Journal of Multivariate Analysis 107 (2012) 53–63

Contents lists available at SciVerse ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

A study of the effect of kurtosis on discriminant analysis under elliptical populations

Jorge M. Arevalillo *, Hilario Navarro

Department of Statistics, Operational Research and Numerical Analysis, University Nacional Educación a Distancia (UNED), Paseo Senda del Rey 9, 28040, Madrid, Spain

ARTICLE INFO

Article history:
Received 5 July 2011
Available online 12 January 2012

AMS 2010 subject classifications:
primary 62H30
secondary 60E15

Keywords:
Kurtosis ordering
Elliptically contoured distribution
Bayes error
Discriminant analysis

ABSTRACT

This paper is concerned with the role some parameters indexing four important families within the multivariate elliptically contoured distributions play as indicators of multivariate kurtosis. The problem is addressed for the exponential power family, for a subclass of the Kotz family and for the Pearson type II and type VII distributions. Once such a problem is analyzed, we study the effect these parameters have, as kurtosis indicators, on binary discriminant analysis by exploring their relationship with the error rate of the Bayes discriminant rule. The effect is analyzed under mild conditions on the kernel function generating the elliptical density. Some numerical examples are given in order to illustrate our theoretical insights and findings.

1. Introduction

The word kurtosis comes from the Greek $\kappa\upsilon\rho\tau\acute{o}\varsigma$ which means bulging. The statistical concept behind it is concerned with the curvature, the amount of peakedness and tailweight of a distribution; it dates back to the renowned Karl Pearson [15] who defined it as a measure of departure from normality and introduced the terms *leptokurtic*, *mesokurtic* and *platikurtic* to classify distributions accordingly. Since then different alternatives to describe this concept have been proposed both for univariate and multivariate distributions; see [13,1,8,16,10] and [11,14,20] respectively, to name but a few.

Rather than using a single statistical quantity to account for kurtosis, Van Zwet [17] described it by means of a partial stochastic order between distribution functions. A kurtosis ordering $\leq_s$ between two symmetric random variables X and Y, with distribution functions F_X and G_Y, was defined in the following way: F_X is said to be less or equal than G_Y in kurtosis, $F_X \leq_s G_Y$, if and only if $G_Y^{-1}(F_X(x))$ is convex for $x > \mu$, where μ is the point of symmetry of F_X.

Therefore, in order that any statistical quantity can be considered a kurtosis indicator, it should retain such ordering. Van Zwet's definition, initially established for symmetric univariate distributions, has been adapted to the non symmetric case in [2]. Recently, Wang [19] extended it to the case of multivariate elliptically contoured distributions.

In this paper we explore the link between the multivariate kurtosis ordering as given in [19] and the binary discriminant analysis under elliptically contoured distributions. The main results are established for four important elliptical families: two subclasses of the Kotz family, the Pearson type II distributions and the multivariate t which is a subfamily of the Pearson type VII distributions. The work is organized as follows: the next section reviews several topics on elliptical distributions that will be used along the paper. In Section 3 we describe and extend the results in [19], putting the focus on the aforementioned families. Section 4 introduces the discriminant analysis in elliptical populations and studies the connection between kurtosis

* Corresponding author.
E-mail addresses: jmartin@ccia.uned.es (J.M. Arevalillo), hnavarro@ccia.uned.es (H. Navarro).

0047-259X/$ – see front matter
doi:10.1016/j.jmva.2012.01.011

and the error of the Bayes discriminant rule when the kernel function generating the elliptical density is non increasing; the theoretical discussion is illustrated with some numerical examples that will shed light on our findings. Finally, we will summarize our advances and will establish some concluding remarks.

2. Background on elliptically contoured distributions

Elliptically contoured distributions are useful tools for modeling multivariate data since they provide an alternative when the normality assumption fails; some survey works that treat them from a theoretical viewpoint are [3,5,4,7]. This section reviews some properties concerned with the stochastic representation of the elliptical distributions; the properties will be stated for the aforementioned subfamilies within the class of the elliptically contoured distributions.

Definition 1. Let $\boldsymbol{X} = (X_1, \dots, X_p)'$ be an absolutely continuous distributed random vector. We say that $\boldsymbol{X}$ follows an elliptically contoured distribution if its density function is given by

$$f(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Lambda}, g) = C_p|\boldsymbol{\Lambda}|^{-1/2} g((\boldsymbol{x}-\boldsymbol{\mu})'\boldsymbol{\Lambda}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})), \tag{1}$$

where $\boldsymbol{\mu} \in \mathbb{R}^p$, $\boldsymbol{\Lambda}$ is a $p \times p$ positive definite matrix, $C_p = \frac{\Gamma(\frac{p}{2})}{\pi^{\frac{p}{2}} \int_0^\infty t^{\frac{p}{2}-1} g(t)dt}$ and g is a non negative real valued function such that $\int_0^\infty t^{\frac{p}{2}-1} g(t)dt < \infty$.

We will write $\boldsymbol{X} \sim EC_p(\boldsymbol{\mu}, \boldsymbol{\Lambda}, g)$ to denote that $\boldsymbol{X}$ is distributed in accordance to a p-dimensional elliptically contoured distribution.

An interesting property of the elliptically contoured distribution is concerned with its stochastic representation which can be derived from Theorems 2.5.3 and 2.5.5 in [5]. The property is stated as follows.

Proposition 1. *A vector $\boldsymbol{X}$ satisfies $\boldsymbol{X} \sim EC_p(\boldsymbol{\mu}, \boldsymbol{\Lambda}, g)$ if and only if*

$$\boldsymbol{X} \overset{\mathrm{d}}{=} \boldsymbol{\mu} + \boldsymbol{A}' R \boldsymbol{U}^{(p)}, \tag{2}$$

where $\boldsymbol{A}$ is a square matrix such that $\boldsymbol{A}'\boldsymbol{A} = \boldsymbol{\Lambda}$, $\boldsymbol{U}^{(p)}$ is a p-dimensional vector with uniform distribution on the unit hypersphere and R is an absolutely continuous non negative random variable, independent of $\boldsymbol{U}^{(p)}$, such that R^2 has density function

$$h_{R^2}(r) = \frac{1}{\int_0^\infty t^{\frac{p}{2}-1} g(t)dt} r^{p/2-1} g(r), \quad r \geq 0. \tag{3}$$

The variable R in (2) is known as the modular variable since it has the same distribution as the modulus of $\boldsymbol{A}'^{-1}(\boldsymbol{X}-\boldsymbol{\mu})$. On the other hand, there is a close connection between the scale matrix $\boldsymbol{\Lambda}$ in (1) and the regular covariance matrix $\boldsymbol{\Sigma}$ of the elliptical distribution. From Theorem 2.6.4 in [5] it follows that $\boldsymbol{\Sigma} = \frac{E(R^2)}{p} \boldsymbol{\Lambda}$, provided that $E(R^2) < \infty$. Note that for the multivariate normal distribution, where the generating function is $g(t) = e^{-t/2}$, we obtain $E(R^2) = p$ and $\boldsymbol{\Sigma} = \boldsymbol{\Lambda}$.

Proposition 2 states a result which stems from the distribution of an affine transformation of a vector $\boldsymbol{X}$ having an elliptically contoured distribution. The proposition will be utilized in Section 4, where the connection between kurtosis and discriminant analysis will be analyzed.

Proposition 2. *Let $\boldsymbol{X} \sim EC_p(\boldsymbol{\mu}, \boldsymbol{\Lambda}, g)$ with stochastic representation $\boldsymbol{X} \overset{\mathrm{d}}{=} \boldsymbol{\mu} + \boldsymbol{A}'R\boldsymbol{U}^{(p)}$. The marginal distribution of the component X_1 of $\boldsymbol{X}$ is elliptical with stochastic representation $X_1 \overset{\mathrm{d}}{=} \mu_1 + a_{11}R_1U^{(1)}$, where $U^{(1)}$ is a Bernoulli random variable in $\{-1, 1\}$ independent of R_1, and $R_1 \overset{\mathrm{d}}{=} R\sqrt{B}$ with B a* Beta $\left(\frac{1}{2}, \frac{p-1}{2}\right)$ *variable independent of R. In addition, the modular variable R_1 has a density function given by*

$$h_{R_1}(r) = \frac{2\Gamma\left(\frac{p}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{p-1}{2}\right)\int_0^\infty t^{\frac{p}{2}-1} g(t)dt} \left(\int_0^\infty s^{\frac{p-1}{2}-1} g(r^2+s)ds\right), \quad r \geq 0. \tag{4}$$

Proof. The statement follows as a particular case of part 2 of Theorem 5 in [7] or by a direct application of Theorems 2.6.1 and 2.6.2 in [5]. □

Now, we review the results given by Propositions 1 and 2 when they are applied to the Kotz, Pearson type II and Pearson type VII families.