Mixtures With Negative Weights

Miguel Martins Felgueiras

CEAUL e ESTG do Instituto Politécnico de Leiria

Abstract

Allowing weights $w \notin [0;1]$, non-convex mixtures increase usual mixtures flexibility. Applications can be found, for instance, in internet traffic or queueing systems. For distribution families closed under minimization, we investigate finite mixtures with negative components. Our main purpose is to define this new mixtures and to study their properties.

keywords: negative weights, closed distributions, exponential mixtures.

AMS: 60E05

1 Introduction

Convex mixtures are indeed the most studied type of mixtures, specially because their weights can be regarded as subpopulations proportions. However, the condition $w \in [0, 1]$ is restrictive, and more flexible models are useful when modelling some data types.

Bartholomew (1969) and Steutel (1967, 1970) developed some preliminary work on this subject for non-convex mixtures of exponentials, and many applications have been developed recently. Gaussian mixtures also have been used in applied problems, but for other distributions non-convex mixtures have not received great attention from researchers.

In this paper we introduce a new class of distributions, called distributions closed under minimization, and use this new distribution family to develop non-convex mixtures with two components. Mixture moments, mode, and random numbers generation are also studied in this work. Finally, some possible extensions are discussed.

2 Distributions Closed Under Minimization

Definition 1.

Let $X_1, ..., X_N$ be a sequence of independent and identically distributed random variables to $X \sim F(x)$. F is closed under minimization $(X_{1:N})$ if

$$X_{1:N} \sim F_{\gamma_N}(\alpha_N x + \beta_N) \tag{1}$$

for some $\alpha_N > 0$ and β_N, $\gamma_N \in \mathbb{R}$.

These distributions have some interesting properties, namely

$$\overline{F}_{\gamma_N}(\alpha_N x + \beta_N) = P[X_{1:N} > x] = \left[\overline{F}(x)\right]^N \tag{2}$$

and

$$\alpha_N f_{\gamma_N}(\alpha_N x + \beta_N) = -\left[\left[\overline{F}(x)\right]^N\right]' = N f(x)\left[\overline{F}(x)\right]^{N-1}. \tag{3}$$

The case $\gamma_N = 1$ corresponds to classical extreme values theory, because only location-scale transforms are allowed. For $\gamma_N \neq 1$, the non-linear shape transformation leads to other types of models aside from GEV, as Generalized Logistic $[GL2(\lambda, \alpha)]$ with

$$F(x) = 1 - \left[\frac{\exp\left(-\frac{x}{\lambda}\right)}{1 + \exp\left(-\frac{x}{\lambda}\right)}\right]^\alpha$$

for $\lambda > 0$ and $\alpha > 0$, or Generalized Pareto $[GP(\lambda, \alpha)]$ with

$$F(x) = 1 - \left[1 + \frac{x-1}{\alpha\lambda}\right]^{-\alpha}$$

for $x > 1$ and $1 + \dfrac{x-1}{\alpha\lambda} > 0$.

3 Non-convex Mixtures for Distributions Closed Under Minimization

Theorem 1.
Let F be a distribution function and f a density function. If $w \in [-1,1]$,

$$f^*(x) = (1-w) f(x) + 2w f(x) F(x) \tag{4}$$

is always a density function.

This is a well known result and appears in a Gumbel (1958) work. Note that if $2f(x) F(x)$ is a density function, than f^* is a mixture density. So, for distributions fulfilling definition 1, non-convex mixtures can be introduced as the next theorem states.

Theorem 2.
X^* is a non-convex mixture with density

$$f^*(x) = (1+w) f(x) - w\alpha_2 f_{\gamma_2}(\alpha_2 x + \beta_2), \tag{5}$$

for $w \in [-1,1]$ and some α_2, β_2, γ_2, if and only if $X \sim F(x)$ is a random variable with distribution function closed under minimization, considering $N=2$.

Proof.
Equating expressions (4) and (5),

$$\begin{aligned}
(1+w) f(x) - w\alpha_2 f_{\gamma_2}(\alpha_2 x + \beta_2) &= (1-w) f(x) + 2w f(x) F(x) \\
[2w - 2wF(x)] f(x) &= w\alpha_2 f_{\gamma_2}(\alpha_2 x + \beta_2) \\
2\overline{F}(x) f(x) &= \alpha_2 f_{\gamma_2}(\alpha_2 x + \beta_2).
\end{aligned}$$

integrating both sides,

$$
\begin{aligned}
-\left[\overline{F}(x)\right]^2 &= F_{\gamma_2}(\alpha_2 x + \beta_2) + c \\
\left[\overline{F}(x)\right]^2 &= \overline{F}_{\gamma_2}(\alpha_2 x + \beta_2) + c'.
\end{aligned}
$$

If $c' \neq 0$,

$$
\lim_{x \to +\infty} \overline{F}_{\gamma_2}(\alpha_2 x + \beta_2) + c' = c' \neq 0
$$

and F cannot be a distribution function. For $c' = 0$,

$$
\left[\overline{F}(x)\right]^2 = \overline{F}_{\gamma_2}(\alpha_2 x + \beta_2)
$$

and X is a distribution closed under minimization, considering $N = 2$. □

The weight w defines the mixture type:

- if $w < 0$ we have a two densities sum, the first one contracted by $1 + w$ and the second one contracted by $-w$, where $0 < 1 + w < 1$ and $0 < -w < 1$ (usual convex mixture);
- if $w > 0$ we have a two densities subtraction, the first one expanded by $1 + w$ and the second one contracted by $-w$, where $1 < 1 + w < 2$ and $0 < -w < 1$ (non-convex mixture).

4 Moments and Mode

4.1 Mixture Moments

Assuming the existence of the involved k raw moments, consider the notation $\mu'_{X;k}$ for the original distribution and $\mu'_{X^+;k}$ for the minimum distribution. Mixture moments will simply be

$$
\mu'_{X^*;k} = (1 + w)\,\mu'_{X;k} - w\mu'_{X^+;k}. \tag{6}
$$

We can now compare mixture moments with the original moments. Note that

$$\begin{aligned}\mu'_{X^*;k} &> \mu'_{X;k} \Longleftrightarrow (1+w)\,\mu'_{X;k} - w\mu'_{X^+;k} > \mu'_{X;k} \\ &\Longleftrightarrow w\left[\mu'_{X;k} - \mu'_{X^+;k}\right] > 0\end{aligned}$$

For positive random variables, minimum moments will always be smaller than distributions moments, so

$$w\left[\mu'_{X;k} - \mu'_{X^+;k}\right] > 0 \Longleftrightarrow w > 0.$$

In this situation,

$$\mu'_{X^*;k} > \mu'_{X;k}$$

if $w > 0$ and

$$\mu'_{X^*;k} < \mu'_{X;k}$$

if $w < 0$.

4.2 Mode of the Mixtures

In the mixtures context, unimodality is always an important question. Computation of the mode (and even establishing mode existence) can be a difficult task, mainly because the involved density functions are complex. Besides, data unimodality or multimodality may be originated by a sample fluctuation. For the above defined non-convex mixtures, the mixture density function can be written as

$$\begin{aligned}f^*(x) &= (1-w)\,f(x) + 2wf(x)\,F(x) = \\ &= f(x)\left[1 - w + 2wF(x)\right] = f(x)\left[1 + w - 2w\overline{F}(x)\right], \qquad (7)\end{aligned}$$

leading to

$$f^{*\prime}(x) = f'(x)\left[1 + w - 2w\overline{F}(x)\right] + 2w\left[f(x)\right]^2. \qquad (8)$$

General conclusions cannot be taken, but besides the boundary of the domain, the mode candidates are the points where $f^{*\prime}(x) = 0$. This equation leads to

$$\begin{aligned} f^{*\prime}(x) &= 0 \Longleftrightarrow f'(x)\left[(1+w) - 2w\overline{F}(x)\right] + 2w\left[f(x)\right]^2 = 0 \Longleftrightarrow \\ &\Longleftrightarrow \frac{\left[f(x)\right]^2}{f'(x)} = \overline{F}(x) - \frac{1+w}{2w}. \end{aligned} \tag{9}$$

According with f density, the above equation may lead to a explicit solution. Otherwise, iterative methods must be applied.

5 Generating Random Numbers

For convex mixtures, random numbers generation is usually a simple task, and in general commercial and user free software generate this kind of numbers for common distributions.

The situation is far more complicated for non-convex mixtures, since negative weights are allowed and usual software does not generate these numbers. Under these circumstances we show that the inverse transform method works well, as the following theorem states.

Theorem 3.
Let X^* be a non-convex mixture as introduced by theorem 2, $X \sim F(x)$ the random variable with distribution function closed under minimization and $Y \sim U(0,1)$. Then

$$X^* \stackrel{d}{=} \overline{F}^{-1}\left[\frac{1+w-\sqrt{(1-w)^2+4wY}}{2w}\right]. \tag{10}$$

Proof.

Since

$$\begin{aligned} F^{*}(x) &= \int_{-\infty}^{x} \left[(1+w) f(t) - w\alpha_2 f_{\gamma_2}(\alpha_2 t+\beta_2)\right] dt = \\ &= (1+w) F(x) - wF_{\gamma_2}(\alpha_2 x+\beta_2) = \\ &= (1+w)\left(1-\overline{F}(x)\right) - w\left(1-\overline{F}_{\gamma_2}(\alpha_2 x+\beta_2)\right) = \\ &= 1+w-(1+w)\overline{F}(x) - w + w\left[\overline{F}(x)\right]^2 = 1-(1+w)\overline{F}(x)+w\left[\overline{F}(x)\right]^2 \end{aligned}$$

then

$$\begin{aligned} y &= 1-(1+w)\overline{F}(x)+w\left[\overline{F}(x)\right]^2 \underset{\overline{F}(x)=t}{\Longleftrightarrow} 1-y-(1+w)t+wt^2=0 \Longleftrightarrow \\ &\Longleftrightarrow\ t=\frac{1+w\pm\sqrt{(1+w)^2-4w(1-y)}}{2w} \underset{0\le\overline{F}(x)\le 1}{\Longrightarrow} \\ &\Longrightarrow\ \overline{F}(x)=\frac{1+w-\sqrt{(1+w)^2-4w(1-y)}}{2w} \Longleftrightarrow \\ &\Longleftrightarrow\ x=\overline{F}^{-1}\left[\frac{1+w-\sqrt{(1-w)^2+4wy}}{2w}\right]. \end{aligned}$$

□

6 Non-convex Exponential Mixtures

To exemplify the previous theory, let us consider $X \sim Exp(\lambda)$. Mixture density function will be

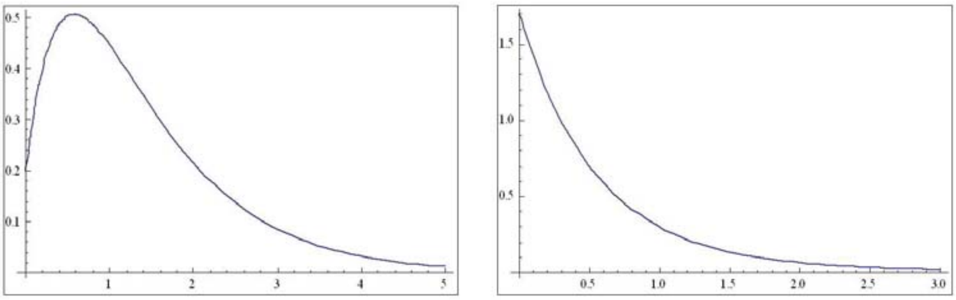
$$f^{*}(x) = (1+w)\lambda e^{-\lambda x} - 2w\lambda e^{-2\lambda x}, \tag{11}$$

and the corresponding X^{*} raw moments

$$\mu'_{X^{*};k} = (1+w)\frac{k!}{\lambda^k} - w\frac{k!}{(2\lambda)^k}. \tag{12}$$

According to λ and w, different density shapes are possible, for instance

Figure 1: Some non-convex exponential mixtures densities



$\lambda = 1$ and $w = 0.8$ $\qquad$ $\lambda = 1$ and $w = -0.7$

As far as the mode is considered, equation (9) originates the solution

$$\begin{aligned} \frac{\left[\lambda e^{-\lambda x}\right]^2}{-\lambda^2 e^{-\lambda x}} &= e^{-\lambda x} - \frac{1+w}{2w} \Longleftrightarrow -e^{-\lambda x} = e^{-\lambda x} - \frac{1+w}{2w} \Longleftrightarrow \\ &\Longleftrightarrow e^{-\lambda x} = \frac{1+w}{4w} \Longleftrightarrow w > 0 \wedge x = -\frac{1}{\lambda}\ln\left(\frac{1+w}{4w}\right). \quad (13) \end{aligned}$$

Noting that $x > 0$, then

$$-\frac{1}{\lambda}\ln\left(\frac{1+w}{4w}\right) > 0 \Longleftrightarrow \frac{1+w}{4w} < 1 \Longleftrightarrow w > \frac{1}{3}.$$

- When $w < 0$ we have an unimodal convex mixture with mode $x = 0$, already studied in literature.

- When $0 < w \leq \frac{1}{3}$ we have an unimodal non-convex mixture with mode $x = 0$. Note that for $0 < w < \frac{1}{3}$

$$\begin{aligned} f^{*\prime}\left(0^+\right) &= \left[-\left(1+w\right)\lambda^2 e^{-\lambda x} + 4w\lambda^2 e^{-2\lambda x}\right]_{x=0} = \\ &= -\left(1+w\right)\lambda^2 + 4w\lambda^2 = \lambda^2\left(3w-1\right) \underset{w<1/3}{<} 0. \end{aligned}$$

If $w = 1/3$,

$$f^{*\prime}\left(x\right) = \frac{4}{3}e^{-\lambda x}\left[e^{-\lambda x} - 1\right] < 0.$$

- When $w > \frac{1}{3}$
$$f^{*\prime}\left(0^{+}\right) = \lambda^{2}\left(3w-1\right) \underset{w>1/3}{>} 0,$$
implying that $x=0$ is not a mode, so we have an unimodal non-convex mixture with mode $x = -\frac{1}{\lambda}\ln\left(\frac{1+w}{4w}\right)$.

Finally, for mixture concavities study, since

$$f^{*\prime}\left(x\right) = \lambda^{2}e^{-\lambda x}\left[-\left(1+w\right)+4we^{-\lambda x}\right]$$

and

$$f^{*\prime\prime}\left(x\right) = \lambda^{3}e^{-\lambda x}\left[\left(1+w\right)-8we^{-\lambda x}\right]$$

then

$$\left[f^{*}f^{*\prime\prime}-\left(f^{*\prime}\right)^{2}\right]\left(x\right) = -2w\left(1+w\right)\lambda^{4}e^{-3\lambda x}, \tag{14}$$

implying that:

- for $w < 0$ expression (14) is always positive and so the mixture is infinitely divisible;[1]
- for $w > 0$ expression (14) is always negative and so the mixture is strongly unimodal (see Medgyessy, 1977).

7 Extensions

In the previous sections we have considered non-convex mixtures with two components, one with density function $f\left(x\right)$ and another with density function $\alpha_{2}f_{\gamma_{2}}\left(\alpha_{2}x+\beta_{2}\right)$. although these models might be interesting by themselves, they can be used to construct richer models, which is perhaps their main interest.

[1] Steutel (1967) had already showed infinite divisibility of all two exponential finite mixtures.

We can considerer a convex mixture of non-convex mixtures, leading to a random variable with density function

$$f^*(x) = \sum_{i=1}^{N} p_i \left[(1+w_i) f_i(x) - w_i \alpha_2 f_{i,\gamma_2}(\alpha_2 x + \beta_2) \right] \tag{15}$$

where $-1 < w_i < 1$, $0 < p_i < 1$ and $\sum_{i=1}^{N} p_i = 1$. This is a more complex model, with modelling advantages, but also with the handicap of increased number of parameters.

It is also possible to relax the condition $N = 2$, as long as f^* is a density function. This leads to a wider interval for w, but introduces a new parameter (N) that we shall probably need to estimate.

Theorem 4.
Let $X \sim F(x)$ be a random variable with distribution function closed under minimization. Then X^* is a non-convex mixture with density function

$$f^*(x) = (1+w) f(x) - w \alpha_N f_{\gamma_N}(\alpha_N x + \beta_N), \tag{16}$$

for $w \in \left[-1, (N-1)^{-1}\right]$, $N > 1$ and some α_N, β_N, γ_N.

Proof.
Density function f^* can be written as

$$f^*(x) = (1+w) f(x) - wN f(x) \left[\overline{F}(x)\right]^{N-1},$$

subject to the conditions $\int_{\mathbb{R}} f^*(x)\, dx = 1$ and $f^*(x) \geq 0$. The first condition is universal,

$$\begin{aligned} \int_{\mathbb{R}} f^*(x)\, dx &= \int_{\mathbb{R}} \left[(1+w) f(x) - wN f(x) \left[\overline{F}(x)\right]^{N-1} \right] dx = \\ &= (1+w) - w = 1. \end{aligned}$$

For the second condition,

$$\begin{aligned} f^*(x) &\geq 0 \Longleftrightarrow (1+w) f(x) - wNf(x) \left[\overline{F}(x)\right]^{N-1} \geq 0 \Longleftrightarrow \\ &\Longleftrightarrow wf(x) \left[1 - N \left[\overline{F}(x)\right]^{N-1}\right] \geq -f(x) \Longleftrightarrow \\ &\Longleftrightarrow w \left[N \left[\overline{F}(x)\right]^{N-1} - 1\right] \leq 1. \end{aligned}$$

When $w > 0$, the above inequality leads to the sufficient condition

$$\begin{aligned} w^{-1} &\geq N - 1 \underset{N>1}{\geq} N \left[\overline{F}(x)\right]^{N-1} - 1 \Longleftrightarrow \\ &\Longleftrightarrow w \leq (N-1)^{-1} \end{aligned}$$

and when $w < 0$

$$\begin{aligned} w^{-1} &\leq -1 \leq N \left[\overline{F}(x)\right]^{N-1} - 1 \Longleftrightarrow \\ &\Longleftrightarrow w \geq -1, \end{aligned}$$

originating the final solution

$$-1 \leq w \leq (N-1)^{-1}.$$

□

References

[1] Bartholomew, D. (1969). Sufficient conditions for a mixture of exponentials to be a probability density function. *Annals of Mathematical Statistics* **40**, 2189–2194.

[2] Frühwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*. New York: Springer.

[3] Gumbel, E. (1958). Distributions à plusieurs variables dont les marges sont donnés. *Comptes Rendus Academie Sciences Paris* **246**, 1717-2720.

[4] Medgyessy, P. (1977). *Decomposition of Superpositions of Density Functions and Discrete Distributions.* Budapest: Adam Hilger.

[5] Mendonça, S., Pestana, D. (2002). Misturas generalizadas e hiperexponenciais. *Actas do X Congresso Anual da SPE*, 427–433.

[6] Steutel, F. (1967). Note on the infinite divisibility of exponential mixtures. *The Annals of Mathematical Statistics* **38**, 1303–1305.

[7] Steutel, F. (1970). Preservation of Infinite Divisibility under Mixing and Related Topics. *Mathematical Center Tracts* **38**. Amsterdam, Mathematisch Centrum.