A method for the Generate a random sample from a finite mixture distributions

Dariush GHORBANZADEH
CNAM- Paris,
département Mathématiques-Statistiques

Philippe Durand
CNAM- Paris,
département Mathématiques-Statistiques

Luan Jaupi
CNAM- Paris,
département Mathématiques-Statistiques

***Abstract*—A finite m ixture m odel i s a c onvex c ombination of more probability density functions. By combining the properties of the individual probability density functions, mixture models are capable of approximating any arbitrary distribution. In this work we propose a method for the Generate a random sample from a finite mixture distribution. The proposed method envelope conventional models: translation, scaling and translation-scaling.**

Keywords: Finite mixture distribution, Mixture of normal distributions, Skew normal distribution.

I. Introduction

A finite m ixture m odel i s a c onvex c ombination o f more probability density functions. By combining the properties of the individual probability density functions, mixture models are capable of approximating any arbitrary distribution. Consequently, finite m ixture m odels a re a p owerful a nd flexible tool for modeling complex data. Mixture models have been used in many applications in statistical analysis and machine learning such as modeling, clustering, classification and latent class and survival analysis. Mixture of normal distributions has provided an extremely exible method of modeling a wide variety of random phenomena and has continued to receive increasing attention Titterington et all [1], Law & Kelton [2], Venkataraman [3] and Castillo & Daoudi [4]. In this work, we apply the model proposed in the mixture of normal distributions and Skew normal distributions who studied by several authors as Azzalini [5], Henze [6] and Ghorbanzadeh et all [7].

II. Mixture Models

we say that a distribution f is a mixture of k component distributions $f_1, \ldots, f_k$ if

$$f(x) = \sum_{i=1}^{k} \theta_i \, f_i(x) \tag{1}$$

with the θ_i being the mixing weights, $0 \leq \theta_i \leq 1$, $\theta_1 + \ldots + \theta_k = 1$. The equation (1) is a complete stochastic model, so it gives us a recipe for generating new data points: first pick a distribution, with probabilities given by the mixing weights, and then generate one observation according to that distribution. In practice, a lot of effort is given over to parametric mixture models, where the f_i are all from the same parametric family, but with different parameters, for example they might all be Gaussians with different centers and variances.

In the litterature , Simulation of a variate from a finite k-mixture distribution is undertaken in two steps. First a multivariate $Y : \theta_1, \ldots, \theta_k$ mixture indicator variate is drawn from the multinomial distribution with k probabilities equal to the mixture weights. Then, given the drawn mixture indicator value, k say, the variate X is drawn from the kth component distribution. The mixture indicator value k used to generate the $X = x$ is then discarded.

In this work we assume that the functions $f_1, \ldots, f_k$ are known and are defined from translation or scaling or translation-scaling of a kernel distribution.

III. Method of simulation

Let g a probability density function with the cumulative density function G. For the mixture distributions obtained by the translation of the kernel g we have the following proposition.

Proposition 1. *Let Y and Z two independent random variables with $Y \sim g$ and for $i = 1, \ldots, k$, $\mathbb{P}(Z = \mu_i) = \theta_i$. The random variable X defined by $X = Y + Z$ has the probability density function:*

$$f(x) = \sum_{i=1}^{k} \theta_i \, g(x - \mu_i) \tag{2}$$

Proof Conditional on $\{Z = \mu_i\}$, we have

$$\mathbb{P}(X \leq x | Z = \mu_i) = \mathbb{P}(Y + Z \leq x | Z = \mu_i) =$$
$$\mathbb{P}(Y \leq x - \mu_i) = G(x - \mu_i)$$

We deduce,

$$F_X(x) = \mathbb{P}(X \leq x) = \sum_{i=1}^{k} \mathbb{P}(X \leq x | Z = \mu_i) \, \mathbb{P}(Z = \mu_i)$$
$$= \sum_{i=1}^{k} \theta_i \, G(x - \mu_i)$$

by deriving , we get the probability density function of X defined in (2). □

6th Annual International Conference on Computational Mathematics, Computational Geometry & Statistics (CMCGS 2017)

ISSN 2251-1911
doi: 10.5176/2251-1911_CMCGS17.3

For the mixture distributions obtained by the scaling of the kernel g we have the following proposition.

Proposition 2. *Let Y and W two independent random variables with $Y \sim g$ and for $i = 1, \ldots, k$, $\mathbb{P}(W = \sigma_i) = \theta_i$ with $\forall i$, $\sigma_i > 0$. The random variable X defined by $X = WY$ has the probability density function:*

$$f(x) = \sum_{i=1}^{k} \theta_i \frac{1}{\sigma_i} g\left(\frac{x}{\sigma_i}\right) \tag{3}$$

Proof Conditional on $\{W = \sigma_i\}$, we have $\text{P}(X \leq x|W = \sigma_i) = \mathbb{P}(WY \leq x|W = \sigma_i) = \mathbb{P}\left(Y \leq \frac{x}{\sigma_i}\right) = G\left(\frac{x}{\sigma_i}\right)$
$We deduce, F_X(x) = \mathbb{P}(X \leq x) = \sum_{i=1}^{k} \mathbb{P}(X \leq x|W = \sigma_i)\, \mathbb{P}(W = \sigma_i) = \sum_{i=1}^{k} \theta_i\, G\left(\frac{x}{\sigma_i}\right)$
by deriving , we get the probability density function of X defined in (3). □
For the mixture distributions obtained by the translation-scaling of the kernel g we have the following proposition.

Proposition 3. *Let Y, Z and W three random variables Y independent of the pair (Z, W) with $Y \sim g$ and for $i = 1, \ldots, k$, $\mathbb{P}(Z = \mu_i, W = \sigma_i) = \theta_i$. The random variable X defined by $X = WY + Z$ has the probability density function:*

$$f(x) = \sum_{i=1}^{k} \theta_i \frac{1}{\sigma_i} g\left(\frac{x - \mu_i}{\sigma_i}\right) \tag{4}$$

Proof Conditional on $\{Z = \mu_i, W = \sigma_i\}$, we have

$$\mathbb{P}(X \leq x|Z = \mu_i, W = \sigma_i) =$$

$$\mathbb{P}(WY + Z \leq x|Z = \mu_i, W = \sigma_i)$$

$$= \mathbb{P}\left(Y \leq \frac{x-\mu_i}{\sigma_i}\right) = G\left(\frac{x-\mu_i}{\sigma_i}\right)$$

We deduce,

$$F_X(x) \quad = \mathbb{P}(X \leq x)$$

$$= \sum_{i=1}^{k} \mathbb{P}(X \leq x|Z = \mu_i, W = \sigma_i)\, \mathbb{P}(Z = \mu_i, W = \sigma_i$$

$$= \sum_{i=1}^{k} \theta_i\, G\left(\frac{x-\mu_i}{\sigma_i}\right)$$

by deriving , we get the probability density function of X defined in (4). □

IV. Simulation Results

In this part we will present the simulation results obtained by our method. We consider the mixture of translation, scaling and translation-scaling for two distributions normal and normal distribution Skew.

A. *Mixture of normal distributions*

For this part we consider the kernel probability density function g the standard normal $\mathcal{N}(0, 1)$ probability density function defined by $g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

In the case of the translation model we simulated a sample of size 5000 for the values $\theta_1 = 0.28$, $\theta_2 = 0.37$, $\theta_3 = 0.35$, $\mu_1 = -4$, $\mu_2 = -1$ and $\mu_3 = 3$. The following figure show the results obtained by the **Proposition 1**.

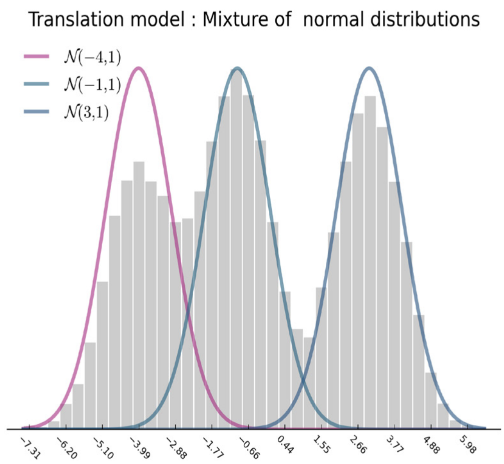

Figure 1. Simulation results for the mixture of three normal distributions with the mixing weights $\theta_1 = 0.28$, $\theta_2 = 0.37$ and $\theta_3 = 0.35$.

In the case of the scaling model we simulated a sample of size 5000 for the values $\theta_1 = 0.28$, $\theta_2 = 0.37$, $\theta_3 = 0.35$, $\sigma_1 = 2$, $\sigma_2 = 3$ and $\sigma_3 = \sqrt{3}$. The following figure show the results obtained by the **Proposition 2**.

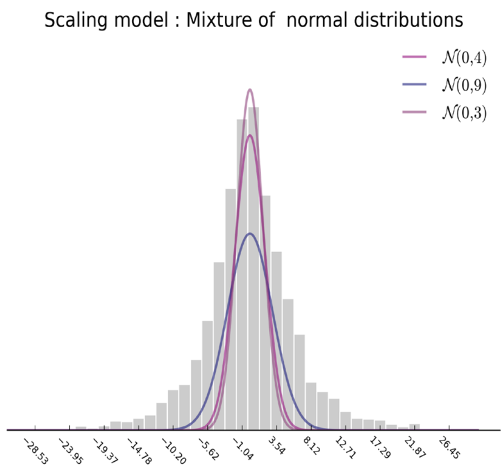

Figure 2. Simulation results for the mixture of three normal distributions with the mixing weights $\theta_1 = 0.28$, $\theta_2 = 0.37$ and $\theta_3 = 0.35$.

In the case of the translation-scaling model we simulated a sample of size 5000 for the values $\theta_1 = 0.25$, $\theta_2 = 0.35$, $\theta_3 = 0.40$, $(\mu_1, \sigma_1) = (-4, 1)$, $(\mu_2, \sigma_2) = (1, \sqrt{3})$ and $(\mu_3, \sigma_3) = (6, \sqrt{2})$. The following figure show the results obtained by the **Proposition 3**.

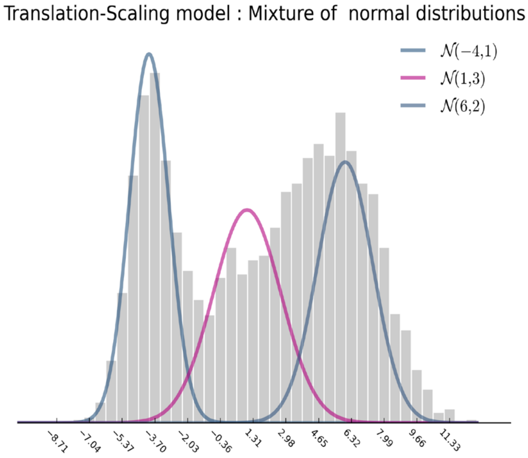

Figure 3. Simulation results for the mixture of three normal distributions with the mixing weights $\theta_1 = 0.25$, $\theta_2 = 0.35$ and $\theta_3 = 0.40$.

B. *Mixture of skew normal distributions*

The probability density function of the skew normal distribution with parameter λ noted $\mathcal{SN}(\lambda)$, is given by $g(x) = 2\,\varphi(x)\,\Phi(\lambda x)$, where φ and Φ denote the standard normal $\mathcal{N}(0, 1)$ probability density function and cumulative distribution function, respectively.

In this part for the simulation of the normal skew distribution , we apply the algorithm developed in Ghorbanzadeh et all [7]. In the case of the translation model we simulated a sample of size 5000 from $\mathcal{SN}(6)$ with translation values $\mu_1 = -1.5$, $\mu_2 = -1$, $\mu_3 = 0.5$ and with the mixing weights $\theta_1 = 0.28$, $\theta_2 = 0.37$, $\theta_3 = 0.35$. The following figure show the results obtained by the **Proposition 1**.

In the case of the scaling model we simulated a sample of size 5000 from $\mathcal{SN}(-7)$ with scaling values $\sigma_1 = 3$, $\sigma_2 = 4.7$, $\sigma_3 = 5.5$ and with the mixing weights $\theta_1 = 0.3$, $\theta_2 = 0.4$, $\theta_3 = 0.3$. The following figure show the results obtained by the **Proposition 2**.

In the case of the translation-scaling model we simulated a sample of size 5000 from $\mathcal{SN}(2)$ with scaling values translation-scaling $(\mu_1, \sigma_1) = (-4, 1)$, $(\mu_2, \sigma_2) = (-2, 3)$ and $(\mu_3, \sigma_3) = (2, 1.3)$ and with the mixing weights $\theta_1 = 0.25$, $\theta_2 = 0.35$, $\theta_3 = 0.4$. The following figure show the results obtained by the **Proposition 3**.

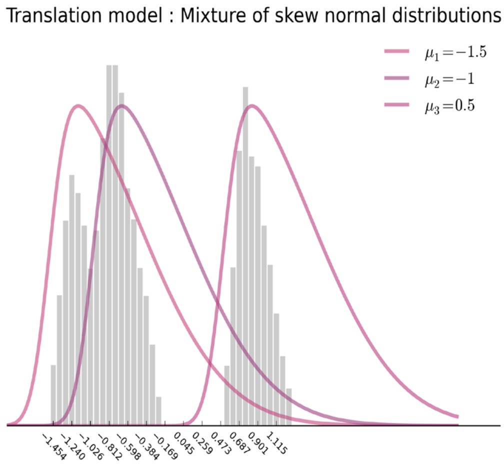

Figure 4. Simulation results for the mixture of three skew normal distributions with the mixing weights $\theta_1 = 0.28$, $\theta_2 = 0.37$ and $\theta_3 = 0.35$.

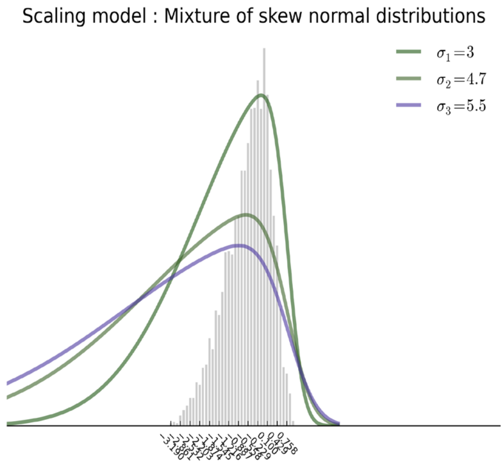

Figure 5. Simulation results for the mixture of three skew normal distributions with the mixing weights $\theta_1 = 0.3$, $\theta_2 = 0.4$ and $\theta_3 = 0.3$.

V. Conclusion

Researchers are faced with homogeneous data in their studies. They can model these data with a known distribution. In the growing research work area, the modeling might be easy like this. There are many data which are heterogeneous in many areas. In these cases, the mixture models can be more appropriate for modeling the data. The method studied in this work is very simple to implement and program. The results obtained by simulations for finite mixture weights with three components, are very satisfactory.

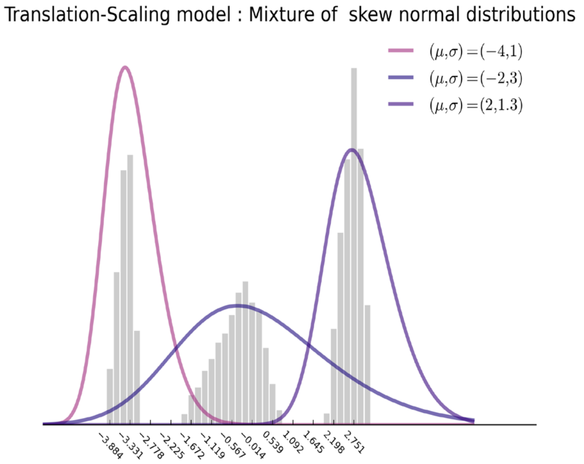

Figure 6. Simulation results for the mixture of three normal distributions with the mixing weights $\theta_1 = 0.25$, $\theta_2 = 0.35$ and $\theta_3 = 0.4$.

REFERENCES

[1] D.M. Titterington, A.F.M. Smith, and U.E. Makov (1985). Statistical Analysis of Finite Mixture Distributions. New York, London, Sydney: John Wiley & Sons.

[2] A. M. Law and W. D. Kelton (2000). Simulation Modeling and Analysis, 3rd Edition, McGraw-Hill, New York, NY.

[3] S. Venkataraman (1997). Value at risk for a mixture of normal distributions: the use of quasi-Bayesian estimation techniques. Economic Perspective, Federal Reserve Bank of Chicago, March/April, 2-13.

[4] J.Castillo J. Daoudi J (2009). The Mixture of Left-Right Truncated Normal Distributions." Journal of Statistical Planning and Inference, 139, 3543-3551.

[5] Azzalini, A. (1985) A class of distributions which includes the normal ones, Scandinavian Journal of Statistics, 12, 171-178.

[6] Henze, N. (1986). A probabilistic representation of the skew-normal distribution, Scandinavian Journal of Statistics, 13, 271-275.

[7] Ghorbanzadeh, D. , Jaupi, L. and Durand, P. (2014). A Method to Simulate the Skew Normal Distribution. Applied Mathematics, 5, 2073-2076. doi: 10.4236/am.2014.513201.