ESTIMATION OF PARAMETERS OF GENERALIZED BETA OF THE SECOND KIND (GB2) DISTRIBUTION BY MAXIMUM LIKELIHOOD ESTIMATION (MLE) AND NEWTON-RAPHSON ITERATION

By Warsono Warsono

ESTIMATION OF PARAMETERS OF GENERALIZED BETA OF THE SECOND KIND (GB2) DISTRIBUTION BY MAXIMUM LIKELIHOOD ESTIMATION (MLE) AND NEWTON-RAPHSON ITERATION

Warsono, Waryoto, Mustofa Usman and Wamiliana

Department of Mathematics
University of Lampung
Indonesia

Abstract

The aim of this study is to derive the estimates of the parameters of generalized beta of the second kind (GB2) distribution by using the maximum likelihood estimation (MLE). Due to the difficulty in finding the analytical solution by MLE approach, the estimate is determined numerically by using iteration and Newton-Raphson methods. The Newton-Raphson method is used to estimate parameters, to find a confidence interval, to estimate the bias, and to estimate the variance for some difference sample size configuration viz. for $n = 20, 30, 50, 100$ and 500. The estimation of the parameters a, b, p and q attains the values close to the real value of the parameters. If the size of the sample increases, the confidence interval becomes narrower and the bias and variance become smaller.

1. Introduction

Many papers which have discussed the distribution of income used beta distribution [1]; gamma distribution [2-5]; and Weibull model [6]. The generalized beta distribution of the second kind (GB2) is a very flexible four-parameter distribution. It is used a lot to analyze income distribution. References [7] and [8] suggested the generalized beta of the second kind (GB2) as a model for the size distribution of income and indicators of poverty. It captures the characteristics of income distribution including skewness, peakness in low-middle range, and long right hand tail. This distribution therefore provides good description of income distribution [7, 9-11]. GB2 is used in mathematical economic, in insurance company, health science and in industry. Although a large number of functional forms have been proposed, the four-parameter generalized beta of the second kind (GB2) model is now widely acknowledged to give an excellent description of income distributions, providing the goodness-of-fit with relative parsimony, while also including many other models as special or limiting cases [7, 12-15]. [16] addressed issues of time-inconsistency in top-coded US Current Population Survey earnings data by fitting GB2 distributions that account for top-coding, and derive a consistent time series of Gini coefficients from the estimates. In [17], the model of optimization of the behavior predicts that the earnings distribution has the GB2 shape.

A random variable has a distribution of generalized beta of the second kind (GB2) with parameter (a, b, p, q) if the probability density function is of the form:

$$f(x) = \frac{ax^{ap-1}}{b^{ap}B(p, q)\left(1 + \left(\frac{x}{b}\right)^a\right)^{p+q}}, \quad x > 0; a, b, p, q > 0. \tag{1}$$

Parameter b is a scale parameter, a, p and q are each shape parameter, $B(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p + q)}$ is the beta function, and $\Gamma(\cdot)$ is the gamma function [7]. The kth moment of the GB2 distribution is $E(x^k) =$

$$\frac{b^k \Gamma\left(p + \frac{k}{a}\right)\Gamma\left(q - \frac{k}{a}\right)}{\Gamma(p)\Gamma(q)}$$ and exists only if $-ap < k < aq$. Tail behavior of the distribution depends on ap (lower tail) and aq (upper tail), with larger values of a reducing the density at both tails, and the relative sizes of p and q affecting skewness [10, 18].

The aim of this study is to estimate the parameters of the GB2 distribution by using the maximum likelihood estimation (MLE) method and then the simulation by using Software R is used to estimate the parameters for some difference sample size configuration, namely for the sample sizes: 20, 30, 50 100 and 500.

2. The Estimation

The estimation of parameters GB2 by MLE

To estimate the parameter by using the maximum likelihood estimation (MLE) method, first we define the likelihood function as follows:

$$L(a, b, p, q \,|\, \underline{x}) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \frac{a(x_i)^{ap-1}}{b^{ap} B(p, q)\left(1 + \left(\frac{x_i}{b}\right)^a\right)^{p+q}}$$

$$= \frac{a^n \left[\prod_{i=1}^{n} (x_i)^{ap-1}\right]}{\left[b^{nap}(B(p, q))^n\right]\left[\prod_{i=1}^{n}\left(1 + \left(\frac{x_i}{b}\right)^a\right)^{p+q}\right]}. \tag{2}$$

Applying the natural logarithm to equation (2) above, we have

$$\ln L(a, b, p, q \,|\, \underline{x}) = \ln\left\{\frac{a^n \left[\prod_{i=1}^{n} (x_i)^{ap-1}\right]}{\left[b^{nap}(B(p, q))^n\right]\left[\prod_{i=1}^{n}\left(1 + \left(\frac{x_i}{b}\right)^a\right)^{p+q}\right]}\right\}$$

$$= n \ln a + (ap - 1)\sum\nolimits_{i=1}^{n} \ln(x_i) - nap1 - n \ln B(p, q)$$

$$- (p + q)\sum\nolimits_{i=1}^{n} \ln\left[1 + \left(\frac{x_i}{b}\right)^{a}\right], \tag{3}$$

where $B(p, q) = \dfrac{\Gamma(p) \cdot \Gamma(q)}{\Gamma(p + q)}$, $B(p, q)$ is the beta function and $\Gamma(p)$ is the gamma function, then equation (3) can be written as:

$$\ln L(a, b, p, q | \underline{x}) = n \ln a + (ap - 1)\sum_{i=1}^{n} \ln(x_i) - nap \ln b - n \ln \Gamma(p)$$

$$+ n \ln \Gamma(q) - n \ln \Gamma(p + q)$$

$$- (p + q)\sum\nolimits_{i=1}^{n} \ln\left[1 + \left(\frac{x_i}{b}\right)^{a}\right]. \tag{4}$$

Next, we set the first derivatives with respect to the parameters of interest equal to zero as follows:

$$\frac{\partial \ln L(\theta)}{\partial \theta} = 0.$$

The first derivative with respect to a is set equal to zero,

$$\frac{\partial \ln L(a, b, p, q | \underline{x})}{\partial a} = 0,$$

$$\frac{n}{\hat{a}} + \hat{p}\sum_{i=1}^{n} \ln(x_i) - n\hat{p} \ln \hat{b} - \frac{\partial}{\partial a}\left\{(p + q)\sum_{i=1}^{n} \ln\left[1 + \left(\frac{x_i}{b}\right)^{a}\right]\right\} = 0,$$

$$\frac{n}{\hat{a}} + \hat{p}\sum\nolimits_{i=1}^{n} \ln(x_i) - n\hat{p} \ln \hat{b} - (\hat{p} + \hat{q})\left\{\sum\nolimits_{i=1}^{n}\left[\frac{\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}} \ln\left(\frac{x_i}{\hat{b}}\right)}{1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}}\right]\right\} = 0. \tag{5}$$

The first derivative with respect to b is set equal to zero,

$$\frac{\partial \ln L(a, b, p, q \mid \underline{x})}{\partial b} = 0,$$

$$-\frac{\hat{a}n\hat{p}}{\hat{b}} - (\hat{p} + \hat{q})\frac{\partial}{\partial b}\left\{\sum_{i=1}^{n} \ln\left[1 + \left(\frac{x_i}{b}\right)^a\right]\right\} = 0.$$

Let

$$S = \ln\left[1 + \left(\frac{x_i}{b}\right)^a\right], \quad \frac{\partial s}{\partial b} = \frac{-\hat{a}x_i\left(\frac{x_i}{b}\right)^{\hat{a}-1}}{\hat{b}^2\left[1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right]},$$

$$\frac{\partial \ln L(a, b, p, q \mid \underline{x})}{\partial b} = -\frac{\hat{a}n\hat{p}}{\hat{b}} - (\hat{p} + \hat{q})\sum_{i=1}^{n}\left\{\frac{-\hat{a} \cdot x_i \cdot \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}}{\hat{b}^2\left[1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right]}\right\} = 0. \quad (6)$$

The first derivative with respect to p is set equal to zero,

$$\frac{\partial \ln L(a, b, p, q \mid \underline{x})}{\partial p} = 0,$$

$$\hat{a}\sum_{i=1}^{n} \ln(x_i) - \hat{a}n\ln\hat{b} - \frac{n\Gamma'(\hat{p})}{\Gamma(\hat{p})} - \frac{n\Gamma'(\hat{p} + \hat{q})}{\Gamma(\hat{p} + \hat{q})} - \sum_{i=1}^{n} n\left[1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right] = 0. \quad (7)$$

Since $\psi(p) = \frac{\partial \ln \Gamma(p)}{\partial p} = \frac{\Gamma'(p)}{\Gamma(p)}$,

$$\psi(p + q) = \frac{\partial \ln \Gamma(p + q)}{\partial(p + q)} \equiv \frac{\Gamma'(\hat{p} + \hat{q})}{\Gamma(\hat{p} + \hat{q})}, \quad p > 0,$$

where $\psi(p)$ is a function psi (digamma).

Equation (7) can be written as follows:

$$\hat{a}\sum_{i=1}^{n}\ln(x_i) - \hat{a}n\ln\hat{b} - n\psi(\hat{p}) - n\psi(\hat{p}+\hat{q}) - \sum_{i=1}^{n}\ln\left[1+\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right] = 0. \quad (8)$$

The first derivative with respect to q is set equal to zero,

$$\frac{\partial \ln L(a, b, p, q\,|\,\underline{x})}{\partial p} = 0,$$

$$\frac{n\Gamma'(\hat{q})}{\Gamma(\hat{q})} - \frac{n\Gamma'(\hat{p}+\hat{q})}{\Gamma(\hat{p}+\hat{q})} - \sum_{i=1}^{n}\ln\left[1+\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right] = 0,$$

$$n\psi(\hat{q}) - n\psi(\hat{p}+\hat{q}) - \sum_{i=1}^{n}\ln\left[1+\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right] = 0. \quad (9)$$

Equations (5), (6), (8) and (9) are very difficult to be solved by analytical method. Thus, the Newton-Raphson method will be used to estimate the parameters a, b, p and q. To estimate the parameters a, b, p and q by using Newton-Raphson method, first we find the gradient vector and the first derivative vector from the logarithm function with respect to a, b, p and q and define $g(\lambda)$ as follows:

$$g(\lambda) = \frac{\partial \ln L(a, b, p, q\,|\,\underline{x})}{\partial \lambda} = \begin{bmatrix} \dfrac{\partial \ln L(a, b, p, q\,|\,\underline{x})}{\partial a} \\[2mm] \dfrac{\partial \ln L(a, b, p, q\,|\,\underline{x})}{\partial b} \\[2mm] \dfrac{\partial \ln L(a, b, p, q\,|\,\underline{x})}{\partial p} \\[2mm] \dfrac{\partial \ln L(a, b, p, q\,|\,\underline{x})}{\partial q} \end{bmatrix}. \quad (10)$$

The Hessian matrix which was found from the second derivative of the logarithm function with respect to the parameters a, b, p and q and is defined as $H(\lambda)$:

$$H(\lambda) = \frac{\partial^2 \ln L(a, b, p, q \mid \underline{x})}{\partial \lambda \partial \lambda'}$$

$$= \begin{bmatrix} \dfrac{\partial^2 \ln L(a, b, p, q \mid \underline{x})}{\partial a \partial a} & \cdots & \dfrac{\partial^2 \ln L(a, b, p, q \mid \underline{x})}{\partial a \partial q} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 \ln L(a, b, p, q \mid \underline{x})}{\partial q \partial a} & \cdots & \dfrac{\partial^2 \ln L(a, b, p, q \mid \underline{x})}{\partial q \partial q} \end{bmatrix}. \quad (11)$$

The iteration process by using the Newton-Raphson method is:

$$\hat{\lambda}_{i+1} = \hat{\lambda}_i - [(H(\lambda))^{-1} \cdot g(\lambda)], \text{ where } i = 1, 2, 3, \dots. \quad (12)$$

To find the gradient vector $(g(\lambda))$, we use the first derivative from respective parameters. From equations (5), (6), (8) and (9), we find the Hessian matrix $H(\lambda)$, by using the second derivative of the logarithm function with respect to the respective parameters.

The second derivative of equation (5) with respect to a is

$$\frac{\partial^2 \ln L(a, b, p, q \mid \underline{x})}{\partial a \partial a} = 0,$$

$$-\frac{n}{\hat{a}^2} - (\hat{p} + \hat{q}) \sum_{i=1}^{n} \frac{\partial}{\partial a} \left[\frac{\left(\frac{x_i}{b}\right)^a \ln\left(\frac{x_i}{b}\right)}{1 + \left(\frac{x_i}{b}\right)^a} \right] = 0,$$

$$-\frac{n}{\hat{a}^2} - (\hat{p} + \hat{q}) \sum_{i=1}^{n} \left[\frac{\left(\ln\left(\frac{x_i}{\hat{b}}\right)\right)^2 \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}}{1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}} - \frac{\left(\frac{x_i}{\hat{b}}\right)^{2\hat{a}} \left(\ln\left(\frac{x_i}{\hat{b}}\right)\right)^2}{\left(1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right)^2} \right] = 0,$$

$$-\frac{n}{\hat{a}^2} - (\hat{p} + \hat{q}) \sum_{i=1}^{n} \left[\frac{\left(\ln\left(\frac{x_i}{\hat{b}}\right)\right)^2 \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}} \left(1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right) - \left(\frac{x_i}{\hat{b}}\right)^{2\hat{a}} \left(\ln\left(\frac{x_i}{\hat{b}}\right)\right)^2}{\left(1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right)^2} \right] = 0. \quad (13)$$

The second derivative of equation (5) with respect to b is

$$\frac{\partial^2 \ln L(a, b, \ p, \ q \,|\, \underline{x})}{\partial a \partial b} = 0,$$

$$-\frac{n\hat{p}}{\hat{b}} - (\hat{p} + \hat{q}) \sum_{i=1}^{n} \left\{ \frac{\partial}{\partial b} \left(\frac{\left(\frac{x_i}{b}\right)^a \ln\left(\frac{x_i}{b}\right)}{1 + \left(\frac{x_i}{b}\right)^a} \right) \right\} = 0.$$

Let $y = \left(\frac{x_i}{b}\right)^a \ln\left(\frac{x_i}{b}\right)$. Then

$$\frac{\partial y}{\partial b} = -\frac{1}{\hat{b}} \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}} + \ln\left(\frac{x_i}{\hat{b}}\right) \left(-\hat{a} x_i^{\hat{a}} \cdot \hat{b}^{-\hat{a}-1}\right)$$

$$= \frac{-\hat{b} \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}}{\hat{b}^2} - \frac{\left(\hat{a} x_i \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}\right) \ln\left(\frac{x_i}{\hat{b}}\right)}{\hat{b}^2}.$$

Let $S = \dfrac{\left(\frac{x_i}{b}\right)^a \ln\left(\frac{x_i}{b}\right)}{\left(1 + \left(\frac{x_i}{b}\right)^a\right)}$. Then

$$\frac{\partial S}{\partial b} = \frac{\left[\frac{\partial}{\partial b}\left(\left(\frac{x_i}{b}\right)^a \ln\left(\frac{x_i}{b}\right)\right)\right]\left(1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right) - \left(\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}} \ln\left(\frac{x_i}{\hat{b}}\right)\right)\left[\frac{\partial}{\partial b}\left(1 + \left(\frac{x_i}{b}\right)^a\right)\right]}{\left(1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right)^2}$$

$$
= \left\{ \left[\left(\frac{-\hat{a}x_i\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}\ln\left(\frac{x_i}{\hat{b}}\right) - \hat{b}\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}}{b^2} \right)\left(1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right) \right. \right.
$$

$$
\left. + \left(\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\ln\left(\frac{x_i}{\hat{b}}\right) \right)\left(\frac{\hat{a}\cdot x_i\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}}{\hat{b}^2} \right) \right\} x
$$

$$
- \frac{n\hat{p}}{\hat{b}} - (\hat{p} + \hat{q})\sum_{i=1}^{n}\left\{ \left[\left(\frac{-\hat{a}x_i\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}\ln\left(\frac{x_i}{\hat{b}}\right) - \hat{b}\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}}{\hat{b}^2} \right)\left(1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right) \right. \right.
$$

$$
\left. + \left(\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\ln\left(\frac{x_i}{\hat{b}}\right) \right)\left(\frac{\hat{a}x_i\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}}{\hat{b}^2} \right) \right\} x \left(1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right)^{-2} = 0. \qquad (14)
$$

The second derivative of equation (5) with respect to p is

$$
\frac{\partial^2 \ln L(a, b, \, p, \, q \,|\, x)}{\partial a \partial p} = 0,
$$

$$
\sum_{i=1}^{n} \ln(x_i) - n \ln \hat{b} - \left\{ \sum_{i=1}^{n} \left[\frac{\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\ln\left(\frac{x_i}{\hat{b}}\right)}{1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}} \right] \right\} = 0. \qquad (15)
$$

The second derivative of equation (5) with respect to q is

$$
\frac{\partial^2 \ln L(a, b, \, p, \, q \,|\, x)}{\partial a \partial q} = 0,
$$

$$-\left\{\sum_{i=1}^{n}\left[\frac{\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\ln\left(\frac{x_i}{\hat{b}}\right)}{1+\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}}\right]\right\}=0. \tag{16}$$

The second derivative of equation (6) with respect to a is

$$\frac{\partial^2 \ln L(a,\,b,\,p,\,q\,|\,\underline{x})}{\partial b\partial a}=0,$$

$$-\frac{n\hat{p}}{\hat{b}}+(\hat{p}+\hat{q})\sum_{i=1}^{n}\frac{\partial}{\partial a}\left\{\frac{ax_i\left(\frac{x_i}{b}\right)^{a-1}}{b^2\left[1+\left(\frac{x_i}{b}\right)^{a}\right]}\right\}=0.$$

Let $w=ax_i\left(\frac{x_i}{b}\right)^{a-1}$. Then

$$\frac{\partial w}{\partial a}=x_i\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}+\hat{a}x_i\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}\ln\left(\frac{x_i}{\hat{b}}\right)-\frac{n\hat{p}}{\hat{b}}$$

$$+(\hat{p}+\hat{q})\sum_{i=1}^{n}\frac{1}{\hat{b}^2}\left\{\left(\frac{x_i\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}+\hat{a}x_i\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}\ln\left(\frac{x_i}{\hat{b}}\right)}{\left(1+\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right)^2}\right)\right.$$

$$\times\left.\left(1+\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right)-\frac{\left(\hat{a}x_i\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}\right)\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\ln\left(\frac{x_i}{\hat{b}}\right)}{\left(1+\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right)^2}\right\}=0,$$

$$-\frac{n\hat{p}}{\hat{b}} + (\hat{p}+\hat{q})\sum_{i=1}^{n}\left\{\frac{\hat{a}x_i\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}\ln\left(\frac{x_i}{\hat{b}}\right)}{\hat{b}^2\left(1+\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right)} + \frac{x_i\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}}{\hat{b}^2\left(1+\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right)}\right.$$

$$\left.-\frac{\ln\left(\frac{x_i}{\hat{b}}\right)\hat{a}x_i\left(\frac{x_i}{\hat{b}}\right)^{2\hat{a}-1}}{\left[1+\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right]^2}\right\} = 0. \qquad (17)$$

The second derivative of equation (6) with respect to b is

$$\frac{\partial^2 \ln L(a,b,p,q|\underline{x})}{\partial b \partial b} = 0,$$

$$\frac{\hat{a}n\hat{p}}{\hat{b}^2} + (\hat{p}+\hat{q})\sum_{i=1}^{n}\frac{\partial}{\partial b}\left\{\frac{ax_i\left(\frac{x_i}{b}\right)^{a-1}}{b^2\left[1+\left(\frac{x_i}{b}\right)^{a}\right]}\right\} = 0.$$

Let $w = \dfrac{ax_i\left(\frac{x_i}{b}\right)^{a-1}}{b^2\left[1+\left(\frac{x_i}{b}\right)^{a}\right]}$. Then

$$\frac{\partial w}{\partial b} = \frac{\hat{a}x_i\left\{\left[\frac{\partial}{\partial b}\left(\frac{x_i}{b}\right)^{a-1}\right]\cdot\hat{b}^2\left[1+\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right] -\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}\cdot\left[\frac{\partial}{\partial b}b^2\left(1+\left(\frac{x_i}{b}\right)^{a}\right)\right]\right\}}{\hat{b}^4\left[1+\left(\frac{x_i}{\hat{b}}\right)^{a}\right]^2}$$

$$= \hat{a}x_i \left\{ \left[-(\hat{a}-1)(x_i)^{\hat{a}-1} \right] \cdot \hat{b}^{-\hat{a}} \cdot \hat{b}^2 \left[1 + \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}} \right] - \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}-1} \right.$$

$$\left. \cdot \left[\frac{\partial}{\partial b} b^2 \left(1 + \left(\frac{x_i}{b} \right)^a \right) \right] \right\} \hat{b}^4 \left[1 + \left(\frac{x_i}{\hat{b}} \right)^a \right]^{-2}$$

$$= \frac{\hat{a}x_i \left\{ \left[-(\hat{a}-1)\left(\frac{x_i}{\hat{b}} \right)^{\hat{a}-2} \right] x_i \left[1 + \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}} \right] - \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}-1} \cdot \left[2\hat{b} \left(1 + \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}} \right) \right] \right\}}{\hat{b}^4 \left[1 + \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}} \right]^2}$$

$$+ \frac{\hat{a}x_i \left\{ -\left(\frac{x_i}{\hat{b}} \right)^{\hat{a}-1} \left[-\hat{a} \cdot \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}-1} \cdot x_i \right] \right\}}{\hat{b}^4 \left[1 + \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}} \right]^2}$$

$$= \frac{-\hat{a}(\hat{a}-1)x_i^2 \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}-2}}{\hat{b}^4 \left[1 + \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}} \right]} - \frac{2\hat{a}x_i \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}-1}}{\hat{b}^3 \left[1 + \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}} \right]} + \frac{\hat{a}^2 x_i^2 \left(\frac{x_i}{\hat{b}} \right)^{2\hat{a}-2}}{\hat{b}^4 \left[1 + \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}} \right]^2}$$

$$= \frac{\hat{a}^2 x_i^2 \left(\frac{x_i}{\hat{b}} \right)^{2\hat{a}-2}}{\hat{b}^4 \left[1 + \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}} \right]^2} - \frac{\hat{a}(\hat{a}-1)x_i^2 \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}-2}}{\hat{b}^4 \left[1 + \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}} \right]} - \frac{2\hat{a}x_i \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}-1}}{\hat{b}^3 \left[1 + \left(\frac{x_i}{\hat{b}} \right)^{\hat{a}} \right]} \frac{\hat{a}n\hat{p}}{\hat{b}^2}$$

$$+ (\hat{p} + \hat{q}) \sum_{i=1}^{n} \left\{ \frac{\hat{a}^2 x_i^2 \left(\frac{x_i}{\hat{b}}\right)^{2\hat{a}-2}}{\hat{b}^4 \left[1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}} \right]^2} - \frac{\hat{a}(\hat{a}-1) x_i^2 \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-2}}{\hat{b}^4 \left[1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}} \right]} \right.$$

$$\left. - \frac{2\hat{a} x_i \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}}{\hat{b}^3 \left[1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}} \right]} \right\} = 0. \tag{18}$$

The second derivative of equation (6) with respect to p is

$$\frac{\partial^2 \ln L(a, b, p, q \,|\, \underline{x})}{\partial b \partial p} = 0,$$

$$-\frac{\hat{a} n}{\hat{b}^2} + \sum_{i=1}^{n} \left\{ \frac{\hat{a} \cdot x_i \cdot \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}}{\hat{b}^2 \left[1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}} \right]} \right\} = 0. \tag{19}$$

The second derivative of equation (6) with respect to q is

$$\frac{\partial^2 \ln L(a, b, p, q \,|\, \underline{x})}{\partial b \partial q} = 0,$$

$$\sum_{i=1}^{n} \left\{ \frac{\hat{a} \cdot x_i \cdot \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}}{\hat{b}^2 \left[1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}} \right]} \right\} = 0. \tag{20}$$

The second derivative of equation (8) with respect to a is

$$\frac{\partial^2 \ln L(a, b, p, q \,|\, \underline{x})}{\partial p \partial a} = 0,$$

$$\sum_{i=1}^{n} \ln(x_i) - n \ln \hat{b} - \sum_{i=1}^{n} \frac{\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}} \ln\left(\frac{x_i}{\hat{b}}\right)}{\left(1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right)} = 0. \tag{21}$$

The second derivative of equation (8) with respect to b is

$$\frac{\partial^2 \ln L(a, b, p, q \mid \underline{x})}{\partial p \partial b} = 0.$$

Let $w = \ln\left[1 + \left(\frac{x_i}{b}\right)^a\right]$. Then

$$\frac{\partial w}{\partial b} = \frac{-\hat{a} x_i^{\hat{a}} \cdot \hat{b}^{(-\hat{a}-1)}}{\left[1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right]} = \frac{-\hat{a} x_i^{\hat{a}} \cdot \hat{b}^{(-\hat{a}-1)} \cdot \hat{b}^2 x_i^1 x_i^{-1}}{\hat{b}^2\left[1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right]}$$

$$= \frac{-\hat{a} x_i\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}}{\hat{b}^2\left[1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right]} \frac{\hat{a} n}{\hat{b}} + \sum_{i=1}^{n} \frac{\hat{a} x_i\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}}{\hat{b}^2\left[1 + \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right]} = 0. \tag{22}$$

The second derivative of equation (8) with respect to p is

$$\frac{\partial^2 \ln L(a, b, p, q \mid \underline{x})}{\partial p \partial p} = 0,$$

$$n\psi'(\hat{p}) - n\psi'(\hat{p} + \hat{q}) = 0. \tag{23}$$

The second derivative of equation (8) with respect to q is

$$\frac{\partial^2 \ln L(a, b, p, q \mid \underline{x})}{\partial p \partial q} = 0,$$

$$-n\psi'(\hat{p} + \hat{q}) = 0. \tag{24}$$

The second derivative of equation (9) with respect to a is

$$\frac{\partial^2 \ln L(a,b,p,q\,|\,\underline{x})}{\partial q \partial a} = 0,$$

$$-\sum_{i=1}^n \frac{\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}} \cdot \ln\left(\frac{x_i}{\hat{b}}\right)}{1+\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}} = 0. \tag{25}$$

The second derivative of equation (9) with respect to b is

$$\frac{\partial^2 \ln L(a,b,p,q\,|\,\underline{x})}{\partial q \partial b} = 0,$$

$$\sum_{i=1}^n \frac{\hat{a} x_i \left(\frac{x_i}{\hat{b}}\right)^{\hat{a}-1}}{\hat{b}^2\left[1+\left(\frac{x_i}{\hat{b}}\right)^{\hat{a}}\right]} = 0. \tag{26}$$

The second derivative of equation (9) with respect to is

$$\frac{\partial^2 \ln L(a,b,p,q\,|\,\underline{x})}{\partial q \partial p} = 0,$$

$$n\psi'(\hat{p}+\hat{q}) = 0. \tag{27}$$

The second derivative of equation (9) with respect to q is

$$\frac{\partial^2 \ln L(a,b,p,q\,|\,\underline{x})}{\partial q \partial q} = 0,$$

$$n\psi'(\hat{q}) - n\psi'(\hat{p}+\hat{q}) = 0. \tag{28}$$

3. Simulation Result and Discussion

To find the Hessian matrix, we substitute the result from the second derivative of logarithm function with respect to the respective parameters a, b, p and q into equation (11). For the iteration process by Newton-Raphson method, we use the software R. In this simulation, we use the initial

values for the respective parameters as $a = 2$; $b = 1, 2$; $p = 15$ and $q = 0, 75$. In generating the sample, we use the sample sizes for n equals to 20, 30, 50, 100 and 500 and the number of iterations 100 for respective sample sizes. From the simulation, we calculate the mean, bias, variances and confidence interval (CI) for respective sample sizes $n = 20, 30, 50, 100$ and 500.

Table 1. Estimation value, confidence interval (CI), bias, and variances of generalized beta II (GB2) for the parameters $a = 2$; $b = 1.2$; $p = 15$ and $q = 0.75$, sample sizes $n = 20, 30, 50, 100$ and 500

Sample size	Parameters	$a = 2$	$b = 1.2$	$p = 15$	$q = 0.75$
$n = 20$	Estimation	2.02889	1.23003	15.02581	0.76278
	CI	[2.01148; 2.04630]	[1.21727; 1.24279]	[14.99674; 15.05488]	[0.74040; 0.78516]
	Bias	0.02889	0.03003	0.02581	0.01278
	Variance	0.0000789261	0.0000423958	0.0002199300	0.0001304063
$n = 30$	Estimation	2.02609	1.22833	15.02016	0.76178
	CI	[2.01521; 2.03697]	[1.21715; 1.23951]	[15.01211; 15.02822]	[0.75515; 0.76841]
	Bias	0.02609	0.02833	0.02016	0.01178
	Variance	0.0000307998	0.0000325219	0.0000168876	0.0000114491
$n = 50$	Estimation	2.01791	1.22602	15.01780	0.76107
	CI	[2.01428; 2.02154]	[1.22457; 1.22747]	[15.01527; 15.02033]	[0.75810; 0.76403]
	Bias	0.01791	0.02602	0.01780	0.01107
	Variance	0.0000034302	0.0000005464	0.0000016640	0.0000022841
$n = 100$	Estimation	2.01088	1.20308	15.00697	0.75039
	CI	[2.00914; 2.01261]	[1.20208; 1.20407]	[15.00651; 15.00743]	[0.75003; 0.75075]
	Bias	0.01088	0.00308	0.00697	0.00039
	Variance	0.0000007826	0.0000002569	0.0000000550	0.0000000338
$n = 500$	Estimation	2.00999	1.20301	15.00400	0.75020
	CI	[2.00990; 2.01008]	[1.20281; 1.20321]	[15.00388; 15.00413]	[0.75019; 0.75021]
	Bias	0.00999	0.00301	0.00400	0.00020
	Variance	0.0000000021	0.0000000102	0.0000000039	0.0000000001

From the results of the simulation given in Table 1, we can conclude that when the sample size increases, the estimation values and the real values are close. Further, when the sample size is larger, the confidence interval will be narrower, the bias and variances will be smaller and close to zero.

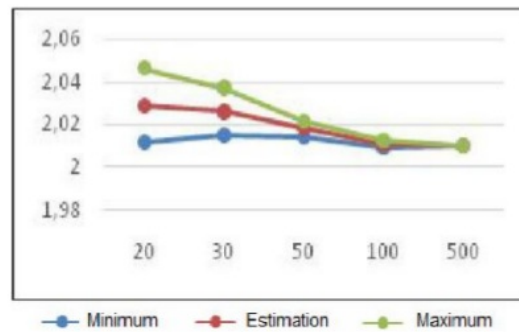Figures 1 to 6 describe the estimation of parameters and its confidence intervals, bias and variances.

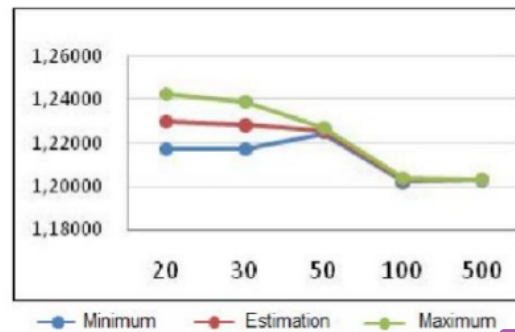**Figure 1.** The graph of the *a* estimation and its confident interval (CI).

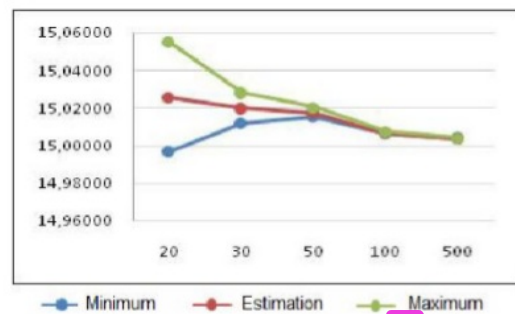**Figure 2.** The graph of the *b* estimation and its confident interval (CI).

Figure 3. The graph of the *p* estimation and its confidence interval (CI).
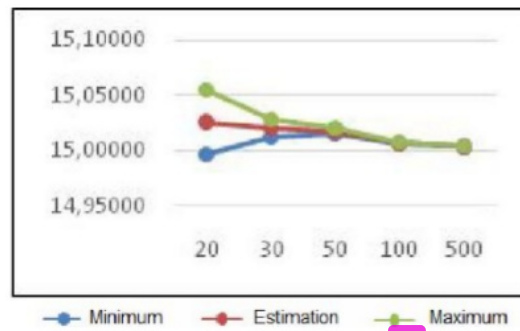
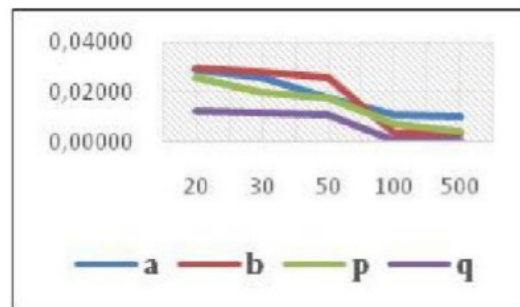**Figure 4.** The graph of the q estimation and its confidence interval (CI).

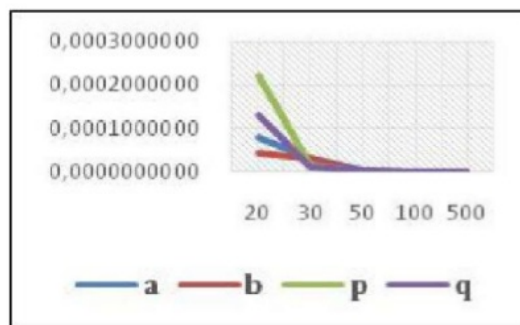**Figure 5.** The graph of bias for the parameters a, b, p and q of the (GB2).

Figure 6. The graph of variances for the parameters a, b, p and q of the (GB2).

4. Conclusion

The estimation of parameters of generalized beta of the second kind (GB2) uses maximum likelihood estimation (MLE) method. Then the estimation is continued by iteration and Newton-Raphson methods. The results show that the estimation of the parameters a, b, p and q attains the values close to the real value of the parameters, if the size of the sample increases. The larger the sample size, the narrower the confidence interval is. The larger the sample size is, the smaller the bias and the variance are.

Acknowledgement

References

[1]   L. C. Thurow, Analyzing the American income distribution, paper and proceedings, American Economics Association 60 (1970), 261-269.

[2]   C. Taille, Lorenz ordering within the generalized gamma family of income distributions, Statistical Distributions in Scientific Work, C. Taille, G. P. Patil and B. Balderssari, eds., Reidel, Boston, Vol. 6, 1981, pp. 182-192.

[3]   J. B. McDonald and B. C. Jensen, An analysis of estimators of alternative measures of income inequality associated with the gamma distribution function, J. Amer. Statist. Assoc. 74 (1979), 856-860.

[4]   A. B. Salem and T. D. Mount, A convenient descriptive model of income distribution: the gamma density, Econometrica 44 (1974), 963-970.

[5]   S. K. Singh and G. S. Maddala, A function for the size distribution of income, Econometrica 44 (1976), 963-970.

[6]   C. P. A. Bartels and H. Van Metele, Alternative Probability Density Functions of Income, Vrije University Amsterdam: Research Memorandum 29, 1975.

[7]   J. B. McDonald, Some generalized functions for the size distribution of income, Econometrica 52(3) (1984), 647-663.

[8] M. Graf and D. Nedyalkova, Modeling of income and indicators of poverty and social exclusion using the generalized beta distribution of the second kind, Review of Income and Wealth 60 (2014), 821-842. Doi: 10.1111/roiw.12031.

[9] Stephen P. Jenkins, Inequality and the GB2 income distribution, ISER, University of Essex, DIW Berlin and IZA Discussion Paper No. 2831, 2007.

[10] Stephen P. Jenkins, Distributionally-sensitive inequality indices and the GB2 income distribution, Review of Income and Wealth Series 55(2) (2009), 392-398.

[11] C. Kleiber, On the Lorenz order within parametric families of income distributions, Sankhya 61 (1999), 314-317.

[12] R. F. Bordley, J. B. McDonald and A. Mantrala, Something new, something old: parametric models for the size distribution of income, Journal of Income Distribution 6 (1996), 91-103.

[13] K. Brachmann, A. Stich and M. Trede, Evaluating parametric income distribution models, Allg. Stat. Arch. 80 (1996), 285-298.

[14] J. B. Butler and J. B. McDonald, Using incomplete moments to measure inequality, J. Econometrics 42 (1989), 109-119.

[15] J. B. McDonald and Y. J. Xu, A generalization of the beta distribution with applications, J. Econometrics 66 (1995), 133-152.

[16] S. Feng, R. V. Burkhauser and J. S. Butler, Levels and long-term trends in earnings inequality: overcoming current population survey censoring problems using the GB2 distribution, J. Bus. Econom. Statist. 24 (2006), 57-62.

[17] S. C. Parker, The generalized beta as a model for the distribution of earnings, Econom. Lett. 62 (1999), 197-200.

[18] C. Kleiber and S. Kotz, Statistical Size Distributions in Economics and Actuarial Sciences, Hoboken, Wiley-Interscience, NJ, USA, 2003.

Warsono: warsono.1963@fmipa.unila.ac.id

ESTIMATION OF PARAMETERS OF GENERALIZED BETA OF THE SECOND KIND (GB2) DISTRIBUTION BY MAXIMUM LIKELIHOOD ESTIMATION (MLE) AND NEWTON-RAPHSON ITERATION

ORIGINALITY REPORT

22%

SIMILARITY INDEX

PRIMARY SOURCES

1   meetings.sis-statistica.org
    Internet   40 words — 1%

2   www.econstor.eu
    Internet   36 words — 1%

3   Fabio Clementi, Mauro Gallegati. "The Distribution of
    Income and Wealth", Springer Nature, 2016   32 words — 1%
    Crossref

4   hal.archives-ouvertes.fr
    Internet   29 words — 1%

5   2013.isiproceedings.org
    Internet   28 words — 1%

6   Vijay P. Singh. "Systems of frequency distributions for
    water and environmental engineering", Physica A: 26 words — 1%
    Statistical Mechanics and its Applications, 2018
    Crossref

7   Haijun Hu, Guangxu Cheng, Yun Li, Yiping Tang. "Risk-
    based maintenance strategy and its applications in a   21 words — 1%
    petrochemical reforming reaction system", Journal of Loss
    Prevention in the Process Industries, 2009
    Crossref

8   Dubé, . "Appendix", Spatial Econometrics Using
    Microdata, 2014.   20 words — 1%

Crossref

9   J.S.K. Chan, S.T.B. Choy, U.E. Makov, Z. Landsman. "MODELLING INSURANCE LOSSES USING CONTAMINATED GENERALISED BETA TYPE-II DISTRIBUTION", ASTIN Bulletin, 2018
Crossref   20 words — 1%

10   Calderín-Ojeda, Enrique. "On the Composite Weibull–Burr Model to describe claim data", Communications in Statistics Case Studies Data Analysis and Applications, 2015.
Crossref   19 words — 1%

11   mpra.ub.uni-muenchen.de
Internet   19 words — 1%

12   www.researchgate.net
Internet   18 words — 1%

13   umbra.nascom.nasa.gov
Internet   18 words — 1%

14   José María Sarabia, Vanesa Jordá. "Explicit expressions of the Pietra index for the generalized function for the size distribution of income", Physica A: Statistical Mechanics and its Applications, 2014
Crossref   18 words — 1%

15   cejeme.eu
Internet   18 words — 1%

16   sankhya.isical.ac.in
Internet   17 words — 1%

17   afni.nimh.nih.gov
Internet   16 words — < 1%

18   A. Khosravi Tanak, G.R. Mohtashami Borzadaran, J. Ahmadi. "Entropy maximization under the constraints on the generalized Gini index and its application in modeling income distributions", Physica A: Statistical Mechanics and its Applications, 2015
Crossref   16 words — < 1%

19 J.David Cummins, Georges Dionne, James B. McDonald, B.Michael Pritchett. "Applications of the GB2 family of distributions in modeling insurance loss processes", Insurance: Mathematics and Economics, 1990
Crossref

16 words — < 1%

20 New Economic Windows, 2016.
Crossref

16 words — < 1%

21 Pavel V. Shevchenko. "Modelling Large Losses", Modelling Operational Risk Using Bayesian Inference, 2011
Crossref

16 words — < 1%

22 Cummins, J.. "Applications of the GB2 family of distributions in modeling insurance loss processes", Insurance Mathematics and Economics, 199012
Crossref

15 words — < 1%

23 www.sebastianpawlak.com
Internet

15 words — < 1%

24 Shuaizhang Feng. "Levels and Long-Term Trends in Earnings Inequality", Journal of Business and Economic Statistics, 01/2006
Crossref

15 words — < 1%

25 Graf, Monique, and Desislava Nedyalkova. "Modeling of Income and Indicators of Poverty and Social Exclusion Using the Generalized Beta Distribution of the Second Kind", Review of Income and Wealth, 2013.
Crossref

15 words — < 1%

26 Masbou, Jérémy, David Point, Jeroen E. Sonke et al. "Hg stable isotope time trend in ringed seals registers decreasing sea ice cover in the Alaskan Arctic", Environmental Science & Technology
Crossref

14 words — < 1%

27 "Handbook of Reliability Engineering", Springer Nature, 2003
Crossref

14 words — < 1%

28 McDonald, J.B.. "Regression models for positive random variables", Journal of Econometrics, 199001/02
Crossref

13 words — < 1%

29 Guanghan Liu. "Sample Size in Epidemiologic Studies", Wiley, 2014
Crossref

12 words — < 1%

30 Christian Kleiber. "A Guide to the Dagum Distributions", Modeling Income Distributions and Lorenz Curves, 2008
Crossref

12 words — < 1%

31 www.math.ualberta.ca
Internet

11 words — < 1%

32 web.mit.edu
Internet

11 words — < 1%

33 www.economicpolicyresearch.org
Internet

11 words — < 1%

34 Andrea Vinh, William E. Griffiths, Duangkamon Chotikapanich. "Bivariate income distributions for assessing inequality and poverty under dependent samples", Economic Modelling, 2010
Crossref

11 words — < 1%

35 link.springer.com
Internet

10 words — < 1%

36 S. Mirzaei, G. R. Mohtashami Borzadaran, M. Amini, H. Jabbari. "A new generalized Weibull distribution in income economic inequality curves", Communications in Statistics - Theory and Methods, 2018
Crossref

9 words — < 1%

37 www.lisproject.org
Internet

9 words — < 1%

38	www.tandfonline.com Internet	8 words — < 1%
39	sebgod.free.fr Internet	8 words — < 1%
40	José María Sarabia. "Parametric Lorenz Curves: Models and Applications", Modeling Income Distributions and Lorenz Curves, 2008 Crossref	8 words — < 1%
41	C S M Currie, R C H Cheng, H K Smith. "Dynamic pricing of airline tickets with competition", Journal of the Operational Research Society, 2017 Crossref	8 words — < 1%
42	aip.scitation.org Internet	8 words — < 1%
43	Duangkamon Chotikapanich. "Estimating Lorenz Curves Using a Dirichlet Distribution", Journal of Business and Economic Statistics, 04/2002 Crossref	8 words — < 1%
44	Yang, X.. "A generalized beta copula with applications in modeling multivariate long-tailed data", Insurance Mathematics and Economics, 201109 Crossref	7 words — < 1%
45	M. Franic, M. Kujundzic Tiljak, M. Pozar, D. Romic, M. Mimica, J. Petrak, D. Ivankovic, M. Pecina. "Anterior versus posterior approach in 3D correction of adolescent idiopathic thoracic scoliosis: A meta-analysis", Orthopaedics & Traumatology: Surgery & Research, 2012 Crossref	6 words — < 1%
46	John F. Kottas, Hon-Shiang Lau. "The Use of Versatile Distribution Families in Some Stochastic Inventory Calculations", Journal of the Operational Research Society, 2017 Crossref	6 words — < 1%