THE UNIFORM MIXTURE OF GENERALIZED ARC-SINE DISTRIBUTIONS

M.C. Jones[1]

Abstract

A single, tractable, special case of the problem of continuous mixtures of beta distributions over their parameters is considered. This is the uniform mixture of generalized arc-sine distributions which, curiously, turns out to be linked by transformation to the Cauchy distribution.

Let Beta(a, b) stand for the usual beta distribution on $[0, 1]$ with parameters $a, b > 0$. Mixtures of beta distributions are commonly found in the statistical literature, but arise almost always in the form of discrete mixtures over a and/or b or else, very occasionally, as mixtures over an additional scale parameter (Bertin *et al.*, 1997). Continuous mixtures over a and b are conspicuous by their absence. This is unsurprising on tractability grounds, given the need to integrate over functions of a and b including a beta function.

In this note, it is observed that one very special case of a continuous mixture of beta distributions is simple and tractable. This special case is the distribution of U where $U|\Theta = \theta \sim \text{Beta}(\theta, 1 - \theta)$, $0 < \theta < 1$, and $\Theta \sim U(0, 1)$, the continuous uniform distribution on $(0, 1)$. The Beta$(\theta, 1 - \theta)$ distribution is also known as the generalized arc-sine distribution. Its density is

$$f(u|\theta) = \frac{u^{\theta-1}(1-u)^{-\theta}}{B(\theta, 1-\theta)} = \frac{\sin(\pi\theta)}{\pi u}\left(\frac{u}{1-u}\right)^{\theta}, \qquad 0 < u < 1,$$

[1]Department of Statistics, The Open University, Walton Hall, Milton Keynes, MK7 6AA, U.K. (e-mail: m.c.jones@open.ac.uk)

so that the density of its uniform mixture is

$$f(u) = \frac{1}{\pi u} \int_0^1 \sin(\pi\theta) \exp\{\theta \log(u/(1-u))\} d\theta$$

$$= \frac{1}{u(1-u)} \frac{1}{\left\{\pi^2 + \log^2\left(\frac{u}{1-u}\right)\right\}}, \qquad 0 < u < 1. \qquad (1)$$

Five representative densities of the generalised arc-sine family (θ = 0.1(0.2)0.9) are plotted in Figure 1 along with density (1) (the dashed line). (All densities tend to ∞ as $u \to 0, 1$, although u has to be very close to 0 or 1 for this to be apparent for θ close to 1 or 0, respectively.) Note that the uniform mixture density shares the property of being U-shaped with all individual generalized arc-sine distributions.

The uniform "averaged generalized arc-sine density" (1) can be immediately recognised as the distribution of $U = e^{\pi X}/(1 + e^{\pi X})$ where $X \sim$ standard Cauchy. As such, it arises from the logistic transformation sub-case of the transformation approach of Johnson (1949) applied to Cauchy rather than normal random variables. This, in turn, means that distribution (1) is also one of a tractable family of distributions on [0,1] that form alternatives to the beta distribution; Johnson and Tadikamalla (1982) provide, through similar transformation of the logistic, another attractive family.
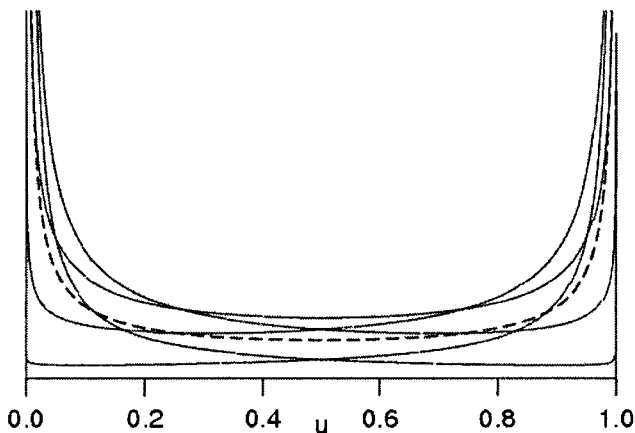

FIGURE 1 *Five generalized arc-sine densities (solid lines) and the uniform mixture density (1) (dashed line). To identify individual generalized arc-sine densities, note that at $u = 0.4$, the densities correspond to $\theta = 0.9, 0.1, 0.7, 0.3$ and 0.5 in increasing values of $f(0.4)$.*

The uniform mixture distribution can be contrasted with the "median generalized arc-sine density", that of Beta$(1/2, 1/2)$, the ordinary arc-sine density, which is the symmetric member of the set of generalized arc-sine densities shown in Figure 1 This is the distribution of $V = (1/2)(1 + X/\sqrt{1 + X^2})$, a simple consequence of the relationship of Cacoullos (1965). Note that the latter has more probability mass than the former in the interior of [0,1] and less probability mass very close to the boundaries. This is a consequence of the relative spreads of the respective transformations.

The author's motivation for this study was his interest in perhaps replacing a beta kernel estimator of a density on $[0, 1]$, due to Chen (2000), which is of the form

$$n^{-1} \sum_{i=1}^{n} K\left(Y_i; (y/w) + 1, ((1-y)/w) + 1\right)$$

by its natural converse

$$n^{-1} \sum_{i=1}^{n} K\left(y; (Y_i/w) + 1, ((1-Y_i)/w) + 1\right).$$

Here, $Y_1, ..., Y_n$ are the data, y is the point at which estimation takes place, w is a smoothing parameter and $K(z; a, b)$ is the beta density in z with parameters a and b. The expectation of the latter is a continuous mixture of beta distributions over a distribution for their parameters.

However, it is not clear that any continuous mixtures of beta distributions over a and b other than the simple one discussed in this note afford such explicit solution. This and the intriguing link between the uniform mixture of generalized arc-sine distributions and the Cauchy distribution are what make the note stand on its own.

ACKNOWLEDGEMENT

REFERENCES

BERTIN, E. M. J., CUCULESCU, I. AND THEODORESCU, R. (1997). *Unimodality of Probability Measures*, Dordrecht: Kluwer.

CACOULLOS, T. (1965). "A relation between t and F distributions", *Journal of the American Statistical Association*, **60**, 528–531, Corrigenda, **60**, 1249.

CHEN, S. X. (1999). "Beta kernel estimators for density functions", *Computational Statistics and Data Analysis*, **31**, 131–145.

JOHNSON, N. L. (1949). "Systems of frequency curves generated by methods of translation", *Biometrika*, **36**, 149–176.

TADIKAMALLA, P. K. AND JOHNSON, N. L. (1982). "Systems of frequency curves generated by transformations of logistic variables", *Biometrika*, **69**, 461–465.