Available online at www.sciencedirect.com

SCIENCE DIRECT®

Statistics & Probability Letters 71 (2005) 71–84

STATISTICS & PROBABILITY LETTERS

www.elsevier.com/locate/stapro

Expressions for Rényi and Shannon entropies for multivariate distributions

K. Zografos[a], S. Nadarajah[b,*]

[a]Department of Mathematics, University of Ioannina, 45110 Ioannina, Greece

[b]Department of Statistics, University of Nebraska, Lincoln, Nebraska 68583, USA

Received 5 December 2002; received in revised form 12 October 2004
Available online 20 November 2004

Abstract

Exact forms of Rényi and Shannon entropies are determined for several multivariate distributions, including multivariate t, multivariate Cauchy, multivariate Pearson type VII, multivariate Pearson type II, multivariate symmetric Kotz type, multivariate logistic, multivariate Burr, multivariate Pareto type I, multivariate Pareto type II, multivariate Pareto type III, multivariate Pareto type IV, Dirichlet, inverted Dirichlet, multivariate Liouville, multivariate exponential, multivariate Weinman exponential, multivariate ordered Weinman exponential, bivariate gamma exponential, bivariate conditionally specified exponential, multivariate Weibull and multivariate log-normal. Monotonicity properties of Rényi and Shannon entropies for these distributions are also studied. We believe that the results presented here will serve as an important reference for scientists and engineers in many areas.

MSC: primary 62A10; 62E99; 62B10; 94A17

Keywords: Elliptically contoured distributions; Rényi entropy; Shannon entropy

*Corresponding author.
E-mail address: snadaraj@unlserve.unl.edu (S. Nadarajah).

0167-7152/$ - see front matter
doi:10.1016/j.spl.2004.10.023

1. Introduction

Let $(\mathfrak{X}, \mathfrak{B}, P)$ be a probability space. Consider a probability density function (pdf) f associated with P, dominated by a σ-finite measure μ on $\mathfrak{X}$. Denote by $\mathscr{H}_{\mathrm{Sh}}(f)$ the well-known Shannon entropy introduced in Shannon's (1948) paper on the mathematical theory of communication. It is defined by

$$\mathscr{H}_{\mathrm{Sh}}(f) = -\int_{\mathfrak{X}} f(x) \log f(x) \, \mathrm{d}\mu. \tag{1}$$

Since Shannon's (1948) pioneering work, entropy (1) has been used as a major tool in information theory and in almost every branch of science and engineering. Numerous entropy and information indices (cf. Burbea and Rao (1982) for the definition of the generalized φ-entropy, Vajda (1989) and Arndt (2001)) have been developed and used in various disciplines and contexts. Information theoretic principles and methods have become integral parts of probability and statistics and have been applied in various branches of statistics and related fields (cf. Soofi (2000) and references therein).

One of the main extensions of Shannon entropy was defined by Rényi (1961). This generalized entropy measure is given by

$$\mathscr{I}_{\mathrm{R}}(\lambda) = \mathscr{I}_{\mathrm{R}}(\lambda, f) = \frac{\log G(\lambda)}{1-\lambda} \tag{2}$$

(for $\lambda > 0$ and $\lambda \neq 1$), where

$$G(\lambda) = \int_{\mathfrak{X}} f^{\lambda} \, \mathrm{d}\mu. \tag{3}$$

The additional parameter λ is used to describe complex behavior in probability models and the associated process under study. Rényi entropy $\mathscr{I}_{\mathrm{R}}(\lambda)$ is monotonically decreasing in λ, while Shannon entropy (1) is obtained from (2) for $\lambda \uparrow 1$. These entropy measures have been compared by several authors (cf. Golan and Perloff (2002) and references therein). They share the following properties:

(i) Rényi's entropy is non-negative for any arbitrary f. Shannon entropy is non-negative only in the discrete case, i.e. when μ is a counting measure.

(ii) $\mathscr{H}_{\mathrm{Sh}}(f)$ is concave in f. $\mathscr{I}_{\mathrm{R}}(\lambda)$ is also concave in f because we restrict the parameter λ to be strictly positive. If $\lambda < 0$ then Rényi entropy is convex in f.

(iii) The two entropy measures differ in terms of their additivity properties. Shannon entropy satisfies the additive decomposition property, i.e. for two random vectors X and Y with the joint pdf $f_{X,Y}$ and associated marginal and conditional pdfs $f_X, f_Y, f_{X|Y}$ and $f_{Y|X}$,

$$\mathscr{H}_{\mathrm{Sh}}(f_{X,Y}) = \mathscr{H}_{\mathrm{Sh}}(f_Y) + E_Y[\mathscr{H}_{\mathrm{Sh}}(f_{X|Y})] = \mathscr{H}_{\mathrm{Sh}}(f_X) + E_X[\mathscr{H}_{\mathrm{Sh}}(f_{Y|X})].$$

However this property does not hold for Rényi's entropy $\mathscr{I}_{\mathrm{R}}(\lambda)$. If X and Y are independent then the above equation reduces to the property of standard additivity $\mathscr{H}_{\mathrm{Sh}}(f_{X,Y}) = \mathscr{H}_{\mathrm{Sh}}(f_X) + \mathscr{H}_{\mathrm{Sh}}(f_Y)$ which holds for both Rényi and Shannon entropies.

(iv) Maximum entropy principle is a widely used estimation method applied for linear and nonlinear estimation models (cf. Golan et al. (1996) among others). Maximum entropy