Statistics & Decisions 19, 71 – 82 (2001)

REGRESSIONAL IDENTIFIABILITY AND IDENTIFICATION FOR BETA MIXTURES

Arjun K. Gupta and Jacek Wesołowski

Received: Revised version: March 27, 2000

Abstract

First kind beta mixtures with the mixing parameter being the right extremity of the support are considered. Identifiability of such a mixture by the posterior mean is established under the assumption of absolute continuity. For linear posterior means a complete description of possible prior (consequently joint) distributions is given without the absolute continuity assumption - the class of priors includes the gamma and both kinds of beta laws.

1 Introduction

Identification and identifiability are the basic questions in statistical modelling. A wide review on this subject can be found in Prakasa Rao (1992). Here we are interested in mixture models. For a random vector (X, Y) denote by $\mu_{X|Y}$ the conditional distribution of X given Y. In such a setting X (or its distribution μ_X) is called a mixture with respect to Y (or its distribution μ_Y).

The classical mixture identifiability question is concerned with a one to one correspondence between μ_X and μ_Y under given $\mu_{X|Y}$ - see for instance Teicher (1961), Barndorff-Nielsen (1965) or Patil and Bildikar (1966).

Here we study identifiability and identification taking into account the posterior mean $E(Y|X)$ instead of the marginal μ_X. Investigations of such problems go back to Johnson (1957), where a Poisson mixture with linearity of posterior mean was studied. Since then different mixture models were considered in numerous papers, some dealing with uniqueness of μ_Y (identifiablity), others with characterization of classes of distributions μ_Y for a given class of posterior means, often assumed to be linear functions (identification). See for instance: Krishnaji (1974), Korwar (1975, 1977), Xekalaki (1983), Cacoullos and Papageorgiou (1983, 1984), Papageorgiou (1984, 1985), Kyriakoussis and Papageorgiou (1991), Johnson and Kotz (1992), Arnold et al. (1993), Sapatinas (1995), Wesołowski

AMS (1991) subject classification: Primary 62H05; Secondary 62E10, 62F15.
Key words and phrases: mixtures, posterior mean, identifiability, characterization of probability distribution, beta conditional distribution, conditional moments, linearity of regression, gamma distribution, beta type distributions.

(1995a, b, 1996a, b), Gupta and Wesołowski (1997, 1999), Papageorgiou and Wesołowski (1997).

In this paper we are interested in beta mixtures with a mixing variable being the right end of the support. A uniform mixture is a special kind of such a model. Identifiability and identification problems for uniform mixtures were considered, using other methods, in Gupta and Wesołowski (1997). Here we characterize completely the family of models with linear posterior mean - this is done in Section 2. It consists of the beta distributions of both kinds and the gamma distribution. The models obtained in that way possess an intriguing continuity property as the slope tends to 1, allowing to aproximate gamma density by beta type densities. If it is assumed additionally that the prior is absolutely continuous then the question of identifiability by any consistent posterior mean can be settled - see Section 3. Here an approach often applied in earlier papers - see for example Cacoullos and Papageorgiou (1984) - was based on studying the distribution of X and then making use of identifiability of a mixture via the posterior distribution. Instead we are interested directly in deriving the uniqueness of the distribution of the prior. Our approach is based on the Laplace transform technique and the method of moments and not on solving functional equations for densities, as it was done in the paper Gupta and Wesołowski (1997) in the case of the uniform mixture.

2 Identification by linearity of posterior mean

Introduce first the notation for three distributions which play the crucial role in this paper. Denote by:
- $B_I(p,q,r)$ the first kind beta distribution defined by the density

$$f(x) = \frac{x^{p-1}(r-x)^{q-1}}{B(p,q)r^{p+q-1}} I_{(0,r)}(x),$$

where p, q, r are positive numbers;
- $B_{II}(p,q,r)$ the second kind beta distribution defined by the density

$$f(x) = \frac{r^q x^{p-1}}{B(p,q)(r+x)^{p+q}} I_{(0,\infty)}(x),$$

where p, q, r are positive numbers;
- $G(\alpha,p)$ the gamma distribution defined by the density

$$f(x) = \frac{\alpha^p}{\Gamma(p)} x^{p-1} e^{-\alpha x} I_{(0,\infty)}(x),$$

where α, p are positive numbers.

Let $X = V_1$ and $Y = V_1 + V_2$, where V_1, V_2 are i.i.d. gamma $G(\alpha, p_i)$, $i = 1, 2$, r.v's. Then it is not difficult to observe that X is a first kind beta mixture of Y of the form

$$\mu_{X|Y} = B_I(p_1, p_2, Y)\ , \tag{2.1}$$

and

$$E(Y|X) = X + p_2/\alpha\ .$$

Here we are interested in the converse to the above observations without assuming the summation scheme.

At first let us consider the question of identification of the beta mixture (2.1) via a general linear form of the posterior mean

$$E(Y|X) = aX + b\ , \tag{2.2}$$

where a and b are some real constants. When $\mu_{X|Y}$ is known it suffices to know the prior distribution to determine the bivariate measure.

Such problems, involving linearity of regression, were discussed for other mixtures: binomial and Pascal in Korwar (1975), quasi-binomial in Korwar (1977), normal in Cacoullos and Papageorgiou (1984), Pareto in Wesołowski (1995a), second kind beta in Wesołowski (1996b), uniform in Gupta and Wesołowski (1997).

From (2.1) observe that $E(X) = 0$ iff $E(Y) = 0$ iff $P(X = Y = 0) = 1$. Hence to get rid off that degenerate case we will assume that $E(X)$ is non-zero.

Theorem 1 *Let* (X, Y) *be a random vector satisfying* (2.1) *and* (2.2) *for some constants* a *and* b. *Assume additionally that* $0 < E(X) < \infty$. *Then* $b > 0$ *and only the following cases are possible:*
(i) $1 < a < 1 + p_2/p_1$ *and*

$$Y \sim B_{II}\left(p_1 + p_2, 1 + \frac{p_2}{a-1} - p_1, \frac{b}{a-1}\right); \quad X \sim B_{II}\left(p_1, 1 + \frac{p_2}{a-1} - p_1, \frac{b}{a-1}\right);$$

(ii) $a = 1$ *and*

$$Y \sim G\left(\frac{p_2}{b}, p_1 + p_2\right); \quad X \sim G\left(\frac{p_2}{b}, p_1\right);$$

(iii) $0 < a < 1$ *and*

$$Y \sim B_I\left(p_1 + p_2, \frac{ap_2}{1-a}, \frac{b}{1-a}\right); \quad X \sim B_I\left(p_1, \frac{p_2}{1-a}, \frac{b}{1-a}\right);$$

(iv) $a = 0$ *and*

$$P(Y = b) = 1, \quad X \sim B_I(p_1, p_2, b).$$

Proof. Observe that it follows immediately from (2.1) that $supp(X) = [0, T] \subseteq [0, \infty)$ (if $T = \infty$ then take $[0, T)$ instead of $[0, T]$, obviously) and $\inf\{y : F_Y(y) = 1\} = T$, where F_Y is the distribution function of Y. Note also that since $X \leq Y$ a.s. then for $T < \infty$ it follows that $E(Y|X = T) = T$ - the fact we will use quite often in the sequel. Now (2.1) and (2.2), via the generalized Bayes rule, result in the equation

$$\int_x^\infty (y - x)^{p_2}\, dH(y) = [(a-1)x + b] \int_x^\infty (y - x)^{p_2 - 1}\, dH(y), \tag{2.3}$$

where $dH(y) = y^{1 - p_1 - p_2}\, dF_Y(y)$, which holds μ_X a.s. Since its both sides are continuous in x, it can be assumed that (2.3) is satisfied for all $x \in [0, T]$.

Before the detailed analysis of the above equation let us derive some basic limitations for the coefficients a and b.

Observe that $b \leq 0$ is impossible: for $b < 0$ one gets negative values at the rhs of (2.2) with positive probability (X close to zero); that $b = 0$ is impossible can be noticed due to (2.3) by taking $x = 0$, then the lhs has to be zero which contradicts the assumption $E(X) > 0$.

Find now the relations between $E(X)$ and $E(Y)$ using (2.1) and then (2.2):

$$E(Y) = aE(X) + b, \qquad (p_1 + p_2)E(X) = p_1 E(Y).$$

Consequently

$$E(X)(1 + p_2/p_1 - a) = b > 0$$

and thus $a < 1 + p_2/p_1$. To see that $a < 0$ is impossible consider first the case $T = \infty$. Then the rhs of (2.2) is negative with positive probability which is contradictory. If $T < \infty$ then (2.2) implies $a = (T - b)/T < 0$, which means that $0 < T < b$. Hence $\exists\, \epsilon > 0$ such that $\forall\, x \in (0, \epsilon)$ it follows that $ax + b > T$, which is contradictory to $E(Y|X) \leq T$, μ_X a.s.

The case $1 < a < 1 + p_2/p_1$. Observe that in this case $T = \infty$ since if $T < \infty$ we have by (2.2) the contradition: $T = E(Y|X = T) = aT + b > T$, where the last inequality follows from the fact that $a > 1$.

Now let us apply the following transformation of variables to (2.3):

$$z = \frac{cy}{c+y}, \quad y \in [0, \infty), \quad y = \frac{cz}{c-z}, \quad z \in [0, c),$$

$$t = \frac{cx}{c+x}, \quad x \in [0, \infty), \quad x = \frac{ct}{c-t}, \quad t \in [0, c),$$

where $c = b/(a-1)$. Consequently (2.3) takes the form

$$\int_t^c \frac{(z-t)^{p_2}(c-z)^{p_1-1}}{z^{p_1+p_2-1}}\, dF_Y(y(z)) = (a-1)\int_t^c \frac{(z-t)^{p_2-1}(c-z)^{p_1}}{z^{p_1+p_2-1}}\, dF_Y(y(z)). \tag{2.4}$$

Multiplying both sides of (2.4) by $t^{\alpha-1}$, $\alpha \geq p_1$, and integrating with respect to t one gets

$$\int_0^c t^{\alpha-1}\left(\int_t^c (z-t)^{p_2}(c-z)^{p_1-1} z^{1-p_1-p_2}\, dF_Y(y(z))\right) dt$$
$$= (a-1)\int_0^c t^{\alpha-1}\left(\int_t^c (z-t)^{p_2-1}(c-z)^{p_1} z^{1-p_1-p_2}\, dF_Y(y(z))\right) dt.$$

Now upon changing the order of integration we get

$$\int_0^c \frac{(c-z)^{p_1-1}}{z^{p_1+p_2-1}}\left(\int_0^z \frac{t^{\alpha-1}(z-t)^{p_2}}{z^{\alpha+p_2}}\, dt\right) z^{\alpha+p_2}\, dF_Y(y(z))$$

$$= (a-1)\int_0^c \frac{(c-z)^{p_1}}{z^{p_1+p_2-1}}\left(\int_0^z \frac{t^{\alpha-1}(z-t)^{p_2-1}}{z^{\alpha+p_2-1}}\, dt\right) z^{\alpha+p_2-1}\, dF_Y(y(z)).$$

The inner integrals can be evaluated as beta functions:

$$B(\alpha, p_2+1)\int_0^c (c-z)^{p_1-1} z^{\alpha-p_1+1}\, dF_Y(y(z)) = (a-1)B(\alpha, p_2)\int_0^c (c-z)^{p_1} z^{\alpha-p_1}\, dF_Y(y(z)).$$

Inserting in the above identity $\alpha = k + p_1$ for any $k = 0, 1, \ldots$, it follows that

$$\frac{p_2}{k+p_1+p_2}\int_0^c (c-z)^{p_1-1}z^{k+1}\,dF_Y(y(z)) = (a-1)\int_0^c (c-z)^{p_1}z^k\,dF_Y(y(z))$$

$$= (a-1)c\int_0^c (c-z)^{p_1-1}z^k\,dF_Y(y(z)) - (a-1)\int_0^c (c-z)^{p_1-1}z^{k+1}\,dF_Y(y(z)).$$

Hence

$$E(V^{k+1}) = c\frac{p_1+p_2+k}{p_1+p_2+\frac{p_2}{a-1}+k}E(V^k), \quad k = 0, 1, 2, \ldots,$$

where V is a rv with a df G (concentrated on $(0, c)$) defined by

$$(c-z)^{p_1-1}\,dF_Y(y(z)) = K\,dG(z).$$

Then by the general recurrence relation for the moments of the first kind beta distribution (which is determined by the sequence of moments) it follows that $V \sim B_I(p_1+p_2, p_2/(a-1), c)$. Consequently

$$(c-z)^{p_1-1}\,dF_Y(y(z)) \propto z^{p_1+p_2-1}(c-z)^{\frac{p_2}{a-1}-1}\,dz,$$

and, returning to y's, since $dz = (c/(c+y))^2\,dy$, one gets for $y \in (0, c)$

$$dF_Y(y) \propto \frac{y^{p_1+p_2-1}}{(c+y)^{\frac{ap_2}{a-1}+1}}\,dy.$$

Observe that we have to have $\frac{ap_2}{a-1}+1-p_1-p_2+1 > 1$ or equivalently $a(1-p_1) > 1-p_1-p_2$. Now if $p_1 < 1$ then $a > 1 - \frac{p_2}{1-p_1}$, which is obvious. If $p_1 > 1$ then the above inequality becomes $a < 1 + \frac{p_2}{p_1-1}$ which holds since $a < 1 + \frac{p_2}{p_1}$. If $p_1 = 1$ then again the inequality is obvious.

The case $0 < a < 1$. Then $T = b/(1-a) < \infty$. To see this we first observe that T cannot be infinite for, otherwise, with positive probability, $aX + b < X$ contradicting the fact that $X < E(Y|X) = aX + b$ with probability 1. Since $T < \infty$ then $T = E(Y|X = T) = aT + b$ gives $T = b/(1-a)$.

Now (2.3) takes the form

$$\int_x^T (y-x)^{p_2}\,dH(y) = (1-a)(T-x)\int_x^T (y-x)^{p_2-1}\,dH(y).$$

Multiply now both sides by $x^{\alpha-1}$, $\alpha > 0$, and integrate with respect to x over $[0, T]$ to get

$$\int_0^T x^{\alpha-1}\left(\int_x^T (y-x)^{p_2}\,dH(y)\right)dx$$

$$= (1-a)T\int_0^T x^{\alpha-1}\left(\int_x^T (y-x)^{p_2-1}\,dH(y)\right)dx - (1-a)\int_0^T x^{\alpha}\left(\int_x^T (y-x)^{p_2-1}\,dH(y)\right)dx.$$

Change the order of integration then:

$$\int_0^T \frac{1}{y^{p_1+p_2-1}}\left(\int_0^y \frac{x^{\alpha-1}(y-x)^{p_2}}{y^{\alpha+p_2}}\,dx\right)y^{\alpha+p_2}\,dF_Y(y)$$

$$= (1-a)T\int_0^T \frac{1}{y^{p_1+p_2-1}}\left(\int_0^y \frac{x^{\alpha-1}(y-x)^{p_2-1}}{y^{\alpha+p_2-1}}\,dx\right) y^{\alpha+p_2-1}\,dF_Y(y)-$$

$$-(1-a)\int_0^T \frac{1}{y^{p_1+p_2-1}}\left(\int_0^y \frac{x^{\alpha}(y-x)^{p_2-1}}{y^{\alpha+p_2}}\,dx\right) y^{\alpha+p_2}\,dF_Y(y).$$

After evaluating the inner integrals we get

$$B(\alpha,p_2+1)\int_0^T y^{\alpha-p_1+1}\,dF_Y(y)$$

$$= (1-a)TB(\alpha,p_2)\int_0^T y^{\alpha-p_1}\,dF_Y(y) - (1-a)B(\alpha+1,p_2)\int_0^T y^{\alpha-p_1+1}\,dF_Y(y).$$

Hence for $\alpha = k+p_1$, $k=0,1,\ldots$, it follows that

$$E(Y^{k+1}) = T\frac{p_1+p_2+k}{p_1+\frac{p_2}{1-a}+k}E(Y^k), \quad k=0,1,\ldots$$

Now the recurrence relation for the moments of the first kind beta distribution implies that $Y \sim B_I(p_1+p_2, ap_2/(1-a), T)$.

The case $a=1$. Then again $T=\infty$ since $b>0$. Also (2.3) has the form

$$\int_x^\infty (y-x)^{p_2}\,dH(y) = b\int_x^\infty (y-x)^{p_2-1}\,dH(y), \quad x\geq 0. \tag{2.5}$$

Take now the Laplace transform of both sides of (2.5). Then for the lhs of (2.5) we have

$$\int_0^\infty e^{-sx}\left(\int_x^\infty (y-x)^{p_2}\,dH(y)\right)dx = \int_0^\infty e^{-sy}\left(\int_0^y (y-x)^{p_2}e^{s(y-x)}\,dx\right)dH(y)$$

$$= \int_0^\infty e^{-sy}\left(\int_0^y x^{p_2}e^{sx}\,dx\right)dH(y), \quad s>0.$$

Now for the inner integral we have

$$\int_0^y x^{p_2}e^{sx}\,dx = \frac{x^{p_2}}{s}e^{sx}\Big|_0^y - \frac{p_2}{s}\int_0^y x^{p_2-1}e^{sx}\,dx = \frac{y^{p_2}}{s}e^{sy} - \frac{p_2}{s}\int_0^y (y-x)^{p_2-1}e^{s(y-x)}\,dx.$$

Consequently

$$\int_0^\infty e^{-sx}\left(\int_x^\infty (y-x)^{p_2}\,dH(y)\right)dx = \frac{L}{s} - \frac{p_2}{s}\int_0^\infty e^{-sx}\left(\int_x^\infty (y-x)^{p_2-1}\,dH(y)\right)dx,$$

where $L = \int_0^\infty y^{p_2}\,dH(y)$ (observe that the above identity implies that L is finite). Hence (2.5) yields

$$\int_0^\infty e^{-sx}\left(\int_x^\infty (y-x)^{p_2-1}\,dH(y)\right)dx = \frac{L}{p_2+bs}, \quad s>0.$$

Since at the rhs we have the Laplace transform of the exponential function it follows that

$$\int_x^\infty (y-x)^{p_2-1}\,dH(y) = Ke^{-\gamma x}, \quad x\geq 0,$$

for $\gamma = p_2/b$ and some constant K. Hence upon multiplying both sides by $x^{\alpha-1}$ and integrating with respect to x we have

$$\int_0^\infty x^{\alpha-1}\left(\int_x^\infty (y-x)^{p_2-1}\,dH(y)\right)\,dx = K\int_0^\infty x^{\alpha-1}e^{-\gamma x}\,dx = K\frac{\Gamma(\alpha)}{\gamma^\alpha}.$$

On the other hand, the first expression in (2.6) can be viewed as

$$\int_0^\infty \frac{1}{y^{p_1+p_2-1}}\left(\int_0^y \frac{x^{\alpha-1}(y-x)^{p_2-1}}{y^{\alpha+p_2-1}}\,dx\right) y^{\alpha+p_2-1}\,dF_Y(y) = B(\alpha,p_2)E\left(Y^{\alpha-p_1}\right).$$

Put now $\alpha = k+p_1$, $k=0,1,2,\ldots$, to get from (2.6) and (2.7)

$$E(Y^k) = K\frac{\Gamma(k+p_1+p_2)}{\Gamma(p_2)\gamma^{k+p_1}}, \quad k=0,1,\ldots.$$

Now taking $k=0$ we find that

$$K = \frac{\Gamma(p_2)}{\Gamma(p_1+p_2)}\gamma^{p_1}.$$

Finally

$$E(Y^k) = \frac{\Gamma(k+p_1+p_2)}{\gamma^k\Gamma(p_1+p_2)}, \quad k=0,1,\ldots,$$

and consequently $Y \sim G(b/p_2, p_1+p_2)$.

The case $a=0$. Then by (2.2) it follows that $supp(X) \subset [0,b]$ since $P(Y \geq X) = 1$. On the other hand, by (2.1) also $supp(Y) \subset [0,b]$. Consequently $E(Y) = b$ (which follows from (2.2)) yields $P(Y=b)=1$. □

Remark. Put F_a for the distribution function of the respective Y from Theorem 1, for a slope coefficient of the regression function equal to a. Then the model we considered is continuous in distribution with respect to a at $a=1$, i.e.

$$\lim_{a\uparrow 1} F_a(x) = \lim_{a\downarrow 1} F_a(x) = F_1(x), \quad \forall\, x \in \mathbf{R}.$$

Observe first that the beta densities defined in points (i) and (iii), respectively, converge for a tending to 1, to the gamma density defined in (ii). To see this it suffices to use the Stirling formula for the Gamma function. Hence the convergence in distribution is a consequence of the Scheffé theorem.

3 Identifiability in absolutely continuous case

Now we study the question of identifiability of the model defined by the mixture (2.1) by any consistent posterior mean, i.e. we are concerned with a unique determination of μ_Y by $E(Y|X)$. Here we need an additional smoothness assumption, similarly as in Gupta and Wesołowski (1997), where the uniform mixture model was considered, however the method of the proof is completely different.

Theorem 2 *Let (X,Y) be a random vector with the mixture property (2.1). Assume additionally that the distribution of Y is absolutely continuous (with respect to the Lebesgue measure). Then it is uniquely determined by $E(Y|X)=m(X)$ and*

$$E\left(\frac{(Y-x)^{p_2-1}}{Y^{p_1+p_2-1}}I_{[x,\infty)}(Y)\right)=E(Y^{1-p_1})\exp\left(\int_0^x\frac{1-p_2-m'(t)}{m(t)-t}\,dt\right),\ \forall\, x\in supp(X).$$

Proof. The assumptions, via the Bayes theorem, imply that

$$\int_x^T\frac{(y-x)^{p_2}}{y^{p_1+p_2-1}}\,dF_Y(y)=(m(x)-x)\int_x^T\frac{(y-x)^{p_2-1}}{y^{p_1+p_2-1}}\,dF_Y(y) \tag{3.1}$$

for P_X a.a. x's, where $m(x)=E(Y|X=x)$ and $T=\inf\{y:\ F_Y(y)=1\}$. Observe that the lhs of (3.1) is a continuous (also differentiable since F_Y has a density) function in x, and since it is defined in $[0,T)$ everywhere except for a set of the Lebesgue measure zero, then it can be extended to a continuous (differentiable) function on $[0,T)$. Consequently the rhs of (3.1) is also a differentiable in $[0,T)$. Finally we conclude that m is defined on $[0,T)$, it is differentiable and $m(0)>0$. Then it follows that T can be computed by

$$T=\inf\{x\geq 0:\ m(x)=x\}.$$

Take now the Laplace transform of the lhs of (3.1). Similarly as in the proof of Th. 1, we get

$$\int_0^T e^{-sx}\left[\int_x^T(y-x)^{p_2}\,dH(y)\right]\,dx=\frac{L}{s}-\frac{p_2}{s}\int_0^T e^{-sx}M(x)\,dx,\quad s>0,$$

where H is defined in Section 2, $L=\int_0^T y^{p_2}\,dH(y)$ (observe that the above identity implies that $L=E(Y^{1-p_1})$ is finite) and

$$M(x)=\int_x^T(y-x)^{p_2-1}\,dH(y).$$

But, integrating by parts, one gets

$$\frac{1}{s}\int_0^T e^{-sx}M(x)\,dx=-\int_0^T e^{-sx}\left(\int_0^x M(y)\,dy\right)\,dx-\frac{Le^{-sT}}{sp_2}.$$

Consequently (3.1) implies (observe that $m(x)-x>0$ for any $x\in[0,T)$)

$$\frac{L}{s}\left(1-e^{-sT}\right)=\int_0^T e^{-sx}\left[p_2\int_0^x M(y)\,dy+(m(x)-x)M(x)\right]\,dx,$$

and by the uniqueness property of the Laplace transform it follows that

$$L=p_2\int_0^x M(y)\,dy+(m(x)-x)M(x),\quad x\in[0,T). \tag{3.2}$$

Observe that each term at the rhs of (3.2) is differentiable, for the function M it follows from the absolute continuity assumption. So after differentiation of (3.2) we obtain

$$0=p_2M(x)+(m'(x)-1)M(x)+(m(x)-x)M'(x),\quad x\in[0,T).$$

Solving the above differential equation we get

$$M(x) = K \exp\left(\int_0^x \frac{1-p_2-m'(t)}{m(t)-t}\,dt\right), \quad \forall\, x \in [0,T),$$

where $K = E(Y^{1-p_1}) = L/m(0)$. Consequently m uniquely determines M (and the eplicit formula is given).

Assume now that $T < \infty$. Then for any $\alpha \geq p_1$

$$N(\alpha) = \int_0^T x^{\alpha-1} M(x)\,dx = \int_0^T x^{\alpha-1}\left(\int_x^T \frac{(y-x)^{p_2-1}}{y^{p_1+p_2-1}}\,dF_Y(y)\right)dx$$

$$= \int_0^T \frac{1}{y^{p_1+p_2-1}}\left(\int_0^y \frac{x^{\alpha-1}(y-x)^{p_2-1}}{y^{\alpha+p_2-1}}\,dx\right) y^{\alpha+p_2-1}\,dF_Y(y) = B(\alpha,p_2)E(Y^{\alpha-p_1}).$$

Take now $\alpha = k + p_1$ for any $k = 0,1,\ldots$. Then the above identity yields

$$E(Y^k) = \frac{N(k+p_1)}{B(k+p_1,p_2)}, \quad k = 0,1,\ldots.$$

Since the function N is uniquely detrmined by M, and consequently by m, it follows that all the moments of Y are also uniquely determined. Finally the distribution of Y is also unique, since it is determined by a sequence of its moments.

Now let's consider the case $T = \infty$. Then transform the variables in the definition of M similarly as in the proof of Th. 1,

$$z = \frac{y}{1+y}, \quad y \in [0,\infty), \quad y = \frac{z}{1-z}, \quad z \in [0,1),$$

$$t = \frac{x}{1+x}, \quad x \in [0,\infty), \quad x = \frac{t}{1-t}, \quad t \in [0,1).$$

Consequently

$$M_1(t) = M(x(t)) = \int_t^1 \left(\frac{z-t}{(1-z)(1-t)}\right)^{p_2-1} dH(y(z)), \quad t \in [0,1).$$

Then we have

$$\int_t^1 (z-t)^{p_2-1}\,dG(z) = (1-t)^{p_2-1}M_1(t), \quad t \in [0,1), \tag{3.3}$$

where $dG(z) = (1-z)^{1-p_2}dH(y(z))$. Now similarly as for $T < \infty$, changing the starting point of the argument from the definition of M to the identity (3.3), we find that for any $\alpha > 0$

$$B(\alpha,p_2)\int_0^1 z^{\alpha-1}z^{p_2}\,dG(z) = \int_0^1 t^{\alpha-1}(1-t)^{p_2-1}M_1(t)\,dt.$$

Consequently

$$E(V^k) = K\frac{\int_0^1 t^k(1-t)^{p_2-1}M_1(t)\,dt}{B(k+1,p_2)}, \quad k = 0,1,\ldots,$$

where V is a rv with a df G_1 defined by $dG_1(z) = Kz^{p_2}dG(z)$, $z \in (0,1)$, and

$$K = \frac{B(1,p_2)}{\int_0^1 (1-t)^{p_2-1} M_1(t)\, dt}.$$

Hence G_1, and then G is uniquely determined by M_1, and then by M, and then by m. But G uniquely defines H and finally we conclude that F_Y is uniquely determined by m. □

Remark 1. Observe that if (2.1) holds for the random vector (X, Y), then introducing $Z = XY/(Y - X)$ one easily gets

$$\mu_{Z|Y} = B_{II}(p, q, Y), \tag{3.4}$$

i.e. a second kind beta mixture model. Such models were characterized in terms of $E(Y|Z)$, using other methods, in Wesołowski (1996b), i.e. the mixture (2.1) is identifiable by $E(Y|(XY)/(Y-X))$. On the other it follows immediately from our Theorem 1, that the second kind beta mixture (3.4) is uniquely determined, not only by $E(Y|Z)$, but also by $E(Y|(YZ)/(Y+Z))$.

Remark 2. A first kind beta mixture model with two beta conditional distributions of the form:

$$\mu_{X|Y} = B_I(p_1(Y), q_1(Y), 1-Y)\,, \quad \mu_{Y|X} = B_I(p_2(X), q_2(X), 1-X)$$

was considered in James (1975). Then it was shown that, while q's are necessarily constants, the p's are the logaritmic functions. Then this case is not covered by the model discussed in the present paper.

Example. Consider a bivariate (X, Y) model with

$$\mu_{X|Y} = B_I(2,1,Y), \quad E(Y|X) = \frac{2(1+X+X^2)}{3(1+X)}.$$

Then some quantities appearing in the proof of Theorem 2 can be computed explicitly

$$T = \inf\{x > 0:\ m(x) = x)\} = 1, \quad m(x) - x = \frac{2-x-x^2}{3(x+1)}, \quad m'(x) = \frac{2x(x+2)}{3(x+1)^2},$$

$$M(x) = K \exp\left(-\int_0^x \frac{2t}{1-t^2}\, dt\right) = K(1-x^2); \quad x \in (0,1).$$

Hence

$$N(\alpha) = \frac{2K}{\alpha(\alpha+2)}, \quad \alpha > -2$$

and finally $E(Y^k) = \frac{4}{4+k}$, $k = 1, 2, \ldots$, which implies that $Y \sim B_I(4,1,1)$.

Acknowledgement. Thanks to the referee's remarks the presentation of the material was improved.

References

[1] Arnold, B.C., Castillo, E., Sarabia, J.M. (1993) Conditionally specified models: structure and inference. In: *Multivariate Analysis: Future Directions, 2* (C.M. Coudras, C.R. Rao, eds.), Elsevier Science B.V., Amsterdam, 441-450.

[2] Barndorff-Nielsen, O. (1965) Identifiability of mixtures of exponential families. *Journal of Mathematical Analysis and Applications* **12**, 115-121.

[3] Cacoullos, T., Papageorgiou, H. (1983) Characterizations of discrete dsitributions by a conditional distribution and a regression function. *Annals of the Institute of Statistical Mathematics* **35**, 95-103.

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

[5] Gupta, A.K., Wesołowski, J. (1997) Uniform mixtures via posterior means. *Annals of the Institute of Statistical Mathematics* **49**, 171-180.

[6] Gupta, A.K., Wesołowski, J. (1999) Discrete uniform mixtures via posterior means. *Test* - to appear.

[7] James, I.R. (1975) Multivariate distributions which have beta conditional distributions. *Journal of the American Statistical Association* **70**, 681-684.

[8] Johnson, N.L. (1957) Uniqueness of a result in the theory of accident proneness. *Biometrika* **44**, 530-531.

[9] Johnson, N.L., Kotz, S. (1992) Non-existence of certain polynomial regression in random sum theory. *Sankhyā, Ser. A* **54**, 455-459.

[10] Korwar, R.M. (1975) On characterizing some discrete distributions by linear regression. *Communications in Statistics* **4**, 1133-1147.

[11] Korwar, R.M. (1977) On characterizing Lagrangian-Poisson and quasi-binomial distributions. *Communications in Statistics* **6**, 1409-1416.

[12] Krishnaji, N. (1974) Characterization of some discrete distributions based on a damage model. *Sankhyā, Ser. A* **36**, 204-213.

[13] Kyriakoussis, A., Papageorgiou, H. (1991) Characterizations of logarithmic series distribution. *Statistica Neerlandica* **45**, 1-8.

[14] Papageorgiou, H. (1984) Characterizations of continuous binomial and negative binomial mixtures. *Biometrical Journal* **26**, 795-798.

[15] Papageorgiou, H. (1985) On characterizing some discrete distributions by a conditional distribution and a regression function. *Biometrical Journal* **27**, 473-479.

[16] Papageorgiou, H., Wesołowski, J. (1997) Posterior mean identifies the prior distribution in NB and related models. *Statistics and Probability Letters* **36**, 127-134.

[17] Patil, G.P., Bildikar, S. (1966) Identifiability of countable mixtures of discrete probability distributions using methods of infinite matrices. *Proceedings of the Cambridge Philosophical Society* **62**, 485-494.

[18] Prakasa Rao, B.L.S. (1992) *Identifiability in Stochastic Models: Characterization of Probability Distributions.* Academic Press, Boston.

[19] Sapatinas, T. (1995) Identifiability of mixtures of power series distributions and related characterizations. *Annals of the Institute of Statistical Mathematics* **47**, 447-459.

[20] Teicher, H. (1961) Identifiability of mixtures. *Annals of Mathematical Statistics* **32**, 244-248.

[21] Wesołowski, J. (1995a) Bivariate distributions via a Pareto conditional distribution and a regression function. *Annals of the Institute of Statistical Mathematics* **47**, 177-183.

[22] Wesołowski, J. (1995b) Bivariate discrete measures via power series conditional distribution and a regression function. *Journal of Multivariate Analysis* **55**, 219-229.

[23] Wesołowski, J. (1996a) A new conditional specification of the bivariate Poisson conditionals distribution. *Statistica Neerlandica* **50**, 390-393.

[24] Wesołowski, J. (1996b) Bivariate distributions via the second kind beta conditional distribution and a regression function. *Journal of Mathematical Sciences* **78**, 125-129.

[25] Xekalaki, E. (1983) Infinite divisibility, completness and regression properties of the univariate generalized Waring distributions. *Annals of the Institute of Statistical Mathematics* **35**, 279-289.

Arjun K. Gupta
Department of Mathematics and Statistics
Bowling Green State University
Bowling Green, OH 43403-0221, U.S.A.
e-mail: gupta@andy.bgsu.edu

Jacek Wesołowski
Wydział Matematyki i Nauk Informacyjnych
Politechnika Warszawska
00-661 Warszawa, Poland
e-mail: wesolo@alpha.mini.pw.edu.pl