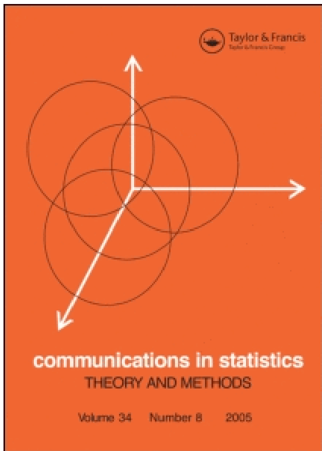

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597238

Gamma Mixture: Bimodality, Inflexions and L-Moments

S. E. Ahmed [a]; M. N. Goria [b]; A. Hussein [a]
[a] Department of Mathematics and Statistics, University of Windsor, Windsor, Ontario, Canada
[b] Department of Computer and Management Sciences, University of Trento, Trento, Italy

Online Publication Date: 01 January 2008

To cite this Article: Ahmed, S. E., Goria, M. N. and Hussein, A. (2008) 'Gamma Mixture: Bimodality, Inflexions and L-Moments', Communications in Statistics - Theory and Methods, 37:8, 1147 — 1161

To link to this article: DOI: 10.1080/03610920701713286
URL: http://dx.doi.org/10.1080/03610920701713286

Communications in Statistics—Theory and Methods, 37: 1147–1161, 2008

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610920701713286

Taylor & Francis
Taylor & Francis Group

Distributions and Applications

Gamma Mixture: Bimodality, Inflexions and L-Moments

S. E. AHMED[1], M. N. GORIA[2], AND A. HUSSEIN[1]

[1]Department of Mathematics and Statistics, University of Windsor, Windsor, Ontario, Canada
[2]Department of Computer and Management Sciences, University of Trento, Trento, Italy

We study some characteristics of the mixture of two gamma distributions. Specifically, we characterize regions of the parameter space where the mixture density is bimodal and/or has four inflexion points and further, we provide formulae for the L-moments of such mixtures. These characteristics may be useful in density and parameter estimation methods, as well as in analytical and graphical modality testing procedures. With a simple example, we illustrate the estimation of the parameters of the gamma mixture by the method of L-moments. Such estimators are in some circumstances more efficient than those based on the conventional method of moments.

Keywords Bimodal; Gamma mixture; L-moments; Moments; Probability weighted moments.

Mathematics Subject Classification Primary 62E10, 62F10; Secondary 62G30.

1. Introduction

The history of mixture dates back to Karl Pearson, more than a century ago. However, its use as an effective tool for modeling the real life data is quite recent (Crawford et al., 1992; McLachlan and Peel, 2000). Most of the attention in the literature has been focused on the analysis and applications of normal mixtures. Mixture of other distributions such as gamma and others, though relevant to modeling real data, has received relatively insignificant treatment in the literature.

Besides many other areas, the gamma mixtures have been applied in cure rate models (Peng et al., 2001), in pattern recognition problems (Webb, 2000), and in economics (Sternberg, 1994).

Received September 29, 2006; Accepted July 31, 2007
Address correspondence to A. Hussein, Department of Mathematics and Statistics, University of Windsor, Windsor, Ontario N9B 3P4, Canada; E-mail: ahussein@uwindsor.ca

Characteristics of a mixture density such as regions of bimodality, inflexion points, and moments provide necessary tools for many density and parameter estimation methods. These characteristics of the mixture are useful also in graphical and analytical modality testing procedures (Roeder, 1994). Therefore, it is important to study such characteristics of mixture distributions in detail.

In the case of two-component normal mixtures, Eisenberger (1969) and most extensively Robertson and Fryer (1969) studied the regions of bimodality and inflexion points of the parametric space associated with the normal mixture. Our first objective in this article is similar to that of Robertson and Fryer (1969), although the analysis is more complicated due to the presence of more parameters than in the normal case. In fact, we shall specify the regions where the density of the gamma mixture is bimodal and where it has four inflexion points for a set of four parameters. Note that if the density is bimodal then it may have up to four inflexion points. Our aim in this article is to characterize bimodal gamma mixtures which necessarily have four inflection points. On the other hand, the density may have four inflexion points and not be bimodal. However, in this latter case, apart from the x-axis, there will a tangent touching the density at two points, a characteristic known as bitangentiality (Robertson and Fryer, 1969). Thus, the difference between region of four inflexion points and that of bimodality is the region of bitangentiality.

The second objective of this article is to provide formulae for the L-moments of the gamma mixture.

Throughout the article

$$g(x; \theta_1, \alpha_1, \theta_2, \alpha_2) = pf(x; \theta_1, \alpha_1) + (1-p)f(x; \theta_2, \alpha_2) \tag{1}$$

is the density of two component gamma mixtures, where $0 < p < 1$ is the mixing proportion and for $i = 1, 2$,

$$f(x; \theta_i, \alpha_i) = \left[\theta_i^{\alpha_i}\Gamma(\alpha_i)\right]^{-1} x^{(\alpha_i - 1)} \exp -(x/\theta_i) \quad x \geq 0, \;\; \alpha_i \geq 0, \;\; \theta_i \geq 0$$

is the pdf of gamma random variable with scale θ_i and shape α_i.

The rest of the article is organized as follows: in Sec. 2 we identify regions of the parameter space corresponding to bimodality. In Sec. 3 we shall identify regions of the same space where the mixture density has four inflexions. In Sec. 4 we calculate the first four L-moments of the gamma mixture in (1).

2. Bimodality

In this section, we shall identify regions of parameter space where the density in (1) is bimodal. Since modality of g is unaffected by scale changes, without loss of generality, we shall consider the mixture density $g_1(x; 1, \alpha_1, \theta, \alpha_2)$, rescaled with respect to the larger scale parameter so that $\theta = \theta_2/\theta_1 \leq 1$, and therefore, we shall focus only on values of θ in [0, 1].

Now if g_1 is bimodal then between its two modes there is a minimum at which $g_1'(x) = 0$ and $g_1'' > 0$. In the light of (1), $g_1'(x) = 0$ reduces to

$$-[(1-p)/p]f(x; \theta, \alpha_2)/f(x; 1, \alpha_1) = [\alpha_1 - 1 - x]/[\alpha_2 - 1 - x/\theta],$$

which is meaningful provided that $\alpha_1 > 1$, $\alpha_2 > 1$, and either $v_1 < x < v_2\theta$ or $v_2\theta < x < v_1$, where $v_i = \alpha_i - 1$. Therefore, we need to treat the two cases $v_1 < \theta v_2$ and

$\theta v_2 < v_1$ separately in combination with values of θ. Also, after eliminating p, the system $g_1'(x) = 0$ and $g_1'' > 0$ becomes

$$h(x) = (x/\theta - v_2)\big[(x - v_1)^2 - v_1\big] - (x - v_1)\big[(x/\theta - v_2)^2 - v_2\big] < 0, \qquad (2)$$

the right side of which is a cubic equation in x.

2.1. Case I ($v_1 < \theta v_2$)

Case $\theta < 1$. Since $h(v_1) = v_1(v_2\theta - v_1)/\theta > 0$ and $h(v_2\theta) = v_2(v_2\theta - v_1) > 0$, for g_1 to be bimodal the cubic equation, $h(x)$, in (2) must have two and hence three distinct real roots.

On dividing $h(x)$ throughout by $(1 - \theta)$ and letting

$$y = x - \big[v_2\theta^2 + v_1\theta - 2v_2\theta - v_1][3(1 - \theta)\big]^{-1},$$

(2) is equivalent to

$$h_1(y) = y^3 + py + q > 0, \quad h_1(y) = -h(y).$$

The roots y_i, $i = 1, 2, 3$ of $h_1(x)$ are real and distinct if and only if the discriminant function is positive, i.e.,

$$D = -108\big(p^3/27 + q^2/4\big) > 0;$$

for details see Kurosh (1988).

By cumbersome but straightforward computations, we find that $D > 0$ is equivalent to

$$\begin{aligned} h_2(v_1, \theta, d) &= 27v_1^2(1 - \theta)^2\theta + 2v_1 d(2 - \theta)[9\theta(1 - \theta) + d(2\theta^2 + \theta - 1)] \\ &\quad - 4d(1 - \theta)\theta^2 - \theta d^2(\theta^2 + 8\theta - 8) + 2d^3(\theta^2 + 2\theta - 2) - \theta d^4 < 0, \qquad (3) \end{aligned}$$

where $d = (v_2\theta - v_1) > 0$. Note that Eq. (3) is quartic in d, cubic in θ, and quadratic in v_1. It is therefore convenient to work with v_1. As a function of v_1, h_2 is negative between its two real roots provided

$$16d(3\theta(1 - \theta) + d(1 - \theta + \theta^2)^3 > 0,$$

which holds true for all values of θ and d. The two roots of h_2 in (3) are

$$a = \frac{-(2-\theta)d[9(1-\theta)\theta + d(1+\theta)(2\theta^2 + \theta - 1)] - 2d^{1/2}(3\theta(1-\theta) + d(1-\theta+\theta^2))^{3/2}}{27(1-\theta)^2}$$

and

$$b = \frac{-(2-\theta)d[9(1-\theta)\theta + d(2\theta^2 + \theta - 1)] + 2d^{1/2}(9\theta(1-\theta) + d(1-\theta+\theta^2))^{3/2}}{27(1-\theta)^2\theta}.$$

Note that b is positive provided $d \neq \theta$, whereas the first root is always negative. The location of the minimum, m, of h_2 in (3) clearly depends on θ and d. In fact, for

$\theta < (d + .05)/10$ this minimum is positive and otherwise negative. The bound on θ increases with d till it reaches around .45. From this it follows that the admissible region of v_1 is $(0, b)$ if $\theta > (d + .05)/10$. Otherwise, the admissible region is (m, b), where

$$m = \frac{-(2-\theta)d[9(1-\theta)\theta + d(2\theta^2 + \theta - 1)]}{27\theta(1-\theta)^2}.$$

Furthermore, since

$$h'(v_1) = d(\theta - d) > 0, \quad h'(v_1 + d) = \theta d(d+1) > 0,$$

if v_1 lies in the above intervals, then h has two roots, say $x_1 < x_2$, in the admissible range, where it is negative. On the other hand, it is not hard to show that the relation between p and x is one to one in the above range. Consequently, using these roots together with $g_1' = 0$, we can find intervals (p_1, p_2), (henceforth called P-intervals) containing the values of the mixing proportion p for which the gamma mixture density is bimodal. Explicitly, p_i is such that

$$\frac{p_i}{1-p_i} = \frac{x_i^{(v_2-v_1)} e^{\left(\frac{(\theta-1)x_i}{\theta}\right)} \Gamma(v_1+1)[\theta v_2 - x_i]\theta^{-(v_2+2)}}{\Gamma(v_2+1)(x_i - v_1)}. \tag{4}$$

In summary, if v_1 lies in the above admissible region (i.e., either $(0, b)$ or (m, b) defined above), $d = \theta v_2 - v_1 > 0$ and $\theta < 1$, and moreover, $x_1 < x_2$ are the roots of h in (2) corresponding to these parameter regions, then the mixture g is bimodal if and only if p is in the P-interval (p_1, p_2) with endpoints obtained by using x_1, x_2 in (4). Outside these restrictions, however, g is unimodal and cannot have minimum.

Case $\theta = 1$. For $\theta = 1$, if $d = v_2 - v_1 > 1$, $\sqrt{v_2} > \sqrt{v_1} + 1$ and $x_1 < x_2$ are the roots of

$$h(x) = (v_2 - v_1)\left[x^2 + x(v_2 + v_1 - 1) + v_1 v_2\right] \tag{5}$$

corresponding to these parameter regions, then g is bimodal if and only if p is in the P-interval (p_1, p_2) obtained from (4) using the roots x_1, x_2.

To see this, note that for $\theta = 1$ the function $h(x)$ in (2) reduces to (5), a quadratic polynomial. This polynomial has two distinct real roots provided its discriminant function is positive, i.e., $(v_1 + v_2 - 1)^2 - 4v_1 v_2 > 0$. Letting $d = v_2 - v_1$, we find that $0 < v_1 < (d-1)^2/4$ or equivalently, $0 < v_1^{1/2} < |d-1|/2$.

Now if $d > 1$ then $v_2 > \left(1 + v_1^{1/2}\right)^2$ or equivalently, $v_2^{1/2} > v_1^{1/2} + 1$. Whereas if $d < 1$, then $v_2 < \left(1 - v_1^{1/2}\right)^2$, which is not possible. Similar arguments as in the previous case ($\theta < 1$) would again lead to the existence of the P-interval, (p_1, p_2).

2.2. *Case II* ($\theta v_2 < v_1$)

Case $\theta < 1$. Note that in this case v_2 can be larger than v_1. Here the inequality (2) modifies to

$$\begin{aligned} h(x) &= (1-\theta)x^3 + (v_2\theta^2 - 2v_2\theta + 2v_1\theta - v_1)x^2 \\ &\quad + \theta x\left(\theta v_2^2 - 2v_1 v_2\theta - v_2\theta + 2v_1 v_2 + v_1 - v_1^2\right) - v_1 v_2(v_2 - v_1)\theta^2 < 0, \end{aligned} \tag{6}$$

and again, $h(v_2\theta) = v_2(v_1 - v_2\theta) > 0$, $h(v_1) = v_1(v_1 - v_2\theta)/\theta > 0$. Consequently, for g_1 to be bimodal, the cubic equation in (6) must have two, hence three distinct real roots. Next, as in the previous section, we reduce the function h in (6) to the equivalent one

$$h_1(y) = y^3 + py + q < 0$$

by putting

$$y = x - (3(1-\theta))^{-1}(v_2\theta^2 - 2v_2\theta + 2v_1\theta - v_1).$$

Now h_1 has real roots provided its discriminant function is positive, which is the case if

$$\begin{aligned} h_2(v_1, \theta, d) = {} & 27\theta(1-\theta)^2 v_1^2 + 2v_1(2-\theta)d(-9\theta(1-\theta) + d(2\theta^2 + \theta - 1)) \\ & + 4\theta^2(1-\theta)d + d^2\theta(8 - 8\theta - \theta^2) + d^3(4 - 4\theta - 2\theta^2) - \theta d^4) < 0, \quad (7) \end{aligned}$$

where $d = v_1 - \theta v_2$. As in the previous case, h_2 in (7) is quadratic function in v_1 and has two real roots provided its discriminant is positive, i.e.,

$$16d(-3\theta(1-\theta) + d(1 - \theta + \theta^2)^3 > 0,$$

or equivalently, if

$$d > 3\theta(1-\theta)(1 - \theta + \theta^2) = d(\theta).$$

Thus, for these values of d, the two roots of the above quadratic polynomial, h_2, in v_1 are

$$a_1 = \frac{(1-2\theta)d[9\theta(1-\theta) + d(1+\theta)(2-\theta)] - 2d_1^{1/2}(3 - 3\theta + d - d\theta + d\theta^2)^{3/2}}{27(1-\theta)^2}$$

and

$$b_1 = \frac{(1-2\theta)d[9\theta(1-\theta) + d(1+\theta)(2-\theta)] - 2d^{1/2}(3 - 3\theta + d - d\theta + d\theta^2)^{3/2}}{27(1-\theta)^2}.$$

The root b_1 is positive whereas a_1 and the minimum, m_1, of h_2 in (7) can be negative. In fact, if we let $d = d(\theta) + d_1$ then as d_1 increases the range where these quantities are positive shrinks for decreasing θ. More precisely, $a_1 > 0$ for $0 < d_1 < 1$ provided $\theta < .77$ whereas for $1 \le d_1 < 10$, a_1 is positive if $\theta < .15$. The minimum m_1 is positive for d_1 in the above intervals with θ bounded above by .84 and .55, respectively, and thereafter it remains positive for all values of d_1. Thus, for large values of d_1 we can take (m_1, b_1) as the admissible range of v_1 provided $\theta < .55$; otherwise, the admissible range is $(0, b_1)$. Also note that for moderate values of d_1 we can still use the same interval provided a_1 is negative and $\theta > .55$. However, for small values of θ we shall use (a_1, b_1) provided a_1 is positive.

Furthermore, since

$$h'(v_1 - d) = d(1-d)/\theta < 0, \quad h'(v_1) = d(\theta + d)/\theta^2 > 0,$$

the cubic function, h, has two roots in the admissible range of v_1 and between these two roots h is negative. Assuming these roots are $x_1 < x_2$, the corresponding P-interval $= (p_1, p_2)$ containing the mixing proportions for which the mixture is bimodal are obtained from (4) by using these roots.

In summary, if $\theta < 1$ and $d = v_1 - \theta v_2 > d(\theta) = 3\theta(1-\theta)(1-\theta+\theta^2) > 0$, and v_1 lies in the above admissible ranges, and moreover, $x_1 < x_2$ are the roots of h in (2) corresponding to these parameter regions, then g is bimodal if and only if p lies inside a P-interval whose endpoints, p_1, p_2, are obtained from (4) using these two roots. Outside these constraints, g cannot have a minimum, which implies g is unimodal.

The special case $v_1 = v_2 = v$ simplifies bimodality regions quite a lot. In fact, in this case the cubic function $h(x)$ in (2) has one zero root and its other roots are obtained from

$$x^2 - v(1+\theta)x + v(1+v)\theta = 0. \tag{8}$$

The quadratic equation in (8) has real roots provided $v > 4\theta(\theta-1)^{-2}$, which in turn identifies the values of v for which the mixture is bimodal.

Again, in this case, both roots are within an admissible range, and hence, using these roots in (4), we can determine the end-points of the P-interval containing the mixing proportion p for which the mixture is bimodal.

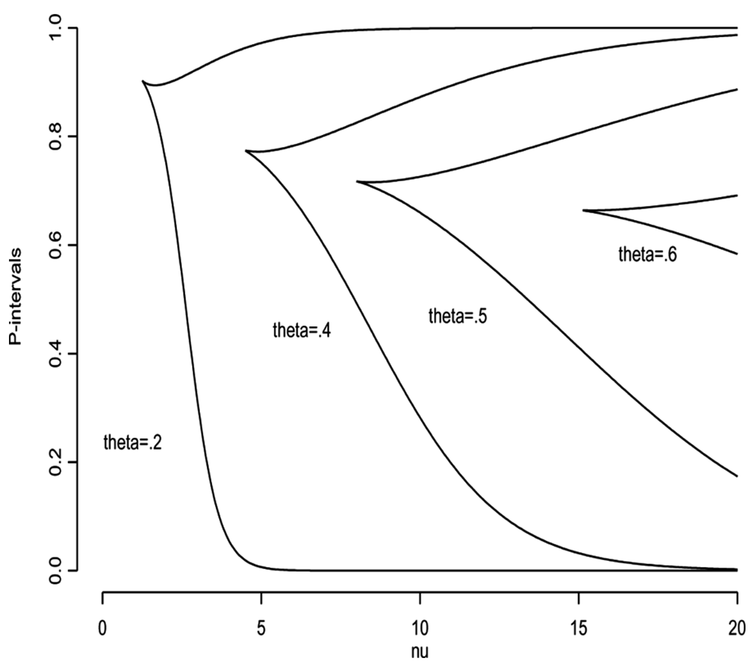

Figure 1. P-intervals (intervals of mixing proportion, p) corresponding to bimodal mixture, depicted as functions of $\alpha_1 - 1 = \alpha_2 - 1 = v$ for several values of $\theta = \theta_2/\theta_1$. In each curve, the upper and lower arms represent the upper and lower end-points, respectively, of the intervals containing mixing proportions for which (1) is bimodal.

This special case is depicted in Figure 1 showing the P-interval as a function of v for several values of $\theta < 1$. For example, for $\theta = .2$, the admissible range is $v > 1.25$ which marks the beginning of the first curve in the figure.

Case $\theta = 1$. Here the function $h(x)$ in (2) becomes quadratic and the inequality (2) reduces to

$$h(x) = [x^2 + x(1 - v_2 - v_1) + v_1 v_2] < 0. \tag{9}$$

The left-hand-side of (9) has two real roots provided that $(v_1 + v_2 - 1)^2 - 4v_1 v_2 > 0$, or equivalently, $2v_1^{1/2} < (1 + d), d = v_1 - v_2$, that is, provided that either $v_1^{1/2} > v_2^{1/2} + 1$, or $v_2^{1/2} + v_2^{1/2} < 1$ for $v_1 < 1$. Both roots, say $x_1 < x_2$, of the quadratic polynomial in (2) or, equivalently, in (9) correspond to admissible parameter values. Therefore, for $\theta = 1$ and v_1 and v_2 satisfying either $2v_1^{1/2} < v_1 - v_2 + 1$ or $v_2^{1/2} + v_2^{1/2} < 1$, g is bimodal if and only if p is in the interval (p_1, p_2), where the end-points are obtained from (4) using the roots x_1, x_2.

3. Inflexion Points

Our objective in this section is to identify the region of the parameter space where the mixture of gamma densities in (1) or equivalently $g_1(x; 1, \alpha_1, \theta, \alpha_2)$ has four inflexion points. Note that, whether the mixture density g_1 has two or four inflexion points is determined by its four parameters $p, \alpha_1, \alpha_2, \theta$, as compared to normal mixture case which has only three parameters, thus complicating the analysis further.

Now as we vary parameter values, g_1 goes from having two to having four inflexion points and as g_1 passes from one state to the other, it must pass through a state where not only the second but also the third derivative is zero, i.e., $g_1'' = g_1''' = 0$. Eliminating p in this system, we end up with a fifth degree polynomial equation

$$\begin{aligned} h(x, v_1, v_2, \theta,) = &-[(x/\theta - v_2)^2 - v_2] \quad [(x - v_1)^3 - 3v_1(x - v_1) - 2v_1)] \\ &+ [(x - v_1)^2 - v_1] \quad [(x/\theta)^3 - 3v_2(x/\theta - v_2) - 2v_2] = 0. \end{aligned} \tag{10}$$

Admissible roots of (10) are those leading to values of p satisfying $0 < p < 1$. Consequently, from $g'' = 0$ and $0 < p < 1$ we find that an admissible real root must satisfy the condition

$$C : [(x - v_1)^2 - v_1]/[(x/\theta - v_2)^2 - v_2] < 0.$$

Real roots satisfying this condition shall be called C-roots. The condition C implies that the admissible region for the C-roots is given by two disjoint intervals depending on the values of θ, v_1, v_2. Here we list all possible cases of admissible regions, I_1 and I_2:

a) if $(\theta v_2)^{1/2} > v_1^{1/2} + 1$, then

$$I_1 = \left(v_1 - v_1^{1/2}, v_1 + v_1^{1/2}\right), \quad I_2 = \left(\theta\left(v_2 - v_2^{1/2}\right), \theta\left(v_2 + v_2^{1/2}\right)\right);$$

b) if $v_1^{1/2} < (\theta v_2)^{1/2} < v_1^{1/2} + 1$, then

$$I_1 = \left(v_1 - v_1^{1/2}, \theta\left(v_2 - v_2^{1/2}\right)\right), \quad I_2 = \left(v_1 + v_1^{1/2}\right), \theta\left(v_2 + v_2^{1/2}\right);$$

c) if $\theta v_2 < v_1$, then

$$I_1 = \left(\theta\left(v_2 - v_2^{1/2}\right), v_1 - v_1^{1/2}\right), \quad I_2 = \left(\theta\left(v_2 + v_2^{1/2}\right), v_1 + v_1^{1/2}\right);$$

d) if $(\theta v_2)^{1/2} < (v_1)^{1/2} + 1$, then

$$I_1 = \left(\theta\left(v_2 - v_2^{1/2}\right), v_1 - v_1^{1/2}\right), \quad I_2 = \left(\theta\left(v_2 + v_2^{1/2}\right), v_1 + v_1^{1/2}\right);$$

e) if $(\theta v_2)^{1/2} < v_1^{1/2} - 1$, then

$$I_2 = \left(\theta\left(v_2 - v_2^{1/2}\right), \theta\left(v_2 + v_2^{1/2}\right)\right), \quad I_2 = \left(v_1 - v_1^{1/2}\right), \left(v_1 + v_1^{1/2}\right).$$

Next we show that the number of roots in the admissible region is even. Since the proof in the other cases is very similar, we illustrate only case (a). It is straightforward to see that

$$\begin{aligned} h\left(v_1 - v_1^{1/2}\right) &= 2v_1\left(v_1^{1/2} - 1\right)\left[v_1 - v_1^{1/2} - \theta\left(v_2 - v_2^{1/2}\right)\right]\left[v_1 - v_1^{1/2} - \theta\left(v_2 + v_2^{1/2}\right)\right] > 0, \\ h\left(v_1 + v_1^{1/2}\right) &= -2v_1\left(v_1^{1/2} - 1\right)\left[v_1 + v_1^{1/2} - \theta\left(v_2 - v_2^{1/2}\right)\right]\left[v_1 + v_1^{1/2} - \theta\left(v_2 + v_2^{1/2}\right)\right] > 0, \\ h\left(\theta\left(v_2 - v_2^{1/2}\right)\right) &= -2v_2\left(v_2^{1/2} - 1\right)\left[\theta\left(v_2 - v_2^{1/2}\right) - \left(v_1 + v_1^{1/2}\right)\right] \\ &\quad \times \left[\theta\left(v_2 - v_2^{1/2}\right) - \left(v_1 + v_1^{1/2}\right)\right] < 0 \end{aligned}$$

and

$$h\left(\theta\left(v_2 - v_2^{1/2}\right)\right) = 2v_2\left(v_2^{1/2} + 1\right)\left[\theta\left(v_2 + v_2^{1/2}\right) - \left(v_1 + v_1^{1/2}\right)\right]\left[\theta\left(v_2 + v_2^{1/2}\right) - \left(v_1 - v_1^{1/2}\right)\right] > 0.$$

Consequently, h has odd number of zeros in I_1 and also in I_2, i.e., an overall even number of roots.

3.1. *Transition From Two to Four Inflexion Points*

The case $v_1 = v_2 = v$ or equivalently, $d = v_2 - v_1 = 0$, is simple and helps find a cut-off-point, θ_0, in the θ-space above and below which there are different patterns of C-roots. Here, we argue that since g_1 is nondegenerate at $p = 0$ and $p = 1$ and has two inflexion points, by continuity it will continue to have two inflexion points. That is, there is no turning point where the number of inflexions change from 2 to 4. Under the condition $v_2 - v_1 = d = 0$, the polynomial in (10) has one root at zero and the other roots are obtained from

$$\begin{aligned} h(x, v_1, v_2, \theta,) = x^4 - 2v(1+\theta)x^3 + v(v\theta^2 - \theta^2 + 4v\theta + 2\theta + v - 1)x^2 \\ - 2v\theta(v^2 - 1)(1+\theta)x + (v\theta)^2(v^2 - 1) = 0 \end{aligned} \tag{11}$$

The desired cut-off point is the minimum θ, $\theta_0 = \theta_0(v)$, such that the above polynomial has no real roots in the admissible region. This can be determined by studying the discriminant function of the above fourth degree polynomial discussed in Kurosh (1988). However, it is more handy, though laborious, to do it numerically a software package. Thus for fixed value of v we search minimum value of θ such that (11) has only zero root. It turns out that the numerical values of θ_0 found can be very accurately approximated by the regression $\theta(v) = a + b\log(v)$, where intercept and angular coefficients are given by $(0.267949, 0.146597)$ for $1 \le v \le 10$, $(0.322616, 0.1222855)$ for $10 \le v \le 30$, $(0.444892, 0.0869041)$ for $30 \le v \le 100$ and $(0.60116, 0.0529709)$ for $100 \le v \le 300$. The difference between exact and interpolated values lies in the interval $(-0.0055, 0.0055)$.

Figure 2 shows the range of θ for which sensible P-intervals (interval of values of p corresponding to density with four inflexions) exist for various values of $v_1 = v_2 = v$. It is clear that $\theta_0(v)$ are the foremost peeks of the curves. For example, for the curve labeled by $v = 1.5$, $\theta_0(1.5) = 0.327389$ is the cut-off point represented by the far most right point on the curve. For instance, note that P-interval for bimodality in Figure 1 for $\theta = .5$ and $v = 15$ say, is contained in the P-interval of four inflexions for the same values of parameters in Figure 2.

3.2. *Region of Four Inflexion Points*

Depending on the cut-off point, $\theta_0 = a + b\log(v_1)$, established in the previous subsection, it turns out that we have two distinct patterns of C-roots according to $\theta > \theta_0$ or $\theta < \theta_0$. For each of these cases, we shall describe the C-roots in detail for $d = v_2 - v_1 > 0$ and briefly mention the necessary modifications for $d < 0$.

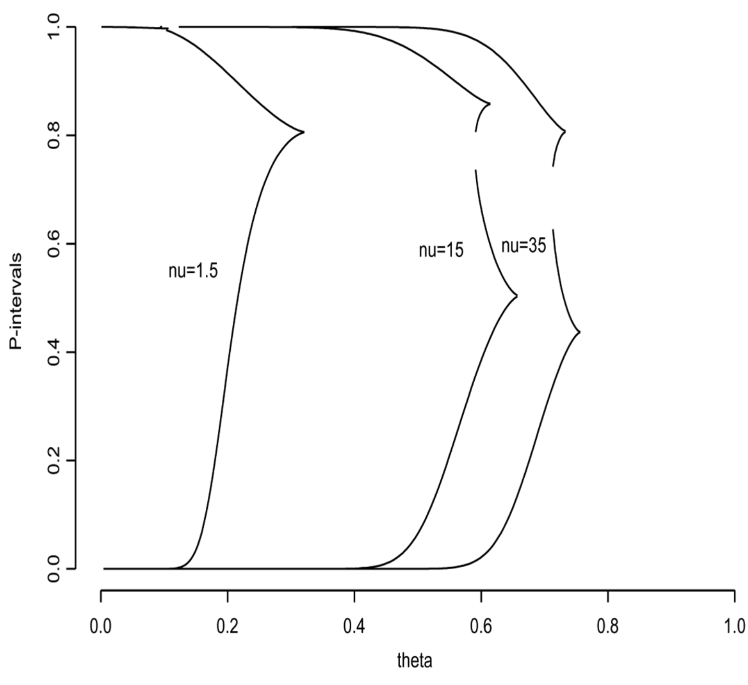

Figure 2. Intervals of p, P-intervals, corresponding to gamma mixtures density with four inflexions, as functions of θ and $v_1 = v_2 = v$. The upper and the lower arms of each curve represent the upper and the lower end-points, respectively, of the P-intervals.

The procedure of finding the regions of four inflexion points is as follows. For fixed v_1, we identify whether $\theta > \theta_0$ or $\theta < \theta_0$. Secondly, for the given θ, v_1, v_2 we find C-roots, x_i, by solving numerically the 5th degree polynomial equation in (10) with respect to x. Finally, we find the mixing proportions p_i corresponding to these C-roots from the condition $g_1'' = 0$ and construct the P-interval $= (p_1, p_2)$ containing values of p for which the mixture density has four inflexion points.

Case $\theta > \theta_0$. For fixed value of v_1, we calculate θ from $\theta = .01 + a + b\log(v_1) > \theta_0$ and then by increasing $d = v_2 - v_1 > 0$ from low to quite high values we observe the pattern of C-roots. As a result we find, numerically, constants $d_1 < d_2 < d_3$ such that

1. for $d < d_1$ the density is unimodal (i.e., there are no valid P-intervals);
2. for $d_1 \leq d < d_2$ there are two C-roots in the first interval, leading to one P interval;
3. for $d_2 \leq d < d_3$ there are two C-roots in each interval, giving rise to two P intervals;
4. for $d \geq d_3$ there is one root in each interval leading to one P interval.

For example, for $v_1 = 1.5$ we have $d_1 = 4.092$, $d_2 = 9.797$, $d_3 = 9.964$, for $v_1 = 10$, we have $d_1 = 12.82$, $d_2 = 15.9739$, $d_3 = 16.54$ and for $v_1 = 9$ we have $d_1 = 10.748$, $d_2 = 13.978$, $d_3 = 14.488$.

Figure 3 shows the P-intervals of four inflexions as functions of $d > 0$ for $v_1 = 1.5, 5, 9$. These curves show clearly the above-mentioned values, d_i, which form

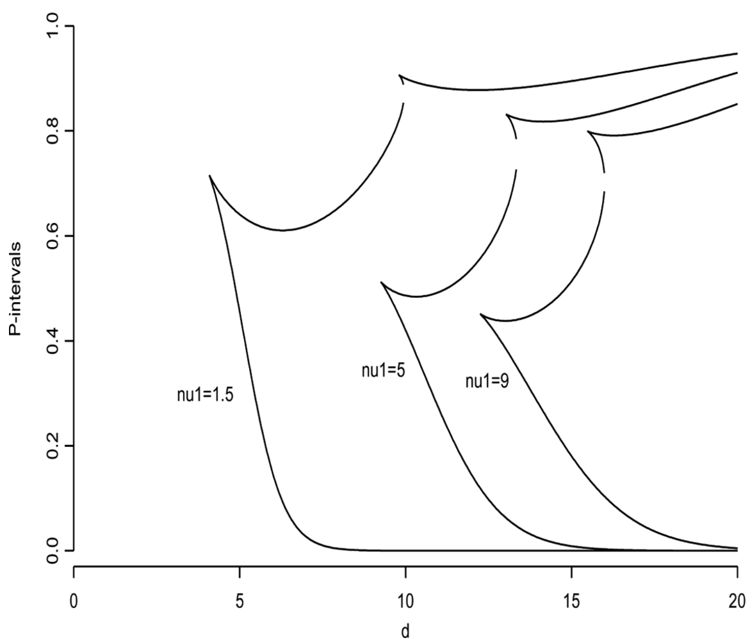

Figure 3. Intervals of p, P-intervals, corresponding to gamma mixture density with four inflexions depicted as functions of $d = v_2 - v_1$ for fixed values of v_1. The upper and lower arms of each curve represent the upper and lower end-points, respectively, of the P-intervals.

the cut-off points for the existence of none, one, two, and again one P-intervals, respectively.

For $v_1 > v_2$, the above (1)–(4) will still hold with the only change that the two C-roots in (2) are in the second interval, instead of the first interval.

Case $\theta < \theta_0$. Proceeding as before with

$$\theta = -.01 + a + b \log v_1$$

and provided that

$$(\theta v_2)^{1/2} < v_1^{1/2} - 1,$$

we have constants d_i, $i = 1, 2 \ldots, 6$, in an increasing order such that

1. for $d \leq d_1$ there is one root in each interval leading to one P-interval;
2. for $d_1 < d \leq d_2$ there are two C-roots in each interval leading to two P-interval;
3. for $d_2 < d \leq d_3$, there are two roots in the second interval leading to one P-interval;
4. for $d_3 < d \leq d_4$, the density is unimodal;
5. for $d_4 < d \leq d_5$, there are two roots in the smaller interval, hence one P-interval;
6. for $d_5 < d < d_6$, there are two roots in each interval, hence two P-interval;
7. for $d \geq d_6$ there is one root in each interval, hence one P-interval.

In case

$$(\theta v_2)^{1/2} > v_1^{1/2} - 1,$$

there are only four cases of admissible regions and, consequently, only 3) to 7) hold.

For example with $v_1 = 3$, we find

$$d_3 = 3.208, \quad d_4 = 10.1, \quad d_5 = 18.4, \quad d_6 = 18.618, \quad d_7 = 18.68,$$

whereas if

$$(\theta v_2)^{1/2} < v_1^{1/2} - 1;$$

then, for example, with $v_1 = 10$ we find

$$d_1 = 1.616, \quad d_2 = 2.05, \quad d_3 = 4.67, \quad d_4 = 18.64, \quad d_5 = 24.36, \quad d_6 = 24.807.$$

For $v_1 > v_2$ only e) or d) and e) of the disjoint intervals containing C-roots described above will be valid. In the latter case, i.e, $v_2 \geq 5$, we have only one C-root in each interval leading to one P-interval. In the former case, we have two situations; if the upper limit of the lower interval (say I_1) is not far away from the lower limit of the upper interval such as $3 \leq v_2 < 5$, then we first have two C-roots in each interval followed by one C-root in each interval, leading, respectively, to two P-intervals followed by one P-interval.

Finally, for $1 < v_2 < 3$, when the lower limit of the upper interval is well below the upper limit of the lower interval, we have two C-roots in one interval, followed by the above pattern.

4. L-Moments of the Gamma Mixture

The L-moments of a random variable, X, are linear combinations of probability weighted moments (PWMs), which in turn are expectations of certain functions (e.g., $\beta_i = \int xF^i f dx$) of the random variable. The PWMs, first introduced by Greenwood et al. (1979), and hence the L-moments based on them, can be defined for any random variable whose mean exists. The theory of the L-moments was unified and popularized by Hosking (1990). The L-moments are known to be robust to outliers and sample variability. Therefore, L-moments are superior to the method of conventional moments and sometimes even to the likelihood-based methods in describing properties of a distribution such as skewness and kurtosis and in estimating and testing hypotheses about distribution parameters.

In this section, we shall provide formulae for the L-moments of the mixture of two gamma distributions as in (1). For simplicity we limit our discussion to the first four of these L-moments and therefore, we consider the rescaled mixture with $\theta_2/\theta_1 = \theta$,

$$g_1(x; 1, \alpha_1, \theta, \alpha_2) = pf(x; \alpha_1) + (q/\theta)f(x/\theta; \alpha_2) \tag{12}$$

with cdf $G_1(x; 1, \alpha_1, \theta, \alpha_2)$, where $f(x; \alpha_i)$ is the density of standard gamma variable with cdf $F(x; \alpha_i)$ and shape α_i and $p = 1 - q$. The L-moments of such gamma mixtures are linearly related to their PWMs,

$$\beta_i = \int xG_1^i(x; 1, \alpha_1, \theta, \alpha_2)g_1(x; 1, \alpha_1, \theta, \alpha_2)dx;$$

therefore, we first compute these PWMs and relate them to the component PWMs $\beta_{ij} = \int xF^i(x; \alpha_j)f(x; \alpha_j)dx$, which are available from Hosking (1990). Obviously,

$$\begin{aligned}
\beta_0 &= p\alpha_1 + q\theta\alpha_2, \\
\beta_1 &= p^2\beta_{11} + q^2\theta\beta_{12} + pq[\alpha_1 I_{1/(1+\theta)}(\alpha_2, \alpha_1 + 1) + \theta\alpha_2 I_{\theta/(1+\theta)}(\alpha_1, \alpha_2 + 1)],
\end{aligned}$$

where $I_x(a, b)$ is the normalized beta function Gradshteyn and Ryzhik (1980). For β_2, β_3 we need the following notations:

$$\begin{aligned}
A_{ij}(\tau) &= E\big[(1 - X_{ij})I_{\frac{(1-X_{ij})\tau}{1+\tau(1-X_{ij})}}(\alpha_i, \alpha_i + \alpha_j + 1)\chi_{[0,\tau/(1+\tau)]}(X_{ij})\big], \\
B_{ij}(\tau) &= E\big[(1 - X_{jj})I_{\frac{(1-X_{jj})\tau}{1+\tau(1-X_{jj})}}(\alpha_i, 2\alpha_j + 1)\chi_{[0,1/2]}(X_{jj})\big], \\
C_{ij}(\tau) &= E\big[(1 - X_{ij})(1 - Y_{ij})I_{\frac{(1-X_{ij})(1-Y_{ij}\tau)}{1+\tau(1-X_{ij})(1-Y_{ij})}}(\alpha_i, 2\alpha_i + \alpha_j + 1)\chi_{[0,\tau/(1+\tau)]*[0,1]}(X_{ij}, Y_{ij})\big], \\
D_{ij}(\tau) &= E\big[(1 - X_{jj})(1 - Y_{jj})I_{\frac{(1-X_{jj})(1-Y_{jj}\tau)}{1+\tau(1-X_{jj})(1-Y_{jj})}}(\alpha_i, 3\alpha_j + 1)\chi_{[0,1/2]*[0,(1-x)/(2-x)]}(X_{jj}, Y_{jj})\big], \\
E_{ij}(\tau) &= E\big[(1 - X_{jj})(1 - Y_{ij}^0)I_{\frac{(1-X_{jj})(1-Y_{ij}^0\tau)}{1+\tau(1-X_{jj})(1-Y_{ij}^0)}}(\alpha_i, 2\alpha_i + 2\alpha_j + 1)\chi_{[0,1/2]*[0,1]}(X_{jj}, Y_{ij}^0)\big]
\end{aligned}$$

where, $i \neq j$; $i, j = 1, 2$, χ is an indicator function and X_{ij}, Y_{ij}, Y_{ij}^0 are independently beta random variables with densities, $Beta(x; \alpha_i, \alpha_j)$, $Beta(y; \alpha_i, \alpha_i + \alpha_j)$ and

$Beta(y^0; \alpha_i, 2\alpha_j)$, respectively. Now, after some algebra, we find that

$$\beta_2 = p^3\beta_{21} + q^3\theta\beta_{22} + pq(\alpha_1 + \alpha_2)[p\theta A_{12}(\theta) + qA_{21}(1/\theta)], \tag{13}$$

$$\begin{aligned}\beta_3 = {} & p^4\beta_{31} + q^4\theta\beta_{32} + pq[p^2\theta(2\alpha_1 + \alpha_2)C_{12}(\theta) + q^2(2\alpha_2 + \alpha_1)C_{21}(1/\theta)] \\ & + 9pq[q^2\theta\alpha_2 D_{12}(\theta) + p^2\alpha_1 D_{21}(1/\theta)] \\ & + 3p^2q^2[\theta(\alpha_1 + 2\alpha_2)E_{12}(\theta) + (\alpha_2 + 2\alpha_1)E_{21}(1/\theta)].\end{aligned} \tag{14}$$

Finally, by definition (Hosking, 1990), the $(r+1)$th L-moment is written as

$$\lambda_{r+1} = \sum_{i=0}^{r} p_{r,i}\beta_i$$

where $r = 0, 1, \ldots$ and $p_{r,i} = (-1)^{r-i}\binom{r}{i}\binom{r+i}{i}$. Hence, the first four L-moments are

$$\begin{aligned}\lambda_1 &= \beta_0, \\ \lambda_2 &= 2\beta_1 - \beta_0, \\ \lambda_3 &= 6\beta_2 - 6\beta_1 + \beta_0, \\ \lambda_4 &= 20\beta_3 - 30\beta_2 + 12\beta_1 - \beta_0.\end{aligned} \tag{15}$$

In general, by equating these theoretical L-moments to their sample counterparts, estimation of the parameters ($\alpha_1, \alpha_2, \theta$, and p) can be carried out as one would do in the case of the classical method of moments (Hosking, 1990).

The sample L-moments are obtained by replacing the β_i in the above expressions by their sample counterparts

$$b_i = \frac{1}{n}\sum_{k=i+1}^{n}\binom{k-1}{i}\binom{n-1}{i}^{-1}X_{k,n}$$

where $i = 0, 1, 2, \ldots$ and $X_{k,n}$ is the kth order statistic of the data.

The fifth L-moment is quite complicated and thus omitted here. However, one can always use the expression

$$\begin{aligned}5\beta_4 = {} & p\sum_{i=0}^{4}\binom{4}{i}q^i p^{4-i}\int_0^\infty xF^{4-i}(x; \theta_1, \alpha_1)F^i(x; \theta_1\theta, \alpha_2)f(x; \theta_1, \alpha_1)dx \\ & + q\sum_{i=0}^{4}\binom{4}{i}q^i p^{4-i}\int_0^\infty xF^{4-i}(x; \theta_1, \alpha_1)F^i(x; \theta_1\theta, \alpha_2)f(x; \theta_1\theta, \alpha_2)dx\end{aligned}$$

where for the general gamma pdf and cdf, the relation $\theta_2/\theta_1 = \theta$ has been used. As in the expression of β_3, one can further expand the integrals in terms of beta densities.

It is also possible to obtain the L-moment estimators by direct numerical integration using the definition of the PWM, β_i, given above.

In general, there are no closed-form expressions for the solution of the system of equations resulting from the L-moments method. However, numerical optimization methods, implemented in Maple, Mathematica, and similar software, provide numerical solutions to the problem.

For the sake of completeness, we mention that the conventional noncentral moments of the gamma mixture in (1) is quite simple and straightforward, since they are just mixtures of the corresponding moments of the component distributions. These are considered in John (1970) for the simple case where the two component distributions have a common and known shape parameter.

Using the mgf of the gamma distribution, the rth noncentral moments of (12) are

$$\begin{aligned} m_r = E[x^r] &= p\int_0^\infty x^r f(x;\alpha_1)dx + q/\theta\int_0^\infty x^r f(x/\theta;\alpha_2)dx \\ &= p[\alpha_1(\alpha_1+1)\ldots(\alpha_1+r-1)] + q\theta^r[\alpha_2(\alpha_2+1)\ldots(\alpha_2+r-1)] \end{aligned}$$

which is more general than in John (1970). The conventional central moments of the mixture can be derived from the m_r by using conversion formulae (Papoulis, 1984).

Example. For illustration purposes, we take an example of a mixture of a completely known gamma distribution and a gamma distribution with unknown and eventually different scale parameter. Such situation is encountered for instance in mixed effects linear models (see Lehmann, 1999 for a similar example of normal mixture). Thus, without loss of generality, consider (12) with $\alpha_1 = \alpha_2 = 1$ and θ unknown. In this simple case, the system of equations resulting from equating β_0 and β_1 to their sample counterparts, b_0 and b_1, respectively, has the following explicit solution:

$$\hat{\theta} = \frac{b_0^2 - 4(b_0 - b_1)}{4(b_0 - b_1) - 1}; \quad \hat{p} = \frac{(b_0+1)(4b_1 - 3b_0)}{b_0^2 - 8b_0 + 8b_1 + 1}. \tag{16}$$

For this example, the results of a small simulation comparing these estimators to those generated by the conventional method of moments have shown that L-moment estimators are more efficient than their conventional moment counterparts for $p < .7$ and for any value of θ. Both types of estimators become inefficient for higher values of p, i.e., when information about θ is carried by a small portion of the whole sample. However, in this latter case, the conventional method of moments estimators are less affected in terms of efficiency as compared to the L-moment estimators. In any case, further detailed investigations of efficiency are needed.

References

Crawford, S. L., DeGroot, M. H., Kadane, J. B., Small, M. J. (1992). Modeling lake-chemistry distributions: Approximate Bayesian methods for estimating a finite-mixture model. *Technometrics* 34:441–453.

Eisenberger, I. (1969). Genesis of bimodal distribution. *Technometrics* 6:357–363.

Gradshteyn, I. S., Ryzhik, M. (1980). *Tables of Integrals, Series and Products*. New York: Academic Press.

Greenwood, J. A., Landwehr, J. M., Matalas, N. C., Wallis, J. R. (1979). Probability weighted moments: Definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research* 15:1049–1054.

Hosking, J. R. M. (1990). L-moments: Analysis and estimation of distributions using linear combinations of order statistics. *Journal of the Royal Statistical Society, Series B, Methodological* 52:105–124.

John, S. (1970). On identifying the population of origin of each observation in a mixture of observations from two gamma populations. *Technometrics* 12:565–568.

Kurosh, A. G. (1988). *Higher Algebra*. Moscow: MiR Publishers.

Lehmann, E. L. (1999). *Elements of Large-Sample Theory*. New York: Springer-Verlag.

McLachlan, G., Peel, D. (2000). *Finite Mixture Models*. Wiley.

Papoulis, A. (1984). *Probability, Random Variables, and Stochastic Processes*. 2nd ed., McGraw-Hill.

Peng, Y. W., Dear, K. B. G., Carriere, K. C. (2001). Testing for the presence of cured patients: a simulation study. *Statistics in Medicine* 20:1783–1796.

Robertson, C. A., Fryer, J. G. (1969). Some descriptive properties of normal mixture. *Skand. Aktuar.Tidskr.* 52:137–146.

Roeder, K. (1994). A graphical technique for determining the number of components in a mixture of normals. *Journal of the American Statistical Association* 89:487–495.

Sternberg, T. D. (1994). The duration of the rental housing vacancies. *Journal of Urban Economics* 36:143–160.

Webb, A. R. (2000). Gamma mixture models for target recognition. *Pattern Recognition* 33:2045–2054.