A note on rescalings, reparametrizations and classes of distributions

M.C. Jones

Department of Statistics, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK

Abstract

Families of distributions on the real line have location, scale and shape parameters. Reparametrizations of such distributions, particularly involving rescalings, do not produce different families of distributions. However, from time to time, 'extended' or 'generalized' distributions are proposed that are nothing other than such reparametrizations. This note was triggered by a specific instance recently published in this journal but other examples are briefly mentioned as well.

MSC: 60E05; 62E10.

Keywords: Generalized distributions; Two-piece distributions.

Abbreviated Title: Rescalings and Reparametrizations.

1. Introduction

Let f be the density of a simple univariate distribution having support the whole real line, $\Re$, or sometimes a subset thereof. A popular problem in distribution theory is as follows. Suppose that $f(x;\mu,\sigma) = \sigma^{-1}f(\sigma^{-1}(x-\mu))$ has no parameters other than the ubiquitous location, $\mu \in \Re$, and scale, $\sigma > 0$. Then, introduce one or more further, shape, parameters, β say, into f with a view to generating a family of distributions with a wide variety of features, particularly skewness and/or kurtosis. (Actually, f can already have other shape parameters, but let us keep things simple.) Canonically, one should work in terms of f alone, generate one's family as $f(x;\beta)$ and reintroduce μ and σ — which are essential for all practical purposes involving data — through $f(x;\mu,\sigma,\beta) = \sigma^{-1}f(\sigma^{-1}(x-\mu);\beta)$ later. It is clear that any two proposed distributions which differ only in terms of scale are essentially (i.e. canonically) the same. Indeed, in general, one can reparametrize $f(x;\mu,\sigma,\beta)$ as $f(x;\mu',\sigma',\beta')$ through an invertible function $\{\mu',\sigma',\beta'\} = p(\mu,\sigma,\beta)$, say, and one still has precisely the same distribution. In particular, we obtain precisely the same answer when either form of the distribution is fitted to data by any method that is invariant to reparametrization e.g. maximum likelihood.

With this background, it is surprising that, from time to time, reparametrized families of distributions which differ from existing families of distributions only by a scale change (in addition to the reparametrization) are claimed to be 'extensions' or 'generalizations' of the existing families. They are not 'new' families of distributions; they are just different parametrizations — perhaps helpfully different — of the same family of distributions. Some examples of this error form the remainder of this note.

2. Two-piece distributions

This note was triggered by the following example from Arellano-Valle, Gómez and Quintana (2005). It concerns two-piece or split distributions of the form

$$f(x;\beta) = \frac{2}{1+\beta}\left\{f(x)I(x<0) + f\left(\frac{x}{\beta}\right)I(x \geq 0)\right\} \tag{1}$$

where f is an arbitrary symmetric distribution and $\beta > 0$ is a one-dimensional parameter the role of which is to introduce skewness. Formulation (1) itself represents an existing but arguably sub-optimal parametrization of these two-piece distributions. An attractive alternative parametrization was given by Fernandez and Steel (1998):

$$f(x;\gamma) = \frac{2\gamma}{1+\gamma^2}\left\{f(\gamma x)I(x<0) + f\left(\frac{x}{\gamma}\right)I(x\geq 0)\right\}, \tag{2}$$

$\gamma > 0$. Another interesting parametrization, proposed specifically in the normal case but generally applicable, is Mudholkar and Hutson's (2000):

$$f(x;\epsilon) = f\left(\frac{x}{1+\epsilon}\right)I(x<0) + f\left(\frac{x}{1-\epsilon}\right)I(x\geq 0), \tag{3}$$

$-1 < \epsilon < 1$. And finally, Arellano-Valle et al. (2005) actually work with

$$f(x;a(\alpha),b(\alpha)) = \frac{2}{a(\alpha)+b(\alpha)}\left\{f\left(\frac{x}{a(\alpha)}\right)I(x<0) + f\left(\frac{x}{b(\alpha)}\right)I(x\geq 0)\right\}, \tag{4}$$

$a(\alpha), b(\alpha) > 0$.

Arellano-Valle et al. claim that (4) — or indeed the location-scale version thereof, in their Definition 2 — *extends* (2) and (3). I claim that all four formulations are simply rescaled reparametrized versions of one another, and hence all represent the same three-parameter location/scale/skewness family. The match-up is readily calculated to be

$$\beta = \gamma^2 = \frac{1-\epsilon}{1+\epsilon} = \frac{b(\alpha)}{a(\alpha)}$$

with rescalings (relative to formulation (1), but readily modified to be relative to whichever formulation one prefers)

$$1 \;:\; \gamma : \;\frac{1}{1+\epsilon}\;:\;\frac{1}{a(\alpha)}.$$

Note also that the skewness parameter in formulation (4) is simply $b(\alpha)/a(\alpha)$; neither the two functions nor their dependence on α introduces any more generality into the distributions.

Arellano-Valle et al., Remark 1, protest that (2), (3) and (4) are not reparametrizations of one another, citing two specific arguments. First, they note that when f is the $U(-1,1)$ density, (2) yields $U(-1/\gamma,\gamma)$ while (3) gives $U(-2/(1+\gamma^2), 2\gamma^2/(1+\gamma^2))$. They say "These are clearly different". I would agree only if the general formulation were truly one-parameter and location and scale parameters are ignored. But introduce one of these parameters back into the distribution and both are simply 'the uniform distribution'. (In fact, one of the three parameters is redundant and the other two are equivalent to the ends of the interval.) Arellano-Valle et al. also note, with respect to (2) and (3), "that the asymptotic behavior when the corresponding asymmetry parameter approaches the extremes is radically different". But this is simply a reflection of the standard property of different limits arising depending on different normalizations (here, rescalings) of the same distribution.

3. Other examples

Briefly, the above discussion can be compared with earlier cases where rescaled distributions are erroneously referred to as 'generalized' distributions.

The first example lives on $(0, \infty)$. There, it is well known that the F distribution and the 'beta distribution of the second kind' are the same distribution except for a particular rescaling (e.g. Johnson, Kotz and Balakrishnan, 1995, Chapter 27). Pham-Gia and Duong (1989) scale the F distribution in a third way and call it the generalized F distribution. (Relatedly, of course, chi-squared and gamma distributions are the same distribution with different scalings.)

The other example is on $\Re$, but with immediate application to multivariate versions through its elliptically symmetric extension. The t and Pearson Type VII distributions are the same except for both a rescaling and a reparametrization (e.g. Johnson, Kotz and Balakrishnan, 1995, Chapter 28). And so is the distribution at the heart of the paper by Arellano-Valle and Bolfarine (1995) ... where it is called the generalized t distribution.

References

Arellano-Valle, R.B., Bolfarine, H., 1995. On some characterizations of the t-distribution. Statist. Probab. Lett. 25, 79–85.

Arellano-Valle, R.B., Gómez, H.W., Quintana, F.A., 2005. Statistical inference for a general class of asymmetric distributions. J. Statist. Planning Inference 128, 427–443.

Fernández, C., Steel, M.J.F., 1998. On Bayesian modelling of fat tails and skewness. J. Amer. Statist. Assoc. 93, 359–371.

Johnson, N.L., Kotz, S., Balakrishnan, N., 1995. Continuous Univariate Distributions, Volume 2 (Second Edition). Wiley, New York.

Mudholkar, G.S., Hutson, A., 2000. The epsilon-skew-normal distribution for analyzing near-normal data. J. Statist. Planning Inference 83, 291–309.

Pham-Gia, T., Duong, Q.P., 1989. The generalized beta- and F-distributions in statistical modelling. Math. Comput. Modelling 12, 1613–1625.