J. Basic. Appl. Sci. Res., 2(8)8401-8407, 2012
© 2012, TextRoad Publication

ISSN 2090-4304
Journal of Basic and Applied Scientific Research
www.textroad.com

Randomly Weighted Averages as: TSP Generation

Monireh Hamel Darbandi[1], Hajir Homei[2] and Aziz Hamel Darbandi[3]

[1,2] Department of Statistics, Faculty of Mathematical Sciences, University of Tabriz, Tabriz 5166617766, Iran.
[3] Department of Statistics, Faculty of Science, Ankara University, Ankara, Turkey

ABSTRACT

A new rich class of generalized two-sided power (TSP) distributions, where their density functions are expressed in terms of the Gauss hypergeometric functions, is introduced and studied. In this class, the symmetric distributions are supported by finite intervals and have normal shape densities. Our study on TSP distributions also leads us to a new class of distributions.

KEYWORDS: and Phrases: Randomly weighted averages, Two-sided power distribution.

1 INTRODUCTION

The paper of Nadarajah (1999) initiated research work on distribution functions supported by finite intervals, say (0,1), that assume different formulations on subintervals of their supports, say (0, θ), (θ, 1). Two-sided power distributions (TSP) are of this type and were introduced by van Dorp and Kotz (2002a) as underlying statistical distributions for certain monthly interest rates.The potential, flexibility and applicability of TSP distributions in applied fields have been examined and explored in a series of papers by van Dorp and Kotz, (2002a, 2002b, 2003), Nadarajah (2005) and Perez et al. (2005), among others.
In this articles we show how a TSP random variables can be deduced from a bivariate Dirichlet random vector. Then we derive equivalent form for the kth moment of the TSP distributions about zero. According, a weighted average of the first and the last order statistics of a uniform [0,1] random sample has a TSP distribution. This naturally leads to a new generalization of TSP distribution (HTSP). Our HTSP distributions exhibit interesting features, and overcome some deficiencies of the TSP distributions. Our symmetric HTSP distribution form a rich class of symmetric distributions on finite intervals with normal shape densities. Also, we introduced other new family with named N-sided power distribution.

2 Conditional distribution for random weighted averages

In this section we provide the conditional distribution of $S_{n^*}=\sum_{j=1}^{n} R_{kj} X_j$ for given $(X_1,\dots,X_n)=(x_1,\dots,x_n)$ at z, denoted by $k(z|x_1,\dots,x_n)$. At first we assume $x_1>x_2>\cdots>x_n>0$, but later we will remove this restriction on $x_1,\dots,x_n$. We recall that $(U_{(1)},\dots,U_{(n)})$ is the order statistics of a random sample $U_1,\dots,U_n$ for the uniform $[0,1]$. For the sequence of indices $\{k_1,\dots,k_{n-1}\}$ is an ordered sequence in $\{1,\dots,n^*-1\}$ Thus $\{U_{k_1},..,U_{k_{n-1}}\}\subset\{U_{(1)},\dots,U_{(n^*-1)}\}$ the increments R_{k_j} are defined by $R_{k_j}=U_{k_j}-U_{k_{j-1}}$, $j=1,\dots,n-1$, where $U_{k_0}=0$ and $U_{k_n}=1$. Since can be $\sum_{j=1}^{n} R_{k_j}$, $k(z|x_1,\dots,x_n)$ expressed as

$$P(\sum_{j=1}^{n-1} c_j R_{kj} \le z-x_n),\quad c_j=x_j-x_n,\quad j=1,\dots,n-1,$$

the distribution $\sum_{j=1}^{n-1} c_j R_{kj}$ was derived by Weisberg (1971):

$$P(\sum_{j=1}^{n-1} c_j R_{kj} \le z-x_n)=1-\sum_{j=1}^{r}\frac{h_j^{m_j-1}(x_j;z)}{(m_j-1)},\qquad (2.1)$$

where $m_j=k_j-k_{j-1},\ \ j=1,\dots,n,\ \ k_n=n^*,\ \ \sum_{j=1}^{n} m_j=n^*,\ \ h_j^{(m_j-1)}(c_j)$ is the $(m_j-1)-$th derivative of

***Corresponding Author:** Monireh Hamel Darbandi, Department of Statistics, Faculty of Mathematical Sciences, University of Tabriz, Tabriz 5166617766, Iran. Email: darbandi89@ms.tabrizu.ac.ir, homei@tabrizu.ac.ir

$$h_j(x;z)=\frac{(x-z)^{n^*-1}}{c_j\prod_{i\neq j}^{n}(x-x_i)^{m_i-1}}$$

at x evaluated at $x_1,\dots,x_n$, where r is the largest positive integer such that $z < x_r$.

The distribution of $\sum_{j=1}^{n-1}c_jR_{kj}$ in equation (2.1), alternatively can be expressed as

$$P(\sum_{j=1}^{n-1}c_jR_{kj}\le z-x_n)=\sum_{j=r^*+1}^{n}\frac{f_j^{(m_{j-1})}(x_j;z)}{(m_j-1)!},\qquad (2.2)$$

where r^* is the largest positive integer such that $x_{r^*}\ge z$, and f_j $(x_j;z)$ is the $(m_j-1)-th$ derivatives of

$$f_j(x;z)=\frac{(x-z)^{n^*-1}}{\prod_{i\neq j}^{n}(x-x_i)^{m_i}},$$

at $x = x_j$.

By using the Heaviside function

$$U(x)=\begin{cases}0 & x<0\\ 1 & x>0\end{cases}$$

the distribution in equation (2.2) can be expressed as

$$k(z\mid x_1,\dots,x_n)=\sum_{j=1}^{n}\frac{f_j^{(m_j-1)}(x_j;z)U(z-x_j)}{(m_j-1)},$$

for any set of distinct values $x_1,\dots,x_n$, and any $z\in[min(x_1,\dots,x_n),max(x_1,\dots,x_n)]$. The conditional

$$k(z\mid x_1,\dots,x_n)=\begin{cases}0 & z<\min(x_1,\dots,x_n)\\ 1 & z\ge\min(x_1,\dots,x_n),\end{cases}$$

$k(z;x_1,\dots,x_n)$ is a big family of distribution that includes two-sided power (TSP) distributions and hyper two-sided power (HTSP) distributions.

3 TSP Random Variables as weighted averages

We recall from Van Dorp and Kotz (2002) that a TSP random variable $Z\sim TSP(a,m,b,n)$ is defined on an interval (a,b) in the real line with the probability density function (p.d.f.) $f(z|a,m,b,n)$ given by

$$k(z\mid a,m,b,n)=\begin{cases}\frac{n}{b-a}\left(\frac{z-a}{m-a}\right)^{n-1}, & a<z<m,\\ \frac{n}{b-a}\left(\frac{b-z}{b-m}\right)^{n-1}, & m<z<b.\end{cases}$$

A TSP random variable with $a = 0$, $b = 1$ is called standard and is denoted by $X\sim STSP(\theta,n)$, the parameter m in this case is denoted by θ. It can be readily verified that $Z = (b-a)X + a$.

Theorem 3.1. Assume that (W,V) is a bivariate standard Dirichlet random vector with parameters $\propto= 1$, $\beta = n-1$, $\gamma = 1$; $n > 1$, $i.e.$ (W,V) possesses the joint density function

$$f_{W,V}(w,v)=n(n-1)v^{n-2}I_{[0,1]}(w)I_{[0,1]}(v)I_{[0,1]}(w+v).$$

Then for every θ, $0 < \theta < 1$, $X = W + \theta V$ is a STSP random variable with parameters θ,n.

Proof. The joint density function of $X = W + \theta V$ and W is given by

$$f_{X,W}(x,w) = \frac{n(n-1)}{\theta}(\frac{x-w}{\theta})^{n-2} I_{[0,1]}(w) I_{[w,w+\theta(1-w)]}(x)$$

Thus

$$f_X(x) = \int_0^x \frac{n(n-1)}{\theta}(\frac{x-w}{\theta})^{n-2} dw, \quad 0 < x < \theta,$$

And

$$f_X(x) = \int_{\frac{x-\theta}{1-\theta}}^{x} \frac{n(n-1)}{\theta}(\frac{x-w}{\theta})^{n-2} dw, \quad \theta < x < 1,$$

Giving

$$f_x(x) = n(\frac{x}{\theta})^{n-1} I_{(0,\theta)}(x) + n(\frac{1-x}{1-\theta})^{n-1} I_{(0,\theta)}(x)$$

which is the p.d.f. of a STSP random variable with parameters θ, n. The proof of the theorem is complete.
Remark 3.1. It follow from Theorem 3.1 that X is for an integer $n \geq 1$, $X \sim STSP(\theta, n)$ if

$$X \overset{d}{=} U_{(1)} + \theta(U_{(n)} - U_{(1)}), \quad 0 < \theta < 1$$

where $U_{(1)} = min(U_1, \dots, U_n)$, $U_{(n)} = max(U_1, \dots, U_n)$; and $U_1, \dots, U_n$ are i.i.d uniform (0, 1).
Theorem 3.2. Let $Z \sim STSP(\theta, n)$, $n > 1$, $\theta \in [0,1]$. Then for any integer $k \geq 0$,

$$EX^k = \frac{\Gamma(n+1)}{n+k+1} \sum_{i=0}^{k} \binom{k}{i} \frac{\Gamma(i+1)}{\Gamma(n+i)} \theta^{k-i} (1-\theta)^i,$$

Proof. By using Theorem 3.1, we obtain that

$$EX^k = E(W + \theta V)^k = \sum_{i=0}^{k} (1-\theta)^i \theta^{k-i} E\left[W^i (V+W)^{k-i}\right].$$

The result will follow from Theorem 3.1.

4 GTSP random variables

According, a weighted average of the first and the last order statistics of a uniform [0, 1] random sample has a TSP distribution (Remark 2.1.). This naturally leads to a new generalization of TSP distribution (HTSP).

Definition 4.1. A random variable $X = W + \theta V$, where (W, V) possesses the joint p.d.f. given in equation (4.1) called standard generalized TSP random variable with parameters θ, n, k_1, k_2; $X \sim STSP(\theta, n, k_1, k_2)$. Note that $0 < \theta < 1, n, k_1, k_2$ are real numbers subject to $0 < k_1 < k_2 < n + 1$.

$$f(w,v) = \frac{\Gamma(n+1)}{\Gamma(k_1)\Gamma(k_2-k_1)\Gamma(n-k_2+1)} w^{k_1-1} v^{k_2-k_1-1} (1-w-v)^{n-k_2}. \qquad (4.1)$$

The p.d.f of X can be expressed in terms of certain Gauss hypergeometric function. Indeed the joint p.d.f. of (W, X) is given by

$$f_{W,X}(w,x) = \kappa w^{k_1-1} (x-w)^{k_2-k_1-1} (\theta - x + w(1-\theta))^{n-k_2}, \qquad (4.2)$$

where $0 < w < 1$ and $w \leq x \leq w + \theta(1-w)$, and

$$\kappa = \frac{\Gamma(n+1)}{\Gamma(k_1)\Gamma(k_2-k_1)\Gamma(n-k_2+1)\theta^{n-k_1}}.$$

Therefore,

$$f(x) = \begin{cases} \int_0^x f(w,x)dw, & 0 < x < \theta \\ \int_{\frac{(x-\theta)}{(1-\theta)}}^x f(w,x)dw, & \theta < x < 1 \end{cases} \quad (4.3)$$

The density function $f(x)$ can be expressed in terms of the Gauss hypergeometric function $F(a,b,c;z)$, which is a well-known special function. Indeed according to the Euler's formula, the Gauss hypergeometric function assumes the integral representation

$$F(a,b,c;z) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tz)^{-a}\,dt$$

where a,b,c are parameters subject to $-\infty < a < +\infty$, $c > b > 0$, whenever they are real and z is the variable. By using the Euler's formula, the density function in equation (4.3) can be expressed as follows. For $0 < x < \theta$, $f(x)$ is giving by

$$k_1(\theta)^{k_1-1}(\frac{x}{\theta})^{k_2-1}(1-\frac{x}{\theta})^{n-k_2}F(k_2-n,k_1,k_2;-\frac{x(1-\theta)}{\theta-x}); \quad (4.4)$$

For $\theta < x < 1$, $f(x)$ is giving by

$$k_1(1-\theta)^{n-k_2}(\frac{1-x}{1-\theta})^{n-k_1}(1-\frac{1-x}{1-\theta})^{k_1-1}F(1-k_1,n-k_2+1,n-k_1+1;-\frac{\theta(1-x)}{x-\theta}); \quad (4.5)$$

where

$$\kappa_i = \frac{\Gamma(n+1)}{\Gamma(k_i)\Gamma(n-k_i+1)}, \quad i = 1,2.$$

The density f at θ exits if $k_2 < n + k_1$ only. Indeed,

$$f(\theta) = \frac{n\Gamma(n+k_1-k_2)}{\Gamma(k_1)\Gamma(n-k_2+1)}\theta^{k_1-1}(1-\theta)^{n-k_2}, \quad k_2 < n+k_1 \quad (4.6)$$

If $k_1 = 1, k_2 = n$, then $F(0,1,n,z) = 1$, and the density function $f(x)$ givin by Equation (4.4) and (4.5) readily reduces to the STSP density function given in section 2.

Remark 4.1. For the case that k_1, k_2, and n are integers, a HTSP random variable can be expressed as

$$Z = bU_{(k_1)} + m(U_{(k_2)} - U_{(k_1)}) + a(1-U_{(k2)}), \quad (4.7)$$

Where $m = (b-a)\theta + a$, and $U_{(1)},\dots,U_{(n)}$ is the order statistic of n i.i.d uniform (0,1) random variables. Note equation (4.7) for the case that $k_1 = 1, k_2 = n$ gives a TSP random variable.

Remark 4.2 As we stated in Remark 4.1, a TSP random variable Z, for integer n, is the random weighted average of b, m, and a, $[b > m > a]$, with random variables $1 - U_{(1)}, U_{(n)} - U_{(1)}$ and $1 - U_{(n)}$, respectively. The intermediate point m receives more mass than the ending points b,a that statistically receive equal masses. For large n, the intermediate point receives substantially larger weight than the ending points; and the density is too tall at θ, $f(\theta) = n$. The HTSP distribution overcome this illusion. By putting more random weights on the ending points, $f(\theta)$ will be of reasonable size, even for large n. Indeed it is evident from equation (4.2) that $f(\theta) \to 0$, as $n \to \infty$ for fixed k_1, k_2.
Let us record that if $X \sim SHTSP(\theta, n, k_1, k_2)$, then

$$E(X) = \frac{k_1 + \theta(k_2 - k_1)}{n+1}$$

and

$$V(X) = \frac{k_1(n-k_1+1)(1-\theta)^2 + k_2(n-k_2-+1)\theta^2 + 2k_1((n-k_2-+1)\theta(1-\theta)}{(n+1)^2(n+2)}$$

In the special case that $k_1 = 1, k_2 = n$, it gives the variance for STPD derived by Van Dorp and Kotz (2002a).

Let us pay a special attention to the symmetric SHTSP distributions and random variables. Details are given in the following theorem.

Theorem 4.1. Let $X \sim SHTSP(\theta, n, k_1, k_2)$, $0 < k_1 < k_2$, $k_2 < \min\{n+1, n+k_1\}$, then X_2 is symmetric if and only if $\theta = \frac{1}{2}$ and $k_2 = n - k_1 + 1$.

Proof. The "if" part is straight forward. For the "only if" part, we notice that for $x = \theta$, from equation (4.6), $f(\theta) = f(1-\theta)$ implies that $k_2 = n - k_1 + 1$. From the symmetry $E(X) = \frac{1}{2}$, which gives $k_1(1-\theta) + k_2 = \frac{(n+1)}{2}$. But for a given $k_1 < \frac{(n+1)}{2}$, the line $k_1(1-\theta) + (n - k_1 + 1)\theta$ joins the points $(0, n - k_1 + 1)$ and $(1, k_1)$ as θ varies from 0 to 1. Hence it intersects the line $\frac{(n+1)}{2}$ at exactly one point attained at $\theta = \frac{1}{2}$. The proof of the theorem is complete.

If X is a symmetric SHTSP random variable with parameters (n, k) then it follows from Theorem 4.1 that $\frac{1}{2} < k < \frac{n+1}{2}$. density function of X can be deduced from equations (4.4) and (4.5), namely,

$$f(x) = \begin{cases} 2^{n-2k+1}\kappa_3 x^{n-k}(1-2x)^{k-1} G\left(-\dfrac{x}{1-2x}\right), & 0 < x < \dfrac{1}{2} \\ 2^{n-2k+1}\kappa_3 x^{n-k}(2x-1)^{n-k} G\left(-\dfrac{1-x}{2x-1}\right), & \dfrac{1}{2} < x < 1, \end{cases}$$

Where $G(z) = F(1-k, k, n-k+1; z)$, $\kappa_3 = \dfrac{\Gamma(n+1)}{\Gamma(k)\Gamma(n-k+1)}$. Also

$$f(1/2) = n\frac{\Gamma(2k-1)}{[\Gamma(k)]^2}(1/4)^{k-1}.$$

Clearly $f(x) = f(1-x)$, $0 < x < 1$, as expected. Let us also record from equation (4.8) that

$$E(x) = \frac{1}{2}, \qquad V(x) = (1/2)\left[\frac{k}{(n+1)(n+2)}\right].$$

5 N-Sided Power Distribution

Let X be a random variable with cumulative distribution function given by

$$f(x \mid \theta_1, \ldots, \theta_n) = \begin{cases} \displaystyle\sum_{j=0}^{i} c_j (x - \theta_{n-j})^{n-1}, \\ \theta_{n-j} < x \le \theta_{n-i-1}, \quad i = 0, \ldots, n-2 \end{cases}$$

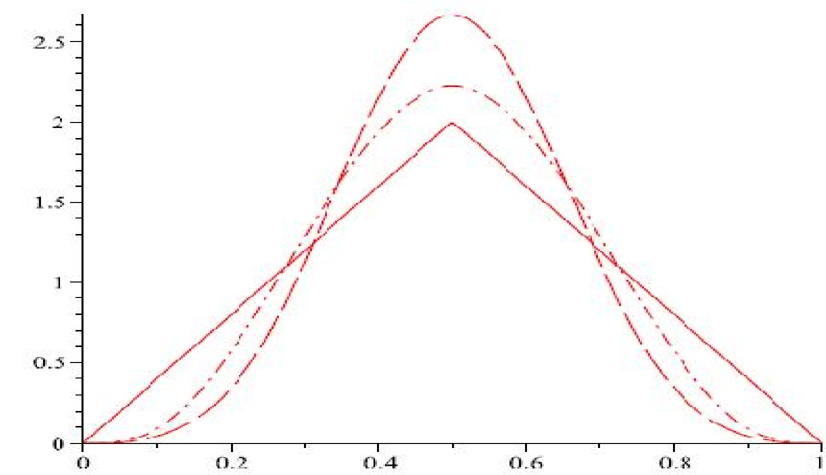

Figure1:Symmetric NSP(θ1,θ2,θ3,θ4,θ5) distribution:—·—·— , θ1=1,θ2=0.9 θ3=0.5,θ4=0.1,θ5=0 ; ——— , θ1=1,θ2=0.75,θ3=0.5,θ4=0.25,θ5=0

Where

$$c_j = \left[\prod_{i=1,\ i \neq n-j}^{n} \left(\theta_{n-j} - \theta_i \right) \right]^{-1}, \quad j = 0, \ldots, n-1$$

X will be said to follow a N-sided power distribution, $NSP(\theta_1, \ldots, \theta_n), \theta_1 > \theta_2 > \cdots > \theta_n, n > 0$, where n is an integer. The density of (5.1) is unimodal with the mode at $\overline{\theta}$.

For $n = 3$, $F(.|\theta_1, \ldots, \theta_n)$ simplifies to a TSP distribution. Figure 1 provides examples of symmetric $NSP(\theta_1, \ldots, \theta_n)$ distributions i.e. $\overline{\theta}$ =0.5, including triangular distribution. Figure 2 presents examples of positively and negatively skewed $NSP(\theta_1, \ldots, \theta_n)$ distributions, including examples of triangular distributions.

The density function of a $NSP(\theta_1, \theta_2, \theta_3)$ distribution follows from expression (5.1) as

$$f(z \mid \theta_1, \theta_2, \theta_3) = \begin{cases} \dfrac{2(x - \theta_3)}{\theta_{13}\theta_{23}}, & \theta_3 < x \leq \theta_2 \\ \dfrac{2(\theta_1 - x)}{\theta_{13}\theta_{12}}, & \theta_2 < x \leq \theta_1 \end{cases}$$

$f(x|\theta_1, \theta_2, \theta_3)$ is two sided power distribution.

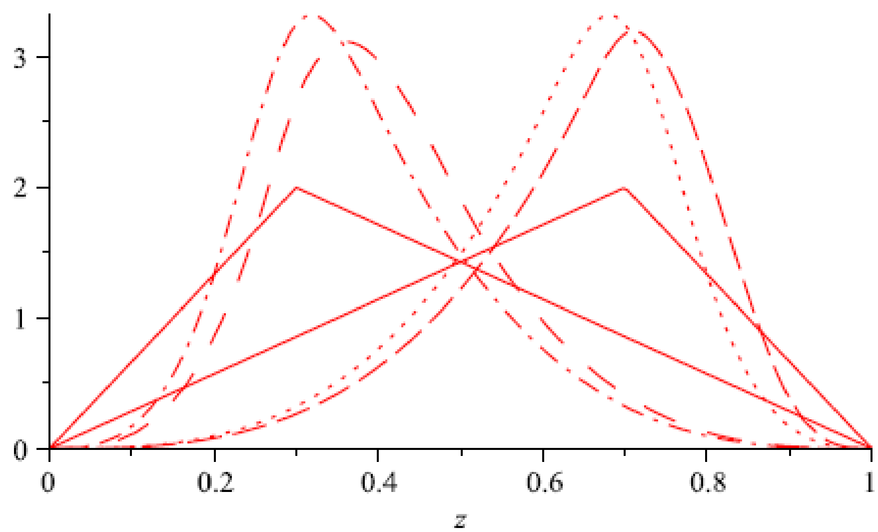

Figure 2:Positively skewed NSP(θ1,θ2,θ3,θ4,θ5) distributions: ···· , θ1=1,θ2=0.8,θ3=0.7,θ4=0.6,θ5=0 ——— , θ1=1,θ2=0.9 ,θ3=0.7 ,θ 4=0.65,θ5=0 and negatively skewed NSP(θ1,θ2,θ3, θ4,θ5) distrbutions:-·-·-· ,θ1=1,θ2=0.4,θ3=0.3,θ4=0.2,θ5=0,- - - , θ1=1, θ2=0.5,θ3=0.3,θ4=0.25,θ5=0

the expressions for the mean and the variance can be obtained from expression (5.1) and

$$E(X) = \overline{\theta}$$

$$Var(X) = \frac{S_\theta^2}{n+1}$$

The meaning of the parameters is as follows: θ_1, θ_n are the end points of the support, n is the shape parameter and $\theta_2, \dots, \theta_{n-1}$ are the threshold parameters for a change in the form of the probability density function.
In the following section, more classical estimation procedure for the NSP $(\theta_1, \theta_2, \theta_3)$ distribution are discussed using data.

REFERENCES

[1] Nadarajah, S.(2005). On the two-sided power distribution. Metrika, 61, 309-321.

[2] Nadarajah, S. (1999). A polynomial model for bivariate extreme value distributions. Statist. Probab. Lett., 42, 15-25.

[3] Van Dorp, J.R., Kotz, S. (2003). Generalized trapezoidal distributions. Metrika, 58 (1), 85-97. [4] van Dorp J.R., Kotz, S. (2002a). The standard two-sided power distribution and its properties: with applications in financial engineering. The Amer. Statist., 56, 9099.

[5] van Dorp J.R., Kotz, S. (2002b). A novel extension of the triangular distribution and its parameter estimation. The Statistician 51, 63-79.

[6] Perez, J.G., Rambaud, S.C., Garcia, L.G. (2005). The two-sided power distribution for the treatment of the uncertainty in PERT. Stat. Methods Appl., 14, 209-222.

[7] Weisberg, H. (1971). The distribution of linear combinations of order statistics from the uni- form distribution. Ann. Math. Statist., 42, 704-709.