

On the Kumaraswamy Fisher Snedecor Distribution

Adepoju, K.A*, Chukwu, A.U, Shittu, O.I

Department of Statistics, University of Ibadan, Nigeria

Copyright©2016 by authors, all rights reserved. Authors agree that this article remains permanently open access under the terms of the Creative Commons Attribution License 4.0 International License

Abstract We propose the Kumaraswamy-F (KUMAF) distribution which is a generalization of the conventional Fisher Snedecor (F-distribution). The new distribution can be used even when one or more of the regular assumptions are violated. It is obtained with the addition of two shape parameters to a continuous F-distribution which is commonly used to test the null hypothesis in the Analysis of Variance (ANOVA test). The statistical properties of the proposed distribution such as moments, moment generating function, the asymptotic behavior among others were investigated. The method of maximum likelihood is used to estimate the model parameters and the observed information matrix is derived. The distribution is found to be more flexible and robust to regular assumptions of the conventional F-distribution. In future research, the flexibility of this distribution as well as its robustness using a real data set will be examined. The new distribution is recommended for used in most applications where the assumption underlying the use of conventional F distribution for one-way analysis of variance are violated such as homogeneity of variance or normality assumption probably as result of the presence of outlier(s). It is instructive to note that the new distribution preserves the originality of the data without transformation.

Keywords Fisher-Snedecor Distribution, Kumaraswamy-F Distribution, One Way ANOVA, Outlier, Maximum Likelihood Method

1. Introduction

An **F-test** is any statistical test in which the test statistic has an F-distribution under the null hypothesis. It is most often used when comparing statistical models that have been fitted to a data set, in order to identify the model that best fits the population from which the data were sampled. Exact *F-tests* mainly arise when the models have been fitted to the data using least squares. The name was coined by George W. Snedecor, in honour of Sir Ronald A. Fisher. Fisher initially developed the statistic as the variance ratio in the 1920s. The efficient use of classical F-test to compare several (k)

population means relies heavily on a good number of assumptions of the Analysis of Variance (ANOVA).

For the F-ratio statistic there are two fundamental assumptions: the variances of the compared populations are the same; the estimates of the population variance are independent.

Therefore before we proceed with an analysis of the data we have collected we have to make sure that these assumptions have been met.

These assumptions include independent of k populations being tested, equality of the population variances, and absence of outlier among others. When a number of the above assumptions are violated, the use of F-test result to test for the equality of the regresand becomes incorrect or misleading. For example, if the assumption of independence is violated, then the one-way ANOVA is simply not appropriate. Similarly, when the assumption of normality or unequal variances is violated, the classical F-test fails to reject the null hypothesis even if the data actually provide strong evidence for it. A potentially more damaging violation of assumption occurs when one are more of the populations being tested are not normally distributed probably due to the presence of outliers. This occurs more especially when the sample sizes are not equaled (unbalanced). Often, the effect of an assumption violation on the result of one-way ANOVA depends on the extent of the violation (such as how unequal the population variances are, or how heavy-tailed one or another population distribution is). Some small violations may have little practical effect on the analysis, while other violations may render the one-way ANOVA result uselessly incorrect. Krutchkoff [10] discussed some misconceptions about the F-test and provided a simulation based solution to overcome drawbacks of the test. ANOVA under heteroscedasticity is a Behrens-Fisher's type problem. The Behrens-Fisher problem, named after Ronald Fisher and W. V. Behrens, is the problem of interval estimation and hypothesis testing concerning the difference between the means of two normally distributed populations when the variances of the two populations are not assumed to be equal, based on two independent samples. Tsui and Weerahandi [14] generalized the conventional definition of the p-value from the F-distribution so that

problems such as the Behrens-Fisher problem can be resolved. Weerahandi [15] discussed the numerical equivalence of this test with the representation of Behrens-Fisher solution given in Barnard [1]. Rice and Gaines [12] extended the p-value given in Barnard [1] to the one-way ANOVA case. Samaradasa and Weerahandi [16] extended the representation of the two-sample test given in Tsui and Weerahandi [14] to the one-way ANOVA case and provide exact tests for making multiple comparisons for means and variances. This test referred to as the generalized F-test for one-way ANOVA is numerically equivalent in performance to the test in Rice and Gaines [12].

This paper therefore focuses on developing a generalized F-distribution that is capable of handling data that are non-normal and non-Gaussian and probably infected with outliers. The proposed distributions will be used to develop p-value that is less sensitive to any serious model assumption violations.

2. The Kumaraswamy-F Distribution (KUMA-F)

In this section Kumaraswamy-F distribution is developed by compounding the tractable link function developed by Kumaraswamy [8] with conventional F-distribution by Fisher Snedecor to obtain the Kumaraswamy-F distribution.

A random variable X is distributed as the Kumaraswamy-F distribution if its PDF is derived as follows:

$$f(x) = \frac{r_1^{\frac{r_1}{2}} x^{\frac{r_1-1}{2}}}{B\left(\frac{r_1}{2}, \frac{r_2}{2}\right) r_2^{\frac{r_1}{2}} \left(1 + \frac{r_1 x}{r_2}\right)^{\frac{r_1+r_2}{2}}} \quad (1)$$

$$F(x) = I_{\frac{r_1 x}{r_2}} = \frac{\beta\left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2}\right)}{\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right)} = A \quad (2)$$

The Kumaraswamy generator (link function) as given by Jones [7] as

$$g(x) = bc [F(X)]^{c-1} [1 - F(X)^c]^{b-1} f(x) \quad (3)$$

then using (1) and (2) in (3), we obtain the probability density function of the Kumaraswamy-F distribution as:

$$g(x) = \frac{bc I_{\frac{r_1 x}{r_2}}^{c-1} \left(1 - I_{\frac{r_1 x}{r_2}}^c\right)^{b-1} r_1^{\frac{r_1}{2}} x^{\frac{r_1}{2}} - 1}{B\left(\frac{r_1}{2}, \frac{r_2}{2}\right) r_2^{\frac{r_1}{2}} \left(1 + \frac{r_1 x}{r_2}\right)^{\frac{r_1+r_2}{2}}} \quad (4)$$

$$b > 0, c > 0, x > 0, r_1 > 0, r_2 > 0$$

The graph of the above pdf is given in Fig. 1 below:

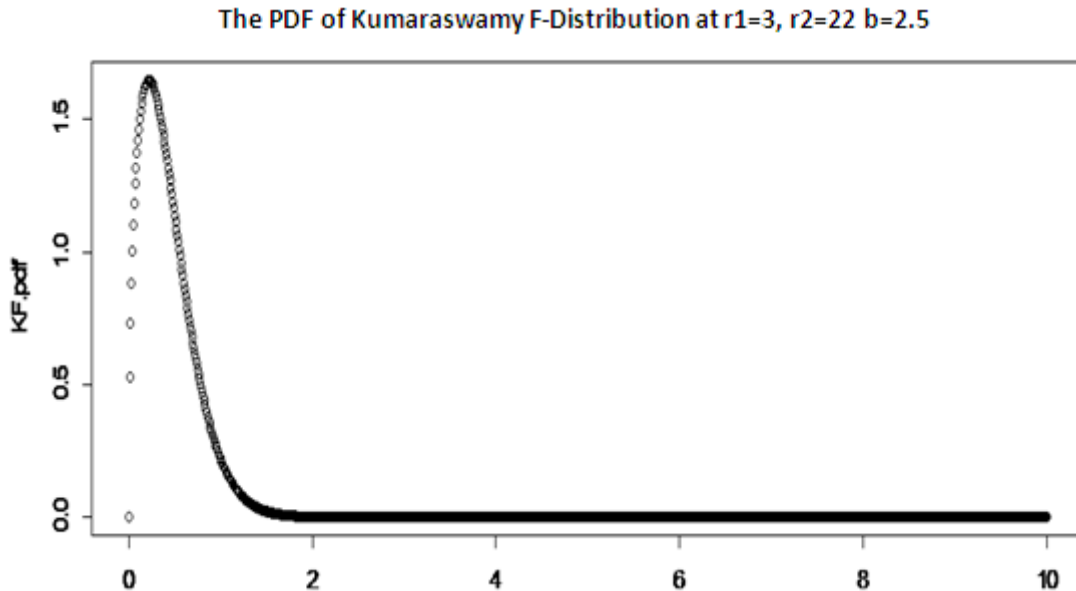


Figure 1. The PDF of Kumaraswamy F-Distribution at r1=3, r2=22 b=2.5

Using the appropriate transformation, one can show that

$$\int_0^{\infty} g(x)dx = 1$$

Let

$$A = \frac{I_{r_1 x}}{r_2} \text{ such that } \frac{dA}{dx} = \frac{\left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}} x^{\frac{r_1-1}{2}}}{\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right) \left(1 + \frac{r_1 x}{r_2}\right)^{\frac{r_1+r_2}{2}}}$$

Then

$$g(x) = bc A^{c-1} (1 - A^c)^{b-1} \frac{dA}{dx}$$

$$\int_{-\infty}^{\infty} g(x)dx = \int_0^1 bc A^{c-1} (1 - A^c)^{b-1} dA = 1$$

Let $M = A^c$ $\frac{dM}{dA} = c A^{c-1}$

$$dA = \frac{dM}{c A^{c-1}} \quad A = M^{\frac{1}{c}}$$

$$\begin{aligned} \int_{-\infty}^{\infty} g(x)dx &= \int_0^1 bc \left(M^{\frac{1}{c}}\right)^{c-1} (1 - M^c)^{b-1} \frac{dM}{c A^{c-1}} \\ &= b \int_0^1 M^{1-\frac{1}{c}} (1 - M)^{b-1} \frac{dM}{\left(M^{\frac{1}{c}}\right)^{c-1}} \\ &= b \int_0^1 (1 - M)^{b-1} dM \\ &= 1 \end{aligned}$$

This verifies that $g(x)$ is indeed a probability density function of a continuous distribution.

3. Cumulative Distribution Function

Lemma1: Given that $X \sim KUMA - F(b, c, r_1, r_2)$, its distribution function is expressed as

$$G(x) = 1 - \left[1 - \frac{\beta\left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2}\right)}{\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right)} \right]^c \Bigg]^b$$

Proof

$$G(x) = P(X \leq x) = \int_0^x bc I_{\frac{r_1 t}{r_2}}^{c-1} \left(1 - I_{\frac{r_1 t}{r_2}}^c\right)^{b-1} \frac{\left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}} t^{\frac{r_1-1}{2}}}{B\left(\frac{r_1}{2}, \frac{r_2}{2}\right) \left(1 + \frac{r_1 t}{r_2}\right)^{\frac{r_1+r_2}{2}}} dt$$

Let $P = I_{\frac{r_1 t}{r_2}}$ and $\frac{dP}{dt} = \frac{\left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}} t^{\frac{r_1-1}{2}}}{B\left(\frac{r_1}{2}, \frac{r_2}{2}\right) \left(1 + \frac{r_1 t}{r_2}\right)}$

We can be simplified as follows

$$\begin{aligned} G(x) &= b \int_0^M (1 - P^c)^{b-1} dP \\ G(x) &= 1 - (1 - P^c)^b \end{aligned} \tag{5}$$

Recall that $P = \frac{\beta\left(\frac{r_1 x}{r_2}; \frac{r_1}{2}, \frac{r_2}{2}\right)}{\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right)}$ then

$$G(x) = 1 - \left[1 - \frac{\beta\left(\frac{r_1 x}{r_2}; \frac{r_1}{2}, \frac{r_2}{2}\right)^c}{\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right)^c} \right]^b \quad (6)$$

The plot of the CDF for various values of b, c, r_1 and r_2 is plotted below

The CDF of Kumaraswamy-F distribution at $b=2.5, c=5, r_1=3, r_2=22$

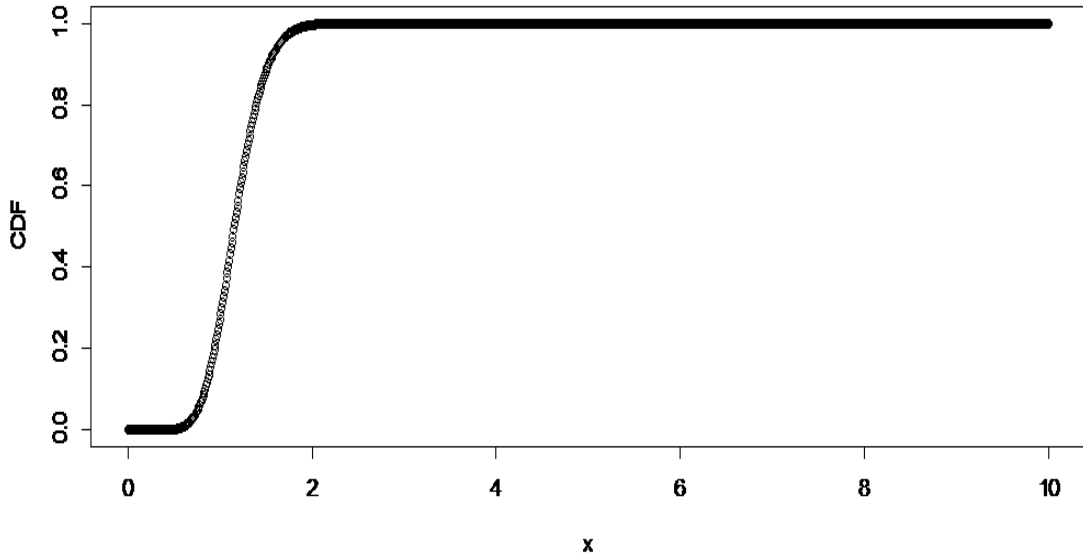


Figure 2. CDF of Kumaraswamy distribution

4. Limit Behaviour

In this section, we investigate the limit behavior of Kumaraswamy-F distribution as $x \rightarrow \infty$ and as $x \rightarrow 0$. This can be achieved by taking the limit of the equation (4).

For $x \rightarrow \infty$

$$\lim_{x \rightarrow \infty} g(x) = \lim_{x \rightarrow \infty} bc \left(I_{\frac{r_1 x}{r_2}} \right)^{c-1} \left(1 - I_{\frac{r_1 x}{r_2}}^c \right)^{b-1} \frac{1}{\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right)} \left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}} x^{\frac{r_1}{2}-1} \frac{1}{\left(1 + \frac{r_1}{r_2} x\right)^{\frac{r_1+r_2}{2}}}$$

This tends to zero because

$$\lim_{x \rightarrow \infty} \frac{1}{\left(1 + \frac{r_1 x}{r_2}\right)^{\frac{r_1+r_2}{2}}} = 0$$

Similarly, as $x \rightarrow 0$, $\lim g(x) = 0$, this is because $\lim_{x \rightarrow 0} x^{\frac{r_1}{2}-1} = 0$

The above indicates that the proposed distribution has a mode.

5. Hazard Rate Function

The hazard rate function of a random variable X with the probability density function is obtained using

$$h(x) = \frac{g(x)}{1 - G(x)}$$

then using (4) and (6) in the above expression ,the hazard rate function of the Kumaraswamy-F (KUMA-F)distribution is

$$h(x) = \frac{\left[\frac{bc I_{\frac{r_1 x}{r_2}}^{c-1} \left(1 - I_{\frac{r_1 x}{r_2}}^c \right)^{b-1} r_1^{\frac{r_1}{2}} x^{\frac{r_1}{2}-1}}{B\left(\frac{r_1}{2}, \frac{r_2}{2}\right) r_2^{\frac{r_1}{2}} \left(1 + \frac{r_1 x}{r_2} \right)^{\frac{r_1+r_2}{2}}} \right]}{\left[1 - \frac{\left(\beta\left(\frac{r_1 x}{r_2}; \frac{r_1}{2}, \frac{r_2}{2}\right) \right)^c}{\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right)} \right]^b} \tag{7}$$

6. Estimation and Information Matrix

Let X be a random variable with the KUMA-F distribution (4.). The log-likelihood function is

$$\begin{aligned} l(b, c, r_1, r_2) = & n \log C - n \log b - n \log \beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right) + (c-1) \sum_{i=1}^n \log \frac{\beta\left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2}\right)}{\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right)} - (c-1) \log \beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right) + \\ & (b-1) \sum_{i=1}^n \log \left[\beta^c\left(\frac{r_1}{2}, \frac{r_2}{2}\right) - \beta\left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2}\right) \right] - (b-1) \log \beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right) + \frac{nr_1}{2} \log r_1 - \frac{nr_1}{2} \log r_2 \\ & + \left(\frac{r_1}{2} - 1\right) \sum_{i=1}^n \log x - \left(\frac{r_1+r_2}{2}\right) \sum_{i=1}^n \log \left(1 + \frac{r_1 x}{r_2} \right) \end{aligned} \tag{8}$$

By differentiating (8) with respect to parameters b, c, r_1 and r_2

$$\frac{\partial l}{\partial b} = -n \frac{\Gamma(b)'}{\Gamma(b)} + n \frac{\Gamma(1+b)'}{\Gamma(1+b)} - c \log \beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right) + \sum_{i=1}^n \frac{\delta}{\delta c} \log \left[\beta^c\left(\frac{r_1}{2}, \frac{r_2}{2}\right) - \beta^c\left(\frac{r_1}{2}, \frac{r_2}{2}\right) \right] \tag{9}$$

$$\frac{\partial l}{\partial c} = \frac{n}{c} + \sum_{i=1}^n \log \beta \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) - \log \beta \left(\frac{r_1}{2}, \frac{r_2}{2} \right) + (b-1) \sum_{i=1}^n \frac{\delta}{\delta c} \log \left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right] - (b-1) \log \beta \left(\frac{r_1}{2}, \frac{r_2}{2} \right) \quad (10)$$

$$\begin{aligned} \frac{\partial \delta}{\partial r_1} = & -n \frac{\left(\frac{r_1}{2} \right)^{-1}}{\left(\frac{r_1}{2} \right)} + n \frac{\left(\frac{r_1+r_2}{2} \right)^{-1}}{\left(\frac{r_1+r_2}{2} \right)} + \frac{(c-1) \sum_{i=1}^n \frac{\delta}{\delta r_1} \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)}{\beta \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)} - (c-1) \left[\frac{\left(\frac{r_1}{2} \right)^{-1}}{\left(\frac{r_1}{2} \right)} - \frac{\left(\frac{r_1+r_2}{2} \right)^{-1}}{\left(\frac{r_1+r_2}{2} \right)} \right] + (b-1) \\ & \sum_{i=1}^n \left\{ \frac{\delta \left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right]}{\delta r_1 \left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right]} \right\} - (b-1) \left[\frac{\left(\frac{r_1}{2} \right)^{-1}}{\left(\frac{r_1}{2} \right)} - \frac{\left(\frac{r_1+r_2}{2} \right)^{-1}}{\left(\frac{r_1+r_2}{2} \right)} \right] + \frac{n}{2} \log r_1 + \frac{n}{2} - \frac{n}{2} \log r_2 + \frac{1}{2} \sum_{i=1}^n \log x - \\ & \left[\frac{1}{2} \sum_{i=1}^n \log \left(1 + \frac{r_1 x}{r_2} \right) + \left(\frac{r_1+r_2}{2} \right) \sum_{i=1}^n \frac{x}{\left(1 + \frac{r_1 x}{r_2} \right)} \right] \quad (11) \end{aligned}$$

$$\begin{aligned} \frac{\partial \delta}{\partial r_2} = & -n \frac{\left(\frac{r_1}{2} \right)^{-1}}{\left(\frac{r_2}{2} \right)} + n \frac{\left(\frac{r_1+r_2}{2} \right)^{-1}}{\left(\frac{r_1+r_2}{2} \right)} + (c-1) \sum_{i=1}^n \frac{\delta}{\delta r_2} \frac{\beta \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)}{\beta \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)} - (c-1) \left[\frac{\left(\frac{r_2}{2} \right)^{-1}}{\left(\frac{r_2}{2} \right)} - \frac{\left(\frac{r_1+r_2}{2} \right)^{-1}}{\left(\frac{r_1+r_2}{2} \right)} \right] \\ & + (b-1) \sum_{i=1}^n \frac{\delta \left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right]}{\delta r_2 \left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right]} - (b-1) c \left[\frac{\left(\frac{r_2}{2} \right)^{-1}}{\left(\frac{r_2}{2} \right)} - \frac{\left(\frac{r_1+r_2}{2} \right)^{-1}}{\left(\frac{r_1+r_2}{2} \right)} - \frac{nr_1}{2r_2} \right] \\ & - \left[\frac{1}{2} \sum_{i=1}^n \log \left(1 + \frac{r_1 x}{r_2} \right) + \left(\frac{r_1+r_2}{2} \right)^n \sum_{i=1}^n \frac{-\left(\frac{r_1}{r_2^2} x \right)}{1 + \frac{r_1 x}{r_2}} \right] \quad (12) \end{aligned}$$

noting that $\phi(k) = \sqrt[k]{k}$ equations(9), (10), (11) and (12) can be re-expressed as

$$\frac{\delta l}{\delta b} = -n\phi(b) + n\phi(1+b) + \sum_{i=1}^n \log \left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right] - C \log \beta \left(\frac{r_1}{2}, \frac{r_2}{2} \right) \quad (13)$$

$$\frac{\delta \ell}{\delta c} = \frac{n}{c} + \sum_{i=1}^n \log \beta \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) - \log \beta \left(\frac{r_1}{2}, \frac{r_2}{2} \right) + (b-1) \sum_{i=1}^n \frac{\delta}{\delta c} \log \left[\beta \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right] - (b-1) \log \beta \left(\frac{r_1}{2}, \frac{r_2}{2} \right) \tag{14}$$

$$\frac{\delta \ell}{\delta r_1} = -n \varphi \left(\frac{r_1}{2} \right) + n \varphi \left(\frac{r_1 + r_2}{2} \right) + (b-1) \sum_{i=1}^n \frac{\delta}{\delta r_1} \frac{\left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right]}{\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)} - (b-1) \left[\varphi \left(\frac{r_1}{2} \right) - \varphi \left(\frac{r_1 + r_2}{2} \right) \right] - \left(\frac{r_1 + r_2}{2} \right) \sum_{i=1}^n \left(\frac{x}{r_2 + r_1 x} \right) \tag{15}$$

$$\begin{aligned} \frac{\delta \ell}{\delta r_2} = & -n \varphi \left(\frac{r_2}{2} \right) + n \varphi \left(\frac{r_1 + r_2}{2} \right) + (c-1) \sum_{i=1}^n \frac{\delta}{\delta r_2} \frac{\beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)}{\beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)} - (c-1) \varphi \left(\frac{r_2}{2} \right) + (c-1) \varphi \beta \left(\frac{r_1 + r_2}{2} \right) \\ & + (b-1) \sum_{i=1}^n \frac{\frac{\delta}{\delta r_2} \beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)}{\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)} - (b-1) \varphi \beta^c \left(\frac{r_2}{2} \right) + (b-1) \varphi \left(\frac{r_1 + r_2}{2} \right) - \frac{n r_1}{\delta r_2} \\ & - \frac{1}{2} \sum_{i=1}^n \log \left(1 + \frac{r_1 x}{r_2} \right) + \left(\frac{r_1 + r_2}{2} \right) \sum_{i=1}^n \left(\frac{r_1 x}{r_2^2 + r_1 r_2 x} \right) (b-1) \end{aligned} \tag{16}$$

Equations (13), (14), (15) and (16) above are solved numerically for b, c, r_1 and r_2 to obtain the respective estimates $\hat{b}, \hat{c}, \hat{r}_1$, and \hat{r}_2 .

7. Fisher Information

The Fisher Information Matrix is obtained by finding the derivatives of the equations (13), (14), (15) and (16) which in turn give the diagonal elements of the Fisher's Information Matrix. They are as follows:

$$\frac{\delta^2 \ell}{\delta b^2} = -n \Psi(b) + n \Psi(a + b) \tag{17}$$

$$\frac{\delta^2 \ell}{\delta b \delta c} = \sum_{i=1}^n \frac{\delta}{\delta c} \frac{\left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right]}{\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)} - \log B^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) \tag{18}$$

$$\frac{\delta l}{\delta b \delta r_1} = \sum_{i=1}^n \frac{\delta}{\delta r_1} \frac{\left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right]}{\left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right]} - C \varphi \left(\frac{r_1}{2} \right) + C \varphi \left(\frac{r_1 + r_2}{2} \right) \quad (19)$$

$$\frac{\delta^2 l}{\delta b \delta r_1} = \sum_{i=1}^n \frac{\delta}{\delta r_2} \frac{\left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right]}{\left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right]} - C \varphi \left(\frac{r_2}{2} \right) + C \varphi \left(\frac{r_1 + r_2}{2} \right) \quad (20)$$

$$\frac{\delta^2 l}{\delta c^2} = \frac{n}{C^2} + (b-1) \sum_{i=1}^n \frac{\delta^2}{\delta c^2} \left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right] \quad (21)$$

$$\frac{\delta^2 l}{\delta c r_1} = a \sum_{i=1}^n \frac{\delta^2}{\delta r_1} \frac{\left[\beta^c \left(\frac{r_1 x}{r}, \frac{r_1}{r}, \frac{r_2}{r} \right) \right]}{\beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)} - a \varphi \left(\frac{r_1}{2} \right) + a \varphi \left(\frac{r_1 + r_2}{2} \right) + (b-1)$$

$$\sum_{i=1}^n \frac{\delta^2}{\delta r_1 \delta c} \log \left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right] - (b-1) \varphi \left(\frac{r_1}{2} \right) + (b-1) \varphi \left(\frac{r_1 + r_2}{2} \right) \quad (22)$$

$$\frac{\delta l}{\delta c \delta r_2} = \sum_{i=1}^n \frac{\delta}{\delta r_2} \frac{\beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)}{\beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)} - \varphi \left(\frac{r_2}{2} \right) + \varphi \left(\frac{r_1 + r_2}{2} \right) - (b-1) \varphi \left(\frac{r_2}{2} \right) + (b-1) \varphi \left(\frac{r_1 + r_2}{2} \right) \quad (23)$$

$$\begin{aligned} \frac{\delta^2 l}{\delta r_1^2} &= n \varphi \left(\frac{r_1}{2} \right) + n \varphi \left(\frac{r_1 + r_2}{2} \right) + (b-1) \sum_{i=1}^n \frac{\delta^2}{\delta r_2} \frac{\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)}{\left[\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right) \right]} \\ &- (b-1) \varphi \left(\frac{r_1}{2} \right) + (b-1) \varphi \left(\frac{r_1 + r_2}{2} \right) + \frac{n}{2r_1} - \frac{1}{2} \sum_{i=1}^n \left(\frac{x}{r_2 + r_1 x} \right) - \left(\frac{r_1 + r_2}{2} \right) \sum_{i=1}^n \frac{x^2}{(r_2 + r_1 x)} \end{aligned} \quad (24)$$

$$\begin{aligned} \frac{\delta^2 l}{\delta r_1 \delta r_2} &= -n \varphi \left(\frac{r_1}{2} \right) + n \varphi \left(\frac{r_1 + r_2}{2} \right) + (b-1) \sum_{i=1}^n \frac{\delta^2}{\delta r_1 \delta r_2} \frac{\left[\beta^c \left(\frac{r_1}{r}, \frac{r_2}{r} \right) - \beta^c \left(\frac{r_1 x}{r}, \frac{r_1}{r}, \frac{r_2}{r} \right) \right]}{\beta^c \left(\frac{r_1}{2}, \frac{r_2}{2} \right) - \beta^c \left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2} \right)} \\ &+ (b-1) \varphi \left(\frac{r_1 + r_2}{2} \right) - \frac{n}{2} \frac{r_1}{r_2^2} \sum_{i=1}^n \frac{x}{\left(1 + \frac{r_1 x}{r_2} \right)} - \frac{1}{2} \sum_{i=1}^n \frac{x}{r_2 + r_1 x} + \left(\frac{r_1 + r_2}{2} \right) \sum_{i=1}^n \frac{x}{(r_2 + r_1 x)^2} \end{aligned} \quad (25)$$

$$\frac{\delta \ell^2}{\delta r_2^2} = -n \phi\left(\frac{r_2}{2}\right) + n \phi\left(\frac{r_1+r_2}{2}\right) + (c-1) \sum_{i=1}^n \frac{\delta^2}{\delta r_2^2} \frac{\beta^c\left(\frac{r_1}{2}, \frac{r_2}{2}\right) - \beta^c\left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2}\right)}{\left[\beta^c\left(\frac{r_1}{2}, \frac{r_2}{2}\right) - \beta^c\left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2}\right)\right]} - (b-1) \phi\left(\frac{r_2}{2}\right) \quad (26)$$

8. Moments and Generating Functions

8.1. Generating Functions

We derive the moment generating function and characteristic function for a random variable X having the KUMA-F density Function given in equation as follows.

By definition, the moment generating function of a random variable X is defined as

$$M_x(t) = E(e^{tx}), \text{ where } |t| < 1.$$

Using equation (1), we have

$$M_x(t) = bc \left(\frac{r_1}{r_2}\right)^{\frac{r_1}{r_2}} \frac{1}{B\left(\frac{r_1}{2}, \frac{r_2}{2}\right)} \frac{\int_0^\infty e^{tx} I_{\frac{r_1 x}{r_2}}^{c-1} \left(1 - I_{\frac{r_1 x}{r_2}}^c\right)^{b-1} x^{\frac{r_1}{2}-1}}{\left(1 + \frac{r_1 x}{r_2}\right)^{\frac{r_1+r_2}{2}}} dx$$

For real integer $b > 0$

$$\left(1 - I_{\frac{r_1 x}{r_2}}^c\right)^{b-1} = \sum_{i=1}^{b-1} \binom{b-1}{i} (-1)^i I_{\frac{r_1 x}{r_2}}^{ci}$$

$$M_x(t) = bc \left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}} \frac{1}{B\left(\frac{r_1}{2}, \frac{r_2}{2}\right)} \sum_{i=0}^{b-1} (b-1)(-1)^i \frac{\int_0^\infty e^{tx} I_{\frac{r_1 x}{r_2}}^{c-1} x^{\frac{r_1}{2}-1}}{\left(1 + \frac{r_1 x}{r_2}\right)^{\frac{r_1+r_2}{2}}} dx$$

Recall that

$$A = I_{\frac{r_1 x}{r_2}} \frac{dI}{dx} = \frac{\left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}} x^{\frac{r_1}{2}-1}}{B\left(\frac{r_1}{2}, \frac{r_2}{2}\right) \left(1 + \frac{r_1 x}{r_2}\right)^{\frac{r_1+r_2}{2}}}$$

Since

$$e^{tx} = \sum_{k=0}^{\infty} \frac{(tx)^k}{K_1^1} ; \quad dx = \frac{\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right) \left(1 + \frac{r_1 x}{r_2}\right)^{\frac{r_1+r_2}{2}}}{\left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}} x^{\frac{r_1-1}{2}}} dI$$

$$M_x(t) = bc \left(\frac{r_1}{r_2}\right)^{\frac{r_1}{r}} \sum_{i=0}^{b-1} \sum_{k=0}^{\infty} (-1)^i \binom{b-1}{i} t^k \int_0^{\frac{r_1 x}{r_2}} x^k I^{\frac{c+ci-1}{r_2}} x^{\frac{r_1-1}{2}} dI$$

Since

$$I = \beta\left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2}\right) \Big| \beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right)$$

From which

$$\beta\left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2}\right) = \int_0^{\frac{r_1 x}{r_2}} y^{\frac{r_1}{r_2}} (1+y)^{-\left(\frac{r_1+r_2}{2}\right)} dy$$

$$(1+y)^{-\left(\frac{r_1+r_2}{2}\right)} = 2^{-\left(\frac{r_1+r_2}{2}\right)} \left[1 - \left(\frac{1-y}{2}\right)\right]^{-\left(\frac{r_1+r_2}{2}\right)}$$

$$= 2^{-\left(\frac{r_1+r_2}{2}\right)} \sum_{j=0}^{\infty} (-1)^j \frac{(1-y)^j}{2^j} \frac{\left(\frac{r_1+r_2}{2} + j\right)}{\left(\frac{r_1+r_2}{2}\right)}$$

Since $y = \frac{r_1 x}{r_2}$, we have

$$\left(1 + \frac{r_1 x}{r_2}\right)^{-\left(\frac{r_1+r_2}{r}\right)} = 2^{-\left(\frac{r_1+r_2}{r}\right)} \sum_{j=0}^{\infty} (-1)^j \frac{\left(1 - \frac{r_1 x}{r_2}\right)^j}{2^j} \frac{\left(\frac{r_1+r_2}{2} + j\right)}{\left(\frac{r_1+r_2}{2}\right)}$$

Noting that

$$\left(1 - \frac{r_1 x}{r_2}\right)^j = \sum_{m=0}^j (-1)^m \left(\frac{r_1 x}{r_2}\right)^m \binom{j}{m}$$

$$\left(1 + \frac{r_1 x}{r_2}\right)^{-\left(\frac{r_1+r_2}{2}\right)} = 2^{-\left(\frac{r_1+r_2}{2}\right)} \sum_{j=0}^{\infty} \sum_{m=0}^j (-1)^{m+j} \frac{\left(\frac{r_1 x}{r_2}\right)^m \binom{j}{m} \left(\frac{r_1+r_2}{2} + j\right)}{2^j}$$

Then

$$\begin{aligned} \beta\left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2}\right) &= \int_0^{\frac{r_1 x}{r_2}} y^{\frac{r_1 x}{r_2} - 1 + m} \frac{2^{-\left(\frac{r_1+r_2}{r}\right)} \sum_{j=0}^{\infty} \sum_{m=0}^j (-1)^{m+j} \binom{j}{m} \left(\frac{r_1+r_2}{2} + j\right)}{\left(\frac{r_1+r_2}{2}\right)^{2^j}} dy \\ &= \frac{2^{-\left(\frac{r_1+r_2}{2}\right)} \sum_{j=0}^{\infty} \sum_{m=0}^j (-1)^{m+j} \binom{j}{m} \left(\frac{r_1+r_2}{2} + j\right)}{\left(\frac{r_1+r_2}{2}\right)^{2^j}} \int_0^{\frac{r_1 x}{r_2}} y^{\frac{r_1 x}{r_2} - 1 + m} dy \\ &= \frac{2^{-\left(\frac{r_1+r_2}{2}\right)} \sum_{j=0}^{\infty} \sum_{m=0}^j (-1)^{m+j} \binom{j}{m} \left(\frac{r_1+r_2}{2} + j\right)}{\left(\frac{r_1+r_2}{2}\right)^{2^j}} \left(\frac{r_1}{r_2}\right)^{\frac{r_1+m}{2}} \frac{x^{\frac{r_1+m}{2}}}{\left(\frac{r_1}{2} + m\right)} dy \end{aligned}$$

Since

$$I = \frac{\beta\left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2}\right)}{\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right)}; \frac{dI}{dx} = \frac{\left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}} x^{\frac{r_1-1}{2}}}{\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right) \left(1 + \frac{r_1 x}{r_2}\right)^{\frac{r_1+r_2}{2}}}$$

This implies

$$\begin{aligned} dI &= \frac{\left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}} x^{\frac{r_1-1}{2}} dx}{\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right) \left(1 + \frac{r_1 x}{r_2}\right)^{\frac{r_1+r_2}{2}}} \\ M_x(t) &= bc \left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}} \sum_{i=0}^{b-1} \sum_{r=0}^{\infty} (-1)^i \binom{b-1}{i} t^r \int_0^1 x^r \frac{x^{\frac{r_1-1}{2}}}{B\left(\frac{r_1}{2}, \frac{r_2}{2}\right) \left(1 + \frac{r_1 x}{r_2}\right)^{\frac{r_1+r_2}{2}}} dx \frac{\left[\beta\left(\frac{r_1 x}{r_2}, \frac{r_1}{2}, \frac{r_2}{2}\right)\right]^{c+c_i-1}}{\left[\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right)\right]^{c+c_i-1}} \\ &= bc \left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}} \frac{1}{\left[\beta\left(\frac{r_1}{2}, \frac{r_2}{2}\right)\right]^{a+c-1}} \sum_{i=0}^{b-1} \sum_{r=0}^{\infty} \frac{(-1)^i \binom{b-1}{i} t^r}{r!} \left[\sum_{j=0}^{\infty} \sum_{m=0}^j \frac{(-1)^{m+j} \left(\frac{r_1+r_2}{2} + j\right)}{2^{j+\frac{r_1+r_2}{2}}} \right] \end{aligned}$$

$$* \left(\frac{r_1}{r_2} \right)^{\frac{r_1+m}{2}} \frac{1}{\left(\frac{r_1}{2} + m \right)} \left[\frac{r_1+r_2}{2} \right]^{c+ci-1} \int_0^\infty x^{\left(\frac{r_1+m}{2} \right)(c+ci-1)} + \frac{r_1}{2} - 1 \left(1 + \frac{r_1 x}{2} \right)^{-\frac{r_1+r_2}{2}} dx$$

If we consider

$$\int_0^\infty x^{\left(\frac{r_1+m}{2} \right)(c+ci-1)+r+\frac{r_1}{2}-1} \left(1 + \frac{r_1 x}{r_2} \right)^{-\frac{r_1+r_2}{2}} dx$$

$$\text{Let } y = \frac{r_1 x}{r_2} \quad \frac{dy}{dx} = \frac{r_1}{r_2} \quad dx = \frac{r_2}{r_1} dy$$

$$x = \frac{r_2}{r_1} y$$

$$\int_0^\infty \left(\frac{r_2}{r_1} y \right)^{\left(\frac{r_1+m}{2} \right)(c+ci-1)+r+\frac{r_1}{2}-1} \left(1 + \frac{r_1 x}{r_2} \right)^{-\frac{r_1+r_2}{2}} \frac{r_2}{r_1} dy$$

$$\left(\frac{r_2}{r_1} \right)^{\left(\frac{r_1+m}{2} \right)(c+ci-1)+r+\frac{r_1}{2}-1} \int_0^\infty y^{\left(\frac{r_1+m}{2} \right)(c+ci-1)+r+\frac{r_1}{2}-1} (1+y)^{-\left(\frac{r_1+r_2}{2} \right)} dy$$

Recall from the Beta distribution of the second kind that

$$\int_0^\infty y^{\alpha-1} (1+y)^{-(\alpha+\beta)} dy = \beta(\alpha, \beta)$$

$$\alpha + \beta = \frac{r_1+r_2}{2}$$

$$\alpha = \left(\frac{r_1}{2} + m \right) (c+ci-1) + r + \frac{r_1}{2}$$

Then

$$\int_0^\infty y^{\left(\frac{r_1+m}{2} \right)(c+ci-1)+r+\frac{r_1}{2}-1} (1+y)^{-\left(\frac{r_1+r_2}{2} \right)} dy$$

We have

$$M_x(t) = bc \left(\frac{r_1}{r_2} \right)^{\frac{r_1}{2}} \frac{1}{\left[\beta \left(\frac{r_1}{2}, \frac{r_2}{2} \right) \right]^{c(1+i)}} \sum_{i=0}^{b-1} \sum_{r=0}^{\infty} \frac{(-1)^i \binom{b-1}{i} t^r}{r!}$$

$$\left[\sum_{j=0}^{\infty} \sum_{m=0}^j \frac{(-1)^{m+j} \left(\frac{r_1+r_2}{2} + j \right)^{\frac{r_1}{2}+m}}{2^{\frac{r_1+r_2}{2}+j}} \binom{r_1}{r_2} \frac{1}{\left(\frac{r_1}{2} + m \right)^{\frac{r_1+r_2}{2}}} \right]^{c(1+i)-1} \left(\frac{r_2}{r_1} \right)^{\left(\frac{r_1}{2} + m \right) (c(1+i)-1) + r + \frac{r_1}{2}}$$

$$\beta \left(\left(\frac{r_1}{2} + m \right) (c(1+i)-1) + \frac{r_1}{2} + r, \frac{r_2}{2} - \left(\frac{r_1}{2} + m \right) (c(1+i)-1) \right) \tag{27}$$

9. Moments

The rth moment of the Kumaraswamy-F is given as follows;

Using the that $E(e^{tx}) = \sum_{r=0}^{\infty} \frac{t^r}{r!} E(X^r)$

$$E(X^r) = bc \left(\frac{r_1}{r_2} \right)^{\frac{r_1}{2}} \frac{1}{\left[\beta \left(\frac{r_1}{2}, \frac{r_2}{2} \right) \right]^c} \left[\sum_{j=0}^{\infty} \sum_{m=0}^j \frac{(-1)^{m+j} \left(\frac{r_1+r_2}{2} + j \right)^{\frac{r_1}{2}+m}}{2^{\frac{r_1+r_2}{2}+j}} \binom{r_1}{r_2} \frac{1}{\left(\frac{r_1}{2} + m \right)^{\frac{r_1+r_2}{2}}} \right]^{c(a+i)-1}$$

$$\left(\frac{r_2}{r_1} \right)^{\left(\frac{r_1}{2} + m \right) (c(a+i)-1) + r + \frac{r_1}{2}} \beta \left(\left(\frac{r_1}{2} + m \right) (c-1) + \frac{r_1}{2} + r, \frac{r_2}{2} - \left(\frac{r_1}{2} + m \right) (c-1) - r \right) \tag{28}$$

10. Conclusions

The Kumaraswamy distribution was used to study the Kumaraswamy-Fisher Snedecor KUMA-F distribution and some of its properties which include the moments from which its mean, variance, skewness and kurtosis can be derived. The graph of the new F-distribution is positively skewed and appears to be fairly leptokurtic while the shape of the CDF of KUMA-F is similar to that of the conventional F- distribution. It is instructive to note that if the values of the parameters b and c were set to unity, the proposed KUMA-F distribution reverts to the conventional F-distribution. The proposed KUMA-F distribution is better in testing the equality of means across populations, when there is violation of assumptions due to the presence of outlier without transformation of data. Further work concerns application to real life data that would demonstrate the capabilities of the proposed distribution.

REFERENCES

[1] Barnard, G. A. (1984). Comparing the means of two independent samples. *Appl. Statist.* 33, 266-271.

[2] Brown, M.B Forsythe, A.B. (1974). The small sample behavior of some statistics which test the equality of several means. *Technometrics* 16, 129-132.

[3] Chen, S., Chen, J.H. (1998). Single-Stage analysis of variance under heteroscedasticity. *Communications in Statistics and Simulations* 27(3), 641-666.

[4] Gamage, J, Weerahandi, S. (1998). Size performance of some tests in one-way ANOVA. *Communications in Statistics and Simulations* 27(3), 625-640.

[5] Gupta, R. D. and Kundu, D. (1999). Generalized exponential distributions. *Australian and New Zealand Journal of Statistics* 41, 173-188.

[6] Gupta, R. D. and Kundu, D. (2001). Exponentiated exponential family: an alternative to gamma and Weibull. *Biometrical Journal* 43, 117-130.

[7] Jones, M.C. (2009). "Kumaraswamy distribution: A beta-type distribution with some tractability advantages". *Statistical Methodology* 6 (1): 70–81

[8] Kumaraswamy, P. (1980). "A generalized probability density functions for double-bounded random processes". *Journal of Hydrology* 46 (1-2): 79–88.

[9] K.W. Tsui, S. Weerahandi Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters *J. Amer. Statist. Assoc.*, 84 (1989), pp. 602–607

[10] Krutchkoff, R. G. (1988). One-way fixed effects analysis of

- variance when the error variances may be unequal. *J. Statist. Comput. Simul.* 30, 259-271.
- [11] Lomax Richard G (2007): "Statistical Concept: A second Course", P. 10 ISBN 0-8058-5850-4'
- [12] Rice, W. R. and Gains, S. D. (1989). One-way analysis of variance with unequal variances. *Proc. Natl. Acad. Sci.* 86, 8183-8184.
- [13] Scott, A.J, Smith, T.M.F. (1971). Interval estimates for linear combinations of means. *Applied Statistics* 20(3), 276-285.
- [14] Tsui, K. and Weerahandi, S. (1989). Generalized p-values in significance testing of hypotheses in the presence of nuisance parameters. *Journal of the American Statistical Association* 84, 602-607.
- [15] Weerahandi, S. (1993). ANOVA under unequal error variances. *Biometrika* 38, 330-336.
- [16] Weerahandi, S. (1995). *Exact statistical method for data analysis*. Springer-Verlag, New York, 2-50.
- [17] Welch, B.L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika* 38, 330-336.