The distribution of Nadarajah and Kotz revisited

Hazhir Homei[†, 1], **Manizheh Jalilvand**[†]

[†] Department of Mathematics, University of Tabriz, Iran,

[‡] Department of Mathematics, University of Tabriz, Iran.

Abstract

A new model for real lifetimes is proposed, here That can be used in topics such as vehicle speed, asphalt, etc. The distributional properties of this model are discussed, also the Nadarajah and Kotz distributions are generalized by using it.

Keywords: Real Lifetime, Product, Deferential Equation

Mathematics Subject Classification [2010]: 15A03, 15A23, 15B36

1 Introduction

In statistics, the generalized distributions play a very important role in the lifetime. In this paper, we use the generalized distribution of the Nadarajah and Kotz (2005) if it is possible, otherwise, we suggest the approximation of Nadarajah (2006a,2006b) by using MLE. We will answer the questions posed at the conclusion of two separate articles about the lifetime in Homei and Nadarajah (2018) and Hadad et al. (2021) and generalize some of the results obtained by solving some differential equations.

2 The distribution of a real lifetime and some properties

The product of random variables has found many interesting applications theoretically. There are various examples of random variables in the literature that their products are analyzed theoretically and practically (see section 2 in Adamska et al. 2022) of which are being reviewed in this section and generalized to the multivariate cases; see Nadarajah and Kotz (2005).

Theorem 2.1. *Let the random vector X, effective coefficient, and the random variable Y, lifetime in the laboratory, be with $CE(\alpha_1,\ldots,\alpha_r)$ and $L(\alpha,\beta)$ distributions. Then the distribution of the real lifetime, $\mathbf{Z}=\mathbf{X}Y$, can be expressed by:*

$$f(z_1,\ldots,z_r)=\frac{\beta^{-\alpha}\Gamma(\sum_{i=1}^r\alpha_i)}{\Gamma(\alpha)\prod_{i=1}^r\Gamma(\alpha_i)}(\sum_{i=1}^r z_i)^{\alpha-\sum_{i=1}^r\alpha_i}exp\{-\frac{\sum_{i=1}^r z_i}{\beta}\}\prod_{i=1}^r z_i^{\alpha_i-1}. \tag{1}$$

Where $z_i>0,\ i=0,\cdots,r$.

Proof. The distribution of $\mathbf{Z}$ can be found by using classical methods of transformations. Thus we will have

$$J=(\frac{1}{\sum_{i=1}^r z_i})^{r-1},\quad y=\sum_{i=1}^r z_i,\quad x_i=\frac{z_i}{\sum_i z_i} \tag{2}$$

We also know that the probability density function $\mathbf{Z}=(z_1,\ldots,z_r)$ is as follows:

$$f_{\mathbf{Z}}(z_1,\ldots,z_r)=f_Y(\sum_{i=1}^r z_i)f_X(\frac{z_i}{\sum_1 z_i},\ldots,\frac{z_r}{\sum_i z_i})|J| \tag{3}$$

$$=\frac{1}{\beta^\alpha\Gamma(\alpha)}(\sum_{i=1}^r z_i)^{\alpha-1}e^{-\frac{\sum_{i=1}^r z_i}{\beta}}\frac{\Gamma(\sum_{i=1}^r\alpha)}{\prod_{i=1}^r\Gamma(\alpha_i)}\prod_{i=1}^r(\frac{z_i}{\sum_{i=1}^r z_i})^{\alpha_i-1}(\frac{1}{\sum_{i=1}^r z_i})^{r-1}, \tag{4}$$

the proof is completed. □

The Nadarajah and Kotz (2005) distribution will be obtained exactly for $r=2$ and thus the distribution of Z_1 should be calculated (marginal distribution). Theorem 2.2 shows the distribution of the product of the random variables Gamma and Dirichlet random vector. Moreover, if $\alpha=\sum\alpha_i$ then the distribution of $\mathbf{Z}$ in (2) shows that $Z_1,\ldots,Z_k$ are independent with Gamma distribution.

Theorem 2.2. *If $\mathbf{X}\sim Ga*D(\alpha,\beta,\alpha_1,\cdots,\alpha_n)$ then we have*

$$E(Z_i)=\frac{\alpha_i}{\sum_{i=1}^n\alpha_i}\alpha\beta$$

and

$$\Sigma=\begin{pmatrix}\sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n}\\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n}\\ \vdots & \vdots & & \vdots\\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2\end{pmatrix}$$

Where Σ denotes the covariance matrix of Z_j's and

$$\sigma_j=\frac{\beta^\alpha\alpha(\alpha+1)\alpha_j(\alpha_j+1)}{(\sum_{i=1}^n\alpha_i)(\sum_{i=1}^n\alpha_i+1)}-\left(\frac{\alpha_j\alpha\beta}{\sum_{i=1}^n\alpha_i}\right)^2$$

$$\sigma_{jk}=\frac{\alpha\alpha_j\alpha_k}{(\alpha+\beta)^2(\sum_{i=1}^n\alpha_i)(\sum_{i=1}^n\alpha_i+1)}(\frac{\beta}{\alpha+\beta+1}-\frac{\alpha}{\sum_{i=1}^n\alpha_i}).$$

3 Lifetime in the independent r-position

By the most important properties of the distribution quoted in Nadarajah et al (2022), the following theorems are stated.

Theorem 3.1. *Let $X_1,\cdots,X_r$ be the independent random vectors, effective coefficient, with $CE(n_{11},\cdots,n_{1k}),\cdots,CE(n_{r1},\cdots,n_{rk})$ distributions respectively and that the random variable Y_j, $j=1,2,\cdots,r$, lifetime in the laboratory, is independent from $\mathbf{X}_1,\ldots,\mathbf{X}_r$ with $L(\alpha_j,1)$, $j=1,2,\cdots,r$, distribution. Then the product moments in $(s_1,\ldots,s_k)$ of the real lifetime in the independent r-position, $\boldsymbol{T}=\sum_{j=1}^r Y_j\mathbf{X}_j$, are*

$$E(T_1^{s_1}T_2^{s_2}\ldots T_k^{s_k})=\frac{\Gamma(\alpha+s)}{\Gamma(\alpha+h)}\sum_{h_1}\cdots\sum_{h_k}\left(\prod_{j=1}^k\binom{s_j}{h_{1j}\ldots h_{rj}}\right)\times\prod_{i=1}^r\frac{\Gamma(\alpha_i+h_i)}{\Gamma(\alpha_i)}$$

$$\frac{\Gamma(n_i)}{\Gamma(n_i+h_i)}\prod_{i=1}^r\prod_{j=1}^k\frac{\Gamma(n_{ij}+h_{ij})}{\Gamma(n_{ij})}),$$

where T_j's are the components of vector T, $\sum_{i=1}^r h_i=h$, $\sum_{i=1}^r\alpha_i=\alpha$ and $\sum_{i=1}^r s_i=s$.

Proof.

$$\begin{aligned}E(T_1^{s_1}T_2^{s_2}\ldots T_k^{s_k})&=E\left(\prod_{j=1}^k(\sum_{i=1}^r Y_j\mathbf{X}_{ij})^{s_j}\right)\\&=E\left(\prod_{j=1}^k(\sum_{j=1}^r Y_j\frac{\sum_{i=1}^r Y_j}{\sum_{j=1}^r Y_j}\mathbf{X}_{ij})^{s_j}\right)\\&=E\left(\prod_{j=1}^k(\sum_{j=1}^r Y_j\sum_{i=1}^r\frac{Y_j}{\sum_{j=1}^r Y_j}\mathbf{X}_{ij})^{s_j}\right)\\&=E\left((\sum_{j=1}^r Y_j)^{\sum_{j=1}^r s_j}\right)\left(\prod_{j=1}^k(\sum_{i=1}^r\frac{Y_j}{\sum_{j=1}^r Y_j}\mathbf{X}_{ij})^{s_j}\right)\\&=E\left((\sum_{j=1}^r Y_j)^s\right)E\left(\prod_{j=1}^k(\sum_{i=1}^r\frac{Y_j}{\sum_{j=1}^r Y_j}\mathbf{X}_{ij})^{s_j}\right)\end{aligned}$$

The previously mentioned mathematical expectations can be easily calculated separately.

$$E(L_1^{s_1}L_2^{s_2}\ldots L_k^{s_k})=\frac{\Gamma(\alpha)}{\Gamma(\alpha+h)}\sum_{h_1}\cdots\sum_{h_k}\left(\prod_{j=1}^k\binom{s_j}{h_{1j}\ldots h_{rj}}\right)\times\prod_{i=1}^r\frac{\Gamma(\alpha_i+h_i)}{\Gamma(\alpha_i)}$$

where L_j's are components of vector $Z(=\sum_{i=1}^n\frac{Y_i}{\sum_{i=1}^n Y_i}\mathbf{X}_i)$

The result is obtained by placing the moments in the above expressions. □

The most important properties of any distribution are the first moment and variance.

Theorem 3.2. *Suppose $\boldsymbol{X}_1,\ldots,\boldsymbol{X}_r$ are some independent effective coefficients with $CE(n_{11},\ldots,n_{1k}),\ldots,CE(n_{r1},\ldots,n_{rk})$ distributions and $Y_1,\ldots,Y_r$ are some independent lifetime in the laboratory with $L(n,\frac{1}{\theta})$ distributions and also independent from $\boldsymbol{X}_i$'s. Then the distribution of the real lifetime $\bar{T}=\frac{\sum_{i=1}^r\boldsymbol{X}_iY_i}{r}$ equals to*

$$f(z_1,\ldots,z_r)=\frac{r^{rn-r+k}\theta^{rn}\Gamma(\sum_{i=1}^r\alpha_i)}{\Gamma(rn)\prod_{i=1}^r\Gamma(\alpha_i)}(\sum_{i=1}^r t_i)^{rn-\sum_{i=1}^r\alpha_i}e^{-\theta\sum_{i=1}^r rt_i}\prod_{i=1}^r t_i^{\alpha_i-1}. \tag{5}$$

Where $t_i>0,\sum_{j=1}^k n_{ij}=n, i=1,\ldots,r$ *and* $\sum_{j=1}^r n_{ji}=\alpha_i,\ j=1,\ldots,r.$

Proof. Here $\sum_{i=1}^n Y_i$ has a gamma distribution and the main theorem of Hadad et al. (2021) concludes that $\sum_{i=1}^n Y_i\mathbf{X}_i$ has a Dirichlet distribution. Of course, it is easy to prove that $\sum_{i=1}^n Y_i$ and $\sum_{i=1}^n\frac{Y_i}{\sum_{i=1}^n Y_i}\mathbf{X}_i$ are independent. Therefore, we can use the theorem **??** and obtain the desired distribution with a simple variable change.

$$\mathbf{T}=\sum_{i=1}^n Y_i\mathbf{X}_i=\sum_{i=1}^n Y_i.\sum_{i=1}^n\frac{Y_i}{\sum_{i=1}^n Y_i}\mathbf{X}_i.$$

□

The following theorem is another characterization on the real lifetime in Homei et al.(2022). It has been tried not to use Stieljes transformation for proof.

4 Approximation of Nadarajah by MLE

It is not easy to calculate the distribution of T and it requires a lengthy calculation to find the distribution of T. Inspired by the proof of the previous theorems, we suggest approximating the distribution of $\sum_{i=1}^n\frac{Y_i}{\sum_{j=1}^n Y_i}\mathbf{X}_i$ first and then obtaining an approximation of the distribution of T. In this section, we propose an approximation for the distribution of $\sum_{i=1}^n\frac{Y_i}{\sum_{i=1}^n Y_i}\mathbf{X}_i$. By Theorem 2 and the fact that the support of Z is Dirichlet distribution, we are interested in approximating its distribution by the Dirichlet family of distributions. The main idea of this approximation is taken from Nadarajah's works; see e.g. Nadarajah et al.(2013), Nadarajah (2006a) and Nadarajah (2006b). The idea of approximating distributions involving complicated formulas by the beta distribution are very well established in the statistics literature. The next procedure is based on the simulation and the Kolmogorov-Smirnov test and thus, the following steps are applied:

step 1 : Generate random numbers of $\sum_{i=1}^n\frac{Y_i}{\sum_{i=1}^n Y_i}\mathbf{X}_i$ by simulating the $\mathbf{X}_i'$s and Y's.
step 2 : Obtain MLE for unknown parameters.
step 3 : We do the forth step in Homei and Nadarajah (2018).

4.1 A comparison with the work of others

The distribution of Z can be approximated by the new method similar to Homei and Nadarajah (2018) and then compared with the method in Homei and Nadarajah (2018) showing to be better very close to it. Moreover, the result can be improved by increasing n. It is worth noting that the previous method is applicable to a univariate problem, but the presented method is applicable to the vectors. To evaluate the approximation of the distribution of Z, we used the p-values given in table 1, which show to be robust for most of the chosen parameters. We illustrate another example from Homei and Nadarajah (2018) graphically. Thus, let W,X_1,X_2 be independent with uniform $[0,1]$ distribution. The density function of Z is $f_Z(z)=-2(log(1-z)^{(1-z)}z^z)I_{(0,1)}(z)$. By using the new method, $B(1.78,1.71)$ will be a good approximation for Z.Of course, another approximation that was obtained for the distribution is $B(1.7,1.7)$. Figure 1 shows the exact and approximate density of Z.

(n_1,m_1)	(n_2,m_2)	(α_1,α_2)	p	q	$P-Value$
(1,1)	(1,1)	(1,1)	1.78	1.71	0.50
(1,1)	(1,1)	(0.2,0.3)	1.24	1.22	0.65
(2,4)	(3,5)	(0.2,0.3)	3.13	5.61	0.58
(2,4)	(3,5)	(1.75,2)	4.3	7.49	0.13
(3,7)	(5,9)	(0.2,0.3)	4.95	9.67	0.23
(3,7)	(5,9)	(1.75,2)	6.92	13.99	0.97
(4,10)	(7,13)	(0.2,0.3)	6.03	12.5	0.26
(4,10)	(7,13)	(1.75,2)	8.81	18.53	0.91
(5,13)	(9,17)	(0.2,0.3)	7.91	16.87	0.25
(5,13)	(9,17)	(1.75,2)	11.26	24.62	0.70

Table 1: Checking robustness of the approximation for r=2.

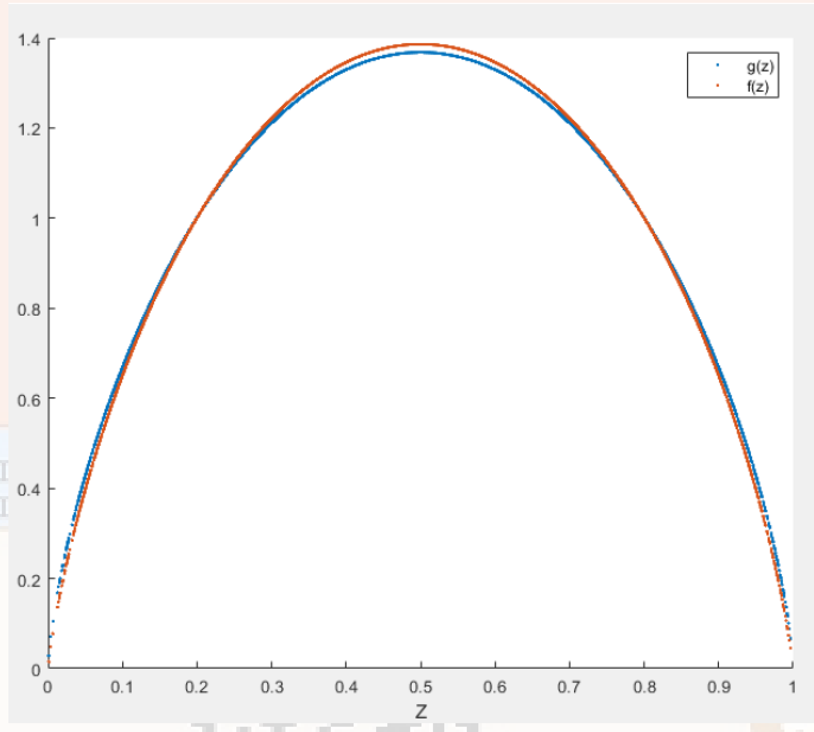

Figure 1: Exact and approximated distribution of Z when n=8000

[1] Presenter. Namer, Email address: m.jalilvand.sarand@gmail.com

5 Conclusions

By considering the environmental conditions, a model is introduced for analyzing a real lifetime in this paper and some of its distributional properties are discussed. Also, an approximation of the distribution is proposed if it is very complicated. In order to compare better the model with the work of Nadarajah, the beta distribution has been considered, but, using the quality of the approximation for Dirichlet distribution directly for evaluation is under investigation.

References

[1] Hadad, H., Homei, H., Behzadi, M., and Farnoosh, R. Solving some stochastic differential equation using Dirichlet distributions, *Computational Methods for Differential Equations.*, 9 (2), (2021), :393-398.

[2] Homei,H. 2021. The stochastic linear combination of dirichlet distributions,*Communications in Statistics: Theory and Methods,* 50(10):2354-2359.

[3] Homei,H., and Nadarajah, S. On products and mixed sums of Gamma and Beta random variables motivated by availability, *Methodology and Computing in Applied Probability.*, 20(2), (2018), 799810.

[4] Homei, H., Nadarajah, S., and Taherkhani, A. 2022 Randomly weighted averages on multivariate dirichlet distributions with generalized parameters, REVSTAT Statistical Journal, Submited

[5] Nadarajah, S., and Kotz, S. 2005. On the Product and Ratio of Gamma and Beta random variables, *Allgemeines Statistisches Archiv.*, 89(4), (2005), :43549.