}.$$

If we change the variables in the k th integral from \mathbf{y}_k to \mathbf{y}_1 then $f(\mathbf{y}_k)$ becomes $f(\mathbf{y}_1)$ and $U(\mathbf{y}_1)$ will be transformed correspondingly to $U(\mathbf{y}_m)$ for some $m \in \{1, 2, \dots, N!\}$. The terms in the sums will remain unchanged since each \mathbf{y}_j corresponds to a distinct \mathbf{y}_n under the k th transformation ($k = 1, \dots, N!$). Furthermore, the Jacobian of the k th transformation is unity for every $k = 1, \dots, N!$. Thus, after interchanging the order of summation and integration we see that the (r, s) -entropy of the ordered sequence is given by

$$H_r^s(\mathbf{Y}) = (2^{1-s} - 1)^{-1} \left\{ \left[\int_{\mathbf{R}^N} f^r(\mathbf{y}_1) \left(\sum_{m=1}^{N!} U(\mathbf{y}_m) \right) \left[\frac{(\sum_{j=1}^{N!} f(\mathbf{y}_j))^r}{\sum_{j=1}^{N!} f^r(\mathbf{y}_j)} \right] d\mathbf{y}_1 \right]^{\frac{s-1}{r-1}} - 1 \right\}.$$

As $\sum_{m=1}^{N!} U(\mathbf{y}_m) = 1$ almost everywhere, we conclude that

$$H_r^s(\mathbf{Y}) = (2^{1-s} - 1)^{-1} \left\{ \left[\int_{\mathbf{R}^N} f^r(\mathbf{y}_1) \left[\frac{(\sum_{j=1}^{N!} f(\mathbf{y}_j))^r}{\sum_{j=1}^{N!} f^r(\mathbf{y}_k)} \right] d\mathbf{y}_1 \right]^{\frac{s-1}{r-1}} - 1 \right\}.$$

By taking limits we can easily check the corresponding expressions for $(r = 1, s \neq 1)$, $(r \neq 1, s = 1)$ and $(r = 1, s = 1)$. \square

Remark 1. The result obtained for $r = 1$ and $s = 1$ is the one proved by K.M. Wong and S. Chen [9].

Theorem 2. *The (r, s) -entropy of a sequence of any N random variables is decreased if the sequence is ordered. The decrease in entropy is given by*

$$0 \leq \mathcal{E}_r^s(\mathbf{X}) - \mathcal{E}_r^s(\mathbf{Y}) \leq \begin{cases} (1 - N!^{s-1})(H_r^s(\mathbf{X}) + (2^{1-s} - 1)^{-1}), & s \neq 1 \\ \log_2 N!, & s = 1. \end{cases}$$

Equality on the left hand side holds iff $f(\mathbf{y}_j) = 0, j = 2, \dots, N!$ almost everywhere in Ω , where Ω is the region in which $f(\mathbf{y}_j)$ is defined. Equality on the right hand side holds iff $f(\mathbf{y}_1) = f(\mathbf{y}_2) = \dots = f(\mathbf{y}_{N!})$.

Proof. Let $r \neq 1, s \neq 1$. Since

$$\frac{(\sum_{j=1}^{N!} f(\mathbf{y}_j))^r}{\sum_{j=1}^{N!} f^r(\mathbf{y}_j)} \begin{cases} \leq 1, & 0 < r < 1 \\ \geq 1, & r > 1, \end{cases}$$

multiplying by $f^r(\mathbf{y}_1)$ on both sides and integrating we get

$$\int_{\mathbb{R}^N} f^r(\mathbf{y}_1) \frac{(\sum_{j=1}^{N!} f(\mathbf{y}_j))^r}{\sum_{j=1}^{N!} f^r(\mathbf{y}_j)} d\mathbf{y}_1 \begin{cases} \leq \int_{\mathbb{R}^N} f^r(\mathbf{y}_1) d\mathbf{y}_1, & 0 < r < 1 \\ \geq \int_{\mathbb{R}^N} f^r(\mathbf{y}_1) d\mathbf{y}_1, & r > 1. \end{cases}$$

Let us consider the function

$$\eta(x) = (2^{1-s} - 1)^{-1} (x^{\frac{s-1}{r-1}} - 1), \quad r \neq 1, s \neq 1, r > 0.$$

It is easy to verify that η is increasing in $x > 0$ for $0 < r < 1$ and decreasing in $x > 0$ for $r > 1$. Thus, applying η to both sides we obtain

$$H_r^s(\mathbf{Y}) \leq H_r^s(\mathbf{X})$$

for all $r > 0$ and any s . Therefore, by continuity of $H_r^s(\mathbf{Y})$ and $H_r^s(\mathbf{X})$ with respect to r and s , we have

$$\mathcal{E}_r^s(\mathbf{Y}) \leq \mathcal{E}_r^s(\mathbf{X})$$

for all $r > 0$ and any s . Equality holds iff $f(\mathbf{y}_j) = 0 (j = 2, \dots, N!)$ almost everywhere in Ω , where Ω is the region in which $f(\mathbf{y}_j)$ is defined.

On the other hand, applying Jensen's inequality we have

$$\left(\frac{\sum_{j=1}^{N!} f(\mathbf{y}_j)}{N!} \right)^r \begin{cases} \geq \frac{\sum_{j=1}^{N!} f^r(\mathbf{y}_j)}{N!}, & 0 < r < 1 \\ \leq \frac{\sum_{j=1}^{N!} f^r(\mathbf{y}_j)}{N!}, & r > 1 \end{cases}$$

i.e.,

$$\frac{(\sum_{j=1}^{N!} f(\mathbf{y}_j))^r}{\sum_{j=1}^{N!} f^r(\mathbf{y}_j)} \begin{cases} \geq N!^{r-1}, & 0 < r < 1 \\ \leq N!^{r-1}, & r > 1. \end{cases}$$

Multiplying by $f^r(\mathbf{y}_1)$ on both sides and integrating we get

$$\int_{\mathbb{R}^N} f^r(\mathbf{y}_1) \frac{(\sum_{j=1}^{N!} f(\mathbf{y}_j))^r}{\sum_{j=1}^{N!} f^r(\mathbf{y}_j)} d\mathbf{y}_1 \begin{cases} \geq N!^{r-1} \int_{\mathbb{R}^N} f^r(\mathbf{y}_1) d\mathbf{y}_1, & 0 < r < 1 \\ \leq N!^{r-1} \int_{\mathbb{R}^N} f^r(\mathbf{y}_1) d\mathbf{y}_1, & r > 1. \end{cases}$$

Thus, applying η to both sides, we obtain

$$H_r^s(\mathbf{Y}) \geq N!^{s-1} H_r^s(\mathbf{X}) + (2^{1-s} - 1)^{-1} (N!^{s-1} - 1)$$

for all $r > 0$ ($r \neq 1$) and any s .

Therefore, by continuity of $H_r^s(\mathbf{Y})$ and $H_r^s(\mathbf{X})$ with respect to r and s , we obtain the announced result. Finally, equality holds iff

$$f(\mathbf{y}_1) = f(\mathbf{y}_2) = \dots = f(\mathbf{y}_{N!}).$$

□

3. THE (r, s) -ENTROPY MEASURE OF THE ORDER STATISTICS

Now we consider a sequence of N random variables X_1, \dots, X_N that are i.i.d. with p.d.f. $f(x)$. Then the (r, s) -entropy measure of each of the random variables is defined as

$$\mathcal{E}_r^s(X_i) = \begin{cases} H_r^s(X_i) = (2^{1-s} - 1)^{-1} \left[\left(\int_{\mathbb{R}} f(x)^r dx \right)^{\frac{s-1}{r-1}} - 1 \right], & r > 0, r \neq 1, s \neq 1, \\ H_1^s(X_i) = (2^{1-s} - 1)^{-1} \left\{ \exp_2 \left((s-1) \int_{\mathbb{R}} f(x) \log_2 f(x) dx \right) - 1 \right\}, & r = 1, s \neq 1, \\ H_r^1(X_i) = (1-r)^{-1} \log_2 \int_{\mathbb{R}} f(x)^r dx, & r > 0, r \neq 1, s = 1, \\ H(X_i) = - \int_{\mathbb{R}} f(x) \log_2 f(x) dx, & r = 1, s = 1. \end{cases}$$

The p.d.f. of the k th-order statistic Y_k , after the sequence X_1, \dots, X_N has been observed, is given by

$$f_k(y) = N \binom{N-1}{k-1} [F(y)]^{k-1} [1-F(y)]^{N-k} f(y).$$

Note that we now write $f_k(y)$ instead of $f_k^*(y)$. We define the entropy of the k th-order statistic Y_k as

$$\mathcal{E}_r^s(Y_k) = \begin{cases} H_r^s(Y_k) = (2^{1-s} - 1)^{-1} \left[\left(\int_{\mathbb{R}} f_k(y)^r dy \right)^{\frac{s-1}{r-1}} - 1 \right], & r > 0, r \neq 1, s \neq 1, \\ H_1^s(Y_k) = (2^{1-s} - 1)^{-1} \left\{ \exp_2 \left((s-1) \int_{\mathbb{R}} f_k(y) \log_2 f_k(y) dy \right) - 1 \right\}, & r = 1, s \neq 1, \\ H_r^1(Y_k) = (1-r)^{-1} \log_2 \int_{\mathbb{R}} f_k(y)^r dy, & r > 0, r \neq 1, s = 1, \\ H(Y_k) = - \int_{\mathbb{R}} f_k(y) \log_2 f_k(y) dy, & r = 1, s = 1, \end{cases}$$

and we also define, see Wong [9], the average unified (r, s) -entropy of the order statistics Y_k as

$$\overline{\mathcal{E}}_r^s(Y) = \frac{1}{N} \sum_{k=1}^N \mathcal{E}_r^s(Y_k).$$

In the following theorem we establish an upper bound for the difference between the average unified (r, s) -entropy measure of the order statistics and the unified (r, s) -entropy measure of a member of the original random variables.

Theorem 3. Consider N i.i.d. random variables X_1, \dots, X_N . Then

$$\mathcal{E}_r^s(X) - \overline{\mathcal{E}}_r^s(Y) \begin{cases} \leq (1 - N^{s-1})(H_r^s(X) + (2^{1-s} - 1)^{-1}), & r \geq s, s \neq 1 \\ \leq \log_2 N, & r \geq s, s = 1. \end{cases}$$

Proof. Let $s \neq 1, r \neq 1$. Consider the random variable Z , taking the values

$$z_k = \int_{\mathbb{R}} N^r \binom{N-1}{k-1}^r (F(y)^{k-1} (1-F(y))^{N-k})^r f(y)^r dy, \quad k = 1, \dots, N$$

with probability $\frac{1}{N}$, and consider the function

$$\Phi(t) = t^{\frac{s-1}{r-1}}, \quad s \neq 1, r \neq 1.$$

Then

$$E[\Phi(Z)] = \frac{1}{N} \sum_{k=1}^N \left\{ \int_{\mathbb{R}} N^r \binom{N-1}{k-1}^r (F(y)^{k-1} (1-F(y))^{N-k})^r f(y)^r dy \right\}^{\frac{s-1}{r-1}},$$

$$\Phi(E[Z]) = \left\{ \frac{1}{N} \sum_{k=1}^N \int_{\mathbb{R}} N^r \binom{N-1}{k-1}^r (F(y)^{k-1} (1-F(y))^{N-k})^r f(y)^r dy \right\}^{\frac{s-1}{r-1}}.$$

Applying Jensen's inequality, we have

$$\begin{aligned} & \left\{ \frac{1}{N} \sum_{k=1}^N \int_{\mathbb{R}} N^r \binom{N-1}{k-1}^r (F(y)^{k-1} (1-F(y))^{N-k})^r f(y)^r dy \right\}^{\frac{s-1}{r-1}} \\ & \leq \frac{1}{N} \sum_{k=1}^N \left\{ \int_{\mathbb{R}} N^r \binom{N-1}{k-1}^r (F(y)^{k-1} (1-F(y))^{N-k})^r f(y)^r dy \right\}^{\frac{s-1}{r-1}}, \end{aligned}$$

when $\frac{s-1}{r-1} < 0$ or $\frac{s-1}{r-1} > 1$. Subtracting 1, multiplying by $(2^{1-s} - 1)^{-1}$ ($s \neq 1$) on both sides and simplifying, we get

$$\begin{aligned} \frac{1}{2^{1-s} - 1} \left[\left\{ \frac{1}{N} \sum_{k=1}^N \int_{\mathbb{R}} N^r \binom{N-1}{k-1}^r (F(y)^{k-1} (1-F(y))^{N-k})^r f(y)^r dy \right\}^{\frac{s-1}{r-1}} - 1 \right] \\ \leq \overline{H}_r^s(Y), \quad r > s. \end{aligned}$$

As

$$\begin{aligned} & \sum_{k=1}^N \left[N \binom{N-1}{k-1} F(y)^{k-1} (1-F(y))^{N-k} f(y) \right]^r \\ & \begin{cases} \leq \left[\sum_{k=1}^N N \binom{N-1}{k-1} F(y)^{k-1} (1-F(y))^{N-k} f(y) \right]^r, & r > 1, \\ \geq \left[\sum_{k=1}^N N \binom{N-1}{k-1} F(y)^{k-1} (1-F(y))^{N-k} f(y) \right]^r, & 0 < r < 1 \end{cases} \end{aligned}$$

and $\sum_{k=1}^N N \binom{N-1}{k-1} F(y)^{k-1} (1-F(y))^{N-k} f(y) = N f(y)$, dividing by N on both sides and integrating over \mathbb{R} we get

$$\frac{1}{N} \sum_{k=1}^N \int_{\mathbb{R}} \left[N \binom{N-1}{k-1} F(y)^{k-1} (1-F(y))^{N-k} f(y) \right]^r dy \begin{cases} \leq \frac{1}{N} \int_{\mathbb{R}} N^r f(y)^r dy, & r > 1, \\ \geq \frac{1}{N} \int_{\mathbb{R}} N^r f(y)^r dy, & 0 < r < 1. \end{cases}$$

Thus, applying η to both sides, we obtain

$$\begin{aligned} & \frac{1}{2^{1-s} - 1} \left[\left\{ \frac{1}{N} \sum_{k=1}^N \int_{\mathbb{R}} N^r \binom{N-1}{k-1}^r (F(y)^{k-1} (1-F(y))^{N-k})^r f(y)^r dy \right\}^{\frac{s-1}{r-1}} - 1 \right] \\ & \geq \frac{1}{2^{1-s} - 1} \left\{ \left[N^{r-1} \int_{\mathbb{R}} f(y)^r dy \right]^{\frac{s-1}{r-1}} - 1 \right\} \\ & = \frac{N^{s-1}}{2^{1-s} - 1} \left\{ \left[\int_{\mathbb{R}} f(y)^r dy \right]^{\frac{s-1}{r-1}} - 1 \right\} + \frac{N^{s-1} - 1}{2^{1-s} - 1} \\ & = N^{s-1} H_r^s(X) + \frac{N^{s-1} - 1}{2^{1-s} - 1}, \quad s \in \mathbb{R}, r \in \mathbb{R}. \end{aligned}$$

Therefore we conclude that

$$\overline{H}_r^s(Y) \geq N^{s-1} H_r^s(X) + \frac{N^{s-1} - 1}{2^{1-s} - 1}, \quad r > s$$

and the result follows by taking appropriate limits for r and s . \square

Theorem 4. Suppose we have N i.i.d. random variables X_1, \dots, X_N . Then

$$\mathcal{E}_r^s(X) - \overline{\mathcal{E}}_r^s(Y) \geq 0 \quad \forall (r, s) \in \left\{ (r, s)/r > 0, s \geq 2 - \frac{1}{r} \right\}.$$

Proof. Let us define $h_k(y) = N \binom{N-1}{k-1} F(y)^{k-1} (1 - F(y))^{N-k}$. From Minkowski's inequality we get

$$\left(\int_{\mathbf{R}} \left(\sum_{k=1}^N \frac{1}{N} h_k(y) f(y) \right)^r dy \right)^{1/r} \begin{cases} \leq \sum_{k=1}^N \left(\int_{\mathbf{R}} \left(\frac{1}{N} \right)^r h_k(y)^r f(y)^r dy \right)^{1/r} = \sum_{k=1}^N \frac{1}{N} \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{1/r}, & r > 1, \\ \geq \sum_{k=1}^N \frac{1}{N} \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{1/r}, & 0 < r < 1. \end{cases}$$

Then

$$(3.1) \quad \left(\int_{\mathbf{R}} \left(\sum_{k=1}^N \frac{1}{N} h_k(y) f(y) \right)^r dy \right)^{\frac{s-1}{r-1}} \begin{cases} \leq \left(\sum_{k=1}^N \frac{1}{N} \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{1/r} \right)^{r \frac{s-1}{r-1}}, & (r > 1, \frac{s-1}{r-1} > 0) \text{ or } (0 < r < 1, \frac{s-1}{r-1} < 0) \\ \geq \left(\sum_{k=1}^N \frac{1}{N} \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{1/r} \right)^{r \frac{s-1}{r-1}}, & (r > 1, \frac{s-1}{r-1} < 0) \text{ or } (0 < r < 1, \frac{s-1}{r-1} > 0). \end{cases}$$

Now we consider the function

$$t(x) = x^{r \frac{s-1}{r-1}}, \quad r \neq 1.$$

This function is convex when $r \frac{s-1}{r-1} > 1$ or $r \frac{s-1}{r-1} < 0$, and concave when $0 < r \frac{s-1}{r-1} < 1$. If we consider the random variable Z , taking the values

$$z_k = \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{1/r}, \quad k = 1, 2, \dots, N$$

with probabilities $\frac{1}{N}$, it follows that

$$E[Z] = \frac{1}{N} \sum_{k=1}^N \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{1/r},$$

$$t(E[Z]) = \left(\frac{1}{N} \sum_{k=1}^N \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{1/r} \right)^{r \frac{s-1}{r-1}},$$

$$t(Z) = \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{\frac{r-1}{s-1}},$$

and

$$E[t(Z)] = \frac{1}{N} \sum_{k=1}^N \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{\frac{s-1}{r-1}}.$$

Applying Jensen's inequality, we have

$$\left(\frac{1}{N} \sum_{k=1}^N \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{1/r} \right)^{r \frac{s-1}{r-1}}$$

$$\begin{cases} \leq \frac{1}{N} \sum_{k=1}^N \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{\frac{s-1}{r-1}}, & r \frac{s-1}{r-1} > 1 \text{ or } r \frac{s-1}{r-1} < 0 \\ \geq \frac{1}{N} \sum_{k=1}^N \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{\frac{s-1}{r-1}}, & 0 < r \frac{s-1}{r-1} < 1. \end{cases}$$

3.2

Combining inequalities (3.1) and (3.2), we get

$$\left(\int_{\mathbf{R}} \left(\sum_{k=1}^N \frac{1}{N} h_k(y) f(y) \right)^r dy \right)^{\frac{s-1}{r-1}}$$

$$\begin{cases} \leq \frac{1}{N} \sum_{k=1}^N \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{\frac{s-1}{r-1}}, & s \geq 1, s > 2 - \frac{1}{r} \\ \geq \frac{1}{N} \sum_{k=1}^N \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{\frac{s-1}{r-1}}, & s < 1, s > 2 - \frac{1}{r}. \end{cases}$$

Subtracting 1 and multiplying by $(2^{1-s} - 1)^{-1}$ ($s \neq 1$) on both sides of the inequality, we have

$$H_r^s(X) = (2^{1-s} - 1)^{-1} \left(\int_{\mathbf{R}} \left(\sum_{k=1}^N \frac{1}{N} h_k(y) f(y) \right)^r dy \right)^{\frac{s-1}{r-1}}$$

$$\geq \frac{1}{N} \sum_{k=1}^N \frac{1}{2^{1-s} - 1} \left\{ \left(\int_{\mathbf{R}} h_k(y)^r f(y)^r dy \right)^{\frac{s-1}{r-1}} - 1 \right\} = \overline{H}_r^s(Y)$$

for all $(r, s) \in \{(r, s)/r > 0, s \geq 2 - \frac{1}{r}\}$. Therefore, by continuity of $H_r^s(X)$ and $\overline{H}_r^s(Y)$ with respect to r and s , we get

$$\mathcal{E}_r^s(X) - \overline{\mathcal{E}}_r^s(Y) \geq 0 \quad \forall (r, s) \in \{(r, s)/r > 0, s \geq 2 - \frac{1}{r}\}.$$

□

Now we study some properties of the (r, s) -entropy measure of the order statistics when the p.d.f. of the original i.i.d. random variables is symmetric about the mean.

Theorem 5. For an i.i.d. sequence of random variables X_1, \dots, X_N , N being odd, whose members have a p.d.f. that is symmetric about its mean μ , the (r, s) -entropy measure of the order statistics has the following properties:

- (1) $\mathcal{E}_r^s(Y_k) = \mathcal{E}_r^s(Y_{N-k+1})$,
- (2) $\frac{\partial \mathcal{E}_r^s(Y_k)}{\partial k} = -\frac{\partial \mathcal{E}_r^s(Y_{N-k+1})}{\partial k}$,
- (3) $\frac{\partial \mathcal{E}_r^s(Y_k)}{\partial k} = 0$ if $k = \frac{N+1}{2}$,

(4) Let $f(x)$ be the original p.d.f. of the random variables X_1, \dots, X_N and $\mathcal{E}_r^s(Y_k)$ the (r, s) -entropy measure of the k th-order statistic. Let us define $X_n^* = aX_n + b$ ($a > 0$), then the corresponding (r, s) -entropy measure of the k th-order statistic is given by

$$\mathcal{E}_r^s(Y_k^*) = \begin{cases} a^{1-s} H_r^s(Y_k) + \frac{a^{1-s} - 1}{2^{1-s} - 1}, & s \neq 1 \\ H_r^1(Y_k) + \log_2 a, & s = 1. \end{cases}$$

Proof. (1) Since

$$H_r^s(Y_{N-k+1}) = (2^{1-s} - 1)^{-1} \left\{ \left(\int_{\mathbb{R}} f_{N-k+1}^r(y) dy \right)^{\frac{s-1}{r-1}} - 1 \right\},$$

taking $y = \mu + z$ we get

$$H_r^s(Y_{N-k+1}) = (2^{1-s} - 1)^{-1} \left\{ \left(\int_{\mathbb{R}} f_{N-k+1}^r(\mu + z) dz \right)^{\frac{s-1}{r-1}} - 1 \right\}.$$

As the p.d.f. of the original i.i.d. random variables is symmetric about the mean, we have

$$f_k(\mu + y) = f_{N-k+1}(\mu - y).$$

Therefore

$$\begin{aligned} H_r^s(Y_{N-k+1}) &= (2^{1-s} - 1)^{-1} \left\{ \left(\int_{\mathbb{R}} f_{N-k+1}^r(\mu - z) dz \right)^{\frac{s-1}{r-1}} - 1 \right\} \\ &= (2^{1-s} - 1)^{-1} \left\{ \left(\int_{\mathbb{R}} f_k^r(t) dt \right)^{\frac{s-1}{r-1}} - 1 \right\} = H_r^s(Y_k). \end{aligned}$$

By continuity of H_r^s with respect to r and s we get

$$\mathcal{E}_r^s(Y_k) = \mathcal{E}_r^s(Y_{N-k+1}).$$

(2) Using the definition of $H_r^s(Y_k)$ and differentiating with respect to k , we have

$$\frac{\partial H_r^s(Y_k)}{\partial k} = (2^{1-s} - 1)^{-1} \left\{ \left(\int_{\mathbb{R}} f_k^r(y) dy \right)^{\frac{s-1}{r-1}-1} \frac{s-1}{r-1} r \int_{\mathbb{R}} f_k^{r-1}(y) \frac{\partial f_k(y)}{\partial k} dy \right\}.$$

We know that (Wong and Chen [9], p. 281)

$$\frac{\partial f_k(y)}{\partial k} = \left(-D(k) + \log \frac{F(y)}{1-F(y)} \right) f_k(y)$$

where

$$D(k) = \sum_{i=1}^{k-1} \frac{1}{i} - \sum_{i=1}^{N-k} \frac{1}{i}$$

with the property that $D(N-k+1) = -D(k)$ for $k < \frac{N+1}{2}$.

Suppose $k_0 \leq \frac{N+1}{2}$. Let $y = \mu + z$, then

$$\begin{aligned} \left[\frac{\partial H_r^s(Y_k)}{\partial k} \right]_{k=k_0} &= \frac{(s-1)r}{(2^{1-s}-1)(r-1)} \left\{ \left(\int_{\mathbb{R}} f_{k_0}^r(\mu+z) dz \right)^{\frac{s-1}{r-1}-1} \right. \\ &\quad \left. \int_{\mathbb{R}} f_{k_0}^r(\mu+z) \left(-D(k_0) + \log \frac{F(\mu+z)}{1-F(\mu+z)} \right) dy \right\} \\ &= \frac{(s-1)r}{(2^{1-s}-1)(r-1)} \left\{ \left(\int_{\mathbb{R}} f_{N-k_0+1}^r(\mu-z) dz \right)^{\frac{s-1}{r-1}-1} \right. \\ &\quad \left. \int_{\mathbb{R}} f_{N-k_0+1}^r(\mu-z) \left(D(N-k_0+1) - \log \frac{1-F(\mu-z)}{F(\mu-z)} \right) dy \right\} \\ &= \left[\frac{\partial H_r^s(Y_{N-k+1})}{\partial k} \right]_{k=k_0}. \end{aligned}$$

By continuity of $H_r^s(Y_k)$ with respect to r and s we get

$$\frac{\partial \mathcal{E}_r^s(Y_k)}{\partial k} = -\frac{\partial \mathcal{E}_r^s(Y_{N-k+1})}{\partial k}.$$

(3) If $k = \frac{N+1}{2}$, then (2)

$$\frac{\partial \mathcal{E}_r^s(Y_{\frac{N+1}{2}})}{\partial k} = -\frac{\partial \mathcal{E}_r^s(Y_{\frac{N+1}{2}})}{\partial k}.$$

Therefore

$$\frac{\partial \mathcal{E}_r^s(Y_k)}{\partial k} = 0 \text{ if } k = \frac{N+1}{2}.$$

(4) If $g(x)$ and $G(x)$ are the p.d.f. and the c.d.f. of X_n^* , respectively, $g(x) = \frac{1}{a} f\left(\frac{x-b}{a}\right)$ and $G(x) = F\left(\frac{x-b}{a}\right)$, where F is the c.d.f. of the random variables X_1, \dots, X_N , then

$$\begin{aligned} g_k(x) &= N \binom{N-1}{k-1} G^{k-1}(x) [1 - G(x)]^{N-k} g(x) \\ &= N \binom{N-1}{k-1} F^{k-1}\left(\frac{x-b}{a}\right) \left[1 - F\left(\frac{x-b}{a}\right)\right]^{N-k} \frac{1}{a} f\left(\frac{x-b}{a}\right) = \frac{1}{a} f_k\left(\frac{x-b}{a}\right). \end{aligned}$$

Therefore

$$\begin{aligned} H_r^s(Y_k^*) &= (2^{1-s} - 1)^{-1} \left\{ \left(\int_{\mathbb{R}} \frac{1}{a^r} f_k^r\left(\frac{x-b}{a}\right) dx \right)^{\frac{s-1}{r-1}} - 1 \right\} \\ &= (2^{1-s} - 1)^{-1} \left\{ \left(\int_{\mathbb{R}} \frac{1}{a^{r-1}} f_k^r(y) dy \right)^{\frac{s-1}{r-1}} - 1 \right\} \\ &= \frac{(2^{1-s} - 1)^{-1}}{a^{s-1}} \left\{ \left(\int_{\mathbb{R}} f_k^r(y) dy \right)^{\frac{s-1}{r-1}} - 1 \right\} + (2^{1-s} - 1)^{-1} \left(\frac{1}{a^{s-1}} - 1 \right) \\ &= a^{1-s} H_r^s(Y_k) + \frac{a^{1-s} - 1}{2^{1-s} - 1}. \end{aligned}$$

Now the result follows by continuity of H_r^s with respect to r and s . □

An example illustrates the result obtained above.

4. NUMERICAL EXAMPLE

Consider the logistic p.d.f. with location parameter μ and scale parameter a , i.e.

$$f(x/\mu, a) = \frac{1}{a} \frac{\exp\left(-\frac{x-\mu}{a}\right)}{\left(1 + \exp\left(-\frac{x-\mu}{a}\right)\right)^2}, \quad x \in \mathbb{R}, \mu \in \mathbb{R}, a > 0,$$

so that the variation of the parameter a results in a family of p.d.f.'s symmetric about μ . It can be easily checked that

$$E(X/\mu, a) = \mu, \quad V(X/\mu, a) = \frac{a^2 \pi^2}{3}, \quad F(x/\mu, a) = \left(1 + \exp\left(-\frac{x-\mu}{a}\right)\right)^{-1}$$

and

$$f_{Y_k}(x/\mu, a) = \frac{1}{a} N \binom{N-1}{k-1} \exp\left(-\frac{(N-k+1)(x-\mu)}{a}\right) \left(1 + \exp\left(-\frac{x-\mu}{a}\right)\right)^{-N-1}.$$

Let us define $H(s, r, k, N, a, \mu) = H_r^s(Y_k)$, where Y_k is the k th-order statistics of a sequence of N random variables that are i.i.d. with p.d.f. $f(\cdot/\mu, a)$. First, let us suppose $\mu = 0$ and $a = 1$. Gradshtein [3], p. 305, shows that

$$\int_{-\infty}^{+\infty} \frac{e^{-\mu x} dx}{(e^{\beta/\gamma} + e^{-x/\gamma})^\nu} = \gamma \exp \left\{ \beta \left(\mu - \frac{\nu}{\gamma} \right) \right\} \frac{\Gamma(\gamma\mu)\Gamma(\nu - \gamma\mu)}{\Gamma(\nu)},$$

provided $\operatorname{Re}(\frac{\nu}{\gamma}) > \operatorname{Re}(\mu) > 0$ and $|\operatorname{Im}(\beta)| < \pi \operatorname{Re}(\gamma)$. Taking $\mu = (N - k + 1)$, $\nu = (N + 1)r$, $\beta = 0$ and $\gamma = 1$, we get

$$\begin{aligned} \int_{\mathbf{R}} f_{Y_k}^r(x) dx &= N^r \binom{N-1}{k-1}^r \int_{\mathbf{R}} \frac{e^{-(N-k+1)rx} dx}{(1 + e^{-x})^{(N+1)r}} \\ &= N^r \binom{N-1}{k-1}^r \frac{\Gamma(r(N-k+1))\Gamma(rk)}{\Gamma(r(N+1))}, \end{aligned}$$

which yields

$$H(s, r, k, N, 1, 0) = (2^{1-s} - 1)^{-1} \left\{ \left(\frac{N^r \binom{N-1}{k-1}^r \Gamma(r(N-k+1))\Gamma(rk)}{\Gamma(r(N+1))} \right)^{\frac{s-1}{r-1}} - 1 \right\},$$

$s \neq 1, r \neq 1,$

and

$$H(1, r, k, N, 1, 0) = \frac{1}{(1-r)\log 2} \log \left\{ \frac{N^r \binom{N-1}{k-1}^r \Gamma(r(N-k+1))\Gamma(rk)}{\Gamma(r(N+1))} \right\},$$

$r \neq 1,$

where $\log = \log_e$ and $\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx$. Furthermore

$$\begin{aligned} H(1, 1, k, N, 1, 0) &= \frac{-1}{\log 2} \int_{-\infty}^{+\infty} f_{Y_k}(x) \log f_{Y_k}(x) dx \\ &= \frac{1}{\log 2} \left\{ -\log \left[N \binom{N-1}{k-1} \right] + (N-k+1) \int_{-\infty}^{+\infty} x f_{Y_k}(x) dx \right. \\ &\quad \left. + (N+1) \int_{-\infty}^{+\infty} \log(1 + e^{-x}) f_{Y_k}(x) dx \right\}. \end{aligned}$$

Now, from the relation

$$\Psi(n) = -\gamma + \sum_{i=1}^{n-1} \frac{1}{i},$$

where $\Psi(z) = \frac{d}{dz} \log \Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}$ is the psi (or digamma) function and $\gamma = -\int_0^\infty e^{-x} \log x \, dx = 0.577, 215 \dots$ is Euler's constant, we obtain after substituting $u = \frac{1}{1+e^{-x}}$ and $u = -\log(1 + e^{-x})$ in the integrals $\int_{-\infty}^{+\infty} x f_{Y_k}(x) \, dx$ and $\int_{-\infty}^{+\infty} \log(1 + e^{-x}) f_{Y_k}(x) \, dx$, respectively, that

$$H(1, 1, k, N, 1, 0) = \frac{1}{\log 2} \left\{ -\log \left[N \binom{N-1}{k-1} \right] + (N-k+1) \left(\sum_{i=1}^{k-1} \frac{1}{i} - \sum_{i=1}^{N-k} \frac{1}{i} \right) \right. \\ \left. + (N-1) N \binom{N-1}{k-1} \sum_{j=0}^{N-k} \binom{N-k}{j} (-1)^{N-k-j} \frac{1}{(N-j)^2} \right\}$$

and

$$H(s, 1, k, N, 1, 0) = (2^{1-s} - 1)^{-1} \{ \exp_2((1-s)H(1, 1, k, N, 1, 0)) - 1 \}, \quad s \neq 1.$$

For more results about order statistics of the logistic distribution see the book of Balakrishnan N. and Cohen A.C. [2].

Finally, we have

$$H(s, r, k, N, a, \mu) = \begin{cases} a H(s, r, k, 1, 0) + \frac{a^{1-s} - 1}{2^{1-s} - 1}, & s \neq 1 \\ H(s, r, k, 1, 0) + \frac{\log a}{\log 2}, & s = 1. \end{cases}$$

The (r, s) -entropies $H(s, r, k, N, a, \mu)$ of the order statistics are plotted in Figures 1 and 2 for various values of a . The number of input samples N is taking to be seven. The median ($k = 4$) is the point that has globally maximum entropy and the first-order and seventh-order statistics have globally minimum entropy.

ACKNOWLEDGEMENTS

The authors thank the referee for his valuable comments.

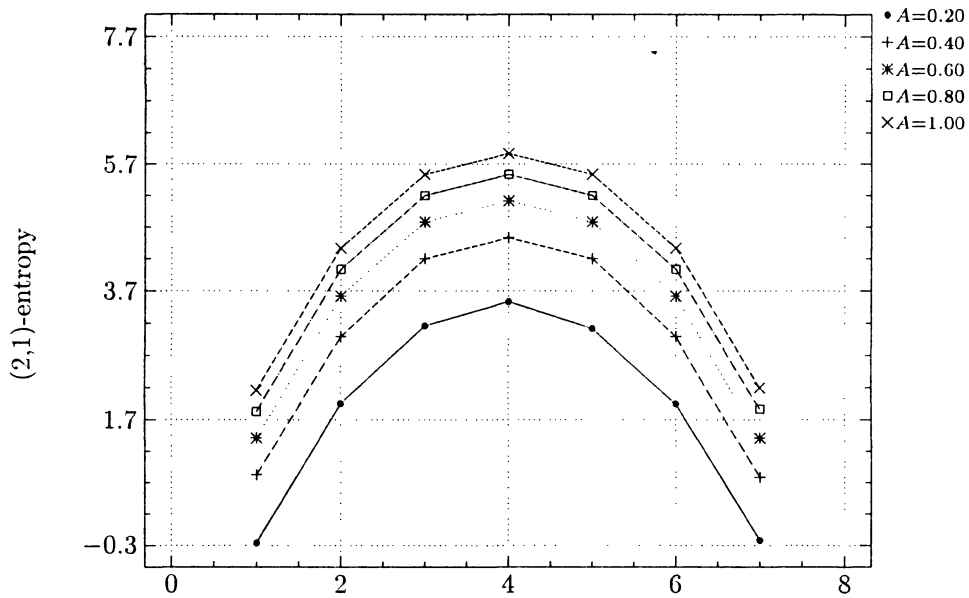


Figure 1. (r, s) -entropy of the order statistics with logistic distribution

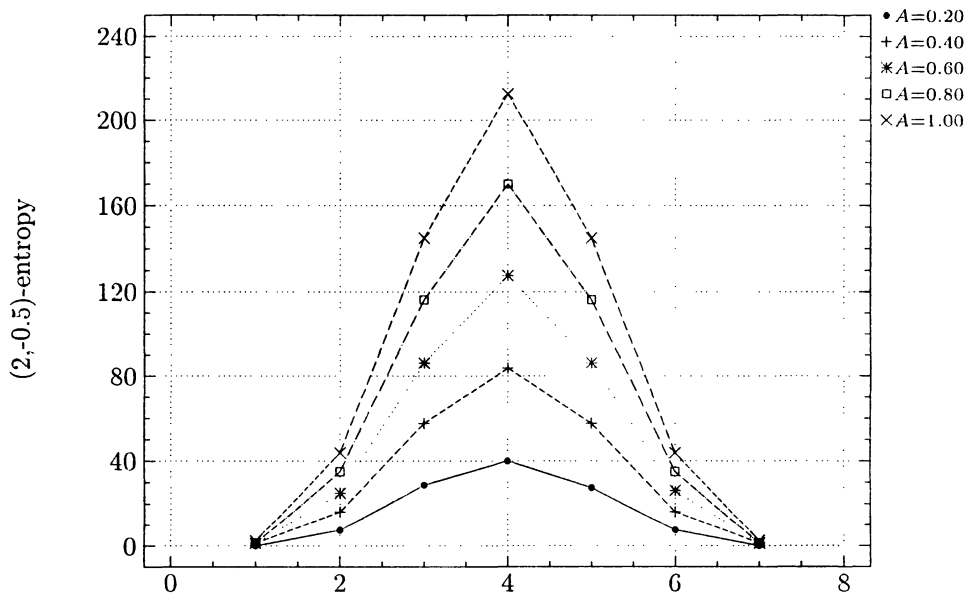


Figure 2. (r, s) -entropy of the order statistics with logistic distribution

References

- [1] *S. Arimoto*: Information-theoretic consideration on estimation problems. *Information and Control* 19 (1971), 181–194.
- [2] *N. Balakrishnan and A. C. Cohen*: Order statistics and inference, Estimation methods. Academic Press, 1991.
- [3] *I. S. Gradshteyn and I. M. Ryzhik*: Table of integrals, series and products. Academic Press, 1980.
- [4] *I. Havrda and F. Charvat*: Quantification method of classification processes: concept of structural α -entropy. *Kybernetika* 3 (1967), 30–35.
- [5] *A. Renyi*: On measures of entropy and information. Proc. 4th Berkeley Symp. Math. Statist. and Prob. 1 (1961), 547–561.
- [6] *C. E. Shannon*: A mathematical theory of communications. *Bell. Syst. Tech. J.* 27 (1948), 379–423.
- [7] *B. D. Sharma and D. P. Mittal*: New nonadditive measures of entropy for discrete probability distribution. *J. Math. Sci.* 10 (1975), 28–40.
- [8] *I. J. Taneja*: On generalized information measures and their applications. *Adv. Elect. and Elect. Phis.* 76 (1989), 327–413.
- [9] *K. M. Wong and S. Chen*: The entropy of ordered sequences and order statistics. *IEEE Transactions on Information Theory* 36(2) (1990), 276–284.

Authors' addresses: M. D. Esteban, D. Morales, L. Pardo, Departamento de Estadística e I.O., Facultad de Matemáticas, Universidad Complutense de Madrid, 28040 Madrid (Spain); *M. L. Menéndez*, Departamento de Matemática Aplicada, E.T.S. de Arquitectura, Universidad Politécnica de Madrid, 28040 Madrid (Spain).