Far East Journal of Mathematical Sciences (FJMS)

Published Online: November 2013
Available online at http://pphmj.com/journals/fjms.htm
Volume 81, Number 1, 2013, Pages 95-102

ENTROPIES AND FISHER INFORMATION MATRIX FOR THE BETA TYPE 3 DISTRIBUTION

Daya K. Nagar and Yohana Tabares-Herrera

Instituto de Matemáticas
Universidad de Antioquia
Calle 67, No. 53-108
Medellín, Colombia
e-mail: dayaknagar@yahoo.com
yohanat88@yahoo.es

Abstract

In this article, we derive Fisher information matrix and entropies such as Rényi and Shannon for the beta type 3 distribution.

1. Introduction

The beta type 1 distribution with parameters (a, b) is defined by the probability density function (p.d.f.)

$$f_{B1}(u; a, b) = \frac{u^{a-1}(1-u)^{b-1}}{B(a, b)}, \quad 0 < u < 1, \tag{1}$$

where $a > 0$, $b > 0$, and $B(a, b)$ is the beta function defined by

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad \mathrm{Re}(a) > 0, \quad \mathrm{Re}(b) > 0.$$

Received: September 6, 2013; Accepted: September 14, 2013
2010 Mathematics Subject Classification: 33E99, 62H99.
Keywords and phrases: beta distribution, entropy, Gauss hypergeometric function, information matrix, psi function.

The beta type 1 distribution is well known in Bayesian methodology as a prior distribution on the success probability of a binomial distribution. The random variable V with the p.d.f.

$$f_{B2}(v;\, a,\, b) = \frac{v^{a-1}(1+v)^{-(a+b)}}{B(a,\, b)}, \quad v > 0, \tag{2}$$

where $a > 0$ and $b > 0$, is said to have a *beta type* 2 *distribution* with parameters (a, b). Since (2) can be obtained from (1) by the transformation $V = U/(1-U)$, some authors call the distribution of V an *inverted beta distribution*. The beta type 1 and beta type 2 are very flexible distributions for positive random variables and have wide applications in statistical analysis, e.g., see Johnson et al. [5]. Systematic treatment of matrix variate generalizations of beta distributions is given in Gupta and Nagar [2]. By using the transformation $W = U/(2-U)$, the beta type 3 density is obtained as (Gupta and Nagar [3, 4], Cardeño et al. [1]),

$$f_{B3}(w;\, a,\, b) = \frac{2^a w^{a-1}(1-w)^{b-1}(1+w)^{-(a+b)}}{B(a,\, b)}, \quad 0 < w < 1, \tag{3}$$

where $a > 0$ and $b > 0$.

It is well known that if X and Y are independent random variables having a standard gamma distribution with shape parameters a and b, respectively, then $X/(X+Y)$, X/Y and $X/(X+2Y)$ follow the beta type 1, beta type 2 and beta type 3 distributions, respectively.

In this article, we derive Fisher information matrix, Rényi and Shannon entropies for the beta type 3 distribution. The distributions of the product and the ratio of two independent random variables when at least one of them is beta type 3 are available in Sánchez and Nagar [13]. Recently, Nagar and Joshi [10] have derived densities of sum and difference of two independent beta type 3 variables. For results on non-central beta type 3 distribution, the reader is referred to Nagar and Ramirez-Vanegas [8, 9].

2. Entropies

In this section, exact forms of Rényi and Shannon entropies are determined for the beta type 3 distribution.

Let $(\mathscr{X}, \mathscr{B}, \mathscr{P})$ be a probability space. Consider a p.d.f. f associated with $\mathscr{P}$, dominated by σ-finite measure μ on $\mathscr{X}$. Denote by $H_{SH}(f)$ the well-known Shannon entropy introduced in Shannon [12]. It is defined by

$$H_{SH}(f) = -\int_{\mathscr{X}} f(x) \ln f(x) d\mu. \tag{4}$$

One of the main extensions of the Shannon entropy was defined by Rényi [11]. This generalized entropy measure is given by

$$H_R(\eta, f) = \frac{\ln G(\eta)}{1-\eta} \quad (\text{for } \eta > 0 \text{ and } \eta \neq 1), \tag{5}$$

where

$$G(\eta) = \int_{\mathscr{X}} f^{\eta} d\mu.$$

The additional parameter η is used to describe complex behavior in probability models and the associated process under study. Rényi entropy is monotonically decreasing in η, while Shannon entropy (4) is obtained from (5) for $\eta \uparrow 1$. For the details, see Nadarajah and Zografos [7], Zografos and Nadarajah [15] and Zografos [14].

First, we give some definitions and results useful in deriving these entropies.

The integral representation of the Gauss hypergeometric function is given as (Luke [6, Eq. 3.6(1)]),

$${}_2F_1(a, b; c; z) = \frac{1}{B(a, c-a)} \int_0^1 \frac{t^{a-1}(1-t)^{c-a-1}}{(1-zt)^b} dt, \tag{6}$$

where $\operatorname{Re}(c) > \operatorname{Re}(a) > 0$, $|\arg(1-z)| < \pi$. Expanding $(1-zt)^{-b}$, $|zt| < 1$, in (6) and integrating *t*, the series expansion for ${}_2F_1$ is derived as

$$ {}_2F_1(a, b; c; z) = \frac{\Gamma(c)}{\Gamma(a)\Gamma(b)} \sum_{j=0}^{\infty} \frac{\Gamma(a+j)\Gamma(b+j)}{\Gamma(c+j)} \frac{z^j}{j!}, \quad |z| < 1. \tag{7} $$

From (6), it easily follows that

$$ \int_0^1 \frac{w^{a-1}(1-w)^{b-1}}{(1+Kw)^c} dw = \frac{B(a, b)}{(1+K)^c} {}_2F_1\left(b, c; a+b; \frac{K}{1+K}\right). \tag{8} $$

Also, from (7), it is easy to see that ${}_2F_1(a, b; b; z) = (1-z)^{-a}$.

Lemma 2.1. *Let* $g(a, b) = \lim_{\eta \to 1} h(\eta)$, *where*

$$ h(\eta) = \frac{d}{d\eta} {}_2F_1\left(\eta(b-1)+1, \eta(a+b); \eta(a+b-2)+2; \frac{1}{2}\right). $$

Then

$$ g(a, b) = \sum_{r=1}^{\infty} \frac{\Gamma(b+r)}{\Gamma(b)} \frac{(1/2)^r}{r!} $$

$$ \times [(b-1)\psi(b+r) - 2\psi(a+b) - (b-1)\psi(b) + 2\psi(a+b+r)], \tag{9} $$

where $\psi(\alpha) = \Gamma'(\alpha)/\Gamma(\alpha)$ *is the digamma function.*

Proof. Expanding ${}_2F_1$ in series form, we write

$$ h(\eta) = \frac{d}{d\eta} \sum_{r=0}^{\infty} \Delta_r(\eta) \frac{(1/2)^r}{r!} = \sum_{r=0}^{\infty} \left[\frac{d}{d\eta} \Delta_r(\eta)\right] \frac{(1/2)^r}{r!}, \tag{10} $$

where

$$ \Delta_r(\eta) = \frac{\Gamma[\eta(b-1)+1+r]\Gamma[\eta(a+b)+r]\Gamma[\eta(a+b-2)+2]}{\Gamma[\eta(b-1)+1]\Gamma[\eta(a+b)]\Gamma[\eta(a+b-2)+2+r]}. $$

Now, differentiating the logarithm of $\Delta_r(\eta)$ w.r.t. to η, one obtains

$$\frac{d}{d\eta}\Delta_r(\eta) = \Delta_r(\eta)[(b-1)\psi(\eta(b-1)+1+r) + (a+b)\psi(\eta(a+b)+r)$$

$$+ (a+b-2)\psi(\eta(a+b-2)+2) - (b-1)\psi(\eta(b-1)+1)$$

$$- (a+b)\psi(\eta(a+b)) - (a+b-2)\psi(\eta(a+b-2)+2+r)]. \quad (11)$$

Finally, substituting (11) in (10) and taking $\eta \to 1$, one obtains the desired result. □

Now, we derive the Rényi and the Shannon entropies for the beta type 3 distribution.

Theorem 2.1. *For the beta type* 3 *distribution defined by the p.d.f.* (3), *the Rényi and the Shannon entropies are given by*

$$H_R(\eta,\ f) = \frac{1}{1-\eta}\Big[\eta\ln\Gamma(a+b) - \eta\ln\Gamma(a) - \eta\ln\Gamma(b) - \eta b\ln 2$$

$$+ \ln\Gamma[\eta(a-1)+1] + \ln\Gamma[\eta(b-1)+1]$$

$$- \ln\Gamma[\eta(a+b-2)+2]$$

$$+ \ln{}_2F_1\left(\eta(b-1)+1,\ \eta(a+b);\ \eta(a+b-2)+2;\ \frac{1}{2}\right)\Big] \quad (12)$$

and

$$H_{SH}(f) = -\ln\Gamma(a+b) + \ln\Gamma(a) + \ln\Gamma(b) + b\ln 2$$

$$- [(a-1)\psi(a) + (b-1)\psi(b) - (a+b-2)\psi(a+b)]$$

$$- 2^{-b}g(a,\ b), \quad (13)$$

respectively, *where* $\psi(\alpha) = \Gamma'(\alpha)/\Gamma(\alpha)$ *is the digamma function and* $g(a,\ b)$ *is given by* (9).

Proof. For $\eta > 0$ and $\eta \neq 1$, using the p.d.f. of *W* given by (3), we have

$$G(\eta) = \int_0^1 [f_{B3}(w;\, a,\, b)]^\eta dw$$

$$= \frac{2^{\eta a}}{[B(a,\, b)]^\eta} \int_0^1 \frac{w^{\eta(a-1)}(1-w)^{\eta(b-1)}}{(1+w)^{\eta(a+b)}} dw$$

$$= \frac{2^{\eta a}}{[B(a,\, b)]^\eta} \frac{B(\eta(a-1)+1,\, \eta(b-1)+1)}{2^{\eta(a+b)}}$$

$$\times\ {}_2F_1\left(\eta(b-1)+1,\, \eta(a+b);\, \eta(a+b-2)+2;\, \frac{1}{2}\right),$$

where the last line has been obtained by using (8). Now, taking logarithm of $G(\eta)$ and using (5), we get (12). The Shannon entropy (13) is obtained from (12) by taking $\eta \uparrow 1$ and using L'Hopital's rule. □

3. Fisher Information Matrix

In this section, we calculate the Fisher information matrix for the beta type 3 distribution. The information matrix plays a significant role in statistical inference in connection with estimation, sufficiency and properties of variances of estimators. For a given observation *w*, the Fisher information matrix for the beta type 3 distribution is defined as

$$-\begin{bmatrix} E\left(\dfrac{\partial^2 \ln L(a,\, b)}{\partial a^2}\right) & E\left(\dfrac{\partial^2 \ln L(a,\, b)}{\partial b \partial a}\right) \\ E\left(\dfrac{\partial^2 \ln L(a,\, b)}{\partial b \partial a}\right) & E\left(\dfrac{\partial^2 \ln L(a,\, b)}{\partial b^2}\right) \end{bmatrix}, \tag{14}$$

where $L(a,\, b) = \ln f_{B3}(w;\, a,\, b)$. From (3), the natural logarithm of $L(a,\, b)$ is obtained as

$$\ln L(a,\, b) = a \ln 2 - \ln \Gamma(a) - \ln \Gamma(b) + \ln \Gamma(a+b) + (a-1)\ln w$$

$$+ (b-1)\ln(1-w) - (a+b)\ln(1+w), \quad 0 < w < 1. \tag{15}$$

Now, differentiating (15) appropriately, we obtain

$$\frac{\partial^2 \ln L(a, b)}{\partial^2 a} = -\psi_1(a) + \psi_1(a+b), \tag{16}$$

$$\frac{\partial^2 \ln L(a, b)}{\partial b \partial a} = \psi_1(a+b) \tag{17}$$

and

$$\frac{\partial^2 \ln L(a, b)}{\partial^2 b} = -\psi_1(b) + \psi_1(a+b), \tag{18}$$

where $\psi_1(\alpha)$ is the trigamma function defined as the derivative of the digamma function $\psi(\alpha) = \Gamma'(\alpha)/\Gamma(\alpha)$.

Now, noting that expected value of a constant is the constant itself and observing that (16), (17) and (18) are constants, we have the information matrix as

$$\begin{bmatrix} \psi_1(a) - \psi_1(a+b) & -\psi_1(a+b) \\ -\psi_1(a+b) & \psi_1(b) - \psi_1(a+b) \end{bmatrix}.$$

Acknowledgment

The research work of DKN was supported by the Sistema Universitario de Investigación, Universidad de Antioquia under the project no. IN10182CE.

References

[1] Liliam Cardeño, Daya K. Nagar and Luz Estela Sánchez, Beta type 3 distribution and its multivariate generalization, Tamsui Oxford J. Math. Sci. 21(2) (2005), 225-241.

[2] A. K. Gupta and D. K. Nagar, Matrix Variate Distributions, Chapman & Hall/CRC, Boca Raton, 2000.

[3] A. K. Gupta and D. K. Nagar, Matrix variate beta distribution, Int. J. Math. Math. Sci. 23(7) (2000), 449-459.

[4] Arjun K. Gupta and Daya K. Nagar, Properties of matrix variate beta type 3 distribution, Int. J. Math. Math. Sci. 2009 (2009), Article ID308518, 18 pages (doi:10.1155/2009/308518).

[5] N. L. Johnson, S. Kotz and N. Balakrishnan, Continuous Univariate Distributions -2, 2nd ed., John Wiley & Sons, New York, 1994.

[6] Y. L. Luke, The Special Functions and their Approximations, Vol. 1, Academic Press, New York, 1969.

[7] S. Nadarajah and K. Zografos, Expressions for Rényi and Shannon entropies for bivariate distributions, Inform. Sci. 170(2-4) (2005), 173-189.

[8] Daya K. Nagar and Yeison Arley Ramirez-Vanegas, Noncentral beta type 3 distribution, Progress in Applied Mathematics 4(2) (2012), 29-43.

[9] Daya K. Nagar and Yeison Arley Ramirez-Vanegas, Distributions of sum, difference, product and quotient of independent noncentral beta type 3 variables, British J. Mathematics and Computer Science 3(1) (2013), 12-23.

[10] Daya K. Nagar and Lata Joshi, Sum and difference of independent beta type 3 variables, (submitted).

[11] A. Rényi, On measures of entropy and information, Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Vol. I, Univ. California Press, Berkeley, Calif., 1961, pp. 547-561.

[12] C. E. Shannon, A mathematical theory of communication, Bell System Technical J. 27 (1948), 379-423; 623-656.

[13] Luz Estela Sánchez and Daya K. Nagar, Distributions of the product and quotient of independent beta type 3 variables, Far East J. Theo. Stat. 17(2) (2005), 239-251.

[14] K. Zografos, On maximum entropy characterization of Pearson's type II and VII multivariate distributions, J. Multivariate Anal. 71(1) (1999), 67-75.

[15] K. Zografos and S. Nadarajah, Expressions for Rényi and Shannon entropies for multivariate distributions, Statist. Probab. Lett. 71(1) (2005), 71-84.