A General Family Of Probability Distributions

Gilbert Saporta, Vasile Preda

Pub. Inst. Stat. Univ. Paris
XXXIX, fasc. 2, 1995, 71 à 93

A GENERAL FAMILY OF PROBABILITY DISTRIBUTIONS

Gilbert SAPORTA
Conservatoire National des Arts et Métiers
Chaire de Statistique Appliquée
292, Rue Saint-Martin
75141 Paris, Cedex 03
FRANCE

Vasile PREDA
University of Bucharest
Mathematics Faculty
14, Academiei str.
Bucharest - ROMANIA

Abstract

A new family of probability distributions is defined and Bayes estimation, Fisher information, Shannon entropy, an information differential metric and a James estimator are studied. This family is an extension of some families considered in statistics. Also some procedures to obtain the moment and maximum likelihood estimates are given.

1 Introduction.

The purpose of this paper is to define a new family of compound probability distributions and to stated some theoretical results and procedures of this family. This family of probability distributions is an extension of the families

of distributions considered by Holla and Bhattacharya (1968) and Luceño (1992). Also our family of distributions is intended, as in Luceño (1992), as a tool which could allow data analysis to be performed in the original scale. This family is found by compounding two distributions and has the normal, the gamma, the exponential and the double exponential distributions as limiting cases.

Let us suppose that X is a continuous random variable having a normal distribution with mean ξ and variance τ^2 and that ξ is a continuous random variable with a probability density function given by

$$g(\xi) = \begin{cases} \theta \frac{\lambda_1^{n+1}}{n!}(\xi - v)^n \cdot \exp\left(-\lambda_1(\xi - v)\right) & , \text{ if } \quad \xi > v \\ \eta \frac{\lambda_2^{n+1}}{n!}(v - \xi)^n \cdot \exp\left(-\lambda_2(v - \xi)\right) & , \text{ if } \quad \xi < v \end{cases} \tag{1}$$

where $0 \leq \theta \leq 1$, $\eta = 1 - \theta$, $\lambda_1 \geq 0$, $\lambda_2 \geq 0$, $\lambda_1 + \lambda_2 > 0$ and n is a positive integer. We note that for $n = 0$ we get the probability density family given by Luceño(1992).

Compounding these two distributions, a new family is found whose probability density function $f(x)$ and cumulative distribution function $F(x)$ are given for any real x. Further, in this part some properties for $g(\xi)$ and $f(\xi)$ are given. For example Bayes estimation, Fisher information, Shannon entropy and Burbea-Rao information differential metric are studied. Our results are defined relative to $\Phi(\cdot)$, the standard normal cumulative distribution function, and $\Psi_n(\cdot)$ a transform of order n relative to standard normal distribution function. For this transform $\Psi_n(\cdot)$ some recurrence properties are stated. If $n = 1$ we obtain the transform $\Psi(\cdot)$ which appear in Raiffa and Schlaifer (1961) and De Groot (1970, pp.246).

The definition of Ψ_n transform and some important properties of this transform are described in Section 2. Many interesting properties of probability density function $g(\xi)$ are presented in section 3. Thus for $g(\xi)$ we obtain the moment generating function, the mean, the variance, the semivariance, the central moments, the relations for skewness and kurtosis, the Fisher information matrix for some parameters, the information differential metric (Burbea and Rao (1982)) and Shannon entropy (Guiaşu (1987)). In Section 4, the compound probability density function $f(x)$, the cumulative distribution function $F(x)$, the moment generating function, the skewness and the kurtosis are stated. Also the posterior distribution of ξ and Bayes estimation of ξ based on an observation x from a normal distribution with

mean ξ and variance τ^2, under a prior density $g(\xi)$ and quadratic loss function is given in Section 4. The moment and maximum likelihood estimates of some parameters can be found using standard algorithms presented in Section 5. Using a result of Ahmad, Giri and Sinha (1983), in Section 6, a James estimator for proportion θ is given. Finally, certain useful results are included in three appendices.

2 Ψ_n transform

In this section we define the Ψ_n transform and some properties of Ψ_n are given. Also in Appendices A and B some recurrence relations are stated.

Let n be a positive integer and $\varphi(\cdot)$ the standard normal density function. We put

$$\Psi_n(s) = \int_s^\infty (t-s)^n \varphi(t)dt$$

for any real s.

We note that $\Psi_n(t) \geq 0$ for any t and $\lim_{n\to\infty} \Psi_n(s) = +\infty$. According to Appendix A we have that Ψ_n is a derivable function and according to Lemma 2A, $\Psi_n'(s) = -n\Psi_{n-1}(s)$, for any $n \geq 1$ and s real. Also for any real s and $n \geq 1$, according to Lemma 1A, we have

$$\Psi_{n+1}(s) = -s\Psi_n(s) + n\Psi_{n-1}(s).$$

Further on other relations concerning Ψ_n are given by Lemmas 5A, 6A, 7A (Appendix A) and the following

Theorem 2.1 We have:

a) $\Psi_n(\cdot)$ is a convex and a decreasing function;

b) $\Psi_n(s) \geq (\leq) \sum_{k=0}^{n} (-1)^k \binom{n}{k} s^k a_{n-k}$, if n is odd (even);

c) $\lim_{s\to\infty} \Psi_n(s) = 0$;

d) $\lim_{s\to-\infty} [\Psi_n(s) - \sum_{k=0}^{n} (-1)^k \binom{n}{k} s^k a_{n-k}] = 0$,

where a_k is defined in Lemma 3A (Appendix A).

Proof. a) According to Lemma 2A (Appendix A) we have

$$\Psi_n'(s) \leq 0 \quad \text{and} \quad \Psi_n''(s) = n(n-1) \cdot \Psi_{n-2}(s) \geq 0.$$

For b) and d) we have in view that

$$\Psi_n(s) = \int_{-\infty}^{\infty} (t-s)^n \varphi(t)\, dt - \int_{-\infty}^{s} (t-s)^n \varphi(t)\, dt$$

and further we use Lemma 3A (Appendix A).
Relation c) is immediate.

Remark 2.1 According to Johnson and Kotz (1970), in Fisher (1930) defined the Hh_n functions, $Hh_n(s) = (n!)^{-1} \int_s^{\infty} (t-s)^n \exp(-t^2/2) dt$. Thus $Hh_n(s) = (n!)^{-1}(2\pi)^{1/2} \cdot \Psi_n(s)$.

3 Properties of random variable ξ.

Let us suppose that ξ has a probability density function defined by $g(\xi)$ given by (1). The moment generating function of ξ is given by $m_\xi(t) = E_\xi(e^{t\xi})$. Using Lemma 1B and Lemma 2B (Appendix B) we obtain

$$m_\xi(t) = \left[\theta \left(\frac{\lambda_1}{\lambda_1 - t}\right)^{n+1} + \eta \left(\frac{\lambda_2}{\lambda_2 - t}\right)^{n+1}\right] \exp(tv)$$

for $-\lambda_2 < t < \lambda_1$.

The mean and variance of ξ are given by

$$\mu = E_\xi(\xi) = m'_\xi(0) = v + (n+1)[\theta\lambda_1^{-1} - \eta\lambda_2^{-1}]$$

$$Var_\xi(\xi) = E_\xi(\xi - \mu)^2 = m''_\xi(0) - (m'_\xi(0))^2 = \alpha_{2,\mu}(\xi)$$
$$= (n+1)(n+2)[\theta\lambda_1^{-2} - \eta\lambda_2^{-2}] - (n+2)^2[\theta\lambda_1^{-1} - \eta\lambda_2^{-1}]^2$$

where $\eta = 1 - \theta$, and $\alpha_{2,\mu}(\xi) = E_\xi(\xi - \mu)^2$.

The k - th moments of ξ about a given constant a is $\alpha_{k,a}(\xi) = E_\xi(\xi - a)^k$. Using (1) , the moments about v are given by $\alpha_{k,v}(\xi) = \frac{(n+k)!}{n!}[\theta\lambda_1^{-k} - \eta\lambda_2^{-k}]$. Now by Lemma 3C (Appendix C) we obtain the formula for central moments of ξ ,

$$\mu_k(\xi) = \sum_{i=1}^{k} (-1)^i \binom{k}{i} (n+1)^i [\theta\lambda_1^{-1} - \eta\lambda_2^{-1}]^i \alpha_{k-i,v}(\xi)$$

We remark that $\mu_0(\xi) = 1$ and $\mu_1(\xi) = 0$. Because

$$\mu_2(\xi) = (n+1)(n+2)[\theta\lambda_1^{-2} - \eta\lambda_2^{-2}] - (n+2)^2[\theta\lambda_1^{-1} - \eta\lambda_2^{-1}]^2 \quad ,$$

$$\mu_3(\xi) = (n+3)(n+2)(n+1)[\theta\lambda_1^{-3} - \eta\lambda_2^{-3}] - 3(n+2)(n+1)^2 \cdot$$
$$\cdot[\theta\lambda_1^{-1} - \eta\lambda_2^{-1}] \cdot [\theta\lambda_1^{-2} - \eta\lambda_2^{-2}] + 2(n+1)^3[\theta\lambda_1^{-1} - \eta\lambda_2^{-1}]^3 \quad ,$$

$$\mu_4(\xi) = (n+4)(n+3)(n+2)(n+1)[\theta\lambda_1^{-4} - \eta\lambda_2^{-4}] - 4(n+3) \cdot$$
$$\cdot(n+2)(n+1)^2[\theta\lambda_1^{-1} - \eta\lambda_2^{-1}][\theta\lambda_1^{-3} - \eta\lambda_2^{-3}] + 6(n+2)(n+1)^3 \cdot$$
$$\cdot[\theta\lambda_1^{-1} - \eta\lambda_2^{-1}]^2[\theta\lambda_1^{-2} - \eta\lambda_2^{-2}] - 3(n+1)^4[\theta\lambda_1^{-1} - \eta\lambda_2^{-1}]^4 \quad ,$$

we obtain the skewness and the kurtosis $\alpha_3(\xi) = \mu_3(\xi)\,(\mu_2(\xi))^{-3/2}$ and $\alpha_4(\xi) = \mu_4(\xi)\,(\mu_2(\xi))^{-2} - 3$ respectively.

In the following we consider Fisher information relative to some parameters of $g(\xi)$, Shannon entropy of $g(\xi)$ and information metric relative to $g(\xi)$.

a). *Fisher information.* If θ , λ_1 , λ_2 and n are known we obtain Fisher information for v . According to Saporta (1990, pp.291) we can write

$$I(v) = -E_\xi\left(\tfrac{\partial^2}{\partial v^2}\ln g(\xi)\right) = n\int_{-\infty}^{v}\eta\lambda_2^{n+1}(n!)^{-1}(v-\xi)^{n-2}\exp\left(-\lambda_2(v-\xi)\right)\mathrm{d}\xi +$$
$$+n\int_{v}^{\infty}\theta\lambda_1^{n+1}(n!)^{-1}(\xi-v)^{n-2}\exp\left(-\lambda_1(\xi-v)\right)d\xi = (n-1)^{-1}[\eta\lambda_2^2 + \theta\lambda_1^2]$$

When θ , λ_2 , v and $n \geq 2$ are known then Fisher information for λ_1 is $I(\lambda_1) = \theta(n+1)\lambda_1^{-2}$.

Generally, if v is a given constant and $\alpha = (\alpha_1, \alpha_2, \alpha_3)$ where $\alpha_1 = \theta$, $\alpha_2 = \lambda_1$, $\alpha_3 = \lambda_2$, where n is known, using Lemma 2B (Appendix B) we obtain $I(\alpha)$, the Fisher information matrix for α , $I(\alpha) = -E\left(\frac{\partial^2 \ln g}{\partial\alpha_i\partial\alpha_j}\right)_{i,j}$

$$I(\alpha) = \begin{pmatrix} (\theta(1-\theta))^{-1} & 0 & 0 \\ 0 & \theta(n+1)\lambda_1^{-2} & 0 \\ 0 & 0 & (1-\theta)(n+1)\lambda_2^{-2} \end{pmatrix}$$

Thus we have the metric tensor field components (Burbea and Rao, 1982) $\epsilon_{11} = (\theta(1-\theta))^{-1}$, $\epsilon_{12} = \epsilon_{13} = \epsilon_{21} = \epsilon_{23} = \epsilon_{31} = \epsilon_{32} = 0$, $\epsilon_{22} = \theta(n+1)\lambda_1^{-2}$, $\epsilon_{33} = (1-\theta)(n+1)\lambda_2^{-2}$.

Hence, the information metric $ds^2 = \sum_{i,j=1}^{3}\epsilon_{ij}d\alpha_i d\alpha_j$, may be expressed as

$$ds^2 = (\theta(1-\theta))^{-1}(d\theta)^2 + \theta(n+1)\lambda_1^{-2}(d\lambda_1)^2 + (1-\theta)(n+1)\lambda_2^{-2}(d\lambda_2)^2$$

b). *Shannon entropy.* According to Guiaşu (1978), Shannon entropy of $g(\xi)$ is $H(g) = -\int_{-\infty}^{\infty} g(\xi)\ln g(\xi)d\xi$. Thus we have

$$H(g) = -\int_{-\infty}^{v} \eta\frac{\lambda_2^{n+1}}{n!}(v-\xi)^n \exp\left(-\lambda_2(v-\xi)\right)\cdot\left[\ln\left(\eta\frac{\lambda_2^{n+1}}{n!}\right) + n\ln(v-\xi)\right.$$
$$\left.-\lambda_2(v-\xi)\right]d\xi - \int_{v}^{\infty} \theta\frac{\lambda_1^{n+1}}{n!}(\xi-v)^n \exp\left(-\lambda_1(\xi-v)\right)\cdot$$
$$\cdot\left[\ln\left(\theta\frac{\lambda_1^{n+1}}{n!}\right) + n\ln(\xi-v) - \lambda_1(\xi-v)\right]d\xi$$

Using Lemma 1B and Lemma 2B (Appendix B), Lemma 3A (Appendix A) and the relation $\Gamma'(t) = \int_0^{\infty} z^{t-1}\ln z\cdot\exp(t)dt$, where $\Gamma'(t)$ is the derivative of gamma function $\Gamma(t)$, we obtain

$$H(g) = n + 1 + \ln(n!) - (\eta\ln\eta + \theta\ln\theta) - \left(\eta\ln\lambda_2 + \theta\ln\lambda_1 + \frac{\Gamma'(n+1)}{(n-1)!}\right)$$

where $\eta = 1-\theta$. If $\theta \to 1$ we obtain the Shannon entropy for gamma distribution. Further on if $n = 0$ we obtain the exponential case.

We see that $g(\xi)$ given by (1) is a mixture of two gamma distributions ,each having different intervals as ranges of variation. Plucinska (1966) used two generalized gamma distributions, one for negative and one for positive values of the argument, to construct a new class of distribution functions. Mixtures of such distributions are used in Plucinska (1967).

4 The compound distribution. Derivation and properties.

By compounding the distribution of ξ, given by probability density function $g(\xi)$, and X given by a normal distribution with mean ξ and variance τ^2, by Lemma 3B (Appendix B) we obtain

$$f(x) = \theta\cdot(2\pi\tau^2)^{-1/2}\cdot\frac{\lambda_1^{n+1}}{n!}\cdot L_1(x,\lambda_1,v,\tau) + \eta\cdot(2\pi\tau^2)^{-1/2}\cdot\frac{\lambda_2^{n+1}}{n!}\cdot L_2(x,\lambda_2,v,\tau)$$

where

$$L_1(x,\lambda_1,v,\tau) = \tau^{n+1}(2\pi)^{1/2}\exp\left(-\lambda_1\tau\tfrac{x-v}{\tau}\right)\cdot\exp\left(\tfrac{(\lambda_1\tau)^2}{2}\right)\cdot\Psi_n\left(-\tfrac{x-v}{\tau}+\lambda_1\tau\right)$$

$$L_2(x,\lambda_2,v,\tau)=\tau^{n+1}(2\pi)^{1/2}\exp\left(\lambda_2\tau\tfrac{x-v}{\tau}\right)\cdot\exp\left(\tfrac{(\lambda_2\tau)^2}{2}\right)\cdot\Psi_n\left(\tfrac{x-v}{\tau}+\lambda_2\tau\right)$$

We put $K(\omega)=\exp(\omega^2/2)$, $h_n(u,\omega)=\exp(u\omega)\cdot\Psi_n(u+\omega)$, $\omega_1=\lambda_1\tau$, $\omega_2=\lambda_2\tau$. Thus we get the probability density function

$$f(x)=(n!\tau)^{-1}\left[\theta\omega_1^{n+1}K(\omega_1)h_n(-\tfrac{x-v}{\tau},\omega_1)+\eta\omega_2^{n+1}K(\omega_2)h_n()\tfrac{x-v}{\tau},\omega_2\right] \quad (2)$$

for $-\infty<x<\infty$.

The cumulative distribution function $F(x)$ is given, for any $-\infty<x<\infty$ by

$$F(x)=\int_{-\infty}^{x}f(t)dt=\theta\omega_1^{n+1}(n!)^{-1}K(\omega_1)\cdot\int_{-\infty}^{\frac{x-v}{\tau}}h_n(-u,\omega_1)\,du+$$
$$+\eta\omega_2^{n+1}(n!)^{-1}K(\omega_2)\cdot\int_{-\infty}^{\frac{x-v}{\tau}}h_n(u,\omega_2)\,du.$$

According to Lemma 2A (Appendix A) we have

$$\int_{-\infty}^{\frac{x-v}{\tau}}h_n(u,\omega_2)\,du=\omega_2^{-1}\exp\left(\tfrac{x-v}{\tau}\omega_2\right)\Psi_n\left(\tfrac{x-v}{\tau}+\omega_2\right)+n\omega_2^{-1}\int_{-\infty}^{\frac{x-v}{\tau}}h_{n-1}(u,\omega_2)\,du$$

for $n\geq 1$, with

$$\int_{-\infty}^{\frac{x-v}{\tau}}h_0(u,\omega_2)\,du=\int_{-\infty}^{\frac{x-v}{\tau}}\exp(u\omega_2)\Psi_0(u+\omega_2)\,du= \quad (3)$$
$$=\omega_2^{-1}\exp(\tfrac{x-v}{\tau}\omega_2)\cdot\Phi(-\tfrac{x-v}{\tau}-\omega_2)+\omega_2^{-1}\exp(-\omega_2^2/2)\cdot\Phi(\tfrac{x-v}{\tau})$$

and

$$\int_{-\infty}^{\frac{x-v}{\tau}}h_n(-u,\omega_1)\,du=-\omega_1^{-1}\exp\left(-\tfrac{x-v}{\tau}\omega_1\right)\Psi_n\left(-\tfrac{x-v}{\tau}+\omega_1\right)+$$
$$+n\omega_1^{-1}\int_{-\infty}^{\frac{x-v}{\tau}}h_{n-1}(-u,\omega_1)\,du$$

for $n\geq 1$, with

$$\int_{-\infty}^{\frac{x-v}{\tau}}h_0(-u,\omega_1)\,du=-\omega_1^{-1}\exp(-\tfrac{x-v}{\tau}\omega_1)\cdot\Phi(\tfrac{x-v}{\tau}-\omega_1)+ \quad (4)$$
$$+\omega_1^{-1}\exp(-\omega_1^2/2)\cdot\Phi(\tfrac{x-v}{\tau})$$

If $u_n = \int_{-\infty}^{\frac{x-v}{\tau}} h_n(u,\omega_2)\,du$, for $n \geq 0$, by (3) we obtain u_0 and by Lemma 1C (Appendix C), for $n \geq 1$, we get

$$u_n = \beta_2\beta_3\cdots\beta_n \cdot [\beta_1 u_0 + A(\alpha_1 + \alpha_2/\beta_2 + \alpha_3/(\beta_2\beta_3) + \cdots + \alpha_n/(\beta_2\beta_3\cdots\beta_n)]$$

where $A = \omega_2^{-1}\exp(\frac{x-v}{\tau}\omega_2)$, $\alpha_n = \Psi_n(\frac{x-v}{\tau} + \omega_2)$,and $\beta_n = n\omega_n^{-1}$.

Also if $u'_n = \int_{-\infty}^{\frac{x-v}{\tau}} h_n(-u,\omega_1)\,du$, for $n \geq 0$, by (4) we get u'_0, and by Lemma 1C (Appendix C), for $n \geq 1$, we get

$$u'_n = \beta'_2\beta'_3\cdots\beta'_n \cdot [\beta'_1 u'_0 + A'(\alpha'_1 + \alpha'_2/\beta'_2 + \alpha'_3/(\beta'_2\beta'_3) + \cdots + \alpha'_n/(\beta'_2\beta'_3\cdots\beta'_n)]$$

where $A' = -\omega_1^{-1}\exp(-\frac{x-v}{\tau}\omega_1)$, $\alpha'_n = \Psi_n(-\frac{x-v}{\tau} + \omega_1)$,and $\beta'_n = n\omega_1^{-1}$.

Thus we obtain the following formula for cumulative distribution function $F(x)$,

$$F(x) = \theta\frac{\omega_1^{n+1}}{n!}K(\omega_1)u'_n + \eta\frac{\omega_2^{n+1}}{n!}K(\omega_2)u_n.$$

For $n = 0$ we obtain the cumulative distribution function $F(x)$ obtained by Luceño (1992). For $n = 1$ we have

$$u'_1 = -\omega_1^{-1}\exp(-\tfrac{x-v}{\tau}\omega_1)\left[\varphi(-\tfrac{x-v}{\tau} + \omega_1) + (\tfrac{x-v}{\tau} - \omega_1)\Phi(\tfrac{x-v}{\tau} + \omega_1)\right] + (\omega_1^2 K(\omega_1))^{-1}\Phi(\tfrac{x-v}{\tau}) - \omega_1^{-2}\exp(-\tfrac{x-v}{\tau}\omega_1)\Phi(\tfrac{x-v}{\tau} - \omega_1) \quad ,$$

and

$$u_1 = \omega_2^{-1}\exp(\tfrac{x-v}{\tau}\omega_2)\left[\varphi(\tfrac{x-v}{\tau} + \omega_2) - (\tfrac{x-v}{\tau} + \omega_2)\Phi(-\tfrac{x-v}{\tau} - \omega_2)\right] + (\omega_2^2 K(\omega_2))^{-1}\Phi(\tfrac{x-v}{\tau}) + \omega_2^{-2}\exp(\tfrac{x-v}{\tau}\omega_2)\Phi(-\tfrac{x-v}{\tau} - \omega_2) \quad .$$

The moment generating function and the moments of X can be obtained as in Luceño (1992), using the properties of conditional expectation Rao (1973, pp. 86-98). Thus the moment generating function of X is given by, according Luceño (1992, pp. 406) $m_X(t) = m_\xi(t)\cdot\exp(t^2\tau^2/2)$, i.e. X is the sum of two independent random variables, the first ξ, and the second one having a normal distribution with mean zero and variance τ^2. Using $m_\xi(t)$ given by section 3, we obtain

$$m_X(t) = \left[\eta\left(\frac{\lambda_2}{\lambda_2 + t}\right)^{n+1} + \theta\left(\frac{\lambda_1}{\lambda_1 - t}\right)^{n+1}\right]\cdot\exp(tv + t^2\tau^2/2)$$

for $-\lambda_2 < t < \lambda_1$.

Using Section 3, the mean and the variance of X are given by

$$E(X) = E_\xi(E_X(X/\xi)) = E_\xi(\xi) = \mu = v + (n+1)\left[\theta\lambda_1^{-1} - \eta\lambda_2^{-1}\right]$$
$$\mu_2(X) = Var(X) = Var_\xi(E_X(X/\xi)) - E_\xi(Var_X(X/\xi))$$
$$= Var_\xi(\xi) + E_\xi(\tau^2) = Var_\xi(\xi) + \tau^2$$
$$= (n+1)(n+2)\left[\theta\lambda_1^{-2} - \eta\lambda_2^{-2}\right] - -(n+1)^2\left[\theta\lambda_1^{-1} - \eta\lambda_2^{-1}\right]^2 + \tau^2.$$

The central moments of X are given by (Appendix C)

$$\mu_k(X) = \sum_{j=0}^{[k/2]} \binom{k}{2j} (2j-1)!! \cdot \tau^{2j} \cdot \mu_{k-2j}(\xi)$$

where $[k/2] = \begin{cases} k/2 & \text{if } k \text{ is even} \\ (k-1)/2 & \text{if } k \text{ is odd} \end{cases}$, and $\mu_j(\xi)$ is the central moment of order j of ξ and is defined by Lemma 3C (Appendix C). Also we have $\mu_3(X) = \mu_3(\xi)$ and $\mu_4(X) = \mu_4(\xi) + 6\tau^2\mu_2(\xi) + 3\tau^4$, where $\mu_2(\xi)$, $\mu_3(\xi)$ and $\mu_4(\xi)$ are given in section 3.

The skewness and kurtosis satisfy the equations (Luceño (1992))

$$\alpha_3(X) = \alpha_3(\xi)\left[\mu_2(\xi)/\mu_2(X)\right]^{3/2}$$
$$\alpha_4(X) = \alpha_4(\xi)\left[\mu_2(\xi)/\mu_2(X)\right]^2$$

where $\alpha_3(\xi)$ and $\alpha_4(\xi)$ are defined by section 3. Thus the following relations holds

$$(\alpha_3(X))^4(\alpha_4(X))^{-3} = (\alpha_3(\xi))^4(\alpha_4(\xi))^{-3}$$

Remark 4.1 The compound family of distribution has five parameters, namely θ, λ_1, λ_2, v and τ^2. Sometimes is desirable to restrict this set of parameters. For example, if $\theta = 0$ (or $\theta = 1$) the family has only three parameters because $\eta = 1 - \theta$, and then this skewness and kurtosis are negative and positive respectively.

Remark 4.2 If $\theta = \eta$ and $\lambda_1 = \lambda_2$, a different family with three parameters is found for which α_3 and α_3 are easily to be obtained.

Generally, for n fixed, when the available information permits it, we consider as in Luceño (1992) a set of covariates $c_1, ..., c_k$ and a new set of parameters $\delta_1, ..., \delta_p$ affecting the location parameter v, in the following way:

$v = g_n(\delta_1, ..., \delta_p; c_1, ..., c_k)$, where g_n is a known function. This equation would be a predictor that could be used together with a link function specific for the characteristic of the distribution to be estimated.

Remark 4.3 From the computational point of view, the probability density function and the cumulative distribution function depend on Ψ_n for which Lemma 5A is useful. For $\varphi(\cdot)$, the standard normal density function, very good approximations are known. Thus the quantiles of X and the log-likelihood function for a given sample, are easily computational. Also see Remark 5.4.

Remark 4.4 If $n = 0$ we obtain the family of compound probability distributions defined by Luceño (1992).

Remark 4.5 If $\theta = \eta$ and $n = 0$ we obtain a symmetrical family of probability distributions studied by Holla and Bhattacharya (1968). Also in this case, in Holla and Bhattacharya (1968), an expression for the distribution of the sum of k independent random variables each having this compound distribution is given.

Remark 4.6 We note that for $n = 0$ and $\theta = \eta$, with $\lambda_1 \neq \lambda_2$, the relation (1) described the probability density functions which are sometimes used (see McGill (1962)). The case $\lambda_1 = \lambda_2$ correspond to double exponential distribution (Johnson and Kotz (1970)).

5 Bayes estimation.

The posterior distribution of ξ, $\hat{f}(\xi|x) = g(\xi)f(x|\xi)/f(x)$, where $f(x|\xi)$ is a normal density function with mean ξ and variance τ^2, $g(\xi)$ is defined by (1) and $f(x)$ is defined by (2), is given by

$$\hat{f}(\xi|x) = \begin{cases} \theta \frac{\lambda_1^{n+1} \cdot \exp(\lambda_1^2\tau^2/2)\exp(\lambda_1(v-x)) \cdot (\xi - v)^n}{n!\sqrt{\tau^2 2\pi} \cdot f(x) \cdot \exp([\xi - (x - \lambda_1\tau^2)]^2/(2\tau^2))} & , \text{ if } \xi > v \\ \eta \frac{\lambda_2^{n+1} \cdot \exp(\lambda_2^2\tau^2/2)\exp(\lambda_2(x-v)) \cdot (v - \xi)^n}{n!\sqrt{\tau^2 2\pi} \cdot f(x) \cdot \exp([\xi - (x + \lambda_2\tau^2)]^2/(2\tau^2))} & , \text{ if } \xi < v \end{cases} \quad (5)$$

We note that the family (1) is a family of piecewise conjugate priors of order 2 (Meeden (1992)) relative to the mean of normal distribution with given variance. See Dalal and Hall (1983) , Diaconis and Ylvisaker (1984) and Meeden (1992) for some elicitation procedures using piecewise conjugate priors. From standard decision theory, e.g. Berger (1980), the Bayes estimate of ξ based on observation x from $N(\xi, \tau^2)$ under a prior density $g(\xi)$ given

by (1) and quadratic loss function is given by the mean $m(x)$ of $\hat{f}(\xi|x)$, the posterior distribution of ξ. Using (5) and Lemma 3B (Appendix B) we obtain

$$m(x) = \int_{-\infty}^{\infty} \xi \hat{f}(\xi|x)\, d\xi = \theta \frac{(\lambda_1\tau)^{n+1}\exp((\lambda_1\tau)^2/2)}{n!\tau\exp(\lambda_1\tau\frac{x-v}{\tau})} \cdot (f(x))^{-1}$$
$$[\tau\Psi_{n+1}(-\tfrac{x-v}{\tau} + \lambda_1\tau) + v\Psi_n(-\tfrac{x-v}{\tau} + \lambda_1\tau)]$$
$$-\eta\frac{(\lambda_2\tau)^{n+1}\exp((\lambda_2\tau)^2/2)}{n!\tau\exp(\lambda_2\tau\frac{v-x}{\tau})} \cdot (f(x))^{-1}.$$
$$[\tau\Psi_{n+1}(\tfrac{x-v}{\tau} + \lambda_2\tau) - v\Psi_n(\tfrac{x-v}{\tau} + \lambda_2\tau)].$$

Using Berger(1980, Theorem 3 ,pp.122), we get that this Bayes estimate $m(x)$ is admissible. . Also $m(x)$ is the regression function of ξ on X (Cacoullos and Papagiorgiou (1984)).Recently, Diebolt and Robert(1994) present approximation methods which evaluate the posterior distribution and Bayes estimators by Gibbs sampling, relying on the missing data structure of the mixture model.

6 Moment and ML estimates.

The estimation of the parameters of finite mixture distributions has recently considered. Mixture models provide an interesting alternative to nonparametric modelling, while being less restrictive than the usual distributional assumptions. The maximum likelihood estimates (ML estimates) of the parameters can be found using a standard algorithm for minimizing or maximizing a function of several variables. For the convergence of the algorithm, some good starting values for the parameters are necessary. In Radhakrishna et al (1992) an attempt to derive the moment and ML estimators of the parameters of two-component mixture is made.

In the following we consider these techniques to derive moment and ML estimators of the parameters λ_1, λ_2, θ, v and τ^2.

a) *Moment estimators of the parameters.*

Let μ_j' be the j^{th} row moment of X about the origin. Thus we consider the first sample moments and equate them to the corresponding population moments to obtain the moment equations for the five parameters λ_1, λ_2, θ, v and τ^2.

Let m_j, $1 \leq j \leq 5$, be the first five samples moments. The moment equations to be solved simultaneously for λ_1, λ_2, θ, v and τ^2 are $h_j(\omega) = 0$,

for $1 \leq j \leq 5$, where $h_j(\omega) = \mu'_j - m_j$ and $\omega = (\lambda_1, \lambda_2, \theta, v, \tau^2)^\top$ (here $\top$ denotes the transpose).

This system of equations can be solved using Newton-Raphson or the method of steepest descent. By these procedures we obtain the solution in the following way. Using the Newton-Raphson method we have

$$\omega^{i+1} = \omega^i - \left(W(\omega^i)\right)^\top \cdot h(\omega^i) \quad , \quad i = 0, 1, 2, ...$$

where ω^0 is the zeroth approximation, $h = (h_1, ..., h_5)^\top$ and $W(\omega^i)$ is the Jacobian matrix calculated at ω^i. For the method of steepest descent we have

$$\omega^{i+1} = \omega^i - \alpha_i \left(W(\omega^i)\right)^\top \cdot h(\omega^i) \quad , \quad i = 0, 1, 2, ...$$

where

$$\alpha_i = \frac{\langle h(\omega^i), W(\omega^i)\,(W(\omega^i))^\top h(\omega^i)\rangle}{\langle W(\omega^i)\,(W(\omega^i))^\top h(\omega^i), W(\omega^i)\,(W(\omega^i))^\top h(\omega^i)\rangle}$$

with $\langle\ ,\rangle$ for inner product.

A procedure is terminated as soon as the norm of ω^i is less than a pre-assigned small positive quantity ϵ.

b) ML *estimators of the parameters.*

In the following we derive ML estimator of vector ω . Let $x_1, x_2, ..., x_m$ be a random sample of size m draw from X. Then the likelihood function is given by $L = \prod_{i=1}^{m} f(x_i)$ and log-likelihood function is $\log L = \sum_{i=1}^{m} \log f(x_i)$, where we note that f is a mixture of two probability density functions, $f(x) = \theta f_1(x) + \eta f_2(x)$, with

$$f_1(x) = f_1(x|\omega) = \frac{(\lambda_1\tau)^{n+1}}{n!\tau} K(\lambda_1\tau) \cdot h_n(-\tfrac{x-v}{\tau}, \lambda_1\tau) \quad ,$$
$$f_2(x) = f_2(x|\omega) = \frac{(\lambda_2\tau)^{n+1}}{n!\tau} K(\lambda_2\tau) \cdot h_n(\tfrac{x-v}{\tau}, \lambda_2\tau) \quad .$$

We put $u^i(\omega) = f(x_i)$, $1 \leq i \leq m$, and $u^i_j(\omega)$, $1 \leq j \leq 5$, the derivatives of $u^i(\omega)$ with respect to the component of order j for vector ω, respectively. Thus the ML equations to be solved are given by

$$\frac{\partial \log L}{\partial \omega_j} = \sum_{i=1}^{m} \frac{u^i_j(\omega)}{u^i(\omega)} = 0 \, , \quad 1 \leq j \leq 5 \quad .$$

Because f_1 and f_2 are not dependent of all parameters, this system have a special form. Detailed expressions of $u_j^i(\omega)$ are omitted.

Solving simultaneously these equations, using for example one of the iterative techniques of Newton-Raphson or the method of steepest descent presented above, we obtain ML estimates of the vector parameter ω .

Remark 6.1 For the calculation of ML estimators, also the EM algorithm (Dempster et al (1977)) is it appropriate to be apply. Thus on this line is it important to bear in mind the Newton-Raphson version and the method of scoring presented, for example in Titterington et al (1985).

Remark 6.2 The information matrix can be obtained by evaluating the Jacobian matrix W at ML estimators of the parameters and changing the sign of all the elements. Further, by inverting the information matrix we get the approximate asymptotic variance-covariance matrix of the estimators.

Remark 6.3 An analysis of models with fully categorized data (incomplete data) can be given. In this case the data are represented as $\{x_i,\ i = 1,2,...,m\} = \{(y_i, z_i),\ i = 1,2,...,m\}$, where each $z_i = (z_{ij})_{1\leq j\leq q}$ is an indicator vector of length q with 1 in the position corresponding to the appropriate category and zeros elsewhere. For the likelihood function corresponding to $(x_1,...,x_m)$ see Titterington et al (1985, pp.84).

Remark 6.4 When computing the log-likelihood function, the value of $\ln \Psi_n(s)$ should be evaluated iteratively. Using Lemma 2A we have

$$w_{n+1} = w_n + \ln[-s - w_n'(s)]$$

where $w_n(s) = \ln \Psi_n(s)$ and $w_n'(s) = \Psi_n'(s)/\Psi_n(s)$. Here $w_n'(s)$ is approximate by $(w_n(s+\Delta) - w_n(s)) \cdot \Delta^{-1}$, for small $\Delta > 0$.

Because log-likelihood function takes the form $\log f = \log(\theta f_1 + \eta f_2)$, for $\theta \in (0,1)$, the computations should proceed as in Luceño (1992), in order to avoid rounding errors. Thus $\ln f = \ln(\theta f_1) + \ln[1 + (\eta f_2/\theta f_1)]$, $\eta f_2/\theta f_1 = \exp[\ln(\eta f_2) - \ln(\theta f_1)]$ where $\ln(\theta f_1)$ and $\ln(\eta f_2)$ are evaluated directly using w_n.

Remark 6.5 Also when the predictor $v = g_n(\delta_1,...,\delta_p, c_1,...,c_k)$ (see Remark 3.2) is given, it is possible to used a parametric estimation procedure defined in Luceño (1992, pp.379).

Remark 6.6 Often it is convenient to restrict the values of the parameters in some specific way depending on the characteristics of the data (see

Remarks 3.1 and 3.2). For example $\theta = 0$; $\theta = 1$; $\theta = \eta$ and $\lambda_1 = \lambda_2$. In these cases the moment estimates and ML estimates follow easily.

7 An James estimator for proportion θ.

Mixtures of distributions have received considerable attention in recent years due to the fact that practical problems involving mixtures arise in biology, engineering, fisheries, medicine, social sciences (see Choi and Bulgren (1968) and McDonald (1971)) .James (1978) proposed certain estimators based on some rapid measurements like (i) the number of observations smaller than a fixed point m ,(ii) the number of observations smaller than n_2 and greater than n_1 , (iii) the sample mean (using the method of moments). Although estimators of θ based on (i), (ii) and (iii) have been previously considered by McDonald (1971), Odell and Basu (1976), James (1978) discussed the choice of m in (i) and (n_1, n_2) in (ii) and efficiencies of these estimators relative to the maximum likelihood estimator $\hat{\theta}$ of θ . In this section we obtain an estimator based on (i) through minimax criterion. According to Ahmad et al (1983) we adopt the criterion that an estimate $\tilde{\theta} = \tilde{\theta}(x_1, x_2, ..., x_m)$ of θ in a class of estimators Θ is optimum if it minimizes the supremum with respect to θ of the asymptotic variance of $\tilde{\theta}$ over Θ. Here we consider the problem of estimating the proportion θ in (2) when λ_1 ,λ_2 and v are known and $n = 1$, on the basis of independent observations from $f(x)$.

Let $x_1, x_2, ..., x_m$ be a sequence of m independent observations from the population with density function $f(x) = \theta f_1(x) + \eta f_2(x)$ (with $\eta = 1 - \theta$), where f_1 and f_2 are defined by

$$f_1(x) = \tau^{-1}(\lambda_1\tau)^2 K(\lambda_1\tau) h_1(-\tfrac{x-v}{\tau}, \lambda_1\tau)$$
$$f_2(x) = \tau^{-1}(\lambda_2\tau)^2 K(\lambda_2\tau) h_1(\tfrac{x-v}{\tau}, \lambda_2\tau)$$

for $-\infty < x < \infty$. Let us consider the following set

$$\mathcal{X} = \{x | F_1(x) \neq F_2(x)\}$$

where F_i is the distribution function corresponding to f_i , $i = 1, 2$. (Using Section 4 it is easy to obtain F_i, $i = 1, 2$).

Thus the James' estimator $\tilde{\theta}_x$ for $x \in \mathcal{X}$ is defined as (Ahmad et al (1983))

$$\tilde{\theta}_x = \begin{cases} \theta_x^* & , \quad \text{if } 0 < \theta_x^* < 1; \\ 0 & , \quad \text{if } \theta_x^* \leq 0; \\ 1 & , \quad \text{if } \theta_x^* \geq 1, \end{cases}$$

where $\theta_x^* = \frac{\hat{F}(x) - F_2(x)}{F_1(x) - F_2(x)}$, with $\hat{F}(x) = m^{-1} \cdot \sum_{j=1}^{m} Y_j$, $Y_j = 1$, if $x_j \leq x$, and is zero otherwise. Also the asymptotic variance of $\tilde{\theta}_x$ is given by

$$\frac{z_\theta(x)}{n} = \frac{1}{n} \cdot \frac{[\theta F_1(x) + \eta F_2(x)] \cdot [\theta \bar{F}_1(x) + \eta \bar{F}_2(x)]}{[F_1(x) - F_2(x)]^2}$$

where $\bar{F}_i(x) = 1 - F_i(x)$, $i = 1, 2$, and $\eta = 1 - \theta$.

According to Ahmad et al (1983) we consider an optimal estimator in the class $\Theta(\mathcal{X}) = \{\tilde{\theta}_x | x \in \mathcal{X}\}$. Using some transformations we obtain

$$F_1(x) = \Phi(\tfrac{x-v}{\tau}) - \{\omega_1 K(\omega_1) \exp(-\tfrac{x-v}{\tau}\omega_1) \cdot [\varphi(-\tfrac{x-v}{\tau} + \omega_1) + \\ + (\tfrac{x-v}{\tau} - \omega_1)\Phi(\tfrac{x-v}{\tau} - \omega_1)] + \exp(-\tfrac{x-v}{\tau}\omega_1)\Phi(\tfrac{x-v}{\tau} - \omega_1)K(\omega_1)$$

$$F_2(x) = \Phi(\tfrac{x-v}{\tau}) + \{\omega_2 K(\omega_2) \exp(\tfrac{x-v}{\tau}\omega_2) \cdot [\varphi(\tfrac{x-v}{\tau} + \omega_2) - \\ - (\tfrac{x-v}{\tau} + \omega_2)\Phi(-\tfrac{x-v}{\tau} - \omega_2)] + \exp(\tfrac{x-v}{\tau}\omega_2)\Phi(-\tfrac{x-v}{\tau} - \omega_1)K(\omega_2)$$

Because $\delta(t) = \varphi(t) - t\Phi(-t)$, $-\infty < t < \infty$, the derivative $\delta'(t) = -\Phi(-t) < 0$, $\lim_{t \to -\infty} \delta(t) = +\infty$, and $\lim_{t \to \infty} \delta(t) = 0$ we obtain $\delta(t) > 0$ for any t. Hence $F_2(x) > F_1(x)$ for any x.

Now according to Ahmad et al (1983, Lemma 2.1, case *iii*) we have

$$\sup_{0<\theta<1} z_\theta(t) = \begin{cases} 4^{-1}(F_1(t) - F_2(t))^{-2} & , \quad \text{for } t_2 \leq t \leq t_1; \\ F_1(t)\bar{F}_1(t)(F_1(t) - F_2(t))^{-2} & , \quad \text{for } t > t_1; \\ F_2(t)\bar{F}_2(t)(F_1(t) - F_2(t))^{-2} & , \quad \text{for } t < t_2, \end{cases}$$

where t_1 and t_2 are the respective medians for F_1 and F_2. Also, using Ahmad et al (1983, Remark 2) we obtain a good choice by finding a t^0 such that $f_1(t^0) = f_2(t^0)$. Generally is not easy to obtain t^0 directly. Some numerical approximations are possible However, if $\lambda_1 = \lambda_2$, we may take $t^0 = v$. Hence James' estimator is $\tilde{\theta}_{t^0}$ defined as above.

A similar study is possible for case $n = 0$.

Appendix A

In this section we consider some properties relative to Ψ_n defined in Section 2, Let H be a normal distribution with mean μ and variance τ^2. Then the transform of order n associate to H is

$$\Psi_n(s) = \int_s^\infty (x-s)^n \varphi(x)dx$$

where $\varphi(x)$ is the standard normal density function.

Lemma 1A. For any real s and $n \geq 1$ (integer number) we have

$$\Psi_{n+1}(s) = -s\Psi_n(s) + n\Psi_{n-1}(s)$$

with $\Psi_0(s) = 1 - \Phi(s)$ $(= \Phi(-s))$, $\Psi_1(s) = \varphi(s) - s[1 - \Phi(s)]$.

Proof. We obtain easily Ψ_0 and Ψ_1. For $n \geq 1$,

$$\Psi_{n+1}(s) = \int_s^\infty (t-s)^n (t-s)\varphi(t)dt = -s\int_s^\infty (t-s)^n \varphi(t)dt + \int_s^\infty t(t-s)^n \varphi(t)dt =$$
$$= -s\Psi_n(s) - \int_s^\infty (t-s)^n \varphi'(t)dt = -s\Psi_n(s) + n\int_s^\infty (t-s)^{n-1}\varphi(t)dt =$$
$$= -s\Psi_n(s) + n\Psi_{n-1}(s).$$

Lemma 2A. $\Psi_n(s)$ is a derivable function for any real s and $\Psi_n'(s) = -n\Psi_{n-1}(s)$, where $n \geq 1$.

Proof. Using a change of variable we obtain $\Psi_n(s) = \int_0^\infty z^n \varphi(s+z)dz$. Because $\varphi'(t) = -t\varphi(t)$ we get $\Psi_n'(s) = -\int_0^\infty z^n(s+z)\varphi(s+z)dz =$ $-s\int_0^\infty z^n\varphi(s+z)dz - \int_0^\infty z^{n+1}\varphi(s+z)dz = -s\Psi_n(s) - \Psi_{n+1}(s)$. Finally we use Lemma 1A.

Lemma 3A. We put $a_k = \int_{-\infty}^{\infty} t^k \varphi(t)dt$, where k is a positive integer. Then $a_{2p} = (2p-1)!!$ and $a_{2p+1} = 0$, where p is a positive integer, $(2p-1)!! = (2p-1)(2p-3)\cdot\cdots\cdot 3\cdot 1$, with $(-1)!! = 1$.

Lemma 4A. Let $I_k(s) = \int_{-\infty}^{s} t^k \varphi(t)dt$, where $-\infty < s < \infty$ and k is a positive integer. Then

a) $I_k(s) = (k-1)I_{k-2}(s) - s^{k-1}\varphi(s)$, $k \geq 2$, with $I_0(s) = \Phi(s)$, $I_1(s) = -\varphi(s)$;

b) $I_{2p}(s) = (2p-1)!!I_0(s) - \varphi(s)[s(2p-1)!! + s^3(2p-1)(2p-3)\cdots 7\cdot 5 + \cdots + s^{2p-3}(2p-1) + s^{2p-1}]$;

c) $I_{2p+1}(s) = (2p)!!I_1(s) - \varphi(s)[s^2(2p)(2p-2)\cdots 6\cdot 4 + s^4(2p)(2p-2)\cdots 8\cdot 6 + \cdots + s^{2p-2}(2p) + s^{2p}]$.

The proofs of Lemma 3A and Lemma 4A are simple exercises.

Lemma 5A. We have $\Psi_n(s) = \sum_{k=0}^{n}(-1)^k \binom{n}{k} s^k [a_{n-k} - I_{n-k}(s)]$.

Proof. We can write $\Psi_n(s) = \int_{-\infty}^{\infty}(t-s)^n\varphi(t)dt - \int_{-\infty}^{s}(t-s)^n\varphi(t)dt = \sum_{k=0}^{n}(-1)^k\binom{n}{k}s^k\left[\int_{-\infty}^{\infty} t^{n-k}\varphi(t)dt - \int_{-\infty}^{s} t^{n-k}\varphi(t)dt\right]$. Now we use Lemmas 3A and 4A.

Lemma 6A. We have $\Psi_n(-s) = \sum_{k=0}^{n}(-1)^k\binom{n}{k}s^{n-k}a_k - (-1)^n\Psi_n(s)$.

Proof. We see that $\Psi_n(-s) = \int_{-s}^{\infty}(t+s)^n\varphi(t)dt = \int_{-\infty}^{s}(s-t)^n\varphi(t)dt = = \int_{-\infty}^{\infty}(s-t)^n\varphi(t)dt - \int_{s}^{\infty}(s-t)^n\varphi(t)dt = \sum_{k=0}^{n}(-1)^k\binom{n}{k}s^{n-k}\int_{-\infty}^{\infty} t^k\varphi(t)dt - (-1)^n\Psi_n(s)$ and we use Lemma 3A.

Lemma 7A. Let us consider an integer number $n \geq 0$, and m, τ real numbers such that $\tau \neq 0$. Then

$$\int_0^{\infty} z^n \exp\left(-\frac{(z-m)^2}{2\tau^2}\right) dz = \tau^{n+1}\sqrt{2\pi}\Psi_n(-m/\tau)\,.$$

Proof. We use a change of variable $z = m+\tau t$. Thus the integral reduces to $\tau \int_{-m/\tau}^{\infty} (m+t\tau)^n \exp(-t^2/2)dt = \tau^{n+1} \int_{-m/\tau}^{\infty} [t-(-m/\tau)]^n \exp(-t^2/2)dt = \tau^{n+1}\sqrt{2\pi}\Psi_n(-m/\tau)$.

Appendix B

Lemma 1B. Let be $J_n = \int_0^\infty z^n \exp(-\lambda z)dz$, where n is a positive integer and $\lambda > 0$. Then $J_n = n! \cdot \lambda^{-(n+1)}$.

Proof. We have $J_0 = \lambda^{-1}$ and $J_n = n\lambda^{-1} \cdot J_{n-1}$.

Lemma 2B. For $-\lambda_2 < t < \lambda_1$ we have

a) $\int_v^\infty (\xi - v)^n \exp[\xi(t - \lambda_1) + \lambda_1 v]d\xi = \exp(vt) \cdot n! \cdot (\lambda_1 - t)^{-n-1}$;

b) $\int_{-\infty}^v (v - \xi)^n \exp[\xi(t + \lambda_2) - \lambda_2 v]d\xi = \exp(vt) \cdot n! \cdot (t + \lambda_2)^{-n-1}$.

Proof. It results from Lemma 1B.

Lemma 3B. We have

a) $\int_{-\infty}^v (v - \xi)^n \exp[-\lambda_2(v - \xi)] \exp[-(x - \xi)^2/2\tau^2]d\xi =$

$= \tau^{n+1}\sqrt{2\pi} \exp(\lambda_2 \tau \frac{x-v}{\tau}) \exp(\frac{(\lambda_2\tau)^2}{2}) \cdot \Psi_n(\frac{x-v}{\tau} + \lambda_2\tau)$;

b) $\int_v^\infty (\xi - v)^n \exp[-\lambda_1(\xi - v)] \exp[-(x - \xi)^2/2\tau^2]d\xi =$

$= \tau^{n+1}\sqrt{2\pi} \exp(-\lambda_1(x - v)) \cdot \exp(\frac{(\lambda_1\tau)^2}{2}) \cdot \Psi_n(-\frac{x-v}{\tau} + \lambda_1\tau)$.

Proof. It follows from Lemma 7A.

Appendix C

Lemma 1C. Let $(\alpha_n)_{n\geq 1}$ and $(\beta_n)_{n\geq 1}$ be two real numbers sequences with $\beta_n > 0$ for any n. Also let A be a positive real number. We define $u_n = A\alpha_n + \beta_n u_{n-1}$, for $n \geq 1$, with u_0 a given real number. Then

$$u_n = \beta_2\beta_3\cdots\beta_n \cdot [\beta_1 u_0 + A(\alpha_1 + \alpha_2/\beta_2 + \alpha_3/(\beta_2\beta_3) + \cdots + \alpha_n/(\beta_2\beta_3\cdots\beta_n)] \,.$$

Proof. It is immediate by induction.

According to Luceño (1992, pp. 407, 408) we have

Lemma 2C. The k- th moments of X about a given constant a is

$$\alpha_{k,a}(X) = E\left[(X-a)^k\right] = \sum_{j=0}^{[k/2]} (2j-1)!! \binom{k}{2j} \tau^{2j} \alpha_{k-2j,a}(\xi) \quad ,$$

where $\alpha_{k,a}(\xi) = \sum_{j=0}^{k} \binom{k}{j} \alpha_{k-j,v}(\xi)(v-a)^j$ with $\alpha_{j,v}(\xi) = \frac{(n+j)!}{n!}\left[\theta\lambda_1^{-j} + (-1)^j \eta \lambda_2^{-j}\right]$.
Here $\alpha_{k,a}(\xi)$ is the k- th moment of ξ about a given constant a.

Lemma 3C. The central moments for ξ and X are

$$\mu_k(\xi) = \sum_{i=0}^{k} (-1)^i \binom{k}{i} (\alpha_{1,v}(\xi))^i \alpha_{k-i,v}(\xi) \quad ,$$

and

$$\mu_k(X) = \sum_{i=0}^{k} (-1)^i \binom{k}{i} (\alpha_{1,v}(X))^i \alpha_{k-i,v}(X) \quad .$$

Acknowledgments We are grateful to an anonymous referee for the critical and several helpful comments on the original version of this paper.

References

[1] M. Ahmad, N. Giri and B. K. Sinha (1983). *Estimation of the mixing proportion of two known disributions*, Sankhya, Series A, 45, Part 3, 337-371.

[2] J. Berger (1980). Statistical Decision Theory, Springer-Verlag, New-York.

[3] J. Burbea and C. R. Rao (1982). *Entropy differential metric, distance and divergence measures in probability spaces - a unified approach*, J. Multivariate Anal., 12, 575-596.

[4] T. Cacoullos and H. Papageorgiou (1984). *Characterizations of mixtures of continuous distributions by their posterior means*, Scand. Actuarial Journal,23-30.

[5] K. Choi and W. G. Bulgren (1968). *An estimation for the mixture of distribution*, J. Roy. Stat. Soc. B, 30, 444-460.

[6] S. R. Dalal and W. J. Hall (1983). *Approximating priors by mixtures of natural conjugate priors*, J. Roy. Stat. Soc. B, 45, 278-286.

[7] M. H. De Groot (1970). *Optimal Statistical Decisions*, Mc Graw-Hill, New York.

[8] A. P. Dempster, N. M. Laird and D. B. Rubin (1977). *Maximum likelihood estimation from incomplete data via the EM algorithm (with discussion)*, J. R. Statist. Soc. B 39, 1-38.

[9] P. Diaconis and D. Ylvisaker (1984). *Quantifying prior opinions*. In *Bayesian Statistics II*, J. M. Bernardo, M. H. De Groot, D. V. Lindley and A. F. M. Smith (Eds). North -Holland, Amsterdam.

[10] J. Diebolt and C. P. Robert (1994). *Estimation of finite mixture distributions through Bayesian sampling*, J. Roy. Statist. Soc.,B,56, no.2, 363-375.

[11] S. Guiaşu (1978). *Information Theory and Applications*, Mc Graw-Hill.

[12] M. S. Holla and S. K. Bhattacharya (1968). *On a compound Gaussian distribution*, Annals of the Institute of Statistical Mathematics, Tokyo, 20, 331-336.

[13] I. R. James (1978). *Estimation of the mixing proportion in a mixture of two normal distributions for simple, rapid measurements*, Biometrics, 34, 265-275.

[14] N. L. Johnson and S. Kotz (1970). *Continuous Univariate Distributions* - 1, 2, New York.

[15] A. Luceño (1992). *A new family of probability distribution with applications to data analysis*, Commun. Statist. - Theory Meth., 21(2), 391-409.

[16] P. D. M. McDonald (1971). *Comment on an estimation procedure for mixture of distributions by Choi and Bulgren*, J. Roy. Statist. Soc.,B, 33, 326-329.

[17] W. J. McGill (1962). *Random fluctuations of response rate*, Psychometric, 27, 3-17.

[18] G. Meeden (1992). *An elicitation procedure using piecewise conjugate priors.* In *Bayesian Analysis in Statistics and Econometrics*,P. K. Goel and L. Iyengar (Eds).

[19] A. Plucinska (1966). On a general form of the probability density function and its application to the investigation of the distribution of rheostat resistance, Zastosowania Matematyki, 9, 9-19.(In Polish)

[20] A. Plucinska (1967). The reliability of a compound system under consideration of the system elements prices, Zastosowania Matematyki, 9, 123-134.(In Polish)

[21] C. Radhakrishna, A. V. Dattatreya Rao and G. V. S. R. Anjaneyulu (1992). *Estimation of parameters in a two-component mixture generalized gamma distribution*, Commun. Statist. - Theory Meth., 21(6), 1799-1805.

[22] C. R. Rao (1973). *Linear Statistical Inference and its Applications*, New York, John Wiley & Sons Inc.

[23] G. Saporta (1990). *Probabilites, Analyse des Donnees et Statistique*, Editions Technip, Paris.

[24] D. M. Titterington, A. F. M. Smith and U. E. Makov (1985). *Statistical Analysis of Finite Mixture Distributions*, John Wiley & Sons.