Statistics and Applications {ISSN 2454-7395 (online)}
Volume 19, No. 1, 2021 (New Series), pp 337–365

On the Interrelationships between Lomax, Pareto and Exponentiality with Certain Extensions

Zhixin Lun[1] **and Ravindra Khattree**[2]
[1]*Department of Clinical Pharmacy*
University of California, San Francisco, California, USA

[2]*Department of Mathematics and Statistics and*
Center for Data Science and Big Data Analytics
Oakland University, Rochester, Michigan, USA

Received: 31 October 2020; Revised: 08 March 2021; Accepted: 11 March 2021

Abstract

We start with a recently published connection from Lomax to exponential distribution through the limiting distribution of a Lomax distribution scaled by its shape parameter. Motivated by that observation, we explore several relationships between (generalized) Lomax and other distributions of exponential family such as Gamma, Beta type II, Rayleigh and Weibull. As further extension, we introduce the generalized double Pareto distribution on the entire real line. Various properties of generalized double Pareto distribution are then studied including its representation as a mixture of Student's t and its connection to Laplace (double exponential) distribution. We then provide a simple approach to simulate random numbers from the double Pareto distribution and its implementation in R. Finally, we illustrate an application in a real biomedical research problem.

Key words: Double Pareto distribution; Exponential distribution; Khattree-Bahuguna skewness; Laplace distribution; Lomax distribution; Multivariate Lomax distribution.

AMS Subject Classifications: 62E10, 62E15, 62H05

1. Introduction

It is our distinct privilege to dedicate this paper in honor of Late Prof. Aloke Dey. One of us has learned a lot and personally benefited greatly in his research as well as in teaching from Prof. Dey's landmark book on block designs and although later he never got a chance to thank Prof. Dey in person, we hope our this article serves as our symbolic appreciation of Prof. Dey's contributions to the goal of advancing the knowledge for the betterment of society.

This article came into being due to an earlier simple curiosity described in Lun and Khattree (2020) about univariate and multivariate Lomax distributions. The relationship between Pareto/Lomax and exponential distributions has been well recorded in the literature (see Johnson, Kotz, and Balakrishnan (1994) and Kotz, Kozubowski and Podgórski (2001))

Corresponding Author: Ravindra Khattree
Email: khattree@oakland.edu

yet has not been very well publicized. Harris (1968) provides an approach to generate the Pareto variates through a mixture of exponential variates with parameter having a gamma distribution. Specifically and more generally, let X follow an exponential distribution with rate parameter η and allow η to have a gamma distribution with shape parameter β and scale parameter θ. Then, the unconditional probability density function of X is

$$f(x) = \int_0^\infty \eta e^{-\eta x} \frac{1}{\Gamma(\beta)\theta^\beta} \eta^{\beta-1} e^{-\eta/\theta} d\eta = \frac{\theta\beta}{(1+\theta x)^{\beta+1}}, \quad x > 0, \theta, \beta > 0, \tag{1}$$

which is the density function of Lomax (also called Pareto type II) distribution (hereon denoted by Lomax(β, θ)) with shape parameter β and rate parameter θ. Alternatively, given two independent standard exponential variates W_1 and W_2, the probability density function of $Y = W_1/W_2$ is a standard Lomax distribution (with $\theta = 1, \beta = 1$). On the other hand, assume that X has a Pareto type I distribution with density $g(x) = \lambda x^{-(\lambda+1)}$, $x \geq 1$, then the density function of $Y = \log X$ is $\lambda e^{-\lambda y}, y > 0$. To indicate the other connections, in Table 1, we summarize the expectations, variances, Pearson's coefficients of skewness and Khattree-Bahuguna's skewness [1] of Lomax and exponential distributions.

As $\beta \to \infty$, for Lomax$(\beta, \frac{\theta}{\beta})$, we observe that the variance approaches the square of expectation of an Exponential distribution with the rate parameter θ and the Pearson's coefficient of skewness approaches to 2. As we shall see later, as $\beta \to \infty$ the Khattree-Bahuguna's skewness of the above Lomax distribution also approaches that of exponential distribution. Proof is given in Section 2.3. A natural question then arises as to why does Lomax distribution have its distributional properties so similar to those of exponential distribution?

One of our main objectives in this article is to explore above question by showing and generalizing the connections between Lomax and exponential distributions. These have been motivated by a result by Lun and Khattree (2020) about the limiting distribution of a Lomax random variable scaled by its shape parameter. Based on this result, we discover many more results which establish many other connections between (generalized) Lomax and other distributions of exponential family. We further take upon introducing another connection between Laplace (double exponential) distribution and generalized double Pareto distribution, which is a popular choice of prior distribution in recent years for robust Bayesian shrinkage estimators.

The article is organized as follows. In Section 2, we show the Lomax-exponential and generalized Lomax-gamma connections. Then, we attempt to quantify the distance of Lomax to exponential distributions in terms of Patil-Patil-Bagkavos's η (2012) and by using the Kullback-Leibler divergence. In Section 3, we give some results pertaining to relationships between multivariate Lomax and other distributions of exponential family such as Gamma, Beta type II, Rayleigh and Weibull. In Section 4, we discuss a three-parameter generalized double Pareto distribution, including its simulation and connection to Laplace distribution. Section 5 includes an approach to simulation of generalized double Pareto variates. In Section 6, we give a real-world application of bivariate Lomax distribution. Section 7 includes some concluding remarks.

[1]Khattree-Bahuguna's skewness for a random variable is defined in Khattree and Bahuguna (2019) and reproduced here in Section 2.3. The values of this skewness for Lomax and Exponential distributions are also computed there.

Table 1: Comparison of summary parameters for Lomax(β, θ) and Exp(λ)

Summary Parameter	Lomax(β, θ)	Exp(λ)
E(X)	$\frac{1}{\theta(\beta-1)}, \quad \beta > 1$	$\frac{1}{\lambda}$
Var(X)	$\frac{\beta}{\theta^2(\beta-1)^2(\beta-2)}, \quad \beta > 2$	$\frac{1}{\lambda^2}$
Pearson's coefficient of skewness	$\frac{2(\beta+1)}{\beta-3}\sqrt{\frac{\beta-2}{\beta}}, \quad \beta > 3$	2
Khattree-Bahuguna's skewness	$\frac{1}{2}\left[1+\left(B(1-\frac{1}{\beta},1-\frac{1}{\beta})-\frac{\beta^2}{(\beta-1)^2}\right)\frac{(\beta-1)^2(\beta-2)}{\beta}\right], \quad \beta > 2$	$1-\frac{\pi^2}{12} \simeq 0.177533$

2. Connections Between Generalized Lomax and Gamma Distributions

In this section, we first state a surprisingly simple connection between Lomax and exponential distributions partially given in Lun and Khattree (2020). This result is then generalized to connect the generalized Lomax with Gamma distribution in a straight forward way. To begin with, we first address the limiting distributions.

2.1. Limiting distributions

The following two theorems about limiting distributions can be easily proved.

Theorem 1: Let X be a univariate Lomax(β, θ) random variable with probability density function defined in (1). Define $Y = \beta X$. Then

(i) the distribution of Y is Lomax($\beta, \frac{\theta}{\beta}$);

(ii) as $\beta \to \infty$, the distribution of Y approaches an exponential distribution with rate parameter θ;

(iii) as $\beta \to \infty$, the hazard function of Y approaches θ.

The result (ii) is given in Lun and Khattree (2020). It may be pointed out that an exponential distribution is characterized by its constant hazard function. Thus, (iii) in fact, provides an alternative proof of (ii). We leave it to the reader to calculate the hazard function of the indicated Lomax distribution and then verify the assertion in (iii). The importance of above theorem is that it allows us to be able to conveniently substitute, for large β, one distribution for another by approximating Lomax($\beta, \frac{\theta}{\beta}$) by Exp($\theta$). Figure 1 shows a series of density function plots for Lomax($\beta, \frac{\lambda}{\beta}$) for parameter $\beta = 3.01, 10, 20$ and for a fixed $\lambda = 0.25$, along with an exponential density with rate parameter λ. The closeness of two distributions for large β values is self evident.

Nayak (1987) has introduced a k-dimensional multivariate Lomax distribution by mixing k independent univariate exponential distributions with different failure rates with the mixing parameter η that has a gamma distribution with certain shape parameter β and the scale parameter 1. The Theorem 1 is easily extended to connect the multivariate Lomax and exponential distributions. Again, see Lun and Khattree (2020).

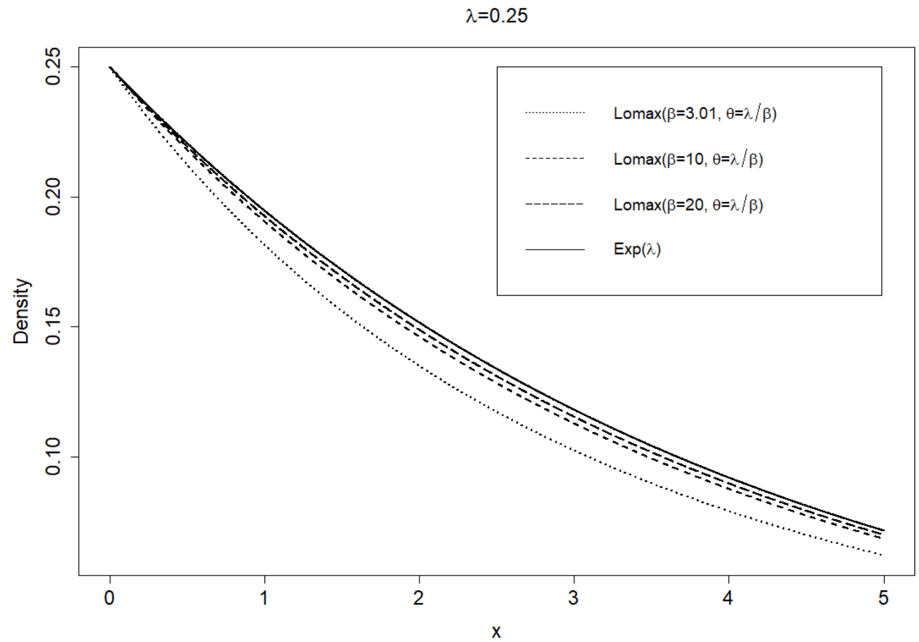


Figure 1: Density plots of a series of Lomax density for various β values ($= 3.01, 10, 20$) and with $\theta = \lambda/\beta$ where $\lambda = 0.25$, along with exponential density with rate parameter $\lambda = 0.25$.

Theorem 2: Let $X_1, X_2, \ldots, X_k$ jointly have k-dimensional multivariate distribution with probability density function (Nayak, 1987),

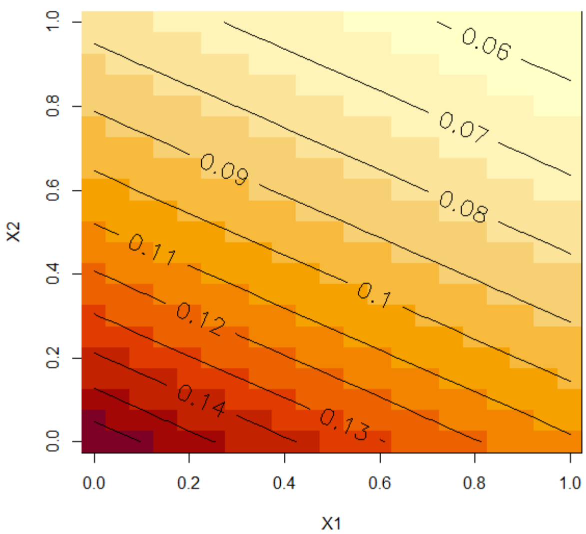
$$f(x_1, x_2, \cdots, x_k) = \frac{\left(\prod_{i=1}^k \theta_i\right) \prod_{i=1}^k (\beta + i - 1)}{\left(1 + \sum_{i=1}^k \theta_i x_i\right)^{\beta+k}}, \quad \beta > 0, \tag{2}$$

where $\theta_i, x_i > 0, i = 1, \cdots, k$. Define $Y_i = \beta X_i, i = 1, \cdots, k$. Then as $\beta \to \infty$, the joint probability distribution of $Y_1, Y_2, \ldots, Y_k$ approaches to that of k independent exponential variates with rate parameters $\theta_i, i = 1, 2, \cdots, k$, respectively.

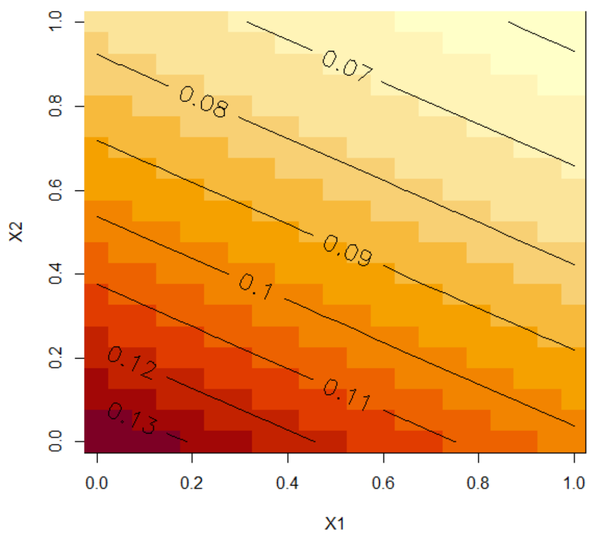
A series of density contour plots for bivariate Lomax distributions with parameters β ($= 3.01, 10, 200$) and $\theta_1 = \lambda_1/\beta, \theta_2 = \lambda_2/\beta$ where $\lambda_1 = 0.25, \lambda_2 = 0.50$ are shown in Figure 2 (a), (b) and (c). As one can see, as β increases, the density contour plots of bivariate Lomax distribution more and more resemble the bivariate independent exponential density contour plot given in Figure 2 (d) with respective rate parameters $\lambda_1 = 0.25, \lambda_2 = 0.50$.

Note that the resulting limiting distribution involves independent exponential variates. In some ways, this is somewhat surprising even though the correlation matrix of a multivariate Lomax distribution has a compound symmetric structure with $\mathrm{Corr}(X_i, X_i) = \frac{1}{\beta}$ which clearly goes to 0 when $\beta \to \infty$. On the other hand, this fact also underscores the well known difficulty that researchers have encountered to satisfactorily define a suitable multivariate exponential distribution with some kind of dependence among variables.

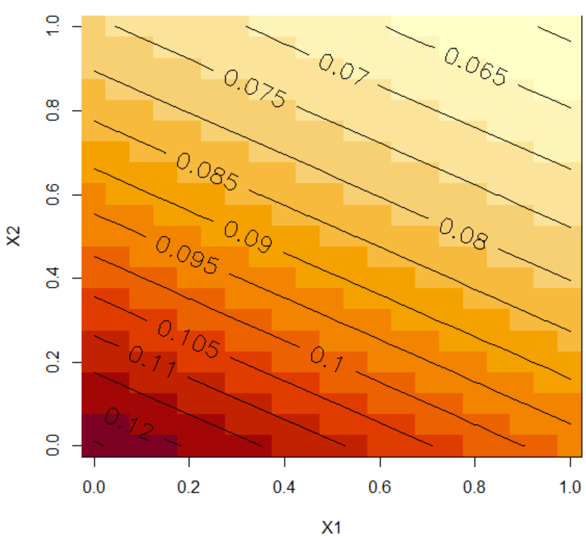
Nayak (1987) also generalized the distribution in (2) by mixing conditionally independent X_i having the Gamma$(l_i, \eta\theta_i)$ distribution, with a mixing variable $\eta \sim$ Gamma$(\beta, 1)$,



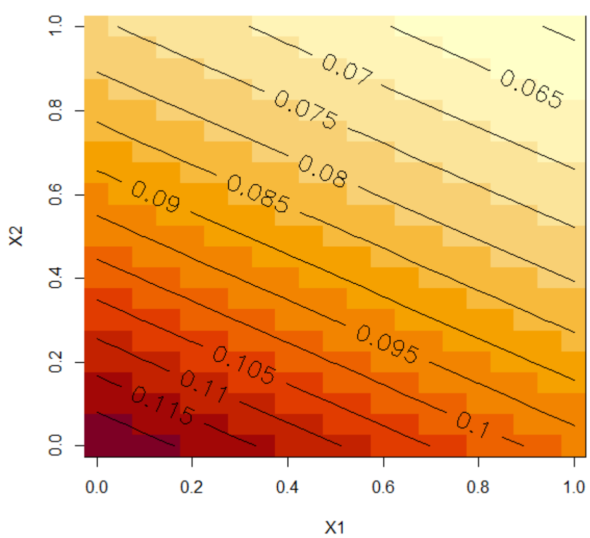
(a) Bivariate Lomax $(\beta = 3.01, \theta_1 = \frac{\lambda_1}{\beta}, \theta_2 = \frac{\lambda_2}{\beta})$



(b) Bivariate Lomax $(\beta = 10, \theta_1 = \frac{\lambda_1}{\beta}, \theta_2 = \frac{\lambda_2}{\beta})$



(c) Bivariate Lomax $(\beta = 200, \theta_1 = \frac{\lambda_1}{\beta}, \theta_2 = \frac{\lambda_2}{\beta})$



(d) Independent Bivariate Exponential (λ_1, λ_2)

Figure 2: Contour plots of a series of bivariate Lomax distributions with changing parameters $\beta, \theta_1 = \lambda_1/\beta, \theta_2 = \lambda_2/\beta$ **where** $\beta = 3.01, 10, 200$ **and** $\lambda_1 = 0.25, \lambda_2 = 0.50$ **and independent bivariate exponential density function with rate parameters** $\lambda_1 = 0.25$ **and** $\lambda_2 = 0.50$**, respectively.**

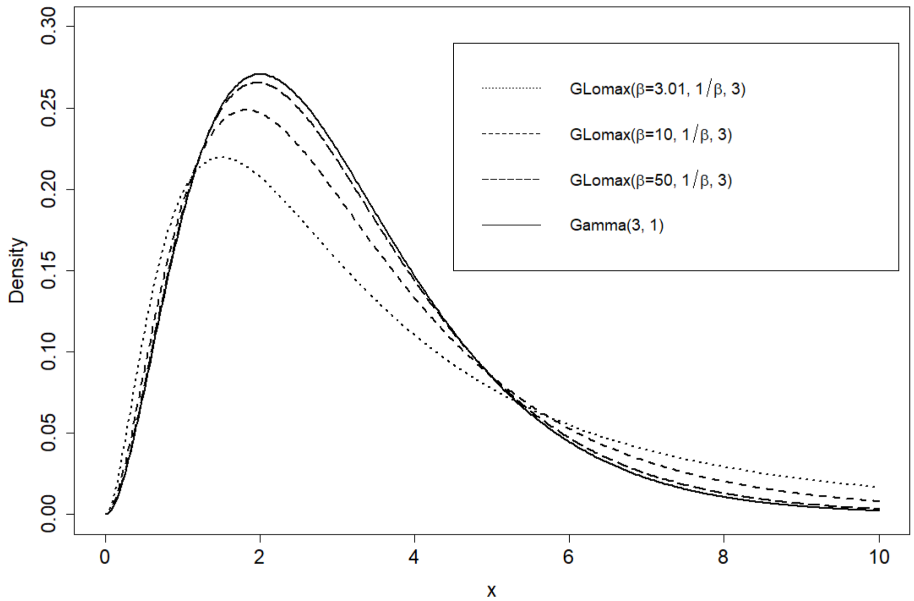


Figure 3: Density plots of a series of generalized Lomax distributions with changing parameters $\beta(= 3.01, 10, 50)$, $\theta = 1.0/\beta$ and a fixed parameter $l = 3.0$, and gamma density function with shape parameter $l = 3.0$ and rate parameter $\lambda = 1.0$.

$i = 1, \ldots, k$. The generalization of above multivariate Lomax-exponential connection to this generalized multivariate Lomax-gamma connection is stated in Theorem 3.

Theorem 3: Let $X_1, X_2, \ldots, X_k$ be a k-dimensional generalized multivariate Lomax random variable with probability density function,

$$f(x_1, x_2, \cdots, x_k) = \frac{\left(\prod_{i=1}^k \theta_i^{l_i}\right) \Gamma\left(\sum_{i=1}^k l_i + \beta\right) \left(\prod_{i=1}^k x_i^{l_i - 1}\right)}{\Gamma(\beta) \left[\prod_{i=1}^k \Gamma(l_i)\right] \left(1 + \sum_{i=1}^k \theta_i x_i\right)^{\sum_{i=1}^k l_i + \beta}},$$

where $\beta, l_i, \theta_i, x_i > 0, i = 1, \cdots, k$. Define $Y_i = \beta X_i, i = 1, \cdots, k$. Then as $\beta \to \infty$, the joint distribution of $Y_1, Y_2, \ldots, Y_k$ approaches that of k independent gamma random variables with shape parameter l_i and rate parameters θ_i, $i = 1, 2, \cdots, k$, respectively.

Clearly, in the special case of $l_i = 1, i = 1, \cdots, k$, the above generalized Lomax-gamma connection reduces to the previous Lomax-exponential connection. Again, to underscores the closeness, we give the density function plots in Figure 3 for a series of univariate generalized Lomax distributions for various values of the parameters $\beta, \theta = \lambda/\beta$ where $\beta = 3.01, 10, 50$, $\lambda = 1.0$, and for a fixed parameter $l = 3.0$, along with Gamma(3,1). To avoid being repetitive, we suppress the contour plots.

2.2. Some measures of closeness to exponential distribution

We have shown that the similarity between Lomax and exponential is due to the fact that the limiting distribution of Lomax$(\beta, \frac{\theta}{\beta})$ is Exp(θ) as $\beta \to \infty$. That begs the question: How to quantify the closeness between a given Lomax distribution and an exponential distribution? To do so, we adopt several approaches. Specifically, we first use Patil-Patil-Bagkavos's η (2012) to measure the closeness of Lomax to exponential and then employ the Kullback-Leibler divergence measure. Later In Section 2.3, we also evaluate their Khattree-Bahuguna skewnesses for this purpose. In order to do all of this, we first define the Patil-Patil-Bagkavos's η (2012).

2.2.1. Patil-Patil-Bagkavos's η

Patil, Patil and Bagkavos (2012) attempted to propose a measure of (a)symmetry of a random variable X as,

$$\eta = \begin{cases} -\mathrm{Corr}(f(X), F(X)) & \text{if } 0 < \mathrm{Var}(f(x)) < \infty, \\ 0 & \text{if } \mathrm{Var}(f(x)) = 0, \end{cases}$$

where $f(x)$ and $F(x)$ are the probability density function and the cumulative distribution function of X, respectively. Clearly, $-1 \leq \eta \leq 1$. However, these authors incorrectly claimed that the η defined above is a measure of the degree of (a)symmetry of a distribution and hence can be used as a measure of skewness of a distribution. Eberl and Klar (2019) disputed their claim and via several examples, they demonstrated that above as a measure of asymmetry is indeed a misleading measure. They further pointed out that η is instead, a measure of the closeness of a given distribution to the exponential distribution and based on their extensive discussion, we readily agree! They further pointed out that η equal to zero indicates a complete departure from exponential distribution while values of $+1$ (-1) show a complete similarity with the positive (negative) exponential. Therefore, in our context, η can be deemed as a tailor-made measure to evaluate the closeness of a Lomax distribution to the exponential distribution. The following theorem gives an explicit expression for η for the Lomax distribution.

Theorem 4: For the Lomax distribution, the Patil-Patil-Bagkavos's η is given by, $\eta = \frac{\sqrt{3\beta(3\beta+2)}}{(3\beta+1)}$.

Proof:. Let X be a Lomax random variable with pdf given in (1). The cumulative distribution function of X is then given by $F(x) = 1 - \frac{1}{(1+\theta x)^{\beta}}$. The covariance between $f(x)$ and $F(x)$ is $\mathrm{Cov}\left[f(X), F(X)\right] = \int_0^{\infty} \left[\frac{\theta\beta}{(1+\theta x)^{\beta+1}}\right]^2 \left[1 - \frac{1}{(1+\theta x)^{\beta}}\right] dx - \int_0^{\infty} \left[\frac{\theta\beta}{(1+\theta x)^{\beta+1}}\right]^2 dx \cdot \frac{1}{2} = -\frac{\theta\beta^2(\beta+1)}{2(2\beta+1)(3\beta+1)}$. Also, $\mathrm{Var}(F(x)) = \frac{1}{12}$ and $\mathrm{Var}(f(x)) = \frac{\theta^2\beta^3(\beta+1)^2}{(3\beta+2)(2\beta+1)^2}$. Thus,

$$\eta = -\mathrm{Corr}(f(X), F(X)) = \frac{\frac{\theta\beta^2(\beta+1)}{2(2\beta+1)(3\beta+1)}}{\sqrt{\frac{\theta^2\beta^3(\beta+1)^2}{(3\beta+2)(2\beta+1)^2}}\sqrt{\frac{1}{12}}} = \frac{\sqrt{3\beta(3\beta+2)}}{(3\beta+1)}. \qquad \square$$

As one would anticipate, η does not depend on scale parameter θ since correlation is invariant of any such scaling.

Straight forward calculations show that even for β as small as 3, $\eta = 0.995$, which indicates that Lomax distribution is generally very similar to exponential distribution even for small β value. Clearly, as $\beta \to \infty$, $\eta \to 1$, thereby reaffirming the previous result about a β multiple of Lomax distribution converging to the exponential distribution .

2.2.2. Kullback-Leibler divergence

Kullback-Leibler divergence is a measure of how different a given probability distribution described by a probability density function $f(x)$ is from another reference distribution with the probability density function $g(x)$ and is defined as

$$D_{KL}(f:g) = \mathrm{E}_f\left(\log\frac{f(x)}{g(x)}\right) = \int_{-\infty}^{\infty} f(x)\log\left(\frac{f(x)}{g(x)}\right)dx.$$

A zero value indicates that the two distributions are identical. In our context, the following theorem delivers the quantification of closeness.

Theorem 5: Let $f(x) = \frac{\theta\beta}{(1+\theta x)^{\beta+1}}$ and $g(x) = \lambda e^{-\lambda x}$. Then the Kullback-Leibler divergence measure for (Lomax : Exponential) pair is given by, $D_{KL}(f:g) = \ln\left(\frac{\beta\theta}{\lambda}\right) - \frac{\beta+1}{\beta} + \frac{\lambda}{\theta(\beta-1)}$.

Proof:. With $f(x)$ and $g(x)$ as given above, we have,

$$\begin{aligned} D_{KL}(f:g) &= \int_0^{\infty} \frac{\theta\beta}{(1+\theta x)^{\beta+1}} \ln\left(\frac{\frac{\theta\beta}{(1+\theta x)^{\beta+1}}}{\lambda e^{-\lambda x}}\right)dx \\ &= \int_0^{\infty} \frac{\theta\beta}{(1+\theta x)^{\beta+1}} \ln\left(\frac{\theta\beta}{(1+\theta x)^{\beta+1}}\right)dx - \int_0^{\infty} \frac{\theta\beta}{(1+\theta x)^{\beta+1}}[\ln(\lambda) - \lambda x]dx \\ &= \ln(\beta\theta) - \frac{\beta+1}{\beta} - \ln(\lambda) + \frac{\lambda}{\theta(\beta-1)} = \ln\left(\frac{\beta\theta}{\lambda}\right) - \frac{\beta+1}{\beta} + \frac{\lambda}{\theta(\beta-1)}. \end{aligned}$$

□

The measure does depend on the ratio of scale parameters $\frac{\lambda}{\theta}$. When $\frac{\lambda}{\beta} = \theta$ and thus $f(x)$ is the pdf of a Lomax$(\beta, \frac{\lambda}{\beta})$, we have $D_{KL}(f:g) = \frac{\beta}{\beta-1} - \frac{\beta+1}{\beta} = \frac{1}{\beta(\beta-1)}$. Clearly, the convergence to zero is of order $\frac{1}{\beta^2}$ which again reaffirms our assertion of the considerable closeness of the two distributions.

2.3. Khattree-Bahuguna's skewness for Lomax/Pareto and exponential distributions

Khattree and Bahuguna (2019) recently defined a measure of skewness of a probability distribution which for a quick reference, we state below.

Definition 6: Let X be a random variable possibly assumed to have been centered by mean and let $F(\cdot)$ be its cumulative distribution function. The Khattree-Bahuguna's skewness of X is defined as

$$\delta = \frac{\int_0^1\left(\frac{F^{-1}(\alpha)+F^{-1}(1-\alpha)}{2}\right)^2 d\alpha}{\int_0^1\left(\frac{F^{-1}(\alpha)+F^{-1}(1-\alpha)}{2}\right)^2 d\alpha + \int_0^1\left(\frac{F^{-1}(\alpha)-F^{-1}(1-\alpha)}{2}\right)^2 d\alpha}.$$

When the second moment exists, the above simplifies to

$$\delta = \frac{1}{2}\left[1 + \frac{\int_0^1 F^{-1}(\alpha)F^{-1}(1-\alpha)d\alpha}{\mu_2}\right], \tag{3}$$

where μ_2 is the second central moment of the distribution. The sample skewness (after the sample has been scaled to have zero mean) can be computed as (see Khattree and Bahuguna, 2019),

Definition 7: Given a random sample of size n consist of observations $x_1, x_2, \ldots, x_n$, let $x_{(1)} \leq x_{(2)} \leq \cdots \leq x_{(n)}$ be the order statistics of $x_1, x_2, \ldots, x_n$ after being centered by their sample mean. Define $y_i = \frac{x_{(i)}+x_{(n-i+1)}}{2}$ and $w_i = \frac{x_{(i)}-x_{(n-i+1)}}{2}$. The sample Khattree-Bahuguna univariate skewness is then defined as $\hat{\delta} = \frac{\sum y_i^2}{\sum y_i^2+\sum w_i^2}$.

Clearly being the measures of skewness, δ or $\hat{\delta}$ have no reference to any other distribution from which the distance of a given distribution can be measured. Nonetheless, δ is essentially a function of the (inverse) cumulative distribution function of the random variable and hence, intuitively speaking, if the two probability distributions are very similar, we expect that it would be reflected in their respective expressions of Khattree-Bahuguna's skewness. With this in mind, we evaluate the Khattree-Bahuguna's skewness of Lomax and exponential distributions and indicate how the former converges to the later. For this we first give, in Theorem 8, the explicit expressions for the two skewnesses. Theorem 9 then establishes the convergence.

Theorem 8: Let $X_1 \sim \text{Lomax}(\beta, \theta)$ and $X_2 \sim \text{Exp}(\lambda)$. Then the Khattree-Bahuguna's skewness of X_1 and X_2 are respectively given by,

Lomax: $\delta_{X_1} = \frac{1}{2}\left[1 + \left(B(1-\frac{1}{\beta}, 1-\frac{1}{\beta}) - \frac{\beta^2}{(\beta-1)^2}\right)\frac{(\beta-1)^2(\beta-2)}{\beta}\right], \quad \beta > 2;$

Exponential: $\delta_{X_2} = 1 - \frac{\pi^2}{12}$.

Note that neither δ_{X_1} nor δ_{X_2} depends on respective scale parameters.

Proof: Proof involves the corresponding evaluations of the expression given in (3). For exponential distribution this evaluation is straightforward. Thus, we will work out the details only for the Lomax distribution. We know that for the Lomax random variable X_1, $F(x_1) = 1-(1+\theta x_1)^{-\beta}$ and hence $F^{-1}(\alpha) = \frac{(1-\alpha)^{-\frac{1}{\beta}}-1}{\theta}$. For $\beta > 2$, the mean and variance of X_1 are $\frac{1}{\theta(\beta-1)}$ and $\frac{\beta}{\theta^2(\beta-1)^2(\beta-2)}$, respectively. Therefore, $\int_0^1 \left[F^{-1}(\alpha)-\mu\right]\left[F^{-1}(1-\alpha)-\mu\right] d\alpha = \int_0^1 \left[\frac{(1-\alpha)^{-\frac{1}{\beta}}-\frac{\beta}{\beta-1}}{\theta}\right]\left[\frac{\alpha^{-\frac{1}{\beta}}-\frac{\beta}{\beta-1}}{\theta}\right] d\alpha = \frac{1}{\theta^2}\left[B(1-\frac{1}{\beta}, 1-\frac{1}{\beta}) - \frac{\beta^2}{(\beta-1)^2}\right]$. Thus,

$$\delta_{X_1} = \frac{1}{2}\left[1 + \left(B(1-\frac{1}{\beta}, 1-\frac{1}{\beta}) - \frac{\beta^2}{(\beta-1)^2}\right)\frac{(\beta-1)^2(\beta-2)}{\beta}\right], \quad \beta > 2.$$

□

Theorem 9: As $\beta \to \infty$, Khattree-Bahuguna's skewness of Lomax distribution approaches that of exponential distribution.

Proof: We have,

$$\begin{aligned}
\lim_{\beta\to\infty}\delta_{X_1} &= \lim_{\beta\to\infty}\frac{1}{2}\left[1+\left(B(1-\frac{1}{\beta},1-\frac{1}{\beta})-\frac{\beta^2}{(\beta-1)^2}\right)\frac{(\beta-1)^2(\beta-2)}{\beta}\right]\\
&= \lim_{u\to 1}\frac{1}{2}\left[1+\left(B(u,u)\frac{u^2}{(1-u)^2}-\frac{1}{(1-u)^2}\right)(2u-1)\right] \qquad \left(u=1-\frac{1}{\beta}\right)\\
&= \frac{1}{2}+\frac{1}{2}\lim_{u\to 1}\frac{u^2B(u,u)-1}{(1-u)^2}\\
&= \frac{1}{2}+\frac{1}{2}\lim_{u\to 1}\frac{2uB(u,u)+u^2B'(u,u)}{2(u-1)}\\
&= \frac{1}{2}+\frac{1}{2}\lim_{u\to 1}\frac{2B(u,u)+2uB'(u,u)+2uB'(u,u)+u^2B''(u,u)}{2}\\
&= \frac{1}{2}+\frac{1}{2}\cdot\frac{2+2(-2)+2(-2)+8-\frac{\pi^2}{3}}{2}=1-\frac{\pi^2}{12}=\delta_{X_2}.
\end{aligned}$$

The calculation of the last limit in above is rather complex and therefore is evaluated by using the Lemmas 21 and 22 which are given in ANNEXURE A. ☐

3. Relationships of Multivariate Lomax to Other Distributions in Exponential Family of Distributions

We now again consider the multivariate Lomax distribution of Nayak (1987). As pointed out by him, multivariate Lomax distribution is related to many other multivariate distributions such as Mardia's Pareto type I, Burr, Logistic, Cook-Johnson's uniform (alternatively called Clayton copula), and F. We will further observe here that Lomax distribution is also related to a few other univariate distributions of exponential family through the linear combinations of multivariate Lomax or via one-to-one transformation from univariate Lomax. We convey these facts via following Theorems. We skip the proofs for the sake of brevity.

Theorem 10: Let $(X_1, X_2, \cdots, X_k)$ follow a k-dimensional multivariate Lomax distribution as given by probability density function in (2). Define $X^* = \sum_{i=1}^{k}\theta_i X_i$. Then X^* is distributed as beta type II (also called inverted beta or beta prime) with shape parameters k and β, and therefore its probability density function is given by,

$$f^*(x^*) = \frac{x^{*k-1}(1+x^*)^{-(\beta+k)}}{B(k,\beta)}. \tag{4}$$

Theorem 11: Let $(X_1, X_2, \cdots, X_k)$ be k random variables jointly following multivariate Lomax distribution as defined in (2). Define $Y^* = \beta\sum_{i=1}^{k}\theta_i X_i$. Then

(i) Y^* is distributed as beta type II with shape parameters k and β and scale parameter β and therefore,

$$f(y^*) = \frac{y^{*k-1}(1+\frac{y^*}{\beta})^{-(\beta+k)}}{\beta^k B(k,\beta)}. \tag{5}$$

(ii) as $\beta \to \infty$, the distribution of Y^* approaches gamma distribution with shape parameter k and scale parameter 1.

Theorem 12: Let X be a random variable following Lomax distribution defined in (1). Define $Y = (\theta X/d)^{1/c}$. Then

(i) the probability density function of Y is a Burr density with shape parameters β and c, and rate parameter $d^{\frac{1}{c}}$, that is, $f(y) = \frac{cd\beta y^{c-1}}{(1+dy^c)^{\beta+1}}, y > 0$. See Nayak (1987);

(ii) let $W = \beta^{\frac{1}{c}} Y$, then the probability density function of W is $f(w) = \frac{cdw^{c-1}}{(1+\frac{d}{\beta}w^c)^{\beta+1}}, w > 0$;

(iii) as $\beta \to \infty$, the distribution of W approaches a Weibull distribution with shape parameter c and rate parameter d, density which is given by, $f(w) = cdw^{c-1}e^{-dw^c}, w > 0$.

It may be noted that when $c = 2$ and $d = \theta$, the distribution of W approaches to that of a Rayleigh random variable.

4. A Generalized Double Pareto-Laplace Connection

We take the previous discussion one step further by making the support of the respective random variables as the entire real line. Specifically, the role of Lomax will now be played by double Pareto distribution and that of exponential is now played by the generalized Laplace distribution. The probability density function for the two are given below.

$$\text{Double Pareto}: \quad f(x) = \begin{cases} \frac{\theta_1\theta_2\beta}{(\theta_1+\theta_2)(1-\theta_1 x)^{\beta+1}}, & \text{if } x \leq 0,\ \beta, \theta_1, \theta_2 > 0 \\ \\ \frac{\theta_1\theta_2\beta}{(\theta_1+\theta_2)(1+\theta_2 x)^{\beta+1}}, & \text{if } x > 0,\ \beta, \theta_1, \theta_2 > 0. \end{cases}$$

$$\text{Generalized Laplace}: \quad g(x) = \frac{1}{\sigma}\frac{\kappa}{1+\kappa^2} \begin{cases} e^{\frac{1}{\sigma\kappa}(x-\mu)}, & \text{if } x \leq 0,\ \sigma, \kappa > 0,\ \mu \in \mathbb{R} \\ \\ e^{-\frac{\kappa}{\sigma}(x-\mu)}, & \text{if } x > 0,\ \sigma, \kappa > 0,\ \mu \in \mathbb{R}. \end{cases}$$

A particular connection between the above two distributions is given by Kotz *et al.* (2001) who indicate that a double Pareto random variable can be generated by taking the ratio of two independent Laplace variates. Two especially attractive properties of double Pareto distribution are (i) its Laplace like spike of density function at zero and (ii) its Student's t-like heavy tails. See Armagan, Dunson and Lee (2013) and Pal, Khare and Hobert (2017). The double Pareto has recently received considerable attention as a choice of the prior distribution in the context of Bayesian robust shrinkage estimation (Armagan *et al.*, 2013) and thus its connection to generalized Laplace distribution is of special interest.

We will show that as in the case of Lomax-exponential connection, the properties (i) and (ii) stated above can be interpreted through the double Pareto-Laplace connection. We will also demonstrate that the double Pareto can be represented as a mixture of several t-distributions. However, to do so, we must first define a three-parameter generalized double Pareto distribution, which allows the possibility of asymmetry in the density. For this, it is convenient to pursue an approach where the bivariate Lomax distribution plays a central role. This is given by Theorem 13 that follows.

4.1. Three-parameter generalized double Pareto distribution

To set the stage, we observe that similar to the case of Laplace, a classical symmetric double Pareto distribution can be obtained by the difference of two independent Lomax variates. In order to incorporate asymmetric double Pareto distributions, we propose a three-parameter generalized double Pareto distribution defined via a bivariate Lomax distribution, where the variates naturally exhibit dependence.

Theorem 13: Assume that X_1 and X_2 are jointly distributed as the bivariate Lomax variables with parameters $\beta, \theta_1, \theta_2$ as given in (2). Then the probability density function of $X = X_2 - X_1$ is given by,

$$h_1(x) = \begin{cases} \frac{\theta_1\theta_2\beta}{(\theta_1+\theta_2)(1-\theta_1 x)^{\beta+1}}, & \text{if } x \leq 0,\ \beta, \theta_1, \theta_2 > 0 \\ \\ \frac{\theta_1\theta_2\beta}{(\theta_1+\theta_2)(1+\theta_2 x)^{\beta+1}}, & \text{if } x > 0,\ \beta, \theta_1, \theta_2 > 0. \end{cases} \tag{6}$$

Proof: Let $X = X_2 - X_1$ and $Y = \theta_1 X_1 + \theta_2 X_2$, then the Jacobian of the transformation is $|J| = \frac{1}{\theta_1+\theta_2}$. The joint probability density of X and Y is thus, $h(x, y) = \frac{\theta_1\theta_2\beta(\beta+1)}{(\theta_1+\theta_2)(1+y)^{\beta+2}}, \quad y > \max\{\theta_2 x, -\theta_1 x\}$. Integrating over y gives the marginal density function of X as given above. □

Note that for bivariate Lomax variate (X_1, X_2), $X = X_1 + X_2$ does not follow a Lomax distribution. We state the result about this sum as follows.

Theorem 14: Let (X_1, X_2) follow a bivariate Lomax distribution with parameters $\beta, \theta_1, \theta_2$ and let $X = X_1 + X_2$. Assuming $\theta_1 \neq \theta_2$, the probability density function of X is

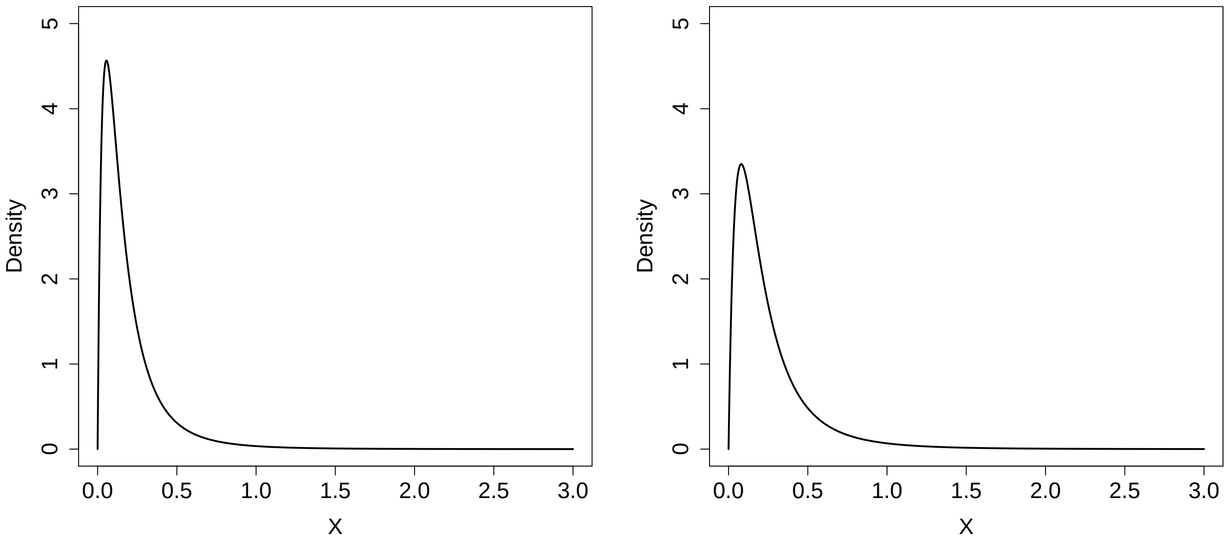
$$h_2(x) = \frac{\theta_1\theta_2\beta}{(\theta_2 - \theta_1)} \left[\frac{1}{(1+\theta_1 x)^{\beta+1}} - \frac{1}{(1+\theta_2 x)^{\beta+1}} \right], \quad x > 0.$$

When $\theta_1 = \theta_2 = \theta$, the probability density function of X is beta type II distribution given in (4) with shape parameters $k = 2$ and β, and rate parameter θ.

Proof follows by letting $Y = X_1$ and $X = X_1 + X_2$ and integrating over y. For the case $\theta_1 = \theta_2 = \theta$, the distribution of $Y = X_1 + X_2$ is a beta type II distribution with rate parameter θ, as already stated in Theorem 10. Figure 4 gives the density plots for the sum of the components of a bivariate Lomax vector with parameters $\beta = 4, \theta_1 = 2.5, \theta_2 = 5$ in (a) and parameters $\beta = 4, \theta_1 = \theta_2 = 2.5$ in (b). Clearly, both density plots exhibit a very different shape compared to Lomax distribution. The nth raw moment of $X = X_1 + X_2$ is given by $E(X^n) = \frac{\theta_1\theta_2 n!}{(\theta_2-\theta_1)(\beta-1)\cdots(\beta-n)} \left[\frac{1}{\theta_1^{n+1}} - \frac{1}{\theta_2^{n+1}} \right], n < \beta$, and $E(X^n) = \infty$ when $n \geq \beta$. Specifically, for $\beta > 2$, we have, $E(X) = \frac{\theta_1+\theta_2}{\theta_1\theta_2(\beta-1)}$ and $\text{Var}(X) = \frac{\beta\theta_1^2+\beta\theta_2^2+2\theta_1\theta_2}{\theta_1^2\theta_2^2(\beta-1)^2(\beta-2)}$.

Definition 15: We define a generalized double Pareto distribution as that for a real-valued random variable X whose probability density function is given by (6). We will denote it by $\mathbf{GDP}(\beta, \theta_1, \theta_2)$.

Thus, in the case of generalized double Pareto distribution, β is the shape parameter and θ_1, θ_2 are two rate parameters. It is clear that if X is $\mathbf{GDP}(\beta, \theta_1, \theta_2)$ then $-X$ is



(a) $\beta = 4, \theta_1 = 2.5, \theta_2 = 5.0$

(b) $\beta = 4, \theta_1 = \theta_2 = 2.5$ **(Beta type II: shape parameters** $k = 2$ **and** $\beta = 4$ **and rate parameter** $\theta = 2.5$**)**

Figure 4: Examples of sum variable for bivariate Lomax distribution

$\mathbf{GDP}(\beta, \theta_2, \theta_1)$. When $\theta_1 = \theta_2 = \theta$, the generalized double Pareto distribution reduces to the classical symmetric double Pareto distribution, denoted by $\mathbf{CDP}(\beta, \theta)$ and with the density function $f(x) = \frac{\theta}{2}\frac{1}{(1+\theta|x|)^{\beta+1}}, -\infty < x < \infty, \beta, \theta > 0$. Letting $\theta = \frac{1}{\xi}$ and scaling X by shape parameter β in this density results in the density of the double Pareto distribution defined by Armagan *et al.* (2013) with probability density function, $f(x) = \frac{1}{2\xi}\left(1 + \frac{|x|}{\beta\xi}\right)^{-(\beta+1)}, -\infty < x < \infty, \beta, \xi > 0$.

Figure 5 contrasts the behavior of the density function of the generalized double Pareto distribution random variables for the symmetric ($\theta_1 = \theta_2$) and asymmetric ($\theta_1 \neq \theta_2$) cases.

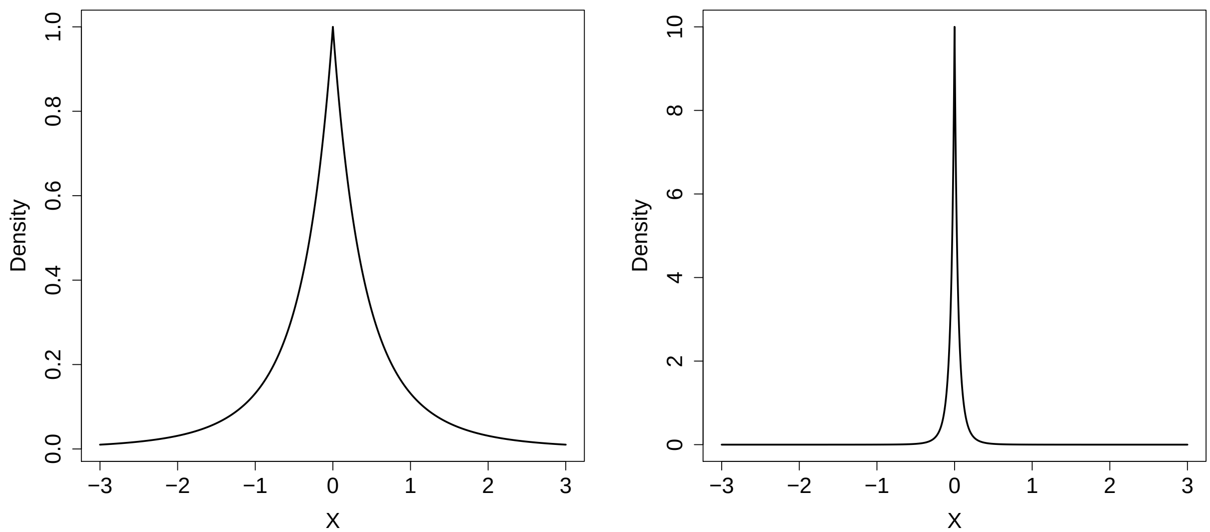
The cumulative distribution function of $\mathbf{GDP}(\beta, \theta_1, \theta_2)$ is given by

$$F(x) = \begin{cases} \frac{\theta_2}{(\theta_1+\theta_2)(1-\theta_1 x)^\beta}, & \text{if } x \leq 0,\ \beta, \theta_1, \theta_2 > 0 \\ 1 - \frac{\theta_1}{(\theta_1+\theta_2)(1+\theta_2 x)^\beta}, & \text{if } x > 0,\ \beta, \theta_1, \theta_2 > 0. \end{cases}$$

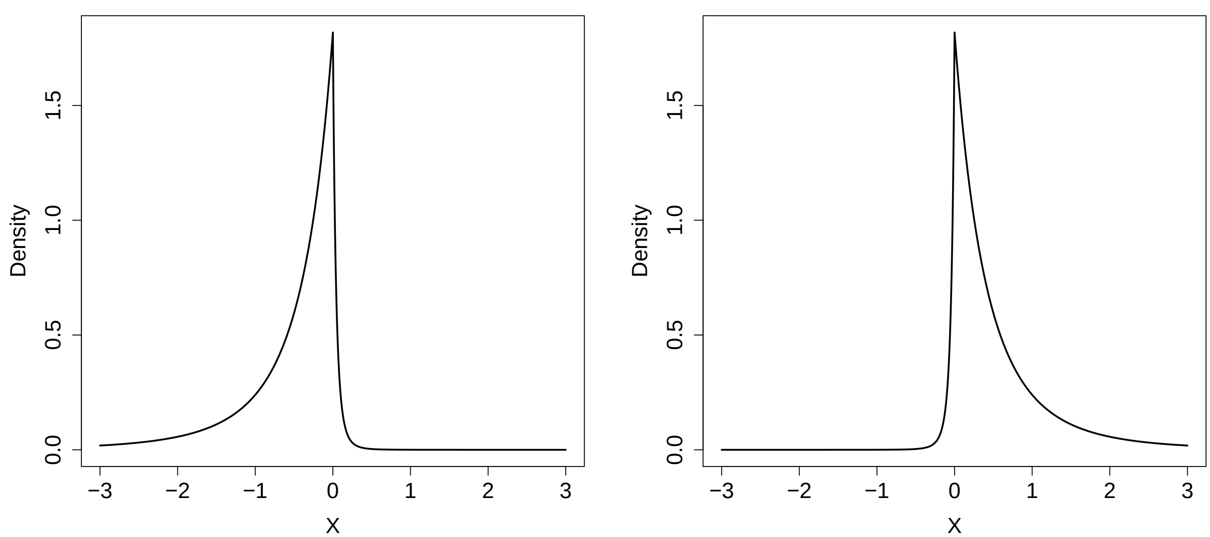
It is easy to observe that $P(X \leq 0) = \frac{\theta_2}{\theta_1+\theta_2}$ and $P(X > 0) = \frac{\theta_1}{\theta_1+\theta_2}$, which can be interpreted as the weights of the two rate parameters. There is a larger proportion of negative values whenever $\theta_2 > \theta_1$. This observation is evident in Figure 5 (c).

The quantile function $F^{-1}(\alpha)$ of $\mathbf{GDP}(\beta, \theta_1, \theta_2)$ is,

$$F^{-1}(\alpha) = \begin{cases} \frac{1}{\theta_1}\left[1 - \left(\frac{\theta_2}{\alpha(\theta_1+\theta_2)}\right)^{\frac{1}{\beta}}\right], & \text{if } 0 < \alpha \leq \frac{\theta_2}{\theta_1+\theta_2}, \\ \frac{1}{\theta_2}\left[\left(\frac{\theta_1}{(1-\alpha)(\theta_1+\theta_2)}\right)^{\frac{1}{\beta}} - 1\right], & \text{if } \frac{\theta_2}{\theta_1+\theta_2} < \alpha < 1. \end{cases} \tag{7}$$



(a) $\beta = 4, \theta_1 = \theta_2 = 0.5$ **(b)** $\beta = 4, \theta_1 = \theta_2 = 5$



(c) $\beta = 4, \theta_1 = 0.5, \theta_2 = 5$ **(d)** $\beta = 4, \theta_1 = 5, \theta_2 = 0.5$

Figure 5: Examples of symmetric (a)-(b) and asymmetric (c)-(d) double Pareto distributions

The nth raw moment of $\mathbf{GDP}(\beta, \theta_1, \theta_2)$ is given by $E(X^n) = \frac{n!}{(\theta_1+\theta_2)(\beta-1)\cdots(\beta-n)}\left[\frac{(-1)^n\theta_2}{\theta_1^n} + \frac{\theta_1}{\theta_2^n}\right]$, if $n < \beta$ and $E(X^n) = \infty$ when $n \geq \beta$. We defer this straightforward yet a bit tedious calculation to ANNEXURE B. When $\beta > 2$, $E(X) = \frac{\theta_1-\theta_2}{\theta_1\theta_2(\beta-1)}$ and $\mathrm{Var}(X) = \frac{\beta\theta_1^2+\beta\theta_2^2-2\theta_1\theta_2}{\theta_1^2\theta_2^2(\beta-1)^2(\beta-2)} = \frac{\beta(\theta_1-\theta_2)^2+2\theta_1\theta_2(\beta-1)}{\theta_1^2\theta_2^2(\beta-1)^2(\beta-2)}$. Also assuming $\beta > 3$, the Pearson's coefficient of skewness is given by,

$$\gamma = \frac{2(\beta+1)\sqrt{(\beta-2)}(\theta_1-\theta_2)[\beta\theta_1^2+\beta\theta_2^2+(\beta-3)\theta_1\theta_2]}{(\beta-3)(\beta\theta_1^2+\beta\theta_2^2-2\theta_1\theta_2)^{3/2}}, \qquad \beta > 3.$$

Clearly, $\gamma = 0$ for the symmetric double Pareto as in that case, $\theta_1 = \theta_2$. By letting $\frac{\theta_1}{\theta_2} = \kappa$, we can further simplify the Pearson's γ as $\gamma = \frac{2(\beta+1)\sqrt{(\beta-2)}(\kappa-1)[\beta\kappa^2+\beta+(\beta-3)\kappa]}{(\beta-3)(\beta\kappa^2+\beta-2\kappa)^{3/2}}$. When $\kappa \to 0$, that is, when $\theta_1 \ll \theta_2$, the skewness approaches to that of negative univariate Lomax distribution. That is,

$$\lim_{\kappa\to 0}\gamma = \frac{-2(\beta+1)}{\beta-3}\sqrt{\frac{\beta-2}{\beta}}.$$

Similarly, let $\frac{\theta_2}{\theta_1} = \kappa^*$, we have $\gamma = \frac{2(\beta+1)\sqrt{(\beta-2)}(1-\kappa^*)[\beta\kappa^{*2}+\beta+(\beta-3)\kappa^*]}{(\beta-3)(\beta\kappa^{*2}+\beta-2\kappa^*)^{3/2}}$. When $\kappa^* \to 0$, that is, $\theta_2 \ll \theta_1$, the skewness approaches that of positive univariate Lomax distribution. Specifically, $\lim\limits_{\kappa^*\to 0}\gamma = \frac{2(\beta+1)}{\beta-3}\sqrt{\frac{\beta-2}{\beta}}$.

By using (7), Khattree-Bahuguna's skewness is evaluated to be

$$\delta = \frac{1}{2}\left[1 + \frac{2\mathcal{I}_1 + \mathcal{I}_2 - \mu^2}{\mu_2}\right],$$

where

$$\mathcal{I}_1 = \frac{d_1}{\theta_1\theta_2}\left[\frac{1}{\beta-1} - \left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{(\beta-1)(\beta-2)}\right]$$

$$\mathcal{I}_2 = \frac{1}{\min(\theta_1,\theta_2)^2}\left\{d_2^{\frac{2}{\beta}}\left[B\left(d_2; 1-\frac{1}{\beta}, 1-\frac{1}{\beta}\right) - B\left(d_1; 1-\frac{1}{\beta}, 1-\frac{1}{\beta}\right)\right] + 2\left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{\beta-1}d_1 - 2\frac{\beta}{\beta-1}d_2 + d_2 - d_1\right\}$$

$$\mu = \frac{\theta_1-\theta_2}{\theta_1\theta_2(\beta-1)}, \qquad \mu_2 = \frac{\beta(d_2-d_1)^2}{d_1^2d_2^2(\beta-1)^2(\beta-2)} + \frac{2}{\theta_1\theta_2(\beta-1)(\beta-2)}$$

$$d_1 = \min\left(\frac{\theta_1}{\theta_1+\theta_2}, \frac{\theta_2}{\theta_1+\theta_2}\right), \qquad d_2 = 1 - d_1,$$

and $B(x; a, b) = \int_0^x t^{a-1}(1-t)^{b-1}dt$. For the symmetric double Pareto distribution, $d_1 = d_2 = \frac{1}{2}$, and then $\mathcal{I}_1 = \frac{-1}{\theta_1\theta_2(\beta-1)(\beta-2)}, \mathcal{I}_2 = 0$. Accordingly, $\delta = 0$. Detailed and cumbersome calculations for all of these facts are deferred to ANNEXURE B.

Like Laplace distribution, the probability density function shown in Figure 5 is also spiked. Thus, it is natural to explore any double Pareto-Laplace connection by using the similar technique as used earlier to obtain the Lomax-exponential connection. In fact, this double Pareto-Laplace connection was indirectly hinted by Armagan *et al.* (2013) where their Laplace prior (Bayesian lasso) was treated as the limiting case of double Pareto prior. The following theorem formalizes it.

Theorem 16: Let X be a $\mathbf{GDP}(\beta, \theta_1, \theta_2)$ random variable. Define $Y = \beta X$. Then

(i) the distribution of Y is $\mathbf{GDP}(\beta, \frac{\theta_1}{\beta}, \frac{\theta_2}{\beta})$;

(ii) as $\beta \to \infty$, the distribution of Y approaches a two-parameter Laplace distribution with parameters θ_1, θ_2, ($\mathbf{Laplace}(\theta_1, \theta_2)$) with probability density function

$$g^*(y) = \begin{cases} \frac{\theta_1\theta_2}{\theta_1+\theta_2} e^{\theta_1 y}, & \text{if } y \leq 0, \\ \frac{\theta_1\theta_2}{\theta_1+\theta_2} e^{-\theta_2 y}, & \text{if } y > 0. \end{cases} \tag{8}$$

Therefore, we surmise that $\mathbf{GDP}(\beta, \frac{\theta_1}{\beta}, \frac{\theta_2}{\beta})$ can be approximated by $\mathbf{Laplace}(\theta_1, \theta_2)$ if the shape parameter β is large. Moreover, both $\mathbf{GDP}(\beta, \frac{\theta_1}{\beta}, \frac{\theta_2}{\beta})$ and $\mathbf{Laplace}(\theta_1, \theta_2)$ have the probability density spiked at zero. As pictorially demonstrated in Figure 6, as β increases, the density plot of $\mathbf{GDP}(\beta, \theta_1, \theta_2)$ approaches that of Laplace distribution. Also with reparameterization $\theta_1 = \frac{1}{\sigma\kappa}$ and $\theta_2 = \frac{\kappa}{\sigma}$ the Laplace density in (8) reduces to the form introduced by Hinkley and Revankar (1977) with zero location parameter ($\mu = 0$) as

$$g^*(y) = \frac{1}{\sigma}\frac{\kappa}{1+\kappa^2} \begin{cases} e^{\frac{1}{\sigma\kappa} y}, & \text{if } y \leq 0,\ \sigma, \kappa > 0 \\ e^{-\frac{\kappa}{\sigma} y}, & \text{if } y > 0,\ \sigma, \kappa > 0. \end{cases}$$

4.2. A representation of double Pareto distribution

The next result shows that any symmetric double Pareto random variable can be thought of as a Student's t random variable when scaled by an independent Lomax random variate. This results in a symmetric yet a Student's t-like heavy tails of double Pareto distribution. Due to this heavy-tail property, it has been widely used in Bayesian shrinkage as a choice of prior. More formally,

Theorem 17: A symmetric double Pareto random variate X with shape parameter $(\nu - 1)$ and scale parameter $\sqrt{\nu}$ can be represented as

$$X \stackrel{d}{=} \sqrt{Y} T, \tag{9}$$

where the random variable Y has a standard Lomax distribution with shape parameter $\beta > 0$ and T has an independent Student's t distribution with degrees of freedom $\nu = 2\beta + 1$. The notation $\stackrel{d}{=}$ indicates the equivalence of distributions.

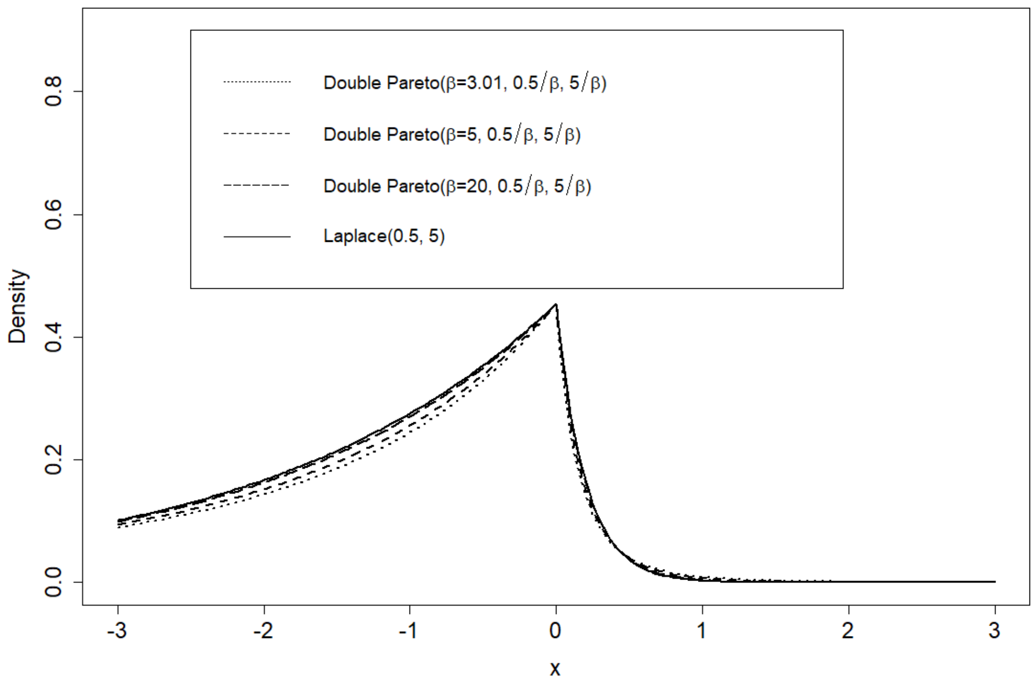


Figure 6: Density plots of GDP$(\beta, \frac{0.5}{\beta}, \frac{5}{\beta})$ **with changing parameter** $\beta = 3.01, 5, 20$ **and Laplace**$(\theta_1 = 0.5, \theta_2 = 5)$

Proof:. Given $Y \sim$ Lomax$(\beta, 1)$ and $T \sim$ Student's t with degrees of freedom ν, the probability density function of X, is

$$\begin{aligned} f_X(x) &= \int_0^\infty f_T(\frac{x}{\sqrt{y}})\frac{1}{\sqrt{y}} f_Y(y)dy = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})}\int_0^\infty \frac{1}{(1+\frac{x^2}{\nu y})^{\frac{\nu+1}{2}}}\frac{1}{\sqrt{y}}\frac{\beta}{(1+y)^{\beta+1}}dy \\ &= \frac{2\beta\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})}\int_0^\infty \frac{1}{(1+\frac{x^2}{\nu u^2})^{\frac{\nu+1}{2}}(1+u^2)^{\beta+1}}du \qquad (u = \sqrt{y}). \\ &= \frac{2\beta}{\sqrt{\nu}B(\frac{\nu}{2}, \frac{1}{2})}\int_0^\infty \frac{1}{(1+\frac{x^2}{\nu u^2})^{\frac{\nu+1}{2}}(1+u^2)^{\beta+1}}du. \end{aligned}$$

The above integral is difficult to evaluate for general values of ν and β. However, when $\nu = 2\beta + 1$ and hence $\frac{\nu+1}{2} = \beta + 1$, (as stated in Theorem) simplification occurs. In this case, by using the Lemma 23 which is stated and proved in ANNEXURE, we have,

$$f_X(x) = \frac{(\nu-1)}{\sqrt{\nu}B(\frac{\nu}{2}, \frac{1}{2})}\frac{B(\frac{\nu+1}{2} - \frac{1}{2}, \frac{1}{2})}{2(1+\frac{|x|}{\sqrt{\nu}})^{2\cdot\frac{\nu+1}{2}-1}} = \frac{1}{2\sqrt{\nu}}\frac{(\nu-1)}{(1+\frac{|x|}{\sqrt{\nu}})^{\nu}}, \quad -\infty < x < \infty,$$

which is a symmetric double Pareto distribution with shape parameter $(\nu - 1)$ and scale parameter $\sqrt{\nu}$. □

Theorem 18: A symmetric double Pareto random variable X with shape parameter $(\nu-1)$ and scale parameter ν can be represented as $X \stackrel{d}{=} \sqrt{\nu Y}T$, where the random variable Y has a standard Lomax distribution with shape parameter $\beta > 0$ and T has an independent Student's t distribution with degrees of freedom $\nu = 2\beta + 1$. Moreover, as $\nu \to \infty$, X approaches the standard Laplace distribution.

Proof: Through the linear transformation used on Theorem 17, we easily obtain the probability density function of X as $f_X(x) = \frac{1}{2\nu}\frac{\nu-1}{(1+\frac{|x|}{\nu})^\nu}, -\infty < x < \infty$. Clearly $\lim_{\nu\to\infty} f_X(x) = \frac{1}{2}e^{-|x|}, -\infty < x < \infty$. $\square$

The above two results actually describe another remarkable feature of double Pareto-Laplace connection where Laplace distribution also has the similar representation but of mixture of normal distributions instead of Student's t. For completeness and for the sake of comparison, we restate the representation of Laplace distribution mentioned in Kotz *et al.* (2001), in the following result.

Theorem 19: A standard classical (symmetric) Laplace random variable X has the representation $X \stackrel{d}{=} \sqrt{2W}Z$, where the random variables W and Z are independent and have the standard exponential and normal distributions, respectively.

The above result establishes the Laplace distribution as a mixture of normal distribution with a scale parameter having exponential distribution. Actually, the above proposition can be viewed as the limiting case of both sides of (9) as $\beta, \nu \to \infty$ where the double Pareto, Lomax and Student's t distributions respectively approach Laplace, exponential and normal. This is perhaps the reason as to why the double Pareto appears to be a better choice for prior distribution than the Laplace distribution in Bayesian shrinkage estimation when we require a prior with heavy tails.

5. Random Number Generation from Double Pareto Distribution

The R (R Core Team, 2019) package **NonNorMvtDist** (Lun and Khattree, 2020) is a recent versatile package which implements the simulation and probability computations for a large number of non-normal multivariate distributions including the Lomax. See Figure 10. By Theorem 13 and by using the aforementioned package, random numbers from the **GDP**$(\beta, \theta_1, \theta_2)$ can be easily generated in barely two steps as follows.

1. Generate a sample of size n bivariate Lomax random vector (X_1, X_2) with shape parameter β and the vector of rate parameters (θ_1, θ_2) using the function `rmvlomax()`.
2. Return $X = X_2 - X_1$.

As an example, we generate **GDP**$(3.5, 1.5, 5)$ of size 5000, using the following R code.

```
library(NonNorMvtDist)
beta = 3.5; theta1 = 1.5; theta2 = 5
set.seed(2020)
bivLomax = rmvlomax(n = 5000, parm1 = beta, parm2 = c(theta1, theta2))
x = bivLomax[,2] - bivLomax[,1]
hist(x, breaks=30, freq = FALSE)
```

With suitable similar extensions of other distributions as shown in Figure 10, one can implement random number generations in other cases as well.

6. An Illustrative Biomedical Application of Bivariate Lomax distribution

We consider a data set from a breast cancer study from University of California Irvine Machine Learning Repository to highlight the usefulness of Lomax distribution for modeling the non-negative skewed data. The data are attributed to Patrício *et al.* (2018) and consist of nine quantitative clinical features (age, BMI, glucose, insulin, HOMA, leptin, adiponectin, resistin, and MCP-1), and a binary classification variable (Patients *vs.* Healthy controls) observed for 64 patients with breast cancer and 52 healthy control subjects recruited from the University Hospital Centre of Coimbra. With substantial skewness present in all clinical features and hence the lack of normality assumptions, Crisóstomo *et al.* (2016) analyzed the data by applying the nonparametric methods (specifically the Kruskal-Wallis test). Among other things, insulin was identified as a significant discriminator between the two groups but only for corresponding subsets with BMI $> 25\text{kg/m}^2$. However, for the group with BMI $\leq 25\text{kg/m}^2$, the significance of insulin seemed inconclusive. Nonparametric approach with relatively low power may be one of the reason for not so clear a conclusion.

We will choose insulin as the variable of interest for our work. Instead of choosing a nonparametric approach which usually has low power especially when the sample size is not very large, and forgoing the normality based methods due to absence of normality, we here propose a Lomax model for this data. The high skewness in the data, as shown in Figures 7 (a) and (b) for healthy group and the breast cancer group respectively and the shape of the distributions justify our use of this model. Using BMI as a matching variable to match pairs of one healthy subject and one breast cancer subject, from each of the two groups, we obtain the bivariate data on insulin measurements for $n = 52$ such pairs. These values after discarding unmatched subjects are presented in Table 4 in ANNEXURE. The corresponding R code for this application can be obtained from the authors.

Sample descriptive statistics summary for the respective marginal distributions are given in the columns 2-4 of Table 2. With skewed marginal distributions as shown in Figure 7 (a) and (b), we fit the bivariate Lomax distribution on this data, using the maximum likelihood (ML) approach. This results in, $\hat{\beta} = 283.8444, \hat{\theta}_{health} = 0.000508, \hat{\theta}_{cancer} = 0.000299$. The values of descriptive statistics based on these estimates are given as columns 5-7 of Table 2. The agreement between the sample descriptive statistics and ML based descriptive statistics is quite good, even though standard deviation for the latter is somewhat higher. Also the ML estimates of correlation between the paired variables, is equal to $1/\hat{\beta} = 1/283.8444 = 0.0035$ (the sample correlation = 0.0465), which is low (as we must expect since the patients as well as two samples are independent). The bivariate Lomax seems to fit the data very well. This is further justified by Lomax Q-Q plots given in Figure 8.

Clearly as the ML estimates of mean and variance in Table 2, show, breast cancer group does indeed have not only much higher mean value of insulin, its values also vary much more greatly within the group, compared to those for healthy subject group. Further, we note that large variability is persistent even in the group with BMI $< 25\text{kg/m}^2$ – a fact obscured and hence lost in the nonparametric analysis done by the original authors.

Table 2: Descriptive and theoretical statistics for healthy controls and breast cancer patients, respectively.

	Descriptive Statistics ($n = 52$)					
	Sample			ML-Based		
Group	Mean	SD	Skewness	Mean	SD	Skewness
Healthy controls	6.9340	4.86	2.2765	6.9540	6.98	2.0213
Breast cancer patients	11.8420	10.20	1.4753	11.8194	11.86	2.0213

It may also be mentioned that in view of large $\hat{\beta} = 283.8444$, the exponential model, for insulin levels scaled by $\hat{\beta}$, may be applicable for both groups – healthy as well as the breast cancer group. Thus, the estimates of the two rate parameter are

$$\hat{\lambda}_{health} = \hat{\beta}\hat{\theta}_{health} = 0.1442 \quad (1/\bar{X}_{health} = 0.1441 \text{ if fitting the exponential distribution})$$
$$\hat{\lambda}_{cancer} = \hat{\beta}\hat{\theta}_{cancer} = 0.0849 \quad (1/\bar{X}_{cancer} = 0.0845 \text{ if fitting the exponential distribution})$$

Clearly the estimated value of λ_{cancer} is smaller than that for λ_{health} then again reconfirming the higher levels of insulin for the cancer group.

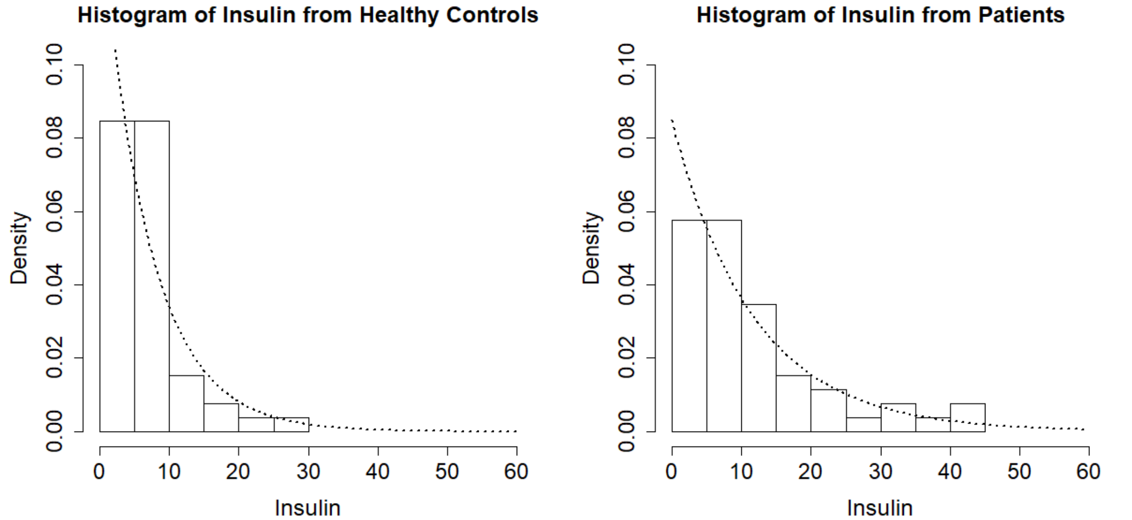
Returning to Lomax context, we may be interested in formally testing the null hypothesis $H_0 : \theta_{health} = \theta_{cancer}$ vs. $H_a : \theta_{health} > \theta_{cancer}$ which aims to test if two mean insulin levels are same for the two groups against the alternative that it is higher for healthy control groups. To do so, we consider the variable representing the difference $Y = X_{cancer} - X_{health}$. With bivariate Lomax assumption on (X_1, X_2) in place, under the null hypothesis Y must follow the symmetric double Pareto distribution.

We take Khattree-Bahuguna's skewness $\hat{\delta}$ as the test statistic. Clearly, under H_0 and hence under symmetry, $\delta = 0$. Thus we must reject H_0 for large values of $\hat{\delta}$, where $\hat{\delta}$ is the estimate of δ obtained by using the sample Khattree-Bahuguna's univariate skewness. For our data, $\hat{\delta} = 0.1164$, which is considerably larger than the one-sided cutoff value $\delta_{0.95} = 0.0742$ calculated under the null hypothesis via large number of simulations ($nsim = 1000$) and by using the R packages of Lun and Khattree (2020). The null hypothesis is thus rejected.

We may also be interested in those pairs with BMI $< 25\text{kg/m}^2$ (this is the data, which original authors had discarded as they analyzed only subjects with BMI $\geq 25\text{kg/m}^2$). Thus, we may try to fit the bivariate Lomax distribution only on $n = 17$ pairs of healthy control

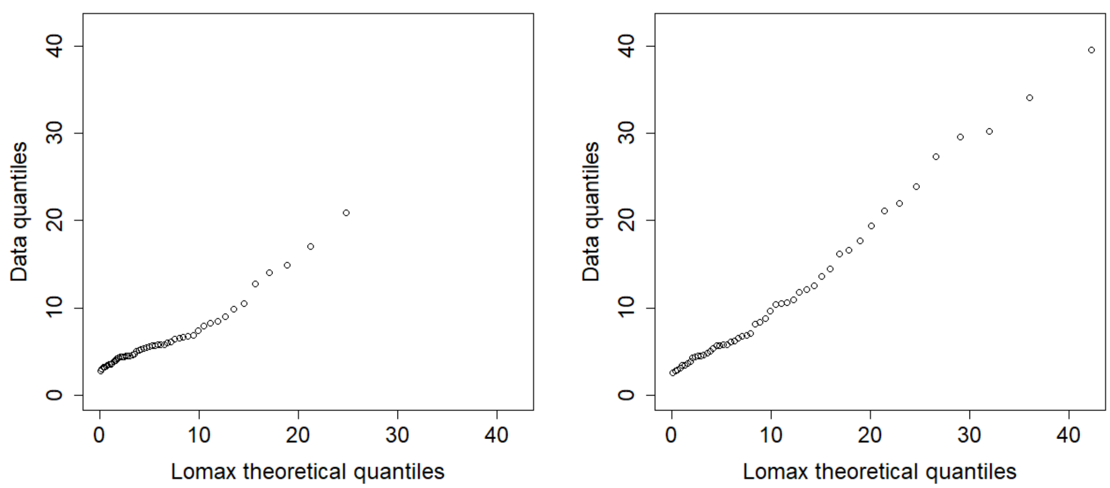
Table 3: Descriptive and theoretical statistics for healthy controls and breast cancer patients with BMI < 25kg/m^2, respectively.

	Descriptive Statistics ($n = 17$)					
	Sample			ML-Based		
Group	Mean	SD	Skewness	Mean	SD	Skewness
Healthy controls	4.4304	1.44	0.5955	4.4100	4.4194	2.0128
Breast cancer patients	8.2047	8.76	2.2348	8.2512	8.2689	2.0128



(a) Healthy Controls (skewness = 2.2765) **(b) Breast Cancer Patients (skewness = 1.4753)**

Figure 7: Empirical distributions of healthy controls and breast cancer patients and density plots (dashed lines) obtained by maximum likelihood estimation based on bivariate Lomax model

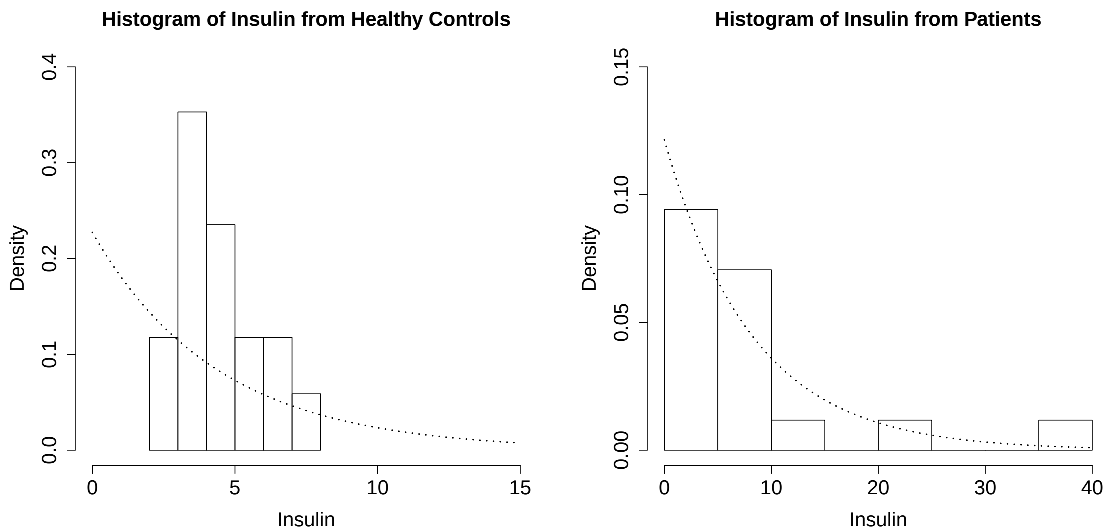


(a) Healthy controls **(b) Breast cancer patients**

Figure 8: Quantile-Quantile plot of insulin levels for healthy controls and for breast cancer patients.

and breast cancer patient, using the maximum likelihood (ML) approach. This results in, $\hat{\beta} = 471.0522, \hat{\theta}_{health} = 0.000482, \hat{\theta}_{cancer} = 0.000258$.

However, as the histograms and summary statistics show, the bivariate Lomax distribution may not fit this subset of data as satisfactorily as the whole data since the group of healthy controls has relatively low skewness while the breast cancer group is highly skewed. See the sample descriptive statistics in Table 3 and histogram in Figure 9. Therefore, as-



(a) Healthy Controls (skewness $= 0.5955$**)** **(b) Breast Cancer Patients (skewness** $= 2.2348$**)**

Figure 9: Empirical distributions of healthy controls and breast cancer patients with BMI < 25**kg/m**^{2} **and density plots (dashed lines) obtained by maximum likelihood estimation based on bivariate Lomax model**

sumptions for the corresponding hypothesis testing are not met and hence hypothesis testing is not performed for this subset of data. It is difficult to determine if the poor fit to Lomax distribution is due to small number ($n = 17$) observations.

7. Concluding Remarks

As the title suggests, this article revolves around connections between Lomax and exponential distributions and between the extensions thereof. Various relationships between multivariate Lomax and several other univariate and multivariate distributions are known to exist and these relationships are graphically reproduced in Figure 10. Via these interrelationships one can possibly establish many more similar connections. For example, generalized double Pareto distribution can also be conveniently obtained via appropriate transformations of many of these bivariate distributions. The same can be said about the representation of generalized double Pareto by a mixture of Student's t distributions. Compared with scale mixture of normal distributions, Generalized double Pareto distribution provides a possibly more robust and more flexible choice of prior in practice, such as robust Bayesian shrinkage estimation and biomedical data modeling.

Data analysis presented here exemplifies the potential applications which distributions presented in this work may have. Comparisons such as that presented in our illustration, require the distributions of random variables which are linear functions of such variables and may not result in nice symmetric distributions with support on the entire real line. Generalized double Pareto distribution with asymmetry, skewness and fat tail is one such distribution which may be a flexible enough choice to accommodate such situations.

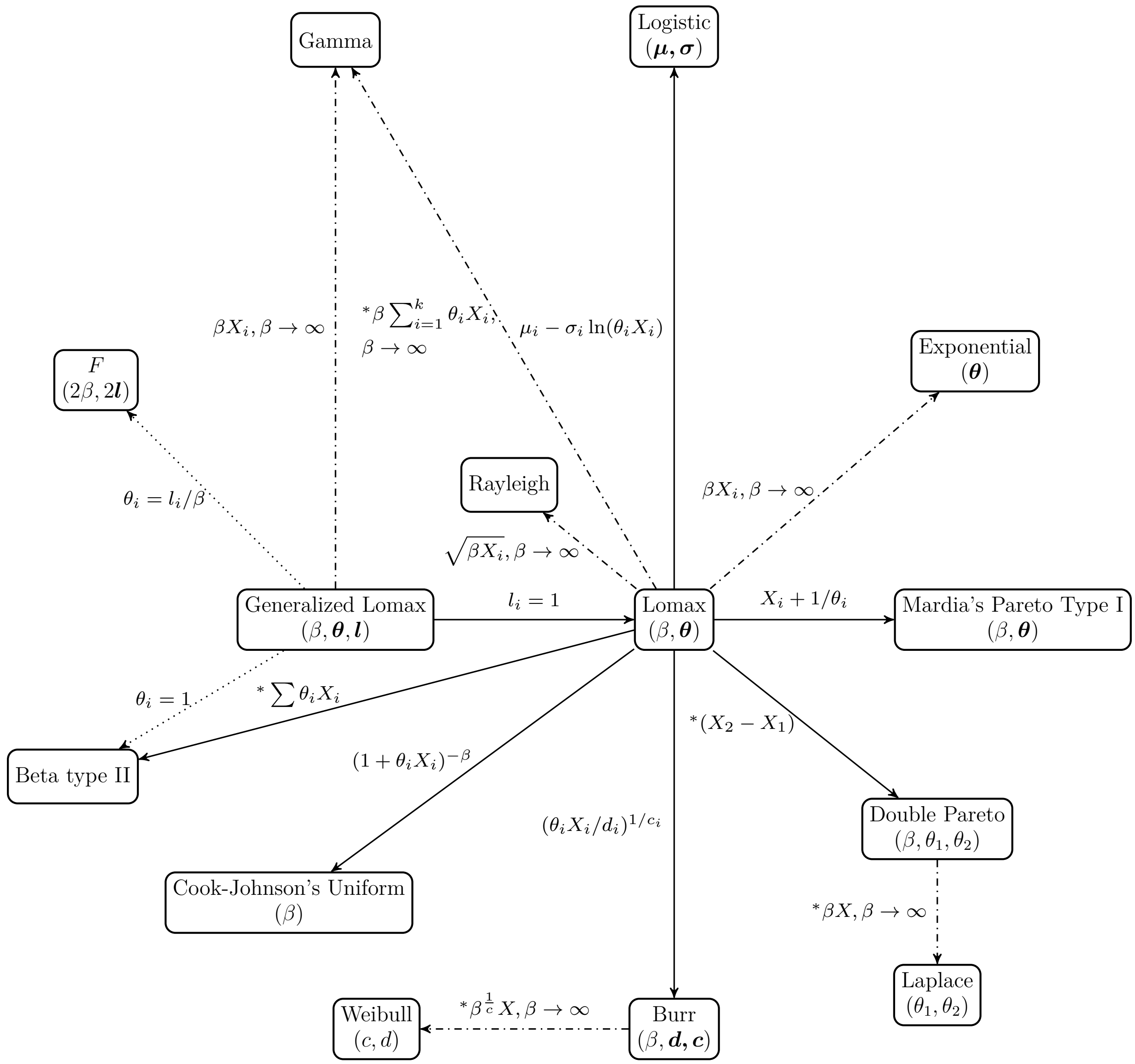


Figure 10: Lomax and its relationships through transformation (solid lines), parameter substitution (dotted lines) and limiting distribution (dash-dotted lines). Univariate distributions are marked with *.

Acknowledgments

Content of this article constitutes a part of the PhD dissertation of Zhixin Lun at Oakland University. Financial research support from Oakland University during the summers of Years 2018 and 2019 is acknowledged. We are also grateful to an anonymous referee and editor for their valuable comments resulting in several improvements in this article.

References

Armagan. A, Dunson, D. and Lee J. (2013). Generalized double Pareto shrinkage. *Statistica Sinica*, **23**, 119–143.

Crisóstomo, J., Matafome, P., Santos-Silva, D., Gomes, A. L., Gomes, M., Patrício, M., Letra, L., Sarmento-Ribeiro, A. B., Santos, L. and Seiça, R. (2016). Hyperresistinemia and metabolic dysregulation: a risky crosstalk in obese breast cancer. *Endocrine*, **53**, 433–442.

Eberl, A. and Klar, B. (2019). A note on a measure of asymmetry. *Statistical Papers*. https://doi.org/10.1007/s00362-019-01145-4

Harris, C. M. (1968). The Pareto distribution as a queue service discipline. *Operations Research*, **16**, 307–313.

Hinkley, D. V. and Revankar, N. S. (1977). Estimation of the Pareto law from underreported data: A further analysis. *Journal of Econometrics*, **5**, 1–11.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994). *Continuous Univariate Distributions, Volume 1, 2nd Edition*. John Wiley & Sons, New York.

Khattree, R. and Bahuguna, M. (2019). An alternative data analytic approach to measure the univariate and multivariate skewness. *International Journal of Data Science and Analytics*, **7**, 1–16.

Kotz, S., Kozubowski, T. J., and Podgórski, K. (2001). *The Laplace Distribution and Generalizations*. Birkhäuser, Boston.

Lun, Z. and Khattree, R. (2020). *KbMvtSkew: Khattree-Bahuguna's Univariate and Multivariate Skewness*. R package version 1.0.2. https://CRAN.R-project.org/package=KbMvtSkew.

Lun, Z. and Khattree, R. (2020). *NonNorMvtDist: Multivariate Lomax (Pareto Type II) and Its Related Distributions*. R package version 1.0.2. https://CRAN.R-project.org/package=NonNorMvtDist.

Lun, Z. and Khattree, R. (2020). An overlooked Lomax-exponential connection. *Communications in Statistics - Theory and Methods*. https://doi.org/10.1080/03610926.2020.1734836

Nayak, T. K. (1987). Multivariate Lomax distribution: properties and usefulness in reliability theory. *Journal of Applied Probability*, **24(1)**, 170–177.

Pal, S., Khare, K., and Hobert, J. P. (2017). Trace class markov chains for Bayesian inference with generalized Double Pareto shrinkage priors. *Scandinavian Journal of Statistics*, **44**, 307–323.

Patil, P. N., Patil, P. P. and Bagkavos, D. (2012). A measure of asymmetry. *Statistical Papers*, **53**, 971–985.

Patrício, M., Pereira, J., Crisóstomo J., Matafome, P., Gomes, M., Seiça, R. and Caramelo, F. (2018). Using Resistin, glucose, age and BMI to predict the presence of breast cancer. *BMC Cancer*, **18(29)**, 1–8.

R Core Team. (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

ANNEXURE

A. Proofs of Some of the Lemmas Used

Lemma 20: For $0 < \alpha < 1$, $\int_0^1 \ln(1-\alpha)\ln(\alpha)d\alpha = 2 - \frac{\pi^2}{6}$.

Proof: Clearly, the integral is convergent. The orders of integration and summation can be interchanged. Using Taylor series to expand $\ln(1-\alpha) = -\sum_{n=1}^{\infty} \frac{\alpha^n}{n}$, we have

$$\begin{aligned}\int_0^1 \ln(1-\alpha)\ln(\alpha)d\alpha &= -\int_0^1 \ln(\alpha)\sum_{n=1}^{\infty}\frac{\alpha^n}{n}d\alpha \\ &= -\sum_{n=1}^{\infty}\frac{1}{n}\left[-\frac{1}{(n+1)^2}\right] \text{(by using integration by parts)} \\ &= \sum_{n=1}^{\infty}\left[\frac{1}{n} - \frac{1}{n+1} - \frac{1}{(n+1)^2}\right] \\ &= \sum_{n=2}^{\infty}\left[\frac{1}{n-1} - \frac{1}{n}\right] - \sum_{n=2}^{\infty}\left[\frac{1}{n^2}\right] = 1 - (\frac{\pi^2}{6} - 1) = 2 - \frac{\pi^2}{6}. \quad \square \end{aligned} \tag{10}$$

$\square$

Lemma 21: For the first derivative of beta function $B(u,u)$, with respect to u, say $B'(u,u)$, $\lim_{u\to1} B'(u,u) = -2$.

Proof. Given $B(u,u) = \int_0^1 x^{u-1}(1-x)^{u-1}dx$, we have by using the Leibniz's Rule, the derivative of $B(u,u)$ with respect to u,

$$B'(u,u) = \int_0^1 \left[x^{u-1}\ln(x)(1-x)^{u-1} + x^{u-1}(1-x)^{u-1}\ln(1-x)\right]dx.$$

Accordingly, $\lim_{u\to1} B'(u,u) = \int_0^1 \ln(x)dx + \int_0^1 \ln(1-x)dx = -1 + -1 = -2$. $\square$

Lemma 22: For the second derivative of beta function $B(u,u)$, say $B''(u,u)$, $\lim_{u\to1} B''(u,u) = 8 - \frac{\pi^2}{3}$.

Proof: Again, by applying Leibniz's Rule, we have

$$\begin{aligned}B''(u,u) = \int_0^1 &[x^{u-1}\ln(x)^2(1-x)^{u-1} + x^{u-1}\ln(x)(1-x)^{u-1}\ln(1-x) \\ &+ x^{u-1}\ln(x)(1-x)^{u-1}\ln(1-x) + x^{u-1}(1-x)^{u-1}\ln(1-x)^2]dx\end{aligned}$$

Thus,

$$\lim_{u\to 1} B''(u,u) = \int_0^1 \ln(x)^2 dx + 2\int_0^1 \ln(1-x)\ln(x)dx + \int_0^1 \ln(1-x)^2 dx$$
$$= 2 + 2(2-\frac{\pi^2}{6}) + 2 = 8 - \frac{\pi^2}{3}. \quad \square$$

The middle integral is evaluated by using the Lemma 20 and integration by parts. $\square$

Lemma 23:

$$\int_0^\infty \frac{1}{(1+\frac{x^2}{y^2})^p(1+y^2)^p}dy = \frac{B(p-\frac{1}{2},\frac{1}{2})}{2(1+|x|)^{2p-1}}, \quad -\infty < x < \infty, \quad p > \frac{1}{2}.$$

Proof:. Let $\mathcal{I} = \int_0^\infty \frac{1}{(1+\frac{x^2}{y^2})^p(1+y^2)^p}dy$. Define $t = \frac{|x|}{y}$, then $dy = -\frac{|x|}{t^2}dt$. So $\mathcal{I} = \int_0^\infty \frac{1}{(1+\frac{x^2}{y^2})^p(1+y^2)^p}dy = \int_0^\infty \frac{\frac{|x|}{t^2}}{(1+t^2)^p(1+\frac{x^2}{t^2})^p}dt$. Thus,

$$2\mathcal{I} = \int_0^\infty \frac{1}{(1+\frac{x^2}{t^2})^p(1+t^2)^p}dt + \int_0^\infty \frac{\frac{|x|}{t^2}}{(1+t^2)^p(1+\frac{x^2}{t^2})^p}dt$$
$$= \int_0^\infty \left(1+\frac{|x|}{t^2}\right)\frac{1}{(1+\frac{x^2}{t^2})^p(1+t^2)^p}dt = \int_0^\infty \left(1+\frac{|x|}{t^2}\right)\frac{1}{(1+\frac{x^2}{t^2}+x^2+t^2)^p}dt.$$

Now define, $s = t - \frac{|x|}{t}$, then $ds = 1 + \frac{|x|}{t^2}dt, \quad s^2 = t^2 - 2|x| + \frac{x^2}{t^2}$, and hence we have

$$2\mathcal{I} = \int_{-\infty}^\infty \frac{1}{(1+s^2+2|x|+x^2)^p}ds = \int_{-\infty}^\infty \frac{1}{[s^2+(1+|x|)^2]^p}ds.$$

Thus, $\mathcal{I} = \int_0^\infty \frac{1}{[s^2+(1+|x|)^2]^p}ds = \frac{B(p-\frac{1}{2},\frac{1}{2})}{2(1+|x|)^{2p-1}}$. $\square$

B. Moments and Skewness of Generalized Double Pareto Distribution

B.1. Moments

Consider the nth raw moment for univariate Lomax distribution with shape parameter β and rate parameter θ:

$$\mu_n' = \int_0^\infty x^n f(x)dx = \int_0^\infty x^n \frac{\beta\theta}{(1+\theta x)^{\beta+1}}dx$$
$$= \frac{\beta}{\theta^n}\int_0^1 u^{(\beta-n)-1}(1-u)^{(n+1)-1}du \quad \left(\text{Let } u = \frac{1}{1+\theta x}, \text{ then } dx = -\frac{1}{\theta u^2}du\right)$$
$$= \frac{\beta}{\theta^n}B(\beta-n, n+1) \quad = \frac{n!}{\theta^n(\beta-1)\cdots(\beta-n)}, \qquad n < \beta.$$

Consider the nth raw moment for generalized double Pareto distribution with probability density function given by

$$h(x) = \begin{cases} h_1(x) = \frac{\theta_1\theta_2\beta}{(\theta_1+\theta_2)(1-\theta_1 x)^{\beta+1}}, & \text{if } x \le 0,\ \beta,\theta_1,\theta_2 > 0 \\ h_2(x) = \frac{\theta_1\theta_2\beta}{(\theta_1+\theta_2)(1+\theta_2 x)^{\beta+1}}, & \text{if } x > 0,\ \beta,\theta_1,\theta_2 > 0. \end{cases}$$

Then we have

$$\begin{aligned}\mu_n' &= \int_{-\infty}^{0} x^n h_1(x)dx + \int_0^\infty x^n h_2(x)dx \\ &= \frac{\theta_2}{\theta_1+\theta_2}(-1)^n \frac{n!}{\theta_1^n(\beta-1)\cdots(\beta-n)} + \frac{\theta_1}{\theta_1+\theta_2}\frac{n!}{\theta_2^n(\beta-1)\cdots(\beta-n)} \\ &= \frac{n!}{(\theta_1+\theta_2)(\beta-1)\cdots(\beta-n)}\left[\frac{(-1)^n\theta_2}{\theta_1^n} + \frac{\theta_1}{\theta_2^n}\right], \qquad n < \beta.\end{aligned}$$

B.2. Pearson's Coefficient of Skewness

The third central moment is given by $\mu_3 = \mu_3' - 3\mu_2\mu_1' - \mu_1'^3$ while the third raw moment is given by

$$\mu_3' = \frac{3!}{(\theta_1+\theta_2)(\beta-1)(\beta-2)(\beta-3)}\left[\frac{(-1)^3\theta_2}{\theta_1^3} + \frac{\theta_1}{\theta_2^3}\right] = \frac{6(\theta_1^2+\theta_2^2)(\theta_1-\theta_2)}{\theta_1^3\theta_2^3(\beta-1)(\beta-2)(\beta-3)}.$$

Also, $\mu_2 = \mathrm{Var}(X) = \frac{\beta\theta_1^2+\beta\theta_2^2-2\theta_1\theta_2}{\theta_1^2\theta_2^2(\beta-1)^2(\beta-2)}$ and $\mu_1' = \frac{\theta_1-\theta_2}{\theta_1\theta_2(\beta-1)}$. Upon substitution, the third central moment is given by

$$\begin{aligned}\mu_3 &= \frac{6(\theta_1^2+\theta_2^2)(\theta_1-\theta_2)}{\theta_1^3\theta_2^3(\beta-1)(\beta-2)(\beta-3)} - \frac{3(\beta\theta_1^2+\beta\theta_2^2-2\theta_1\theta_2)}{\theta_1^2\theta_2^2(\beta-1)^2(\beta-2)}\frac{(\theta_1-\theta_2)}{\theta_1\theta_2(\beta-1)} - \frac{(\theta_1-\theta_2)^3}{\theta_1^3\theta_2^3(\beta-1)^3} \\ &= \frac{2(\theta_1-\theta_2)(\beta+1)[\beta\theta_1^2+\beta\theta_2^2+(\beta-3)\theta_1\theta_2]}{\theta_1^3\theta_2^3(\beta-1)^3(\beta-2)(\beta-3)}.\end{aligned}$$

Finally, the Pearson's skewness becomes

$$\begin{aligned}\gamma = \frac{\mu_3}{\mu_2^{3/2}} &= \frac{2(\theta_1-\theta_2)(\beta+1)[\beta\theta_1^2+\beta\theta_2^2+(\beta-3)\theta_1\theta_2]}{\theta_1^3\theta_2^3(\beta-1)^3(\beta-2)(\beta-3)} \times \left(\frac{\theta_1^2\theta_2^2(\beta-1)^2(\beta-2)}{\beta\theta_1^2+\beta\theta_2^2-2\theta_1\theta_2}\right)^{3/2} \\ &= \frac{2(\beta+1)\sqrt{(\beta-2)}(\theta_1-\theta_2)[\beta\theta_1^2+\beta\theta_2^2+(\beta-3)\theta_1\theta_2]}{(\beta-3)(\beta\theta_1^2+\beta\theta_2^2-2\theta_1\theta_2)^{3/2}}, \qquad \beta > 3.\end{aligned}$$

B.3. Khattree-Bahuguna's Skewness

Recall the quantile function for $\mathrm{GDP}(\beta,\theta_1,\theta_2)$

$$F^{-1}(\alpha) = \begin{cases} \frac{1}{\theta_1}\left[1 - \left(\frac{\theta_2}{\alpha(\theta_1+\theta_2)}\right)^{\frac{1}{\beta}}\right], & \text{if } 0 < \alpha \le \frac{\theta_2}{\theta_1+\theta_2}, \\ \frac{1}{\theta_2}\left[\left(\frac{\theta_1}{(1-\alpha)(\theta_1+\theta_2)}\right)^{\frac{1}{\beta}} - 1\right], & \text{if } \frac{\theta_2}{\theta_1+\theta_2} < \alpha < 1. \end{cases}$$

For simplification, we let $d_1 = \frac{\theta_2}{\theta_1+\theta_2}$, $\quad d_2 = \frac{\theta_1}{\theta_1+\theta_2}$. Clearly, $d_1 + d_2 = 1$. Thus,

$$F^{-1}(\alpha) = \begin{cases} h_1(\alpha) = \frac{1}{\theta_1}\left[1 - \left(\frac{d_1}{\alpha}\right)^{\frac{1}{\beta}}\right], & \text{if } 0 < \alpha \le d_1, \\ h_2(\alpha) = \frac{1}{\theta_2}\left[\left(\frac{d_2}{1-\alpha}\right)^{\frac{1}{\beta}} - 1\right], & \text{if } d_1 < \alpha < 1, \end{cases}$$

and

$$F^{-1}(1-\alpha) = \begin{cases} g_1(1-\alpha) = \frac{1}{\theta_2}\left[\left(\frac{d_2}{\alpha}\right)^{\frac{1}{\beta}} - 1\right], & \text{if } 0 < \alpha \le d_2, \\ g_2(1-\alpha) = \frac{1}{\theta_1}\left[1 - \left(\frac{d_1}{1-\alpha}\right)^{\frac{1}{\beta}}\right], & \text{if } d_2 < \alpha < 1. \end{cases}$$

As the above quantile functions have not been centered, the Khattree-Baguhuna's skewness is computed by $\delta = \frac{1}{2}\left[1 + \frac{\int_0^1 F^{-1}(\alpha)F^{-1}(1-\alpha)d\alpha - \mu^2}{\mu_2}\right]$. We consider the computation of the kernel term $\int_0^1 F^{-1}(\alpha)F^{-1}(1-\alpha)$. If $d_1 < d_2$, we have $\int_0^1 F^{-1}(\alpha)F^{-1}(1-\alpha)d\alpha = \int_0^{d_1} h_1(\alpha)g_1(1-\alpha)d\alpha + \int_{d_1}^{d_2} h_2(\alpha)g_1(1-\alpha)d\alpha + \int_{d_2}^1 h_2(\alpha)g_2(1-\alpha)d\alpha$. Now, consider the first and third integrals,

$$\begin{aligned}\mathcal{I}_1 &= \int_0^{d_1} h_1(\alpha)g_1(1-\alpha)d\alpha = \frac{1}{\theta_1\theta_2}\int_0^{d_1}\left[\left(\frac{d_2}{\alpha}\right)^{\frac{1}{\beta}} - 1 - \left(\frac{d_1d_2}{\alpha^2}\right)^{\frac{1}{\beta}} + \left(\frac{d_1}{\alpha}\right)^{\frac{1}{\beta}}\right]d\alpha \\ &= \frac{1}{\theta_1\theta_2}\left[\left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{\beta-1}d_1 - d_1 - \left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{\beta-2}d_1 + \frac{\beta}{\beta-1}d_1\right] = \frac{d_1}{\theta_1\theta_2}\left[\left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\left(\frac{\beta}{\beta-1} - \frac{\beta}{\beta-2}\right) + \frac{1}{\beta-1}\right] \\ &= \frac{d_1}{\theta_1\theta_2}\left[\frac{1}{\beta-1} - \left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{(\beta-1)(\beta-2)}\right].\end{aligned}$$

Similarly,

$$\begin{aligned}\mathcal{I}_3 &= \int_{d_2}^1 h_2(\alpha)g_2(1-\alpha)d\alpha = \frac{1}{\theta_1\theta_2}\int_{d_2}^1\left[\left(\frac{d_2}{1-\alpha}\right)^{\frac{1}{\beta}} - \left(\frac{d_1d_2}{(1-\alpha)^2}\right)^{\frac{1}{\beta}} - 1 + \left(\frac{d_1}{1-\alpha}\right)^{\frac{1}{\beta}}\right]d\alpha \\ &= \frac{1}{\theta_1\theta_2}\left[\left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{\beta-1}d_1 - \left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{\beta-2}d_1 - d_1 + \frac{\beta}{\beta-1}d_1\right] = \frac{d_1}{\theta_1\theta_2}\left[\frac{1}{\beta-1} - \left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{(\beta-1)(\beta-2)}\right].\end{aligned}$$

Thus, $\mathcal{I}_1 = \mathcal{I}_3$. Now we consider the second integral term:

$$\begin{aligned}\mathcal{I}_2 &= \int_{d_1}^{d_2} h_2(\alpha)g_1(1-\alpha)d\alpha = \int_{d_1}^{d_2}\frac{1}{\theta_2}\left[\left(\frac{d_2}{1-\alpha}\right)^{\frac{1}{\beta}} - 1\right]\frac{1}{\theta_2}\left[\left(\frac{d_2}{\alpha}\right)^{\frac{1}{\beta}} - 1\right]d\alpha \\ &= \frac{1}{\theta_2^2}\int_{d_1}^{d_2}\left[d_2^{\frac{2}{\beta}}(1-\alpha)^{-\frac{1}{\beta}}\alpha^{-\frac{1}{\beta}} - \left(\frac{d_2}{1-\alpha}\right)^{\frac{1}{\beta}} - \left(\frac{d_2}{\alpha}\right)^{\frac{1}{\beta}} + 1\right]d\alpha \\ &= \frac{1}{\theta_2^2}\left\{d_2^{\frac{2}{\beta}}\left[B\left(d_2; 1-\frac{1}{\beta}, 1-\frac{1}{\beta}\right) - B\left(d_1; 1-\frac{1}{\beta}, 1-\frac{1}{\beta}\right)\right] - d_2^{\frac{1}{\beta}}\frac{\beta}{\beta-1}\left[-d_1^{-\frac{1}{\beta}+1} + d_2^{-\frac{1}{\beta}+1}\right] - d_2^{\frac{1}{\beta}}\frac{\beta}{\beta-1}\left[d_2^{-\frac{1}{\beta}+1} - d_1^{-\frac{1}{\beta}+1}\right] + d_2 - d_1\right\} \\ &= \frac{1}{\theta_2^2}\left\{d_2^{\frac{2}{\beta}}\left[B\left(d_2; 1-\frac{1}{\beta}, 1-\frac{1}{\beta}\right) - B\left(d_1; 1-\frac{1}{\beta}, 1-\frac{1}{\beta}\right)\right] + 2\left(\frac{d_2}{d_1}\right)^{\frac{1}{\beta}}\frac{\beta}{\beta-1}d_1 - 2\frac{\beta}{\beta-1}d_2 + d_2 - d_1\right\}\end{aligned}$$

Recall that $\mu = \frac{\theta_1-\theta_2}{\theta_1\theta_2(\beta-1)} = \frac{d_2-d_1}{\theta_2 d_2(\beta-1)}$ and $\mu_2 = \frac{\beta(d_2-d_1)^2}{d_1^2 d_2^2(\beta-1)^2(\beta-2)} + \frac{2}{\theta_1\theta_2(\beta-1)(\beta-2)}$. Therefore, the Khattree-Bahuguna's skewness is $\delta = \frac{1}{2}\left[1 + \frac{2\mathcal{I}_1+\mathcal{I}_2-\mu^2}{\mu_2}\right]$. We omit the similar derivation for $d_2 < d_1$.

Table 4: Data set of healthy controls and breast cancer patients (Suitably adjusted from the source data reported at *https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Coimbra***)**

	Healthy controls		Breast cancer patients	
Pair	BMI	Insulin	BMI	Insulin
1	18.67	6.11	18.37	6.03
2	20.69	3.12	20.83	4.56
3	20.76	7.55	20.83	3.42
4	21.11	3.55	21.08	6.20
5	21.37	3.23	21.30	13.85
6	21.47	3.47	21.36	3.00
7	22.00	3.35	21.51	6.68
8	22.03	2.87	22.21	36.94
9	22.70	4.69	22.22	5.70
10	22.85	3.23	22.50	5.26
11	22.86	4.09	22.66	3.48
12	23.00	4.95	22.83	6.86
13	23.01	5.66	22.89	2.74
14	23.12	4.50	23.14	4.90
15	23.34	5.78	23.62	4.42
16	23.50	2.71	24.22	3.73
17	23.80	6.47	24.24	21.70
18	25.30	3.51	25.51	10.39
19	25.70	8.08	25.59	2.82
20	25.90	4.58	26.56	10.55
21	26.35	5.14	26.56	6.52
22	26.60	4.46	26.67	41.61
23	27.10	26.21	26.67	22.03
24	27.20	14.07	26.84	4.53
25	27.30	5.20	26.85	3.33
26	27.69	3.85	27.18	19.91
27	27.70	6.04	27.64	2.43
28	28.58	4.34	28.44	8.81
29	29.22	5.38	29.14	10.95
30	29.40	10.70	29.15	16.58
31	29.61	5.82	29.30	4.17
32	30.28	4.38	29.38	4.71
33	30.30	8.34	29.67	14.65
34	30.48	5.54	29.78	8.40
35	31.24	4.18	30.48	7.01
36	31.45	9.24	30.80	30.21
37	31.98	4.53	30.84	41.89
38	32.04	18.08	30.92	10.49
39	32.27	5.81	31.22	18.08
40	32.50	5.43	31.23	30.13
41	34.17	6.59	31.25	4.33
42	34.42	23.19	31.25	12.16
43	34.53	4.43	31.64	9.67
44	35.09	5.65	31.98	16.64
45	35.25	6.82	32.05	5.73
46	35.59	3.88	32.46	28.68
47	35.86	8.58	32.46	24.89
48	36.21	15.53	33.18	5.75
49	36.51	14.03	34.84	12.55
50	36.79	10.18	35.56	8.15
51	37.04	6.76	36.05	11.91
52	38.58	6.70	37.11	5.64