Pak.j.stat.oper.res. Vol.17 No.1 2021 pp 227-234 DOI: https://doi.org/10.18187/pjsor.v17i1.3702

Pakistan Journal of Statistics and Operation Research

Remarks on and Characterizations of 2S-Lindley and 2D-Lindley Distributions Introduced by Chesneau et al.(2020)

G.G. Hamedani[1*], Mahrokh Najaf[2]

*Corresponding author

1. Department of Mathematical and Statistical Sciences, Marquette University, Milwaukee, WI 53201-1881, USA, g.hamedani@mu.edu
2. Department of Electrical Engineering, University of Science and Culture, Tehran, Iran mahi.najaf@gmail.com

Abstract

Chesneau et al.(2020) considered the distributions of sum and differences of two independent and identically distributed random variables with the common Lindley distribution. They derived, very nicely, the above mentioned distributions and provided certain important mathematical and statistical properties as well as simulations and applications of the new distributions. In this short note, we like to show that the assumption of "independence" can be replaced with a much weaker assumption of "sub-independence". Then we present certain characterizations of the proposed distributions to complete, in someway, their work.

Key Words: Lindley Distribution; Independence; Sub-Independence; Identically Distributed Random Variables; Characterizations of Distributions.

1. Introduction

To make this short note self-contained, we will copy some parts of our previous work Hamedani(2013) here. We may in some occasions have asked ourselves if there is a concept between "uncorrelatedness" and "independence" of two random variables. It seems that the concept of "sub-independence" is the one: it is much stronger than uncorrelatedness and much weaker than independence. The notion of sub-independence seems important in the sense that under usual assumptions, Khintchine's Law of Large Numbers and Lindeberg-Levy's Central Limit Theorem as well as other important theorems in probability and statistics hold for a sequence of sub-independence ($s.i.$) random variables. While sub-independence can be substituted for independence in many cases, it is difficult, in general, to find conditions under which the former implies the latter. Even in the case of two discrete identically distributed random variables X and Y, the joint distribution can assume many forms consistent with sub-independence.
Limit theorems as well as other well-known results in probability and statistics are often based on the distribution of the sums of independent (and often identically distributed) random variables rather than the joint distribution of the summands. Therefore, the full force of independence of the summands will not be required. In other words, it is the convolution of the marginal distributions which is needed, rather than the joint distribution of the summands which, in the case of independence, is the product of the marginal distributions. The concept of sub-independence is shown to be sufficient to yield the conclusions of these

theorems and results. This is precisely the reason for the statement: "why assume independence when you can get by with sub-independence."

The concept of sub-independence can help to provide solution for some modeling problems where the variable of interest is the sum of a few components. Examples include household income, the total profit of major firms in an industry, and a regression model $Y = g(X) + \varepsilon$ where $g(X)$ and ε are uncorrelated; however, they may not be independent. For example, in Bazargan et al.(2007), the return value of significant wave height (Y) is modeled by the sum of a cyclic function of random delay D, $\hat{g}(D)$, and a residual term $\hat{\varepsilon}$. They found that the two components are at least uncorrelated, but not independent and used sub-independence to compute the distribution of the return value.

Let X and Y be two random variables with joint and marginal cumulative distribution functions ($cdfs$) $F_{X,Y}$, F_X and F_Y respectively. Then X and Y are said to be independent if and only if

$$F_{X,Y}(x,y) = F_X(x)\,F_Y(y), \qquad (x,y) \in \mathbb{R}^2. \tag{1.1}$$

or equivalently, if and only if

$$\varphi_{X,Y}(s,t) = \varphi_X(s)\,\varphi_Y(t), \qquad (s,t) \in \mathbb{R}^2. \tag{1.2}$$

where $\varphi_{X,Y}(s,t)$, $\varphi_X(s)$ and $\varphi_Y(t)$, respectively, are the corresponding joint and marginal cfs. Note that (1.1) and (1.2) are also equivalent to

$$\begin{aligned} &P(X \in A \text{ and } Y \in B) = P(X \in A)\,P(Y \in B), \\ &\textit{for all Borel sets } A,\, B. \end{aligned} \tag{1.3}$$

The concept of sub-independence, as far as we have gathered, was formally introduced by Durairajan(1979) and developed by Hamedani in the past 40 years, stated as follows: The random variables X and Y with $cdfs$ F_X and F_Y are sub-independent ($s.i.$) if the cdf of $X+Y$ is given by

$$F_{X+Y}(z) = (F_X * F_Y)(z) = \int_{\mathbb{R}} F_X(z-y)\,dF_Y(y), \qquad z \in \mathbb{R}. \tag{1.4}$$

or equivalently if and only if

$$\varphi_{X+Y}(t) = \varphi_{X,Y}(t,t) = \varphi_X(t)\,\varphi_Y(t), \qquad t \in \mathbb{R}. \tag{1.5}$$

The drawback of the concept of sub-independence in comparison with that of independence has been that the former does not have an equivalent definition in the sense of (1.3), which some believe, to be the natural definition of independence. We found such a definition which is stated below. We shall give the definition for the continuous case (Definition 1.1).

We observe that the half-plane $H = \{(x,y) : \ x+y<0\}$ can be expressed as a countable disjoint union of rectangles:

$$H = \cup_{i=1}^{\infty} E_i \times F_i,$$

where E_i and F_i are intervals. Now, let $(X,Y) : \Omega \to \mathbb{R}^2$ be a continuous random vector and for $c \in \mathbb{R}$, let

$$A_c = \{\omega \in \Omega : X(\omega) + Y(\omega) < c\}$$

and

$$A_i^{(c)} = \left\{\omega \in \Omega : \ X(\omega) - \frac{c}{2} \in E_i\right\}, \ B_i^{(c)} = \left\{\omega \in \Omega : Y(\omega) - \frac{c}{2} \in F_i\right\},$$

Definition 1.1. The continuous random variables X and Y are $s.i.$ if for every $c \in \mathbb{R}$

$$P(A_c) = \sum_{i=1}^{\infty} P\left(A_i^{(c)}\right) P\left(B_i^{(c)}\right), \tag{1.6}$$

To see that (1.6) is equivalent to (1.4), observe that (LHS of (1.6))

$$P(A_c) = P(X+Y<c) = P((X,Y)\in H_c), \tag{1.7}$$

where $H_c = \{(x,y) : x+y<c\}$. Now, if X and Y are $s.i.$ then

$$P(A_c) = (P_X \times P_Y)(H_c)$$

where P_X, P_Y are probability measures on $\mathbb{R}$ defined by

$$P_X(B) = P(X\in B) \text{ and } P_Y(B) = P(Y\in B) \ ,$$

and $P_X \times P_Y$ is the product measure.

We also observe that (RHS of (1.6))

$$\begin{aligned}\sum_{i=1}^{\infty} P\left(A_i^{(c)}\right) P\left(B_i^{(c)}\right) &= \sum_{i=1}^{\infty} P\left(X-\frac{c}{2}\in E_i\right) P\left(Y-\frac{c}{2}\in F_i\right)\\ &= \sum_{i=1}^{\infty} P\left(X\in E_i+\frac{c}{2}\right) P\left(Y\in F_i+\frac{c}{2}\right)\\ &= \sum_{i=1}^{\infty} P_X \times P_Y \ \left(E_i+\frac{c}{2}\right)\times\left(F_i+\frac{c}{2}\right).\end{aligned} \tag{1.8}$$

Now, (1.7) and (1.8) will be equal if $H_c = \cup_{i=1}^{\infty}\left\{\left(E_i+\frac{c}{2}\right)\times\left(F_i+\frac{c}{2}\right)\right\}$, which is true since the points in H_c are obtained by shifting each point in H over to the right by $\frac{c}{2}$ units and then up by $\frac{c}{2}$ units.
If X and Y are $s.i.$, then unlike independence, X and αY are not necessarily $s.i.$ for any real $\alpha \neq 1$. This demonstrates how weak is the concept of sub-independence in comparison with that of independence. Please observe the following simple example.

Example 1.1. Let X and Y have the joint cf given by

$$\begin{gathered}\varphi_{X,Y}(t_1,t_2) = \exp\left\{-\left(t_1^2+t_2^2\right)/2\right\}[1+\beta\ t_1t_2\left(t_1-t_2\right)^2\times\\ \exp\left\{\left(t_1^2+t_2^2\right)/4\right\}], \qquad (t_1,t_2)\in\mathbb{R}^2.\end{gathered}$$

where β is an appropriate constant. (The characteristic function is the Fourier Transform of probability density function (pdf), so the corresponding joint pdf is given by

$$\begin{gathered}f(x,y) = \tfrac{1}{2\pi}\exp\left\{-\left(x^2+y^2\right)/2\right\}[1-16\beta\ p(x,y)\times\\ \exp\left\{-\left(x^2+y^2\right)/2\right\}], \qquad (x,y)\in\mathbb{R}^2.\end{gathered}$$

where $p(x,y) = \left\{6xy-2x^2-2y^2+4x^2y^2-2x^3y-2xy^3+1\right\}$).

Then X and Y are $s.i.$ standard normal random variables, and hence $X+Y$ is normal with mean 0 and variance 2, but X and $-Y$ are not $s.i.$ and consequently $X-Y$ does not have a normal distribution.

The concept of sub-independence defined above can be extended to n (>2) random variables as follows.

Definition 1.2. The random variables $X_1,\ X_2,...,X_n$ are $s.i.$ if for each subset $\{X_{\alpha_1},X_{\alpha_2},...,X_{\alpha_r}\}$ of $\{\ X_1,X_2,...,X_n\}$

$$\varphi_{X_{\ \alpha_1},...,X_{\ \alpha_r}}(t,...,t) = \Pi_{i=1}^{r}\varphi_{X_{\ \alpha_i}}(t), \qquad t\in\mathbb{R}.$$

2. Remarks

i) If the random variables X and Y are sub-independent identically distributed ($s.i.i.d.$) with the common Lindley distribution with the parameter θ, the characteristic function of $X+Y$ is

$$\varphi_{X+Y}(t)=\frac{\theta^4(\theta-it+1)^2}{(1+\theta)^2(\theta-it)^4}, \qquad t\in\mathbb{R}.$$

The cf of X is

$$\begin{aligned}\varphi_X(t)&=\int_0^\infty e^{itx}\frac{\theta^2}{1+\theta}(1+x)e^{-\theta x}dx\\&=\frac{\theta^2(\theta-it+1)}{(1+\theta)(\theta-it)^2}, \qquad t\in\mathbb{R}.\end{aligned}$$

and since X and Y are $s.i.$, we have

$$\begin{aligned}\varphi_{X+Y}(t)=\varphi_X(t)\varphi_Y(t)&=\left\{\frac{\theta^2(\theta-it+1)}{(1+\theta)(\theta-it)^2}\right\}^2\\&=\frac{\theta^4(\theta-it+1)^2}{(1+\theta)^2(\theta-it)^4}, \qquad t\in\mathbb{R}.\end{aligned}$$

ii) If the random variables X and Y are identically distributed ($i.d.$) with the common Lindley distribution with the parameter θ, and if X and $-Y$ are $s.i.$, the characteristic function of $X-Y$ is

$$\varphi_{X-Y}(t)=\frac{\theta^4\left((\theta+1)^2+t^2\right)}{(1+\theta)^2(\theta^2+t^2)^2}, \qquad t\in\mathbb{R}.$$

The cf of $X-Y$, under the assumption of $s.i.$ of X and $-Y$, is

$$\begin{aligned}\varphi_{X-Y}(t)&=\varphi_X(t)\varphi_{-Y}(t)=\varphi_X(t)\varphi_Y(-t)\\&=\left\{\frac{\theta^2(\theta-it+1)}{(1+\theta)(\theta-it)^2}\right\}\left\{\frac{\theta^2(\theta+it+1)}{(1+\theta)(\theta+it)^2}\right\}\\&=\frac{\theta^4\left((\theta+1)^2+t^2\right)}{(1+\theta)^2(\theta^2+t^2)^2}, \qquad t\in\mathbb{R}.\end{aligned}$$

iii) For a detailed treatment of the concept of sub-independence, we refer the interested reader to Hamedani(2013).

3. Characterizations of the 2S-Lindley and 2D-Lindley Distributions

Chesneau et al.(2020) introduced the distributions of the sum and differences of two $i.i.d.$ (now, $s.i.i.d.$) Lindley random variables with the parameter $\theta>0$ (called 2S-Lindley and 2D-Lindley) with their respective $pdfs$ given by

$$f_{2S-L}(x)=\frac{\theta^4 x}{(1+\theta)^2}\left(\frac{x^2}{6}+x+1\right)e^{-\theta x}, \quad x>0, \tag{3.1}$$

and

$$f_{2D-L}(x) = \frac{\theta^4}{(1+\theta)^2} e^{-\theta[x-2\inf(x,0)]} \times \left[\begin{array}{c} 2\theta^2 \inf(x,0)^2 - 2\theta^2 \inf(x,0)\,x - 4\theta^2 \inf(x,0) \\ +2\theta^2 x + 2\theta^2 - 2\theta \inf(x,0) + \theta x + 2\theta + 1 \end{array} \right], \tag{3.2}$$

where $x \in \mathbb{R}$.

To understand the behavior of the data obtained through a given process, we need to be able to describe this behavior via its approximate probability law. This, however, requires to establish conditions which govern the required probability law. In other words, we need to have certain conditions under which we may be able to recover the probability law of the data. So, the characterization of a distribution is important in applied sciences, where an investigator is vitally interested to find out if their model follows the selected distribution. Therefore, the investigator relies on conditions under which their model would follow a specified distribution. A probability distribution can be characterized in different directions one of which is based on the truncated moments. This type of characterization initiated by Galambos and Kotz(1978) and followed by other authors such as Kotz and Shanbhag(1980), Glänzel et al.(1984), Glänzel(1987), Glänzel and Hamedani(2001) and Kim and Jeon(2013), to name a few. For example, Kim and Jeon(2013) proposed a credibility theory based on the truncation of the loss data to estimate conditional mean loss for a given risk function. It should also be mentioned that characterization results are mathematically challenging and elegant. In this section, we present characterizations of the 2S-Lindley and 2D-Lindley distributions based on the conditional expectation (truncated moments) of certain functions of the random variable.

We will employ Theorem 1 of Glänzel(1987) given in the Appendix A. As shown in Glänzel(1990), this characterization is stable in the sense of weak convergence.

Proposition 3.1. Let X be a continuous random variable and let $q_1(x) = x^{-1}\left(\frac{x^2}{6} + x + 1\right)^{-1}$ and $q_2(x) = q_1(x)\, e^{-\theta x}$ for $x > 0$. Then X has *pdf* (3.1) if and only if the function ξ defined in Theorem 1 is of the form

$$\xi(x) = \frac{1}{2} e^{-\theta x}, \qquad x > 0.$$

Proof. If X has pdf (3.1), then

$$(1 - F_{2S-L}(x))\, E[q_1(X) \mid X \geq x] = \frac{\theta^3}{(1+\theta)^2} e^{-\theta x}, \qquad x > 0,$$

and

$$(1 - F_{2S-L}(x))\, E[q_2(X) \mid X \geq x] = \frac{\theta^3}{2(1+\theta)^2} e^{-2\theta x}, \qquad x > 0,$$

and hence

$$\xi(x) = \frac{1}{2} e^{-\theta x}, \qquad x > 0.$$

We also have

$$\xi(x)\, q_1(x) - q_2(x) = -\frac{1}{2} q_1(x)\, e^{-\theta x} < 0, \qquad for \quad x > 0.$$

Conversely, if ξ is of the above form, then

$$s'(x) = \frac{\xi'(x)\, q_1(x)}{\xi(x)\, q_1(x) - q_2(x)} = \theta, \qquad x > 0.$$

and

$$s(x) = \theta x.$$

Now, according to Theorem 1, X has density (3.1).

Corollary 3.1. Suppose X is a continuous random variable. Let $q_1(x)$ be as in Proposition 3.1. Then X has density (3.1) if and only if there exist functions q_2 and ξ defined in Theorem 1 for which the following first order differential equation holds

$$\frac{\xi'(x)\, q_1(x)}{\xi(x)\, q_1(x) - q_2(x)} = \theta, \qquad x > 0.$$

Corollary 3.2. The differential equation in Corollary 3.1 has the following general solution

$$\xi(x) = e^{\theta x}\left[-\int \theta e^{-\theta x}\left(q_1(x)\right)^{-1} q_2(x) + D\right],$$

where D is a constant.

Proof. If X has pdf (3.1), then clearly the differential equation holds. Now, if the differential equation holds, then

$$\xi'(x) = \xi(x)\,\theta - \theta\left(q_1(x)\right)^{-1} q_2(x),$$

or

$$\xi'(x)\, e^{-\theta x} - \xi(x)\,\theta\; e^{-\theta x} = -\theta\; e^{-\theta x}\left(q_1(x)\right)^{-1} q_2(x),$$

or

$$\frac{d}{dx}\left\{\, e^{-\theta x}\xi(x)\right\} = -\theta\; e^{-\theta x}\left(q_1(x)\right)^{-1} q_2(x),$$

from which we arrive at

$$\xi(x) = e^{\theta x}\left[-\int \theta e^{-\theta x}\left(q_1(x)\right)^{-1} q_2(x) + D\right].$$

A set of functions satisfying the above differential equation is given in Proposition 3.1 with $D = 0$. Clearly, there are other triplets (q_1, q_2, ξ) satisfying the conditions of Theorem 1.

Remark 3.1. Similar results can be stated for the 2D-Lindley distribution as well.

References

1. Bazargan, H., Bahai, H., and Aminzadeh-Gohari, A. (2007). Calculating the return value using a mathematical model of significant wave height. *J. Marine Science and Technology*, 11:34–42.
2. Chesneau, C., Tomy, L., and Gillatiose, J. (2020). On a sum and difference of two lindley distributions: theory and applications. *Revstat-Statistical Journal*, 18(5):673–695.
3. Galambos, J. and Kotz, S. (1978). *Characterizations of probability distributions: A unified approach with an emphasis on exponential and related Models, Lecture Notes in Mathematics, Vol. 675.* Springer, Berlin.
4. Glänzel, W. (1987). A characterization theorem based on truncated moments and its application to some distribution families. *Mathematical Statistics and Probability Theory (Bad Tatzmannsdorf, 1986)*, B:75–84.
5. Glänzel, W. (1990). Some consequences of a characterization theorem based on truncated moments, statistics:. *A Journal of Theoretical and Applied Statistics*, 21(4):613–618.
6. Glänzel, W. and Hamedani, G. G. (2001). Characterizations of the univariate continuous distributions. *Studia Scientiarum Mathematicarum Hungarica*, 37(1-2):83–118.
7. Glänzel, W., Teles, A., and Schubert, A. (1984). Characterization by truncated moments and its application to Pearson-type distributions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 66:173–182.
8. Hamedani, G. (2013). Sub-independence: an expository perspective. *CSTM,*, 42(20):3615–3638.
9. Kim, J. H. and Jeon, Y. (2013). Credibility theory based on trimming. *Insurance: Mathematics and Economics*, 53(1):36–47.
10. Kotz, S. and Shanbhag, D. N. (1980). Some new approaches to probability distributions. *Advances in Applied Probability*, 12(4):903–921.

Appendix A

Theorem 1. Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a given probability space and let $H = [a, b]$ be an interval for some $a < b$ $(a = -\infty,\ b = \infty$ might as well be allowed$)$. Let $X : \Omega \to H$ be a continuous random variable with the distribution function F and let q_1 and q_2 be two real functions defined on H such that

$$\mathbf{E}\left[q_2(X) \mid X \geq x\right] = \mathbf{E}\left[q_1(X) \mid X \geq x\right]\xi(x), \qquad x \in H,$$

is defined with some real function ξ. Assume that $q_1, q_2 \in C^1(H)$, $\xi \in C^2(H)$ and F is twice continuously differentiable and strictly monotone function on the set H. Finally, assume that the equation $\xi q_1 = q_2$ has no real solution in the interior of H. Then F is uniquely determined by the functions q_1, q_2 and ξ, particularly

$$F(x) = \int_a^x C \left|\frac{\xi'(u)}{\xi(u)\, q_1(u) - q_2(u)}\right| \exp(-s(u))\ du\ ,$$

where the function s is a solution of the differential equation $s' = \frac{\xi'\, q_1}{\xi q_1 - q_2}$ and C is the normalization constant, such that $\int_H dF = 1$.

Note: The goal is to have the function $\xi(x)$ as simple as possible.

We like to mention that this kind of characterization based on the ratio of truncated moments is stable in the sense of weak convergence (see Glänzel(1990)) , in particular, let us assume that there is a sequence $\{X_n\}$ of random variables with distribution functions $\{F_n\}$ such that the functions q_{1n} , q_{2n} and ξ_n $(n \in \mathbb{N})$ satisfy the conditions of Theorem 1 and let $q_{1n} \to q_1$, $q_{2n} \to q_2$ for some continuously differentiable real functions q_1 and q_2. Let, finally, X be a random variable with the distribution F. Under the condition that $q_{1n}(X)$ and $q_{2n}(X)$ are uniformly integrable and the family $\{F_n\}$ is relatively compact, the sequence X_n converges to X in distribution if and only if ξ_n converges to ξ, where

$$\xi(x) = \frac{E[q_2(X) \mid X \geq x]}{E[q_1(X) \mid X \geq x]}.$$

This stability theorem makes sure that the convergence of distribution functions is reflected by corresponding convergence of the functions q_1, q_2 and ξ, respectively. It guarantees, for instance, the 'convergence' of characterization of the Wald distribution to that of the Lévy-Smirnov distribution if $\alpha \to \infty$.

A further consequence of the stability property of Theorem 1 is the application of this theorem to special tasks in statistical practice such as the estimation of the parameters of discrete distributions. For such purpose, the functions q_1, q_2 and, specially, ξ should be as simple as possible. Since the function triplet is not uniquely determined, it is often possible to choose ξ as a linear function. Therefore, it is worth analyzing some special cases which helps to find new characterizations reflecting the relationship between individual continuous univariate distributions and appropriate in other areas of statistics.

In some cases, one can take $q_1(x) \equiv 1$, which reduces the condition of Theorem 1 to $\mathbf{E}[q_2(X) \mid X \geq x] = \xi(x),\ x \in H$. We, however, believe that employing three functions q_1, q_2 and ξ will enhance the domain of applicability of Theorem 1.