SHANNON ENTROPY IN ORDER STATISTICS AND THEIR CONCOMITANS FROM BIVARIATE NORMAL DISTRIBUTION

M. NAGHAVY[1], M. MADADI[2] AND V. AMIRZADEH[3]
[1]YOUNG RESEARCHERS SOCIETY,
[2]MAHANI MATHEMATICAL RESEARCH CENTER,
[3]DEPARTMENT OF STATISTICS,
SHAHID BAHONAR UNIVERSITY OF KERMAN, KERMAN, I.R.IRAN.

E-MAIL: MADADI@UK.AC.IR

(Received: 30 May 2012, Accepted: 28 October 2012)

Abstract. In this paper, we derive first some results on the Shannon entropy in order statistics and their concomitants arising from a sequence of $\{(X_i, Y_i) : i = 1, 2, ...\}$ independent and identically distributed (iid) random variables from the bivariate normal distribution and extend our results to a collection $\mathfrak{C}(X, Y) = \{(X_{r_1:n}, Y_{[r_1:n]}), (X_{r_2:n}, Y_{[r_2:n]}), ..., (X_{r_k:n}, Y_{[r_k:n]})\}$ of order statistics and their concomitants. We finally compute the value of the Shannon entropy in order statistics and their concomitants from a bivariate normal distribution.

AMS Classification: 94A15.
Keywords: Bivariate Normal Distribution; Concomitants of Order Statistics; Shannon Entropy.

1. Introduction

Let $\{(X_i, Y_i) : i = 1, 2, ...\}$ be a sequence of random vectors from a bivariate continuous distribution. If we arrange the X-values in ascending order, the

JOURNAL OF MAHANI MATHEMATICAL RESEARCH CENTER
VOL. 1, NUMBER (2) 111- 118.

corresponding Y-values are called the concomitants of the relevant order statistic. Concomitants of order statistics arise in several applications. In selection procedures, items or subjects may be chosen on the basis of their X characteristic, and an associated characteristic Y that is hard to measure or can be observed only later may be of interest. For example, X may be the score of a candidate on a screening test, and Y the measure of his/her final performance. Concomitants of order statistics have also been used in dealing with the estimation of parameters on the basis of multivariate data sets that are subject to some form of censoring. A comprehensive review of these applications may be found in David and Nagaraja (1998) and Section 9.8 and 11.7 of David and Nagaraja (2003).

The first study of uncertainty (information) measures was undertaken by Nyquist (1924) and Hartley (1928), although the concept was introduced by Clausius in the year 1850 in the context of classical thermodynamics. Later Shannon (1946) studied the properties of information sources and the communication channels used to transmit their output and defined an entropy known as Shannon entropy. The Shannon entropy of an absolutely continuous random variable X having probability density function (pdf) $f_X(x)$, is defined as:

$$H(X) = -\int_{-\infty}^{+\infty} f_X(x) \ln f_X(x) dx, \tag{1}$$

where “ln“ stands for the natural logarithm.
The Shannon entropy of a random variable X is a mathematical measure of information which measures the average reduction of uncertainty of X. Because of its descriptive character, analytical expressions for univariate distributions have been obtained, among others, by Lazo and Rathie (1978) and Cover and Thomas (1991). Recently, several authors have investigated Shannon entropy in record values (Madadi and Tata, 2011). The concept of entropy may be successfully used for quantifying the amount of information regarding the parent distribution that one may obtain by observing an additional record value. In this paper, we compute Shannon entropy in order statistics and their concomitants from a bivariate normal distribution.

2. Entropy of Order Statistics and Their Concomitants

Definition 2.1 Let $X_1, ..., X_n$ be a random sample from a continuous population with cumulative distribution function (cdf) $F_X(x)$ and pdf $f_X(x)$ and $(X_1, Y_1), ...,$ (X_n, Y_n) be a random sample from a joint distribution with cdf $F_{X,Y}(x, y)$ and pdf $f_{X,Y}(x, y)$. We denote the r-th order statistic corresponding to the X-values by $X_{r:n}$ and the corresponding concomitant by $Y_{[r:n]}$. Then, the joint distribution of $(X_{r:n}, Y_{[r:n]})$; $r = 1, 2, ..., n$ is given by [1]:

$$f_{X_{r:n},Y_{[r:n]}}(x, y) = \frac{n!}{(r-1)!(n-r)!}\Big(F_X(x)\Big)^{r-1}\Big(1 - F_X(x)\Big)^{n-r} f_{X,Y}(x, y) \tag{2}$$

Also, the joint pdf of $\mathfrak{C}(X, Y) = \{(X_{r_1:n}, Y_{[r_1:n]}), (X_{r_2:n}, Y_{[r_2:n]}), ..., (X_{r_k:n}, Y_{[r_k:n]})\}$ is [1]:

$$\begin{aligned} f(x_1, .., x_k; y_1, .., y_k) &= \frac{n!}{(r_1-1)!(n-r_k)!}\Big(F_X(x_1)\Big)^{r_1-1}\Big(1 - F_X(x_k)\Big)^{n-r_k} \\ &\times\prod_{i=2}^{k}\frac{\Big(F_X(x_i) - F_X(x_{i-1})\Big)^{r_i-r_{i-1}-1}}{(\,r_i - r_{i-1} - 1)!}\prod_{i=1}^{k} f_{X,Y}(x_i, y_i). \end{aligned} \tag{3}$$

Theorem 2.1 The entropy of $(X_{r:n}, Y_{[r:n]})$ for the bivariate normal distribution, $N_2(0, 0, 1, 1, \rho)$, is:

$$\begin{aligned} H(X_{r:n}, Y_{[r:n]}) &= -\ln c_r + c_r(r-1)I(r:n) + c_r(n-r)I(n-r+1:n) + \frac{1}{2} \\ &+ \ln(2\pi\sqrt{1-\rho^2}) + \frac{c_r}{2}\sum_{m=0}^{n-r}\binom{n-r}{m}(-1)^m\eta(r+m-1), \end{aligned}$$

where

$$\eta(j) = \int_{-\infty}^{+\infty} x^2\Big[\Phi(x)\Big]^j\phi(x)dx,$$

$$I(j:n) = \sum_{m=0}^{n-j}\binom{n-j}{m}\frac{(-1)^m}{(m+j)^2}, \quad c_j = \frac{n!}{(j-1)!(n-j)!}.$$

and $\Phi(x) = \int_{-\infty}^{x}\phi(t)dt$ is the cdf of the standard normal variable.

Proof. Suppose $(X, Y) \sim N_2(0, 0, 1, 1, \rho)$. Then, the pdf of (X, Y) is

$$f_{X,Y}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}}\exp\left\{-\frac{(x^2 - 2\rho xy + y^2)}{2(1-\rho^2)}\right\}.$$

From (1) and (2) we have

$$\begin{aligned} H(X_{r:n}, Y_{[r:n]}) &= E\Big(-\ln f(X_{r:n}, Y_{[r:n]})\Big) \\ &= -\ln c_r + (r-1)E_1 + (n-r)E_2 + E_3, \end{aligned} \tag{4}$$

where

$$E_1 = E\Big(-\ln F_X(X_{j:n})\Big) = c_r I(r:n), \tag{5}$$

$$E_2 = E\Big(-\ln\Big(1-F_X(X_{j:n})\Big)\Big) = c_r I(n-r+1:n) \tag{6}$$

and

$$\begin{aligned}
E_3 &= E\Big(-\ln f_{X,Y}(X_{r:n}, Y_{[r:n]})\Big)\\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} c_r\Big(-\ln f_{X,Y}(x,y)\Big)\Big(F_X(x)\Big)^{r-1}\Big(1-F_X(x)\Big)^{n-r}\\
&\quad \times f_{X,Y}(x,y)dxdy\\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} c_r \ln(2\pi\sqrt{1-\rho^2})\Big(\Phi(x)\Big)^{r-1}\Big(1-\Phi(x)\Big)^{n-r}\\
&\quad \times f_{X,Y}(x,y)dxdy\\
&\quad +\frac{c_r}{2(1-\rho^2)}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^2\Big(\Phi(x)\Big)^{r-1}\Big(1-\Phi(x)\Big)^{n-r}\\
&\quad \times f_{X,Y}(x,y)dxdy\\
&\quad -\frac{\rho c_r}{(1-\rho^2)}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} xy\Big(\Phi(x)\Big)^{r-1}\Big(1-\Phi(x)\Big)^{n-r}\\
&\quad \times f_{X,Y}(x,y)dxdy\\
&\quad +\frac{c_r}{2(1-\rho^2)}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} y^2\Big(\Phi(x)\Big)^{r-1}\Big(1-\Phi(x)\Big)^{n-r}\\
&\quad \times f_{X,Y}(x,y)dxdy\\
&= A_1 + A_2 - A_3 + A_4, \text{ say.} \qquad (7)
\end{aligned}$$

It is easy to see that

$$\begin{aligned}
A_1 &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} c_r \ln(2\pi\sqrt{1-\rho^2})\Big(\Phi(x)\Big)^{r-1}\Big(1-\Phi(x)\Big)^{n-r}\\
&\quad \times f_{X,Y}(x,y)dxdy\\
&= \ln(2\pi\sqrt{1-\rho^2}), \qquad (8)
\end{aligned}$$

and

$$\begin{aligned}
A_2 &= \frac{c_r}{2(1-\rho^2)}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^2\Big(\Phi(x)\Big)^{r-1}\Big(1-\Phi(x)\Big)^{n-r}\\
&\quad \times f_{X,Y}(x,y)dxdy\\
&= \frac{c_r}{2(1-\rho^2)}\sum_{m=0}^{n-r}\binom{n-r}{m}(-1)^m \eta(r+m-1). \qquad (9)
\end{aligned}$$

We compute A_3 and A_4 as follows:

$$\begin{aligned} A_3 &= \frac{\rho c_r}{(1-\rho^2)} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} xy\Big(\Phi(x)\Big)^{r-1}\Big(1-\Phi(x)\Big)^{n-r} \\ &\quad \times f_{X,Y}(x,y)dxdy \\ &= \frac{\rho c_r}{(1-\rho^2)} \int_{-\infty}^{+\infty} x\Big(\Phi(x)\Big)^{r-1}\Big(1-\Phi(x)\Big)^{n-r}.B_1(x)dx, \end{aligned}$$

where

$$\begin{aligned} B_1(x) &= \int_{-\infty}^{+\infty} y f_{X,Y}(x,y)dy \\ &= \rho x \phi(x). \end{aligned}$$

So

$$A_3 = \frac{\rho^2 c_r}{(1-\rho^2)} \sum_{m=0}^{n-r} \binom{n-r}{m} (-1)^m \eta(r+m-1) \tag{10}$$

and

$$\begin{aligned} A_4 &= \frac{c_r}{2(1-\rho^2)} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y^2\Big(\Phi(x)\Big)^{r-1}\Big(1-\Phi(x)\Big)^{n-r} \\ &\quad \times f_{X,Y}(x,y)dxdy \\ &= \frac{c_r}{2(1-\rho^2)} \int_{-\infty}^{+\infty} \Big(\Phi(x)\Big)^{r-1}\Big(1-\Phi(x)\Big)^{n-r}.B_2(x)dx, \end{aligned}$$

where

$$\begin{aligned} B_2(x) &= \int_{-\infty}^{+\infty} y^2 f_{X,Y}(x,y)dy \\ &= (1-\rho^2+\rho^2 x^2)\phi(x). \end{aligned}$$

Therefore

$$A_4 = \frac{1}{2} + \frac{\rho^2 c_r}{2(1-\rho^2)} \sum_{m=0}^{n-r} \binom{n-r}{m} (-1)^m \eta(r+m-1). \tag{11}$$

Substituting (8),(9),(10) and (11) into (7), we obtain

$$E_3 = \frac{1}{2} + \ln(2\pi\sqrt{1-\rho^2}) + \frac{c_r}{2} \sum_{m=0}^{n-r} \binom{n-r}{m} (-1)^m \eta(r+m-1). \tag{12}$$

Substituting (5), (6) and (12) into (4) the proof is complete.■

Lemma 2.1 Suppose $X_{r:n}$ and $Y_{[r:n]}$ respectively denote the r-th order statistic and its concomitant from a random sample of size n from $N_2(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$. Then if $X = \frac{Z+\mu_1}{\sigma_1}$ and $Y = \frac{Z'+\mu_2}{\sigma_1}$, then $\forall \mu_1, \mu_2 \in \Re$ and $\forall \sigma_1, \sigma_2 > 0$,

$$H(X,Y) = \ln(\sigma_1\sigma_2) + H(Z,Z').$$

Proof. Straightforward.

Theorem 2.2 The entropy of $\mathfrak{C}(X,Y)$; $r_1 < r_2 < ... < r_k$ is:

$$\begin{aligned} H\left[\mathfrak{C}(X,Y)\right] &= \sum_{i=1}^{k} H(X_{r_i:n}, Y_{[r_i:n]}) - \sum_{i=2}^{k} c_{r_i}(r_i-1)I(r_i:n) + \sum_{i=1}^{k-1} c_{r_i}(n-r_i) \\ &\quad \times I(n-r_i+1:n) + \sum_{i=2}^{k}(r_i - r_{i-1} - 1)\left[c_{r_i} I\left(r_i:n\right) + (r_i - r_{i-1})\right. \\ &\quad \left. \times \binom{r_i - 1}{r_{i-1} - 1} I\left(r_i - r_{i-1} : r_i - 1\right)\right] + \sum_{i=2}^{k} \ln(r_i - r_{i-1} - 1)! \\ &\quad + \sum_{i=1}^{k} \ln c_{r_i} - \ln c, \end{aligned} \tag{13}$$

where

$$c = \frac{n!}{(r_1-1)!(n-r_k)!}.$$

Proof.

$$\begin{aligned} H\Big(\mathfrak{C}(X,Y)\Big) &= -\ln c + c_{r_1}(r_1-1)I(r_1:n) + c_{r_k}(n-r_k)I(n-r_k+1:n) \\ &\quad + \sum_{i=2}^{k}(r_i - r_{i-1} - 1)\left[c_{r_i} I\left(r_i:n\right) + (r_i - r_{i-1})\binom{r_i - 1}{r_{i-1} - 1}\right. \\ &\quad \left. \times\ I\left(r_i - r_{i-1} : r_i - 1\right)\right] + \sum_{i=2}^{k} \ln(r_i - r_{i-1} - 1)! \\ &\quad + \sum_{i=1}^{k} \Psi(r_i:n), \end{aligned}$$

where

$$c = \frac{n!}{(r_1-1)!(n-r_k)!}\ , \quad \Psi(r_i:n) \ = E\Big(-\ln f_{X,Y}(X_{r_i:n}, Y_{[r_i:n]})\Big).$$

Now (13) follows using a proof similar to that of theorem 2.1. ■

Corollary 2.2 $H\left[\mathfrak{C}(X,Y)\right] - \sum_{i=1}^{k} H(X_{r_i:n}, Y_{[r_i:n]})$ does not depend on the parent distribution.

Table 1 shows values of the Shannon entropy in order statistics and their concomitants from $N_2(0, 0, 1, 1, \rho)$ for $n = 5, 10, 15, 20$ and $\rho = 0.25, 0.50$ and 0.75. In fact for each n and ρ as r increases from 1 to n, the entropy first decreases, reaches a minimum at $r = \left[\frac{n+1}{2}\right]$ and then increases.

Table 1.Shannon entropy in order statistics and their concomitants from $N_2(0, 0, 1, 1, \rho)$

		r									
n	ρ	1	2	3	4	5	6	7	8	9	10
	0.25	2.4	2.22	2.18							
5	0.50	2.29	2.11	2.07							
	0.75	2.02	1.84	1.8							
	0.25	2.26	2.03	1.93	1.88	1.86					
10	0.50	2.15	1.92	1.82	1.77	1.75					
	0.75	1.88	1.65	1.55	1.5	1.48					
	0.25	2.19	1.94	1.82	1.75	1.71	1.68	1.67	1.66		
15	0.50	2.08	1.83	1.71	1.64	1.6	1.57	1.56	1.55		
	0.75	1.81	1.56	1.44	1.37	1.33	1.3	1.29	1.28		
	0.25	2.14	1.88	1.75	1.67	1.62	1.58	1.55	1.55	1.53	1.52
20	0.50	2.03	1.77	1.64	1.56	1.51	1.47	1.44	1.44	1.42	1.41
	0.75	1.76	1.5	1.37	1.29	1.24	1.2	1.17	1.17	1.15	1.14

3. Conclusion

We have derived the exact form of Shannon entropy for the order statistics and their concomitants from a bivariate normal distribution and extended these results for a collection of order statistics and their concomitants. Since normal distribution has application in many fields, we believe that our results will be important as a reference for many areas of study.

Acknowledgment. The authors are very grateful to the anonymous referees for their useful suggestion on an earlier version of this paper. these suggestion have enabled the authors to improve the paper significantly. this research was supported partially by Mahani Mathematical Research Center and Young Researchers Society of Shahid Bahonar University of Kerman.

References

[1] Abo-Eleneen, Z. A. and Nagaraja, H. N., Fisher Informmation in An Order Statistic and Its Concomitant, Annuals Institue Statistics Mathematical, Vol. 54, Number 3 (2002) 667-680.

[2] Cover, T. M. and Thomas, J. A. , Elements of Information Theory, John Wiley & Sons, New York, (1991).

[3] David, H. A. and Nagaraja, H. N. , Concomitants of Order Statistics, Order Statistics: Theory & Methods. Elsevier, Amsterdam, Vol. 16 (1998) 487-513.

[4] David, H. A. and Nagaraja, H. N., Order Statistics, third edition, John Wiley & Sons, New York, (2003).

[5] Hartley, R.T.V. , Transmission of Information, Bell System Technical Journal, Vol. 7, Number 3 (1928) 535-563.

[6] Lazo, A.V. and Rathie, P., On the Entropy of Continuous Probability Distributions, Information Theory, IEEE Transactions , Vol. 24, Number 1 (1978) 120-122.

[7] Madadi, M. and Tata, M., Shannon Information in Record Data, Metrika, Vol. 74, Number 1 (2011) 11-31.

[8] Nyquist, H. , Certain Affecting Telegraph Speed, Bell System Technical Journal, Vol. 3, Number 4 (1924) 324–346.

[9] Shannon, C. E. , A Mathematical Theory of Communication, Bell System Technical Journal, Vol. 27, Number 3 (1948) 379-423.