Uniformly Distributed Between Two Independent Random Variables

Park, Choon il

1. Introduction

The main purpose of this article is to investigate properties of natural *random* mixtures of distributions introduced for the first time (so far as we are aware) by Van Assche (1987) and, incidentally, to demonstrate the usefulness of direct methods of analysis based on calculation of moments. The elementary tools employed are appropriate for classroom use in undergraduate of elementary graduate courses in probability and statistics. The cases studied in this article serve as examples demonstrating that utilization of first principles can sometimes be more advantageous then more sophisticated techniques, such as those employed in Van Assche(1987). The key to our approach is the realization that the class of distributions under consideration can be represented in the form (1) shown in section 2.

2. Definition

To set the stage we briefly review the work of Van Assche (1987). He defined a random variable Z* uniformly distributed between two random variables X_1 and X_2 by the formula

$$
\begin{aligned}
\Pr[Z^* \leq z \mid X_1 = x_1, X_2 = x_2] \\
&= 1 \qquad \text{for } z \ \max(x_1, x_2) \\
&= 0 \qquad \text{for } z \ \min(x_1, x_2) \\
&= \frac{z - x_1}{x_2 - x_1} \qquad \text{for } x_1 \leq z < x_2 \\
&= \frac{z - x_1}{x_2 - x_1} \qquad \text{for } x_2 \leq z < x_1
\end{aligned}
$$

This is equivalent to defining

$$Z^* = \frac{1}{2}(X_1 + X_2) + \left(W - \frac{1}{2}\right)|X_1 - X_2| \qquad (1)$$

With W, independent of X_1 and X_2 ,distributed uniformly over the interval [0,1]. We will consider only cases in which X_1 and X_2 are mutually independent.

If, in addition, X_1 and X_2 have a common distribution, (X_1 -X_2) has a distribution symmetrical about 0, and distribution of Z* is the same as that of

$$
\begin{aligned}
Z^* &= \frac{1}{2}(X_1 + X_2) + \left(W - \frac{1}{2}\right)|X_1 - X_2| \qquad (2a) \\
&= WX_1 + (1 - W)X_2 \qquad (2b)
\end{aligned}
$$

3. Moments

We use the notation μ'_r (Y),(μ'_r (Y)) to denote the *r*th(central) moments of a random variable Y. From (2b), since W, X_1, and X_2 are mutually independent,

$$\mu'_r(Z) = E[Z^r]$$
$$= \sum_{j=0}^{r} \binom{r}{j} E[W^j(1-W)^{r-j}]\,\mu'_j(X)\,\mu'_{r-j}(X), \qquad (3a)$$

Where μ'_r (X) is the *j*th crude moment of the common distribution of X_1 and X_2. An equivalent formula, based on (2a) and using the fact that $E[(W-1/2)^j]=0$ for *j* odd, is

$$\mu'_r(Z) = \sum_{(r-j),\text{even}} \binom{r}{j} E\left[\left(W - \frac{1}{2}\right)^{r-j}\right]$$
$$\times E[2^{-j}(X_1 - X_2)^j (X_1 - X_2)^{r-j}]. \qquad (3b)$$

If We is distributed uniformly over [0,1], then

$$E[W^j(1-W)^{r-j}] = B(j+1, r-j+1)$$
$$= (r-1)^{-1}\binom{r}{j},$$

And (3a) takes the simple form

$$\mu'_r(Z) = (r+1)^{-1}\sum_{j=0}^{r}\mu'_j(X)\,\mu'_{r-j}(X) \qquad (4)$$

$$\left[B(a,b) = \int \omega^{a-1}(1-\omega)^{b-1}d\omega = \Gamma(a)\Gamma(b)/\Gamma(a+b) \text{ is the beta funcion.}\right]$$

Since, for any ξ, W(X_1 + ξ) + (1+ W)(X_1 + ξ) = WX_1 + (1 -W)X_1 + ξ, we can find the central moments of Z, assuming that μ'_r(X) = 0 without loss of generality, so that μ'_r(X)= μ_r (X). With this assumption. We have

$$\mu'_1(Z) = E[Z] = 0 \qquad (5a)$$

$$\text{var}(Z) = \mu_2(Z) = \frac{2}{3}\mu_2(X) = \frac{2}{3}\text{var}(X) \qquad (5b)$$

and (5c)

$$\mu_4(Z) = \frac{1}{5}[2\mu_4(X) + \{\mu_2(X)\}^2],$$

From which we get the measure of kurtosis

$$\beta_2(Z) = \mu_4(Z)/\{\mu_2(Z)\}^2 = \frac{9}{20}(2\beta_2(X)+1). \qquad (5d)$$

As special cases we note (i) X normally distributed ($\beta_2(X)=3$): $\beta_2(Z)=3.15$; (ii) X uniformly distributed ($\beta_2(X)=1.8$): $\beta_2(Z)=2.07$; (iii) x exponentially distributed ($\beta_2(X)=9$): $\beta_2(Z)=4.5$.

Note that for any distribution of X, $\beta_2(X) \geq 1$, so, form (5d), $\beta_2(Z) \geq 1.35$. This value is attained when X is a symmetrical two-point distribution ($\Pr[X=a] = \Pr[X=b]=1/2$, $a \neq b$).

Then Formula (3a) applies for a general W distribution. For example, if W has a symmetrical beta distribution with probability density function

$$\frac{\Gamma(2a)}{\{\Gamma(a)\}^2} w^{a-1}(1-w)^{a-1}, \quad 0 \leq w \leq 1\,; a > a,$$

then $E[W^j(1-W)^{r-j}] = a^{[j]}a^{[r-j]}(2a)^{[r]}$ with $a^{[m]} = a(a+1),\ldots\ldots(a+m-1)$.

From (3a) [taking $\mu'_1(X) = 0$],

$$\mathrm{var}(Z) = 2\,E[W^2]\,\mathrm{var}(X) = \frac{a+1}{2a+1}\mathrm{var}(X). \qquad (6)$$

As a increases from 0 to ∞, var(Z) decreases from var(X) to (1/2) var (X).

If a = 1/2, so that W-distribution is U-shaped, var(Z) = (3/4) var (X), whereas for a = 3/2, with a unimodal W-distribution, var(Z) = (5/8) var (X). In addition,

$$\mu_r(Z) = 2\,E[W^4]\mu_4(X) + 6\,E[W^2(1-W)^2]\{\mu_2(X)\}^2$$
$$= \frac{1}{2(2a+1)(2a+3)}[(a+2)(a+3)\mu_4(X) + 3a(a+1)\{\mu_2(X)\}^2]. \qquad (7)$$

When a =1/2, $\beta_2(Z) = (35/36)\,\beta_2(X) + 1/4$; when a = 3/2, $\beta_2(Z) = (21/25)\,\beta_2(X) + 3/5$.

If the X-distribution is normal ($\beta_2(X)=3$), then for a = 1/2 , $\beta_2(Z)= 3\,{}^1/_6$; for a =3/2, $\beta_2(Z) =$ 3.12.

Note that the U-shaped distribution (a = 1/2) produces the higher value for $\beta_2(Z)$. Both distributions of Z are (slightly) leptokuric.

From (2a), with $(X_1 + X_2)$ and $(X_1 - X_2)$ mutually independent, E[Z] = 0, $\mu_r(Z) = 0$ for r odd, and

$$\mu_r(Z) = \sum_{j=0}^{r/2}\binom{r}{2j} 2^{-2j}[E(X_1+X_2)^{2j}] \times E\left[\left(W - \frac{1}{2}\right)^{r-2j}\right] E[(X_1 - X_2)^{r-2j}], \qquad (8)$$

From which we get

$$\mathrm{var}(Z) = \frac{1}{2} + 2\,\mathrm{var}(X) \qquad (9a)$$

And

$$\mu_r(Z) = \frac{3}{4} + 6\mu_2(W) + 12\mu_4(W). \qquad (9b)$$

Hence

$$\beta_2(X) = 3 + \frac{12\{\beta_2(W) - 1\}}{(1/4)\{\mu_2(W)\}^{-2} + 2\{\mu_2(W)\}^{-1} + 4} \qquad (10)$$

If the support of the distribution of W is [0,1], then $\mu_2(W) \le 1/4$. Noting, in addition , that $\beta_2(W) \ge 1$, we have

$$3 \le \beta_2(Z) \le 3 + \frac{3}{4}(\beta_2(W) - 1\}. \qquad (11)$$

The distribution of Z is always leptokurtic.

If W has a symmetrical triangular [0,1] (tine) distribution with probability density function

$$fw(w) = 4w \qquad 0 \le w \le \frac{1}{2}$$
$$= 4(1-w), \qquad \frac{1}{2} \le w \le 1, \qquad (12)$$

then $\mu_2(W) = 1/24$, $\mu_4(W) = 1/240$ and, from(10), $\beta_2(Z) = 108/35 = 3.086$.

4. Distributions

We suppose that W, X_1, and X_2 are mutually independent and all have uniform [0,1] distributions. In this case, the joint distribution of X_1 and X_2 is uniform over the square with verices (0,0),(0,1), and (1,0). Whe evaluate $F_Z(z) = \Pr[Z \le z]$ as $E_W(\Pr[Z \le z \mid W])$. For $z \le 1/2$,

$$F_Z(z) = \int_0^z \frac{z - (1/2)w}{W} dw + \frac{1}{2} z^2 \int_z^{1-z} \frac{dw}{w(1-w)} + \int_{1-z}^1 \frac{z - (1/2)(1-w)}{w} dw$$
$$= (1-z)^2 \log(1-z) - z^2 \log z + z.$$

Similar calculations lead ot the same result for $z \ge 1/2$.

Hence

$$F_Z(z) = (1-z)^2 \log(1-z) - z^2 \log z + z, \qquad 0 < z < 1. \quad (13a)$$

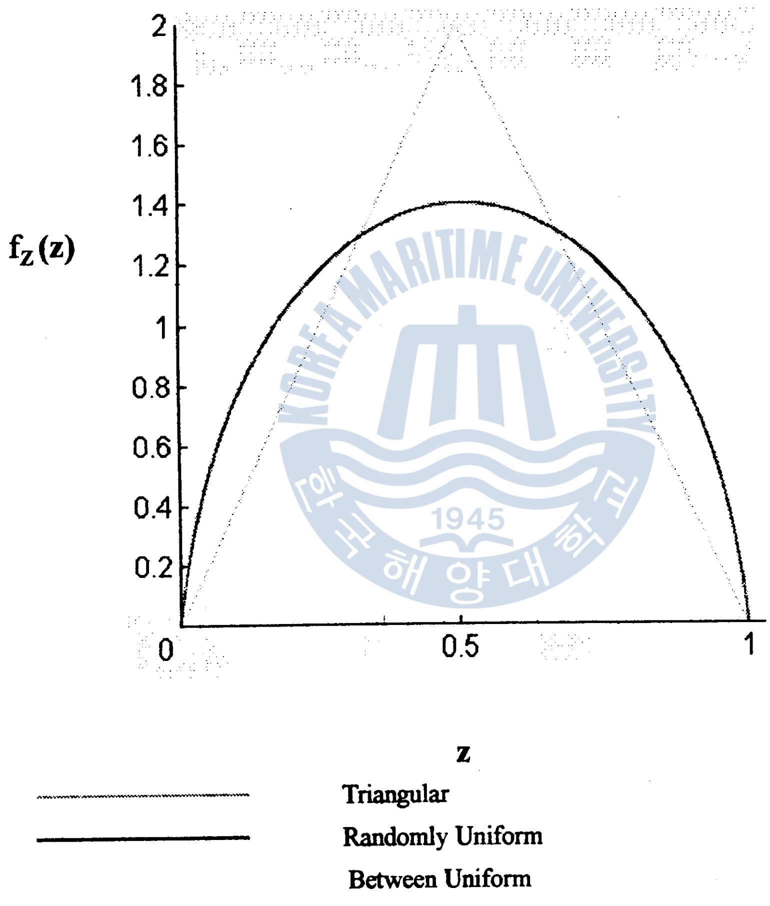
The probability density function of Z is

$$f_Z(z) = F_Z(z) = -2(1-z)\log(1-z) - 2z \log z. \qquad (13a)$$

Figure 1 shows a graph of this function. If W =1/2, the distribution of Z would be a symmetrical triangular("tine") distribution of [0,1], so the distribution (13a and 13b) might be called a "uniformly randomly modified tine" distribution. This distribution would be expected to be "between" the uniform distribution (W= 0 or 1) and the tine distribution (W=12).

Parenthetically, we note that, from (2a), a simple Monte Carlo Procedure, using only simulated uniform variables,

Figure1



We can be used to simulate the distribution (13a and 13b).

Similar calculations for the case in which W has a tine distribution [see (12)] lead to

$$F_Z(z) = 4(1-z)^2 \log(1-z) + 4z^2 \log 2 + 4z - 6z^2 \qquad \text{for } 0 \le z \le \frac{1}{2}$$
$$= -4z^2 \log z - 4(1-z)^2 \log 2 + 3 - 8z + 6z^2 \qquad \text{for } 0 \le z \le \frac{1}{2}$$
$$= 1 - 4z^2 \log z - 4(1-z)^2 \log 2 - 4(1-z) + 6(1-z)^2 \qquad (14a)$$

and

$$f_Z(z) = -8z^* \log z^* - 8(1-\log 2)(1-z^*)$$
$$= -8z^* \log z^* - 2.455(1-z^*), \qquad (14b)$$

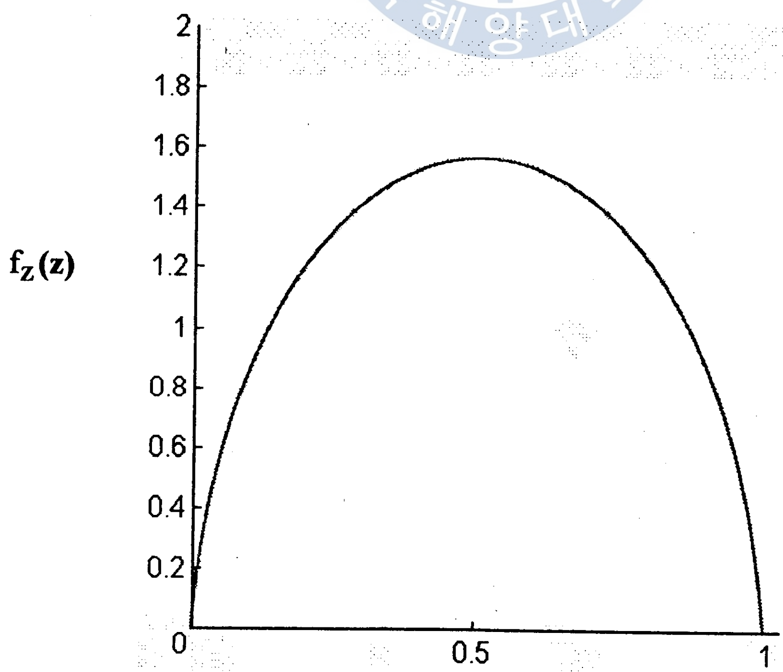
Where $z^* = 1/2 + |1/2 - z|$.

Figure 2 shows a graph of the probability density function. It is similar in shape to that of $f_Z(z)$ in (13b).

A particularly simple case arises if the common distribution of X_1 and X_2 is Cauchy, with probability density function

$$\lambda^{-1}\pi^{-1}\left\{1+\left(\frac{x-\theta}{\lambda}\right)^2\right\}^{-1}, \qquad 0 < \lambda.$$

Figure2



If the distribution of W has support [0,1] [so that neither W nor (1 - W) is negative], the conditional distribution of Z, given W, is also the common distribution of X_1 and X_2.
[See, e.g., Johnson and Kotz (1970,p. 156).] It is also, therefore, the unconditional distribution of Z. this result does not depend on the distribution of W (it need not even be symmetrical), provided only that its support is limited to [0,1].
If X_1 and X_2 have a common standard normal distribution, the conditional distribution of Z, given W, is normal and expected value 0 and variance $\{W^2 + (1 - W)^2\}$. Hence

$$\Pr[Z \le z \mid W] = \Phi\left(\frac{z}{\{W^2 + (1-W)^2\}^{1/2}}\right),$$

where

$$\Phi(y) = (\sqrt{2\pi})^{-1}\int_{-\infty}^{y} e^{(-1/2)u^2}\,\mathrm{du}$$

and

$$\Pr[Z \le z] = \mathrm{E}_W(\Pr[Z \le z \mid W]). \qquad (15)$$

This result is valid for *any* distribution of W. If W has a uniform [0,1] distribution, then

$$\Pr[Z \le z] = \int_0^1 \Phi\left\{\frac{z}{\{W^2 + (1-W)^2\}^{1/2}}\right\} dw. \qquad (16)$$

This can be evaluated by quadrature.

5. Characterization

Equation (3a) can be written in the form

$$\mu_r'(\mathrm{Z}) = 2\mu_r'(\mathrm{W})\mu_r'(\mathrm{X}) + (\text{function of } \mu_1'(\mathrm{X}),\ldots,\mu_{r-1}'(\mathrm{X})). \qquad (17)$$

If the moments of W are known, those of X can be derived from those of Z by using(17) with r =1,2,... successively. If, further, the distribution of X is determined by its moments, then the distribution of X is characterized by that of Z. This is so, in particular, if the support of Xis bounded. This result holds for W having *any* distribution symmetrical about 1/2 with finite moments of all orders.

6. Conclusion

We have described how moment methods can provide a way of obtaining distributions and characterizations of distribution of random mixture variables of form

$$\mathrm{Z} = \mathrm{W}\mathrm{X}_1 + (1 - \mathrm{W})\mathrm{X}_2$$

or, more generally,

$$\mathrm{Z} = \sum_{j=1}^{n} \mathrm{W}_j \mathrm{X}_j,$$

where the X's are mutually independent and have a common distribution and the W's are independent of the X's. The case in which the X's have standard uniform (0,1) distributions leads to an interesting family of symmetric distributions.

The method can be extended straightforwardly to distributions of variables of type

$$Y = \sum_{j=0}^{\infty} (-1)^j \prod_{i=0}^{j} X_i$$

References

Gleser,L,J. (1989), "The Gamma Distribution as a Mixture of Exponential Distributions," *The American Statistician*, 43, 115-117.

Greenwood, M., and Yule, G. U. (1920), "An Enquiry into the Nature of Frequency Distributions of Multiple Happenings With Particular Reference to Occurrence of Multiple Attacks of Disease of Repeated Accidents," *journal of the Royal Statistical Society,* Ser. A, 83, 255-279.

Johnson, N. L., and Kotz, S. (1970), *Distribution in Statistics: Continuous Univariate Distributions-1,* New York: John Wiley.

Kennedy, D. P. (1988), "A Note on stochastic Search Methods for Global Optimization," *Advances in Applied Probability,* 20, 476-478.

Schmittlein, D. C. (1989), "Surprising Inferences From Unsurprising Observations: Do Conditional Expectations Really Regress to the Mean?," *The American Statistician,* 43 176-183.

Titterington, D. M., Smith, A. F. M., and Makov, U. E. (1985), *Statisical Analysis of Finite Mixture Distributions,* New York: John Wiley.

Tukey, J. W. (1960), "A Survey of Sampling From Contaminated Populations," in *Contributions to Probability and Statistics,* ed. I. Olkin, Stanford, CA: Stanford University Press, PP. 445-485.

Van Assche, W. (1986), "Products of 2 × 2 Stochastic Matrices With Random Entries," *journal of Applied Probability,* 23, 1019-1024.

——— (1987), "A Random Variable Uniformly Distributed Between Two Random Variables," *sankhyā*, Ser. A, 49, 207-211.

Department Applied Mathematics,
Korea Maritime University,
Pusan 606-791, Korea