Multivariate exponential power distributions as mixtures of normal distributions with Bayesian applications

E. Gómez-Sánchez-Manzano [a,1], M. A. Gómez-Villegas [a,*,1],
J. M. Marín [b,1]

[a]*Departamento de Estadística e Investigación Operativa I, Fac. de CC. Matemáticas, Universidad Complutense, 28040 Madrid, Spain*

[b]*Departamento de Estadística, Universidad Carlos III, 28903 Getafe, Madrid, Spain*

Abstract

It is shown that a multivariate exponential power distribution is a scale mixture of normal distributions, with respect to a probability distribution function, when its kurtosis parameter belongs to the interval (0,1]. The corresponding mixing probability distribution function is presented. This result is used to design a Bayesian hierarchical model and an algorithm to generate samples of the posterior distribution; these are applied to a problem of Quantitative Genetics.

Key words: Elliptically contoured distribution, Elliptical distribution, Multivariate exponential power distribution, Scale mixture of normal distributions, Stable distribution, Hierarchical Bayesian model, Montecarlo methods.
AMS 2000 subject classification: Primary 62-02, secondary, 62E10, 62F15.

1 Introduction

The multivariate exponential power distribution was introduced by Gómez et al. [10] as a wide class of distributions, extending the class of normal distri-

---

* Corresponding author.
 Email addresses: eusebio_gomez@mat.ucm.es (E. Gómez-Sánchez-Manzano),
ma.gv@mat.ucm.es (M. A. Gómez-Villegas), jmmarin@est-econ.uc3m.es
(J. M. Marín).

butions with others having higher or lower tails.

In this paper it is shown that a multivariate exponential power distribution is a scale mixture of normal distributions, in the strict sense (namely, with respect to a probability distribution function) if and only if its kurtosis parameter (see Gómez et al. [10]) belongs to the interval $(0, 1]$; in this case, the mixing probability distribution function is also shown. This result is used, within a Bayesian framework, to design a hierarchical model and an algorithm to generate samples of the posterior distribution of the parameters; these are applied to a problem of Quantitative Genetics.

The results presented here are derived from the results on the relationship between elliptical distributions and scale mixture of normal distributions shown in Gómez-Sánchez-Manzano et al. [13]. Additional references for this relationship can be found there.

In the unidimensional case, West [23] shows the normal scale mixture property of some exponential power distributions and puts their mixing distributions in connection with stable distributions. These results are extended in the present paper. In the unidimensional case, too, Choy and Smith [5] use the mixture representation to simulate posterior distributions, within the context of random effects linear models.

The characterization of the exponential power distribution as a scale mixture of normal distributions simplifies some aspects of its study.

On the one hand, its marginal distributions and some affine transformations, which are not easy to directly work with, because they have not a closed form but one depending on an integral, become easier within the framework of mixtures of normal distributions, since they result to be mixtures of normals, too (see Gómez-Sánchez-Manzano et al. [13]).

On the other hand, from a Bayesian inference point of view, posterior distributions, when obtained in the usual direct way, are not suitable to work with, because they turn out to be formally complicated; their marginal distributions are cumbersome to obtain, because of difficulties of integration, and, besides, there is no practical way to find algorithms to simulate samples from them. We get around this difficulty, by taking the exponential power distribution as a mixture of normal distributions and using a more suitable hierarchical model. In this way, the conditional posterior distributions of the mean and scale parameters follow the well known scheme of the normal and Wishart distributions, and the posterior distribution of the kurtosis parameter is related to a product of densities of stable distributions. Thus, Gibbs sampler methods can be applied straightforwardly.

As for information about the multivariate exponential power distribution, its

2

definition and properties can be found in Gómez et al. [10]. This distribution is a generalization of the univariate one (see Subbotin [21] and Box and Tiao [4], p. 157) and can be included in the class of Kotz type distributions (see Fang et al. [7], p. 69, and Nadarajah [17]), which, in its turn, is a subset of the more general class of elliptical distributions; a survey on absolutely continuous elliptical distributions, with additional references, can be found in Gómez et al. [12]. A matrix generalization of the exponential power distribution can be found in Gómez et al. [11].

The multivariate exponential power distribution has been used in fields where departures for normality were needed. It has been used by J. K. Lindsey to obtain robust models for nonlinear repeated measurements (Lindsey [15]), to modeling dependencies among responses, as an alternative to models based upon the multivariate t distribution, and also to obtain robust models for the physiology of breathing. Basu et al. [2] use the multivariate exponential power distribution, as a heavy tailed distribution, in the field of speech recognition.

In section 2 we study the normal scale property. In section 3 we design a Bayesian hierarchical model and the algorithm to deal with the posterior distribution of the parameters. Finally, in section 4 we apply this methodology to a problem of Quantitative Genetics.

2   The mixture and the mixing distribution function

We consider the exponential power distribution and use the notation $X \sim \mathrm{EP}_n\left(\mu, \Sigma, \beta\right)$ as shown in Gómez et al. [10]. We also consider scale mixtures of normal distributions and use the notation $X \sim \mathrm{SMN}_n\left(\mu, \Sigma, H\right)$ as shown in Gómez-Sánchez-Manzano et al. [13]. For $\beta \in (0, \infty)$, we put $g_\beta(t) = \exp\left\{-\frac{1}{2}t^\beta\right\}$. It is clear that $X \sim \mathrm{EP}_n\left(\mu, \Sigma, \beta\right)$ if and only if X has the elliptical distribution $\mathrm{E}_n\left(\mu, \Sigma, g_\beta\right)$ (see Gómez-Sánchez-Manzano et al. [13]).

We shall also refer to stable distributions. For $\alpha \in (0, 1)$ and $\sigma > 0$, we denote by $S_\alpha\left(\,\cdot\,; \sigma\right)$ the density of the (positive) stable distribution having characteristic function (see Samorodnitsky and Taqqu [20], p. 8)

$$\varphi(t) = \exp\left\{-\sigma^\alpha \left|t\right|^\alpha \mathrm{e}^{-\mathrm{i}\frac{\pi}{2}\alpha\,\mathrm{sign}(t)}\right\}. \tag{2.1}$$

For $\alpha \in (0, 1)$, the Laplace transform of the distribution function F of the density $S_\alpha\left(\,\cdot\,; \sigma\right)$ is (see Samorodnitsky and Taqqu [20], p. 15) $\mathcal{L}_F(t) = \exp\left\{-\sigma^\alpha t^\alpha\right\}$. In particular, for $\sigma = 2^{-1/\alpha}$ we have the density $S_\alpha\left(\,\cdot\,; 2^{-1/\alpha}\right)$ and then $\mathcal{L}_F(t) = \exp\left\{-\frac{1}{2}t^\alpha\right\} = g_\alpha(t)$.

Theorem 1 (The mixture) *The exponential power distribution* $\mathrm{EP}_n(\mu, \Sigma, \beta)$ *is a scale mixture of normal distributions if and only if* $\beta \in (0, 1]$.

If $X \sim \mathrm{EP}_n(\mu, \Sigma, \beta)$, *with* $\beta \in (0, 1]$, *then* $X \sim \mathrm{SMN}_n(\mu, \Sigma, H_\beta)$, *where, for* $\beta \in (0, 1)$, H_β *is the absolutely continuous distribution function whose density* h_β *is given, for* $v > 0$, *by*

$$h_\beta(v) = \frac{2^{1+\frac{n}{2}-\frac{n}{2\beta}}\Gamma\left(1+\frac{n}{2}\right)}{\Gamma\left(1+\frac{n}{2\beta}\right)} v^{n-3} S_\beta\left(v^{-2}; 2^{1-\frac{1}{\beta}}\right); \tag{2.2}$$

and for $\beta = 1$, H_1 *is the distribution function degenerate in 1.*

PROOF. First we prove that for $\beta \in (1, \infty)$ the distribution $\mathrm{EP}_n(\mu, \Sigma, \beta)$ is not a scale mixture of normal distributions. Since function g_β is continuous, it is sufficient to prove that it does not satisfy condition (3.8) shown in theorem 8 of Gómez-Sánchez-Manzano et al. [13]. We will see that there exists a $t > 0$ such that $(-1)^2 g_\beta^{(2)}(t) < 0$. We have: $(-1)^2 g_\beta^{(2)}(t) = (\frac{1}{2}\beta t^{\beta-2} g_\beta(t))(\frac{1}{2}\beta t^\beta - \beta + 1)$. The factor $\frac{1}{2}\beta t^{\beta-2} g_\beta(t)$ is positive for any $t > 0$. The limit of the second factor when t tends to 0 is $\lim_{t\to 0}\left(\frac{1}{2}\beta t^\beta - \beta + 1\right) = -\beta + 1 < 0$; then there is a $t > 0$ for which this second factor is negative and so $(-1)^2 g_\beta^{(2)}(t) < 0$.

If $\beta \in (0, 1)$, then $g_\beta = \mathcal{L}_F$, where F is the distribution function of the density $S_\beta\left(\cdot; 2^{-1/\beta}\right)$; from theorem 10 and corollary 11 of Gómez-Sánchez-Manzano et al. [13] we obtain that $X \sim \mathrm{SMN}_n(\mu, \Sigma, H_\beta)$, where H_β has density h_β given, for $v > 0$, by

$$h_\beta(v) = \frac{2^{\frac{n}{2}\left(1-\frac{1}{\beta}\right)}\Gamma\left(1+\frac{n}{2}\right)}{\Gamma\left(1+\frac{n}{2\beta}\right)} v^{n-3} S_\beta\left(\frac{1}{2}v^{-2}; 2^{-\frac{1}{\beta}}\right). \tag{2.3}$$

On the other hand, we see from (2.1) that if a random variable W has density $f_W(w) = S_\beta(w; \sigma)$ then $Z = 2W$ has density $f_Z(z) = S_\beta(z; 2\sigma)$; but

$$f_Z(z) = \frac{1}{2}f_W\left(\frac{z}{2}\right) = \frac{1}{2}S_\beta\left(\frac{1}{2}z; \sigma\right);$$

therefore, $S_\beta\left(\frac{1}{2}z; \sigma\right) = 2S_\beta(z; 2\sigma)$. By applying this to (2.3) we get (2.2).

For the case $\beta = 1$, distribution $\mathrm{EP}_n(\mu, \Sigma, 1)$ is the normal distribution $\mathrm{N}_n(\mu, \Sigma)$. $\square$

Note that if V is the "mixing variable", having density h_β, then the variable $W = 2^{(1/\beta)-1}V^{-2}$ has density proportional to $w^{-n/2}S_\beta(w, 1)$.

As for the marginal distributions, they are very simple now: if $X \sim EP_n(\mu, \Sigma, \beta)$, with $\beta \in (0, 1]$, and $X_{(p)} = (X_1, ..., X_p)'$, with $p \leq n$, then $X_{(p)} \sim SMN_p(\mu_{(p)}, \Sigma_{(p)}, H_\beta)$, where $\mu_{(p)}$ and $\Sigma_{(p)}$ are the corresponding subvector and submatrix of μ an Σ, and H_β is the same distribution function shown in theorem 1 for vector X (see Gómez-Sánchez-Manzano et al. [13]).

As an example, we consider the case when $\beta = \frac{1}{2}$. We have the $EP_n\left(\mu, \Sigma, \frac{1}{2}\right)$ distribution, which is the double exponential distribution; from theorem 1 we obtain that this is the $SMN_n\left(\mu, \Sigma, H_{1/2}\right)$ distribution with

$$h_{\frac{1}{2}}(v) = \frac{\Gamma\left(\frac{n}{2}\right)}{2^{\frac{n}{2}}\Gamma(n)} v^{n-3} S_{\frac{1}{2}}\left(v^{-2}; \frac{1}{2}\right). \tag{2.4}$$

Now (see Samorodnitsky and Taqqu [20], p. 10), $S_{1/2}(y; \sigma) = \frac{1}{2}\pi^{-1/2}\sigma^{1/2}y^{-3/2} \exp\left(-\frac{1}{4}\sigma y^{-1}\right)$, for $y > 0$; this density, often called Lévy density, is an inverted-gamma distribution $IG\left(\frac{1}{2}, \frac{\sigma}{4}\right)$. Thus, function (2.4) is

$$h_{\frac{1}{2}}(v) = \frac{1}{2^{\frac{3n+1}{2}}\Gamma\left(\frac{n+1}{2}\right)} v^n \exp\left\{-\frac{1}{8}v^2\right\} I_{(0,\infty)}(v).$$

This is a generalized gamma density (Johnson et al. [14]); if $n = 1$, this is a Rayleigh density (Johnson et al. [14]). The variable V^2 has the gamma distribution $G(\frac{n+1}{2}, \frac{1}{8})$. If $n = 1$, the variable $V^2/8$ has the exponential distribution $Exp(1)$; this agrees with the results shown for this case in Andrews and Mallows [1] and West [23].
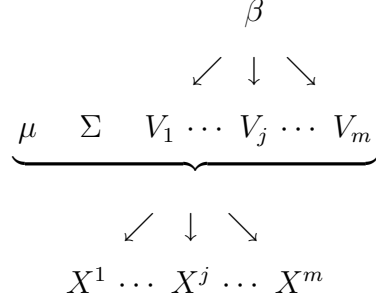
3   A hierarchical Bayesian model

Now we apply the former results, within a Bayesian framework, to deal with the posterior distribution of parameters μ, Σ, β of an exponential power distribution, once a prior distribution was assigned to them and a sample $\mathbf{X} = (X^1, ..., X^m)$ from this distribution has been observed to take the value $\mathbf{x} = (x^1, ..., x^m)$.

The point is that, when obtained in the usual direct way, the posterior distributions are not suitable to work with, as we said in section 1. So, in the case when $\beta \in (0, 1)$, we consider the $EP_n(\mu, \Sigma, \beta)$ distribution as the mixture $SMN_n(\mu, \Sigma, H_\beta)$, and use a hierarchical model instead. In fact, we introduce a vector $V = (V_1, \ldots, V_m)$ of m new parameters and consider each sample component X^j as coming independently from a normal $N_n\left(\mu, V_j^2\Sigma\right)$ distribution, with variance depending on V_j; now, β is an hyperparameter and the

5

density of each V_j conditional to β is h_β; we suppose, as usual, that μ, Σ, β are independent.

From a strictly probabilistic point of view, we have the set $\{\mu, \Sigma, \beta, V_1, \ldots, V_m, X^1, \ldots, X^m\}$ of $2m+3$ variables, and the dependency relationship among them is represented by this diagram:

$$\beta$$
$$\swarrow \quad \downarrow \quad \searrow$$
$$\underbrace{\mu \quad \Sigma \quad V_1 \cdots V_j \cdots V_m}$$
$$\swarrow \quad \downarrow \quad \searrow$$
$$X^1 \cdots X^j \cdots X^m$$

More precisely, the following conditions are met:

(1) The variables μ, Σ, β are independent and their marginal respective (prior) density functions are denoted by f_μ, f_Σ and f_β; the support of β is a subset of the interval $(0, 1)$.

(2) For all $j \in \{1, \ldots, m\}$, V_j is independent of μ and Σ and, for a fixed β, it is also independent of variables $V_1, \ldots, V_{j-1}, V_{j+1}, \ldots, V_m$. Actually, its conditional density function is $f_{V_j | \mu, \Sigma, \beta, V_1, \ldots, V_{j-1}, V_{j+1}, \ldots, V_m} = h_\beta$, given by (2.2).

(3) For all $j \in \{1, \ldots m\}$, the variable X^j, for a fixed value of V_j, depends only on μ and Σ; actually, its conditional distribution is

$$\left(X^j \mid \mu, \Sigma, \beta, v, x^1, \ldots, x^{j-1}, x^{j+1}, \ldots, x^m\right) \sim \mathrm{N}_n\left(\mu, v_j^2 \Sigma\right).$$

We are interested in the use of a Gibbs sampler to simulate samples of the posterior distribution of the parameters μ, Σ, β. To do this, we need the conditional posterior distributions of all the parameters; these are shown in the next theorem.

Theorem 2 (Conditional distributions) *We assume the three above conditions.*

(i) The density function of μ conditional to the remaining variables is

$$f_{\mu | \Sigma, \beta, V, \mathbf{X}}\left(\mu \mid \Sigma, \beta, v, \mathbf{x}\right) \propto f_\mu\left(\mu\right) N_n\left(\mu; a, S\right), \tag{3.5}$$

where N_n means the normal n-variate density function and

$$a = \left(\sum_{j=1}^{m} v_j^{-2} \right)^{-1} \sum_{j=1}^{m} v_j^{-2} x^j;$$

$$S = \left(\sum_{j=1}^{m} v_j^{-2} \right)^{-1} \Sigma.$$

(ii) The density function of Σ conditional to the remaining variables is

$$f_{\Sigma|\mu,\beta,V,\mathbf{X}} \left(\Sigma \mid \mu, \beta, v, \mathbf{x} \right) \propto f_\Sigma \left(\Sigma \right) |\Sigma|^{\frac{1}{2}(n+1)} IW \left(\Sigma; A, m \right), \qquad (3.6)$$

where IW is the inverse Wishart density function

$$IW \left(\Sigma; A, m \right) \propto |\Sigma|^{-\frac{m+n+1}{2}} \exp \left\{ -\frac{1}{2} tr \left(A^{-1} \Sigma^{-1} \right) \right\},$$

and $A = \sum_{j=1}^{m} v_j^{-2} \left(x^j - \mu \right) \left(x^j - \mu \right)'.$

(iii) The density function of β conditional to the remaining variables is

$$f_{\beta|\mu,\Sigma,V,\mathbf{X}} \left(\beta \mid \mu, \Sigma, v, \mathbf{x} \right) \propto f_\beta \left(\beta \right) \left(\frac{2^{1+\frac{n}{2}-\frac{n}{2\beta}} \Gamma \left(1 + \frac{n}{2} \right)}{\Gamma \left(1 + \frac{n}{2\beta} \right)} \right)^m \prod_{j=1}^{m} S_\beta \left(v_j^{-2}; 2^{1-\frac{1}{\beta}} \right).$$

$$(3.7)$$

(iv) For all $j \in \{1, \ldots m\}$, the density function of V_j conditional to the remaining variables is

$$f_{V_j|\mu,\Sigma,\beta,V_1,\ldots,V_{j-1},V_{j+1},\ldots,V_m,\mathbf{X}} \left(v_j \mid \mu, \Sigma, \beta, v_1, \ldots, v_{j-1}, v_{j+1}, \ldots v_m, \mathbf{x} \right) \propto$$

$$\propto v_j^{n-3} S_\beta \left(v_j^{-2}; 2^{1-\frac{1}{\beta}} \right) N_n \left(x^j; \mu, v_j^2 \Sigma \right).$$

for $v_j > 0$.

PROOF. The joint distribution function of the $2m + 3$ variables is

$$f_{\mu,\Sigma,\beta,V,\mathbf{X}} \left(\mu, \Sigma, \mathbf{x}, v \right) =$$

$$= f_\mu(\mu) f_\Sigma(\Sigma) f_\beta \left(\beta \right) \left(\prod_{j=1}^{m} h_\beta \left(v_j \right) \right) \prod_{j=1}^{m} N_n \left(x^j; \mu, v_j^2 \Sigma \right). \qquad (3.8)$$

The density function of each variable conditional to the remaining variables is proportional to (3.8). From this, the conditional density function of μ is proportional to $f_\mu(\mu) \prod_{j=1}^{m} N_n \left(x^j; \mu, v_j^2 \Sigma \right)$; the conditional density function of Σ is proportional to $f_\Sigma(\Sigma) \prod_{j=1}^{m} N_n \left(x^j; \mu, v_j^2 \Sigma \right)$; the conditional density function of

β is proportional to $f_\beta(\beta)\left(\prod_{j=1}^{m} h_\beta(v_j)\right)$; and, for all j, the conditional density function of V_j is proportional to $h_\beta(v_j) N_n\left(x^j; \mu, v_j^2\Sigma\right)$. The results of the theorem are straightforwardly obtained by operating on these expressions. $\square$

Note that, for all j, the conditional density function of the variable $W_j = 2^{(1/\beta)-1}V_j^{-2}$ is proportional to $w^{-(n/2)}S_\beta(w; 1) N_n\left(x^j; \mu, 2^{1-(1/\beta)}w^{-1}\Sigma\right)$, for $w > 0$.

As for the prior densities of μ and Σ, in many cases it will be useful to take for them the Jeffreys non informative priors, which turn out to be $f_\mu(\mu) = 1$ and $f_\Sigma(\Sigma) \propto |\Sigma|^{-(n+1)/2}$ (see Marín [16]). For the parameter β, for the sake of simplicity, the principle of *insufficient reason* (see Bernardo and Smith [3]) can be applied, which gives $f_\beta(\beta) = I_{(0,1)}(\beta)$. By doing this, the expressions (3.5), (3.6) and (3.7) reduce respectively to $N_n(\mu; a, S)$, $IW(\Sigma; A, m)$ and

$$\left(\frac{2^{1+\frac{n}{2}-\frac{n}{2\beta}}\Gamma\left(1+\frac{n}{2}\right)}{\Gamma\left(1+\frac{n}{2\beta}\right)}\right)^m \prod_{j=1}^{m} S_\beta\left(v_j^{-2}; 2^{1-\frac{1}{\beta}}\right).$$

The following Gibbs sampler algorithm, designed according to theorem 2, can be used to generate samples of the posterior distribution of the parameters, based on the prior suggested above.

ALGORITHM 3 (GIBBS SAMPLER)

(1) Make $t = 0$. Set initial values $(\mu^0, \Sigma^0, V_1^0, ..., V_m^0, \beta^0)$.
(2) Generate $\mu^{t+1} \sim N_n(a, S)$, where

$$a = \left(\sum_{j=1}^{m} \left(v_j^t\right)^{-2}\right)^{-1} \sum_{j=1}^{m} \left(v_j^t\right)^{-2} x^j;$$

$$S = \left(\sum_{j=1}^{m} \left(v_j^t\right)^{-2}\right)^{-1} \Sigma^t.$$

(3) Generate $\Sigma^{t+1} \sim IW(A, m)$, where

$$A = \sum_{j=1}^{m} \left(v_j^t\right)^{-2} \left(x_j - \mu^{t+1}\right)\left(x_j - \mu^{t+1}\right)'.$$

(4) For $j = 1, \ldots, m$,
  (a) make $W^t = 2^{\left(1/\beta^t\right)-1}\left(V^t\right)^{-2}$;
  (b) generate W_j^{t+1} with the next procedure:
    (i) generate $U \sim U(0,1)$.

      (ii) `generate` $Y \sim S_{\beta^t}(\,\cdot\,;1)$.

      (iii) `If` $U < \exp\left\{\lambda\left(W_j^t - Y\right)\right\}$, `where`

$$\lambda = \frac{1}{2}\left(x^j - \mu^{t+1}\right)'\Sigma^{-1}\left(x^j - \mu^{t+1}\right),$$

      `make` $W_j^{t+1} = Y$; `else,` `make` $W_j^{t+1} = W_j^t$.

   (c) `make` $V_j^{t+1} = \left(2^{1-\left(1/\beta^t\right)}W_j^{t+1}\right)^{-1/2}$.

(5) `Generate`

$$\beta^{t+1} \sim \left(\frac{2^{1+\frac{n}{2}-\frac{n}{2\beta^t}}\Gamma\left(1+\frac{n}{2}\right)}{\Gamma\left(1+\frac{n}{2\beta^t}\right)}\right)^m \prod_{j=1}^m S_{\beta^t}\left(\left(v_j^{t+1}\right)^{-2};2^{1-\frac{1}{\beta^t}}\right).$$

(6) `Make` $t = t+1$. `Go to Step 2`.

Note that in step 4b a Metropolis-Hasting step is used. Alternatively, a slice sampler step (see Godsill [9]) could be used.

Simulation of variable β in step 5 may be cumbersome because its distribution is very peaked over the mode; in this case a Metropolis-Hasting algorithm is little efficient because its taxes of acceptance are very low; and a slice sampler method goes very slowly. Thence, a *table method* (see Devroye [6]) may be the good choice.

4   Application on the distribution of mutational effects of genetical quantitative traits

In evolutionary studies of Quantitative Genetics it is important to identify the individual spontaneous mutations affecting the expression of, so called, neutral traits (not related with surviving of species). The effect of mutations is observed on different traits as morphological quantitative characters. In this way, it is relevant to study the distribution of this expressed traits in terms of its moments (see García-Dorado and Marín [8] for further details). The shape and kurtosis of the distribution of traits is relevant to future adaptations; the existence of higher kurtosis than the normal distribution is related with the existence of few mutations with large effects (see Turelli [22] for technical details).

A classical experimental methodology in the field of Quantitative Genetics to deal with the previous problem is based basically on mating among siblings. Thence, all the observed variability, after several generations is due to accumulated mutations.

We deal with flies of the species *Drosophila melanogaster* raised under standard conditions of laboratory, by means of a system of regular mating between siblings, in order to facilitate the accumulation of spontaneous mutations. In each Drosophila two morphological characters, length and width of wings, are measured in units of screen. We assume that these characters are distributed according to a bivariate power exponential distribution $EP_2(\mu, \Sigma, \beta)$ and we undertake the estimation of its parameters, in particular the kurtosis parameter, in order to determine the presence and effects of accumulated mutations.

As experimental data, we use a sample of size 100 (see Santiago et al. [19]) and we try to estimate its parameters by using the methodology described in this work.

We have implemented algorithm 3 by means of a `MatLab` program, with the subroutines related to stable distributions implemented by using the library `fBasics` from the `R` project, which, in its turn, uses an algorithm of Nolan [18]; the program `D(COM)Server` have been used to link all of them. We have run 30000 iterations, rejected the 10000 first ones as burn-in iterations, and taken the last 20000 results as a sample of (μ, Σ, β). We summarize the obtained sample by showing the marginal means, standard deviations and quantiles of μ_1, μ_2, $\text{tr}(\Sigma)$ and β in table 1.

Table 1. Posterior characteristics

	μ_1	μ_2	$\text{tr}(\Sigma)$	β
Mean	1205.44	464.57	1705.53	0.92
Std.Dev.	4.11	1.86	259.50	0.02
5%	1198.75	461.53	1317.85	0.87
25%	1202.66	463.32	1523.97	0.90
50%	1205.46	464.57	1682.61	0.92
75%	1208.21	465.80	1867.21	0.93
95%	1212.18	467.63	2163.55	0.95

Figure 1 shows the histogram and density plots for β.

We see that the posterior mean is less than 1, which is the value corresponding to the normal distribution, and that the whole distribution is located under this value. This implies that accumulation of spontaneous mutations produces a distribution of effects with a little higher kurtosis than the normal distribution in such a way that it is expected that Natural Selection may be weak over this traits to remove mutations.
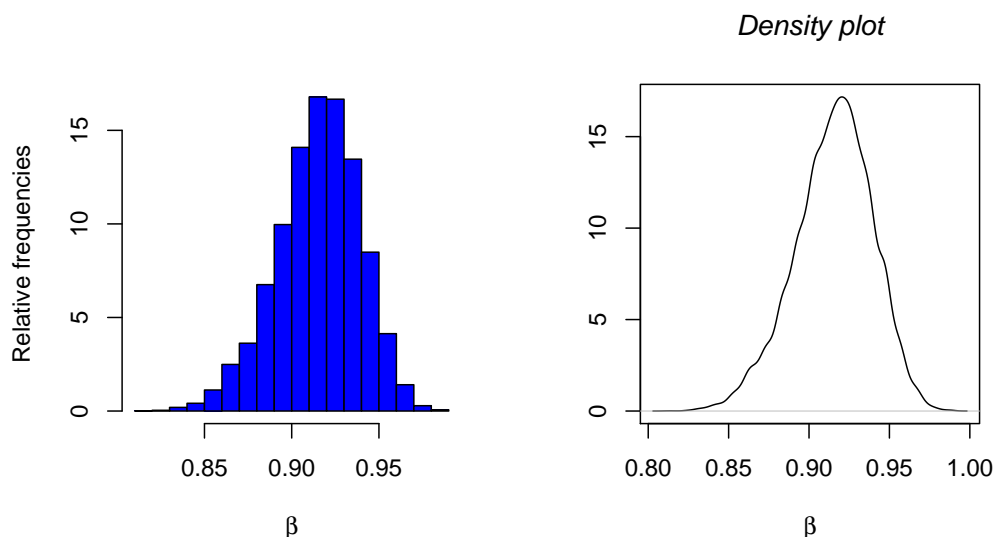
10

Fig. 1. Histogram and density plots of β

References

[1] D. R. Andrews and C. L. Mallows, Scale mixture of normal distributions, *J. R. Stat. Soc. Ser. B Stat. Methodol.* **36** (1974) 99–102.

[2] S. Basu, C. A. Miccheli and P. Olsen, Power exponential densities for the training and classification of acoustic feature vectors in speech recognition, *J. Comput. Graph. Statist.* **10(1)** (2001) 158–84.

[3] J. M. Bernardo and F. M. Smith, *Bayesian Theory*, Wiley and Sons (Chichester, 1994).

[4] G. Box and G. Tiao, *Bayesian Inference in Statistical Analysis*, Addison-Wesley (Reading, 1973).

[5] S. T. B. Choy and F. M. Smith, Hierarchical models with scale mixtures of normal distributions, *Test* **6(1)** (1997) 205–221.

[6] L. Devroye, *Non-Uniform Random Variate Generation*, Springer-Verlag (New York, 1986).

[7] K. T. Fang, S. Kotz, and K. W. Ng, *Symmetric Multivariate and Related Distributions*, Chapman and Hall (London, 1990).

[8] M. A. García-Dorado and J. M. Marín, Minimum Distance Estimation of Mutational Parameters for Quantitative Traits, *Biometrics* **54(3)** (1998) 214–230.

[9] S. J. Godsill, Inference in symmetric alpha-stable noise using MCMC and the slice sampler, *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing* **vol. VI** (2000) 3806–3809.

[10] E. Gómez, M. A. Gómez-Villegas and J. M. Marín, A multivariate generalization of the power exponential family of distributions, *Comm. Statist. Theory Methods* **27** (1998) 589–600.

[11] E. Gómez, M. A. Gómez-Villegas and J. M. Marín, A matrix variate generalization of the power exponential family of distributions, *Comm. Statist. Theory Methods* **31(12)** (2002) 2167–2182.

[12] E. Gómez, M. A. Gómez-Villegas and J. M. Marín, A survey on continuous elliptical vector distributions, *Rev. Mat. Complut.* **16** (2003) 345–361.

[13] E. Gómez-Sánchez-Manzano, M. A. Gómez-Villegas and J. M. Marín, Sequences of elliptical distributions and mixtures of normal distributions, *J. Multivariate Anal.* **97** (2006) 295–310.

[14] N. L. Johnson, S. Kotz and N. Balakrishnan, *Continuous Univariate Distributions. Vol. 1*, Wiley (New York, 1994).

[15] J. K. Lindsey, Multivariate elliptically-contoured distributions for repeated measurements, *Biometrics* **56** (1999) 1277–1280.

[16] J. M. Marín, *Continuous elliptical and power exponential distributions with applications to linear dynamic models*, Ph. D. Thesis (Departamento de Estadística e Investigación Operativa, Universidad Complutense de Madrid, Spain, 1998).

[17] S. Nadarajah, The Kotz-type distribution with applications, *Statistics* **37(4)** (2003) 341–58.

[18] J. P. Nolan, Numerical calculation of stable densities and distribution functions, *Commun. Statist. Stochastic Models* **13** (1997) 759–774.

[19] E. Santiago, J. Albornoz, A. Domínguez, M. A. Toro and C. López-Fanjul, The distribution of spontaneous mutations on quantitative traits and fitness in Drosophila melanogaster, *Genetics* **132** (1992) 771–781.

[20] G. Samorodnitsky, M. S. Taqqu, *Stable Non-gaussian Random Processes*, Chapman and Hall/CRC (Boca Raton, 2000).

[21] M. Subbotin, On the law of frequency of errors, *Mathematicheskii Sbornik* **31** (1923) 296–301.

[22] M. Turelli, Heritable genetic variation via mutation-selection balance: Lerch's zeta meets abdominal bristle, *Theoretical Population Biology* **25** (1984) 138–193.

[23] M. West, On scale mixtures of normal distributions, *Biometrika* **74** (1987) 646–648.