Zbornik radova Filozofskog fakulteta u Nišu
Serija Matematika 6:2 (1992), 343–348

FILOMAT '92, Niš, October 8–10, 1992

A NOTE ON MODELLING MIXED EXPONENTIAL DISTRIBUTION WITH NEGATIVE WEIGHTS

Vesna Jevremović

Abstract. *In this paper we compare three simulation technics for a random variable whose probability density function is a mixture of some exponential distributions with negative weights.*

1. INTRODUCTION.

The finite mixture distributions of the form:

$$\sum_{i=1}^{n} a_i f_i(x), \quad n \in N, n > 1 \tag{1}$$

where $a_i \in [0,1], \sum_{i=1}^{n} a_i = 1$ and $f_i(x)$, $i = 1, \dots, n$ are probability density functions, are videly used in practice. For the coefficients a_i, $i = 1, \dots, n$ the term "weights" is frequently used. When dealing with such a mixtures one has to solve many problems. Namely, using a sample of N realizations of some unknown random variable whose probability density function is $f(x)$, the following questions are to be answered:

(I) does $f(x)$ have the form (1),

(II) if $f(x)$ has the form (1), which is the number of components i.e. the value of n,

(III) if $f(x)$ has the form (1) with n components which are the values of $a_1, \dots, a_n$ and could the mixture be identified, i.e. are there two or more n-tuples $(a_1, \dots, a_n)$ giving the same probability density function (1).

Some of the questions mentioned above are partially solved, or solved for a particular class of probability density functions, see in [1]. If one wants to simulate a sequence of values for a random variable whose probability density function has the form (1), the composition method could be used, see in [2]. But, if some of the weights a_i, $i = 1, \dots, n$ could be negative, apart the questions (I), (II) and (III) there is one more: how to simulate a sequence of values for such a distributions? In this paper we try to answer this question for mixture of exponential distributions.

1991 Mathematics subject classification: 65C05, 62M99

Supported by Grant 0401A of FNS through Math. Inst. SANU

2. MODELLING MIXED EXPONENTIAL DISTRIBUTION WITH NEGATIVE WEIGHTS.

Let the random variable X have the probability density function

$$\sum_{i=1}^{n} a_i f_i(x), \quad n \in N, n > 1 \tag{2}$$

where $f_i(x) = \lambda_i e^{-\lambda_i x}$, $x \leq 0, \lambda_i > 0, \sum_{i=1}^{n} a = 1$, and some of a_i are negative. The function (2) could not be easily recognized as a probability density function, and we have to check the conditions:

$$f(x) > 0 \text{ for } x \in R \text{ and } \int_{-\infty}^{\infty} f(x)dx = 1. \tag{3}$$

2.1. THE APPROXIMATION OF THE INVERSE FUNCTION METHOD.

If the random variable X has the distribution function $F(x)$, and γ is a random number, then $x = F^{-1}(\gamma)$ gives the modelled value of the random variable X. This is the inverse function method, see in [2], which could not be applied when some of coefficients are negative.

Let the random variable X has the distribution function:

$$F(x) = 1 - A_1 e^{-ax} - A_2 e^{-bx}, \quad x > 0,$$

with: $A_1 + A_2 = 1, A_1 > 1, \quad a, b > 0$. The inverse function $F^{-1}(x)$ could not be find explicitly.

Let the random variable γ has the uniform distribution $\mathcal{U}(0,1)$. We have to determine the function $g(\gamma)$ which satisfies the condition:

$$P(g(\gamma), x) = f(x).$$

If we note:

$$e(z) = \begin{cases} 0, z \leq 0 \\ 1, z > 0 \end{cases},$$

then: $P(g(\gamma) < x) = \int_0^1 e(x - g(y))\,dy$. Let $0 < x_1 < x_2 < ... < x_N$, where $F(x_N) = 1 - \epsilon$, with $\epsilon \approx 0, e > 0$. One can choose the 5% , 10% ,... quantile of $F(x)$. From the equation: $\int_0^1 e(x - g(y))\,dy = F(x_k)$, it follows: $g(\gamma) = x_k - \epsilon$, for $\gamma \in (F(x_{k-1}), F(x_k))$, with $\epsilon \approx 0, e > 0$.

2.2. AUTOREGRESSIVE PROCESS WITH THE MIXED EXPONENTIAL MARGINAL DISTRIBUTION.

In [3] we give some examples of non-gaussian autoregressive and moving average processes whose marginal distributions are the mixture of exponential distributions with some negative weights. One of the processes considered there is the following:

$$X_t = \begin{cases} \alpha\xi_t, \text{ w.p. } p_0 \\ \beta\xi_y + X_{t-1}, \text{ w.p. } p_1, t \in D \\ X_{t-1}, \text{ w.p. } q_1 \end{cases} \tag{4}$$

where $\alpha, \beta, p_0, p_1, q_1 \in (0,1)$, with $p_0 + p_1 + q_1 = 1$, while " w.p." means " with probability". The sequence $\{\xi_t, t \in D\}$ is a sequence of independent identically distributed random variables, and we assume also the independence of ξ_t and X_s for $s < t$.

Let the distribution for the ξ_t sequence be exponential with the parameter λ. As we have shown in [3] the marginal distribution for the X_t sequence will be:

$$f_X(x) = a_1\gamma_1 e^{-\gamma_1 x} + a_2\gamma_2 e^{-\gamma_2 x}, \quad x > 0, \tag{5}$$

with:

(6,)
$$a_1 = p_0(\beta - \alpha)/((1 - q_1)\beta - p_0\alpha), \quad a_2 = 1 - a_1$$

$$\gamma_1 = \lambda/\alpha, \quad \gamma_2 = p_0\lambda/(\beta(1 - q_1)).$$

and it is easy to verify that:

(i) $p_0(\beta - \alpha) + p_1\beta < 0$ implies $a_1 > 1$,

(ii) $\beta < \alpha$ and $p_0(\beta - \alpha) + p_1\beta > 0$ implies $a_1 < 0$ and

(iii) in the other cases $a_1 \in [0,1]$.

The probability density function of the form (5) could be unimodal, and we shall prove the following statement:

Let $K = p_0/p_1$, $M = \alpha/\beta$. The mixed exponential distribution (5) has one local extremal value in the cases:

(1^0) If (i) holds with ($K > 1$ and $M > 1 + K$) or $K < 1$

(2^0) If (ii) holds with $K > 1$ or ($K < 1$ and $1 < M < 1 + K$)

(3^0) If $K = 1$ and $M \neq 2$

In the other cases we have $f'(x) < 0, x > 0$.

Proof. If (i) holds, the equation $f'(x) = 0, x > 0$, becomes

$$\frac{a_1}{|a_2} \frac{\gamma_1^2}{\gamma_2^2} = \exp((\gamma_1 - \gamma_2)x).$$

Since we have $\gamma_1 < \gamma_2$, the left hand side has to be less then 1. This inequality and the condition (i) give the result (1^0). In the other cases the proof is similar. Let us notice that the "convex" mixture (with all $a_1 > 0$) of exponential distributions, could not be unimodal because in the case of "convex" mixture we have always $f'(x) < 0, x > 0$.

2.3. THE REJECTION METHOD.

The method described in [4] is one of the rejection methods, and for a random variable X with the probability density function of the form

$$a_1 f_1(x) + a_2 f_2(x), \text{ with } a_2 < 0,$$

where $f_1(x)$ and $f_2(x)$ are probability density functions, consists of following steps:

(a) for a random number z we simulate a value of a random variable whose probability density function is $f_1(x)$,

(b) for a random number r and for a value x obtained in (a) we check if the condition:

$$r \leq \frac{f(x)}{a_1 f_1(x)}$$

is valid. In this case the value x is taken to be the modelled value for X, and if this inequality does not hold we repeat the step (a).

3. ONE EXAMPLE.

Let us consider the function:

$$f(x) = \begin{cases} 0, x \leq 0 \\ \frac{25}{12}e^{-5x/3} - \frac{5}{4}e^{-5x}, x > 0 \end{cases} \tag{7}$$

This function satisfies the conditions (3) and represents a probability density function. If we rewrite (7) in the form:

$$f(x) = \frac{5}{4} \cdot \frac{5}{3}e^{-5x/3} - \frac{1}{4} \cdot 5e^{-5x}, x > 0$$

we can see that this function is a mixture of two exponential distributions: $f_1(x) = \frac{5}{3}e^{-5x/3}, x > 0$, and $f_2(x) = 5e^{-5x}, x > 0$.

If we want to simulate a sequence of values for the random variable whose probability density function is of the form (7) using the approximation of the inverse function method, we write the distribution function:

$$F(x) = 1 - \frac{5}{4}e^{-5x/3} + \frac{1}{4}e^{-5x}, x > 0 \tag{8}$$

For the x_k sequence we shall take the following sequence: 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 3.0, 4.0, 5.0 .

Using a sequence of 40 random numbers (random number generator RAND, from MATLAB version 3.05), and $\epsilon = 0.001$ we obtain the following simulated values:

0.899	0.099	0.299	0.199	0.099	0.699	0.299	1.999	0.699	0.999
0.299	0.899	1.399	0.199	0.099	0.899	0.899	2.399	0.599	0.699
0.099	0.299	0.999	0.299	0.399	0.799	0.299	1.199	0.099	0.899
0.799	0.599	2.399	0.699	1.599	0.099	0.199	0.399	0.199	0.399

Table 1. Simulated values of the random variable with the distribution function (8)

If we want to calculate the corresponding χ^2 value, we consider the intervals: [0,0.199),[0.199,0.399),[0.399,0.699),[0.699,0.899), [0.899, 1.399) and [1.399, +8) and we obtain $\chi^2_5 = 3.996$, so it satisfies the χ^2 test of goodness of fit for the significance level level $\alpha < 0.5$.

There is also a possibility of choosing an appropriate random process whose marginal distribution is equal to the given function.

This second possibility shows also that the mixed distribution with negative weight is a distribution which can occur naturally.

As an example we consider the same distribution function as above and for our example, the probability density function (7) is obtained for $p_o = p_1 = q_1 = 1/3, \alpha = 0.6, \beta = 0.1$ and $\lambda = 1$, and satisfies (i).

In table 2 we give a sequence of 40 modelled values of the process (4) with the parameters $p_o = p_1 = q_1 = 1/3, \alpha = 0.6, \beta = 0.1$, and with exponential distribution $\mathcal{E}(1)$ for the ξ_t sequence. The values of the ξ_t sequence are obtained by the inverse function method. The initial value X_0 is taken to be 1.

0.200	0.200	0.257	0.284	0.284	0.763	0.833	0.833	1.075	0.114
0.134	1.656	1.875	2.044	0.723	0.815	0.815	0.898	1.089	1.117
1.298	0.421	0.350	1.650	1.760	1.767	1.903	1.903	1.903	1.919
0.094	0.611	0.478	0.478	0.316	0.602	0.164	0.668	0.674	0.516

Table 2. Simulated values of the random variable with the distribution function (8) obtained as values of the process (4) with the parameters: $\alpha = 0.6, \beta = 0.1, p_0 = p_1 = 1/3$.

We need 80 random numbers for 40 simulated values. The χ^2_5 value is (for the same intervals like above) 9.83, and satisfies the χ^2 test of goodness of fit for the significance level 0.05.

To avoid the repetitions of the same values in the modelled sample, for the same probability density function one can take another solution from the system of equations (6). Once, a_1, γ_1 and γ_2 given, we have:

$$\alpha = 1/\gamma_1, \beta = a_1/\gamma_2 - a_1/\gamma_1 + 1/\gamma_1, p_1/p_o = 1/(\gamma_2\beta) - 1,$$

so the probabilities p_o and p_1 are not unique . If we take for the process (4) with the marginal distribution (5) the parameters $\alpha = 0.6, \beta = 0.1$ and $p_o = p_1 = 0.45$ we obtain the sample ($\chi^2_5 = 7.875$) :

0.509	0.082	0.700	0.826	0.846	1.828	1.845	0.511	0.913	0.996
1.026	0.101	0.290	0.469	0.667	0.707	0.815	0.611	0.797	2.491
2.491	0.168	0.184	0.261	0.261	0.353	0.263	0.322	0.345	0.394
0.396	1.012	1.086	0.982	0.748	0.754	0.923	0.354	0.393	0.040

Table 3. Simulated values of the random variable with the with the distribution function (8) obtained as values of the process (4) with the parameters: $\alpha = 0.6, \beta = 0.1, p_0 = p_1 = 0.45$.

This second method needs more random numbers than the first one, is less accurate for the small sample size, but is related to the process which can be the model of some real phenomena. Let us notice that while the distribution of the form (2) could be identified, the corresponding autoregressive process could not be identified, but we have the possibility of choice, and if α is greater than b, the results in terms of χ^2 test are better.

Using the rejection method we obtain:

0.173	0.052	0.654	0.158	0.739	1.398	0.817	0.230	1.080	2.177
0.343	1.155	1.648	0.813	0.688	0.301	0.312	0.457	0.115	1.249
0.732	0.560	0.962	0.507	1.280	0.022	0.096	0.218	0.188	0.050
0.061	0.873	1.382	0.787	0.560	2.881	0.628	1.869	0.871	1.314

Table 4. Simulated values of the random variable with the distribution function (8) obtained by the rejection method.

A number of 114 random number is used, and the corresponding χ^2_5 value is 4.175. The procedure is simple, and the step (b) consists in checking the condition $r < 1 - 0.6z^2$.

References

[1] B.S.Everitt, D.J.Hand, *Finite mixture distributions*, Chapman and Hall, London, New York, 1981.

[2] I.M.Sobol', *Čislenny metody Monte-Karlo*, Nauka, Moskva, 1973. (Russian)

[3] V.Jevremović, *A note on mixed exponential distribution with negative weights*, Statistical & Probability Letters (USA) 11 (1991), 259–265.

[4] Bignami A., De Matteis A., *A note on sampling from combinations of distributions*, J. Inst. Maths Applics 8 (1971), 80–81.

Faculty of civil engineering,
University of Belgrade,
Bulevar Revolucije 73,
11000 Beograd, Yugoslavia