Scale Mixtures and Slash Distributions

Miguel Martins Felgueiras

CEAUL e ESTG do Instituto Politécnico de Leiria

Abstract

Pareto scale mixtures are very effective for modeling heavy tailed data. A new class of models is described, generalizing commonly used slash distributions. Mixture properties and possible applications are discussed.

keywords: Pareto distributions, scale mixtures, slash distributions.

AMS: 60E05

1 Introduction

Classical models assume a fixed scale parameter. However, in many situations it is advisable to randomize the scale parameter, with increased variability (Johnson *et al.*, 1992) — for instance, in biostatistical studies the negative binomial model is sometimes referred to as a "more flexible Poisson" since it is the result of modeling the number of eggs laid by females of certain species, the individual being $Poisson(\lambda)$, but considering that the λ's are values from a $Gamma(\alpha, \delta)$ random variable. This procedure leads to a hierarchical model randomizing the former one, and hence more flexible.

In many applications the $Gamma(\alpha, \delta)$ is considered a suitable scale mixing model, because its natural connection with the Laplace transforms brings in a useful toolbox of ready-to-use formulas, and in many cases the resulting mixture is reasonably tractable. But any positive random variables can be used to randomize a scale parameter, although in most cases the resulting mixture is difficult to work with, since usually the corresponding density functions are not expressable in a close form.

The family of Pareto distributions emerges as interesting randomization candidate, for two main reasons. First, it has a simple analytical form, leading to easy mixture densities computation. Second, Pareto's fat tail implies that the resulting densities will have higher kurtosis, useful in heavy tailed data modeling.

The mixture can be defined (following Kelker's (1971) notation) as

$$Y = \Theta X \tag{1}$$

where Θ, X are independent random variables with X absolutely continuous and $\Theta \sim Pareto\,(\alpha)\,,$

$$f_\Theta\,(\theta) = \alpha\theta^{-\alpha-1}, \qquad \theta \geq 1, \alpha > 0.$$

The fact that we use Pareto with left-endpoint $\alpha_\Theta = 1$ is in a sense a severe restriction, since it implies that $\mathbb{P}[|Y| > |X|] = 1$. Pareto random variables $\widetilde{\Theta} = \Theta - 1$ with support $\theta \geq 0$ could also be considered, covering all positive values. However, explicit density functions and interesting mixture distributions were not found in that more general setting. On the other hand, as $\theta > 1$, the above mentioned expansion has important consequences tied to stochastic ordering.

2 Mixture densities and other properties

The probability density function of the mixture $Y = \Theta X$ can be written as

$$f_Y\,(y) = \int_1^\infty \alpha\theta^{-\alpha-2} f_X\left(\tfrac{y}{\theta}\right) d\theta, \tag{2}$$

originating for some usual X distributions the incomplete gamma and beta based densities (Felgueiras, 2008) presented in table 1.

Since the support of Θ is $S_\Theta = [1, \infty[$, multiplying X by Θ implies expansion of the X values. Clearly, the absolute values of the existing moments of such mixtures are always greater than the corresponding X moments. Further,

$$P(Y > t) > P(X > t) \Longleftrightarrow \overline{F}_Y(t) > \overline{F}_X(t), \quad t > 0,$$

i.e. Y stochastic dominates X, a potentially important fact in reliability modeling and in premium computing policies in actuarial applications (Centeno and Andrade e Silva, 2001).

When α increases,

$$\lim_{\alpha \to +\infty} \overline{F}_{\Theta_\alpha}(\theta) = \lim_{\alpha \to +\infty} \theta^{-\alpha} = \begin{cases} 0, & \theta > 1 \\ 1, & \theta = 1 \end{cases}$$

and Θ_α converges to the degenerate random variable at 1.

Convergence in distribution to a constant implies convergence in probability, and by convergence in probability properties, when $\alpha \to +\infty$ then

$$Y = \Theta_\alpha X \underset{\alpha \to \infty}{\overset{\mathrm{d}}{\longrightarrow}} X. \tag{3}$$

Thus, the mixture model can be near the original, for large values of α, or more far apart when α is small, leading to a wide range of solutions.

3 Mixture and slash distribution extensions

The mixture can also be regarded as a random variable quotient,

$$Y = \Theta X = \frac{X}{\Theta^{-1}}, \tag{4}$$

Table 1: Some Pareto scale mixtures densities

Distribution	Density	Mixture density
$X \sim N(0,1)$	$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$	$f_Y(y) = \frac{\alpha 2^{0.5\alpha-1} \gamma\left(\frac{\alpha+1}{2}, \frac{y^2}{2}\right)}{\sqrt{\pi}\, \lvert y \rvert^{\alpha+1}}, \quad y \neq 0$
$\frac{2^{-\frac{3+\beta}{2}} \exp\left[-0.5 \lvert x \rvert^{\frac{2}{1+\beta}}\right]}{\Gamma\left(\frac{3+\beta}{2}\right)}, \quad -1 < \beta \leq 1$		$f_Y(y) = \frac{\alpha(1+\beta)\gamma\left(\frac{\beta+1}{2}(\alpha+1), 0.5\lvert y \rvert^{\frac{2}{1+\beta}}\right)}{2^{-\alpha\frac{\beta+1}{2}} 4\Gamma\left(\frac{3+\beta}{2}\right) \lvert y \rvert^{\alpha+1}}, \; y \neq 0$
$X \sim Cauchy(0,1)$	$f_X(x) = \frac{1}{\pi} \frac{1}{1+x^2}$	$f_Y(y) = \frac{\alpha y^{-\alpha-1}}{\pi} \int_0^y \frac{z^\alpha}{1+z^2} dz, \quad y \neq 0$
$X \sim Gama(\beta,1)$	$f_X(x) = \frac{1}{\Gamma(\beta)} x^{\beta-1} e^{-x}$	$f_Y(y) = \frac{\alpha y^{-\alpha-1}}{\Gamma(\beta)} \gamma(\alpha+\beta, y), \quad y > 0$
$X \sim Beta(p,q)$	$f_X(x) = \frac{(1-x)^{q-1}}{x^{1-p} B(p,q)}$	$f_Y(y) = \begin{cases} \frac{\alpha B(p+\alpha, q, y)}{y^{\alpha+1} B(p,q)}, & 0 < y < 1 \\ \frac{\alpha B(p+\alpha, q)}{y^{\alpha+1} B(p,q)}, & y \geq 1 \end{cases}$
$X \sim Weibull(\beta,1)$	$f_X(x) = \beta x^{\beta-1} e^{-x^\beta}$	$f_Y(y) = \frac{\alpha\gamma\left(\alpha\beta^{-1}+1, y^\beta\right)}{y^{\alpha+1}}, \quad y > 0$
$X \sim Pareto(\beta)$	$f_X(x) = \beta x^{-\beta-1}$	$f_Y(y) = \begin{cases} \alpha^2 y^{-\alpha-1} \ln y, & \alpha = \beta, \quad y > 0 \\ \frac{\alpha\beta\left(y^{-\alpha-1} - y^{-\beta-1}\right)}{\beta-\alpha}, & \alpha \neq \beta, \quad y > 0 \end{cases}$

where

$$f_{\Theta^{-1}}(\theta) = f_{\Theta}\left(\theta^{-1}\right)\theta^{-2} = \alpha\theta^{\alpha-1}, \quad 0 < \theta \leq 1, \alpha > 0,$$

and so

$$\Theta^{-1} \sim Beta(\alpha, 1). \tag{5}$$

When $\alpha = 1$, the expressions above simplify, and since $\Theta^{-1} \sim U(0,1)$ we obtain slash distribution family, often used in reliability and robustness studies (Gómez *et al*, 2007; Johnson *et al.*, 1994).

In this context, it is obvious that Pareto scale mixtures generalize the class of slash distributions, and therefore share their wide range of applications, namely in situations where symmetrical distributions with fat tails are appropriated. For $0 < \alpha < 1$, Pareto scale mixtures have heavier tailweight than the slash distributions, and for $\alpha > 1$ we have the reverse situation.

As a side result, we prove that slash distributions do not have mean value.

Theorem 1.
Let $Y = \Theta X$, where Θ, X are independent random variables, X is absolutely continuous and $\Theta \sim Pareto(1)$. Then Y does not have mean value.

Proof.
When $E(X) = C \neq 0$, then if Y mean exists $E(Y) = E(\Theta)E(X) = cE(\Theta)$. Since $E(\Theta)$ does not exists for $\Theta \sim Pareto(1)$, then it is obvious that also Y mean does not exists. For $E(X) = 0$, note that

$$\begin{aligned} f_Y(y) &= \int_1^{+\infty} \theta^{-3} f_X\left(\frac{y}{\theta}\right) d\theta = \int_{-\infty}^{+\infty} \frac{f_X(x)}{|x|} f_\Theta\left(\frac{y}{x}\right) dx = \\ &= \int_{-\infty}^{+\infty} \frac{f_X(x)}{|x|}\left(\frac{y}{x}\right)^{-2} dx = \frac{1}{y^2}\int_{-\infty}^{+\infty} |x| f_X(x)\, dx, \quad \frac{y}{x} > 1 \end{aligned}$$

leading to

$$f_Y(y) = \begin{cases} \frac{1}{y^2} \int_0^y x f_X(x)\, dx, & y > x > 0, y > 0 \\ \frac{1}{y^2} \int_y^0 -x f_X(x)\, dx, & y < x < 0, y < 0 \end{cases}.$$

The expectation of Y exists if and only if

$$E(|Y|) = \int_{-\infty}^{0} |y| \left[\frac{1}{y^2} \int_y^0 -x f_X(x)\, dx \right] dy + \int_0^{+\infty} |y| \left[\frac{1}{y^2} \int_0^y x f_X(x)\, dx \right] dy$$

is convergent. In what concerns the second integral in the right hand side of that expression

$$\int_0^{+\infty} |y| \left[\frac{1}{y^2} \int_0^y x f_X(x)\, dx \right] dy = \int_0^{+\infty} \frac{1}{y} \left[\int_0^y x f_X(x)\, dx \right] dy,$$

and using straightforward inequalities,

$$\begin{aligned} \int_0^{+\infty} \frac{1}{y} \left[\int_0^y x f_X(x)\, dx \right] dy &> \int_1^{+\infty} \frac{1}{y} \left[\int_1^y x f_X(x)\, dx \right] dy > \\ &> \int_1^{+\infty} \frac{1}{y} \left[\int_1^y f_X(x)\, dx \right] dy = \\ &= \int_1^{+\infty} \frac{1}{y} \left[F_X(y) - F_X(1) \right] dy, \end{aligned}$$

as $\lim_{y \to +\infty} y \times \frac{1}{y} \left[F_X(y) - F_X(1) \right] = 1 - F_X(1) = C > 0$ we conclude that

$$\int_1^{+\infty} \frac{1}{y} \left[F_X(y) - F_X(1) \right] dy$$

is divergent and hence the expectation of Y doesn't exist. □

4 Examples

4.1 Pareto mixtures of normal random variables

Pareto mixtures of normals show the important features of Pareto mixtures of a symmetrical population, and are potentially the more widely useful. In fact, when $X \sim N(0,1)$ we obtain an infinitely divisible mixture (Kelker, 1971) with density

$$f_Y(y) = \alpha 2^{0.5\alpha - 1} |y|^{-\alpha-1} \pi^{-0.5} \gamma\left(\frac{\alpha+1}{2}, \frac{y^2}{2}\right), \quad y \neq 0, \tag{6}$$

where

$$\gamma(a,y) = \int_0^y t^{a-1} e^{-t} dt. \tag{7}$$

For instance, for $\alpha = 1$

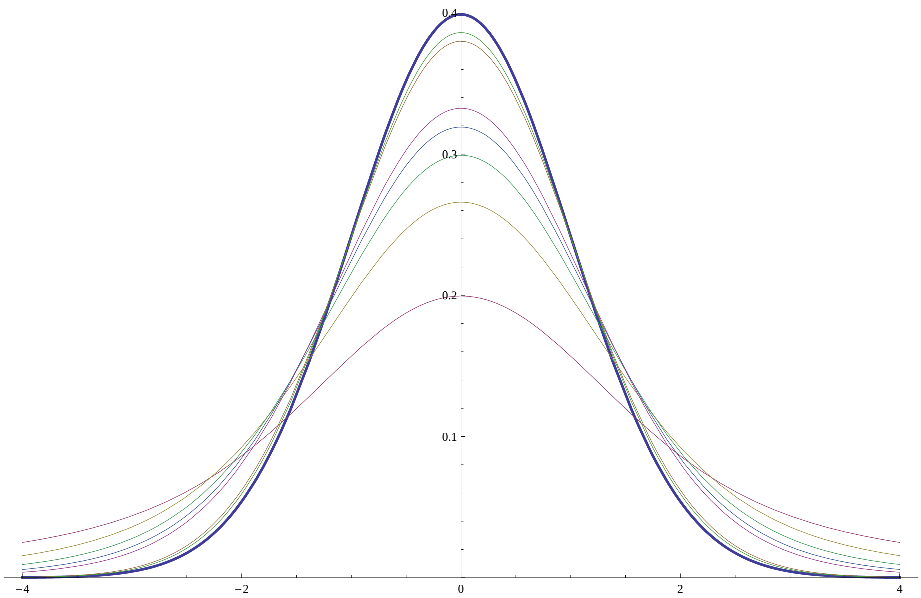
$$f_Y(y) = \frac{1 - e^{-\frac{y^2}{2}}}{\sqrt{2\pi} y^2}, \quad y \neq 0, \tag{8}$$

and for $\alpha = 3$

$$f_Y(y) = \frac{3\left(2 - (2+y^2) e^{-y^2/2}\right)}{\sqrt{2\pi} y^4}, \quad y \neq 0. \tag{9}$$

As previously stated, $\Theta_\alpha X \xrightarrow[\alpha \to \infty]{d} X$. This can be seen in the graphical representation below

Figure 1: Some non convex gaussian mixtures densities



The thick line represents $N(0,1)$ and the other lines the mixture for $\alpha = 1, ..., 5, 20, 30$.

Note that the α parameter works in a rather similar way as the n parameter in *t-Student* distributions. However, in this situation, the Y distribution as heavier tails (for small values of α) and the rate of convergence towards the gaussian limit is slower than in the t family.

Another symmetrical mixture with even heavier tails can be generated for $X \sim Cauchy(0,1)$ and $\alpha = 1$, originating the slash Cauchy density

$$f_Y(y) = \frac{\ln(y^2+1)}{2\pi y^2}, \quad y \neq 0. \tag{10}$$

In the next table, we can observe that Cauchy and slash gaussian quantiles are not far apart, but the slash Cauchy has impressive larger quantiles, and therefore can be useful in modeling very extreme situations.

Table 2: Probability quantiles for the Cauchy, the slash gaussian and the slash Cauchy

α	0.5	0.75	0.90	0.95	0.99	0.999
q_α Cauchy	0	1.00	3.08	6.31	31.82	318.31
q_α slash gaussian	0	1.47	3.99	7.98	39.89	398.94
q_α slash Cauchy	0	2.45	10.75	27.46	200.57	2850.55

4.2 Pareto mixtures of positive random variables

To exemplify Pareto mixtures of positive random variables we choose exponential parent, since it exhibits the more important features of mixtures of a positive support population, and it is the more readily useful in applications.

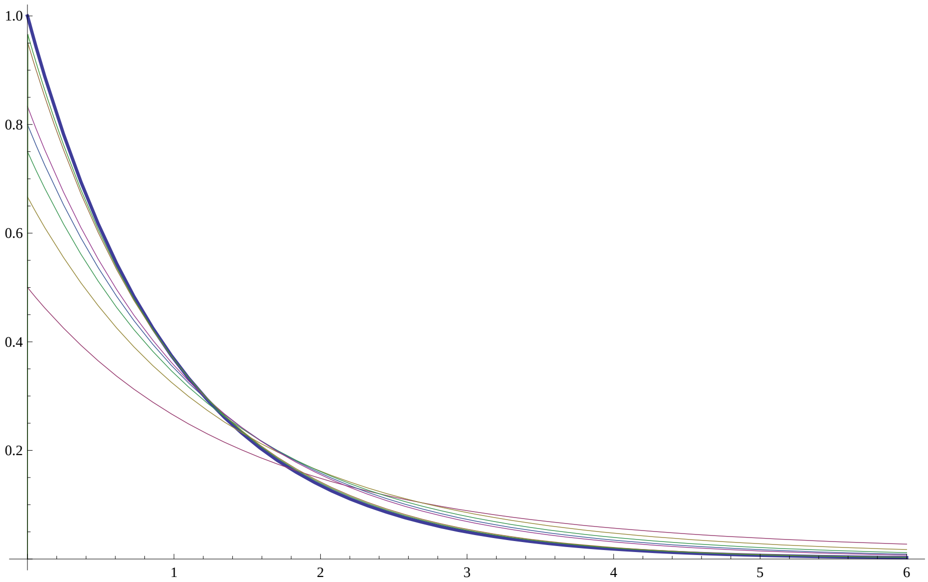
When $X \sim Exp(1)$ we obtain the infinitely divisible mixture (Steutel, 1970) with density[1]

$$f_Y(y) = \frac{\alpha\gamma(\alpha+1, y)}{y^{\alpha+1}}, \quad y > 0. \tag{11}$$

[1] $\beta = 1$ in the gamma mixture density presented in table 1.

Graphically,

Figure 2: Some exponential mixtures densities



The thick line represents $Exp\,(1)$ and the other lines the mixture for $\alpha = 1, ..., 5, 20, 30$.

The procedures and the results are similar to that we presented for the gaussian mixtures. We are performing scale transformations, so mixture density shape always look alike to the original X density shape. As observed previously, varying the α parameter leads to versatile control of the tailweight of the resulting mixtures.

References

[1] Centeno, M. L., and Andrade e Silva, J. (2001). Bonus systems in an open portfolio. *Insurance: Mathematics and Economics* **28**, 341–350.

[2] Felgueiras, M. (2008). Considerações sobre a distribuição Pareto. *Actas do XV Congresso Anual da Sociedade Portuguesa de Estatística*, 193–201.

[3] Gómez, H., Quintana, F., Torres, F. (2007). A new family of slash-distributions with elliptical contours. *Statistics and Probability Letters* **77**, 717–725.

[4] Johnson, N., Kotz, S., Balakrishnan, N. (1994). *Continuous Univariate Distributions, Volume I.* Wiley.

[5] Johnson, N., Kotz, S., Kemp, A. (1992). *Univariate Discrete Distributions.* Wiley.

[6] Kelker, D. (1971). Infinite divisibility and variance mixtures of the normal distribution. *Ann. Math. Statist.* **42**, 802–808.

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