University of São Paulo
"Luiz de Queiroz" College of Agriculture

The new class of Kummer beta generalized distributions: theory and applications

Rodrigo Rossetto Pescim

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor in Science. Area of concentration: Agricultural Statistics and Experimentation

Piracicaba
2014

Rodrigo Rossetto Pescim
Degree in Mathematics

The new class of Kummer beta generalized distributions: theory and applications

versão revisada de acordo com a resolução CoPGr 6018 de 2011

Adviser:
Prof. Dr. CLARICE GARCIA BORGES DEMÉTRIO

Thesis submitted in partial fulfillment of the requirements for the degree of Doctor in Science. Area of concentration: Agricultural Statistics and Experimentation

Piracicaba
2014

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA - DIBD/ESALQ/USP

Pescim, Rodrigo Rossetto
The new class of Kummer beta generalized distributions: theory and applications / Rodrigo Rossetto Pescim.- - versão revisada de acordo com a resolução CoPGr 6018 de 2011. - - Piracicaba, 2013.
130 p: il.

Tese (Doutorado) - - Escola Superior de Agricultura "Luiz de Queiroz", 2013.

1. Análise Bayesiana 2. Distribuição Birnbaum-Saunders 3. Distribuição gama 4. Distribuição normal 5. Matriz de informação observada 6. Razão de verossimilhança I. Título

CDD 519.532
P473n

DEDICATION

To my parents,
José Gilberto Pescim *and* **Luzia Aparecida Rossetto Pescim**, *For their love, patience and unfailing support to me*.

To my grandparents,
João Pescim (in memorian) *and* **Alice Moniz Pescim**, *For their love, tenderness and comprehension*.

To my friend,
Mariana Ragassi Urbano, *because you always give me something*.

To them,
I lovingly dedicate this work.

ACKNOWLEDGMENTS

I would like to express my gratitude to all those who gave me the possibility to complete this thesis, especially my parents José Gilberto Pescim and Luzia Aparecida Rossetto Pescim, my uncles Antonio Edson Pescim, José Gilmar Wolf, Luis Edilberto Pescim, my aunts Neuza Maria Favoreto Pescim and Marlene Aparecida Pescim Wolf, my cousins Guilherme Favoreto Pescim, Patrícia Penatti Pescim and Isabela Pescim Wolf, for their love and supporting me throughout my life.

To my adviser, Prof. Dr. Clarice Garcia Borges Demétrio, for the continuous support of my Doctorate, for her patience, motivation, enthusiasm and immense knowledge.

To Prof. Dr. Gauss Moutinho Cordeiro and Prof. Dr. Edwin Moisés Marcos Ortega, for their scientific contribution, intellectual input and especially for participating in the whole process of the Doctorate training.

To Prof. Dr. Saralees Nadarajah of the University of Manchester and his family, for all their support during my stay in Manchester.

To Prof. Dr. Taciana Villela Savian, Prof. Dr. Cristian Marcelo Villegas Lobos and Prof. Dr. Roseli Aparecida Leandro at ESALQ/USP, and Prof. John Hinde at NUI Galway, Ireland, for their valuable guidance and friendship.

To Sibelle Santanna da Silva, by the moments of happiness.

To my special friends, Mariana Ragassi Urbano, Ana Julia Righetto, Pedro Henrique Ramos Cerqueira and Thiago Gentil Ramires, who were always willing to help, and also, for providing me a more enjoyable life in the last four years.

To my friends from Manchester, Stuart Morrison, Fabiane Trindade, Jane Morakabi, Dan Sadullah and Daniel Ponicke, for their help, support and friendship.

To my friends from Piracicaba and from the Department of Exact Sciences at ESALQ/USP, Cássio Dessotti, Guilherme Biz, Lucas Cunha, Edilan Quaresma, Ezequiel Lopez, Maurício Lordello, Djair (Djavan), Everton da Rocha, Simone Grego, Iuri Ferreira, Thiago Oliveira, Luiz Ricardo Nakamura, Marina Maestre, Alessandra Santos, Ricardo Klein, Maria Cristina Martins, Rafael Moral, Simone Werner, José Nilton, Fernando Mayer, Tiago Santana and Everton de Toledo Hanser.

To employees of the Department of Exact Sciences at ESALQ/USP, the secretaries Solange de Assis Paes Sabadin, Mayara Segatto and Luciane Brajão, and the computer technicians Jorge Alexandre Wiendl and Eduardo Bonilha.

This work was supported by CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil.

SUMMARY

RESUMO

A nova classe de distribuições Kummer beta generalizada: teoria e aplicações

Neste trabalho, foi proposta uma nova classe de distribuições generalizadas, baseada na distribuição Kummer beta (NG; KOTZ, 1995), que contém como casos particulares os geradores exponencializado e beta de distribuições. A principal característica da nova família de distribuições é fornecer grande flexibilidade para as extremidades da função densidade e portanto, ela torna-se adequada para a análise de conjuntos de dados com alto grau de assimetria e curtose. Também foram estudadas duas novas distribuições que pertencem à nova família de distribuições, baseadas nas distribuições Birnbaum-Saunders e gama generalizada, que possuem função de taxas de falhas que assumem diferentes formas (unimodal, forma de banheira, crescente e decrescente). Em todas as pesquisas, propriedades matemáticas gerais como momentos ordinários e incompletos, função geradora, desvios médio, confiabilidade, entropias, estatísticas de ordem e seus momentos foram discutidas. A estimação dos parâmetros é abordada pelo método da máxima verossimilhança e pela análise bayesiana e a matriz de informação observada foi derivada. Considerou-se, também, a estatística de razão de verossimilhanças e testes formais de qualidade de ajuste para comparar todas as distribuições propostas com alguns de seus submodelos e modelos não encaixados. Os resultados desenvolvidos foram aplicados a seis conjuntos de dados.

Palavras-chave: Análise bayesiana; Distribuição Birnbaum-Saunders; Distribuição gama; Distribuição normal; Matriz de informação observada; Razão de verossimilhanças

ABSTRACT

The new class of Kummer beta generalized distributions: theory and applications

In this study, a new class of generalized distributions was developed, based on the Kummer beta distribution (NG; KOTZ, 1995), which contains as particular cases the exponentiated and beta generators of distributions. The main feature of the new family of distributions is to provide greater flexibility to the extremes of the density function and therefore, it becomes suitable for analyzing data sets with high degree of asymmetry and kurtosis. Also, two new distributions belonging to the new class of distributions, based on the Birnbaum-Saunders and generalized gamma distributions, that has as main characteristic the hazard function which assumes different forms (unimodal, bathtub shape, increase, decrease) were studied. In all studies, general mathematical properties such as ordinary and incomplete moments, generating function, mean deviations, reliability, entropies, order statistics and their moments were discussed. The estimation of parameters is approached by the method of maximum likelihood and Bayesian analysis and the observed information matrix is derived. It is also considered the likelihood ratio statistics and formal goodness-of-fit tests to compare all the proposed distributions with some of its sub-models and non-nested models. The developed results for all studies were applied to six real data sets.

Keywords: Bayesian analysis; Birnbaum-Saunders distribution; Gamma distribution; Kummer beta distribution; Likelihood ratio; Normal distribution; Observed information matrix

1 INTRODUCTION

The continuous univariate distributions are fundamental to statistical science and are a powerful indispensable tool for applied statisticians. These distributions have been extensively used over the past decades for fitting data sets in several fields of research such as medical and environmental sciences, biological studies, demography, engineering, actuarial sciences, economics, finance and insurance. However, in many applied areas such as lifetime analysis, reliability, finance and insurance, there is a clear need for extended forms of these univariate distributions. Consequently, a significant progress has been performed for the generalization of some well-known distributions and their applications to a variety of problems in many areas of research.

In this sense, generalized distributions have been widely studied in the last decades. (AMOROSO, 1925) was the precursor of extending continuous distributions, discussing a generalization of the gamma distribution to fit observed distribution of income rate. Since then, numerous authors have developed generalized distributions including (GOOD, 1953) and (WISE, 1975) that extended the inverse normal distribution, (LJUBO, 1965) and (HOSKING; WALLIS, 1987) who generalized the Pareto distribution. Recent developments focus on new techniques for building new meaningful classes of continuous distributions, including the exponentiated generator (EG) approach introduced by (MULDHOLKAR; SRIVASTAVA; FRIEMER, 1995) and the beta generator (BG) approach pioneered by (EUGENE; LEE; FAMOYE, 2002), have been proposed to provide more flexibility and applicability for the new distributions. (LAI, 2013) provided a good review about constructions and applications of the generalized lifetime distributions. Now, we shall give more attention to the generalized distributions based on the exponentiated and beta generators.

Hereafter, we define the exponentiated-G ("EG" for short) distribution for an arbitrary continuous baseline distribution function $G(x)$, say $X \sim$ EG(a), $a > 0$, if X has cumulative distribution function (cdf) given by

$$F_{\mathcal{EG}}(x) = G(x)^a, \tag{1.1}$$

where $a > 0$ is an additional shape parameter. Note that the cdf of EG distribution depends on the shape parameter a and the parameter vector $\boldsymbol{\gamma}$ of the baseline G distribution.

The probability density function (pdf) corresponding to (1.1) can be expressed as

$$f_{\mathcal{EG}}(x) = a\, g(x)\, G(x)^{a-1}. \tag{1.2}$$

The EG distribution (1.2) is also known as alternative Lehmann type I distribution. According to (CORDEIRO; ORTEGA; CUNHA, 2013), "for $a > 1$ and $a < 1$ and for larger values of x, the multiplicative factor $a\, G(x)^{a-1}$ is greater and smaller than one, respectively. The reverse assertion is also true for smaller values of x. The latter immediately implies that the ordinary

moments associated with the density function $f_{\mathcal{EG}}(x)$ are strictly larger (smaller) than those associated with the density $g(x)$ when $a > 1$ ($a < 1$)."

The general properties of EG distributions have been studied by many authors in recent years, see (MULDHOLKAR; SRIVASTAVA; FRIEMER, 1995, 1996) for exponentiated Weibull (EW), (GUPTA; GUPTA; GUPTA, 1998) for exponentiated Pareto (EPa), (GUPTA; KUNDU, 2001) for exponentiated exponential (EE), (NADARAJAH, 2005) for exponentiated Gumbel (EGu), (SHIRKE; KAKADE, 2006) for exponentiated log-normal (ELN), (NADARAJAH; GUPTA, 2007) for exponentiated gamma (EGa) and (CORDEIRO; ORTEGA; SILVA, 2011) for exponentiated generalized gamma (EGG) distributions, among others.

Now, we shall consider the beta generator introduced by (EUGENE; LEE; FAMOYE, 2002) as follows. For any parent distribution and density, $G(.)$ and $g(.)$, respectively, and let $X = G^{-1}(U)$ with $U \sim \text{beta}(a, b)$, the standard beta distribution, the continuous random variable X is said to have a beta generalized (BG) distribution. This, can be characterized by its cdf

$$F_{\mathcal{BG}}(x) = \frac{1}{B(a,b)} \int_0^{G(x)} \omega^{a-1} (1-\omega)^{b-1} \, d\omega, \tag{1.3}$$

where $a > 0$ and $b > 0$ are two extra shape parameters that aim to introduce skewness and to provide greater flexibility of its tails, $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ is the beta function, $\Gamma(a) = \int_0^\infty x^{a-1} \, \mathrm{e}^{-\mathrm{x}} \, \mathrm{dx}$ is the gamma function, $I_y(a,b) = B_y(a,b)/B(a,b)$ is the incomplete beta function ratio and $B_y(a,b) = \int_0^y t^{a-1} (1-t)^{b-1} dt$ is the incomplete beta function. One major benefit of this class of distributions is its ability of fitting skewed data that can not be properly fitted by existing distributions. It has been receiving increased attention over the last decade, in particular after the works of (EUGENE; LEE; FAMOYE, 2002) and (JONES, 2004).

We can also express (1.3) in terms of the hypergeometric function (GRADSHTEYN; RYZHIK, 2007), since the properties of the hypergeometric function are well established in the literature. We can obtain

$$F_{\mathcal{BG}}(x) = \frac{G(x)}{a\, B(a,b)} \, {}_2F_1(a, 1-b, a+1; G(x)). \tag{1.4}$$

The pdf associated to (1.3) takes the form

$$f_{\mathcal{BG}}(x) = \frac{g(x)}{B(a,b)} \, G(x)^{a-1} \, [1 - G(x)]^{b-1}. \tag{1.5}$$

The pdf $f_{\mathcal{BG}}(x)$ will be most tractable when both functions $G(x)$ and $g(x)$ have simple analytic expressions. Except for some special choices of these functions, the density $f_{\mathcal{BG}}(x)$ will be difficult to cope with some generality.

The first distribution of BG class was the beta normal (BN) distribution, introduced by (EUGENE; LEE; FAMOYE, 2002). Since then, many other specific beta-G distributions have been proposed by (NADARAJAH; KOTZ, 2004, 2005), (AKINSETE; FAMOYE; LEE, 2008), (PARANAÍBA et al., 2011) and (CORDEIRO et al., 2013). Some practical applications have been considered, for example, (AKINSETE; FAMOYE; LEE, 2008) fitted the beta Pareto (BPa)

distribution to flood data; (RAZZAGHI, 2009) applied the BN distribution to continuous dose-response modelling; (PESCIM et al., 2010) applied the beta generalized half-normal (BGHN) distribution to myelogenous leukemia data and (CORDEIRO; LEMONTE, 2011) fitted the beta Birnbaum-Saunders (BBS) distribution to fatigue data.

However, the classical beta generator can add a limited structure, depending on the baseline distribution and consequently, those BG distributions do not offer greater flexibility to the extremes of their density functions. Moreover, (ALEXANDER et al., 2012) demonstrated that generators of new distributions with one more shape parameter than the beta generator are necessary to provide additional control over both skewness and kurtosis. From this, the Kummer beta (KB) distribution with three parameters, introduced and studied by (NG; KOTZ, 1995), generalizes the classical beta distribution and provides greater flexibility to extremes (left and right) of the density function giving to it a range of applicability.

In this present work, we propose a new class of generalized distributions based on the KB distribution which is an extension of the exponentiated and beta generators, in order to extend well-known distributions such as normal, Weibull, gamma, Gumbel, Birnbaum-Saunders and generalized gamma for applications in lifetime analysis, reliability, actuarial and environmental sciences. Thus, the thesis is organized as follows. In Chapter 2, we define a new family of distributions so-called the Kummer beta generalized (KB-G) class of distributions. In Section 2.1 is presented the motivation to construct the new family of distributions. Section 2.2 provides some special cases. In Section 2.3, we derive general expansions for the new cdf and pdf in terms of exponentiated and beta generators of distributions. We can apply these expansions to several KB-G distributions. In Section 2.4, we obtain the general properties of the KB-G family of distribution such as moments, generating functions, mean deviations and Rényi entropy. In Section 2.5, we provide some expansions for the pdf of the order statistics. The method of maximum likelihood and a Bayesian procedure are adopted for estimating the model parameters in Section 2.6. In Section 2.7, we analyze two real data sets using special KB-G distributions. Section 2.8 ends with some concluding remarks.

In Chapter 3, we introduce the new fatigue life distribution so-called Kummer beta Birnbaum-Saunders (KBBS) distribution for reliability studies. In Section 3.1, we give a review of the problem related to fatigue process. In Section 3.2, we define the KBBS distribution and plot its density and hazard rate functions. Section 3.3 provides useful expansions for the density and cumulative functions. We obtain explicit expressions for the moments (Section 3.4), generating functions (Section 3.5), incomplete moments (Section 3.6), mean deviations and reliability (Section 3.7) and order statistics (Section 3.8). Some inferential tools are discussed in Section 3.9. An application presented in Section 3.10 reveal the usefulness of the new distribution for fatigue life data. Concluding remarks are addressed in Section 3.11.

In Chapter 4, we propose the Kummer beta generalized gamma (KBGG) distribution. In Section 4.1, we discuss a review of the generalized gamma (GG) distribution. In Section 4.2, we derive more than 32 special distributions from KBGG model. In Section 4.3, we show that

the KBGG distribution can be expressed as a linear combination of EGG density functions. This is an important result to provide some mathematical properties of the EGG distribution. We obtain explicit expressions for the moments and generating function (Section 4.4), mean deviations and Rényi entropy (Section 4.5) and distribution of order statistics (Section 4.6). In Section 4.7, we discuss maximum likelihood estimation and statistical inference. In Section 4.8, three applications are presented to reveal the usefulness of the new distribution for real data sets. Concluding remarks are addressed in Section 4.9.

References

AKINSETE, A.; FAMOYE, F.; LEE, C. The beta-Pareto distribution. **Statistics**, Berlin, v.42, p.547-563, 2008.

ALEXANDER, C.; CORDEIRO, G. M.; ORTEGA, E. M. M.; SARABIA, J. M. Generalized beta-generated distributions. **Computational Statistics and Data Analysis**, New York, v.56, p.1880-1897, 2012.

AMOROSO, L. Ricerche intono alla curva dei redditi. **Annali Mathemática**, Firenze, v.2, p.123-159, 1925.

CORDEIRO, G. M.; CASTELLARES, F.; MONTENEGRO, L. C.; de CASTRO, M. The beta generalized gamma distribution. **Statistics**, Berlin, v.47, p.888-900, 2013.

CORDEIRO, G. M.; ORTEGA, E. M. M.; CUNHA, D. C. C. The exponentiated generalized class of distributions. **Journal of Data Science**, New York, v.11, p.1-27, 2013.

CORDEIRO, G. M.; ORTEGA, E. M. M.; SILVA, G. O. The exponentiated generalized gamma distribution with application to lifetime data. **Journal of Statistical Computation and Simulation**, London, v.81, p.827-842, 2011.

EUGENE, N.; LEE, C.; FAMOYE, F. Beta-normal distribution and its applications. **Communication in Statistics: Theory and Methods**, New York, v.31, p.497-512, 2002.

GOOD, I. J. The population frequencies of the species and the estimation of population parameteres. **Biometrika**, London, v.40, p.237-260, 1953.

GRADSHTEYN I. S.; RYZHIK I. M. **Table of Integrals, Series, and Products**. San Diego: Academic Press, 2007. 315p.

GUPTA, R. D.; KUNDU, D. Exponentiated exponential family: An alternative to gamma and Weibull distributions. **Biometrical Journal**, Berlin, v.43, p.117-130, 2001.

GUPTA, R. C.; GUPTA, P. L.; GUPTA, R. D. Modeling failure time data by Lehman alternatives. **Communications in Statistics—Theory and Methods**, New York, v.27, p.887-904, 1998.

HOSKING, J. R. M.; WALLIS, J. R. Parameter and quantile estimation for the generalized Pareto distribution.**Technometrics**, Alexandria, v.29, p.339-349, 1987.

JONES, M. C. Families of distributions arising from distributions of order statistics. **Test**, Madrid, v.13, p.1-43, 2004.

LAI, C. D. Constructions and applications of lifetime distributions. **Applied Stochastic Models in Business and Industry**, London, v.29, p.127-140, 2013.

LJUBO, M. Curves and concentration indices for certain generalized Pareto distribution. **International Statistical Review**, Amsterdam, v.15, p.257-260, 1965.

MUDHOLKAR, G. S; SRIVASTAVA, D. K.; FRIEMER, M. The exponential Weibull family: A reanalysis of the bus-motor failure data. **Technometrics**, Alexandria, v.37, p.436-445, 1995.

MUDHOLKAR, G. S.; SRIVASTAVA, D. K.; FRIEMER, M. A generalization of the Weibull distribution with application to the analysis of survival data. **Journal of the American Statistical Association**, Boston, v.91, p.1575-1583, 1996.

NADARAJAH, S. The exponentiated Gumbel distribution with climate application. **Environmetrics**, New York, v.17, p.13-23, 2005.

NADARAJAH, S.; GUPTA, A. K. A generalized gamma distribution with application to drought data. **Mathematics and Computers in Simulation**, Amsterdam, v.74, p.1-7, 2007.

NADARAJAH, S.; KOTZ, S. The beta Gumbel distribution. **Mathematical Problems in Engineering**, New York, v.10, p.323-332, 2004.

NADARAJAH, S.; KOTZ, S. The beta exponential distribution. **Reliability Engineering and System Safety**, Barking, v.91, p.689-697, 2005.

NG, K. W.; KOTZ, S. Kummer-gamma and Kummer-beta univariate and multivariate distributions. **Research Report**, Hong Kong, v.84, p.120-141, 1995.

PARANAÍBA, P. F.; ORTEGA, E. M. M.; CORDEIRO, G. M.; PESCIM, R. R. The beta Burr XII distribution with application to lifetime data. **Computational Statistics and Data Analysis**, New York, v.55, p.1118-1136, 2011.

PESCIM, R. R.; DEMÉTRIO, C. G. B.; CORDEIRO, G. M.; ORTEGA, E. M. M.; URBANO, M. R. The beta generalized half-normal distribution. **Computational Statistics and Data Analysis**, New York, v.54, p.945-957, 2010.

RAZZAGHI, M. Beta-normal distribution in dose-response modeling and risk assessment for quantitative responses. **Environmental and Ecological Science**, New York, v.16, p.21-36, 2009.

SHIRKE, D. T.; KAKADE, C. S. On exponentiated lognormal distribution. **International Journal of Agricultural and Statistical Sciences**, New Delhi, v.2, p.319-326, 2006.

WISE, M. E. Skew distributions in biomedicine, incluing some negative powers of time. **Statistical Distributions in Scientific Work**, New York, v.2, p.241-262, 1975.

2 THE KUMMER BETA GENERALIZED FAMILY OF DISTRIBUTIONS

Abstract

(NG; KOTZ, 1995) introduced a probability distribution that provides greater flexibility to extremes. We define and study a new class of distributions so-called the Kummer beta generalized family to extend the normal, Weibull, gamma, Gumbel, Pareto and logistic distributions, among several other well-known distributions. Some special models of this new class of distributions are discussed. The ordinary moments of any distribution in the new family can be written as linear functions of probability weighted moments of the baseline distribution. We also obtain the density function of the order statistics, mean deviations and entropies. We adopt the method of maximum likelihood and Bayesian approach to fit the distributions in the new class and illustrate its potentiality with applications for two real data sets.

Keywords: Gamma distribution; Kummer beta distribution; Likelihood ratio test; Normal distribution; Order statistic

2.1 Introduction

According to (NADARAJAH; KOTZ, 2007), "the beta family distribution, whose origin can be traced to 1676 in a letter from Sir Issac Newton to Henry Oldenbeg, has been used extensively in theoretical and applied statistics for over a century. Originally defined on the unit interval $(0, 1)$ but extended to any finite interval, the beta distribution can take an amazingly great variety of forms. It can be fitted practically to any data representing a phenomenon in almost any field of application."

In fact, the beta distribution is one of the most important models to account for the random phenomena which produce results in the range $(0, 1)$ due to flexibility of its parameters. It is very versatile and can be used to analyze different types of data sets. Many of the finite range distributions encountered in practice can be transformed into the standard beta distribution. In econometrics, for example, quite often the data are analyzed using finite-range distributions. In the statistics literature, there are a plenty of applications for the beta distribution. (BURY, 1999) discussed a number of applications in engineering using the beta model. (BALDING; NICHOLS, 1995) applied the beta distribution in population genetics for a statistical description of the frequencies of alleles. (WILEY; HERSCHOKORU; PADIAU, 1989) developed a statistical model based on the beta distribution to obtain the probability of HIV transmission during sexual contact between an individual infected and a healthy individual.

Generalized beta distributions have been widely studied in statistics and numerous authors have developed various classes of these distributions. (EUGENE; LEE; FAMOYE, 2002) proposed a general class of distributions for a random variable defined from the logit of the beta random variable by employing two parameters whose role is to introduce skewness and to vary tail weight.

Following (EUGENE; LEE; FAMOYE, 2002) who defined the beta normal (BN) distribution, (NADARAJAH; KOTZ, 2004) introduced the beta Gumbel distribution (BGu), provided expressions for the moments, examined the asymptotic distribution of the extreme order statistics and performed maximum likelihood estimation. (NADARAJAH; GUPTA, 2004) defined the beta Fréchet (BF) distribution and derived the analytical shapes of the density and hazard rate functions. Further, (NADARAJAH; KOTZ, 2005) proposed the beta exponential (BE) distribution and obtained the moment generating function (mgf), the first four moments, the asymptotic distribution of the extreme order statistics and discussed maximum likelihood estimation. More recently, (PESCIM et al., 2010), (PARANAÍBA et al., 2011) and (CORDEIRO; LEMONTE, 2011) studied important mathematical properties of the beta generalized half-normal (BGHN), beta Burr XII (BBXII) and beta Birnbaum-Saunders (BBS) distributions, respectively. However, we can note that those distributions do not offer more flexibility to extremes (right and left) of the curves of the density functions and therefore they are not suitable for analyzing data sets with high degree of asymmetry and kurtosis.

(NG; KOTZ, 1995) proposed the Kummer beta (KB) distribution on the unit interval $(0,1)$ with cumulative distribution function (cdf) and probability density function (pdf) given by

$$F_{\mathcal{KB}}(x) = K \int_0^x t^{a-1}\,(1-t)^{b-1}\,\exp(-c\,t)dt,$$

and

$$f_{\mathcal{KB}}(x) = K\,x^{a-1}\,(1-x)^{b-1}\,\exp(-c\,x), \quad 0 < x < 1,$$

where $a > 0$, $b > 0$ and $-\infty < c < \infty$. Here,

$$K^{-1} = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}\,{}_1F_1(a; a+b; -c) \tag{2.1}$$

and

$${}_1F_1(a; a+b; -c) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 t^{a-1}\,(1-t)^{b-1}\exp(-c\,t)dt = \sum_{k=0}^{\infty} \frac{(a)_k(-c)^k}{(a+b)_k\,k!}$$

is the confluent hypergeometric function (ABRAMOWITZ; STEGUN, 1968), $\Gamma(\cdot)$ is the gamma function and $(d)_k = d(d+1)\ldots(d+k-1)$ denotes the ascending factorial. According to (NAGAR; GUPTA, 2002), "(GORDY, 1998) has also defined the Kummer beta distribution in relation to the problem of common value auction. This distribution is an extension of the beta distribution, and for $a < 1$ (and certain values of the parameter c) yields bimodal distributions on finite range." Plots of the KB density function are displayed in Figure 2.1 for selected parameter values.

Consider starting from a parent continuous cdf $G(x)$. A natural way of generating families of distributions on some other support from a simple starting baseline distribution with pdf $g(x) = dG(x)/dx$ is to apply the quantile function to a family of distributions on the interval

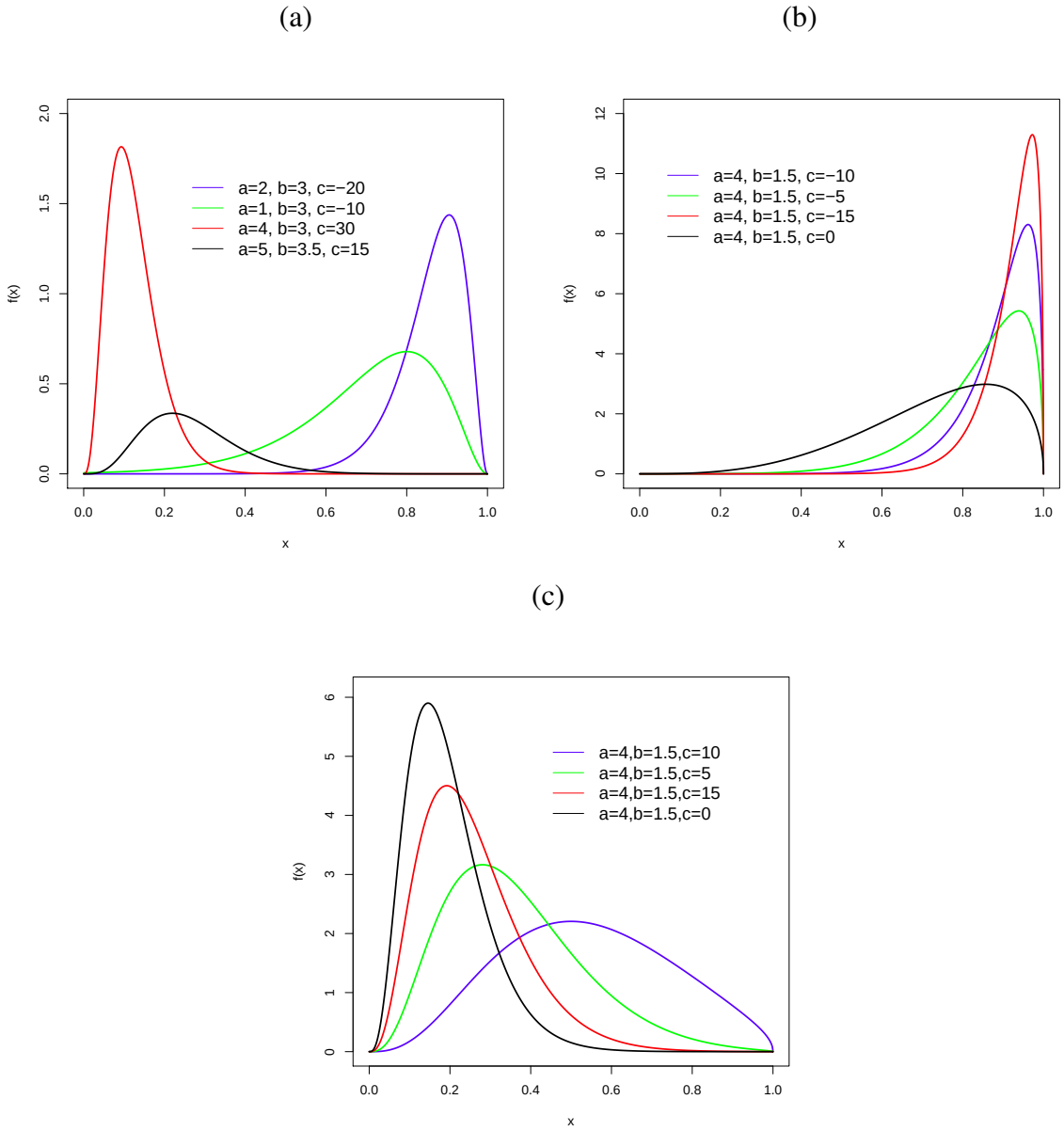


Figure 2.1 – Plots of the Kummer beta pdf for some parameter values

$(0,1)$. In other words, let $X = G^{-1}(U)$ with $U \sim \text{KB}(a,b,c)$, the Kummer beta distribution. Then, the random variable X is said to have a Kummer beta generalized (KB-G) distribution.

From an arbitrary baseline cdf $G(x)$, the KB-G family of cumulative distributions is defined by

$$F_{\mathcal{KBG}}(x) = K \int_0^{G(x)} t^{a-1}\,(1-t)^{b-1}\,\exp(-c\,t)dt, \tag{2.2}$$

where $a > 0$ and $b > 0$ are shape parameters which introduce skewness, and thereby promote weight variation of the tails, whereas the parameter $-\infty < c < \infty$ "squeezes" the pdf to the left or right, i.e., it leads the tail weights of the pdf to extremes of the curves of density functions.

The pdf corresponding to (2.2) can be expressed as

$$f_{\mathcal{KBG}}(x) = K\, g(x)\, G(x)^{a-1}\, [1 - G(x)]^{b-1}\, \exp\left[-c\, G(x)\right], \tag{2.3}$$

where K is defined in (2.1).

The KB-G family of distributions defined by (2.3) is an alternative family of models to the class of distributions proposed by (ALEXANDER et al., 2012). The shape parameter $c > 0$, in (ALEXANDER et al., 2012), together with $a > 0$ and $b > 0$ promote the weight variation of the tails and more flexibility. On the other hand, the parameter $-\infty < c < \infty$ of the proposed family offers more flexibility to the extremes (left and/or right) for the density function curves and therefore the new family of distributions becomes more suitable for analyzing data sets with high degree of asymmetry. For each continuous G distribution (here and henceforth "G" denotes the baseline distribution), we can associate the KB-G distribution with three extra parameters a, b and c defined by the pdf (2.3).

Special generalized distributions can be generated as follow. The KB-normal (KBN) distribution is obtained by taking $G(x)$ in equation (2.2) to be the normal cdf. Analogously, the KB-Weibull (KBW), KB-gamma (KBGa) and KB-Gumbel (KBGu) distributions are obtained by taking $G(x)$ to be the cdf of the Weibull, gamma and Gumbel distributions, respectively. Hence, each new KB-G distribution can be obtained from a specified G distribution. The KB distribution is a clearly example of the KB-G distribution when G is the uniform distribution on $(0, 1)$, whereas the G distribution corresponds to $a = b = 1$ and $c = 0$.

The class of distributions (2.3) includes two important special cases: the beta-generalized (BG) and exponentiated generalized (EG) classes of distributions defined by (EUGENE; LEE; FAMOYE, 2002) and (MUDHOLKAR; SRIVASTAVA; FRIEMER, 1995) when $c = 0$ and for $c = 0$ and $b = 1$, respectively. We can note that the BG distributions can be limited in one aspect. They have only two additional shape parameters and so they can add only a limited structure to the generated distribution. For instance, a BG distribution may have problems to capture the behaviour of random variables with symmetric but highly leptokurtic distributions. While the beta parameters offer explicit control over skewness when the baseline distribution is symmetric, they have less control over higher moments such as kurtosis. Further, the EG distribution still introduces only one extra shape parameter, whereas three parameters may be required to control both tail weights and the distribution of weight in the center. Hence, the generated distribution (2.3) is a more flexible model since it has one more shape parameter than the classical beta or exponentiated generators.

We study some mathematical properties of the KB-G family of distributions because it extends several widely-known distributions in the literature. This chapter is outlined as follows. Section 2.2 provides some special cases. In Section 2.3, we derive general expansions for the new cdf and pdf in terms of exponentiated and beta generators of distributions. We can apply these expansions to several KB-G distributions. In Section 2.4, we obtain the general properties of the KB-G family of distribution such as moments, generating function, mean deviations and

Rényi entropy. In Section 2.5, we provide some expansions for the pdf of the order statistics. The method of maximum likelihood and a Bayesian procedure are adopted for estimating the model parameters in Section 2.6. In Section 2.7, we analyze two real data sets using special KB-G distributions. Section 2.8 ends with some concluding remarks.

2.2 Special KB Generalized Distributions

The KB-G density function (2.3) allows for greater flexibility of its tails and promotes the variation of the tail weights to the extremes of the distribution. It can be widely applied in many areas of engineering and biological sciences. The pdf (2.3) will be most tractable when the cdf $G(x)$ and the pdf $g(x)$ have simple analytic expressions. We have considered six different baselines: normal, Weibull, gamma, Gumbel, Pareto and logistic distributions. In each case, the baseline cdf and pdf of the corresponding KB-G model are summarized in Table 2.1. The KB-G distributions can be applied to the same areas as their corresponding baseline distributions, to offer an improved fit to the data sets.

For brevity, in the remainder of this section, we shall only comment in detail four of the most important KB-G distributions: the Kummer beta normal (KBN), the Kummer beta Weibull (KBW), the Kummer beta gamma (KBGa) and the Kummer beta Gumbel (KBGu) distributions.

2.2.1 KB-normal

The KB-normal (KBN) pdf is obtained from (2.3) by taking $G(\cdot)$ and $g(\cdot)$ to be the cdf and pdf of the normal distribution, N(μ, σ^2), so that

$$f(x) = \frac{K}{\sigma}\phi\left(\frac{x-\mu}{\sigma}\right)\left[\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^{a-1}\left[1-\Phi\left(\frac{x-\mu}{\sigma}\right)\right]^{b-1}\exp\left[-c\,\Phi\left(\frac{x-\mu}{\sigma}\right)\right],$$

where $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ is a location parameter, $\sigma > 0$ is a scale parameter, a and b are positive shape parameters, $c \in \mathbb{R}$, and $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively. A random variable with the above pdf is denoted by X $\sim$ KBGN(a, b, c, μ, σ^2). For $\mu = 0$ and $\sigma = 1$, we have the standard KBN distribution. Following the same methodology proposed by (NADARAJAH, 2008), the nth moment of the KBN distribution can be expressed as a finite sum of the Lauricella functions of type A (EXTON, 1978), when a, b and c are integer numbers.

2.2.2 KB-Weibull

The cdf of the Weibull distribution with parameters $\lambda > 0$ and $\gamma > 0$ is $G(x) = 1-\exp[-(\lambda x)^\gamma]$ for $x > 0$. Correspondingly, the KB-Weibull (KGW) density, say KBW$(a, b, c, \gamma, \lambda)$, reduces to

$$f(x) = K\,\gamma\,\lambda^\gamma\,x^{\gamma-1}\,\{1-\exp[-(\lambda x)^\gamma]\}^{a-1}\,\exp\{-c\,[1-\exp[-(\lambda x)^\gamma]] - b(\lambda x)^\gamma\},$$

where x, a, b, λ, γ are real-positive values and $c \in \mathbb{R}$. For $\gamma = 1$, we obtain the KB-exponential (KBE) distribution. The KBW$(1, 1, 0, 1, \lambda)$ distribution corresponds to the exponential distribution with parameter λ.

2.2.3 KB-gamma

Let Y be a random variable which follows a gamma distribution with cdf $G(y) = \gamma_1(\alpha, \beta\, y) = \gamma(\alpha, \beta\, y)/\Gamma(\alpha)$ for y, α, $\beta > 0$, where $\gamma(a, y) = \int_0^y t^{a-1}\mathrm{e}^{-\mathrm{t}}\mathrm{dt}$ is the incomplete gamma function. The pdf of a random variable X having the KBGa distribution, say X $\sim$ KBGa(a, b, c, α, β), can be expressed as

$$f(x) = \frac{K\,\beta^{\alpha}}{\Gamma(\alpha)}\, x^{\alpha-1}\, \gamma_1(\alpha, \beta x)^{a-1}\, [1 - \gamma_1(\alpha, \beta x)]^{b-1} \exp\left[-c\,\gamma_1(\alpha, \beta x) - \beta x\right].$$

For $\alpha = 1$ and $c = 0$, we obtain the KBE distribution. The KBGa$(1, 1, 0, 1, \beta)$ distribution reduces to the exponential distribution with parameter β.

2.2.4 KB-Gumbel

The pdf and cdf of the Gumbel distribution with location parameter $\mu \in \mathbb{R}$ and scale parameter $\sigma > 0$ are given by

$$g(x) = \sigma^{-1} \exp\left[\frac{x-\mu}{\sigma} - \exp\left(\frac{x-\mu}{\sigma}\right)\right],\ x > 0,$$

and

$$G(x) = 1 - \exp\left[-\exp\left(-\frac{x-\mu}{\sigma}\right)\right],$$

respectively. The mean and variance are equal to $\mu - \gamma\sigma$ and $\pi^2\sigma^2/6$, respectively, where $\gamma \approx 0.57722$ is the Euler's constant. By inserting these equations in (2.3), we obtain the KBGu distribution, say KBGu(a, b, c, μ, σ).

Figure 2.2 represents some of the possible shapes of four KB-G density functions. These plots show the great flexibility achieved with the new distributions.

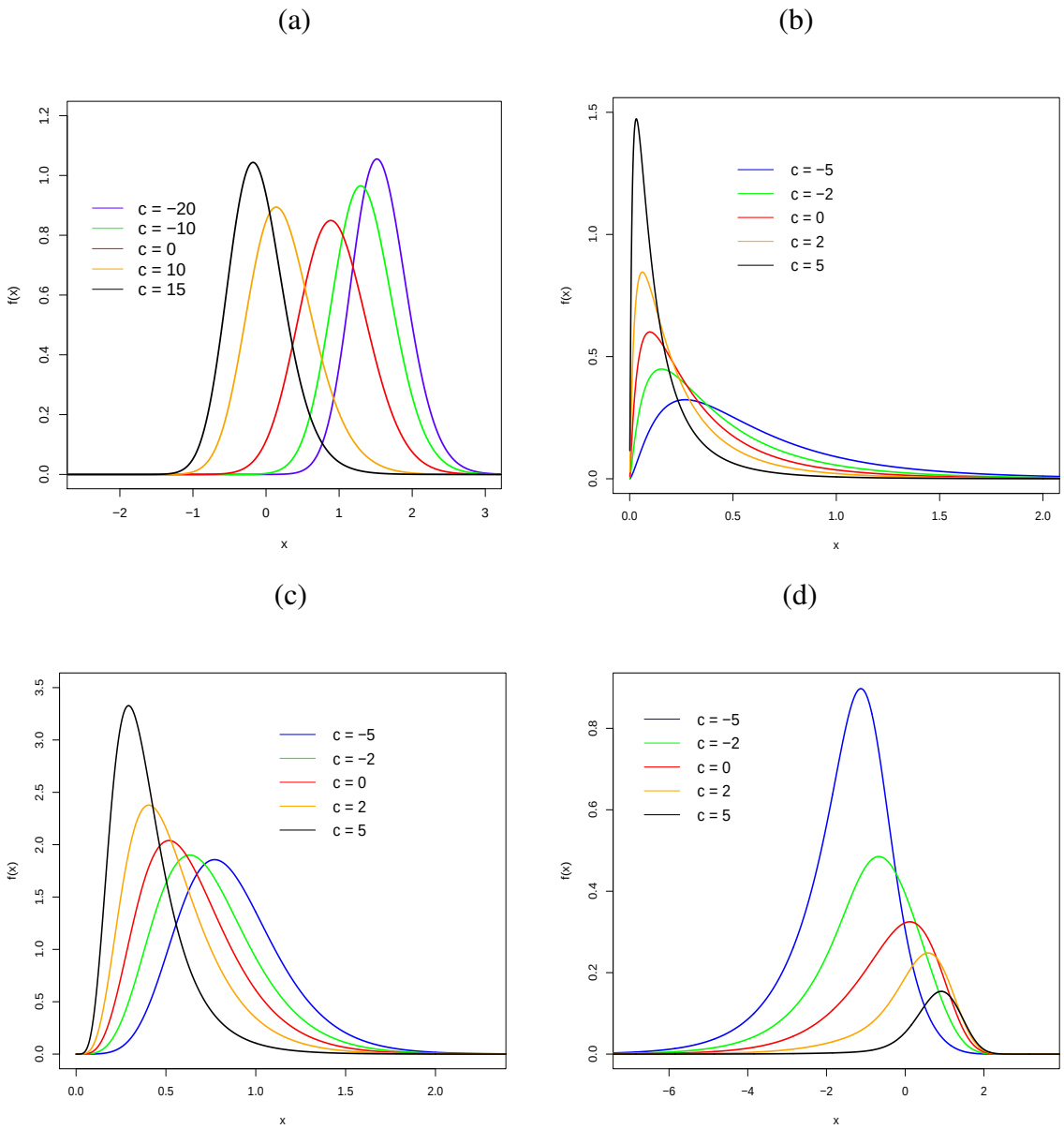


Figure 2.2 – (a) KBN(8,2,c,0,1), (b) KBW(5,3,c,0.5,4), (c) KBGa(3,1.5,c,4,2) and (d) KBGu(0.8,1,c,0,1) pdfs (the red lines represent the BG pdfs)

Table 2.1 – Special KB-G distributions. In the normal distribution, $\Phi(x)$ and $\phi(x)$ denote the standard normal cumulative and density functions. In the Gumbel distribution, $u = \exp\left[-\left(\frac{x-\mu}{\sigma}\right)\right]$

Distribution	Baseline Cumulative Distribution	Density of the Kummer Beta Generalized Distribution
Normal	$G(x) = \Phi(x)$	$f(x) = K\phi(x)\Phi(x)^{a-1}\left[1-\Phi(x)\right]^{b-1} \mathrm{e}^{-\mathrm{c}\,\Phi(\mathrm{x})}$
Weibull	$G(x) = 1 - \mathrm{e}^{-(\lambda \mathrm{x})^{\gamma}}$	$f(x) = K\gamma\lambda^{\gamma}x^{\gamma-1}\left[1-\mathrm{e}^{-(\lambda \mathrm{x})^{\gamma}}\right]^{a-1} \mathrm{e}^{-\mathrm{c}\left[1-\mathrm{e}^{-(\lambda \mathrm{x})^{\gamma}}\right]-\mathrm{b}(\lambda \mathrm{x})^{\gamma}}$
Gamma	$G(x) = \gamma_1(\alpha, \beta x) = \frac{\gamma(\alpha,\beta x)}{\Gamma(\alpha)}$, $\alpha, \beta > 0$	$f(x) = \frac{K\,\beta^{\alpha}\,x^{\alpha-1}}{\Gamma(\alpha)}\gamma_1(\alpha,\beta x)^{a-1}\left[1-\gamma_1(\alpha,\beta x)\right]^{b-1} \mathrm{e}^{-\mathrm{c}\,\gamma_1(\alpha,\beta \mathrm{x})-\beta \mathrm{x}}$
Gumbel	$G(x) = 1 - \mathrm{e}^{-\mathrm{u}}$, $\sigma > 0$, $\mu \in \mathbb{R}$	$f(x) = \frac{K\mathrm{e}^{-\mathrm{c}}}{\sigma u}\left[1-\mathrm{e}^{-\mathrm{u}}\right]^{a-1} \mathrm{e}^{-1/\mathrm{u}-\mathrm{u}(\mathrm{b}-1)+\mathrm{e}^{-\mathrm{u}}}$
Pareto	$G(x) = 1 - \frac{1}{(1+x)^{\nu}}$, $\nu > 0$	$f(x) = \frac{K\nu[(1+x)^{\nu}-1]^{a-1}}{\mathrm{e}^{\mathrm{c}}(1+\mathrm{x})^{\nu(\mathrm{a}+\mathrm{b}-1)+1}} \mathrm{e}^{\mathrm{c}(1+\mathrm{x})^{-\nu}}$
Logistic	$G(x) = \frac{1}{1+\mathrm{e}^{-\mathrm{x}}}$	$f(x) = \frac{K\mathrm{e}^{-\mathrm{bx}}}{(1+\mathrm{e}^{-\mathrm{x}})^{\mathrm{a}+\mathrm{b}}} \mathrm{e}^{-\mathrm{c}(1+\mathrm{e}^{-\mathrm{x}})^{-1}}$

2.3 Expansions for the Density and Cumulative Distribution Functions

The cdf $F(x)$ and pdf $f(x) = dF(x)/dx$ of the KB-G distribution are usually straightforward to compute from $G(x)$ and $g(x) = dG(x)/dx$. However, we provide expansions for these functions in terms of infinite (or finite) weighted sums of cdf's and pdf's of exponentiated-G distributions, respectively. In the next sections, based on these expansions, we obtain some of its structural properties including explicit expressions for the moments, moment generating functions, mean deviations and for the pdf of the order statistics and their moments.

Using the exponential expansion in (2.2), we can write

$$F(x) = \sum_{i=0}^{\infty} w_i \, H_{a+i,b}(x), \tag{2.4}$$

where $w_i = [K\, B(a+i,b)(-c)^i]/i!$ and

$$H_{a,b}(x) = \frac{1}{B(a,b)} \int_0^{G(x)} t^{a-1} \, (1-t)^{b-1} dt$$

denotes the BG cdf with positive shape parameters a and b (EUGENE; LEE; FAMOYE, 2002). Equation (2.4) reveals that the KB-G cdf is a linear combination of BG cdf's. This result is important to derive some properties of any KB-G distribution from those properties of the BG distribution.

For $b > 0$ real non-integer, we have the power series representation

$$[1 - G(x)]^{b-1} = \sum_{j=0}^{\infty} (-1)^j \binom{b-1}{j} G(x)^j, \tag{2.5}$$

where the binomial coefficient is defined for any positive real number. Expanding the term $\exp[-c\, G(x)]$ in power series and using (2.5) in equation (2.2), the KB-G cumulative distribution can be expressed as

$$F(x) = \sum_{i,j=0}^{\infty} w_{i,j} \; G(x)^{a+i+j}, \tag{2.6}$$

where

$$w_{i,j} = \frac{K\,(-1)^{i+j}\, c^i}{i!\,(a+i+j)} \binom{b-1}{j}.$$

If b is an integer, the index i in the previous sum stops at $b-1$. If a is an integer, equation (2.6) reveals that the KB-G pdf can be written by the baseline pdf multiplied by an infinite power series of its cdf.

Otherwise, if a is a real non-integer, we can expand $G(x)^{a+i+j}$ as follows:

$$G(x)^{a+i+j} = \{1 - [1 - G(x)]\}^{a+i+j} = \sum_{k=0}^{\infty} (-1)^k \binom{a+i+j}{k} \, [1 - G(x)]^k \tag{2.7}$$

and using the binomial expansion for $[1-G(x)]^k$, we obtain

$$[1-G(x)]^k = \sum_{r=0}^{k} (-1)^r \binom{k}{r} G(x)^r. \tag{2.8}$$

Then, inserting (2.8) into (2.7), we have

$$G(x)^{a+i+j} = \sum_{k=0}^{\infty} \sum_{r=0}^{k} (-1)^{k+r} \binom{a+i+j}{k} \binom{k}{r} G(x)^r.$$

Further, equation (2.2) can be rewritten as

$$F(x) = \sum_{i,j,k=0}^{\infty} \sum_{r=0}^{k} t_{i,j,k,r}\, G(x)^r, \tag{2.9}$$

where

$$t_{i,j,k,r} = t_{i,j,k,r}(a,b,c) = (-1)^{k+r} \binom{a+i+j}{k} \binom{k}{r} w_{i,j}$$

and $w_{i,j}$ is defined in (2.6). Replacing $\sum_{k=0}^{\infty}\sum_{r=0}^{k}$ by $\sum_{r=0}^{\infty}\sum_{k=r}^{\infty}$ in equation (2.9), we obtain

$$F(x) = \sum_{r=0}^{\infty} b_r\, G(x)^r, \tag{2.10}$$

where the coefficient $b_r = \sum_{i,j=0}^{\infty}\sum_{k=r}^{\infty} t_{i,j,k,r}$ represents a sum of constants.

Expansion (2.10), which holds for any real non-integer a, gives the KB-G cdf as an infinite weighted power series of cdf's of the G distribution. If b is an integer, the index i in (2.9) stops at $b-1$.

We also note that the cdf of the KB-G family of distributions can be expressed in terms of cumulative EG distributions. We have

$$F(x) = \sum_{r=0}^{\infty} b_r\, V_r(x), \tag{2.11}$$

where $V_r = G(x)^r$ is the cdf of the EG distribution with power parameter r.

The corresponding expansions for the KB-G density function are obtained by simple differentiation of (2.6) for $a > 0$ integer as

$$f(x) = g(x) \sum_{i,j=0}^{\infty} w^*_{i,j}\, G(x)^{a+i+j-1}, \tag{2.12}$$

where $w^*_{i,j} = (a+i+j)\, w_{i,j}$. Analogously, from equations (2.10) and (2.11), for $a > 0$ real non-integer, we obtain

$$f(x) = g(x) \sum_{r=0}^{\infty} b^*_r\, G(x)^r, \tag{2.13}$$

and

$$f(x) = \sum_{r=0}^{\infty} c_r \, v_{r+1}(x), \tag{2.14}$$

respectively, where $b_r^* = (r+1)\, b_{r+1}$ and $c_r = b_{r+1}$ for $r = 0, 1 \ldots$, and $v_{r+1} = (r+1)\, g(x)\, G(x)^r$ denotes the EG density function with parameter $r+1$. Equation (2.14) reveals that the KB-G density function is a linear combination of EG densities. This result is important to derive some properties of the KB-G distribution from those of the EG distribution.

Equations (2.12)-(2.14) are the main results of this section. They play an important role in this work.

2.4 General Properties of the KB-G Family

In this section, we derive some mathematical properties such as moments, moment generating function, mean deviation and entropy for any KB-G distribution.

2.4.1 Moments

In a statistical analysis especially in applied statistics, there is a great need and importance in the study of the moments of a probability distribution. Some of the most important features and characteristics of a distribution can be studied through moments (e.g., tendency, dispersion, skewness and kurtosis).

The sth moment of the KB-G distribution can be expressed as an infinite weighted sum of the probability weighted moments (PWM) of order (s, q) of the baseline G distribution from equation (2.12) for a integer and from (2.13) for a real non-integer. We assume that T and X follow the baseline G and KB-G distributions, respectively. The sth moment of X can be expressed in terms of the (s, q)th PWMs of T, say $\tau_{s,q} = \mathrm{E}[T^s\, G(T)^q]$ (for $q = 0, 1, \ldots$), as defined by (GREENWOOD et al., 1979). These weighted moments, $\tau_{s,q}$, can be derived for most baseline distributions.

For an integer a, we have

$$\mu_s' = \mathrm{E}(X^s) = \sum_{i,j=0}^{\infty} w_{i,j}^* \int x^s\, g(x)\, G(x)^{a+i+j-1}\, dx = \sum_{i,j=0}^{\infty} w_{i,j}^* \; \tau_{s,a+i+j-1}.$$

For a real non-integer a, we can write from (2.13)

$$\mu_s' = \sum_{r=0}^{\infty} b_r^* \int x^s\, g(x)\, G(x)^r\, dx = \sum_{r=0}^{\infty} b_r^* \; \tau_{s,r}.$$

So, we can calculate the moments of any KB-G distribution in terms of infinite weighted sums of PWMs of the baseline G distribution.

Alternatively, we can express μ'_s from (2.13) in terms of the baseline quantile function $Q_{\mathcal{G}}(u) = G^{-1}(u)$. We have

$$\mu'_s = \sum_{r=0}^{\infty} b_r^* \int x^s \, g(x) \, G(x)^r \, dx. \tag{2.15}$$

Setting $u = G(x)$ in (2.15), we obtain

$$\mu'_s = \sum_{r=0}^{\infty} b_r^* \int_0^1 u^r \, Q_{\mathcal{G}}(u)^s \, dt.$$

Now, we provide the moments of the KB-G distributions from equation (2.14) in terms of moments of the EG distributions. Suppose Y_{r+1} has the EG density $v_{r+1} = (r+1) \, g(x) \, G(x)^r$ with power parameter $(r+1)$. As a first example, consider for G the Weibull distribution with scale parameter $\lambda > 0$ and shape parameter $\gamma > 0$. If Y_{r+1} has the EW distribution, its moments are

$$\mathrm{E}(Y_{r+1}^s) = (r+1) \, \gamma \, \lambda^\gamma \int_0^\infty y^{s+\gamma-1} \exp\left[-(\lambda \, y)^\gamma\right] \left\{1 - \exp\left[-(\lambda \, y)^\gamma\right]\right\}^r dy. \tag{2.16}$$

Using the binomial expansion in (2.16), we obtain

$$\mathrm{E}(Y_{r+1}^s) = (r+1) \, \gamma \, \lambda^\gamma \sum_{j=0}^{r} (-1)^j \binom{r}{j} \int_0^\infty y^{s+\gamma-1} \exp\left[-(j+1) \, (\lambda \, y)^\gamma\right] dy. \tag{2.17}$$

Replacing $u = (j+1) \, (\lambda \, y)^\gamma$ in equation (2.17), the sth moment of the EW distribution can be expressed as

$$\mathrm{E}(Y_{r+1}^s) = \frac{(r+1)}{\lambda^s} \, \Gamma\left(\frac{s}{\gamma} + 1\right) \sum_{j=0}^{r} \frac{(-1)^j}{(j+1)^{1+s/\gamma}} \binom{r}{j}. \tag{2.18}$$

From equations (2.14) and (2.18), the sth moment of the KBW distribution reduces to

$$\mu'_s = \lambda^{-s} \, \Gamma\left(\frac{s}{\gamma} + 1\right) \sum_{r=0}^{\infty} \sum_{j=0}^{r} \frac{(r+1) \, c_r \, (-1)^j}{(j+1)^{1+s/\gamma}} \binom{r}{j}.$$

As a second example, taking the Gumbel distribution with cdf $G(x) = 1 - \exp[-\exp(-\frac{x-\mu}{\sigma})]$, the moments of Y_{r+1} having the exponentiated Gumbel (EGu) with parameter $(r+1)$ can be obtained from (NADARAJAH; KOTZ, 2006) as

$$\mathrm{E}(Y_{r+1}^s) = \frac{(r+1)}{\sigma} \int_{-\infty}^{\infty} y^s \left\{1 - \exp\left[-\exp\left(-\frac{y-\mu}{\sigma}\right)\right]\right\}^r \exp\left[\frac{y-\mu}{\sigma} - \exp\left(\frac{y-\mu}{\sigma}\right)\right] dy,$$

which, by replacing $u = \exp\left[-\left(\frac{y-\mu}{\sigma}\right)\right]$, reduces to

$$\mathrm{E}(Y_{r+1}^s) = (r+1) \int_0^\infty \left[\mu - \sigma \log(u)\right]^s \left[1 - \exp(-u)\right]^r \exp(-u) du. \tag{2.19}$$

Using the binomial expansion twice in (2.19), we obtain

$$\mathrm{E}(Y_{r+1}^s) \;=\; (r+1)\sum_{k=0}^{s}\sum_{m=0}^{k}(-1)^{k+m}\,\mu^{s-k}\binom{s}{k}\binom{k}{m}I(k,m), \tag{2.20}$$

where $I(k,m)$ denotes the integral

$$I(k,m)=\int_0^\infty [\log(u)]^k\,\exp[-(m+1)u]du.$$

We can note that, by equation (2.6.21.1) in (PRUDNIKOV; BRYCHKOV; MARICHEV, 1986), the integral $I(k,m)$ can be calculated as

$$I(k,m)=\left(\frac{\partial}{\partial p}\right)^k\left[(r+1)^{-p}\Gamma(p)\right]\Big|_{p=1}. \tag{2.21}$$

By combining (2.20) and (2.21), the sth moment of Y_{r+1} is given by

$$\mathrm{E}(Y_{r+1}^s) \;=\; (r+1)\sum_{k=0}^{s}\sum_{m=0}^{k}(-1)^{k+m}\,\mu^{s-k}\binom{s}{k}\binom{k}{m}\left(\frac{\partial}{\partial p}\right)^k\left[(r+1)^{-p}\Gamma(p)\right]\Big|_{p=1} \tag{2.22}$$

From (2.14) and (2.22), the sth moment of the KBGu distribution becomes

$$\mu_s'=\sum_{r=0}^{\infty}c_r\,(r+1)\sum_{k=0}^{s}\sum_{m=0}^{k}(-1)^{k+m}\,\mu^{s-k}\binom{s}{k}\binom{k}{m}\left(\frac{\partial}{\partial p}\right)^k\left[(r+1)^{-p}\Gamma(p)\right]\Big|_{p=1}.$$

2.4.2 Generating Function

Let $X\sim$ KB-G(a,b,c). In this section, we provide four representations for the moment generating function (mgf) of X, say $M(t)=\mathrm{E}[\exp(tX)]$. Clearly, the first one is given by the exponential expansion

$$M(t)=\mathrm{E}[\exp(tX)]=\mathrm{E}\left[\sum_{s=0}^{\infty}\frac{X^s}{s!}\,t^s\right]=\sum_{s=0}^{\infty}\frac{\mu_s'}{s!}\,t^s,$$

where $\mu_s'=\mathrm{E}(X^s)$. The second one comes from equation (2.3) and is given by

$$\begin{aligned}
M(t) &= \mathrm{E}[\exp(tX)]=\int\exp(tx)f(x)dx\\
&= K\int\exp(tx)\,g(x)\,G(x)^{a-1}\,[1-G(x)]^{b-1}\exp\left[-c\,G(x)\right]dx\\
&= K\int\exp\left[tx-c\,G(x)\right]\,G(x)^{a-1}\,[1-G(x)]^{b-1}g(x)dx\\
&= K\,\mathrm{E}\left\{\exp\left[t\,X-c\,G(X)\right]\,G^{a-1}(X)\,\left[1-G(X)\right]^{b-1}\right\}.
\end{aligned} \tag{2.23}$$

Using expansion (2.5) in equation (2.23), we obtain

$$M(t) \;=\; K\sum_{j=0}^{\infty}(-1)^j\binom{b-1}{j}\mathrm{E}\left[\frac{\exp\left(t\,X-Uc\right)}{U^{-(a+j-1)}}\right],$$

where U is a uniform random variable on the unit interval. Note that X and U are not independent.

A third representation for $M(t)$ is obtained from (2.14) as

$$\begin{aligned} M(t) &= \int \exp(tx) f(x) dx \\ &= \sum_{i=0}^{\infty} c_i \int \exp(tx)\, v_{i+1}(x) dx \\ &= \sum_{i=0}^{\infty} c_i\; M_{i+1}(t), \end{aligned}$$

where $M_{i+1}(t)$ is the mgf of $Y_{i+1} \sim \text{EG}(i+1)$. Hence, for any KB-G distribution, $M(t)$ can be immediately determined from the mgf of the baseline G distribution.

A fourth representation for $M(t)$ can be derived from (2.13) as

$$\begin{aligned} M(t) &= \int \exp(tx) f(x) dx \\ &= \sum_{i=0}^{\infty} b_i^* \int \exp(tx)\, g(x)\, G(x)^i dx \\ &= \sum_{i=0}^{\infty} b_i^*\; \rho(t,i), \end{aligned} \tag{2.24}$$

where the function $\rho(t,r) = \int \exp(tx)\, g(x)\, G(x)^r dx$ can be expressed from the baseline quantile function $Q_{\mathcal{G}}(u)$ as

$$\rho(t,a) = \int_0^1 u^a\; \exp\left[t\, Q_{\mathcal{G}}(u)\right]\, du. \tag{2.25}$$

We can obtain the mgf of several KB-G distributions from equations (2.24) and (2.25). For example, the mgf's of the KB-exponencial (KBE) (with parameter λ), KB-logistic (KBL) and KB-Pareto (KBPa) (with parameter $\nu > 0$) distributions are calculated from their respective quantile functions as

$$M_{\mathcal{KBGE}}(t) = \sum_{i=0}^{\infty} b_i^* \int_0^1 u^i\, (1-u)^{-\lambda\, t^{-1}} du = \sum_{i=0}^{\infty} b_i^*\, B\left(i+1, 1-\lambda\, t^{-1}\right),$$

$$M_{\mathcal{KBGL}}(t) = \sum_{i=0}^{\infty} b_i^* \int_0^1 u^{i+t}\, (1-u)^{1-t} du = \sum_{i=0}^{\infty} b_i^*\, B\left(i+t+1, 1-t\right)$$

and

$$M_{\mathcal{KBGPA}}(t) = \sum_{i=0}^{\infty} b_i^* \int_0^1 u^i\; \exp\left[\frac{t}{(1-u)^{1/\nu}}\right] du = \sum_{i,p=0}^{\infty} \frac{b_i^*\, t^p}{p!}\, B\left(i+1, 1-p\,\nu^{-1}\right),$$

respectively.

Clearly, four representations for the characteristic function (chf) $\phi(t) = \text{E}[\exp(\text{i}\, tX)]$ of any KB-G distribution are immediately obtained from the above representations for the mgf by $\phi(t) = M(\text{i}\, t)$, where $\text{i} = \sqrt{-1}$.

2.4.3 Mean Deviations

The amount of scattering in a population may be measured by the totality of the absolute values of the deviations from the mean (in case of a symmetric distribution) or in relation to the median (in case of an asymmetric distribution).

Let $X \sim$ KB-$G(a,b,c)$. The mean deviations about the mean, $\delta_1(X)$, and about the median, $\delta_2(X)$, are defined, respectively, by

$$\delta_1(X) = \int_{-\infty}^{\infty} |x-\mu_1'|\, f(x)\, dx \quad \text{and} \quad \delta_2(X) = \int_{-\infty}^{\infty} |x-M|\, f(x)\, dx.$$

The mean deviation in relation to the mean and to the median can be simplified as

$$\begin{aligned}
\delta_1(X) &= \int_{-\infty}^{\infty} |x-\mu_1'|\, f(x)\, dx \\
&= \int_{-\infty}^{\mu_1'} (\mu_1'-x)\, f(x)\, dx + \int_{\mu_1'}^{\infty} (x-\mu_1')\, f(x)\, dx \\
&= \int_{-\infty}^{\mu_1'} (\mu_1'-x)\, f(x)\, dx + \int_{-\infty}^{\infty} (x-\mu_1')\, f(x)\, dx - \int_{-\infty}^{\mu_1'} (x-\mu_1')\, f(x)\, dx \\
&= 2\int_{-\infty}^{\mu_1'} (\mu_1'-x)\, f(x)\, dx \\
&= 2\mu_1'\, F(\mu_1') - 2\,T(\mu_1') \qquad (2.26)
\end{aligned}$$

and

$$\begin{aligned}
\delta_2(X) &= \int_{-\infty}^{\infty} |x-M|\, f(x)\, dx \\
&= \int_{-\infty}^{M} (M-x)\, f(x)\, dx + \int_{M}^{\infty} (x-M)\, f(x)\, dx \\
&= \int_{-\infty}^{M} (M-x)\, f(x)\, dx + \int_{-\infty}^{\infty} (x-M)\, f(x)\, dx - \int_{-\infty}^{M} (x-M)\, f(x)\, dx \\
&= 2\int_{-\infty}^{M} (M-x)\, f(x)\, dx + \int_{-\infty}^{\infty} x\, f(x)\, dx - M\int_{-\infty}^{\infty} f(x)\, dx \\
&= \mu_1' + 2M\, F(M) - M - 2\,T(M), \qquad (2.27)
\end{aligned}$$

respectively, where $\mu_1' = \mathrm{E}(X)$, $F(\mu_1')$ comes from (2.2), $M = Median(X)$ denotes the median determined from the nonlinear equation $F(M) = 1/2$ and $T(z) = \int_{-\infty}^{z} x\, f(x) dx$.

Applying (2.13) in $T(z)$, we obtain

$$T(z) = \sum_{r=0}^{\infty} b_r^* \int_{-\infty}^{z} x\, g(x)\, G(x)^r. \qquad (2.28)$$

Substituting $u = G(x)$ in (2.28), it yields

$$T(z) = \sum_{r=0}^{\infty} b_r^*\, T_r(z), \qquad (2.29)$$

where the integral $T_r(z)$ can be written in terms of the quantile function, $Q_{\mathcal{G}}(u) = G^{-1}(u)$, by

$$T_r(z) = \int_0^{G(z)} u^r \, Q_{\mathcal{G}}(u) \, du. \tag{2.30}$$

The mean deviations of any KB-G distribution can be computed from equations (2.26)-(2.30). An alternative representation for $T(z)$ is derived from (2.14) as

$$T(z) = \int_{-\infty}^{z} x \, f(x) dx = \sum_{r=0}^{\infty} c_r \, J_{r+1}(z), \tag{2.31}$$

where

$$J_{r+1}(z) = \int_{-\infty}^{z} x \; v_{r+1}(x) dx. \tag{2.32}$$

Equation (2.32) is the basic quantity to compute the mean deviations in terms of the EG distributions. Hence, the KB-G mean deviations depend only on the quantity $J_{r+1}(z)$. So, alternative representations for $\delta_1(X)$ and $\delta_2(X)$ are given by

$$\delta_1(X) = 2\mu_1' F\left(\mu_1'\right) - 2\sum_{r=0}^{\infty} c_r \, J_{r+1}\left(\mu_1'\right) \quad \text{and} \quad \delta_2(X) = \mu_1' - 2\sum_{r=0}^{\infty} c_r \, J_{r+1}(M).$$

A simple application is provided for the KBW distribution. The EW density function with parameters λ, γ and $r+1$ is given by (for $x > 0$)

$$v_{r+1}(x) = (r+1)\,\gamma\,\lambda^{\gamma}\,x^{\gamma-1}\,\exp\left[-(\lambda x)^{\gamma}\right]\,\left\{1 - \exp\left[-(\lambda x)^{\gamma}\right]\right\}^r$$

and then

$$J_{r+1}(z) \;=\; (r+1)\,\gamma\,\lambda^{\gamma} \int_0^z x^{\gamma} \exp\left[-(\lambda x)^{\gamma}\right] \left\{1 - \exp\left[-(\lambda x)^{\gamma}\right]\right\}^r dx.$$

Following the same steps of expression (2.16), we have

$$J_{r+1}(z) \;=\; (r+1)\,\gamma\,\lambda^{\gamma} \sum_{j=0}^{\infty} (-1)^j \binom{r}{j} \int_0^z x^{\gamma} \, \exp\left[-(j+1)(\lambda x)^{\gamma}\right] dx. \tag{2.33}$$

The integral (2.33) can be calculated by incomplete gamma function and then

$$J_{r+1}(z) = (r+1)\,\lambda^{-1} \sum_{j=0}^{\infty} \frac{(-1)^j \binom{r}{j}}{(j+1)^{1+\gamma^{-1}}} \; \gamma\left(1+\gamma^{-1}, (j+1)(\lambda z)^{\gamma}\right).$$

Equations (2.29) and (2.31) are the main results of this section.

2.4.4 Rényi Entropy

Entropies are measures which quantifies the diversity or randomness of a random variable X. They express the expected information content or uncertainty of a probability distribution. Entropy measures provide important tools to indicate variety in distributions at particular moments

in time (for example, market shares) and to analyze evolutionary processes over time (technical change). There are several applications of entropy mainly in innovation studies and income inequality.

One of the most popular measures of entropy is the Rényi entropy defined by

$$\mathcal{J}_R(\xi) = \frac{1}{1-\xi} \log \left[\int_{-\infty}^{\infty} f^{\xi}(x) dx \right], \; \xi > 0 \text{ and } \xi \neq 1.$$

For any KB-G distribution, the integral above can be expressed as

$$\int_{-\infty}^{\infty} f^{\xi}(x) dx = K^{\xi} \int_{-\infty}^{\infty} g^{\xi}(x)\, G^{\xi(a-1)}(x)\; [1 - G(x)]^{\xi(b-1)} \exp\left[-\xi\, c\, G(x)\right]\, dx \tag{2.34}$$

and then, expanding the exponential and the binomial terms in (2.34), we obtain

$$\int_{-\infty}^{\infty} f^{\xi}(x) dx = K^{\xi} \sum_{i,j=0}^{\infty} \frac{(-1)^{i+j} (c\,\xi)^i}{i!} \binom{\xi(b-1)}{j} I_{i,j}(\xi), \tag{2.35}$$

where $I_{i,j}(\xi)$ denotes the integral

$$I_{i,j}(\xi) = \int_0^1 g^{\xi-1}\left(Q_{\mathcal{G}}(u)\right)\; u^{i+j+\xi(a-1)}\, du,$$

to be calculated for each KB-G model. We note that, $Q_{\mathcal{G}}(.)$ represents the quantile function of the baseline G distribution. For the KBE (with parameter λ), KBL and KBPa (with parameter ν) distributions, we obtain

$$I_{i,j}(\xi) = \lambda^{\xi-1} B\left(i + j + \xi(a-1) + 1, \xi\right), \quad I_{i,j}(\xi) = B\left(i + j + \xi\, a\,,\, \xi\right),$$

and

$$I_{i,j}(\xi) = \nu^{\xi-1}\, B\left(i + j + \xi(a-1) + 1,\, \nu^{-1}(\xi - 1) + \xi\right),$$

respectively. Equation (2.35) is the main result of this section.

2.5 Order Statistics

Order statistics have been used in a wide range of problems, including robust statistical estimation and detection of outliers, characterization of probability distributions and goodness-of-fit tests, entropy estimation, analysis of censored data, reliability analysis, quality control and strength of materials.

Suppose $X_1, \ldots, X_n$ is a random sample from a continuous distribution and let $X_{1:n} < \cdots < X_{i:n}$ denote the corresponding order statistics. There has been a large amount of work relating to moments of order statistics $X_{i:n}$, see (ARNOLD; BALAKRISHNAN; NAGARAJA, 1992), (DAVID; NAGARAJA, 2003) and (AHSANULLAH; NEVZOROV, 2005) for excellent accounts. It is well-known that

$$f_{i:n}(x) = \frac{f(x)}{B(i, n-i+1)} F(x)^{i-1} \left[1 - F(x)\right]^{n-i}, \tag{2.36}$$

Using the binomial expansion in (2.36), we have

$$f_{i:n}(x) = \frac{f(x)}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F(x)^{i+j-1}. \tag{2.37}$$

We now provide an expression for the pdf of the KB-G order statistics as a function of the baseline pdf multiplied by infinite weighted sums of powers of $G(x)$. Based on this result we can derive the ordinary moments of the order statistics of any KB-G distribution as infinite weighted sums of the PWMs of the baseline G distribution.

Replacing (2.10) in equation (2.37), we have

$$F(x)^{i+j-1} = \left(\sum_{r=0}^{\infty} b_r \, u^r \right)^{i+j-1}, \tag{2.38}$$

where $u = G(x)$ is the baseline cdf.

We use the identity $(\sum_{k=0}^{\infty} a_k \, x^k)^n = \sum_{k=0}^{\infty} d_{k,n} \, x^k$ (GRADSHTEYN; RYZHIK, 2007), where

$$d_{0,n} = a_0^n \qquad \text{and} \qquad d_{k,n} = (k a_0)^{-1} \sum_{m=1}^{k} [m(n+1) - k] \; a_m \, d_{k-m,n}$$

(for $k = 1, 2, \ldots$) in equation (2.38) to obtain

$$F(x)^{i+j-1} = \sum_{r=0}^{\infty} d_{r,i+j-1} \, G(x)^r, \tag{2.39}$$

where

$$d_{0,i+j-1} = b_0^{i+k-1} \quad \text{and} \quad d_{r,i+j-1} = (k b_r)^{-1} \sum_{m=1}^{r} [(i+j)m - r] \; b_m \, d_{r-m,i+j-1}.$$

For real non-integer $a > 0$, inserting (2.13) and (2.39) into equation (2.37) and changing indices, we can rewrite $f_{i:n}(x)$ for any KB-G distribution in the form

$$f_{i:n}(x) = \frac{g(x)}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} \sum_{u,v=0}^{\infty} b_u^* \, d_{u,i+j-1} \, G(x)^{u+v}. \tag{2.40}$$

For an integer $a > 0$, we can obtain from equations (2.12), (2.37) and (2.39)

$$f_{i:n}(x) = \frac{g(x)}{B(i, n-i+1)} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} \sum_{p,q,u=0}^{\infty} w_{p,q}^* \, d_{u,i+j-1} \, G(x)^{a+p+q+u-1}. \tag{2.41}$$

Equations (2.40) and (2.41) immediately yield the pdf of the KB-G order statistics as a function of the baseline pdf multiplied by infinite weighted sums of powers of $G(x)$. Hence, the moments of the KB-G order statistics can be expressed as infinite weighted sums of PWMs of the G distribution. Clearly, equation (2.41) can be given in terms of linear combinations of EG densities. So, the moments and mgf of the KB-G order statistics can immediately follow from linear combinations of those quantities for the EG distributions.

2.6 Inference

2.6.1 Maximum likelihood method

Let $\boldsymbol{\gamma}$ be the p-dimensional parameter vector of the baseline G distribution in equations (2.2) and (2.3). We consider independent random variables $X_1, \ldots, X_n$, each X_i following a KB-G distribution with parameter vector $\boldsymbol{\theta} = (a, b, c, \boldsymbol{\gamma}^T)^T$. The log-likelihood function $\ell = \ell(\theta)$ for the model parameters obtained from (2.3) is

$$\begin{aligned}\ell(\theta) &= n\log(K) + \sum_{i=1}^{n} \log g\left(x_i;\gamma\right) - c\sum_{i=1}^{n} G\left(x_i;\gamma\right) \\ &\quad + (a-1)\sum_{i=1}^{n} \log\left[G\left(x_i;\gamma\right)\right] + (b-1)\sum_{i=1}^{n} \log\left[1 - G\left(x_i;\gamma\right)\right]. \qquad (2.42)\end{aligned}$$

The elements of score vector are given by

$$\frac{\partial \ell(\theta)}{\partial a} = \frac{n}{K}\frac{\partial K}{\partial a} + \sum_{i=1}^{n} \log\left[G\left(x_i;\gamma\right)\right],$$

$$\frac{\partial \ell(\theta)}{\partial b} = \frac{n}{K}\frac{\partial K}{\partial b} + \sum_{i=1}^{n} \log\left[1 - G\left(x_i;\gamma\right)\right],$$

$$\frac{\partial \ell(\theta)}{\partial c} = \frac{n}{K}\frac{\partial K}{\partial c} - \sum_{i=1}^{n} G\left(x_i;\gamma\right)$$

and

$$\begin{aligned}\frac{\partial \ell(\theta)}{\partial \gamma_j} &= \sum_{i=1}^{n}\left[\frac{1}{g\left(x_i;\gamma\right)}\frac{\partial g\left(x_i;\gamma\right)}{\partial \gamma_j} - c\frac{\partial g\left(x_i;\gamma\right)}{\partial \gamma_j}\right. \\ &\quad \left. + \frac{(a-1)}{G\left(x_i;\gamma\right)}\frac{\partial G\left(x_i;\gamma\right)}{\partial \gamma_j} + \frac{(b-1)}{1-G\left(x_i;\gamma\right)}\frac{\partial G\left(x_i;\gamma\right)}{\partial \gamma_j}\right],\end{aligned}$$

for $j = 1, \ldots, p$, where

$$\frac{\partial K}{\partial a} = -\frac{\left\{\left[\psi(a) - \psi(a+b)\right]\ {}_1F_1(a, a+b, -c) + \frac{\partial\, {}_1F_1(a,a+b,-c)}{\partial a}\right\}}{B(a,b)\left[{}_1F_1(a, a+b, -c)\right]^2},$$

$$\frac{\partial K}{\partial b} = -\frac{\left\{\left[\psi(b) - \psi(a+b)\right]\ {}_1F_1(a, a+b, -c) + \frac{\partial\, {}_1F_1(a,a+b,-c)}{\partial b}\right\}}{B(a,b)\left[{}_1F_1(a, a+b, -c)\right]^2},$$

$$\frac{\partial K}{\partial c} = \frac{a\, {}_1F_1(a+1, a+b+1, -c)}{(a+b)B(a,b){}_1F_1(a, a+b, -c)},$$

$$\frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial a} = -\left[\psi(a) - \psi(a+b)\right]\, {}_1F_1(a, a+b, -c) - \sum_{k=0}^{\infty} \frac{(a)_k(-c)^k}{k!(a+b)_k}\left[\psi(a+b+k) - \psi(a+k)\right]$$

and

$$\frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial b} = \psi(a+b)\, {}_1F_1(a, a+b, -c) + \sum_{k=0}^{\infty} \frac{(a)_k(-c)^k}{k!(a+b)_k}\, \psi(a+b+k).$$

These partial derivatives depend on the specified baseline G distribution. Numerical maximization of the log-likelihood (2.42) is accomplished by using the RS method (RIGBY; STASINOPOULOS, 2005) available in the gamlss package (STASINOPOULOS; RIGBY, 2007) in statistical software R.

For interval estimation of each parameter in $\boldsymbol{\theta} = (a, b, c, \boldsymbol{\gamma}^T)^T$, and tests of hypotheses, we require the observed information matrix. Interval estimation for the model parameters can be obtained with standard likelihood theory. The elements of the information matrix for (2.42) are given in the Appendix A. Under suitable regularity conditions, the asymptotic distribution of the maximum likelihood estimator (MLE) $\widehat{\theta}$ is multivariate normal with the mean vector θ and the variance and covariance matrix that can be estimated by $\{-\partial^2\ell(\theta)/\partial\theta\partial\theta^T\}$ evaluated at $\theta = \widehat{\theta}$. The required second derivatives can be computed numerically.

Consider two nested KB-G distributions: a KB-G_A distribution with corresponding parameters $\theta_1, \ldots, \theta_r$ and maximized log-likelihood $-2\,\ell(\widehat{\theta}_A)$, and a KB-$G_B$ distribution containing the same parameters $\theta_1, \ldots, \theta_r$ plus additional parameters $\theta_{r+1}, \ldots, \theta_p$ and maximized log-likelihood $-2\,\ell(\widehat{\theta}_B)$, the models otherwise being identical. For testing the KB-G_A distribution against the KB-G_B distribution, the likelihood ratio (LR) statistic is simply equal to $w = -2\left[\ell(\widehat{\theta}_A) - \ell(\widehat{\theta}_B)\right]$ and it has an asymptotic χ^2_{p-r} distribution.

We compare non-nested KB-G distributions by penalizing the over-fitting using the Akaike information criterion (AIC) given by AIC $= -2\,\ell(\widehat{\theta}) + 2p^*$ and the Bayesian information criterion (BIC) defined by BIC $= -2\,\ell(\widehat{\theta}) + p^*\log(n)$, where p^* is the number of model parameters and n is the sample size. The distribution with the smallest value of any of these criteria (among all distribution considered) is usually taken as the best choice for describing the given data set.

2.6.2 Bayesian Inference

The Bayesian approach allows the incorporation of previous knowledge of the parameters through informative prior density functions. When this information is not available, we can consider a non-informative prior. In the Bayesian context, the information referring to the model parameters is obtained through a posterior marginal distribution. Thus, two difficulties usually arise. The first refers to attaining marginal posterior distribution, and the second to the calculation of the moments of interest. Both cases require numerical integration that, many times, do not present an analytical solution. To overcome these problems, we use the simula-

tion methods based on the Markov Chain Monte Carlo (MCMC), such as the Gibbs sampler and Metropolis-Hastings algorithms.

Since we have no prior information from historical data or from previous experiment, we assign weakly informative prior distributions to the parameters. Since we assumed informative (but weakly) prior distribution, the posterior distribution is a well-defined proper distribution. We assume the elements of the parameter vector $\boldsymbol{\theta} = (a, b, c, \boldsymbol{\gamma}^T)^T$ to be independent and consider that the joint prior distribution of all unknown parameters has a density function given by

$$\pi(a, b, c, \gamma) \propto \pi(a) \times \pi(b) \times \pi(c) \times \pi(\boldsymbol{\gamma}), \tag{2.43}$$

where $\boldsymbol{\gamma}$ is the p-dimensional parameter vector of the baseline G distribution. We can note, in the literature, that gamma and normal priors are most commonly used priors for positive and real-values parameters.

Combining the likelihood function (2.42) and the joint prior distribution (4.42), the joint posterior distribution for a, b, c and $\boldsymbol{\gamma}$ can be expressed as

$$\begin{aligned} \pi(a, b, c, \boldsymbol{\gamma}|x) \quad \propto \quad & K^n \exp\left[-c \sum_{i=1}^{n} G(x_i; \boldsymbol{\gamma})\right] \\ & \times \prod_{i=1}^{n} g(x_i; \boldsymbol{\gamma})\, G(x_i; \boldsymbol{\gamma})^{a-1} \left[1 - G(x_i; \boldsymbol{\gamma})\right]^{b-1} \times \pi(a, b, c, \boldsymbol{\gamma}). \end{aligned} \tag{2.44}$$

In general, the joint posterior density function (2.44) for any KB-G distribution may be analytically intractable because its integration is not easy to perform. So, the inference on the parameters can be based on MCMC simulation methods to draw samples of the marginal distributions and then, we calculate the features of interest. In this way, we first determine the full conditional distributions of each unknown parameter and after that, we require the use of MCMC computations to obtain the posterior estimates of parameters.

2.7 Applications

In this section, we shall present two applications using well-known data sets to demonstrate the flexibility and applicability of the proposed family of distributions.

2.7.1 USS Halfbeak diesel engine data set

Here, we shall compare the fits of the Kummer beta gamma (KBGa) distribution with those of two sub-models (i.e. the beta gamma (BGa) and gamma distributions) and also to the following non-nested models: the Kumaraswamy generalized gamma (KwGG) (PASCOA; ORTEGA; CORDEIRO, 2011) and the Kumaraswamy Weibull (KwW) (CORDEIRO; ORTEGA; NADARAJAH, 2010) distributions to the data set studied by (ASCHER; FEINGOLD, 1984).

They describe a real data set from USS Halfbeak (submarine) diesel engine. The data are 73 failure times (in hours) of unscheduled maintenance actions for the USS Halfbeak number

4 main propulsion diesel engine over 25.518 operating hours. Table 2.2 gives a descriptive summary for the data and suggest skewed distribution. The KBGa model seems to account very well for the degrees of skewness and kurtosis present in the data set.

Table 2.2 – Descriptive statistics for diesel engine data from the USS Halfbeak

Mean	Median	SD	Variance	Skewness	Kurtosis	Min.	Max.
19.39	21.46	5.81	33.83	-1.54	4.45	1.38	25.51

(i) Maximum Likelihood Estimation

Firstly, in order to estimate the model parameters, we consider the maximum likelihood method discussed in Section 2.6.1. We take the estimates of α and β from the fitted gamma distribution as starting values for the numerical iterative procedure. All computations were performed using statistical software R. Table 2.3 lists the maximum likelihood estimates (MLEs) and the corresponding standard errors (SEs) of the parameters and the values of the following statistics for some models: AIC and BIC as discussed before. The results indicate that the KBGa model has the smallest values of these statistics among all fitted models. So, it could be chosen as the more suitable model.

Table 2.3 – MLEs and the corresponding SEs (given in parentheses) of the model parameters for the diesel engine data and the measures AIC and BIC.

Model	α	β	a	b	c	AIC	BIC
KBGa	32.9915	1.5723	0.0630	4.0543	-11.7926	395.7	407.0
	(0.0097)	(0.0191)	(0.0081)	(0.0121)	(0.0114)		
BGa	28.5945	0.6636	0.1524	114.3634	0	446.1	455.2
	(0.0122)	(0.0036)	(0.0059)	(0.0779)	(-)		
Gamma	5.8340	0.3007	1	1	0	492.8	497.3
	(0.0095)	(0.0051)	(-)	(-)	(-)		
KwGG	α	τ	k	λ	φ		
	25.0291	73.5624	2.2142	0.0134	0.6825	416.8	428.2
	(0.0325)	(0.0841)	(0.0262)	(0.0001)	(0.0011)		
KwW	d	β	a	b			
	8.9196	0.0239	0.4772	17.5320	(-)	457.2	466.2
	(0.0218)	(0.0003)	(0.0047)	(0.0541)	(-)		

A comparison of the proposed distribution with some of its sub-models using LR statistics is given in Table 2.4. We reject the null hypotheses of the two LR tests in favor of the KBGa distribution. The rejection is extremely highly significant for the diesel engine data. This gives a clear evidence of the potential of the three skewness parameters when modeling real data.

Secondly, in order to assess whether the model is appropriate, Figures 2.3a and 2.3b display the histogram of the data and the fitted KBGa density function and some densities of its sub-models and non-nested models, respectively. Further, Figures 2.3c and 2.3d display plots of the empirical and estimated survival functions of the KBGa distribution and of some sub-models

Table 2.4 – LR statistics for the diesel engine data.

Model	Hypotheses	Statistic w	p-value
KBGa vs BGa	$H_0 : c = 0$ vs $H_1 : H_0$ is false	103.13	< 0.0001
KBGa vs Gamma	$H_0 : a = b = 1$ and $c = 0$ vs $H_1 : H_0$ is false	52.46	< 0.0001

and non-nested models, respectively. We can conclude that the KBGa distribution is a very suitable model to fit to these data.

The QQ plots of the normalized quantile residuals were introduced by (DUNN; SMYTH, 1996) and more recently used by (CORDEIRO et al., 2013). Figure 2.4 indicates the improved fit achieved using the KBGa distribution over the other distributions. We also emphasize the gain yielded by the KBGa distribution in relation to the gamma, BGa, KwGG and KwW distributions.

(ii) Bayesian Analysis

In the Bayesian context, if X is a random variable which follows the KBGa(a, b, c, α, β) distribution, we assume that all parameters a, b, c, α and β have independent priors given by equation (2.45),

$$\pi(a, b, c, \alpha, \beta) \propto \pi(a) \times \pi(b) \times \pi(c) \times \pi(\alpha) \times \pi(\beta). \tag{2.45}$$

Here, $a \sim \Gamma(a_1, b_1)$, $b \sim \Gamma(a_2, b_2)$, $c \sim \mathrm{N}(\mu_0, \sigma_0^2)$, $\alpha \sim \Gamma(a_3, b_3)$ and $\beta \sim \Gamma(a_4, b_4)$, where, $\Gamma(a_i, b_i)$ denotes the gamma distribution with mean a_i/b_i, variance a_i/b_i^2 for $a_i > 0$ and $b_i > 0$, and $\mathrm{N}(\mu_0, \sigma_0^2)$ represents the normal distribution with mean μ_0, variance σ_0^2 for $\mu_0 \in \mathbb{R}$ and $\sigma_0^2 > 0$.

Inserting $G(x; \boldsymbol{\gamma})$ and $g(x; \boldsymbol{\gamma})$ to be the cdf and pdf of the gamma distribution in equation (2.44), we obtain the joint posterior density function for a, b, c, α and β as

$$\begin{aligned}\pi(a, b, c, \alpha, \beta|x) \quad \propto \quad & \left(\frac{K\,\beta^{\alpha}}{\Gamma(\alpha)}\right)^n \exp\left[-\beta \sum_{i=1}^{n} x_i - c \sum_{i=1}^{n} \gamma_1(\alpha, \beta\, x_i)\right] \\ & \times \prod_{i=1}^{n} x_i^{\alpha-1}\, \gamma_1(\alpha, \beta\, x_i)^{a-1}\, [1 - \gamma_1(\alpha, \beta\, x_i)]^{b-1} \\ & \times \pi(a, b, c, \alpha, \beta)\end{aligned} \tag{2.46}$$

and then, the full conditional distributions of each unknown quantity are given by

$$\pi(a|x, b, c, \alpha, \beta) \quad \propto \quad K^n \prod_{i=1}^{n} \gamma_1(\alpha, \beta\, x_i)^{a-1} \;\times\; \pi(a),$$

$$\pi(b|x, a, c, \alpha, \beta) \quad \propto \quad K^n \prod_{i=1}^{n} [1 - \gamma_1(\alpha, \beta\, x_i)]^{b-1} \;\times\; \pi(b),$$

$$\pi(c|x, a, b, \alpha, \beta) \quad \propto \quad K^n \exp\left[-c \sum_{i=1}^{n} \gamma_1(\alpha, \beta\, x_i)\right] \times \pi(c),$$

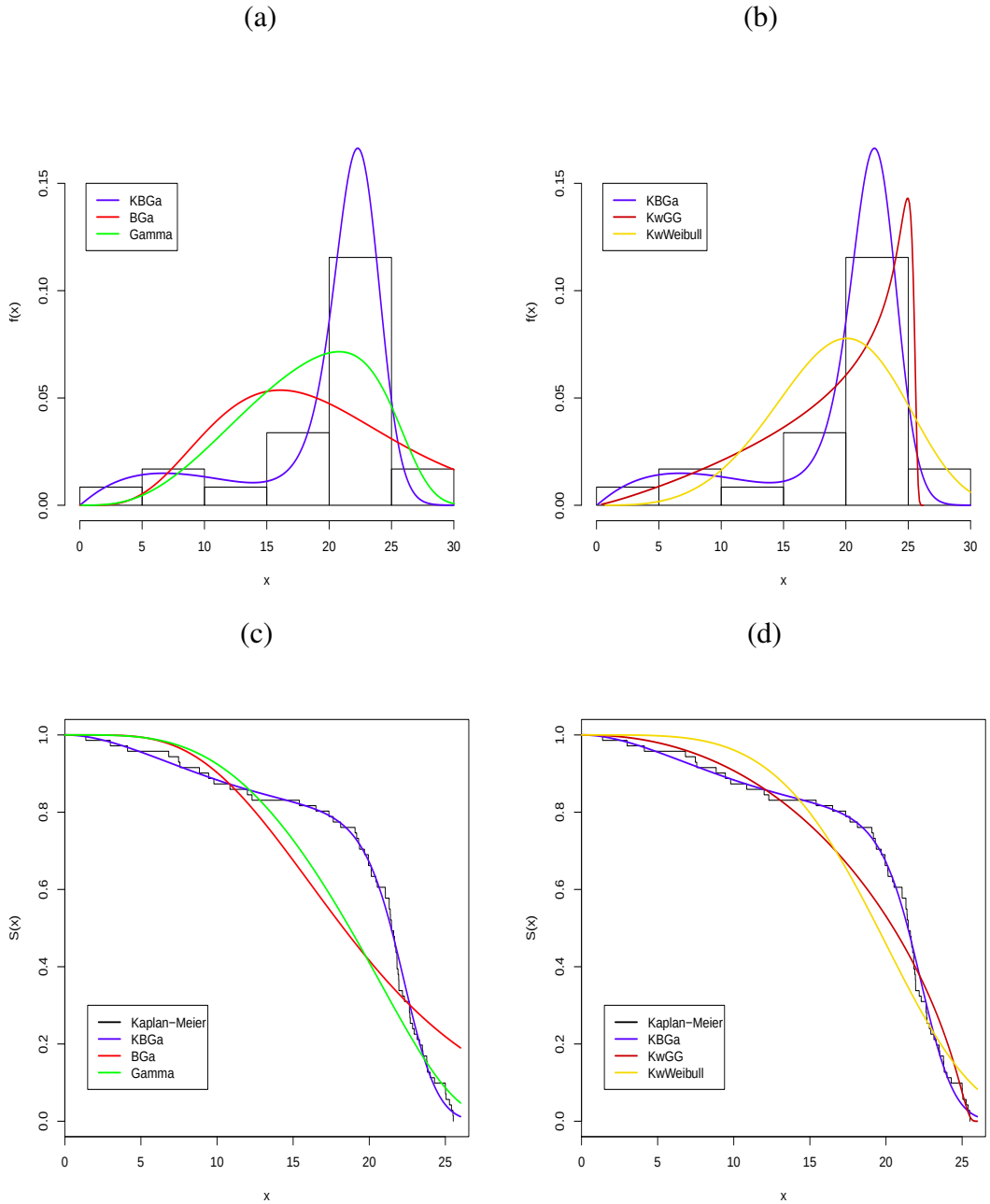


Figure 2.3 – (a) Estimated densities of the KBGa and its sub-models (b) Estimated densities of the KBGa, KwGG and Kw-Weibull models (c) Empirical and estimated survival functions of the KBGa and its sub-models (d) Empirical and estimated survival functions of the KBGa, KwGG and Kw-Weibull models

$$\pi(\alpha|x,a,b,c,\beta) \propto \left(\frac{\beta^{\alpha}}{\Gamma(\alpha)}\right)^{n} \exp\left[-c\sum_{i=1}^{n}\gamma_1(\alpha,\beta\, x_i)\right] \times \prod_{i=1}^{n} x_i^{\alpha-1}\,\gamma_1(\alpha,\beta\, x_i)^{a-1}\,\left[1-\gamma_1(\alpha,\beta\, x_i)\right]^{b-1} \times \pi(\alpha)$$

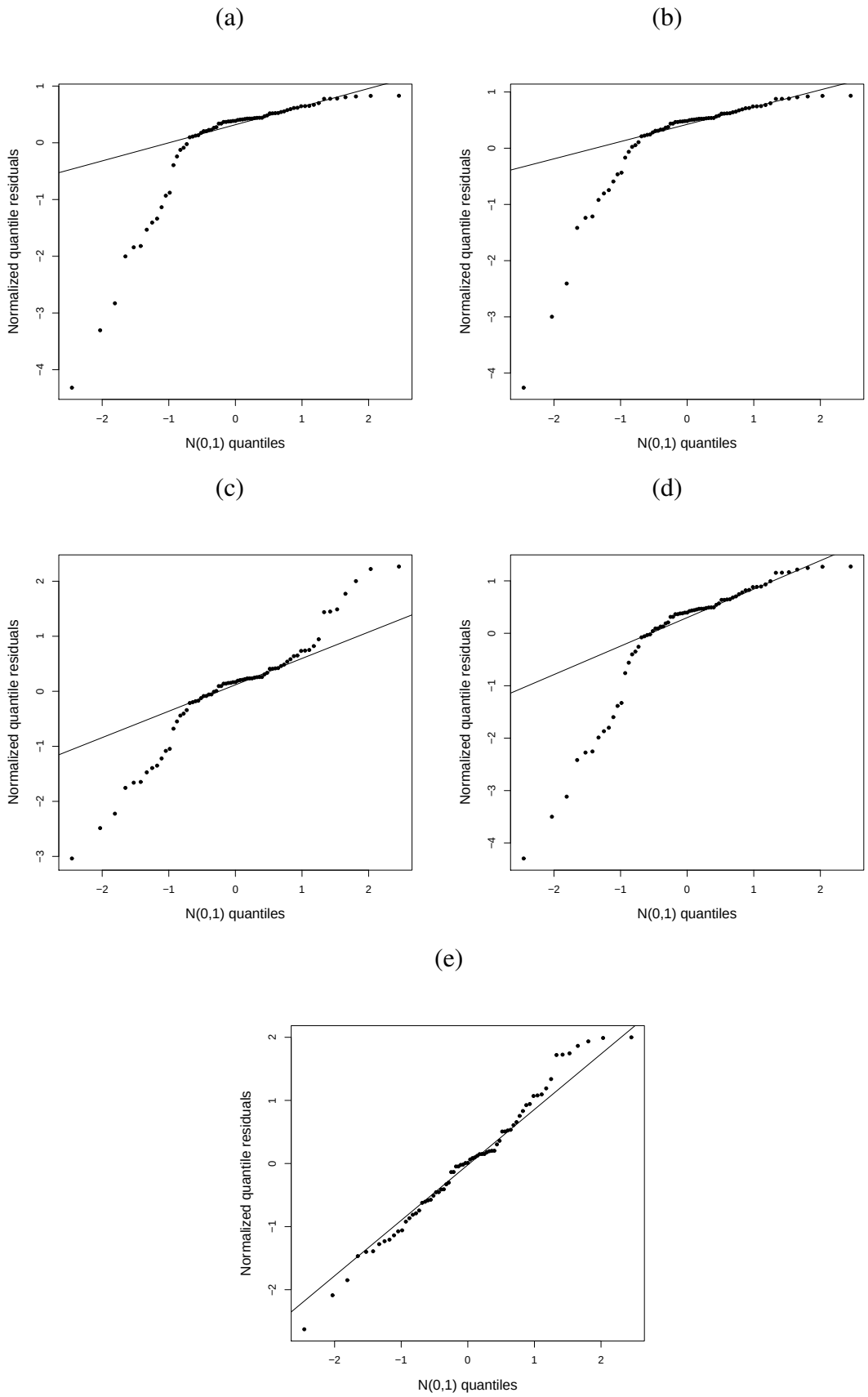


Figure 2.4 – QQ plot of the normalized quantile residuals with an identity line for the distributions: (a) Gamma, (b) BGa, (c) KwGG, (d) KwW and (e) KBGa

and

$$\pi(\beta|x,a,b,c,\alpha) \quad \propto \quad \beta^{\alpha n} \exp\left[-\beta \sum_{i=1}^{n} x_i - c \sum_{i=1}^{n} \gamma_1(\alpha, \beta\, x_i)\right]$$
$$\times \prod_{i=1}^{n} \gamma_1(\alpha, \beta\, x_i)^{a-1} \, [1-\gamma_1(\alpha, \beta\, x_i)]^{b-1} \times \pi(\beta).$$

Since the full conditional distributions for a, b, c, α and β do not have a closed form, we require the use of the Metropolis-Hastings algorithm. All MCMC computations were implemented in statistical software R.

Now, we consider the following independent priors to perform the Metropolis-Hastings algorithm:
$a \sim \Gamma(0.001, 0.001)$, $b \sim \Gamma(0.001, 0.001)$, $c \sim \mathrm{N}(0, 1000)$, $\alpha \sim \Gamma(0.001, 0.001)$ and $\beta \sim \Gamma(0.001, 0.001)$,

so that we have a vague prior distribution. Considering these prior density functions, we generated two parallel independent runs of the Metropolis-Hastings with size 300.000 for each parameter, disregarding the first 30.000 iterations to eliminate the effect of the initial values and, to avoid correlation problems, we considered a spacing of size 10, obtaining a sample of size 27.000 from each chain. To monitor the convergence of the Metropolis-Hastings, we performed the methods suggested by (COWLES; CARLIN, 1996). To monitor the convergence of the samples, we used the between and within sequence information, following the approach developed in (GELMAN; RUBIN, 1992) to obtain the potential scale reduction, $\widehat{R}$. In all cases, these values were close to one, indicating the convergence of the chain.

The histograms with the approximate posterior marginal density functions of the parameters are displayed in Figure 2.5. We report, in Table 4.4, the posterior summaries (posterior means, standard deviation (SD) and the 95% highest posterior density (HPD) intervals) for all parameters of the KBGa distribution. We can note that the values for posterior means (Table 2.5) are in good agreement with the MLEs, as expected.

Table 2.5 – Posterior summaries for the parameters of the KBGa model for the diesel engine data

Parameter	Mean	SD	HPD (95%)	$\widehat{R}$
a	0.0603	0.0099	(0.0406; 0.0793)	0.9997
b	4.0500	0.0100	(4.0305; 4.0698)	0.9998
c	-11.7900	0.0100	(-11.8097; -11.7704)	1.0005
α	32.9901	0.0099	(32.9699; 33.0089)	1.0009
β	1.5700	0.0100	(1.5507; 1.5899)	1.0003

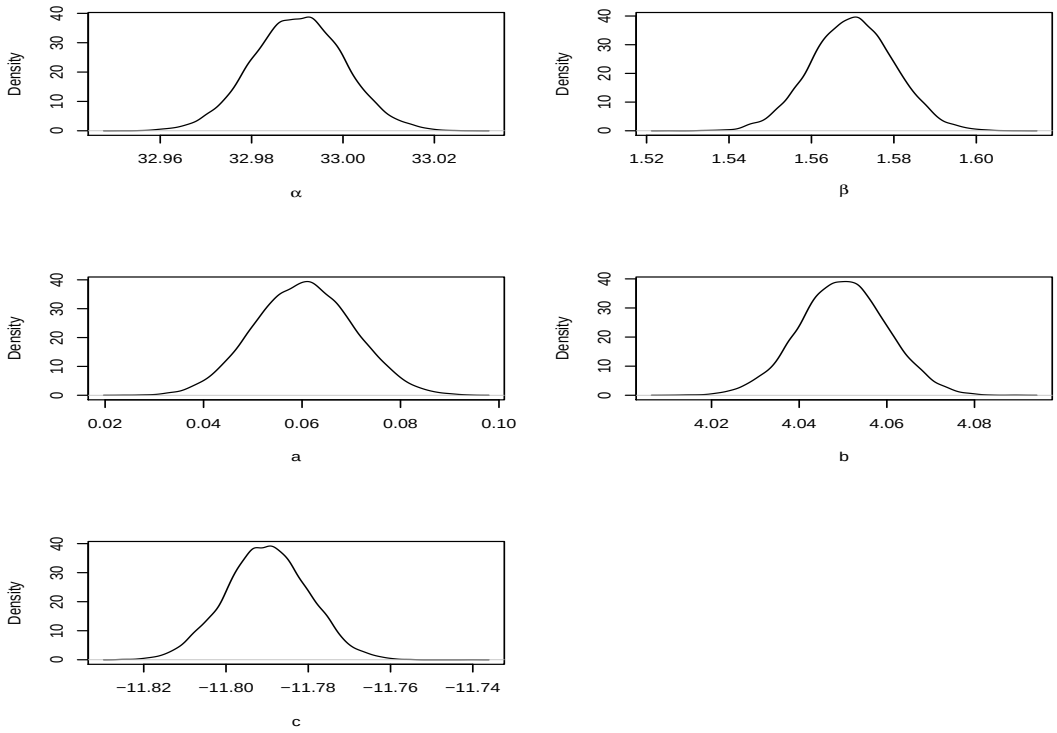


Figure 2.5 – Approximate posterior marginal densities for the parameters of the KBGa model for the diesel engine data

2.7.2 INPC data set

This section contains an application of the Kummer beta normal (KBN) distribution to real data. We also compare the fits of the KBN distribution with those of two sub-models (i.e. the beta normal (BN) and normal distributions) and also to the following non-nested models: the Kumaraswamy normal (KwN) (CORDEIRO; de CASTRO, 2011), the McDonald Normal (McN) (CORDEIRO et al., 2012) and the skew-normal (SN) distributions to INPC data set.

The INPC is a national index of consumer prices of Brazil, released by IBGE (Instituto Brasileiro de Geografia e Estatística). The period of collection goes from day 1 to 30 of the reference month and the target population includes families dwelling in the urban areas, whose head of the household is considered the main employee. The survey was conducted in the metropolitan regions of Belém, Belo Horizonte, Brasília, Curitiba, Fortaleza, Goiânia, Porto Alegre, Recife, Rio de Janeiro, São Paulo and Salvador. The data set was extracted from IBGE database available at $http://www.ibge.gov.br/home/estatistica/indicadores/$. Table 2.6 presents a descriptive summary for the INPC data set and suggest skewed distribution with high degrees of skewness and kurtosis.

(i) Maximum Likelihood Estimation

Table 2.7 gives the MLEs and the corresponding SEs (given in parentheses) of the model parameters and the values of the following statistics for some models: AIC and BIC. The com-

Table 2.6 – Descriptive statistics for INPC data set.

Mean	Median	SD	Variance	Skewness	Kurtosis	Min.	Max.
0.64	0.50	0.60	0.36	1.56	6.59	-0.49	3.39

putations were performed using the statistical software R. The AIC and BIC values for the KBN model are the smallest values among those fitted sub-models and non-nested models.

Table 2.7 – MLEs and the corresponding SEs (given in parentheses) of the model parameters for the INPC data and the measures AIC and BIC

Model	μ	σ	a	b	c	AIC	BIC
KBN	0.4467	0.5573	4.3336	0.2712	9.4513	238.9	254.2
	(0.0547)	(0.0018)	(0.0248)	(0.0040)	(0.0042)		
BN	-0.4391	0.4686	5.3041	0.2905	0	256.0	268.3
	(0.1590)	(0.0028)	(0.0246)	(0.0431)	(-)		
Normal	0.6442	0.5988	1	1	0	288.5	294.6
	(0.0477)	(0.0337)	(-)	(-)	(-)		
KwN	μ	σ	a_1	b_1			
	-0.6987	0.5230	13.2245	0.2899	(-)	252.6	264.8
	(0.0117)	(0.0148)	(0.0205)	(0.0031)	(-)		
McN	μ	σ	a_1	b_1	c_1		
	-1.2530	0.5993	13.9336	0.2858	3.8102	251.1	266.4
	(0.0205)	(0.0178)	(0.0631)	(0.0307)	(0.0412)		
Skew-Normal	μ	σ	α	(-)	(-)		
	-0.0282	0.9005	4.3606	(-)	(-)	250.0	259.1
	(0.0480)	(0.0622)	(0.0970)	(-)	(-)		

A formal test of the need for the third skewness parameter in KB-G distributions is based on the LR statistics. Applying this to INPC data set, the results are shown in Table 2.8. We reject the null hypotheses of the LR test in favor of the KBN distribution. The rejection is extremely highly significant and it gives clear evidence of the potential need for three skewness parameters when modeling real data.

Table 2.8 – LR statistics for the INPC data

Model	Hypotheses	Statistic w	p-value
KBN vs BN	$H_0 : c = 0$ vs $H_1 : H_0$ is false	19.13	< 0.0001
KBN vs Normal	$H_0 : a = b = 1$ and $c = 0$ vs $H_1 : H_0$ is false	55.59	< 0.0001

Figures 2.6a and 2.6b display the histogram of the data and the fitted KBN density function and some densities of its sub-models and non-nested models, respectively. We note that the KBN distribution produces better fit than the other models. The QQ plots of the normalized quantile residuals in Figure 2.7 reveal the improvement in the fit achieved with the KBN distribution over the other distributions.

(a) (b)

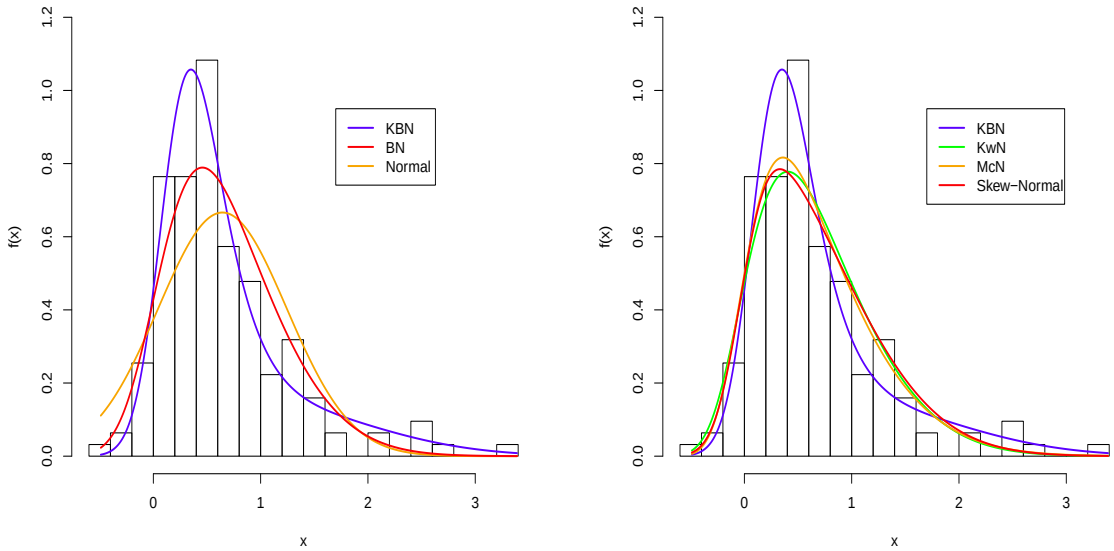


Figure 2.6 – (a) Estimated densities of the KBN and its sub-models (b) Estimated densities of the KBN, KwN, McN and skew-normal models

(ii) Bayesian Analysis

As an alternative analysis, we also use the Bayesian approach. In this context, we assume that the parameters (a, b, c, μ and σ) of the KBN distribution have independence priors, i.e.

$$\pi(a, b, c, \mu, \sigma) \propto \pi(a) \times \pi(b) \times \pi(c) \times \pi(\mu) \times \pi(\sigma). \tag{2.47}$$

Here, $a \sim \Gamma(a_1, b_1)$, $b \sim \Gamma(a_2, b_2)$, $c \sim \mathrm{N}(\mu_0, \sigma_0^2)$, $\mu \sim \mathrm{N}(\mu_1, \sigma_1^2)$ and $\sigma \sim \Gamma(a_4, b_4)$.

Replacing $G(x; \boldsymbol{\gamma})$ and $g(x; \boldsymbol{\gamma})$ to be the cdf and pdf of the normal distribution in equation (2.44), we have the joint posterior density function for a, b, c, μ and σ as

$$\begin{aligned}\pi(a, b, c, \mu, \sigma|x) \quad \propto \quad & \left(\frac{K}{\sigma}\right)^n \exp\left[-c\sum_{i=1}^{n}\Phi\left(\frac{x_i-\mu}{\sigma}\right) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2\right] \\ & \times \prod_{i=1}^{n}\left[\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{a-1}\left[1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{b-1} \\ & \times \pi(a, b, c, \mu, \sigma)\end{aligned} \tag{2.48}$$

and then, the full conditional distributions of each unknown quantity can be expressed as

$$\pi(a|x, b, c, \mu, \sigma) \quad \propto \quad K^n \prod_{i=1}^{n}\left[\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{a-1} \times \pi(a),$$

$$\pi(b|x, a, c, \mu, \sigma) \quad \propto \quad K^n \prod_{i=1}^{n}\left[1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{b-1} \times \pi(b),$$

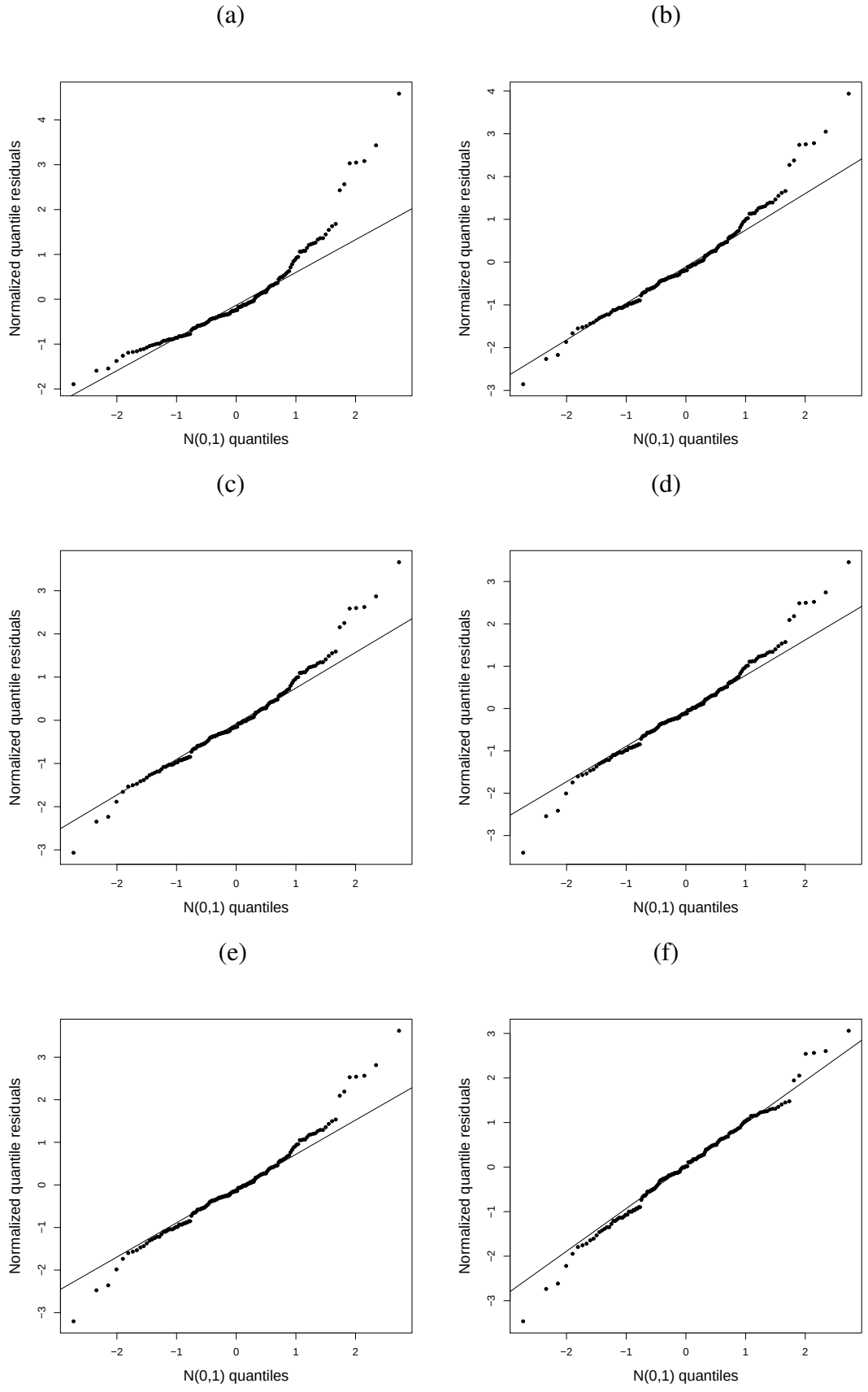


Figure 2.7 – QQ plot of the normalized quantile residuals with an identity line for the distributions: (a) Normal, (b) BN, (c) KwN, (d) McN, (e) Skew-Normal and (f) KBN

$$\pi(c|x,a,b,\mu,\sigma) \propto K^n \exp\left[-c\sum_{i=1}^{n}\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]\times\pi(c),$$

$$\pi(\mu|x,a,b,c,\sigma) \propto \exp\left[-c\sum_{i=1}^{n}\Phi\left(\frac{x_i-\mu}{\sigma}\right)-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2\right]$$
$$\times\prod_{i=1}^{n}\left[\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{a-1}\left[1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{b-1}\times\pi(\mu),$$

and

$$\pi(\sigma|x,a,b,c,\mu) \propto \frac{1}{\sigma^n}\exp\left[-c\sum_{i=1}^{n}\Phi\left(\frac{x_i-\mu}{\sigma}\right)-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2\right]$$
$$\times\prod_{i=1}^{n}\left[\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{a-1}\left[1-\Phi\left(\frac{x_i-\mu}{\sigma}\right)\right]^{b-1}\times\pi(\sigma).$$

We use the Metropolis-Hastings algorithm to generate variables a, b, c, μ and σ from the respective conditional posterior densities since their forms are somewhat complex. All MCMC computations were also implemented in statistical software R.

Now, we consider the following independent priors to perform the Metropolis-Hastings algorithm: $a \sim \Gamma(0.001, 0.001)$, $b \sim \Gamma(0.001, 0.001)$, $c \sim \mathrm{N}(0, 1000)$, $\mu \sim \mathrm{N}(0, 1000)$ and $\sigma \sim \Gamma(0.001, 0.001)$. Considering the density functions of the vague prior distributions above, we generated two parallel independent runs of the Metropolis-Hastings with the same characteristics of the Bayesian analysis performed in Section 2.7.1.

The plots of the histograms with the approximate posterior marginal density functions of the parameters are illustrated in Figure 2.8. We report, in Table 2.9 , the posterior summaries (posterior means, standard deviation (SD) and the 95% highest posterior density (HPD) intervals) for all parameters of the KBN distribution. We can note that the values for posterior means (Table 2.9) are in good agreement with the MLEs.

Table 2.9 – Posterior summaries for the parameters of the KBN model for the INPC data

Parameter	Mean	SD	HPD (95%)	$\widehat{R}$
a	4.3298	0.0098	(4.3107; 4.3493)	0.9999
b	0.2699	0.0102	(0.2505; 0.2898)	1.0007
c	9.4499	0.0102	(9.4301; 9.4696)	0.9998
μ	0.4401	0.0099	(0.4199; 0.4589)	1.0004
σ	0.5515	0.0101	(0.5307; 0.5699)	1.0010

2.8 Concluding Remarks

Following the idea of the class of beta generalized distributions and the distribution introduced by (NG; KOTZ, 1995), we define a new family of Kummer beta generalized (KB-G) distributions to extend several widely known distributions such as the normal, Weibull, gamma,

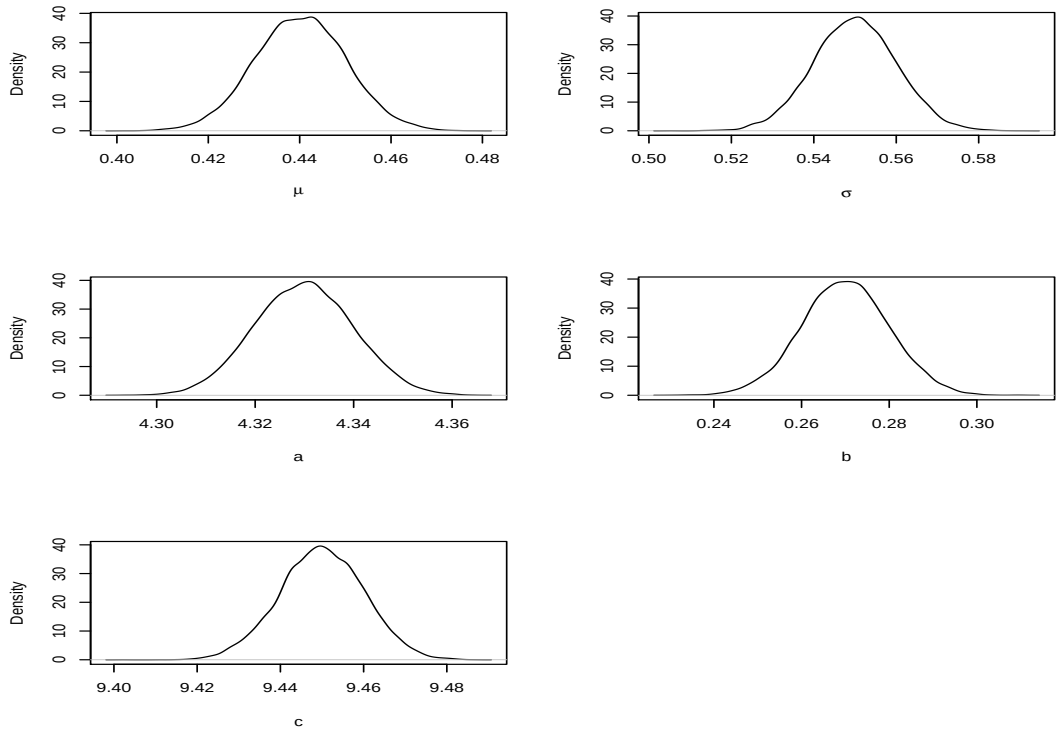


Figure 2.8 – Approximate posterior marginal densities for the parameters of the KBN model for the INPC data

Gumbel, Pareto and logistic distributions. For each continuous G distribution, we can define the corresponding KB-G distribution using simple formulae. Some mathematical properties of the KB-G distributions are readily obtained from those of the baseline G distributions. The moments of any KB-G distribution can be expressed explicitly in terms of infinite weighted sums of probability weighted moments (PWMs) of G distribution. The same happens for the moments of order statistics of the KB-G distributions. The estimation of the parameters is approached by two different methods: maximum likelihood and Bayesian approach. We consider likelihood ratio (LR) statistics and formal goodness-of-fit tests (AIC and BIC) to compare the KBGa and KBN models with some of their sub-models and non-nested models. Two applications to real data sets show the feasibility of the proposed class of models. We hope this generalization may attract wider applications in statistics.

References

ABRAMOWITZ, M.; STEGUN, I. A. **Handbook of Mathematical Functions**. New York: Dover Publications, 1968. 1046p.

AHSANULLAH, M.; NEVZOROV, V. B. **Order Statistics**: examples and exercises. New York: Nova Science Publishers, Hauppauge, 2005. 312p.

ALEXANDER, C.; CORDEIRO, G. M.; ORTEGA, E. M. M.; SARABIA, J. M. Generalized beta-generated distributions. **Computational Statistics and Data Analysis**, New York v.56, p.1880-1897, 2012.

ARNOLD, B. C.; BALAKRISHNAN, N.; NAGARAJA, H. N. **A First Course in Order Statistics**. New York: John Wiley and Sons, 1992. 205p.

ASCHER, H.; FEINGOLD, H. **Repairable systems reliability**. New York: Marcel Dekker, 1984. 300p.

AZZALINI, A. A class of distributions which includes the normal ones. **Scandinavian Journal of Statistics**, Copenhagen, p.31, p.171-178, 1985.

BALDING, D. J.; NICHOLS, R. A. A method for quantifying differentiation between populations at multi-allelic loci and its implications for investigating identity and paternity. **Genetica**, Amsterdam v.96, p.3-12, 1995.

BURY, K. **Statistical Distributions in Engineering**. New York: Marcel Dekker, 1999. 354p.

CORDEIRO, G. M.; CASTELLARES, F.; MONTENEGRO, L. C.; de Castro, M. The beta generalized gamma distribution. **Statistics**, Berlin, v.47, p.888-900, 2013.

CORDEIRO, G. M.; CINTRA, R. J.; RÊGO, L. C.; ORTEGA, E. M. M. The McDonald Normal Distribution. **Pakistan Journal of Statistics and Operation Research**, Islamabad, v.8, p.301-329, 2012.

CORDEIRO, G. M.; de CASTRO, M. A new family of generalized distributions. **Journal of Statistical Computation and Simulation**, London, v.81, p.883-893, 2011.

CORDEIRO, G. M.; LEMONTE, A. J. The beta Birnbaum-Saunders distribution: An improved distribution for fatigue life modeling. **Computational Statistics and Data Analysis**, New York, v.55, p.1445-1461, 2011.

CORDEIRO, G. M.; ORTEGA, E. M. M.; NADARAJAH, S. The Kumaraswamy Weibull distribution with application to failure data. **Journal of the Franklin Institute**, London, v.347, p.1399-1429, 2010.

COWLES, M. K.; CARLIN, B. P. Markov chain Monte Carlo convergence diagnostics: a comparative review. **Journal of the American Statistical Association**, New York, v.91, p.133-169, 1996.

DAVID, H. A.; NAGARAJA, H. N. **Order Statistics**, New Jersey: third edition, John Wiley and Sons, Hoboken, 2003. 201p.

DUNN, P. K.; SMYTH, G. K. Randomized quantile residuals, **Journal of Computational and Graphical Statistics**, New York, v.5, p.236-244, 1996.

EUGENE, N.; LEE, C.; FAMOYE, F. Beta-normal distribution and its applications. **Communication in Statistics: Theory and Methods**, New York, v.31, p.497-512, 2002.

EXTON, H. **Handbook of Hypergeometric Integrals: Theory, Applications, Tables, Computer Programs**. New York: Halsted Press, 1978. 200p.

GELMAN, A.; RUBIN, D.B. Inference from iterative simulation using multiple sequences (with discussion). **Statistical Science**, Washington, v.7, p.457-472, 1992.

GORDY, M. B. Computationally convenient distributional assumptions for common-value actions. **Computational Economics**, Amsterdam, v.12, p.61-78, 1998.

GRADSHTEYN, I. S.; RYZHIK, I. M. **Table of Integrals, Series, and Products**. San Diego: Academic Press, 2007. 315p.

GREENWOOD, J. A.; LANDWEHR, J. M.; MATALAS, N. C.; WALLIS, J. R. Probability weighted moments: Definition and relation to parameters of several distributions expressible in inverse form. **Water Resources Research**, Boston, v.15, p.1049-1054, 1979.

MUDHOLKAR, G. S; SRIVASTAVA, D. K.; FRIEMER, M. The exponential Weibull family: A reanalysis of the bus-motor failure data. **Technometrics**, Boston, v.37, p.436-445, 1995.

NADARAJAH, S. Explicit expressions for moments of order statistics. **Statistics and Probability Letters**, Amsterdam, v.78, p.196-205, 2008.

NADARAJAH, S.; GUPTA, A. K. The beta Fréchet distribution. **Far East Journal of Theoretical Statistics**, New Delhi, v.15, p.15-24, 2004.

NADARAJAH, S.; KOTZ, S. The beta Gumbel distribution. **Mathematical Problems in Engineering**, Boston, v.10, p.323-332, 2004.

NADARAJAH, S.; KOTZ, S. The beta exponential distribution. **Reliability Engineering and System Safety**, London, v.91, p.689-697, 2005.

NADARAJAH, S.; KOTZ, S. The exponentiated type distribution. **Acta Applicandae Mathematicae**, Amsterdam, v.92, p.97-111, 2006.

NADARAJAH, S.; KOTZ, S. Multitude of beta distributions with applications. **Statistics**, Berlin, v.41, p.153-179, 2007.

NAGAR, D. K.; GUPTA, A. K. Matrix-variate Kummer Beta distribution. **Journal of the Australian Mathematical Society**, Sidney, v.73, p.11-26, 2002.

NG, K. W.; KOTZ, S. Kummer-gamma and Kummer-beta univariate and multivariate distributions. **Research Report**, Hong Kong, v.84, p.120-141, 1995.

PARANAÍBA, P. F.; ORTEGA, E. M. M.; CORDEIRO, G. M.; PESCIM, R. R. The beta Burr XII distribution with application to lifetime data. **Computational Statistics and Data Analysis**, Amsterdam, v.55, p.1118-1136, 2011.

PASCOA, M. A. R.; ORTEGA, E. M. M.; CORDEIRO, G. M. The Kumaraswamy generalized gamma distribution with application in survival analysis. **Statistical Methodology**, Amsterdam, v.8, p.411-433, 2011.

PESCIM, R. R.; DEMÉTRIO, C. G. B.; CORDEIRO, G. M.; ORTEGA, E. M. M.; URBANO, M. R. The beta generalized half-normal distribution. **Computational Statistics and Data Analysis**, Amsterdam, v.55, p.945-957, 2010.

PRUDNIKOV, A. P.; BRYCHKOV, Y. A.; MARICHEV, O. I. **Integrals and Series**. Amsterdam: Gordon and Breach Science Publishers, 1986. v.1, 637p.

RIGBY, R. A.; STASINOPOULOS, M. D. Generalized additive models for location, scale and shape (with discussion). **Applied Statistics**, London, v.54, p.507-554, 2005.

STASINOPOULOS, D. M.; RIGBY, R. A. Generalized additive models for location, scale and shape (GAMLSS) in R. **Journal of Statistical Software**, Los Angeles, v.23, p.1-46, 2007.

WILEY, J. A.; HERSCHOKORU, S. J.; PADIAU, N. S. Heterogeneity in the probability of HIV transmission per sexual contact: the case of male-to-female transmission in penile vaginal intercourse. **Statistics in Medicine**, London, v.8, p.93-102, 1989.

3 THE KUMMER BETA BIRNBAUM-SAUNDERS: AN ALTERNATIVE FATIGUE LIFE DISTRIBUTION

Abstract

(BIRNBAUM; SAUNDERS, 1969a) defined a positive continuous distribution commonly used in reliability studies. Based on this probability distribution, we introduce the so-called Kummer beta Birnbaum-Saunders distribution for modeling fatigue life data. Various properties of the new model including explicit expressions for the ordinary and imcomplete moments, generating function, mean deviations, reliability, density function of the order statistics and their moments are derived. We investigate maximum likelihood estimation of the model parameters. The superiority of the new model is illustrated by means of one failure real data set.

Keywords: Birnbaum-Saunders distribution; Fatigue life distribution; Kummer beta distribution; Lifetime data; Maximum likelihood estimation

3.1 Introduction

Fatigue is a structural damage which occurs when a material is exposed to stress and tension fluctuations. When the effect of vibrations on material specimens and structures is studied, the first point to be considered is the mechanism that could cause fatigue to these materials. To understand the fatigue process and the genesis of the fatigue life and cumulative damage distributions, we recall concepts related to crack, cycle, fatigue and load.

In summary, the fatigue process (*fatigue life*) begins with an imperceptible fissure, the initiation, growth, and propagation of which produces a dominant crack in the specimen due to cyclic patterns of stress, whose ultimate extension causes the rupture or failure of this specimen. The failure occurs when the total extension of the crack exceeds a critical threshold for the first time. The partial extension of a crack produced by fatigue in each cycle is modeled by a random variable which depends on the type of material, the magnitude of the stress, and the number of previous cycles, among other factors. More details about the fatigue process can be revised, for example, in (VALLURI, 1963), (BIRNBAUM; SAUNDERS, 1969a), (MURTHY, 1974) and (SAUNDERS, 1976).

Motivated by problems of vibration in commercial aircraft that caused fatigue in the materials, (BIRNBAUM; SAUNDERS, 1969a, 1969b) proposed the two-parameter Birnbaum-Saunders (BS) distribution, also known as the fatigue life distribution, with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$, say BS(α, β). This distribution can be used to lifetime data and it is widely applicable to represent failure times of fatiguing materials. If Z is a standard normal random variable, the random variable X defined by

$$X = \beta \left\{ \frac{\alpha Z}{2} + \left[\left(\frac{\alpha Z}{2} \right)^2 + 1 \right]^{1/2} \right\}^2 ,$$

has a BS(α, β) distribution whose cumulative distribution function (cdf) is given by

$$G(x) = \Phi(\nu),\ x > 0, \tag{3.1}$$

where $\nu = \alpha^{-1}\rho(x/\beta)$, $\rho(z) = z^{1/2} - z^{-1/2}$ and $\Phi(.)$ is the standard normal cumulative function. The parameter β is the median of the distribution, i.e. $G(\beta) = \Phi(0) = 1/2$. For any $k > 0$, $kX \sim$ BS$(\alpha, k\beta)$. (KUNDU; KANNAN; BALAKRISHNAN, 2008) investigated the shape of the BS hazard function. Results on improved statistical inference for this model are discussed by (WU; WONG, 2004), (LEMONTE; CRIBARI-NETO; VASCONCELLOS, 2007) and (LEMONTE; SIMAS; CRIBARI-NETO, 2008). (DÍAZ-GARCIA; LEIVA, 2005) proposed a new family of generalized BS distributions based on contoured elliptical distributions, whereas (GUIRAUD; LEIVA; FIERRO, 2009) introduced a non-central version of the BS distribution. The probability density function (pdf) corresponding to (3.1) is

$$g(x) = r(\alpha, \beta)\, x^{-3/2}\, (x + \beta)\, \exp\left[-\frac{\tau(x/\beta)}{2\alpha^2}\right],\ x > 0, \tag{3.2}$$

where $r(\alpha, \beta) = \exp(\alpha^{-2})(2\alpha\sqrt{2\pi\beta})^{-1}$ and $\tau(z) = z - z^{-1}$. The fractional moments of (3.2) are (RIECK, 1999)

$$\mathrm{E}(\mathrm{X}^p) = \beta^p\ \mathrm{I}(p, \alpha), \tag{3.3}$$

where

$$\mathrm{I}(p, \alpha) = \frac{K_{p+1/2}(\alpha^{-2}) + K_{p-1/2}(\alpha^{-2})}{2K_{1/2}(\alpha^{-2})}, \tag{3.4}$$

and $K_p(z)$ denotes the modified Bessel function of the third kind with p representing its order and z the argument. Its integral representation is $K_p(z) = 0.5 \int_{-\infty}^{\infty} \exp[-z\cosh(t) - p\,t]dt$. A discussion of this function can be found in (WATSON, 1995).

The Kummer beta (KB) distribution may be characterized by the density function (NG; KOTZ, 1995)

$$f_{\mathcal{KB}}(x) = K\, x^{a-1}\, (1 - x)^{b-1}\, \mathrm{e}^{-\mathrm{ct}}, \quad 0 < \mathrm{x} < 1, \tag{3.5}$$

where $a > 0$, $b > 0$ and $-\infty < c < \infty$. Here,

$$K^{-1} = \frac{\Gamma(a)\,\Gamma(b)}{\Gamma(a + b)}\ {}_1F_1(a; a + b; -c)$$

and

$${}_1F_1(a; a + b; -c) = \frac{\Gamma(a + b)}{\Gamma(a\,\Gamma(b)} \int_0^1 t^{a-1}\, (1 - t)^{b-1} \mathrm{e}^{-\mathrm{ct}}\, \mathrm{dt} = \sum_{\mathrm{k}=0}^{\infty} \frac{(\mathrm{a})_\mathrm{k}\, (-\mathrm{c})^\mathrm{k}}{(\mathrm{a} + \mathrm{b})_\mathrm{k}\, \mathrm{k}!}$$

is the confluent hypergeometric function (ABRAMOWITZ; STEGUN, 1968), $\Gamma(\cdot)$ is the gamma function and $(d)_k = d(d+1)\ldots(d+k-1)$ denotes the ascending factorial. An important special model is the classical beta distribution when $c = 0$.

For an arbitrary continuous baseline distribution $G(x)$ with parameter vector $\boldsymbol{\gamma}$ and density function $g(x)$, the Kummer beta generalized (denoted by the prefix "KB-G" for short) cumulative function is defined by

$$F_{\mathcal{KBG}}(x) = K \int_0^{G(x)} t^{a-1} (1-t)^{b-1} \, \mathrm{e}^{-\mathrm{ct}} \, \mathrm{dt}, \tag{3.6}$$

where $a > 0$ and $b > 0$ are shape parameters which induce skewness, and thereby promote weight variation of the tails, whereas the parameter $-\infty < c < \infty$ "squeezes" the pdf to the left or right, i.e., it gives weights to the extremes of the density functions. For more details, see (PESCIM et al., 2012).

The density function corresponding to (3.6) can be expressed as

$$f_{\mathcal{KBG}}(x) = K \, g(x) \, G(x)^{a-1} \, [1 - G(x)]^{b-1} \, \exp\left[-c \, G(x)\right]. \tag{3.7}$$

Clearly, we obtain the classical beta distribution for $c = 0$. Equation (3.7) will be most tractable when both functions $G(x)$ and $g(x)$ have simple analytic expressions. Its major benefit is to offer more flexibility to extremes (right and/or left) of the density functions and therefore it becomes suitable for analyzing data with high degree of asymmetry.

In this work, we introduce a new five-parameter distribution called the Kummer beta Birnbaum-Saunders (KBBS) distribution which contains as sub-models the BS and beta Birnbaum-Saunders (BBS) (CORDEIRO; LEMONTE, 2011) distributions. The main motivation for this extension is that the new distribution is a highly flexible life distribution which admits different degrees of kurtosis and asymmetry. The KBBS distribution comes from (3.7) by taking $G(x)$ and $g(x)$ as the cdf and pdf of the BS(α,β) distribution, respectively. We also provide a comprehensive description of some of its mathematical properties with the hope that it will attract wider applications in reliability, engineering and in other areas of research.

This chapter is outlined as follows. In Section 3.2, we define the KBBS distribution and plot its density and hazard rate functions. Section 3.3 provides useful expansions for the density and cumulative distribution functions. We obtain explicit expressions for the moments (Section 3.4), generating function (Section 3.5), incomplete moments (Section 3.6), mean deviations and reliability (Section 3.7) and order statistics (Section 3.8). Some inferential tools are discussed in Section 3.9. An application presented in Section 3.10 reveal the usefulness of the new distribution for fatigue life data. Concluding remarks are addressed in Section 3.11.

3.2 A New Distribution for Reliability Studies

By taking the cdf (3.1) and pdf (3.2) of the BS distribution with shape parameter $\alpha > 0$ and scale parameter $\beta > 0$, the cdf and pdf of the KBBS distribution are obtained from equations (3.6) and (3.7) as $(x > 0)$

$$F(x) = K \int_0^{\Phi(\nu)} t^{a-1} (1-t)^{b-1} \, \mathrm{e}^{-\mathrm{ct}} \, \mathrm{dt} \tag{3.8}$$

and

$$f(x) = K\, r(\alpha,\beta)\, x^{-3/2}\, (x+\beta)\, \Phi(\nu)^{a-1}\, \left[1-\Phi(\nu)\right]^{b-1} \exp\left\{-\left[\frac{\tau(x/\beta)}{2\alpha^2} + c\,\Phi(\nu)\right]\right\}. \quad (3.9)$$

Hereafter, we denote by X the random variable following (3.9), say $X \sim$ KBBS(a,b,c,α,β). This density has four shape parameters a, b, c and α, which allow for a high degree of flexibility. The parameter c controls tail weights to the extremes of the distribution. The associated hazard rate function becomes

$$\lambda(x) = \frac{K\, r(\alpha,\beta)\, x^{-3/2}\, (x+\beta)\, \Phi(\nu)^{a-1}}{\left[1-F(x)\right]\, \left[1-\Phi(\nu)\right]^{-(b-1)}} \exp\left\{-\left[\frac{\tau(x/\beta)}{2\alpha^2} + c\,\Phi(\nu)\right]\right\}. \quad (3.10)$$

The study of the new distribution is important since it extends some distributions previously considered in the literature. In fact, the BS model (with parameters α and β) arises when $a = b = 1$ and $c = 0$, with a continuous crossover towards models with different shapes (e.g. a specified combination of skewness and kurtosis). The KBBS model contains as sub-models the BBS and the exponentiated Birnbaum-Saunders (EBS) (CORDEIRO; LEMONTE; ORTEGA, 2011) distributions when $c = 0$ and $b = 1$ in addition to $c = 0$, respectively. Plots of the KBBS density and hazard rate functions for selected parameter values are displayed in Figures 3.1 and 3.2, respectively. It is evident that the shapes of the new density function are much more flexible than the BS distribution. Further, it allows four major hazard shapes: increasing, decreasing, bathtub and unimodal failure rates.

3.3 Expansions for Cumulative and Density Functions

Expansions for equations (3.8) and (3.9) can be derived using the concept of exponentiated distributions. (CORDEIRO; LEMONTE; ORTEGA, 2011) defined a random variable Y following the EBS distribution with parameters α, β and $\gamma > 0$, say $Y \sim$ EBS(α,β,γ). The cdf and pdf of Y are denoted by $H(y;\alpha,\beta,\gamma) = \Phi(\nu)^{\gamma}$ and $h(y;\alpha,\beta,\gamma) = \gamma\, g_{\alpha,\beta}(y)\, \Phi(\nu)^{\gamma-1}$, respectively, where ν is defined in (3.1).

By expanding the term $\exp\left[-c\,\Phi(\nu)\right]$ in (3.9), we have

$$f(x) = \sum_{j=0}^{\infty} \frac{(-1)^j\, c^j}{j!\, K^{-1}}\, r(\alpha,\beta)\, x^{-3/2}\, (x+\beta)\, \exp\left[-\frac{\tau(x/\beta)}{2\alpha^2}\right] \Phi(\nu)^{a+j-1}\, \left[1-\Phi(\nu)\right]^{b-1},$$

and then, using the the binomial expansion for $\left[1-\Phi(\nu)\right]^{b-1}$, we obtain the pdf of the KBBS distribution as a linear combination (for $a > 0$ integer) of the EBS densities given by

$$f(x) = \sum_{j,k=0}^{\infty} w_{j,k}\, h(x;\alpha,\beta,a+j+k), \quad (3.11)$$

where $h(x;\alpha,\beta,a+j+k) = (a+j+k)g(x)\Phi(\nu)^{a+j+k-1}$ denotes the EBS$(\alpha,\beta,a+j+k)$ density function and the coefficient $w_{j,k}$ is given by

$$w_{j,k} = \frac{(-1)^{j+k}\, c^j K}{j!(a+j+k)} \binom{b-1}{k}.$$

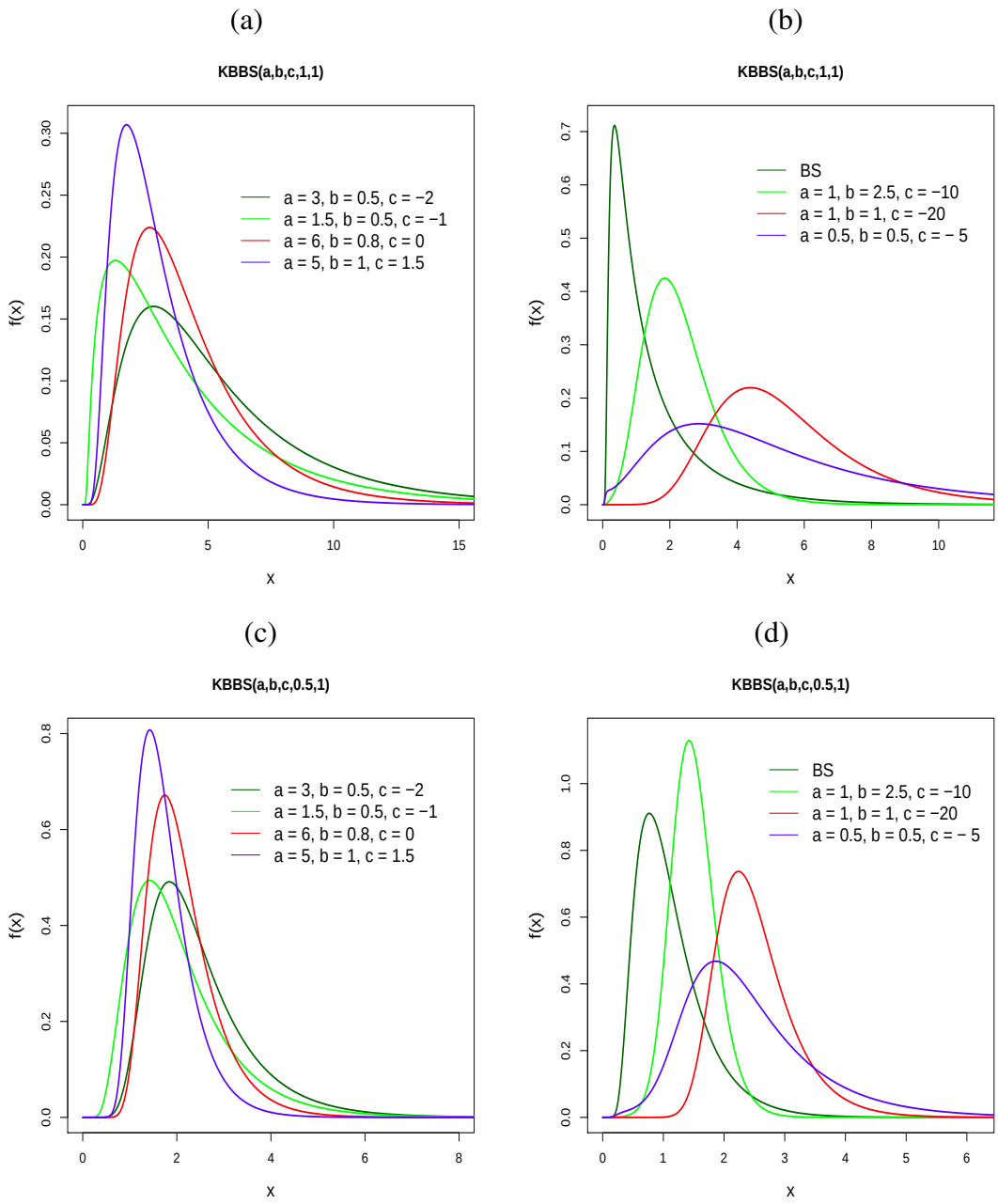


Figure 3.1 – Plots of the density function (3.9) for some parameter values

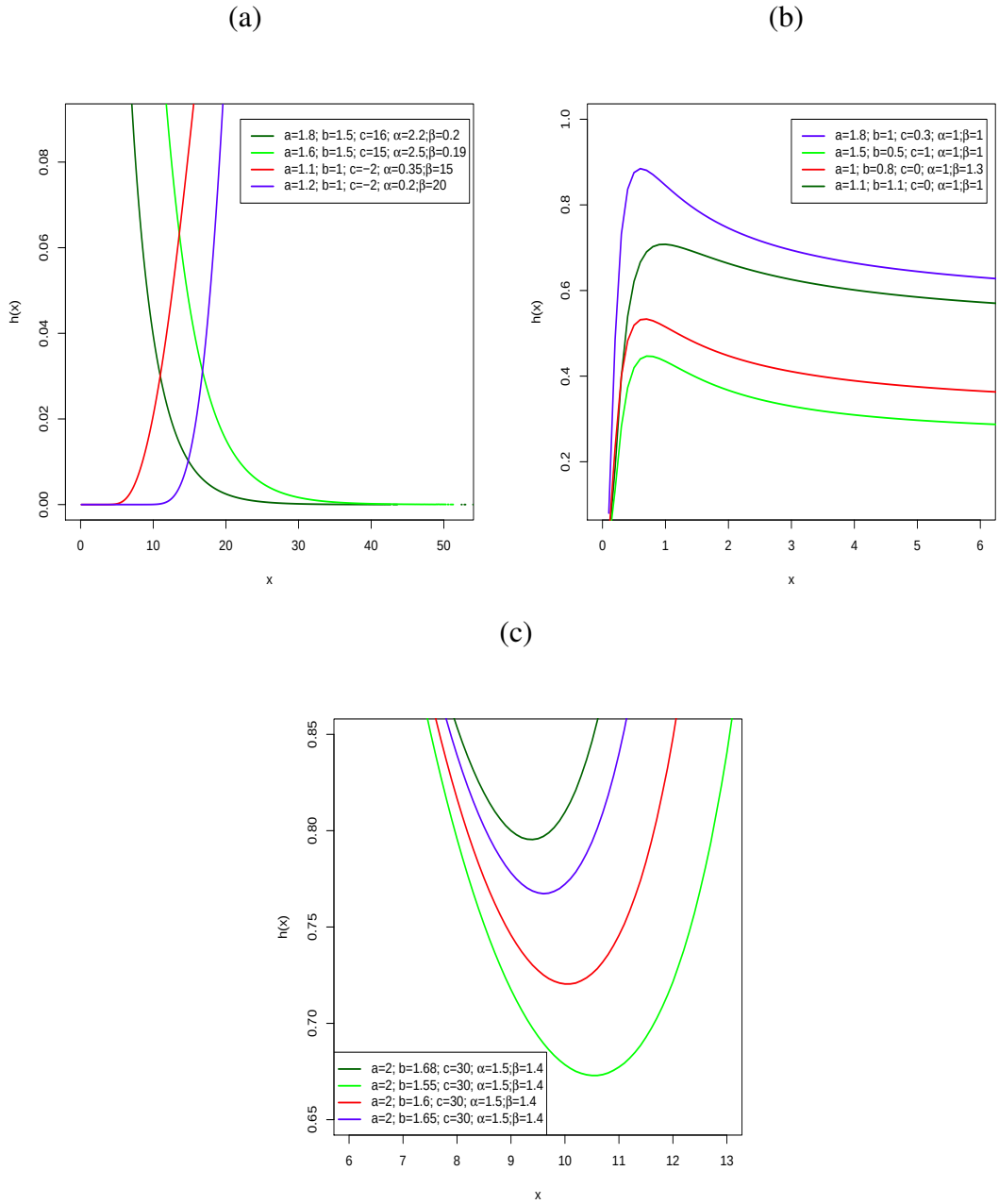


Figure 3.2 – The KBBS hazard rate function (a) Increasing and decreasing hazard rate function (b) Unimodal hazard rate function (c) Bathtub hazard rate function

By integrating (3.11), we obtain

$$F(x) = \sum_{j,k=0}^{\infty} w_{j,k}\, \Phi(\nu)^{a+j+k}. \tag{3.12}$$

If a is a positive non-integer, we can expand $\Phi(\nu)^{a+j+k}$ as

$$\Phi(\nu)^{a+j+k} = \{1 - [1 - \Phi(\nu)]\}^{a+j+k} = \sum_{p=0}^{\infty} (-1)^p \binom{a+j+k}{p} [1 - \Phi(\nu)]^p. \tag{3.13}$$

Using the binomial expansion in (3.13), we have

$$[1-\Phi(\nu)]^p = \sum_{r=0}^{p}(-1)^r \binom{p}{r} \Phi(\nu)^r,$$

and then,

$$\Phi(\nu)^{a+j+k} = \sum_{p=0}^{\infty}\sum_{r=0}^{p}(-1)^{p+r}\binom{a+j+k}{p}\binom{p}{r}\Phi(\nu)^r.$$

Replacing $\sum_{p=0}^{\infty}\sum_{r=0}^{p}$ by $\sum_{r=0}^{\infty}\sum_{p=r}^{\infty}$, we obtain

$$\Phi(\nu)^{a+j+k} = \sum_{r=0}^{\infty}\sum_{p=r}^{\infty}(-1)^{p+r}\binom{a+j+k}{p}\binom{p}{r}\Phi(\nu)^r$$

and

$$\Phi(\nu)^{a+j+k} = \sum_{r=0}^{\infty} s_r(a+j+k)\,\Phi(\nu)^r, \tag{3.14}$$

where the coefficients are

$$s_r(m) = \sum_{p=r}^{\infty}(-1)^{p+r}\binom{a+k+j}{p}\binom{p}{r}.$$

Thus, from equations (3.2), (3.12) and (3.14), the KBBS cumulative distribution can be expressed as

$$F(x) = \sum_{r=0}^{\infty} b_r\,\Phi(\nu)^r, \tag{3.15}$$

where $b_r = \sum_{j,k=0}^{\infty} w_{j,k}\, s_r(a+j+k)$.

For $a > 0$ real non-integer, the KBBS density function expansion corresponding to (3.15) is obtained by simple differentiation

$$f(x) = \sum_{r=0}^{\infty} b_r\, h(x;\alpha,\beta,r). \tag{3.16}$$

Equation (3.16) reveals that the KBBS density function is a linear combination of EBS density functions. This result is important to derive some properties of the KBBS distribution from those of the EBS distribution.

3.4 Moments

The ordinary moments of X can be determined from the probability weighted moments (GREENWOOD et al., 1979) of the BS distribution formally defined for p and r non-negative integers by

$$\tau_{p,r-1} = \int_0^{\infty} x^p\, g(x)\,\Phi(\nu)^{r-1}\,dx. \tag{3.17}$$

The integral (3.17) can be computed numerically in several software such as MAPLE, MATLAB, MATHEMATICA, Ox and R. (CORDEIRO; LEMONTE, 2011) proposed an alternative representation to compute $\tau_{p,r-1}$ given by

$$\begin{aligned}\tau_{p,r-1} &= \frac{\beta^p}{2^{r-1}} \sum_{j=0}^{r-1} \binom{r-1}{j} \sum_{k_1,\ldots,k_j}^{\infty} A(k_1,\ldots,k_j) \\ &\times \sum_{m=0}^{2s_j+j} (-1)^m \binom{2s_j+j}{m} I\left(p + \frac{(2s_j+j-2m)}{2}, \alpha\right), \end{aligned} \tag{3.18}$$

where $s_j = k_1 + \ldots + k_j$, $A(k_1,\ldots,k_j) = \alpha^{-2s_j-j}\, a_{k_1},\ldots,a_{k_j}$, $a_k = (-1)^k 2^{(1-2k)/2}[\sqrt{\pi}(2k+1)]^{-1}$ and $I(p+(2s_j+j-2m)/2,\alpha)$ is determined from (3.4).

The sth moment of X can be expressed from equation (3.16) as

$$\mu'_s = \sum_{r=0}^{\infty} b_r\, \tau_{s,r-1}, \tag{3.19}$$

where $\tau_{s,r-1}$ is obtained from (3.17) and b_r is defined in (3.15).

The four first moments of the KBBS distribution were calculated by numerical integration and through infinite weighted sums in equation (3.19) using the statistical software R. The values from both techniques are usually close when ∞ is replaced by a large number as 500 in (3.19). For selected values $a = 2$, $b = 1.5$, $c = 4$, $\alpha = 0.5$ and $\beta = 1$, Table (3.1) gives some numerical analysis for those moments and for variance, skewness and kurtosis.

Table 3.1 – Values of the four first moments, variance, skewness and kurtosis of the KBBS distribution for $a = 2$, $b = 1.5$, $c = 4$, $\alpha = 0.5$ and $\beta = 1$ obtained by numerical integration and through infinite weighted sums, where $j, k, r = 0, \ldots, n$

Moments	Infinite weighted sums				Numerical integration
	n=50	n=100	n=250	n=500	
μ'_1	0.85967	0.85920	0.85898	0.85893	0.85890
μ'_2	0.83508	0.83355	0.83278	0.83258	0.83242
μ'_3	0.93435	0.92920	0.92633	0.92550	0.92479
μ'_4	1.23327	1.21506	1.20395	1.20042	1.19703
Variance	0.09604	0.0953	0.09492	0.09481	0.09471
Skewness	1.72439	1.6716	1.63790	1.62691	1.61629
Kurtosis	9.18582	8.6644	8.28549	8.14676	7.99257

The skewness and kurtosis measures can be calculated from the ordinary moments using well-known relationships. Plots of the skewness and kurtosis of the KBBS distribution as a function of c for selected values of a and b for $\alpha = 0.5$ and $\beta = 1.0$ are displayed in Figures 3.3 and 3.4, respectively. Figures 3.3a and 3.3b immediately indicate that the additional parameter c promotes high levels of asymmetry.

(a) (b)

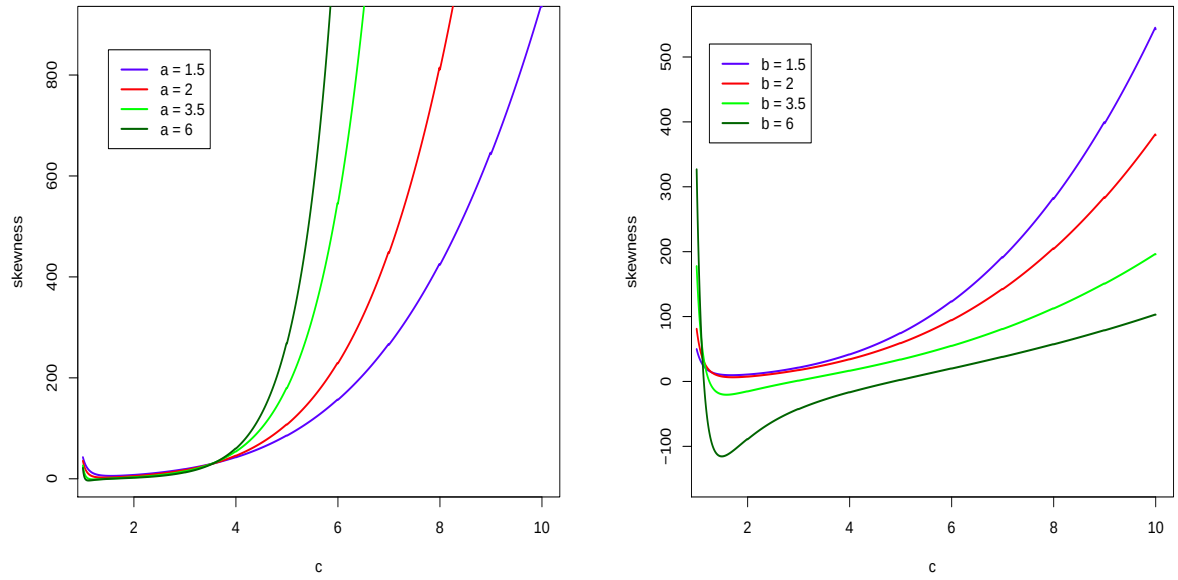


Figure 3.3 – Skewness of the KBBS distribution as a function of c for some values of a and b for $\alpha = 0.5$ and $\beta = 1.0$ (a) $b = 1.5$ and (b) $a = 1.2$

(a) (b)

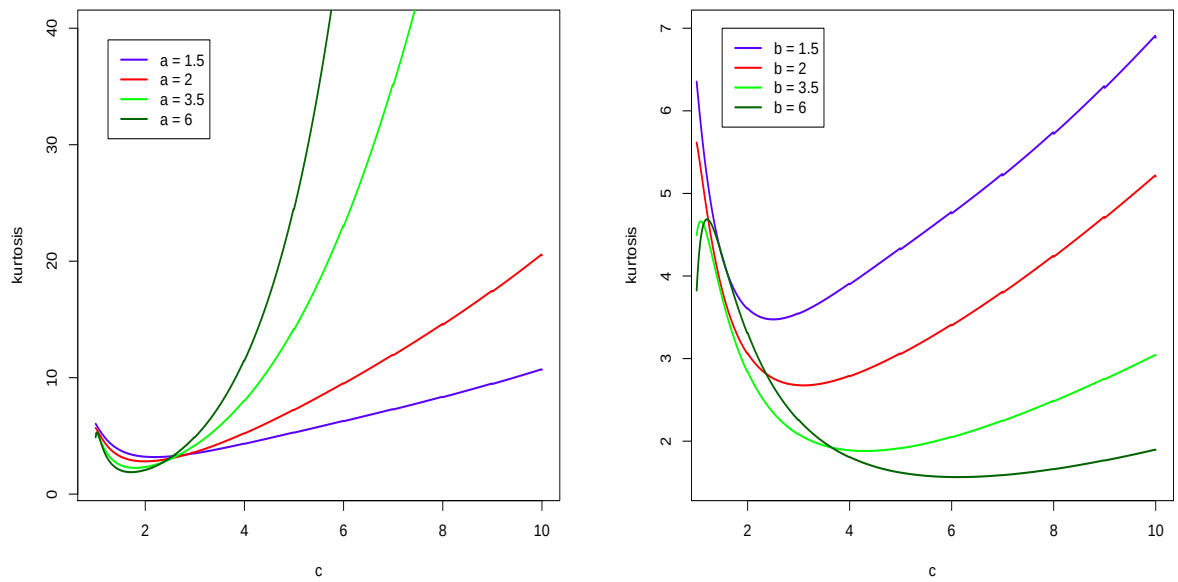


Figure 3.4 – Kurtosis of the KBBS distribution as a function of c for some values of a and b for $\alpha = 0.5$ and $\beta = 1.0$ (a) $b = 1.5$ and (b) $a = 1.2$

3.5 Generating function

In this section, we provide a representation for the moment generating function (mgf) of X, say $M(t) = \mathrm{E}[\exp(tX)]$, which is obtained as a linear combination of the mgf's of the EBS

distributions. From expansion (3.16), we obtain

$$M(t) = \sum_{r=0}^{\infty} b_r \, M_r(t), \tag{3.20}$$

where $M_r(t)$ is the mgf of the EBS(α, β, r) distribution and b_r is defined by (3.15).

Thus, $M_r(t)$ can be expressed as

$$M_r(t) = r \int_0^{\infty} \exp(tx) \, g(x) \, \Phi(\nu)^{r-1} \, dx, \tag{3.21}$$

where $g(x)$ is the BS(α, β) density function. Setting $u = \Phi(\nu)$ in (3.21), we have

$$M_r(t) = r \int_0^{1} u^{r-1} \, \exp[t \, Q_{\mathcal{BS}}(u)] \, du, \tag{3.22}$$

where $x = Q_{\mathcal{BS}}(u)$ is the quantile function of the BS distribution and $u = \Phi(\nu)$ is given by (3.1).

Now, we derive a power series expansion for the quantile function of the EBS distribution which can be useful to calculate the mgf of the KBBS distribution. We use throughout an equation in Section 0.314 of (GRADSHTEYN; RYZHIK, 2007) for a power series raised to a positive integer j given by

$$\left(\sum_{i=0}^{\infty} a_i \, x^i \right)^j = \sum_{i=0}^{\infty} c_{j,i} \, x^i, \tag{3.23}$$

where the coefficients $c_{j,i}$ (for $i = 1, 2, \ldots$) are computed from the recurrence equation

$$c_{j,i} = (i a_0)^{-1} \sum_{m=1}^{i} [m(j+1) - i] \, a_m \, c_{j,i-m} \tag{3.24}$$

and $c_{j,0} = a_0^j$. The coefficient $c_{j,i}$ can be determined from $c_{j,0}, \ldots, c_{j,i-1}$ and hence from the quantities $a_0, \ldots, a_i$. In fact, $c_{j,i}$ can be given explicitly in terms of the coefficients a_i, although it is not necessary for programming numerically our expansions in any algebraic or numerical software.

Following (CORDEIRO; LEMONTE, 2011), we can invert $u = \Phi(\nu)$ if the condition $-2 < (x/\beta)^{1/2} - (\beta/x)^{1/2} < 2$ holds, to express x as a power series expansion of u

$$x = Q_{\mathcal{BS}}(u) = \sum_{q=0}^{\infty} \rho_q \, \nu^q, \tag{3.25}$$

where the coefficients are $\rho_0 = \beta$, $\rho_{2q+1} = \beta \alpha^{2q+1} \binom{1/2}{q} 2^{-2q}$ for $q \geq 0$, $\rho_2 = \beta \alpha^2/2$ and $\rho_{2q} = 0$ for $q \geq 2$. From (3.23), we can write (3.25) as

$$x = Q_{\mathcal{BS}}(u) = \sum_{q=0}^{\infty} \rho_q \, \nu^q = \sum_{q=0}^{\infty} \rho_q \left(\sum_{i=0}^{\infty} d_i \, y^i \right)^q = \sum_{i=0}^{\infty} z_i \, (u - 1/2)^i, \tag{3.26}$$

where $z_i = (2\pi)^{1/2} \sum_{q=0}^{\infty} \rho_q \, e_{q,i}$ and the quantities $e_{q,i}$ follow recursively from equation (3.24) by $e_{q,0} = d_0^q$ and

$$e_{q,i} = (id_0)^{-1} \sum_{m=1}^{q} [m(q+1) - i] \, d_m \, e_{q,i-m}.$$

Here, the quantities d_m are defined by $d_m = 0$ (for $m = 0, 2, 4, \ldots$) and $d_m = j_{(m-1)/2}$ (for $m = 1, 3, 5, \ldots$), where the j_m's are calculated recursively from

$$j_{m+1} = \frac{1}{2(2m+3)} \sum_{v=0}^{m} \frac{(2v+1)\,(2m-2v+1)\, j_v \, j_{m-v}}{(v+1)\,(2v+1)}.$$

We have $j_0 = 1$, $j_1 = 1/6$, $j_2 = 7/120$, $j_3 = 127/7560, \ldots$.

Replacing equation (3.26) in (3.22) and using the exponential expansion, we obtain

$$M_r(t) = \sum_{p=0}^{\infty} \frac{r\, t^p}{p!} \int_0^1 u^{r-1} \left(\sum_{i=0}^{\infty} z_i \, w^i \right)^p du. \tag{3.27}$$

where $w = u - 1/2$. From equations (3.23) and (3.24), we have

$$\left(\sum_{i=0}^{\infty} z_i \, w^i \right)^p = \sum_{i=0}^{\infty} \delta_{p,i} \, w^i = \sum_{i=0}^{\infty} \delta_{p,i} \, (u - 1/2)^i,$$

where $\delta_{p,0} = \rho_0^p$ and

$$\delta_{p,i} = (iz_0)^{-1} \sum_{m=1}^{i} [m(p+1) - i] \, z_m \, \delta_{p,i-m},$$

and then, equation (3.27) becomes

$$M_r(t) = \sum_{p,i=0}^{\infty} \frac{r\, t^p}{p!} \, \delta_{p,i} \int_0^1 u^{r-1} \, (u - 1/2)^i \, du. \tag{3.28}$$

Using the binomial expansion in (3.28), the mgf of the EBS distribution can be expressed as

$$M_r(t) = \sum_{p=0}^{\infty} \delta_{p,r}^* \, t^p, \tag{3.29}$$

where

$$\delta_{p,r}^* = \sum_{i=0}^{\infty} \sum_{q=0}^{i} \binom{i}{q} \frac{(-1)^{i-q} \, r \, \delta_{p,i}}{p! \, (q+r) \, 2^{i-q}}.$$

Finally, substituting (3.29) into (3.20), the mgf of the KBBS distribution reduces to

$$M(t) = \sum_{p=0}^{\infty} \eta_p \, t^p, \tag{3.30}$$

where $\eta_p = \sum_{r=0}^{\infty} b_r \, \delta_{p,r}^*$.

3.6 Incomplete Moments

Many important questions in econometrics require more than just knowing the mean of a distribution, but its shape as well. This is also obvious not only in the study of econometrics and income distributions but in many other areas of research. For empirical purposes, the shape of many distributions can be usefully described by what we call the incomplete moments. These types of moments play an important role for measuring inequality, for example, income quantiles and Lorenz and Bonferroni curves, which depend upon the incomplete moments of a distribution. The nth incomplete moment of X is given by

$$T_n(y) = \int_0^y x^n \, f(x)dx.$$

By inserting (3.16) in $T_n(y)$, we obtain

$$T_n(y) = r(\alpha,\beta)\sum_{r=0}^{\infty} b_r \int_0^y x^{n-3/2}\,(x+\beta)\,\exp\left[-\frac{\tau(x/\beta)}{2\alpha^2}\right]\,\Phi(\nu)^{r-1}dx.$$

From (CORDEIRO; LEMONTE, 2011), we have

$$\Phi(\nu)^{r-1} \;=\; \frac{1}{2^{r-1}}\sum_{j=0}^{r-1}\binom{r-1}{j}\sum_{k_1,\ldots,k_j=0}^{\infty}\frac{A(k_1,\ldots,k_j)}{\beta^{(2s_j+j)/2}}\sum_{m=0}^{2s_j+j}(-\beta)^m\binom{2s_j+j}{m}x^{(2s_j+j-2m)/2},$$

where s_j and $A(k_1,\ldots,k_j)$ are defined in (3.18). Thus,

$$\begin{aligned} T_n(y) \;&=\; r(\alpha,\beta)\sum_{r=0}^{\infty}\frac{b_r}{2^{r-1}}\sum_{j=0}^{r-1}\binom{r-1}{j} \\ &\times \sum_{k_1,\ldots,k_j=0}^{\infty}\beta^{-(2s_j+j)/2}\,A(k_1,\ldots,k_j)\sum_{m=0}^{2s_j+j}(-\beta)^m\binom{2s_j+j}{m} \\ &\times \int_0^y x^{n+(2s_j+j-2m-3)/2}\,(x+\beta)\,\exp\left[-\frac{\tau(x/\beta)}{2\alpha^2}\right]dx. \end{aligned} \tag{3.31}$$

Let

$$D(p,q) = \int_0^q x^q\,\exp\left[-\frac{(x/\beta+\beta/x)}{2\alpha^2}\right]dx = \beta^{p+1}\int_0^{q/\beta} u^q\,\exp\left[-\frac{(u+u^{-1})}{2\alpha^2}\right]du.$$

From (TERRAS, 1981), we can write the integral in (3.31) as

$$D(p,q) = \beta^{p+1}\kappa_{p+1}(\alpha^{-2}) - q^{p+1}K_{p+1}\left(\frac{q}{2\alpha^2\beta},\frac{\beta}{2\alpha^2 q}\right),$$

where $K_p(x_1,x_2)$ denotes the incomplete Bessel function with arguments x_1 and x_2 and order p. For further details, see (JONES, 2007a, 2007b) and (HARRIS, 2008).

Hence, the nth incomplete moment of X can be expressed as

$$\begin{aligned} T_n(y) \;&=\; r(\alpha,\beta)\sum_{r=0}^{\infty}\frac{b_r}{2^{r-1}}\sum_{j=0}^{r-1}\binom{r-1}{j}\sum_{k_1,\ldots,k_j=0}^{\infty}\beta^{-(2s_j+j)/2}\,A(k_1,\ldots,k_j)\sum_{m=0}^{2s_j+j}(-\beta)^m\binom{2s_j+j}{m} \\ &\times\left[D\left(n+\frac{2s_j+j-2m-1}{2},y\right)+\beta\,D\left(n+\frac{2s_j+j-2m-3}{2},y\right)\right]. \end{aligned} \tag{3.32}$$

Equation (3.32) is the main result of this section.

3.7 Other Measures

In this section, we calculate the following measures: mean deviations and the reliability for the KBBS distribution.

3.7.1 Mean Deviations

We can derive the mean deviations about the mean μ_1' (δ_1) and about the median M (δ_2) in terms of the first incomplete moment. The median is obtained by inverting $F(M) = K \int_0^{\Phi(\nu)} t^{a-1}\,(1-t)^{b-1}\,\mathrm{e}^{-\mathrm{c}t}\,\mathrm{d}t = 1/2$ numerically. They can be expressed as

$$\delta_1 = 2\big[\mu_1'\,F(\mu_1') - T_1(\mu_1')\big] \qquad \text{and} \qquad \delta_2 = \mu_1' + 2M\,F(M) - M - 2\,T_1(M),$$

where $T_1(\cdot)$ is the first incomplete moment of X given by (3.32) with $n = 1$. We have

$$\begin{aligned} T_1(\omega) \;=\; & r(\alpha,\beta)\sum_{r=0}^{\infty}\frac{b_r}{2^{r-1}}\sum_{j=0}^{r-1}\binom{r-1}{j}\sum_{k_1,\ldots,k_j=0}^{\infty}\beta^{-(2s_j+j)/2}\,A(k_1,\ldots,k_j)\sum_{m=0}^{2s_j+j}(-\beta)^m\binom{2s_j+j}{m} \\ & \times\left[D\left(\frac{2s_j+j-2m+1}{2},\omega\right)+\beta\,D\left(\frac{2s_j+j-2m-1}{2},\omega\right)\right]. \end{aligned} \tag{3.33}$$

The measures δ_1 and δ_2 are immediately calculated from (3.33) by setting $\omega = \mu_1'$ and $\omega = M$, respectively.

3.7.2 Reliability

In the context of reliability, the stress-strength model describes the life of a component that has a random strength X_1 that is subjected to a random stress X_2. The component fails at the instant that the stress applied to it exceeds the strength, and the component will function satisfactorily whenever $X_1 > X_2$. Hence, $R = Pr(X_1 < X_2)$ is a measure of component reliability which has many applications especially in engineering area (structures, deteriorating of rocket motors and fatigue failure of aircraft structures). According to (PARANAÍBA et al., 2011), in the area of stress-strength models there has been a large amount of work as regards estimation of the reliability R when X_1 and X_2 are independent random variables belonging to the same univariate family of distributions. We derive the reliability R when X_1 and X_2 have independent KBBS$(\alpha,\beta,a_1,b_1,c_1)$ and KBBS$(\alpha,\beta,a_2,b_2,c_2)$ distributions with the same parameters α and β.

The pdf of X_1 and the cdf of X_2 can be written from equations (3.11) and (3.12) as

$$f_1(x) = g(x)\sum_{i,j=0}^{\infty} w_{1i,j}\,(a_1+i+j)\Phi(\nu)^{a_1+i+j} \quad \text{and} \quad F_2(x) = \sum_{k,p=0}^{\infty} w_{2k,p}\,\Phi(\nu)^{a_2+k+p},$$

respectively, where

$$w_{1i,j} = \frac{(-1)^{i+j}\,K_1\,c_1^i}{i!(a_1+i+j)}\binom{b_1-1}{j} \quad \text{and} \quad w_{2k,p} = \frac{(-1)^{k+p}\,K_2\,c_2^k}{k!(a_2+k+p)}\binom{b_2-1}{p}.$$

The reliability, R, is defined by

$$R = \int_0^\infty f_1(x) F_2(x) dx$$

and then

$$R = \sum_{i,j,k,p=0}^{\infty} w_{1i,j}\, w_{2k,p} \int_0^\infty g(x)\, \Phi(\nu)^{a_1+a_2+i+j+k+p-1}\, dx.$$

From equation (3.14), we can write

$$\Phi(\nu)^{a_1+a_2+i+j+k+p-1} = \sum_{r=0}^{\infty} s_r(a_1+a_2+i+j+k+p-1)\, \Phi(\nu)^r,$$

and then R reduces to

$$R = \sum_{i,j,k,p=0}^{\infty} w_{1i,j}\, w_{2k,p} \sum_{r=0}^{\infty} s_r(a_1+a_2+i+j+k+p-1)\, \tau_{0,r-1},$$

where $\tau_{0,r-1}$ can be computed from (3.18).

3.8 Order statistics

Suppose $X_1, \ldots, X_n$ is a random sample from the KBBS distribution and let $X_{1:n} < \cdots < X_{i:n}$ denote the corresponding order statistics. It is well-known that

$$f_{i:n}(x) = \frac{n!\, f(x)}{(i-1)!\,(n-1)!}\, F(x)^{i-1}\, [1-F(x)]^{n-i}\,. \tag{3.34}$$

Using the binomial expansion in (3.34), we have

$$f_{i:n}(x) = \frac{n!\, f(x)}{(i-1)!\,(n-1)!} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F(x)^{i+j-1}. \tag{3.35}$$

Now, using (3.15) and (3.16) in (3.35), the pdf of $X_{i:n}$ can be expressed as

$$f_{i:n}(x) \;=\; \frac{n!}{(i-1)!(n-i)!} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} \left[g(x) \sum_{r=0}^{\infty} b_r\, \Phi(\nu)^r \right] \left[\sum_{k=0}^{\infty} d_r\, \Phi(\nu)^r \right]^{i+j-1}.$$

From equations (3.23) and (3.24), we obtain

$$\left[\sum_{r=0}^{\infty} b_r\, \Phi(\nu)^r \right]^{i+j-1} = \sum_{r=0}^{\infty} c_{i+j-1,r}\, \Phi(\nu)^r,$$

where $c_{i+j-1,0} = b_0^{i+j-1}$ and

$$c_{i+j-1,r} = (r b_0)^{-1} \sum_{m=1}^{r} [m\,(i+j) - r]\, b_m\, c_{i+j-1,r-m}.$$

Hence, the pdf of the ith order statistic for the KBBS distribution can be expressed as

$$f_{i:n}(x) \;=\; \sum_{r=0}^{\infty} m_r\, h(x;\alpha,\beta,2r), \tag{3.36}$$

where

$$m_r = \frac{n!\, b_r}{(i-1)!\,(2r+1)\,(n-i)!} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} c_{i+j-1,r}.$$

Equation (3.36) is the main result of this section. It gives the pdf of the KBBS order statistics as a linear combination of EBS densities with parameters α, β and $2r$. So, several mathematical quantities of the KBBS order statistics such as ordinary and incomplete moments, generating function, mean deviations (and several others) can come immediately from those quantities of the EBS distribution.

3.9 Inference

The estimation of the model parameters of the KBBS distribution will be investigated by the maximum likelihood method. Let $X = (X_1, \ldots, X_n)$ be a random sample of this distribution with unknown parameter vector $\boldsymbol{\theta} = (\alpha, \beta, a, b, c)^T$. The log-likelihood function for $\boldsymbol{\theta}$ is

$$\begin{aligned}\ell(\boldsymbol{\theta}) \;=\; & n\log(K) + n\log\left[r(\alpha,\beta)\right] - \frac{3}{2}\sum_{i=1}^{n}\log(x_i) + \sum_{i=1}^{n}\log(x_i+\beta) \;-\frac{1}{2\alpha^2}\sum_{i=1}^{n}\tau(x_i/\beta) \\ & - c\sum_{i=1}^{n}\Phi(\nu_i) + (a-1)\sum_{i=1}^{n}\log[\Phi(\nu_i)] + (b-1)\sum_{i=1}^{n}\log\left[1-\Phi(\nu_i)\right]. \end{aligned} \tag{3.37}$$

The elements of score vector are given by

$$\begin{aligned} U_\alpha(\boldsymbol{\theta}) \;=\; & -\frac{n}{\alpha}\left(1+\frac{2}{\alpha^2}\right) + \frac{1}{\alpha^3}\sum_{i=1}^{n}\left(\frac{x_i}{\beta}+\frac{\beta}{x_i}\right) \\ & -\frac{1}{\alpha}\sum_{i=1}^{n}\nu_i\,\phi(\nu_i)\left[\frac{(a-1)}{\Phi(\nu_i)} - \frac{(b-1)}{1-\Phi(\nu_i)} - 2\,c\right], \end{aligned}$$

$$\begin{aligned} U_\beta(\boldsymbol{\theta}) \;=\; & -\frac{n}{2\beta} + \sum_{i=1}^{n}\frac{1}{x_i+\beta} + \frac{1}{2\alpha^2\beta}\sum_{i=1}^{n}\left(\frac{x_i}{\beta}-\frac{\beta}{x_i}\right) \\ & -\frac{1}{2\alpha\beta}\sum_{i=1}^{n}\tau(\sqrt{x_i/\beta})\;\phi(\nu_i)\left[\frac{(a-1)}{\Phi(\nu_i)} - \frac{(b-1)}{1-\Phi(\nu_i)} - c\right], \end{aligned}$$

$$U_a(\boldsymbol{\theta}) \;=\; \frac{n}{K}\frac{\partial K}{\partial a} + \sum_{i=1}^{n}\log[\Phi(\nu_i)],\;\; U_b(\boldsymbol{\theta}) = \frac{n}{K}\frac{\partial K}{\partial b} + \sum_{i=1}^{n}\log[1-\Phi(\nu_i)]$$

and

$$U_c(\boldsymbol{\theta}) = \frac{n}{K}\frac{\partial K}{\partial c} + \sum_{i=1}^{n}\Phi(\nu_i)$$

where $\phi(\cdot)$ is the standard normal density, $\nu_i = \alpha^{-1}[(x_i/\beta)^{1/2} - (x_i/\beta)^{-1/2}]$ and $\tau(\sqrt{x_i/\beta}) = (x_i/\beta)^{1/2} + (\beta/x_i)^{1/2}$ for $i = 1, \ldots, n$. The partial derivatives of K in relation to a, b and c are given by

$$\frac{\partial K}{\partial a} = -\frac{\left\{[\psi(a) - \psi(a+b)]\ {}_1F_1(a, a+b, -c) + \frac{\partial\, {}_1F_1(a,a+b,-c)}{\partial a}\right\}}{B(a,b)\left[{}_1F_1(a, a+b, -c)\right]^2},$$

$$\frac{\partial K}{\partial b} = -\frac{\left\{[\psi(b) - \psi(a+b)]\ {}_1F_1(a, a+b, -c) + \frac{\partial\, {}_1F_1(a,a+b,-c)}{\partial b}\right\}}{B(a,b)\left[{}_1F_1(a, a+b, -c)\right]^2},$$

$$\frac{\partial K}{\partial c} = \frac{a\, {}_1F_1(a+1, a+b+1, -c)}{(a+b)B(a,b){}_1F_1(a, a+b, -c)}, \text{ where}$$

$$\begin{aligned}\frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial a} &= -[\psi(a) - \psi(a+b)]\ {}_1F_1(a, a+b, -c) \\ &\quad - \sum_{k=0}^{\infty} \frac{(a)_k(-c)^k}{k!(a+b)_k}\,[\psi(a+b+k) - \psi(a+k)]\end{aligned}$$

and

$$\frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial b} = \psi(a+b)\, {}_1F_1(a, a+b, -c) + \sum_{k=0}^{\infty} \frac{(a)_k(-c)^k}{k!(a+b)_k}\,\psi(a+b+k).$$

Maximization of (3.37) can be performed by using well established routines such as the nlm routine or optim in statistical software R. Setting these equations to zero, $\mathbf{U}(\boldsymbol{\theta}) = \mathbf{0}$, and solving them simultaneously yields the maximum likelihood estimate (MLE) $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$. These equations cannot be solved analytically and statistical software can be used to solve them numerically by means of iterative techniques such as the Newton-Raphson algorithm.

For interval estimation and test of hypothesis on the parameters in $\boldsymbol{\theta}$, we require the 5×5 total observed information matrix $\mathbf{J}(\boldsymbol{\theta}) = -\{U_{rs}\}$, where the elements U_{rs} for $r, s = \alpha, \beta, a, b, c$ are given in Appendix B. The estimated asymptotic multivariate normal $\mathrm{N}_5(0, \mathbf{J}(\widehat{\boldsymbol{\theta}})^{-1})$ distribution of $\widehat{\boldsymbol{\theta}}$ can be used to construct approximate confidence regions for the parameters and for the hazard rate and survival functions. An asymptotic confidence interval with significance level γ for each parameter θ_r is given by

$$\mathrm{ACI}(\theta_r, 100(1-\gamma)\%) = (\hat{\theta}_r - z_{\gamma/2}\sqrt{\hat{\kappa}^{\theta_r,\theta_r}}, \hat{\theta}_r + z_{\gamma/2}\sqrt{\hat{\kappa}^{\theta_r,\theta_r}}),$$

where $\hat{\kappa}^{\theta_r,\theta_r}$ is the rth diagonal element of $\mathbf{J}(\boldsymbol{\theta})^{-1}$ estimated at $\widehat{\boldsymbol{\theta}}$, for $r = 1, \ldots, 4$, and $z_{\gamma/2}$ is the quantile $1 - \gamma/2$ of the standard normal distribution.

The likelihood ratio (LR) statistic is useful for comparing the new distribution with some of its special models. For example, we may adopt the LR statistic to check if the fit using the KBBS distribution is statistically "superior" to a fit using the BS distribution for a given data set. In any

case, considering the partition $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T)^T$, tests of hypotheses of the type $H_0 : \boldsymbol{\theta}_1 = \boldsymbol{\theta}_1^{(0)}$ versus $H_A : \boldsymbol{\theta}_1 \neq \boldsymbol{\theta}_1^{(0)}$ can be performed using the LR statistic $w = 2[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widetilde{\boldsymbol{\theta}})]$, where $\widehat{\boldsymbol{\theta}}$ and $\widetilde{\boldsymbol{\theta}}$ are the estimates of $\boldsymbol{\theta}$ under H_A and H_0, respectively. Under the null hypothesis H_0, $w \xrightarrow{d} \chi_q^2$, where q is the dimension of the vector $\boldsymbol{\theta}_1$ of interest. The LR test rejects H_0 if $w > \xi_\gamma$, where ξ_γ denotes the upper $100\gamma\%$ point of the χ_q^2 distribution.

3.10 Application

In this section, we use a real data set to compare the fits of the KBBS distribution with those of two sub-models (i.e. the BBS and BS distributions) and also to the following non-nested models: the McDonald-Birnbaum-Saunders (McBS) (CORDEIRO; LEMONTE; ORTEGA, 2011) and McDonald-gamma (McGa) (MARCIANO et al., 2012) distributions. All the computations were performed using the statistical software R. Obviously, due to the genesis of the BS and gamma distributions, the fatigue processes are by excellence ideally modeled by these distributions. Thus, the use of the KBBS distribution and its special models and also other lifetime distributions for fitting to the current data set is justified.

3.10.1 Breaking stress of carbon fibres data set

Here, we shall compare the fitted KBBS, BBS, BS, McBS and McGa distributions to the data from (NICHOLS; PADGETT, 2006) on the breaking stress of carbon fibres (in Gba). They described the data from a process which produces carbon fibers to be used in constructing fibrous composite materials. The carbon fiber 50 mm in length were sampled (n=66) from the process, tested and their tensile strength were observed.

Firstly, in order to estimate the model parameters, we consider the maximum likelihood estimation method discussed in Section 3.9. We take the estimates of α and β from the fitted BS distribution as starting values for the numerical iterative procedure. All computations were performed using the statistical software R. Table 3.2 lists the MLEs and the corresponding SEs of the parameters and the values of the following statistics for some models: Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC) and Bayesian Information Criterion (BIC). The results indicate that the KBBS model has the smallest values of these statistics among all fitted models. So, it could be chosen as the more suitable model.

A comparison of the proposed distribution with some of its sub-models using LR statistics is given in Table 3.3. We reject the null hypotheses of the two LR tests in favor of the KBBS distribution. This gives a clear evidence of the potential of the three shape parameters when modeling real data.

In order to assess if the model is appropriate, Figures 3.5a and 3.5b display the histogram of the data and the fitted KBBS density function and some densities of its sub-models and non-nested models, respectively. Further, Figures 3.5c and 3.5d display plots of the empirical and estimated survival functions of the KBBS distribution and of some sub-models and non-nested

Table 3.2 – MLEs, the corresponding SEs (given in parentheses) and information criteria for breaking stress of carbon fibres data

Model	α	β	a	b	c	AIC	BIC	CAIC
KBBS	0.6770	2.9944	0.3430	11.4176	-22.2353	179.6	190.6	181.6
	(0.0631)	(0.0319)	(0.0021)	(0.4447)	(6.3042)			
BBS	1.0452	57.5997	0.1990	1876.8935	0	191.6	200.4	193.0
	(0.0041)	(0.3643)	(0.0279)	(605.85)	(-)			
BS	0.43712	2.51540	1	1	0	204.3	208.7	205.0
	(0.0394)	(0.1432)	(-)	(-)	(-)			
McGa	α_1	β_1	a	b	c_1			
	28.5769	2.3734	0.1240	48.0712	0.2335	182.0	193.0	184.0
	(0.1195)	(0.0972)	(0.0052)	(0.1405)	(0.0988)			
McBS	α	β	a	η	c_1			
	3.8736	0.1487	18.8160	35.5380	29.00002	182.1	193.0	184.0
	(0.0232)	(0.0176)	(0.0549)	(0.4378)	(0.1761)			

Table 3.3 – LR statistics for the breaking stress of carbon fibres data

Model	Hypotheses	Statistic w	p-value
KBBS vs BBS	$H_0 : c = 0$ vs $H_1 : H_0$ is false	30.69	< 0.0001
KBBS vs BS	$H_0 : a = b = 1$ and $c = 0$ vs $H_1 : H_0$ is false	13.08	0.00029

models, respectively. We can conclude that the KBBS distribution is a very suitable model to fit to these data.

Secondly, we shall apply formal goodness-of-fit tests in order to verify which distribution gives the best fit to these data. We consider the Cramér-Von Mises (W^*) and Anderson-Darling (A^*) statistics. In general, the smaller the values of the statistics W^* and A^*, the better the fit to the data. The test statistics W^* and A^* are described in detail in (CHEN; BALAKRISHNAN, 1995). Let $F(x;\boldsymbol{\theta})$ be the cdf, where the form of F is known but $\boldsymbol{\theta}$ (a k-dimensional parameter vector, say) is unknown. To obtain the statistics W^* and A^*, we can proceed as follows:
1. Compute $v_i = F(x_i;\widehat{\boldsymbol{\theta}})$, where the x_i's are in ascending order;
2. Compute $y_i = \Phi^{-1}(v_i)$ is the normal standard quantile function;
3. Compute $u_i = \Phi[(y_i - \overline{y})/s_y]$, where $\overline{y} = n^{-1}\sum_{i=1}^{n} y_i$ and $s_y^2 = (n-1)^{-1}\sum_{i=1}^{n}(y_i - \overline{y})^2$;
4. Calculate

$$W^2 = \sum_{i=1}^{n}\left[u_i - \frac{(2i-1)}{2n}\right]^2 + \frac{1}{12n}$$

and

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}[(2i-1)\log(u_i) + (2n+1-2i)\log(1-u_i)];$$

5. Modify W^2 into $W^* = W^2(1 + 0.5/n)$ and A^* into $A^* = A^2(1 + 0.75/n + 2.25/n^2)$.

The values of the statistics W^* and A^* for all models are given in Table 3.4. Thus, according to these formal tests, the KBBS model fits to the current data better than its sub-models and

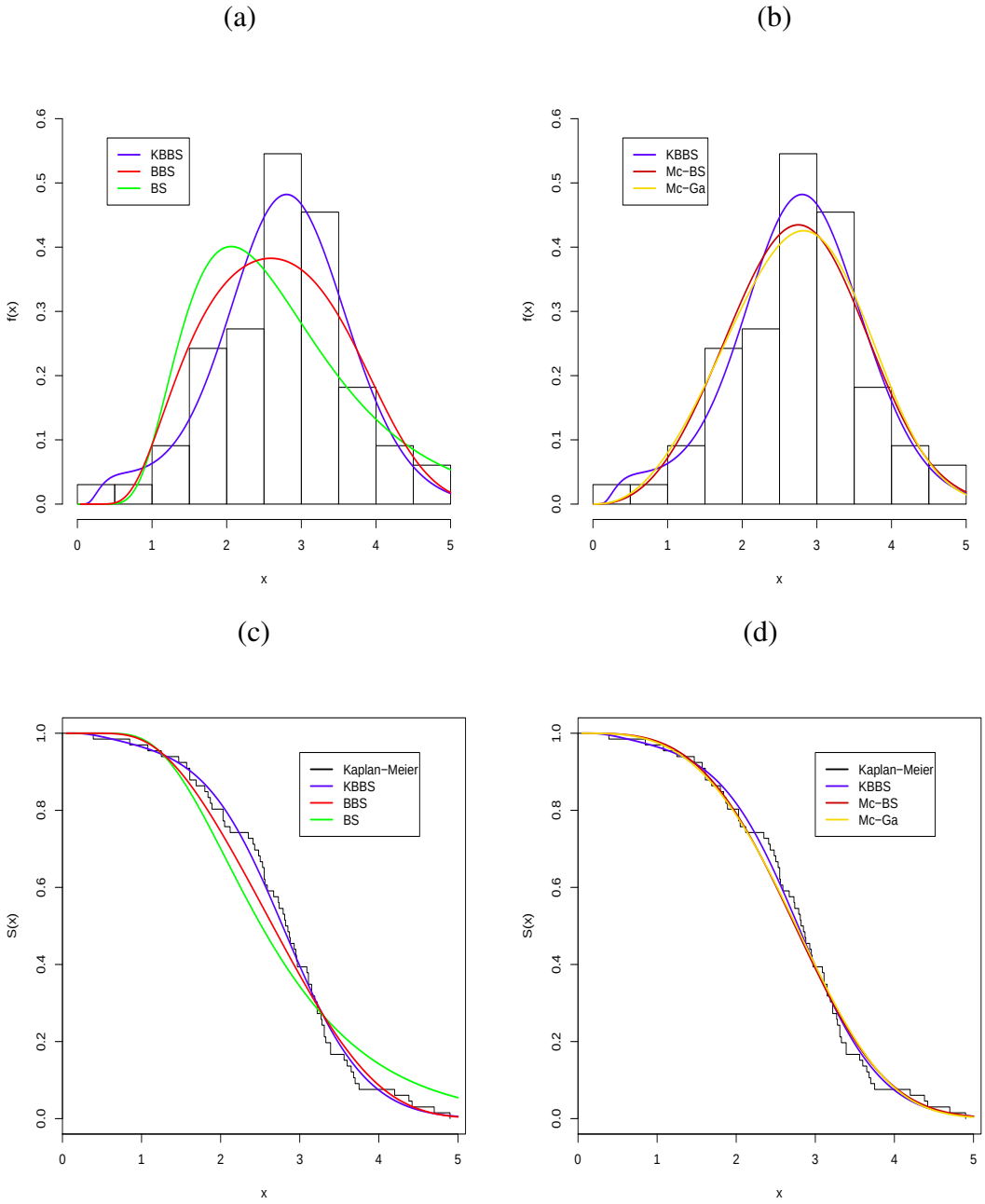


Figure 3.5 – (a) Estimated densities of the KBBS and its sub-models (b) Estimated densities of the KBBS, Mc-BS and McGa models (c) Empirical and estimated survival functions of the KBBS and its sub-models (d) Empirical and estimated survival functions of the KBBS, Mc-BS and McGa models

other lifetime models. These results illustrate the potentiality of the KBBS distribution and the necessity of the additional shape parameters.

The QQ plots of the normalized quantile residuals was introduced by (DUNN; SMYTH, 1996) and more recently used by (CORDEIRO et al., 2013). Figure 3.6 indicates the improved fit achieved using the KBBS distribution over the other distributions. We also emphasize the gain yielded by the KBBS distribution in relation to the BS, BBS, McBS and McGa distribu-

Table 3.4 – Formal tests for breaking stress of carbon fibres data

Model	Statistic	
	W^*	A^*
KBBS	0.0081	0.31232
BBS	0.2115	1.2216
BS	0.4603	2.5896
McBS	0.2522	0.5223
McGa	0.0812	0.5173

tions.

3.11 Concluding Remarks

The Birnbaum-Saunders (BS) distribution is widely used to model times to failure for materials subject to fatigue. We propose the Kummer beta Birnbaum-Saunders (KBBS) distribution to extend the BS distribution introduced by (BIRNBAUM; SAUNDERS, 1969a). We provide a mathematical treatment of the new distribution including expansions for the cumulative and density functions. We derive expansions for the ordinary and incomplete moments, generating function, mean deviations, reliability and the moments of the order statistics. The estimation of parameters is approached by the method of maximum likelihood and the observed information matrix is derived. We consider the likelihood ratio (LR) statistics and formal goodness-of-fit tests to compare the KBBS model with some of its sub-models and non-nested models. An application of the KBBS distribution to a real data set indicates that the new distribution provides consistently better fits than its sub-models and other lifetime models. We hope that this generalization may attract wider applications in the literature of the fatigue life distributions.

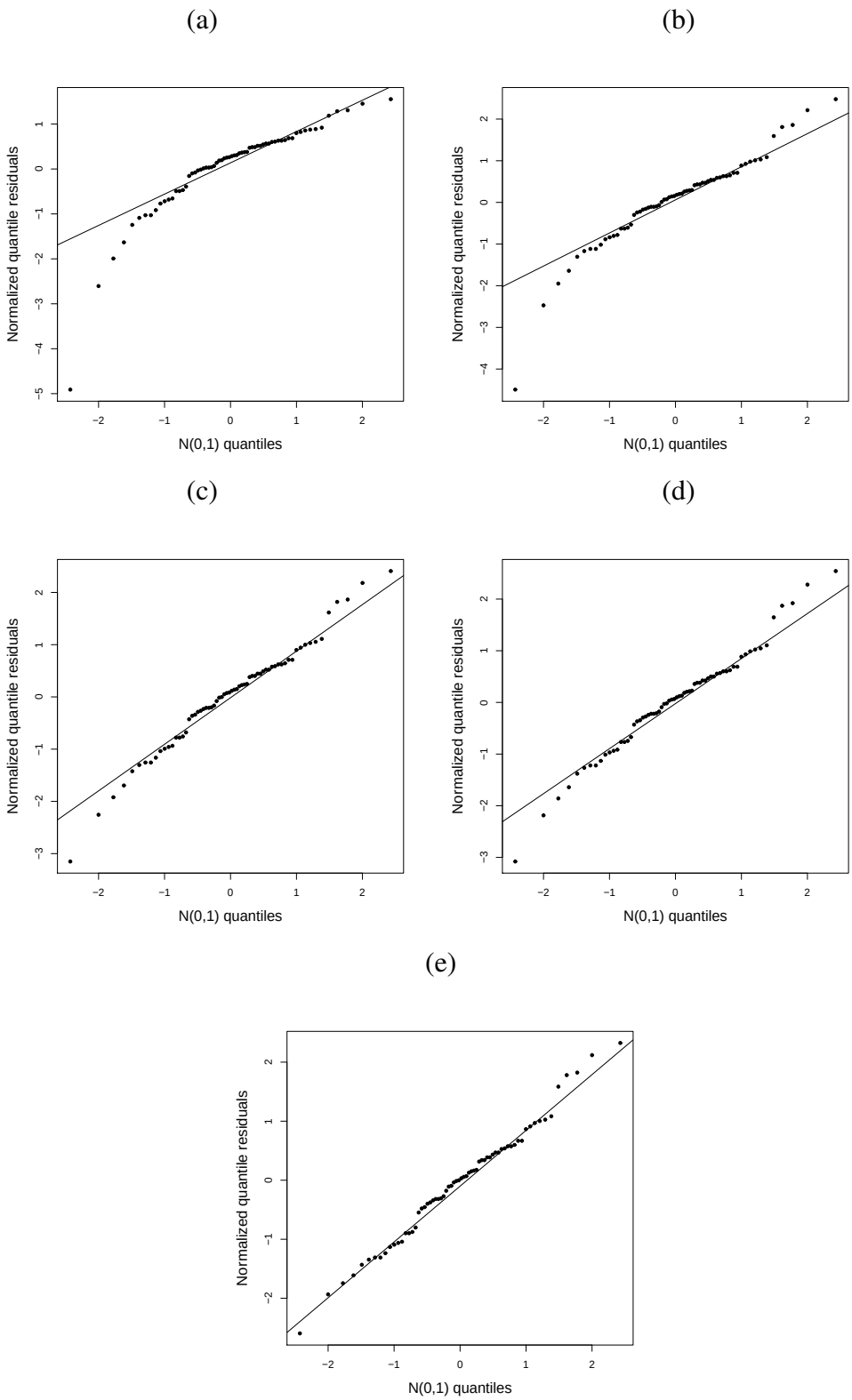


Figure 3.6 – QQ plot of the normalized quantile residuals with an identity line for the distributions: (a) BS, (b) BBS, (c) McBS, (d) McGa and (e) KBBS

References

ABRAMOWITZ, M.; STEGUN, I. A. **Handbook of Mathematical Functions**. New York: Dover Publications, 1968. 1046p.

BIRNBAUM, Z. W.; SAUNDERS, S. C. A new family of life distributions. **Journal of Applied Probability**, London, v.6, p.319-327, 1969a.

BIRNBAUM, Z. W.; SAUNDERS, S. C. Estimation for a family of life distributions with applications to fatigue. **Journal of Applied Probability**, London, v.6, p.328-377, 1969b.

CHEN, G.; BALAKRISHNAN, N. A general purpose approximate goodness-of-fit test. **Journal of Quality Technology**, New York, v.27, p.154-161, 1995.

CORDEIRO, G. M.; CASTELLARES, F.; MONTENEGRO, L. C.; de Castro, M. The beta generalized gamma distribution. **Statistics**, Berlin, v.47, p.888-900, 2013.

CORDEIRO, G. M.; LEMONTE, A. J. The beta Birnbaum-Saunders distribution: An improved distribution for fatigue life modeling. **Computational Statistics and Data Analysis**, Amsterdam, v.55, p.1445-1461, 2011.

CORDEIRO, G. M.; LEMONTE, A. J.; ORTEGA, E. M. M. An extended fatigue life distribution. **Statistics**, Berlin, v.47, p.626-653, 2011.

DÍAZ-GARCÍA, J. A.; LEIVA, V. A new family of life distributions based on elliptically contoured distributions. **Journal of Statistical Planning and Inference**, Amsterdam, v.137, p.1512-1513, 2005.

DUNN, P. K.; SMYTH, G. K. Randomized quantile residuals, **Journal of Computational and Graphical Statistics**, New York, v.5, p.236-244, 1996.

GRADSHTEYN, I. S.; RYZHIK, I. M. **Table of Integrals, Series, and Products**. San Diego: Academic Press, 2007. 315p.

GREENWOOD, J. A.; LANDWEHR, J. M.; MATALAS, N. C.; WALLIS, J. R. Probability weighted moments: Definition and relation to parameters of several distributions expressible in inverse form. **Water Resources Research**, New York, v.15, p.1049-1054, 1979.

GUIRAUD, P.; LEIVA, V.; FIERRO, R. A non-central version of the Birnbaum-Saunders distribution for reliability analysis. **IEEE Transactions on Reliability**, New York, v.58, p.152-160, 2009.

HARRIS, F. E. Incomplete Bessel, generalized incomplete gamma, or leaky aquifer functions. **Journal of Computational and Applied Mathematics**, Amsterdam, v.215, p.260-269, 2008.

JONES, D. S. Incomplete Bessel functions. **Proceedings of the Edinburgh Mathematical Society**, Edinburgh, v.50, p.173-183, 2007a.

JONES, D. S. Incomplete Bessel functions. Asymptotic expansions for large argument. **Proceedings of the Edinburgh Mathematical Society**, Edinburgh, v.50, p.711-723, 2007b.

KUNDU, D.; KANNAN, N.; BALAKRISHNAN, N. On the function of Birnbaum-Saunders distribution and associated inference. **Computational Statistics and Data Analysis**, Amsterdam, v.52, p.2692-2702, 2008.

LEMONTE, A. J.; CRIBARI-NETO, F.; VASCONCELLOS, K. L. P. Improved statistical inference for the two-parameter Birnbaum-Saunders distribution. **Computational Statistical and Data Analysis**, Amsterdam, v.51, p.4656-4681, 2007.

LEMONTE, A. J.; SIMAS, A. B.; CRIBARI-NETO, F. Bootstrap-based improved estimators for the two-parameter Birnbaum-Saunders distribution. **Journal of Statistical Computation and Simulation**, London, v.78, p.37-49, 2008.

MARCIANO, W. P.; NASCIMENTO, A. D. C.; SANTOS-NETOS, M.; CORDEIRO, G. M. The Mc-Γ Distribution and Its Statistical Properties: An Application to Reliability Data. **International Journal of Statistics and Probability**, Ottawa, DOI:10.5539/ijsp.v1n1p53, 2012.

MURTHY, V. K. **The General Point Process**. Boston: Addison-Wesley, 1974. 280p.

NICHOLS, M. .D; PADGETT, W. J. A bootstrap control chart for Weibull percentiles. **Quality and Reliability Engineering International**, London, v.22, p.141-151, 2006.

PARANAÍBA, P. F.; ORTEGA, E. M. M.; CORDEIRO, G. M.; PESCIM, R. R. The beta Burr XII distribution with application to lifetime data. **Computational Statistics and Data Analysis,** Amsterdam, v.55, p.1118-1136, 2011.

PESCIM, R. R.; CORDEIRO, G .M.; DEMÉTRIO, C. G. B.; ORTEGA, E. M. M.; NADARAJAH, S. The new class of Kummer beta generalized distributions. **Statistics and Operations Research Transactions**, Barcelona, v.36, p.153-180, 2012.

RIECK, J. R. A moment-generating function with application to the Birnbaum-Saunders distribution, **Communications in Statistics - Theory and Methods**, London, v.28, p.2213-2222, 1999.

SAUNDERS, S. C. The problems of estimating a fatigue service life with a low probability of failure. **Engineering Fracture Mechanics**, Amsterdam, v.8, p.205-215, 1976.

TERRAS, R. A Miller algorithm for an incomplete Bessel function. **Journal of Computational Physics**, New York, v.39, p.233-240, 1981.

VALLURI, R. S. Some recent developments at galci concerning a theory of metal fatigue. **Acta Metallurgica**, Beijing, v.11, p.759-775, 1963.

WATSON, G. N. **A Treatise on the Theory of Bessel Functions**. London: Cambridge University Press, 1995. 399p.

WU, J.; WONG, A. C. M. Improved interval estimation for the two-parameter Birnbaum-Saunders distribution. **Computational Statistics and Data Analysis**, Amsterdam, v.47, p.809-821, 2004.

4 A NEW EXTENSION OF THE GENERALIZED GAMMA DISTRIBUTION

Abstract

A new extension of the generalized gamma distribution with six-parameter so-called the Kummer beta generalized gamma distribution is introduced and studied. It contains at least 32 special models such as the beta generalized gamma, beta Weibull, beta exponential, generalized gamma, Weibull and gamma distributions and thus could be a better model for analyzing positive skewed data. The new density function can be expressed as a linear combination of generalized gamma densities. Various mathematical properties of the new distribution including explicit expressions for the ordinary and incomplete moments, generating function, mean deviations, entropy, density function of the order statistics and their moments are derived. The elements of the observed information matrix are provided. We discuss the method of maximum likelihood and a Bayesian approach to fit the model parameters. The superiority of the new model is illustrated by means of three real data sets.

Keywords: Generalized gamma distribution; Kummer beta distribution; Lifetime data; Maximum likelihood estimation; Mean deviation; Moment

4.1 Introduction

The generalized gamma (GG) distribution (STACY, 1962) is an important lifetime model since it includes as special models the exponential, Weibull, gamma and Rayleigh distributions, among others. It is suitable for modeling data with hazard rate function of different forms (increasing, decreasing, bathtub and unimodal) and then it is useful for estimating individual hazard functions and both relative hazards and relative times (COX, 2008). The GG distribution has been used in several research areas such as engineering, environment, hydrology and survival analysis. For example, (ORTEGA; BOLFARINE; PAULA, 2003) discussed influence diagnostics in GG regression models, (NADARAJAH; GUPTA, 2007) applied this distribution to drought data, (COX et al., 2007) presented a parametric survival analysis based on GG hazard functions and (COX, 2008) discussed and compared the F-generalized family with the GG model. More recently, (BARKAUSKAS et al., 2009) modeled the noise part of a spectrum as an autoregressive moving average (ARMA) model with the innovations following the GG distribution, (MALHOTRA; SHARMA; KALER, 2009) provided a unified analysis for wireless system over generalized fading channels that is modeled by a two parameter GG model and (XIE; LIU, 2009) analyzed three-moment auto conversion parametrization based on this model. Further, (ORTEGA; CANCHO; PAULA, 2009) proposed a modified GG regression model to allow the possibility that long-term survivors may be presented in the data and (CORDEIRO; ORTEGA; SILVA, 2011) studied the exponentiated generalized gamma (EGG) distribution.

Let $\gamma_1\left(k, x/\alpha\right)$ be the cumulative distribution function (cdf) of the standard gamma distribution, where $\gamma_1(\cdot,\cdot)$ is the incomplete gamma ratio function defined by $\gamma_1(k,x) = \gamma(k,x)/\Gamma(k)$, $\gamma(k,x) = \int_0^x w^{k-1}\,\mathrm{e}^{-w}dw$ and $\Gamma(\cdot)$ are the incomplete and complete gamma functions. The

probability density function (pdf) of GG distribution, with three parameters $\alpha > 0$, $\beta > 0$ and $k > 0$, defined by (STACY, 1962), has the form

$$g(x;\alpha,\beta,k) = \frac{\beta}{\alpha\Gamma(k)}\left(\frac{x}{\alpha}\right)^{\beta k-1}\exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right],\ x>0. \tag{4.1}$$

In the density function (4.1), $\alpha > 0$ is a scale parameter and $\beta > 0$ and $k > 0$ are shape parameters. The cumulative distribution function (cdf) corresponding to (4.1) is

$$G(x;\alpha,\beta,k) = \gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]. \tag{4.2}$$

For an arbitrary baseline distribution $G(x;\boldsymbol{\gamma})$ with parameter vector $\boldsymbol{\gamma}$ and density function $g(x;\boldsymbol{\gamma})$, (PESCIM et al., 2012) proposed the Kummer beta generalized (denoted by the prefix "KB-G" for short) cumulative function defined by

$$F_{\mathcal{KBG}}(x) = K\int_0^{G(x;\boldsymbol{\gamma})} t^{a-1}\,(1-t)^{b-1}\,\mathrm{e}^{-\mathrm{ct}}\,\mathrm{dt}, \tag{4.3}$$

where $a > 0$ and $b > 0$ are shape parameters which induce skewness, and thereby promote weight variation of the tails, whereas the parameter $-\infty < c < \infty$ "squeezes" the pdf to the left or right, i.e., it gives weights to the extremes of the density functions. Here,

$$K^{-1} = \frac{\Gamma(a)\,\Gamma(b)}{\Gamma(a+b)}\,{}_1F_1(a;a+b;-c)$$

and

$${}_1F_1(a;a+b;-c) = \frac{\Gamma(a+b)}{\Gamma(a\,\Gamma(b)}\int_0^1 t^{a-1}\,(1-t)^{b-1}\mathrm{e}^{-\mathrm{ct}}\,\mathrm{dt} = \sum_{\mathrm{k}=0}^{\infty}\frac{(\mathrm{a})_{\mathrm{k}}\,(-\mathrm{c})^{\mathrm{k}}}{(\mathrm{a}+\mathrm{b})_{\mathrm{k}}\,\mathrm{k}!}$$

is the confluent hypergeometric function (ABRAMOWITZ; STEGUN, 1968) and $(d)_k = d(d+1)\ldots(d+k-1)$ denotes the ascending factorial. An important special model is the beta distribution when $c = 0$. The density function corresponding to (4.3) can be expressed as

$$f_{\mathcal{KBG}}(x) = K\,g(x;\boldsymbol{\gamma})\,G(x;\boldsymbol{\gamma})^{a-1}\,\left[1-G(x;\boldsymbol{\gamma})\right]^{b-1}\exp\left[-c\,G(x;\boldsymbol{\gamma})\right]. \tag{4.4}$$

Equation (4.4) will be most tractable when both functions $G(x;\boldsymbol{\gamma})$ and $g(x;\boldsymbol{\gamma})$ have simple analytic expressions. Its major benefit is to offer more flexibility to extremes (right and/or left) of the density functions and therefore it becomes suitable for analyzing data with high degree of asymmetry.

The class of distributions (4.4) includes two important special cases: the beta-generalized (BG) and exponentiated generalized (EG) distributions defined by (EUGENE; LEE; FAMOYE, 2002) and (MUDHOLKAR; SRIVASTAVA; FRIEMER, 1995) when $c = 0$ and $c = 0$ and $b = 1$, respectively.

In this work, we introduce a new six-parameter distribution called the *Kummer beta generalized gamma* (KBGG) distribution which contains at least 32 special sub-models. The main

motivation for this extension is that the new distribution is a highly flexible life distribution which admits different degrees of kurtosis and asymmetry. The KBGG density function is defined from (4.4) by taking (4.2) and (4.1) as the cdf and pdf of the GG(α, β, k) distribution, respectively. The six-parameter KBGG density function can be expressed as

$$\begin{aligned} f(x) &= K\,\frac{\beta}{\alpha\,\Gamma(k)}\left(\frac{x}{\alpha}\right)^{\beta k-1}\exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]^{a-1} \\ &\quad\times\left\{1-\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^{b-1}\exp\left\{-c\,\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}. \end{aligned} \tag{4.5}$$

The associated hazard rate function to (4.5) becomes

$$\tau(x) = \frac{K\,\beta\left(\frac{x}{\alpha}\right)^{\beta k-1}\left\{1-\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^{b-1}}{\alpha\,\Gamma(k)\,[1-F(x)]\;\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]^{1-a}}\exp\left[-\left\{c\,\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^{\beta}\right]+\left(\frac{x}{\alpha}\right)^{\beta}\right\}\right] \tag{4.6}$$

Hereafter, we denote by X the random variable following (4.5), say $X \sim$ KBGG$(a, b, c, \alpha, \beta, k)$. This density has five shape parameters a, b, c, β and k which allow for a high degree of flexibility. The parameter c controls tail weights to the extremes of the distribution. The study of the new distribution is important since it extends some distributions previously considered in the literature. In fact, the generalized gamma (GG) model is clearly a basic exemplar for $a = b = 1$ and $c = 0$, with a continuous crossover towards models with different shapes (e.g. a specified combination of skewness and kurtosis). The KBGG model contains as sub-models the beta generalized gamma (BGG) (CORDEIRO et al., 2012) and the exponentiated generalized gamma (EGG) (CORDEIRO; ORTEGA; SILVA, 2011) distributions when $c = 0$ and $b = 1$ in addition to $c = 0$, respectively. Plots of the new density function for selected parameter values are represented in Figure 4.1. It is evident that the shapes of this density function are much more flexible than the GG distribution. Hence the KBGG can be used in many practical situations. In fact, it can be symmetric, asymmetric and also exhibit bimodality. More details, see Section 4.10. We also provide a comprehensive description of some of its mathematical properties with the hope that it will attract wider applications in reliability, engineering, environment and in other areas of research.

This chapter is outlined as follows. In Section 4.2, we derive more than 32 special distributions from KBGG model. In Section 4.3, we demonstrate that the KBGG density function can be expressed as a linear combination of GG density functions. This is an important result to provide some mathematical properties of the GG distribution. We obtain explicit expressions for the moments and generating function (Section 4.4), incomplete moments (Section 4.5), mean deviations (Section 4.6), Rényi entropy (Section 4.7) and order statistics (Section 4.8). In Section 4.9, we discuss maximum likelihood estimation and statistical inference. In Section 4.10, three applications are presented to illustrate the usefulness of the new distribution for real data. Concluding remarks are addressed in Section 4.11.

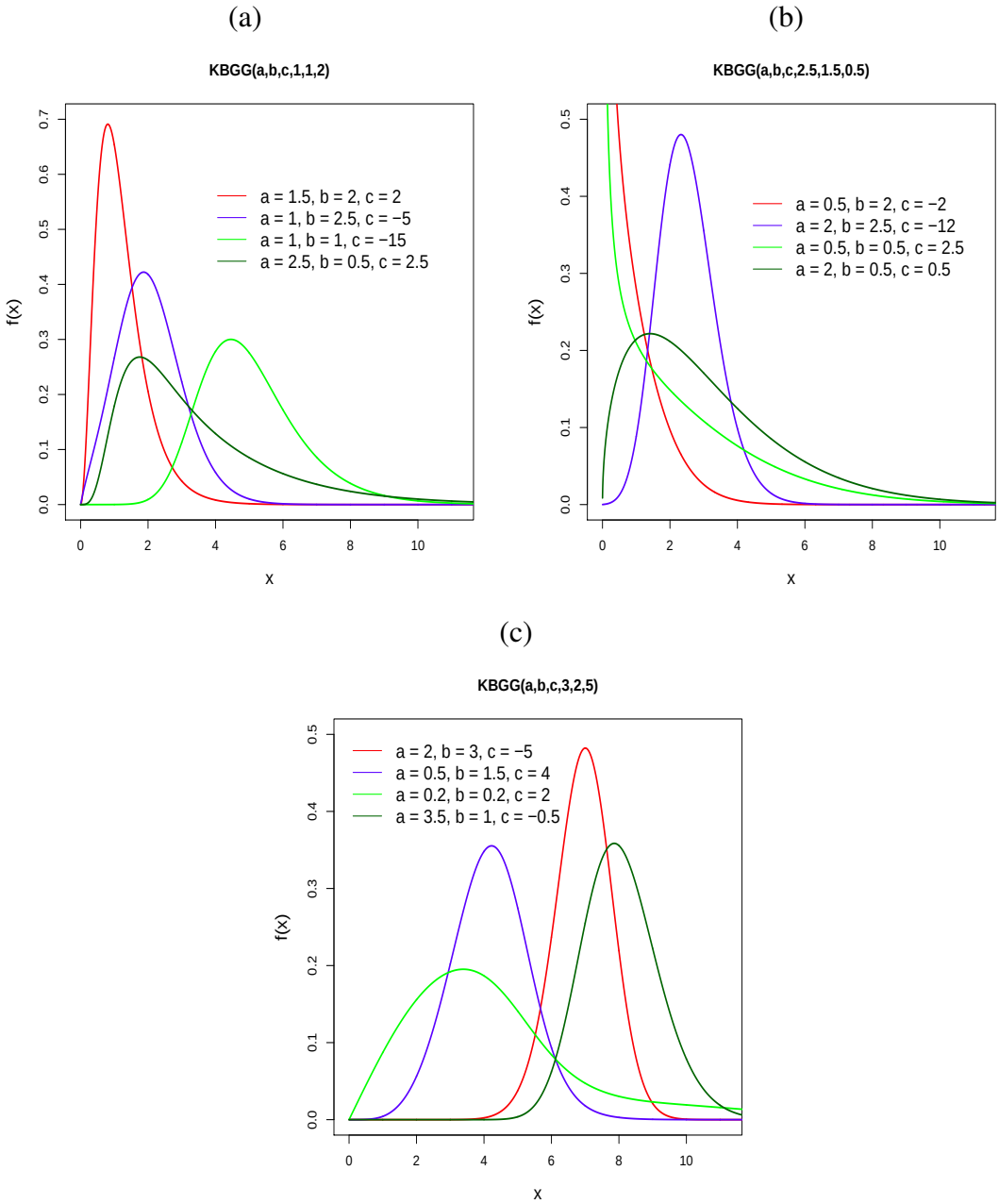


Figure 4.1 – Plots of the KBGG density function (4.5) for some parameter values

4.2 Special Cases of the KBGG Distribution

The following well-known distributions are special sub-models of the KBGG distribution.

4.2.1 Kummer Beta Generator

- For $k = 1$, the KBGG distribution reduces to the Kummer beta Weibull (KBW) distribution. If $k = 1$ and $\beta = 1$, it yields the Kummer beta exponential (KBE) distribution. If $\beta = 2$ in addition to $k = 1$, it gives the Kummer beta generalized Rayleigh (KBGR) distribution. For $\alpha = \sqrt{2}\sigma$, $\beta = 2$ and $k = p/2$, the KBGG distribution corresponds to the

Kummer beta scaled chi-square (KBSchi) distribution. For $\alpha = \sqrt{\theta}$, $\beta = 2$ and $k = 3/2$, the KBGG distribution coincides with the Kummer beta Maxwell (KBMa) distribution.

- For $\beta = 1$, the KBGG distribution coincides with the four parameter Kummer beta gamma (KBGa) distribution. Taking $\alpha = 2$, $\beta = 1$ and $k = p/2$, we obtain the Kummer beta chi-square (KBchi) distribution. Moreover, if $\alpha = 2^{\frac{1}{2\gamma}}\theta$, $\beta = 2\gamma$ and $k = 1/2$, the KBGG distribution becomes the Kummer beta generalized half-normal (KBGHN) distribution. If $\alpha = 2^{\frac{1}{2}}\theta$, $\beta = 2$ and $k = 1/2$, the KBGG model corresponds to the the Kummer beta half-normal (KBHN) distribution. Finally, if $\alpha = \sqrt{w/\mu}$, $\beta = 2$ and $k = \mu$, it yields the Kummer beta Nakagami (KBNa) distribution.

4.2.2 Beta Generator (for $c = 0$)

- For $c = 0$, the KBGG distribution reduces to the five parameter beta generalized gamma (BGG) distribution. If $k = 1$, the BGG distribution corresponds to the beta Weibull (BW) distribution introduced by (FAMOYE; LEE; OLUMOLADE, 2005). If $\beta = 1$ and $k = 1$, it gives the beta exponential (BE) distribution (NADARAJAH; KOTZ, 2005). If $\beta = 2$ in addition to $k = 1$, it yields the beta generalized Rayleigh (BGR) distribution (CORDEIRO et al., 2011). For $\alpha = \sqrt{2}\sigma$, $\beta = 2$ and $k = p/2$, the BGG distribution reduces to the beta scaled chi-square (BSchi) distribution. For $\alpha = \sqrt{\theta}$, $\beta = 2$ and $k = 3/2$, the BGG distribution gives the beta Maxwell (BMa) distribution.

- For $\beta = 1$, the BGG distribution yields the four parameter beta gamma (BGa_3) distribution. If $\alpha = 1$ in addition to $\beta = 1$, the special case corresponds to the beta gamma (BGa_2) distribution. Taking $\alpha = 2$, $\beta = 1$ and $k = p/2$, we obtain the beta chi-square (Bchi) distribution. Further, if $\alpha = 2^{\frac{1}{2\gamma}}\theta$, $\beta = 2\gamma$ and $k = 1/2$, the BGG distribution becomes the beta generalized half-normal (BGHN) distribution proposed by (PESCIM et al., 2010). If $\alpha = 2^{\frac{1}{2}}\theta$, $\beta = 2$ and $k = 1/2$, the BGG model reduces to the distribution which is called the beta half-normal (BHN). (PESCIM et al., 2010).

- Finally, if $\alpha = \sqrt{w/\mu}$, $\beta = 2$ and $k = \mu$, the BGG distribution becomes the Beta Nakagami (BNa) distribution.

4.2.3 Exponentiated Generator (for $b = 1$ and $c = 0$)

- For $b = 1$ and $c = 0$ we obtain from (4.5) the density function of the EGG distribution given by (4.1). If $k = 1$, the EGG distribution reduces to the density of the exponentiated Weibull (EW) distribution introduced by (MUDHOLKAR; SRIVASTAVA; FRIEMER, 1995). If $\beta = 1$ in addition to $k = 1$, the special case corresponds to the exponentiated exponential (EE) distribution (GUPTA; KUNDU, 2001). If $\beta = 2$ in addition to $k = 1$, the special case corresponds to the generalized Rayleigh (GR) distribution. For $\alpha = \sqrt{2}\sigma$, $\beta = 2$ and $k = p/2$, the EGG distribution reduces to the exponentiated scaled chi-square

(ESchi) distribution. For $\alpha = \sqrt{\theta}$, $\beta = 2$ and $k = 3/2$, the EGG distribution corresponds to the exponentiated Maxwell (EMa) distribution.

- For $\beta = 1$, the EGG distribution reduces to the three parameter exponentiated gamma (EGa_3) distribution. If $\alpha = 1$ in addition to $\beta = 1$, the special case corresponds to the exponentiated gamma (EGa_2) distribution. Taking $\beta = 1$ and $\lambda = 1$, the special case corresponds to the two parameter gamma (Ga2p) distribution. Further, if $\beta = \lambda = k = 1$, we obtain the one parameter gamma (Ga1p) distribution. Taking $\alpha = 2$, $\beta = 1$ and $k = p/2$, we obtain the exponentiated chi-square (Echi) distribution. Further, if $\lambda = 1$, in addition to $\alpha = 2$, $\beta = 1$ and $k = p/2$, we obtain the chi-square (Chi) distribution. If $\alpha = 2^{\frac{1}{2\gamma}}\theta$, $\beta = 2\gamma$ and $k = 1/2$, the EGG distribution becomes the exponentiated generalized half-normal (EGHN) distribution.

- If $\alpha = 2^{\frac{1}{2}}\theta$, $\beta = 2$ and $k = 1/2$, the EGG model reduces to the distribution we call the exponentiated half-normal (EHN). Finally, if $\alpha = \sqrt{w/\mu}$, $\beta = 2$ and $k = \mu$, the EGG distribution becomes the exponentiated Nakagami (ENa) distribution.

4.2.4 Baseline distributions (for $a = b = 1$ and $c = 0$)

- For $a = b = 1$ and $c = 0$, the new model reduces to the three parameter generalized gamma (GG) distribution. The case $\beta = 1$ corresponds to the classical two parameter Weibull (W) distribution. If $\beta = 1$ and $\beta = 2$, in addition to $k = 1$, the special case coincides with the exponential (E) and Rayleigh (R) distributions, respectively. For $\alpha = \sqrt{2}\sigma$, $\beta = 2$ and $k = p/2$, the special case corresponds to the scaled chi-square (SChi) distribution. If $\alpha = \sqrt{\theta}$ in addition to $\beta = 2$ and $k = 3/2$, it reduces to the Maxwell (Ma) distribution (BEKKER; ROUX, 2005).

- Taking $\beta = 1$, the special case corresponds to the two parameter gamma (Ga) distribution. If $\alpha = 2$, in addition to $\beta = 1$ and $k = p/2$, we obtain the chi-square (Chi) distribution. If $\alpha = 2^{1/(2\gamma)}\theta$ in addition to $\beta = 2\gamma$, $k = 1/2$, it coincides with the generalized half-normal (GHN) distribution introduced by (COORAY; ANANDA, 2008). Taking $\alpha = 2^{\frac{1}{2}}\theta$ in addition to $\beta = 2$ and $k = 1/2$, it reduces to the well-known half-normal (HN) distribution. Further, if $\alpha = \sqrt{w/\mu}$ in addition to $\beta = 2$ and $k = \mu$, the special case corresponds to the Nakagami (Na) distribution.

Several special sub-models of the KBGG model are illustrated in Table 4.1

4.3 Expansion for the Density Function

A useful expansion for equation (4.5) can be derived using the concept of exponentiated generalized distributions. First, we use an expansion for the general density function (4.4) expressed as a linear combination of EG densities.

Table 4.1 – Some special cases of the KBGG distribution

							$a = b = 1$ and $c = 0$	
Case	α		β		k		Distribution	References
(1)	α		β		k		Generalized gamma	(STACY, 1962)
(2)	α		β		1		Weibull	
(3)	α		1		k		Gamma	
(4)	α		1		1		Exponential	
(5)	α		2		1		Rayleigh	
(6)	$\sqrt{\theta}$		2		$3/2$		Maxwell	
(7)	$\sqrt{2}\alpha$		2		$p/2$		Scaled Chi-Square	
(8)	$2^{\frac{1}{2\alpha}}\theta$		2α		$1/2$		Generalized half-normal	(COORAY; ANANDA, 2008)
							$b = 1$ and $c = 0$	
	α	β	k		a			
(9)	α	β	k		a		Exponentiated generalized gamma	
(10)	α	β	1		a		Exponentiated Weibull	
(11)	α	1	1		a		Exponentiated gamma	(NADARAJAH; GUPTA, 2007)
(12)	α	1	1		a		Exponentiated exponential	(GUPTA; KUNDU, 2001)
(13)	α	2	1		a		Exponentiated Rayleigh	
(14)	$\sqrt{\theta}$	2	$3/2$		a		Exponentiated Maxwell	
(15)	$\sqrt{2}\alpha$	2	$p/2$		a		Exponentiated Scaled Chi-Square	
(16)	$2^{\frac{1}{2\alpha}}\theta$	2α	$1/2$		a		Exponentiated generalized half-normal	
							$c = 0$	
	α	β	k	a	b			
(17)	α	β	k	a	b		Beta generalized gamma	(CORDEIRO et al., 2013)
(18)	α	β	1	a	b		Beta Weibull	
(19)	α	1	k	a	b		Beta gamma	(KONG; LEE; SEPANSKI, 2007)
(20)	α	1	1	a	b		Beta exponential	(NADARAJAH; KOTZ, 2005)
(21)	α	2	1	a	b		Beta generalized Rayleigh	(CORDEIRO et al., 2011)
(22)	$\sqrt{\theta}$	2	$3/2$	a	b		Beta Maxwell	
(23)	$\sqrt{2}\alpha$	2	$p/2$	a	b		Beta Scaled Chi-Square	
(24)	$2^{\frac{1}{2\alpha}}\theta$	2α	$1/2$	a	b		Beta generalized half-normal	(PESCIM et al., 2010)
	α	β	k	a	b	c		
(25)	α	β	1	a	b	c	Kummer beta Weibull	(PESCIM et al., 2012)
(26)	α	1	k	a	b	c	Kummer beta gamma	(PESCIM et al., 2012)
(27)	α	1	1	a	b	c	Kummer beta exponential	New
(28)	α	2	1	a	b	c	Kummer beta generalized Rayleigh	New
(29)	$\sqrt{\theta}$	2	$3/2$	a	b	c	Kummer beta Maxwell	New
(30)	$\sqrt{2}\alpha$	2	$p/2$	a	b	c	Kummer beta Scaled Chi-Square	New
(31)	$2^{\frac{1}{2\alpha}}\theta$	2α	$1/2$	a	b	c	Kummer beta generalized half-normal	New
(32)	$2^{\frac{1}{2}}\theta$	2	$1/2$	a	b	c	Kummer beta half-normal	New

(PESCIM et al., 2012) demonstrated that

$$f_{\mathcal{KBG}}(x) = \sum_{r=0}^{\infty} c_r \, v_{r+1}(x), \tag{4.7}$$

where the coefficients (for $r = 0, 1 \ldots$) are $c_r = \sum_{i,j=0}^{\infty} \sum_{k=r+1}^{\infty} t_{i,j,k,r+1}$,

$$t_{i,j,k,r} = t_{i,j,k,r}(a,b,c) = \frac{K\,(-1)^{i+j+k+r}\,c^i}{i!\,(a+i+j)} \binom{a+i+j}{k} \binom{k}{r} \binom{b-1}{j}$$

and $v_{r+1}(x) = (r+1)\,g(x)\,G(x)^r$ denotes the EG density function with parameter $r+1$.

Equation (4.7) reveals that the KB-G density function is a linear combination of EG densities. This result is important to derive some properties of the KBGG distribution from those of the EGG distribution. This equation holds for any real non-integers a, b and c. If b is an integer, the index i in c_r stops at $b-1$.

Replacing (4.1) and (4.2) in $v_{r+1}(x)$, we obtain the EGG$(\alpha, \beta, k, r+1)$ density function given by

$$v_{r+1}(x) = \frac{(r+1)\,\beta}{\alpha\,\Gamma(k)} \left(\frac{x}{\alpha}\right)^{\beta k-1} \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right] \left\{\gamma_1\left[k, \left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^r. \tag{4.8}$$

We now need to use the series expansion for the incomplete ratio function in (4.8) given by

$$\gamma_1\left[k, \left(\frac{x}{\alpha}\right)^{\beta}\right] = \frac{\left(\frac{x}{\alpha}\right)^{\beta k}}{\Gamma(k)} \sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{x}{\alpha}\right)^{\beta m}}{(k+m)m!}. \tag{4.9}$$

Using the identity for power series raised to powers (GRADSHTEYN; RYZHIK, 2007), we obtain for any r positive integer

$$\left(\sum_{m=0}^{\infty} a_m \, x^m\right)^r = \sum_{m=0}^{\infty} d_{r,m} \, x^m, \tag{4.10}$$

where the coefficients $d_{r,m}$ (for $m = 1, 2, \ldots$) satisfy the recurrence relationship

$$d_{r,m} = (m\,a_0)^{-1} \sum_{p=1}^{m} (rp - m + p)\, a_p \, d_{r,m-p}, \tag{4.11}$$

where $d_{r,0} = a_0^r$. The coefficient $d_{r,m}$ comes from $d_{r,0}, \ldots, d_{r,m-1}$ and hence from $a_0, \ldots, a_m$. The coefficients $d_{r,m}$ can also be written explicitly as functions of the quantities a_m.

Further, using equations (4.9) and (4.10), we obtain the expanded form of an integer raised to power of the GG cumulative distribution given by

$$\begin{aligned}
\left\{\gamma_1\left[k, \left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^r &= \left[\frac{\left(\frac{x}{\alpha}\right)^{\beta k}}{\Gamma(k)} \sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{x}{\alpha}\right)^{\beta m}}{(k+m)m!}\right]^r \\
&= \frac{\left(\frac{x}{\alpha}\right)^{\beta k r}}{\Gamma(k)^r} \left[\sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{x}{\alpha}\right)^{\beta m}}{(k+m)m!}\right]^r \\
&= \frac{\left(\frac{x}{\alpha}\right)^{\beta k r}}{\Gamma(k)^r} \sum_{m=0}^{\infty} d_{r,m} \left(\frac{x}{\alpha}\right)^{\beta m},
\end{aligned} \tag{4.12}$$

where the coefficients $d_{r,m}$ are just obtained from equation (4.11) with $a_p = (-1)^p/(k+p)p!$. Combining equations (4.8) and (4.12), we can rewrite the EGG density function as

$$\begin{aligned} v_{r+1}(x) &= \frac{(r+1)\,\beta}{\alpha\,\Gamma(k)}\left(\frac{x}{\alpha}\right)^{\beta k-1}\exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]\frac{\left(\frac{x}{\alpha}\right)^{\beta kr}}{\Gamma(k)^r}\sum_{m=0}^{\infty} d_{r,m}\left(\frac{x}{\alpha}\right)^{\beta m} \\ &= \sum_{m=0}^{\infty}\frac{d_{r,m}\,(r+1)\,\beta}{\alpha\,\Gamma(k)^{r+1}}\left(\frac{x}{\alpha}\right)^{\beta k+\beta r+\beta m-1}\exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right] \\ &= \sum_{m=0}^{\infty}\frac{d_{r,m}\,\Gamma\left[k(r+1)+m\right]}{\Gamma(k)^{r+1}\,(r+1)^{-1}}\,\frac{\beta\,\mathrm{e}^{-\left(\frac{x}{\alpha}\right)^{\beta}}}{\alpha\,\Gamma\left[k(r+1)+m\right]}\left(\frac{x}{\alpha}\right)^{\beta[k(r+1)+m]-1} \\ &= \sum_{m=0}^{\infty}\eta_{r,m}\;g_{\alpha,\beta,k^\star}(x), \end{aligned} \tag{4.13}$$

where

$$\eta_{r,m} = \frac{d_{r,m}\,\Gamma\left(k^\star\right)}{\Gamma(k)^{r+1}\,(r+1)^{-1}},$$

$k^\star = k(r+1)+m$ and $g_{\alpha,\beta,k^\star}(x)$ is the density function of the GG$(\alpha,\beta,k^\star)$ distribution.

Equation (4.13) reveals that the KBGG density function can be written as a linear combination of GG densities. This equation is the main result of this section. It plays an important role in this chapter. In the next sections, based on this expression, we obtain some of the structural properties for the KBGG distribution including explicit expressions for the ordinary and incomplete moments, generating function, mean deviations and for the pdf of the order statistics.

4.4 Moments and Generating Function

Let X be a random variable having a KBGG distribution. The sth moment of X can be expressed from (4.13) as

$$\mu_s' = \mathrm{E}(X^s) = \sum_{r,m=0}^{\infty}\eta_{r,m}\int_0^{\infty} x^s\,g_{\alpha,\beta,k^\star}(x)dx$$

and then

$$\mathrm{E}(X^s) = \sum_{r,m=0}^{\infty}\eta_{r,m}\;\mathrm{E}(X_{k^\star}^s), \tag{4.14}$$

where $X_{k^\star}\sim$ GG$(\alpha,\beta,k^\star)$.

Equation (4.14) is an important result since it gives the moments of the KBGG distribution as a linear combination of GG moments. So, we have

$$\mathrm{E}(X_{k^\star}^s) = \frac{\beta}{\alpha\,\Gamma(k^\star)}\int_0^{\infty} x^s\left(\frac{x}{\alpha}\right)^{\beta k^\star-1}\exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]dx.$$

Next, by setting $u=\left(\frac{x}{\alpha}\right)^{\beta}$ in last integral, $\mathrm{E}(X_{k^\star}^s)$ reduces to

$$\mathrm{E}(X_{k^\star}^s) = \alpha^s\,\frac{\Gamma\left[k(r+1)+m+s/\beta\right]}{\Gamma(k(r+1)+m)}. \tag{4.15}$$

Replacing (4.15) in (4.14), we obtain the sth moment of X given by

$$\mathrm{E}(X^s) = \alpha^s \sum_{r,m=0}^{\infty} \eta_{r,m} \frac{\Gamma\left[k(r+1)+m+s/\beta\right]}{\Gamma(k(r+1)+m)}, \tag{4.16}$$

where $\eta_{r,m}$ is defined by (4.13).

Equation (4.16) is readily computed numerically using standard statistical software. It (and other expansions in this paper) can also be evaluated in symbolic computation software such as Mathematica and Maple. In numerical applications, a large natural number N can be used in the sums instead of infinity. Several quantities of X (central moments, variance, skewness and kurtosis) can be derived from this result.

The skewness and kurtosis measures can be calculated from the ordinary moments using well-known relationships. Plots of the skewness and kurtosis of the KBGG distribution as a function of c for selected values of a and b for $\alpha = 0.5$, $\beta = 1.0$ and $k = 2.0$ are displayed in Figures 4.2 and 4.3, respectively. Figures 4.2a and 4.2b immediately indicate that the additional parameter c promotes high levels of asymmetry.

(a) (b)

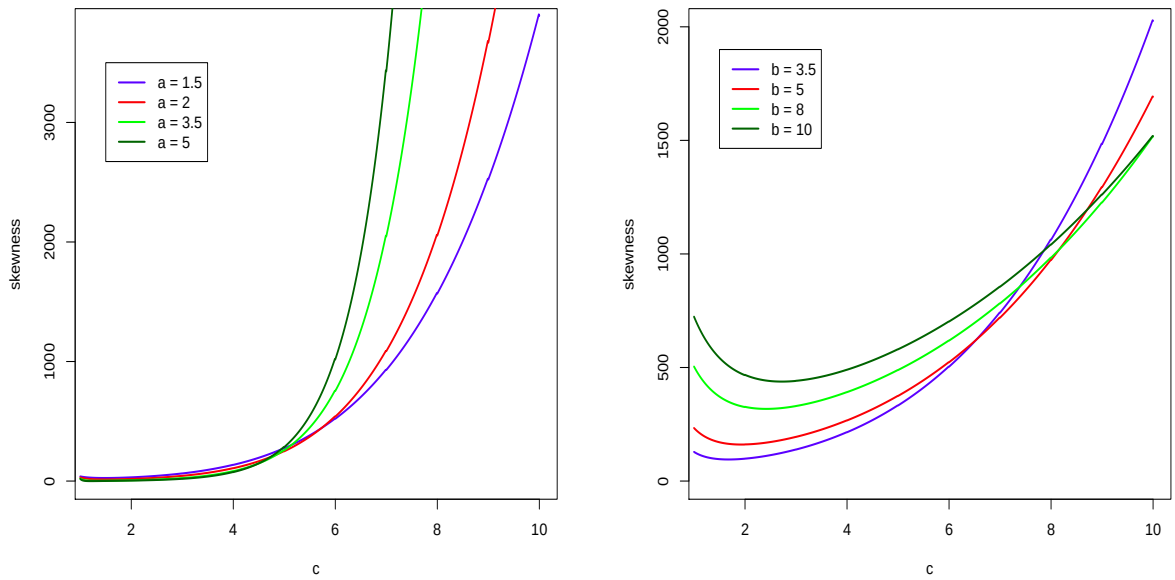


Figure 4.2 – Skewness of the KBGG distribution as a function of c for some values of a and b for $\alpha = 0.5$, $\beta = 1.0$ and $k = 2.0$ (a) $b = 2.0$ and (b) $a = 1.2$

Further, we provide a representation for the moment generating function (mgf) of X, say $M(t) = \mathrm{E}[\exp(tX)]$, which is obtained as a linear combination of GG generating functions.

(a) (b)

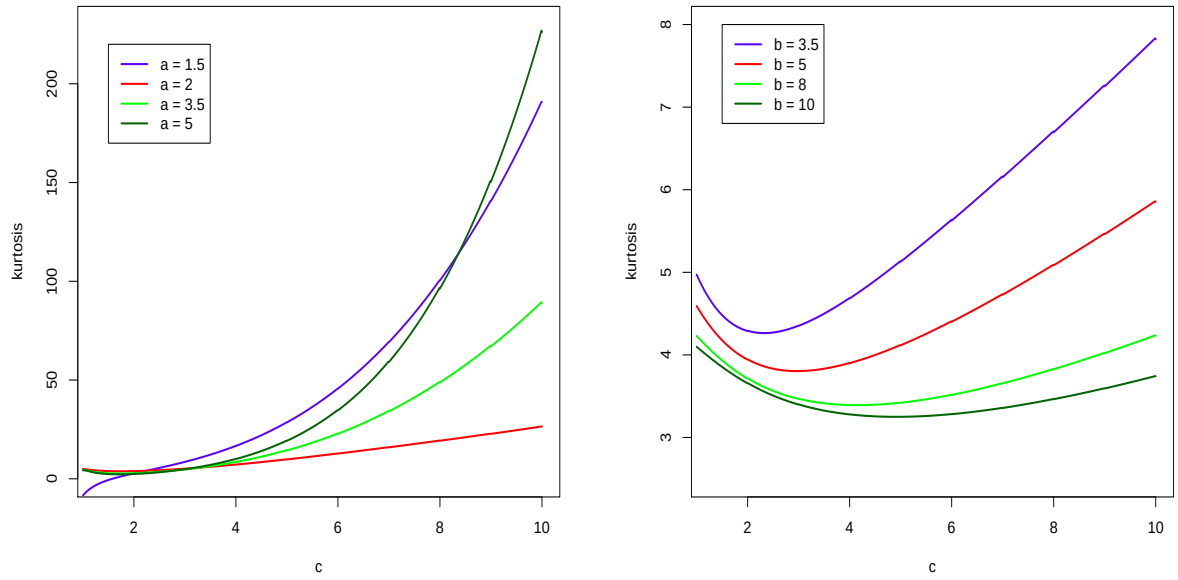


Figure 4.3 – Kurtosis of the KBGG distribution as a function of c for some values of a and b for $\alpha = 0.5$, $\beta = 1.0$ and $k = 2.0$ (a) $b = 2.0$ and (b) $a = 1.2$

From equation (4.13) , we have

$$\begin{aligned} M(t) &= \int_0^\infty \exp(tx)\, f(x)\, dx \\ &= \sum_{r,m=0}^{\infty} \eta_{r,m}\; M_{\alpha,\beta,k^\star}(t), \end{aligned} \tag{4.17}$$

where $M_{\alpha,\beta,k^\star}(t)$ denotes the mgf of the GG$(\alpha, \beta, k^\star)$ distribution.

From (4.17), we derive $M_{\alpha,\beta,k^\star}(t)$ as

$$M_{\alpha,\beta,k^\star}(t) = \frac{\beta}{\alpha\Gamma(k^\star)} \int_0^\infty \exp(tx) \left(\frac{x}{\alpha}\right)^{\beta k^\star - 1} \exp\left[-\left(\frac{x}{\alpha}\right)^\beta\right] dx.$$

Using the exponential expansion and replacing $u = \left(\frac{x}{\alpha}\right)^\beta$ in last integral, $M_{\alpha,\beta,k^\star}(t)$ reduces to

$$M_{\alpha,\beta,k^\star}(t) = \frac{1}{\Gamma(k^\star)} \sum_{\nu=0}^{\infty} \frac{(\alpha t)^\nu}{\nu!} \int_0^\infty u^{\frac{\nu}{\beta}+k^\star-1} \exp(-u)\, du. \tag{4.18}$$

Calculating the integral in (4.18), we obtain

$$M_{\alpha,\beta,k^\star}(t) = \frac{1}{\Gamma(k^\star)} \sum_{\nu=0}^{\infty} \Gamma\left(\frac{\nu}{\beta} + k^\star\right) \frac{(\alpha t)^\nu}{\nu!}. \tag{4.19}$$

Consider the Wright generalized hypergeometric function defined by

$${}_p\Psi_q \left[\begin{array}{c} (\alpha_1, A_1), \ldots, (\alpha_p, A_p) \\ (\beta_1, B_1), \ldots, (\beta_q, B_q) \end{array} ; x \right] = \sum_{n=0}^{\infty} \frac{\prod_{j=1}^{p} \Gamma(\alpha_j + A_j n)}{\prod_{j=1}^{q} \Gamma(\beta_j + B_j n)} \frac{x^n}{n!}. \tag{4.20}$$

Combining (4.19) and (4.20), we can rewrite the mgf of the GG distribution as

$$M_{\alpha,\beta,k^\star}(t) = \frac{1}{\Gamma(k^\star)}\, {}_1\Psi_0 \left[\begin{array}{c} (k^\star, \beta^{-1}) \\ - \end{array} ; \alpha t \right], \tag{4.21}$$

provided that $\beta > 1$.

The KBGG generating function follows by inserting (4.21) in (4.17). For $\beta > 1$, we obtain

$$M(t) = \sum_{r,m=0}^{\infty} \eta_{r,m}\; {}_1\Psi_0 \left[\begin{array}{c} (k^\star, \beta^{-1}) \\ - \end{array} ; \alpha t \right]. \tag{4.22}$$

Equations (4.16) and (4.22) are the main results of this section. The mgf of any KBGG submodel, as those discussed in Section 4.2, can be calculated immediately from (4.22) by substitution of known parameters.

4.5 Incomplete Moments

The answers to many important questions in economics require more than just knowing the mean of the distribution, but its shape as well. This is obvious not only in the study of econometrics but in other areas as well. Incomplete moments of the income distribution form natural building blocks for measuring inequality, for example, the Lorenz curve, Pietra and Gini measures of inequality all depend upon the incomplete moments of the income distribution. The sth incomplete moment of X is defined by $m_s(y) = E(X^s \mid X < y) = \int_0^y x^s\, f(x)dx$. Here, we propose two methods to calculate the KBGG incomplete moments. From the linear combination (4.13)

$$m_s(y) = \sum_{r,m=0}^{\infty} \eta_{r,m}\; t_s^\star(y), \tag{4.23}$$

where $t_s^\star(y) = \int_0^y x^s\, g_{\alpha,\beta,k^\star}(x)dx$ denotes the sth incomplete moment of the GG distribution with parameters α, β and $k^\star$ given by

$$t_s^\star(y) \;=\; \frac{\beta}{\alpha\,\Gamma(k^\star)} \int_0^y x^s \left(\frac{x}{\alpha}\right)^{\beta k^\star - 1} \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right] dx. \tag{4.24}$$

Calculating the integral in (4.24), $t_s^\star(y)$ reduces to

$$t_s^\star(y) \;=\; \alpha^s\, \frac{\gamma\left(k(r+1)+m+s/\beta, (y/\alpha)^\beta\right)}{\Gamma(k(r+1)+m)}.$$

Substituting the last equation in (4.23), we obtain the sth incomplete moment of X given by

$$m_s(y) = \alpha^s \sum_{r,m=0}^{\infty} \eta_{r,m}\, \frac{\gamma\left(k(r+1)+m+s/\beta, (y/\alpha)^\beta\right)}{\Gamma(k(r+1)+m)}. \tag{4.25}$$

4.6 Mean Deviations

We can derive the mean deviations about the mean μ_1' (δ_1) and about the median M (δ_2) in terms of the first incomplete moment. The median is obtained by inverting $F(M) = K \int_0^{\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^\beta\right]} t^{a-1}\,(1-t)^{b-1}\,\mathrm{e}^{-ct}\,dt = 1/2$ numerically. They can be expressed as

$$\delta_1 = 2\big[\mu_1'\,F(\mu_1') - m_1(\mu_1')\big] \qquad \text{and} \qquad \delta_2 = \mu_1' - 2\,m_1(M),$$

where $m_1(\cdot)$ is the first incomplete moment of X given by (4.25) with $s = 1$. We have

$$m_1(\omega) = \alpha \sum_{r,m=0}^{\infty} \eta_{r,m}\, \frac{\gamma\left(k(r+1)+m+1/\beta, (\omega/\alpha)^\beta\right)}{\Gamma(k(r+1)+m)}. \tag{4.26}$$

The measures δ_1 and δ_2 are immediately calculated from (4.26) by setting $\omega = \mu_1'$ and $\omega = M$, respectively.

Bonferroni and Lorenz curves are useful in fields such as reliability, economics, demography, insurance and medicine. For the KBGG distribution, these curves can be calculated (for given $0 < \pi < 1$) from $B(\pi) = (\pi\,\mu_1')^{-1}\,m_1(q)$ and $L(\pi) = (\mu_1')^{-1}\,m_1(q)$, respectively, where $\mu_1' = \mathrm{E}(X)$, $q = F^{-1}(\pi)$ can be computed for a given probability π by inverting (2.2) numerically, when $G(x;\boldsymbol{\gamma})$ is the cdf of the GG distribution. These measures are determined from equation (4.26).

4.7 Rényi Entropy

The entropy of a random variable is a measure of variation of the uncertainty. Entropy has been used in various situations in science and engineering and numerous measures of entropy have been studied and compared in the literature. The Rényi entropy is defined by

$$\mathcal{J}_R(\xi) = \frac{1}{1-\xi}\,\log\left[\int f^\xi(x)dx\right],\ \xi > 0 \ \text{ and } \ \xi \neq 1.$$

Note that the integral above is obtained from (4.5) as

$$\begin{aligned} I(\xi) = \int_0^\infty f^\xi(x)\,dx \;&=\; \left(\frac{K\,\beta}{\alpha\,\Gamma(k)}\right)^\xi \int_0^\infty \left(\frac{x}{\alpha}\right)^{\xi(\beta k-1)} \exp\left[-\xi\left(\frac{x}{\alpha}\right)^\beta\right]\,\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^\beta\right]^{\xi(a-1)} \\ &\quad\times\left\{1-\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^\beta\right]\right\}^{\xi(b-1)} \exp\left\{-c\,\xi\,\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^\beta\right]\right\}dx. \end{aligned} \tag{4.27}$$

Using the exponential and binomial expansions in (4.27), we obtain

$$\begin{aligned} I(\xi) \;&=\; \left[\frac{K\,\beta}{\alpha\,\Gamma(k)}\right]^\xi \sum_{i,j=0}^{\infty} \frac{(-1)^{i+j}}{i!\,(c\,\xi)^{-i}} \binom{\xi(b-1)}{j} \\ &\quad\times\int_0^\infty \left(\frac{x}{\alpha}\right)^{\xi(\beta k-1)} \exp\left[-\xi\left(\frac{x}{\alpha}\right)^\beta\right]\left\{\gamma_1\left[k,\left(\frac{x}{\alpha}\right)^\beta\right]\right\}^{\xi(a-1)+1+j} dx. \end{aligned} \tag{4.28}$$

Noting that $\xi > 0$ and $a > 0$ are real non-integers, we can expand $\left\{\gamma_1\left[k, \left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^{\xi(a-1)+1+j}$ as

$$\begin{aligned}\left\{\gamma_1\left[k, \left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^{\xi(a-1)+1+j} &= \left\{1-\left[1-\gamma_1\left[k, \left(\frac{x}{\alpha}\right)^{\beta}\right]\right]\right\}^{\xi(a-1)+1+j} \\ &= \sum_{p=0}^{\infty}(-1)^p\binom{\xi(a-1)+1+j}{p}\left\{1-\gamma_1\left[k, \left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^{p}\end{aligned}$$

and then

$$\left\{\gamma_1\left[k, \left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^{\xi(a-1)+1+j} = \sum_{p=0}^{\infty}\sum_{r=0}^{p}(-1)^{p+r}\binom{\xi(a-1)+1+j}{p}\binom{p}{r}\left\{\gamma_1\left[k, \left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^{r}.$$

Replacing $\sum_{p=0}^{\infty}\sum_{r=0}^{p}$ by $\sum_{r=0}^{\infty}\sum_{p=r}^{\infty}$, the quantity, $I(\xi)$ can be rearranged in the form

$$\begin{aligned}I(\xi) &= \left(\frac{K\,\beta}{\alpha\,\Gamma(k)}\right)^{\xi}\sum_{r=0}^{\infty}\rho_r \\ &\quad\times\int_0^{\infty}\left(\frac{x}{\alpha}\right)^{\xi(\beta k-1)}\exp\left[-\xi\left(\frac{x}{\alpha}\right)^{\beta}\right]\left\{\gamma_1\left[k, \left(\frac{x}{\alpha}\right)^{\beta}\right]\right\}^{r}dx, \qquad (4.29)\end{aligned}$$

where

$$\rho_r = \sum_{i,j=0}^{\infty}\sum_{r=p}^{\infty}\frac{(-1)^{i+j+p+r}}{i!\,(c\,\xi)^{-i}}\binom{\xi(b-1)}{j}\binom{\xi(a-1)+i+j}{p}\binom{p}{r}. \qquad (4.30)$$

Using expansion (4.12) in (4.29), we obtain

$$I(\xi) = \left[\frac{K\,\beta}{\alpha\,\Gamma(k)}\right]^{\xi}\sum_{r,m=0}^{\infty}\frac{d_{r,m}}{\Gamma(k)^r}\,\rho_r\int_0^{\infty}\left(\frac{x}{\alpha}\right)^{\beta[k(r+\xi)+m]-\xi}\exp\left[-\xi\left(\frac{x}{\alpha}\right)^{\beta}\right]dx. \qquad (4.31)$$

Calculating the integral in (4.31), we have

$$I(\xi) = \frac{K^{\xi}\,\beta^{\xi-1}}{\Gamma(k)^{\xi}\,\alpha^{\xi-1}}\sum_{r,m=0}^{\infty}\rho^{\star}_{r,m}\,\Gamma\left(k(r+\xi)+m+\frac{(1-\xi)}{\beta}\right),$$

where

$$\rho^{\star}_{r,m} = \frac{d_{r,m}\;\rho_r}{\Gamma(k)^r\,\xi^{k(r+\xi)+m-(\xi+1)/\beta}}.$$

Finally, the Rényi entropy reduces to

$$\begin{aligned}\mathcal{J}_R(\xi) &= (1-\xi)^{-1}\left\{\xi\left[\log(K)-\log\Gamma(k)\right]+(\xi-1)\left[\log(\beta)-\log(\alpha)\right]\right. \\ &\quad\left.+\log\left[\sum_{r,m=0}^{\infty}\rho^{\star}_{r,m}\,\Gamma\left(k(r+\xi)+m+\frac{(1-\xi)}{\beta}\right)\right]\right\}.\end{aligned}$$

4.8 Order Statistics

Moments of order statistics play an important role in quality control testing and reliability, where a practitioner needs to predict the failure of future items based on the times of a few early failures. These predictors are often based on moments of order statistics. We now derive an explicit expression for the density of the ith order statistics $X_{i:n}$, say $f_{i:n}(x)$, in a random sample of size n from $X \sim \text{KBGG}(a, b, c, \alpha, \beta, k)$. It is well-known that

$$f_{i:n}(x) = \frac{n!\, f(x)}{(i-1)!\,(n-1)!}\, F(x)^{i-1}\,[1 - F(x)]^{n-i}\,, \tag{4.32}$$

and using the binomial expansion in (4.32), we have

$$f_{i:n}(x) = \frac{n!\, f(x)}{(i-1)!\,(n-1)!} \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} F(x)^{i+j-1}. \tag{4.33}$$

We now demonstrate that $f_{i:n}(x)$ can be written as a linear combination of GG densities. First, we provide an expansion for the cdf of the KBGG distribution. (PESCIM et al., 2012) demonstrated that

$$F_{\mathcal{KBG}}(x) = \sum_{r=0}^{\infty} b_r\, G(x; \boldsymbol{\gamma})^r, \tag{4.34}$$

where the coefficient $b_r = \sum_{i,j=0}^{\infty} \sum_{k=r}^{\infty} t_{i,j,k,r}$ denotes a sum of constants and $t_{i,j,k,r}$ is defined in (4.7).

Equation (4.34) gives the cumulative function for any KB-G distribution as an infinite weighted power series of cdf's of the baseline distribution. Inserting (4.2) in (4.34), we have the KBGG cumulative function expanded as

$$F(x) = \sum_{r=0}^{\infty} b_r \left\{ \gamma_1 \left[k, \left(\frac{x}{\alpha}\right)^{\beta} \right] \right\}^r. \tag{4.35}$$

Combining (4.7) and (4.35), the pdf of the ith order statistic, $X_{i:n}$, can be expressed as

$$f_{i:n}(x) = \sum_{j=0}^{n-i} \frac{n!\,(-1)^j}{(i-1)!(n-i)!} \binom{n-i}{j} \left[\sum_{r=0}^{\infty} c_r\, v_{r+1}(x) \right] \left[\sum_{r=0}^{\infty} b_r \left\{ \gamma_1 \left[k, \left(\frac{x}{\alpha}\right)^{\beta} \right] \right\}^r \right]^{i+j-1} \tag{4.36}$$

Applying the expression (4.10) in (4.36), we have

$$\left[\sum_{r=0}^{\infty} b_r \left\{ \gamma_1 \left[k, \left(\frac{x}{\alpha}\right)^{\beta} \right] \right\}^r \right]^{i+j-1} = \sum_{r=0}^{\infty} d^{\star}_{i+j-1,r} \left\{ \gamma_1 \left[k, \left(\frac{x}{\alpha}\right)^{\beta} \right] \right\}^r, \tag{4.37}$$

where $d^{\star}_{i+j-1,r}$ is given by (4.11). Inserting (4.12) into (4.37), we obtain

$$\left[\sum_{r=0}^{\infty} b_r \left\{ \gamma_1 \left[k, \left(\frac{x}{\alpha}\right)^{\beta} \right] \right\}^r \right]^{i+j-1} = \sum_{r,m=0}^{\infty} \frac{d_{r,m}\, d^{\star}_{i+j-1,r}}{\Gamma(k)^r} \left(\frac{x}{\alpha}\right)^{\beta(kr+m)}. \tag{4.38}$$

Substituting (4.13) and (4.38) in (4.36), $f_{i:n}(x)$ reduces to

$$f_{i:n}(x) = \sum_{r,m=0}^{\infty}\sum_{j=0}^{n-i} \frac{(-1)^j \binom{n-i}{j} n!\, c_r\, e(r,m)\, d_{r,m}\, d^{\star}_{i+j-1,r}}{(i-1)!\,(n-i)!\,\Gamma[k(r+1)+m]\,\Gamma(k)^r}\, g_{\alpha,\beta,k^{\star\star}}(x), \tag{4.39}$$

where $k^{\star\star} = k(2r+1)+2m$ and

$$g_{\alpha,\beta,k^{\star\star}}(x) = \frac{\beta}{\alpha\,\Gamma(k^{\star\star})}\left(\frac{x}{\alpha}\right)^{\beta k^{\star\star}-1} \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]$$

denotes the GG$(\alpha,\beta,k^{\star\star})$ density function.

Equation (4.39) reveals that the density function of the KBGG order statistics is an infinite linear combination of GG densities. Hence, ordinary moments of order statistics can be determined directly from those quantities of the GG distribution.

For $a>0$ and $b>0$ real non-integer, the sth moment of $X_{i:n}$ comes from (4.39) as

$$\mathrm{E}(X^s_{i:n}) = \sum_{r,m=0}^{\infty}\sum_{j=0}^{n-i} \frac{(-1)^j \binom{n-i}{j} n!\, c_r\, e(r,m)\, d_{r,m}\, d^{\star}_{i+j-1,r}}{(i-1)!\,(n-i)!\,\Gamma[k(r+1)+m]\,\Gamma(k)^r}\, \mathrm{E}(X^s_{r,m}), \tag{4.40}$$

where $X_{r,m} \sim$ GG$(\alpha,\beta,k^{\star\star})$. Equation (4.40) gives the sth moment of the KBGG order statistics, which is the main result of this section.

Based upon these moments, we can derive expansions for the L-moments as infinite weighted linear combinations of suitable KBGG means. The L-moments are analogous to the ordinary moments but can be estimated by linear combinations of order statistics. They are linear functions of expected order statistics defined by (HOSKING, 1990) and are relatively robust to the effects of outliers.

4.9 Inference and Estimation

4.9.1 Classical Inference

The estimation of the model parameters of the KBGG distribution will be performed by the maximum likelihood method. Let $X=(X_1,\ldots,X_n)$ be a random sample of this distribution with unknown parameter vector $\boldsymbol{\theta}=(a,b,c,\alpha,\beta,k)^T$. The total log-likelihood function for $\boldsymbol{\theta}$ is

$$\begin{aligned}\ell(\boldsymbol{\theta}) &= n\log\left[\frac{K\,\beta}{\alpha\,\Gamma(k)}\right] + (\beta\,k-1)\sum_{i=1}^{n}\log\left(\frac{x_i}{\alpha}\right) - \sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta} \\ &\quad + (a-1)\sum_{i=1}^{n}\log\left\{\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\} + (b-1)\sum_{i=1}^{n}\log\left\{1-\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\} \\ &\quad - c\sum_{i=1}^{n}\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]. \end{aligned} \tag{4.41}$$

The elements of score vector are given by

$$
\begin{aligned}
U_\alpha(\boldsymbol{\theta}) &= -\frac{n}{\alpha} - \frac{n(\beta\, k-1)}{\alpha} + \frac{\beta}{\alpha}\sum_{i=1}^{n} u_i - \frac{\beta(a-1)}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\gamma(k,u_i)} \\
&\quad + \frac{\beta(b-1)}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\Gamma(k)-\gamma(k,u_i)} + \frac{\beta\, c}{\alpha\Gamma(k)}\sum_{i=1}^{n} v_i, \\
U_\beta(\boldsymbol{\theta}) &= \frac{n}{\beta} + \sum_{i=1}^{n} u_i^{1/\beta} - \sum_{i=1}^{n} u_i\, s_i + (a-1)\sum_{i=1}^{n}\frac{v_i\, s_i}{\gamma(k,u_i)} \\
&\quad + (1-b)\sum_{i=1}^{n}\frac{v_i\, s_i}{\Gamma(k)-\gamma(k,u_i)} - \frac{c}{\Gamma(k)}\sum_{i=1}^{n} v_i\, s_i, \\
U_k(\boldsymbol{\theta}) &= -\psi(k) + \beta\sum_{i=1}^{n} u_i^{1/\beta} - n\,(a-1)\,\psi(k) + (a-1)\sum_{i=1}^{n}\frac{\gamma'(k,u_i)|_k}{\gamma(k,u_i)} \\
&\quad + (1-b)\sum_{i=1}^{n}\frac{\gamma'(k,u_i)|_k}{\Gamma(k)-\gamma(k,u_i)} + \psi(k)\,(b-1)\sum_{i=1}^{n}\frac{\gamma(k,u_i)}{\Gamma(k)-\gamma(k,u_i)} \\
&\quad - \frac{c}{\Gamma(k)}\sum_{i=1}^{n}\gamma'(k,u_i)|_k + \frac{c\,\psi(k)}{\Gamma(k)}\sum_{i=1}^{n}\gamma(k,u_i), \\
U_a(\boldsymbol{\theta}) &= \frac{n}{K}\frac{\partial K}{\partial a} + \sum_{i=1}^{n}\log[\gamma_1(k,u_i)]\,, U_b(\boldsymbol{\theta}) = \frac{n}{K}\frac{\partial K}{\partial b} + \sum_{i=1}^{n}\log[1-\gamma_1(k,u_i)] \\
&\quad \text{and } U_c(\boldsymbol{\theta}) = \frac{n}{K}\frac{\partial K}{\partial c} - \sum_{i=1}^{n}\gamma_1(k,u_i),
\end{aligned}
$$

where $u_i = \left(\frac{x_i}{\alpha}\right)^\beta$, $v_i = \left(\frac{x_i}{\alpha}\right)^{\beta k}\exp\left[\left(\frac{x_i}{\alpha}\right)^\beta\right]$, $s_i = \log\left(\frac{x_i}{\alpha}\right)$, $\gamma'(k,u_i)|_k = \sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\,J(u_i, k+n-1, 1)$, $\psi(.)$ is the digamma function and $J(u_i, k+n-1, 1)$ is defined in Appendix C. The partial derivatives of K with respect to a, b and c are calculated in chapter 1 of this thesis.

The maximum likelihood estimate (MLE) $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained numerically from the nonlinear equations $\mathbf{U_a}(\boldsymbol{\theta}) = \mathbf{U_b}(\boldsymbol{\theta}) = \mathbf{U_c}(\boldsymbol{\theta}) = \mathbf{U}_\alpha(\boldsymbol{\theta}) = \mathbf{U}_\beta(\boldsymbol{\theta}) = \mathbf{U_k}(\boldsymbol{\theta}) = \mathbf{0}$. For interval estimation and hypothesis testing on the model parameters, we require the 6×6 observed information matrix $\mathbf{J} = \mathbf{J}(\boldsymbol{\theta})$ whose elements are given in Appendix C. Under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, the asymptotic distribution of $\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}$ is $\mathrm{N}_6(\mathbf{0}, \mathbf{K}(\widehat{\boldsymbol{\theta}})^{-1})$ where $\mathbf{K}(\boldsymbol{\theta})$ is the expected information matrix. This matrix can be replaced by $\mathbf{J}(\widehat{\boldsymbol{\theta}})$, i.e., the observed information matrix evaluated at $\widehat{\boldsymbol{\theta}}$. The estimated asymptotic multivariate normal $\mathrm{N}_6(\mathbf{0}, \mathbf{J}(\widehat{\boldsymbol{\theta}})^{-1})$ distribution of $\widehat{\boldsymbol{\theta}}$ can be used to construct approximate confidence regions for the parameters and for the hazard and survival functions. We can compute the maximum values of the unrestricted and restricted log-likelihoods to construct LR (likelihood ratio) statistics for testing some sub-models of the KBGG distribution. For example, we may use LR statistics to check whether the fit using the KBGG distribution is statistically "superior" to a fit using the KBGHN, KBW, BGG and GG distributions for a given data set.

4.9.2 Bayesian Inference

Since we have no prior information from historical data or from previous experiment, we assign conjugate but weakly informative prior distributions to the parameters. Since we assumed informative (but weakly) prior distribution, the posterior distribution is a well-defined proper distribution. We assume that the parameters $(a, b, c, \alpha, \beta \text{ and } k)$ have independent priors and consider that the joint prior distribution of all unknown parameters has a density function given by

$$\pi(a, b, c, \alpha, \beta, k) \propto \pi(a) \times \pi(b) \times \pi(c) \times \pi(\alpha) \times \pi(\beta) \times \pi(k), \tag{4.42}$$

where, $a \sim \Gamma(a_1, b_1)$, a_1 and b_1 known; $b \sim \Gamma(a_2, b_2)$, a_2 and b_2 known; $c \sim \mathrm{N}(\mu_0, \sigma_0^2)$, μ_0 and σ_0^2 known; $\alpha \sim \Gamma(a_3, b_3)$, a_3 and b_3 known; $\beta \sim \Gamma(a_4, b_4)$, a_4 and b_4 known; $k \sim \Gamma(a_5, b_5)$, a_5 and b_5 known; where $\Gamma(a_i, b_i)$ denotes the gamma distribution with mean a_i/b_i, variance a_i/b_i^2 for $a_i > 0$ and $b_i > 0$, and $\mathrm{N}(\mu_0, \sigma_0^2)$ represents the normal distribution with mean μ_0, variance σ_0^2 for $\mu_0 \in \mathbb{R}$ and $\sigma_0^2 > 0$. We note that gamma and normal priors are most commonly used priors for positive and real-values parameters.

Combining the likelihood function (4.41) and the prior distribution (4.42), the joint posterior distribution for a, b, c, α, β and k reduces to

$$\begin{aligned}
\pi(a, b, c, \alpha, \beta, k|x) \quad \propto \quad & \left[\frac{K\beta}{\alpha\Gamma(k)}\right]^n \exp\left\{-c\sum_{i=1}^{n}\gamma_1\left[k, \left(\frac{x_i}{\alpha}\right)^{\beta}\right] - \sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta}\right\} \\
& \times \prod_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta k-1}\gamma_1\left[k, \left(\frac{x_i}{\alpha}\right)^{\beta}\right]^{a-1}\left\{1-\gamma_1\left[k, \left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\}^{b-1} \\
& \times \pi(a, b, c, \alpha, \beta, k).
\end{aligned} \tag{4.43}$$

The joint posterior density (4.43) is analytically intractable because the integration of the joint posterior density is not easy to perform. So, the inference can be based on MCMC simulation methods such as the Gibbs sampler and Metropolis-Hastings algorithm, which can be used to draw samples, from which features of the marginal distributions of interest can be inferred. In this direction, we first obtain the full conditional distributions of the unknown quantities given by

$$\pi(a|x, b, c, \alpha, \beta, k) \quad \propto \quad K^n \prod_{i=1}^{n}\gamma_1\left[k, \left(\frac{x_i}{\alpha}\right)^{\beta}\right]^{a-1} \times \pi(a),$$

$$\pi(b|x, a, c, \alpha, \beta, k) \quad \propto \quad K^n \prod_{i=1}^{n}\left\{1-\gamma_1\left[k, \left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\}^{b-1} \times \pi(b),$$

$$\pi(c|x, a, b, c, \alpha, \beta, k) \quad \propto \quad K^n \exp\left\{-c\sum_{i=1}^{n}\gamma_1\left[k, \left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\} \times \pi(c),$$

$$\begin{aligned}\pi(\alpha|x,a,b,c,\beta,k) \;\propto\; & \frac{1}{\alpha^n}\exp\left\{-c\sum_{i=1}^{n}\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]-\sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta}\right\}\\ & \times\prod_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta k-1}\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]^{a-1}\left\{1-\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\}^{b-1}\\ & \times\pi(\alpha),\end{aligned}$$

$$\begin{aligned}\pi(\beta|x,a,b,c,\alpha,k) \;\propto\; & \beta^n\exp\left\{-c\sum_{i=1}^{n}\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]-\sum_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta}\right\}\\ & \times\prod_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta k-1}\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]^{a-1}\left\{1-\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\}^{b-1}\\ & \times\pi(\beta)\end{aligned}$$

and

$$\begin{aligned}\pi(k|x,a,b,c,\alpha,\beta) \;\propto\; & \frac{1}{\Gamma(k)^n}\exp\left\{-c\sum_{i=1}^{n}\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\}\\ & \times\prod_{i=1}^{n}\left(\frac{x_i}{\alpha}\right)^{\beta k-1}\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]^{a-1}\left\{1-\gamma_1\left[k,\left(\frac{x_i}{\alpha}\right)^{\beta}\right]\right\}^{b-1}\\ & \times\pi(k).\end{aligned}$$

Since the full conditional distributions do not have explicit expressions, we require the use of the Metropolis-Hastings algorithm to generate the variables a, b, c, α, β and k for the KBGG distribution.

4.10 Applications

In this section, we use three real data sets which come from diverse fields such as actuarial sciences, environmental studies and engineering to compare the fits of the KBGG distribution with those of three sub-models (i.e. BGG, EGG and GG distributions) and also to the following non-nested model: the Kumaraswamy generalized gamma (KwGG) distribution (PASCOA; ORTEGA; CORDEIRO, 2011). In each case, the parameters are estimated by maximum likelihood and Bayesian methods (Section 4.9) using the statistical software R. The primary reason for choosing these data is that they allow us to show how in different fields it is necessary to have positively skewed distributions with non-negative support. Moreover, these data sets present different degrees of skewness and kurtosis.

(i) Minimum pension data set

It is important for the Mexican Institute of Social Security (IMSS) to study the distributional behaviour of the mortality of retired people on disability because it enables the calculation of long and short term financial estimation, such as the assessment of the reserve required to pay the "minimum pensions". The data set corresponding to 280 lifetimes (in years) of retired

women with temporary disabilities, which are incorporated in the Mexican insurance public system and who died during 2004 were reported and analyzed by (BALAKRISHNAN et al., 2009).

(ii) Ozone data set

These data were analyzed by (LEIVA; BARROS; PAULA, 2009) and correspond to daily ozone level measurements in New York in May-September, 1973, from the New York State Department of Conservation.

(ii) Conductor data set

Failures can occur in microcircuits because of the movement of atoms in the conductors in the circuit, this is referred to the electromigration. The data set refers to an accelerated life test of 59 conductors reported by (LAWLESS, 1982).

4.10.1 Maximum Likelihood Estimation

First, we give the MLEs (and the corresponding standard errors in parentheses) of the parameters and the values of the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) statistics. The smaller the values of these criteria, the better the fit. Note that over-parameterization is penalized in these criteria, so that the three additional parameters in the KBGG model do not necessarily lead to smaller values of the AIC and BIC statistics. Next, we perform LR tests for formal tests of the additional shape parameters. Finally, we provide histograms of the data sets to show a visual comparison of the KBGG fitted density functions.

In order to estimate the model parameters, we take the estimates of α, β and k from the fitted GG distribution as starting values for the numerical iterative procedure. Table 4.2 lists the MLEs (and the corresponding standard errors in parentheses) of the parameters and the values of the AIC and BIC statistics. The results indicate that the KBGG model has the smallest values of the statistics (AIC and BIC) among all fitted models. So, it could be chosen as the most suitable model. A comparison of the proposed distribution with some of its sub-models using LR statistics is shown in Table 4.3. The *p*-values indicate that the proposed model yields the best fit to the three data sets. This gives a clear evidence of the potential of the three parameters when modeling real data. In order to assess if the model is appropriate, Figure 4.4 displays histograms with estimated KBGG density functions for each data sets, respectively. We can conclude that the new distribution is a very suitable model to fit the three data sets.

4.10.2 Bayesian Analysis

For the three real data sets, the following independent priors were considered to perform the Metropolis-Hastings algorithm: $\alpha \sim \Gamma(0.001, 0.001)$, $\beta \sim \Gamma(0.001, 0.001)$, $k \sim \Gamma(0.001, 0.001)$, $a \sim \Gamma(0.001, 0.001)$, $b \sim \Gamma(0.001, 0.001)$ and $c \sim N(0, 1000)$, so that we have vague prior distributions. Considering these prior density functions, we generate two parallel independent runs of the Metropolis-Hastings with size 300.000 for each parameter, disregarding the first 30.000 iterations to eliminate the effect of the initial values and, to avoid correlation problems, we

Table 4.2 – MLEs of the model parameters for the three data sets and the corresponding AIC and BIC statistics

Data	Model	α	β	k	a	b	c	AIC	BIC
(i)	KBGG	7.4489	1.8564	35.6798	0.1621	0.4310	-0.9485	2104.0	2125.8
		(0.0099)	(0.0099)	(0.01003)	(0.00106)	(0.00106)	(0.00107)		
	BGG	32.7248	1.9705	3.9984	0.9872	3.8361	0	2115.1	2133.2
		(0.0209)	(0.0019)	(0.00210)	(0.00238)	(0.00565)	(-)		
	EGG	33.8389	2.8185	3.1002	0.9056	1	0	2113.3	2127.9
		(0.3373)	(0.0177)	(0.00007)	(0.0034)	(-)	(-)		
	GG	34.1730	2.8048	2.8693	1	1	0	2111.3	2122.2
		(0.10391)	(0.00170)	(0.00041)	(-)	(-)	(-)		
	Model	α	τ	k	λ	ϕ	-	AIC	BIC
	KwGG	33.5340	1.4296	2.6525	2.2376	9.2101	-	2114.8	2133.0
		(0.23109)	(0.00811)	(0.00110)	(0.00237)	(0.00677)	(-)		
(ii)	KBGG	3.0409	1.0760	20.2422	0.0807	0.1598	-0.2154	1067.1	1083.6
		(0.0089)	(0.0009)	(0.01002)	(0.0098)	(0.01009)	(0.0009)		
	BGG	4.0775	1.1376	17.5934	0.0923	0.1749	0	1087.7	1101.5
		(0.05950)	(0.00585)	(0.0810)	(0.00964)	(0.00053)	(-)		
	EGG	3.7038	0.6370	4.9592	0.7285	1	0	1090.2	1101.2
		(0.00345)	(0.00122)	(0.0061)	(0.00541)	(-)	(-)		
	GG	3.1291	0.5924	4.3440	1	1	0	1088.3	1096.6
		(0.00104)	(0.00070)	(0.0076)	(-)	(-)	(-)		
	Model	α	τ	k	λ	ϕ	-	AIC	BIC
	KwGG	0.6009	0.5508	11.2001	0.4059	0.7496	-	1091.9	1105.7
		(0.02150)	(0.00175)	(0.09112)	(0.00263)	(0.00148)	(-)		
(iii)	KBGG	7.0954	8.1282	2.0878	0.3840	0.1030	2.7935	221.7	234.1
		(0.00991)	(0.0069)	(0.01001)	(0.00981)	(0.01002)	(0.0008)		
	BGG	4.720	2.0391	3.0389	1.3445	2.1157	0	232.6	243.0
		(0.03850)	(0.0165)	(0.0110)	(0.00842)	(0.0083)	(-)		
	EGG	0.0200	0.5933	28.3765	2.3890	1	0	234.0	242.3
		(0.00005)	(0.00346)	(0.0461)	(0.00247)	(-)	(-)		
	GG	4.1439	2.3300	3.6446	1	1	0	228.6	234.9
		(0.00265)	(0.00270)	(0.00416)	(-)	(-)	(-)		
	Model	α	τ	k	λ	ϕ	-	AIC	BIC
	KwGG	4.1410	1.8808	3.4611	1.3199	2.1071	-	232.6	243.0
		(0.02421)	(0.00314)	(0.09541)	(0.00105)	(0.00112)	(-)		

consider a spacing of size 10, obtaining a sample of size 27.000 from each chain. To monitor the convergence of the Metropolis-Hastings algorithm, we perform the methods suggested by (COWLES; CARLIN, 1996) using the between and within sequence information, following the approach developed in (GELMAN; RUBIN, 1992) to obtain the potential scale reduction, $\widehat{R}$. In all cases, these values were close to one, indicating the convergence of the chain.

The approximate posterior marginal density functions for the parameters are displayed in Figures 4.5, 4.6 and 4.7 for the first, second and third data sets, respectively. In Table 4.4, we report posterior summaries for the parameters of the KBGG model for the three data sets. We note that the values for the *a posterior* means (Table 4.4) are quite close (as expected) to the

Table 4.3 – LR statistics for the three data sets

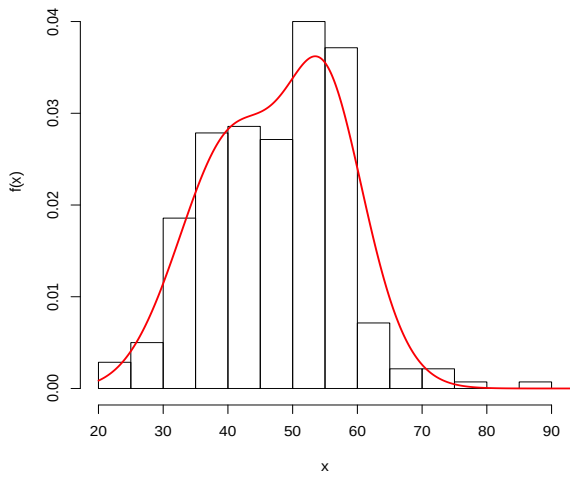
Data	Model	Hypotheses	Statistic w	p-value
D1	KBGG vs BGG	$H_0 : c = 0$ vs $H_1 : H_0$ is false	13.10	0.00029
	KBGG vs EGG	$H_0 : c = 0$ and $b = 1$ vs $H_1 : H_0$ is false	13.36	0.00124
	KBGG vs GG	$H_0 : a = b = 1$ and $c = 0$ vs $H_1 : H_0$ is false	13.38	0.00387
D2	KBGG vs BGG	$H_0 : c = 0$ vs $H_1 : H_0$ is false	22.57	< 0.0001
	KBGG vs EGG	$H_0 : c = 0$ and $b = 1$ vs $H_1 : H_0$ is false	27.03	< 0.0001
	KBGG vs GG	$H_0 : a = b = 1$ and $c = 0$ vs $H_1 : H_0$ is false	27.17	< 0.0001
D3	KBGG vs BGG	$H_0 : c = 0$ vs $H_1 : H_0$ is false	12.92	0.00032
	KBGG vs EGG	$H_0 : c = 0$ and $b = 1$ vs $H_1 : H_0$ is false	16.32	0.00028
	KBGG vs GG	$H_0 : a = b = 1$ and $c = 0$ vs $H_1 : H_0$ is false	12.96	0.00471

MLEs obtained for the KBGG model given in Table 4.2. "SD" denotes the standard deviation from the posterior distributions of the parameters and "HPD" denotes the 95% highest posterior density intervals.

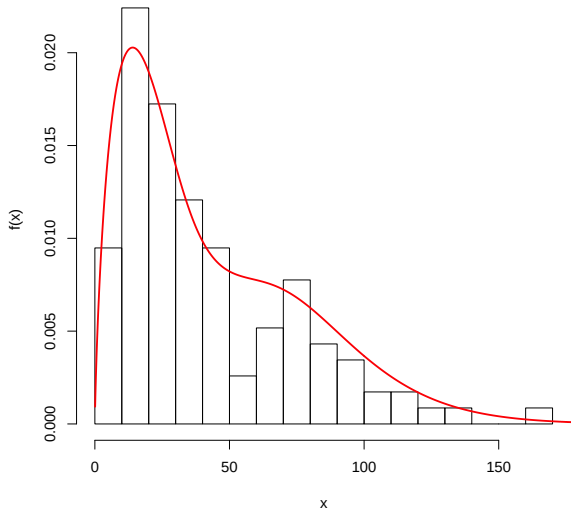
Table 4.4 – Posterior summaries for the parameters from the KBGG model for the three data sets

D1				
Parameter	Mean	SD	HPD (95%)	$\widehat{R}$
α	7.4399	0.0099	(7.4201; 7.4590)	1.0005
β	1.8499	0.0099	(1.8301; 1.8689)	1.0014
k	35.6701	0.01003	(35.6500; 35.6892)	1.0004
a	0.1594	0.0098	(0.1407; 0.1792)	1.0003
b	0.4301	0.01008	(0.4103; 0.4499)	0.9997
c	-0.9401	0.0099	(-0.9594; -0.9204)	1.0002
D2				
Parameter	Mean	SD	HPD (95%)	$\widehat{R}$
α	3.0399	0.0099	(3.0201; 3.0590)	1.0002
β	1.0599	0.0009	(1.0580; 1.0618)	1.0009
k	20.2401	0.01002	(20.2200; 20.2592)	1.0002
a	0.0798	0.0098	(0.0611; 0.0997)	1.0011
b	0.1502	0.01009	(0.1304; 0.1700)	0.9996
c	-0.2100	0.0009	(-0.2119; -0.2080)	1.0005
D3				
Parameter	Mean	SD	HPD (95%)	$\widehat{R}$
α	7.0599	0.0099	(7.0401; 7.0790)	1.0001
β	8.1499	0.0099	(8.1301; 8.1689)	1.0011
k	2.0694	0.0100	(2.0491; 2.0883)	1.0002
a	0.3697	0.0098	(0.3510; 0.3896)	0.9997
b	0.0998	0.0100	(0.0798; 0.1192)	0.9998
c	2.7998	0.0009	(2.7980; 2.8019)	1.0006

Histogram and estimated KBGG pdf for D1



Histogram and estimated KBGG pdf for D2



Histogram and estimated KBGG pdf for D3

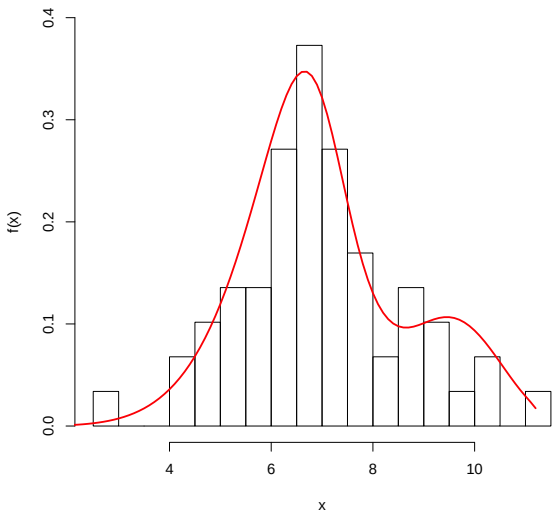


Figure 4.4 – Histogram with estimated KBGG density function for the indicated data sets

4.11 Concluding Remarks

We introduce the Kummer beta generalized gamma (KBGG) distribution with three additional shape parameters because of the wide usage of the GG distribution and the fact that the current generalization provides extensions to its continuous extension to still more complex situations. The new distribution unifies more than 32 distributions and yields a general overview of these distributions for theoretical studies. In fact, the KBGG distribution (4.5) generalizes the Weibull, gamma, exponentiated Weibull, exponentiated gamma, beta Weibull, beta gamma, Kummer beta Weibull and Kummer beta gamma distributions and other important lifetime mo-

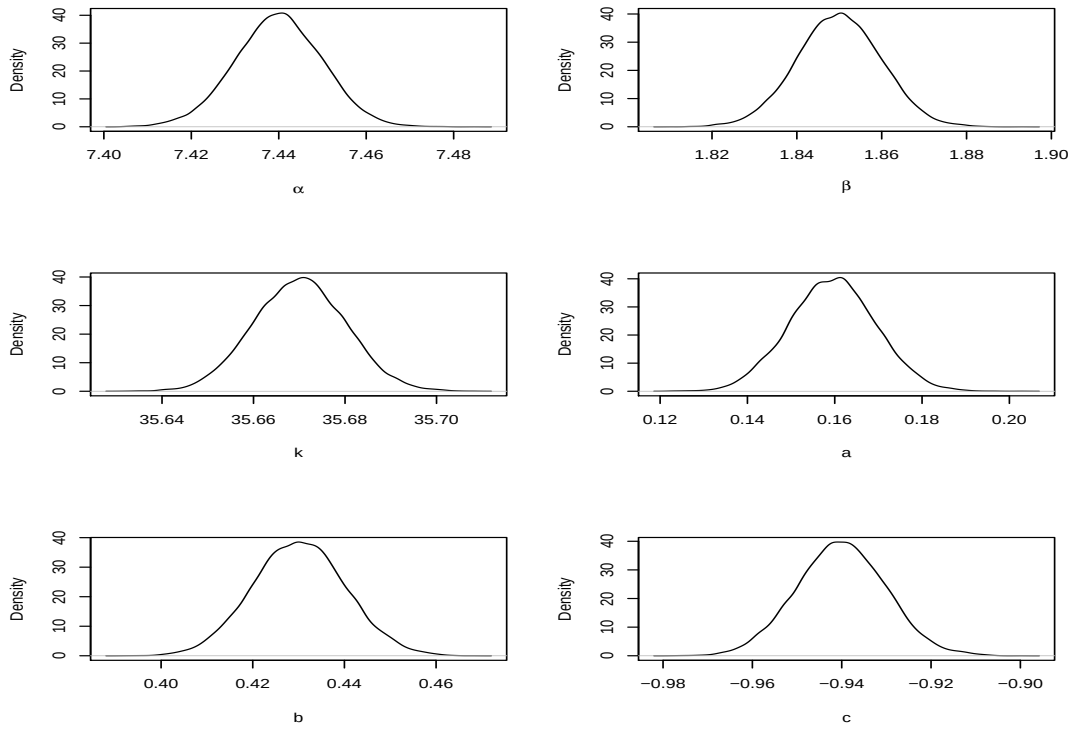


Figure 4.5 – Approximate posterior marginal densities for the parameters of the KBGG model for the first data set

dels. The KBGG density function can be expressed as a linear combination of GG density functions that allows us to derive some of its mathematical properties. The estimation of the model parameters is approached by the method of maximum likelihood and the Bayesian analysis. We consider the likelihood ratio (LR) statistic and other criteria to compare the KBGG model with its sub-models and other non-nested models. The potentiality of the KBGG distribution is illustrated in three applications to real data sets. The new model provides a rather flexible mechanism for fitting a wide spectrum of real world lifetime data in reliability, biology, environmental studies and other areas.

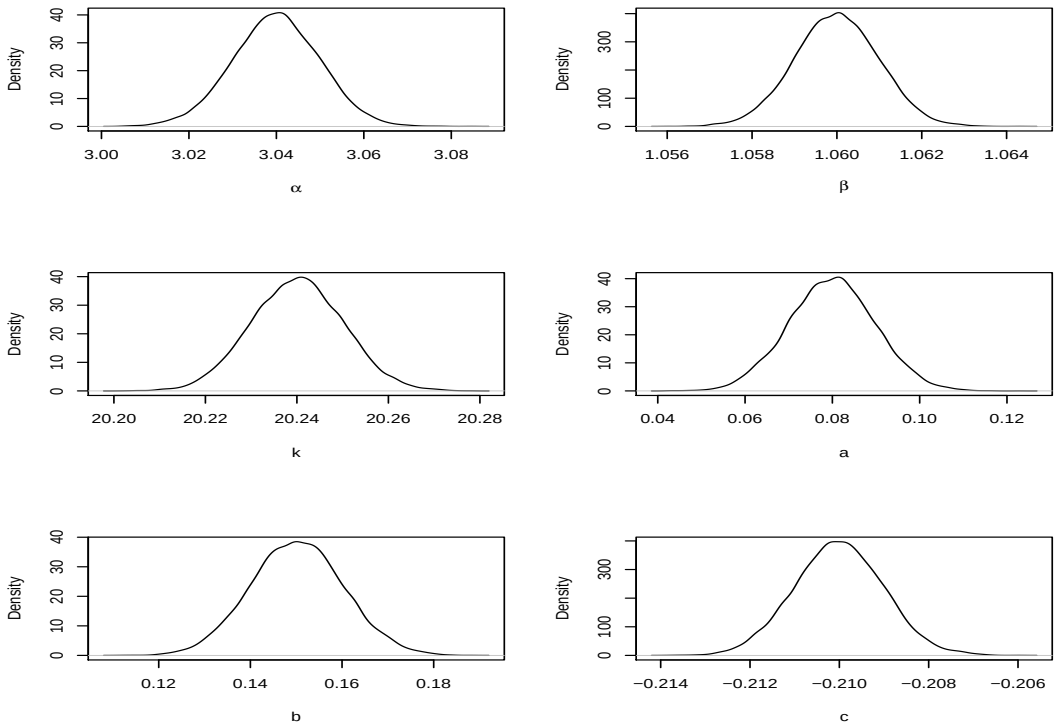


Figure 4.6 – Approximate posterior marginal densities for the parameters of the KBGG model for the second data set

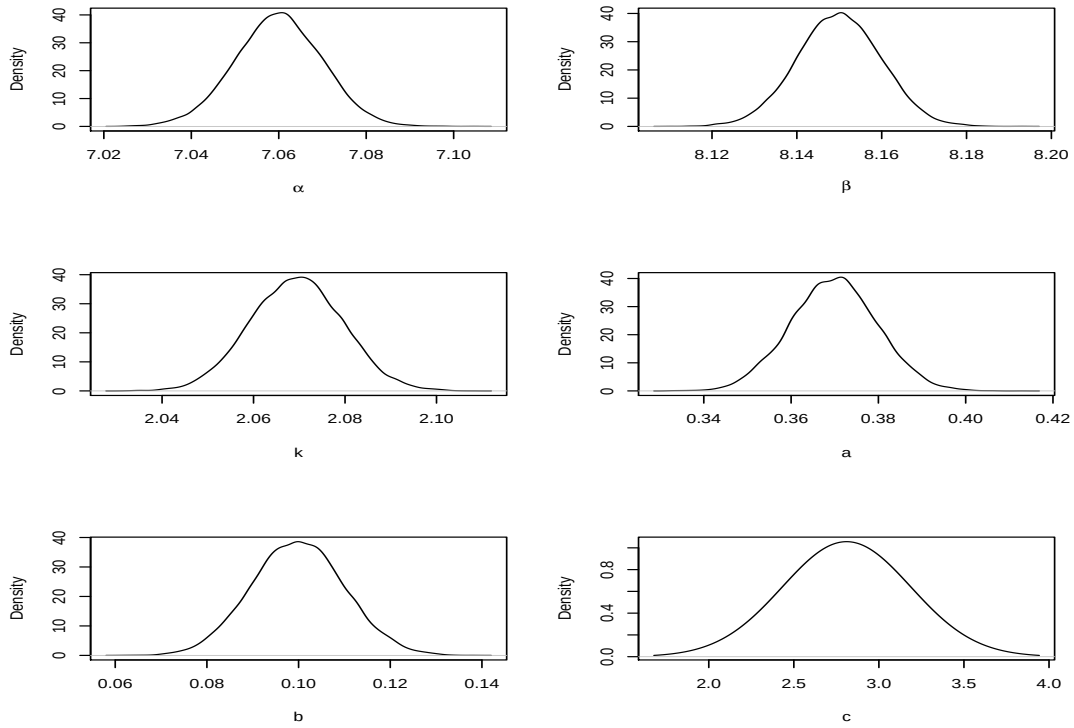


Figure 4.7 – Approximate posterior marginal densities for the parameters of the KBGG model for the third data set

References

ABRAMOWITZ, M.; STEGUN, I. A. **Handbook of Mathematical Functions**. New York: Dover Publications, 1968. 1046p.

BALAKRISHNAN, N.; LEIVA, V.; SANHUEZA, A.; CABRERA, E. Mixture inverse Gaussian distributions and its transformations, moments and applications. **Statistics**, Berlin, v.43, p.91-104, 2009.

BARKAUSKAS, D. A.; KRONEWITTER, S. R.; LEBRILLA, C. B.; ROCKE, D. M. Analysis of MALDI FT-ICR mass spectrometry data: a time series approach. **Analytica Chimica Acta**, Amsterdam, v.648, p.207-214, 2009.

BEKKER, A.; ROUX, J. J. Reliability characteristics of the Maxwell distribution: a Bayes estimation study. **Communications in Statistics - Theory and Methods**, London, v.34, p.2169-2178, 2005.

COORAY, K.; ANANDA, M. M. A. A generalization of the half-normal distribution with applications to lifetime data. **Communications in Statistics - Theory and Methods**, London, v.37, p.1323-1337, 2008.

CORDEIRO, G. M.; CASTELLARES, F.; MONTENEGRO, L. C.; de CASTRO, M. The beta generalized gamma distribution. **Statistics**, Berlin, v.47, p.888-900, 2013.

CORDEIRO, G. M.; CRISTINO, C. T.; HASHIMOTO, E. M.; ORTEGA, E. M. M. The beta generalized Rayleigh distribution with applications to lifetime data. **Statistical Papers**, London, v.54, p.133-161, 2011a.

CORDEIRO, G. M.; ORTEGA, E. M. M.; SILVA, G. O. The exponentiated generalized gamma distribution with application to lifetime data. **Journal of Statistical Computation and Simulation**, London, v.81, p.827-842, 2011b.

COX, C. The generalized F distribution: an umbrella for parametric survival analysis. **Statistics in Medicine**, London, v.27, p.4301-4312, 2008.

COX, C.; CHU, H.; SCHNEIDER, M. F.; MUÑOZ, A. Parametric survival analysis and taxonomy of hazard functions for the generalized gamma distribution. **Statistics in Medicine**, London, v.26, p.4352-4374, 2007.

EUGENE, N.; LEE, C.; FAMOYE, F. Beta-normal distribution and its applications. **Communications in Statistics - Theory and Methods**, London, v.31, p.497-512, 2002.

FAMOYE, F.; LEE, C.; OLUMOLADE, O. The beta-Weibull distribution. **Journal of Statistical Theory and Practice**, New York, v.4, p.121-136, 2005.

GRADSHTEYN, I. S.; RYZHIK, I. M. **Table of Integrals, Series, and Products**. San Diego: Academic Press, 2007. 315p.

GUPTA, R. D.; KUNDU, D. Exponentiated Exponential Family: An Alternative to Gamma and Weibull Distributions. **Biometrical Journal**, Berlin, v.43, p.117-130, 2001.

GUPTA, R. C.; GUPTA, P. L.; GUPTA, R. D. Modeling failure time data by Lehman alternatives. **Communications Statistics - Theory and Methods**, London, v.27, p.887-904, 1998.

KONG, L.; LEE, C.; SEPANSKI, J. H. On the properties of beta-gamma distribution. **Journal of Modern Applied Statistical Methods**, New York, v.6, p.187-211, 2007.

KUNDU, D.; KANNAN, N.; BALAKRISHNAN, N. On the function of Birnbaum-Saunders distribution and associated inference. **Computational Statistics and Data Analysis**, Amsterdam, v.52, p.2692-2702, 2008.

LAWLESS, J. F. **Statistical Models and Methods for Lifetime Data**. New York: John Wiley and Sons, 1982. 580p.

LEIVA, V.; BARROS, M.; PAULA, G.A. Generalized Birnbaum-Saunders Models Using R. In: ESCOLA DE MODELOS DE REGRESSÃO, 11., 2009, Recife. **Generalized Birnbaum-Saunders Models Using R**. Recife: Edufpe, 2009. 329p.

MALHOTRA, J.; SHARMA, A. K.; KALER, R. S. On the performance analysis of wireless receiver using generalized-gamma fading model. **Annals of Telecommunications**, Paris, v.64, p.147-153, 2009.

MUDHOLKAR, G. S.; SRIVASTAVA, D. K.; FRIEMER, M. The exponential Weibull family: A reanalysis of the bus-motor failure data. **Technometrics**, New York, v.37, p.436-445, 1995.

NADARAJAH, S.; GUPTA, A. K. A generalized gamma distribution with application to drought data. **Mathematics and Computer in Simulation**, Amsterdam, v.74, p.1-7, 2007.

NADARAJAH, S.; KOTZ, S. The beta exponential distribution. **Reliability Engineering and System Safety**, London, v.91, p.689-697, 2005.

NG, K. W.; KOTZ, S. Kummer-gamma and Kummer-beta univariate and multivariate distributions. **Research Report**, Department of Statistics, The University of Hong Kong, Hong Kong, v.84, 1995.

ORTEGA, E. M. M.; BOLFARINE, H.; PAULA, G. A. Influence diagnostics in generalized log-gamma regression models. **Computational Statistics and Data Analysis**, Amsterdam, v.42, p.165-186, 2003.

ORTEGA, E. M. M.; CANCHO, V. G.; PAULA, G. A. Generalized log-gamma regression models with cure fraction. **Lifetime Data Analysis**, Amsterdam, v.15, p.79-106, 2009.

PASCOA, M. A. R.; ORTEGA, E. M. M.; CORDEIRO, G. M. The Kumaraswamy generalized gamma distribution with application in survival analysis. **Statistical Methodology**, Amsterdam, v.8, p.411-433,2011.

PESCIM, R. R.; CORDEIRO, G .M.; DEMÉTRIO, C. G. B.; ORTEGA, E. M. M.; NADARAJAH, S. The new class of Kummer beta generalized distributions. **Statistics and Operations Research Transactions**, Barcelona, v.36, p.153-180, 2012.

PESCIM, R. R.; DEMÉTRIO, C. G. B.; CORDEIRO, G. M.; ORTEGA, E. M. M.; URBANO, M. R. The beta generalized half-normal distribution. **Computational Statistics and Data Analysis**, Amsterdam, v.54, p.945-957, 2010.

PRUDNIKOV, A. P.; BRYCHKOV, Y. A.; MARICHEV, O. I. **Integrals and Series**, Amsterdam: Science Publishers, 1986. 392p.

STACY, E. W. A generalization of the gamma distribution, **Annals of Mathematical Statistics**, New York, v.33, p.1187-1192, 1962.

XIE, X.; LIU, X. Analytical three-moment autoconversion parameterization based on generalized gamma distribution. **Journal of Geophysical Research**, New York, v.114, D17201. doi:10.1029/2008JD011633, 2009.

5 CONCLUSION

In this work, we proposed and studied a new family of distributions called the Kummer-beta generalized (KB-G) family, which includes as special cases two classical generators of distributions: the beta-generalized and exponentiated generators. For each parent G distribution, we defined the corresponding KB-G distribution with three additional parameters using simple formulas. Following this idea, we added three shape parameters to extend widely-known distributions such as normal, gamma, Weibull, Gumbel, logistic and Pareto distributions. In fact, for any baseline G distribution, we noted that the corresponding KB-G distribution provides more flexibility, giving it greater applicability. Some characteristics of the KB-G class of distributions, such as ordinary and incomplete moments, moment generating function, mean deviation and order statistics, have tractable mathematical properties. The main role of the generator parameters is related to the skewness and kurtosis of the new class. We adopted the maximum likelihood method and Bayesian approach to estimate the model parameters and determine the observed information matrix for the general family. Inference on the model parameters was conducted based on likelihood ratio statistics for testing nested models and the Akaike information criteria (AIC) and Bayesian information criteria (BIC) statistics for non-nested models. Two applications of the Kummer-beta gamma (KBGa) and Kummer-beta normal (KBN) distributions to real data sets demonstrated that these distributions provide a more appropriate fit than others models in the literature.

In the same way, we introduced and studied an important distribution based on the new family of Kummer-beta generalized distributions: The Kummer-beta Birnbaum-Saunders (KBBS) distribution which is widely applicable to represent failure times of fatiguing materials. The KBBS density function was expressed as a linear combination of Birnbaum-Saunders (BS) density functions which allowed us to derive some of its mathematical properties, such as ordinary and incomplete moments, moment generating function, mean deviations, entropy, reliability, order statistics and their moments. We investigated the maximum likelihood estimation of the model parameters. An application of the KBBS distribution to a real data set indicated that the new distribution provides consistently better fits than its sub-models and other lifetime models.

Further, we also proposed a new six-parameter model called the Kummer beta generalized gamma (KBGG) distribution which contains at least 32 special models such as the beta generalized gamma (BGG), beta Weibull (BW), beta exponential (BE), generalized gamma (GG), Weibull (W) and gamma (Ga) distributions and thus could be a high flexible model for analyzing positive skewed data. The KBGG density function was expressed as a linear combination of GG densities. The estimation of the model parameters was approached by the method of maximum likelihood and the Bayesian analysis. We considered the likelihood ratio (LR) statistic and other criteria to compare the KBGG model with its sub-models and other non-nested model. The potentiality of the KBGG distribution was illustrated in three applications to real

data sets. The new model provided a rather flexible mechanism for fitting a wide spectrum of real world lifetime data in reliability and biological sciences.

APPENDICES

Appendix A - Elements of the information matrix for any KB-G distribution

The elements of this matrix can be worked out as

$$\mathrm{E}\left(-\frac{\partial^2\ell(\theta)}{\partial a^2}\right) = -\frac{n}{K}\,\mathrm{E}\left[\frac{1}{K}\left(\frac{\partial K}{\partial a}\right) - \frac{\partial^2 K}{\partial a^2}\right],$$

$$\mathrm{E}\left(-\frac{\partial^2\ell(\theta)}{\partial b\partial c}\right) = -\frac{n}{K}\,\mathrm{E}\left[\frac{1}{K}\left(\frac{\partial K}{\partial b}\right)\left(\frac{\partial K}{\partial c}\right) - \frac{\partial^2 K}{\partial b\partial c}\right],$$

$$\mathrm{E}\left(-\frac{\partial^2\ell(\theta)}{\partial c^2}\right) = -\frac{n}{K}\,\mathrm{E}\left[\frac{1}{K}\left(\frac{\partial K}{\partial c}\right) - \frac{\partial^2 K}{\partial c^2}\right],$$

$$\mathrm{E}\left(-\frac{\partial^2\ell(\theta)}{\partial a\partial b}\right) = -\frac{n}{K}\,\mathrm{E}\left[\frac{1}{K}\left(\frac{\partial K}{\partial a}\right)\left(\frac{\partial K}{\partial b}\right) - \frac{\partial^2 K}{\partial a\partial b}\right],$$

$$\mathrm{E}\left(-\frac{\partial^2\ell(\theta)}{\partial a\partial c}\right) = -\frac{n}{K}\,\mathrm{E}\left[\frac{1}{K}\left(\frac{\partial K}{\partial a}\right)\left(\frac{\partial K}{\partial c}\right) - \frac{\partial^2 K}{\partial a\partial c}\right],$$

$$\mathrm{E}\left(-\frac{\partial^2\ell(\theta)}{\partial b^2}\right) = -\frac{n}{K}\,\mathrm{E}\left[\frac{1}{K}\left(\frac{\partial K}{\partial b}\right) - \frac{\partial^2 K}{\partial b^2}\right],$$

$$\mathrm{E}\left(-\frac{\partial^2\ell(\theta)}{\partial a\partial\gamma_j}\right) = -\sum_{i=1}^{n}\mathrm{E}\left[\frac{1}{G\left(x_i;\gamma\right)}\frac{\partial G\left(x_i;\gamma\right)}{\partial\gamma_j}\right],$$

$$\mathrm{E}\left(-\frac{\partial^2\ell(\theta)}{\partial b\partial\gamma_j}\right) = -\sum_{i=1}^{n}\mathrm{E}\left[\frac{1}{1-G\left(x_i;\gamma\right)}\frac{\partial G\left(x_i;\gamma\right)}{\partial\gamma_j}\right],$$

$$\mathrm{E}\left(-\frac{\partial^2\ell(\theta)}{\partial c\partial\gamma_j}\right) = \sum_{i=1}^{n}\mathrm{E}\left[\frac{\partial g\left(x_i;\gamma\right)}{\partial\gamma_j}\right]$$

and

$$\begin{aligned}\mathrm{E}\left(-\frac{\partial^2\ell(\theta)}{\partial\gamma_k\partial\gamma_j}\right) &= \sum_{i=1}^{n}\mathrm{E}\left[\frac{1}{g^2\left(x_i;\gamma\right)}\frac{\partial^2 g\left(x_i;\gamma\right)}{\partial\gamma_j\partial\gamma_k}\right] + c\sum_{i=1}^{n}\mathrm{E}\left[\frac{\partial^2 g\left(x_i;\gamma\right)}{\partial\gamma_j\partial\gamma_k}\right] + \\ &\quad \sum_{i=1}^{n}\mathrm{E}\left[\frac{(a-1)}{G^2\left(x_i;\gamma\right)}\frac{\partial^2 G\left(x_i;\gamma\right)}{\partial\gamma_j\partial\gamma_k}\right] + \sum_{i=1}^{n}\mathrm{E}\left[\frac{(1-b)}{\left\{1-G\left(x_i;\gamma\right)\right\}^2}\frac{\partial^2 G\left(x_i;\gamma\right)}{\partial\gamma_j\partial\gamma_k}\right]\end{aligned}$$

for $j = 1, \ldots, p$, where

$$\begin{aligned}
\frac{\partial^2 K}{\partial a^2} &= -\left\{ \frac{[\psi'(a) - \psi'(a+b)]}{{}_1F_1(a, a+b, -c)} + \frac{[\psi(a) - \psi(a+b)]^2}{[{}_1F_1(a, a+b, -c)]^2} \frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial a} \right. \\
&+ \frac{1}{[{}_1F_1(a, a+b, -c)]^2} \frac{\partial^2\, {}_1F_1(a, a+b, -c)}{\partial a^2} + \frac{[\psi(a) - \psi(a+b)]^2}{{}_1F_1(a, a+b, -c)} \\
&+ \frac{2}{[{}_1F_1(a, a+b, -c)]^2} \frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial a} \\
&\left. + \frac{1}{[{}_1F_1(a, a+b, -c)]^3} \left(\frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial a} \right)^2 \right\},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 K}{\partial b^2} &= -\left\{ \frac{[\psi'(b) - \psi'(a+b)]}{{}_1F_1(a, a+b, -c)} + \frac{[\psi(b) - \psi(a+b)]^2}{[{}_1F_1(a, a+b, -c)]^2} \frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial b} \right. \\
&+ \frac{1}{[{}_1F_1(a, a+b, -c)]^2} \frac{\partial^2\, {}_1F_1(a, a+b, -c)}{\partial b^2} + \frac{[\psi(b) - \psi(a+b)]^2}{{}_1F_1(a, a+b, -c)} \\
&+ \frac{2}{[{}_1F_1(a, a+b, -c)]^2} \frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial b} \\
&\left. + \frac{1}{[{}_1F_1(a, a+b, -c)]^3} \left(\frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial b} \right)^2 \right\},
\end{aligned}$$

$$\frac{\partial^2 K}{\partial c^2} = -\left\{ \frac{a(a+1){}_1F_1(a+2, a+b+2, -c)}{(a+b)B(a,b){}_1F_1(a, a+b, -c)} + \frac{a^2\, [{}_1F_1(a+1, a+b+1, -c)]^2}{(a+b)^2 B(a,b)\, [{}_1F_1(a, a+b, -c)]^2} \right\},$$

$$\begin{aligned}
\frac{\partial^2 K}{\partial a \partial b} &= -\left\{ \frac{[\psi'(a+b)]}{{}_1F_1(a, a+b, -c)} + \frac{[\psi(b) - \psi(a+b)]}{[{}_1F_1(a, a+b, -c)]^2} \frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial a} \right. \\
&+ \frac{1}{[{}_1F_1(a, a+b, -c)]^2} \frac{\partial^2\, {}_1F_1(a, a+b, -c)}{\partial a \partial b} + \left[\frac{\psi(b) - \psi(a+b)}{{}_1F_1(a, a+b, -c)} \right]^2 \\
&+ \frac{2\, [\psi(a) - \psi(a+b)]}{[{}_1F_1(a, a+b, -c)]^2} \frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial a} \\
&\left. + \frac{2}{[{}_1F_1(a, a+b, -c)]^3} \frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial b} \frac{\partial\, {}_1F_1(a, a+b, -c)}{\partial a} \right\},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 K}{\partial a \partial c} &= {}_1F_1(a+1, a+b+1, -c) + \frac{a}{(a+b)} + a\, [\psi'(a) - \psi'(a+b)] \\
&+ \frac{a}{{}_1F_1(a, a+b, -c)} \frac{\partial\, {}_1F_1(a+1, a+b+1, -c)}{\partial a} \\
&+ \frac{a}{{}_1F_1(a+1, a+b+1, -c)} \frac{\partial {}_1F_1(a+1, a+b+1, -c)}{\partial a},
\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 K}{\partial b \partial c} &= {}_1F_1(a+1,a+b+1,-c)+\frac{a}{(a+b)}+a\left[\psi'(b)-\psi'(a+b)\right]\\ &+\frac{a}{{}_1F_1(a,a+b,-c)}\frac{\partial\,{}_1F_1(a+1,a+b+1,-c)}{\partial b}\\ &+\frac{a}{{}_1F_1(a+1,a+b+1,-c)}\frac{\partial_1 F_1(a+1,a+b+1,-c)}{\partial b},\end{aligned}$$

$$\begin{aligned}\frac{\partial^2\,{}_1F_1(a,a+b,-c)}{\partial a^2} &= -\left[\psi'(a+b)-\psi'(a)+\{\psi(a)-\psi(a+b)\}^2\right]\,{}_1F_1(a,a+b,-c)\\ &-\sum_{k=0}^{\infty}\frac{(a)_k(-c)^k}{k!(a+b)_k}\,\left[-2\psi(a)\psi(a+k)+2\psi(a+b)\psi(a+k)\right.\\ &+2\,\psi(a)\psi(a+b+k)-2\psi(a+b)\psi(a+b+k)+\psi^2(a+k)\\ &-2\,\psi(a+k)\psi(a+b+k)+\psi^2(a+b+k)\\ &\left.+\psi'(a+k)-\psi'(a+b+k)\right],\end{aligned}$$

$$\begin{aligned}\frac{\partial^2\,{}_1F_1(a,a+b,-c)}{\partial b^2} &= -\left[\psi'(a+b)-\psi^2(a+b)\right]\,{}_1F_1(a,a+b,-c)\\ &-\sum_{k=0}^{\infty}\frac{(a)_k(-c)^k}{k!(a+b)_k}\,\left[-2\psi(a+b)\psi(a+b+k)\right.\\ &\left.-\psi'(a+b+k)+\psi^2(a+b+k)\right]\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^2\,{}_1F_1(a,a+b,-c)}{\partial a \partial b} &= \left[\psi'(a+b)-\psi^2(a+b)-\psi(a)\psi(a+b)\right]\,{}_1F_1(a,a+b,-c)\\ &-\sum_{k=0}^{\infty}\frac{(a)_k(-c)^k}{k!(a+b)_k}\,\left[2\psi(a+b)\psi(a+b+k)-\psi^2(a+b+k)\right.\\ &-\psi(a+k)\psi(a+b)+\psi(a+k)\psi(a+b+k)\\ &\left.-\psi(a)\psi(a+b+k)+\psi(a+b+k)\right].\end{aligned}$$

Appendix B - Elements of the observed information matrix for the KBBS distribution

The elements of the observed information matrix, $\mathbf{J}(\boldsymbol{\theta})$, for the parameters α, β, a, b and c are:

$$\begin{aligned}
U_{\alpha\alpha} &= \frac{n}{\alpha^2} + \frac{6n}{\alpha^4} - \frac{3}{\alpha^4}\sum_{i=1}^{n}\left(\frac{x_i}{\beta} + \frac{\beta}{x_i}\right) + \frac{2(a-1)}{\alpha^2}\sum_{i=1}^{n}\nu_i\,\frac{\phi(\nu_i)}{\Phi(\nu_i)} - \frac{2(b-1)}{\alpha^2}\sum_{i=1}^{n}\nu_i\,\frac{\phi(\nu_i)}{1-\Phi(\nu_i)} \\
&\quad + \frac{(a-1)}{\alpha^3}\sum_{i=1}^{n}\left\{\frac{\nu_i^4\,\phi(\nu_i)}{\Phi(\nu_i)} - \frac{\alpha\,\nu_i^2\,\phi(\nu_i)^2}{\Phi(\nu_i)}\right\} - \frac{(b-1)}{\alpha^3}\sum_{i=1}^{n}\left\{\frac{\nu_i^4\,\phi(\nu_i)}{1-\Phi(\nu_i)} - \frac{\alpha\,\nu_i^2\,\phi(\nu_i)^2}{1-\Phi(\nu_i)}\right\} \\
&\quad - \frac{4\,c}{\alpha}\sum_{i=1}^{n}\nu_i\,\phi(\nu_i)\,\left[1+\phi(\nu_i)\right],
\end{aligned}$$

$$\begin{aligned}
U_{\alpha\beta} &= -\frac{1}{\alpha^3\beta}\sum_{i=1}^{n}\left(\frac{x_i}{\beta} - \frac{\beta}{x_i}\right) + \frac{(a-1)}{2\alpha^2\beta}\sum_{i=1}^{n}\left\{\frac{\alpha\,\nu_i\,\phi(\nu_i)}{\Phi(\nu_i)} + \frac{\nu_i^4\,\phi(\nu_i)}{\Phi(\nu_i)} - \frac{\alpha\,\nu_i^2\,\phi(\nu_i)^2}{\Phi(\nu_i)^2}\right\} \\
&\quad - \frac{(b-1)}{2\alpha^2\beta}\sum_{i=1}^{n}\left\{\frac{\alpha\,\nu_i\,\phi(\nu_i)}{1-\Phi(\nu_i)} + \frac{\nu_i^4\,\phi(\nu_i)}{1-\Phi(\nu_i)} - \frac{\alpha\,\nu_i^2\,\phi(\nu_i)^2}{[1-\Phi(\nu_i)]^2}\right\} \\
&\quad - \frac{c}{\alpha\beta}\sum_{i=1}^{n}\nu_i\,\phi(\nu_i)\,\tau(\sqrt{x_i/\beta}),
\end{aligned}$$

$$\begin{aligned}
U_{\beta\beta} &= \frac{n}{2\beta^2} - \sum_{i=1}^{n}(x_i+\beta)^{-2} - \frac{1}{\alpha^2\beta^3}\sum_{i=1}^{n}x_i + \frac{(a-1)}{2\alpha\beta^2}\sum_{i=1}^{n}\frac{\phi(\nu_i)\,\tau(\sqrt{x_i/\beta})}{\Phi(\nu_i)} \\
&\quad - \frac{(b-1)}{2\alpha\beta^2}\sum_{i=1}^{n}\frac{\phi(\nu_i)\,\tau(\sqrt{x_i/\beta})}{1-\Phi(\nu_i)} - \frac{(a-1)}{4\alpha\beta^2}\sum_{i=1}^{n}\left\{-\frac{\alpha\,\nu_i\,\phi(\nu_i)}{\Phi(\nu_i)} + \frac{\nu_i\,\phi(\nu_i)\,\tau(\sqrt{x_i/\beta})^2}{\alpha\,\Phi(\nu_i)}\right. \\
&\quad \left. + \frac{\nu_i\,\phi(\nu_i)^2\,\tau(\sqrt{x_i/\beta})^2}{\alpha\,\Phi(\nu_i)^2}\right\} - \frac{4c}{4\alpha\beta^2}\sum_{i=1}^{n}\left\{\alpha\,\nu_i\,\phi(\nu_i) + \alpha^{-1}\tau(\sqrt{x_i/\beta})^2\,\phi(\nu_i)^2\right\},
\end{aligned}$$

$$U_{aa} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial a^2} - \frac{1}{K}\left[\frac{\partial K}{\partial a}\right]^2\right\},\ U_{bb} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial b^2} - \frac{1}{K}\left[\frac{\partial K}{\partial b}\right]^2\right\},$$

$$U_{cc} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial c^2} - \frac{1}{K}\left[\frac{\partial K}{\partial c}\right]^2\right\},\ U_{ab} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial a\partial b} - \frac{1}{K}\frac{\partial K}{\partial a}\frac{\partial K}{\partial b}\right\}$$

and

$$U_{ac} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial a\partial c} - \frac{1}{K}\frac{\partial K}{\partial a}\frac{\partial K}{\partial c}\right\},\ U_{bc} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial b\partial c} - \frac{1}{K}\frac{\partial K}{\partial b}\frac{\partial K}{\partial c}\right\},$$

where $\frac{\partial^2 K}{\partial a^2}$, $\frac{\partial^2 K}{\partial b^2}$, $\frac{\partial^2 K}{\partial c^2}$, $\frac{\partial^2 K}{\partial a\partial b}$, $\frac{\partial^2 K}{\partial a\partial c}$ and $\frac{\partial^2 K}{\partial b\partial c}$ are defined in Appendix A.

Appendix C - Elements of the observed information matrix for the KBGG distribution

The elements of the observed information matrix, $\mathbf{J}(\boldsymbol{\theta})$, for the parameters α, β, k, a, b and c are:

$$
\begin{aligned}
U_{\alpha\alpha} &= \frac{n}{\alpha^2} + \frac{n(\beta\,k-1)}{\alpha^2} - \frac{\beta(a-1)}{\alpha^2}\sum_{i=1}^{n}\frac{v_i}{\gamma(k,u_i)} \\
&+ \frac{\beta(a-1)}{\alpha^2}\left\{\frac{\beta\,k}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\gamma(k,u_i)} + \frac{\beta}{\alpha}\sum_{i=1}^{n}\frac{u_i\,v_i}{\gamma(k,u_i)} + \sum_{i=1}^{n}\left[\frac{v_i}{\gamma(k,u_i)}\right]^2\right\} \\
&- \frac{\beta(b-1)}{\alpha^2}\sum_{i=1}^{n}\frac{v_i}{\Gamma(k)-\gamma(k,u_i)} - \frac{\beta(b-1)}{\alpha^2}\left\{\frac{\beta\,k}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\Gamma(k)-\gamma(k,u_i)}\right. \\
&\left.+ \frac{\beta}{\alpha}\sum_{i=1}^{n}\frac{u_i\,v_i}{\Gamma(k)-\gamma(k,u_i)} - \sum_{i=1}^{n}\left[\frac{v_i}{\gamma(k,u_i)}\right]^2\right\} \\
&- \frac{\beta\,c}{\alpha^2\,\Gamma(k)}\sum_{i=1}^{n}v_i - \frac{\beta\,c}{\alpha}\left[\frac{\beta\,k}{\alpha}\sum_{i=1}^{n}v_i + \frac{\beta}{\alpha}\sum_{i=1}^{n}u_i\,v_i\right],
\end{aligned}
$$

$$
\begin{aligned}
U_{\alpha\beta} &= -\frac{nk}{\alpha} + \frac{1}{\alpha}\sum_{i=1}^{n}u_i + \frac{\beta}{\alpha}\sum_{i=1}^{n}u_i\,s_i - \frac{(a-1)}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\gamma(k,u_i)} \\
&- \frac{\beta(a-1)}{\alpha}\left\{k\sum_{i=1}^{n}\frac{v_i\,s_i}{\gamma(k,u_i)} + \sum_{i=1}^{n}\frac{u_i\,v_i\,s_i}{\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{v_i^2\,s_i}{[\gamma(k,u_i)]^2}\right\} \\
&+ \frac{\beta(b-1)}{\alpha}\left\{k\sum_{i=1}^{n}\frac{v_i\,s_i}{\Gamma(k)-\gamma(k,u_i)} + \sum_{i=1}^{n}\frac{u_i\,v_i\,s_i}{\Gamma(k)-\gamma(k,u_i)}\right. \\
&\left.+ \sum_{i=1}^{n}\frac{v_i^2\,s_i}{[\Gamma(k)-\gamma(k,u_i)]^2}\right\} + \frac{c}{\alpha\,\Gamma(k)}\sum_{i=1}^{n}v_i + \frac{\beta\,c\,k}{\alpha\,\Gamma(k)}\sum_{i=1}^{n}v_i\,s_i,
\end{aligned}
$$

$$
\begin{aligned}
U_{\alpha k} &= -\frac{n\beta}{\alpha} - \frac{\beta(a-1)}{\alpha}\left\{\beta\sum_{i=1}^{n}\frac{v_i\,s_i}{\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{v_i\,\gamma'(k,u_i)|_k}{[\gamma(k,u_i)]^2}\right\} \\
&+ \frac{\beta(b-1)}{\alpha}\left\{\beta\sum_{i=1}^{n}\frac{v_i\,s_i}{\Gamma(k)-\gamma(k,u_i)}\right. \\
&\left.- \sum_{i=1}^{n}\frac{\Gamma(k)\,\psi(k)\,v_i}{[\Gamma(k)-\gamma(k,u_i)]^2} + \sum_{i=1}^{n}\frac{\gamma'(k,u_i)|_k}{[\Gamma(k)-\gamma(k,u_i)]^2}\right\} \\
&- \frac{\beta\,c\psi(k)}{\alpha\Gamma(k)}\sum_{i=1}^{n}v_i + \frac{\beta^2 c}{\alpha\Gamma(k)}\sum_{i=1}^{n}v_i\,s_i,
\end{aligned}
$$

$$
\begin{aligned}
U_{\beta\beta} &= -\frac{n}{\beta^2} - \sum_{i=1}^{n} u_i\, s_i^2 + (1-a)\left\{k\sum_{i=1}^{n}\frac{v_i\, s_i^2}{\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{u_i\, v_i\, s_i^2}{\gamma(k,u_i)} - \sum_{i=1}^{n}\left[\frac{v_i\, s_i}{\gamma(k,u_i)}\right]^2\right\} \\
&\quad + (1-b)\left\{k\sum_{i=1}^{n}\frac{v_i\, s_i^2}{\Gamma(k)-\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{u_i\, v_i\, s_i^2}{\Gamma(k)-\gamma(k,u_i)} + \sum_{i=1}^{n}\left[\frac{v_i\, s_i}{\Gamma(k)-\gamma(k,u_i)}\right]^2\right\} \\
&\quad - \frac{c}{\Gamma(k)}\left\{k\sum_{i=1}^{n} v_i\, s_i - \sum_{i=1}^{n} u_i\, v_i\, s_i^2\right\},
\end{aligned}
$$

$$
U_{\beta k} = (a-1)\left\{\beta\sum_{i=1}^{n}\frac{v_i\, s_i^2}{\gamma(k,u_i)} - \sum_{i=1}^{n}\left[\frac{v_i\, s_i}{\gamma(k,u_i)}\right]^2\right\} + \frac{c\,\psi(k)}{\Gamma(k)}\sum_{i=1}^{n} v_i\, s_i - \frac{\beta\, c}{\Gamma(k)}\sum_{i=1}^{n} v_i\, s_i^2,
$$

$$
\begin{aligned}
U_{kk} &= -\psi'(k) - n(a-1)\psi'(k) + (a-1)\left\{\sum_{i=1}^{n}\frac{\gamma''(k,u_i)|_k}{\gamma(k,u_i)} - \sum_{i=1}^{n}\left[\frac{\gamma'(k,u_i)|_k}{\gamma(k,u_i)}\right]^2\right\} \\
&\quad + (1-b)\left\{\sum_{i=1}^{n}\frac{\gamma''(k,u_i)|_k}{\Gamma(k)-\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{\Gamma(k)\,\psi(k)\,\gamma'(k,u_i)|_k}{[\Gamma(k)-\gamma(k,u_i)]^2}\right. \\
&\quad \left. + \sum_{i=1}^{n}\left[\frac{\gamma'(k,u_i)|_k}{\Gamma(k)-\gamma(k,u_i)}\right]^2\right\} + (b-1)\sum_{i=1}^{n}\frac{\psi'(k)\,\gamma(k,u_i)}{\Gamma(k)-\gamma(k,u_i)} \\
&\quad + \psi(k)\,(1-b)\left\{\sum_{i=1}^{n}\frac{\gamma'(k,u_i)|_k}{\Gamma(k)-\gamma(k,u_i)} - \sum_{i=1}^{n}\frac{\Gamma(k)\,\psi(k)\,\gamma(k,u_i)}{[\Gamma(k)-\gamma(k,u_i)]^2}\right. \\
&\quad \left. + \sum_{i=1}^{n}\frac{\gamma(k,u_i)\,\gamma'(k,u_i)|_k}{[\Gamma(k)-\gamma(k,u_i)]^2}\right\} + 2\,c\,\psi(k)\sum_{i=1}^{n}\frac{\gamma'(k,u_i)|_k}{\Gamma(k)} - \sum_{i=1}^{n}\frac{c\,\gamma''(k,u_i)|_k}{\Gamma(k)} \\
&\quad + c\left[\frac{\psi'(k)+\psi^2(k)}{\Gamma(k)}\right]\sum_{i=1}^{n}\gamma(k,u_i),
\end{aligned}
$$

$$
U_{\alpha a} = -\frac{\beta}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\gamma(k,u_i)},\ U_{\alpha b} = \frac{\beta}{\alpha}\sum_{i=1}^{n}\frac{v_i}{\Gamma(k)-\gamma(k,u_i)},\ U_{\alpha c} = \frac{\beta}{\alpha\Gamma(k)}\sum_{i=1}^{n} v_i,\ U_{\beta a} = \sum_{i=1}^{n}\frac{v_i\, s_i}{\gamma(k,u_i)},
$$

$$
U_{\beta b} = -\sum_{i=1}^{n}\frac{v_i\, s_i}{\Gamma(k)-\gamma(k,u_i)},\ U_{\beta c} = \frac{1}{\Gamma(k)}\sum_{i=1}^{n} v_i\, s_i,\ U_{\alpha c} = \frac{\beta}{\alpha\Gamma(k)}\sum_{i=1}^{n} v_i,
$$

$$
U_{ka} = -n\psi(k) + \sum_{i=1}^{n}\frac{\gamma'(k,u_i)|_k}{\gamma(k,u_i)},\ U_{kc} = -\frac{1}{\Gamma(k)}\sum_{i=1}^{n}\gamma'(k,u_i)|_k + \frac{\psi(k)}{\Gamma(k)}\sum_{i=1}^{n}\gamma(k,u_i),
$$

$$
U_{kb} = -\sum_{i=1}^{n}\frac{\gamma'(k,u_i)|_k}{\Gamma(k)-\gamma(k,u_i)} + \psi(k)\sum_{i=1}^{n}\frac{\gamma(k,u_i)}{\Gamma(k)-\gamma(k,u_i)},
$$

$$
U_{aa} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial a^2} - \frac{1}{K}\left[\frac{\partial K}{\partial a}\right]^2\right\},\ U_{bb} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial b^2} - \frac{1}{K}\left[\frac{\partial K}{\partial b}\right]^2\right\},
$$

$$U_{cc} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial c^2} - \frac{1}{K}\left[\frac{\partial K}{\partial c}\right]^2\right\},\ U_{ab} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial a \partial b} - \frac{1}{K}\frac{\partial K}{\partial a}\frac{\partial K}{\partial b}\right\}$$

and

$$U_{ac} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial a \partial c} - \frac{1}{K}\frac{\partial K}{\partial a}\frac{\partial K}{\partial c}\right\},\ U_{bc} = \frac{n}{K}\left\{\frac{\partial^2 K}{\partial b \partial c} - \frac{1}{K}\frac{\partial K}{\partial b}\frac{\partial K}{\partial c}\right\},$$

where

$$\gamma'(k, u_i)|_k = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!}\, J(u_i, k+n-1, 1)$$

and

$$\gamma''(k, u_i)|_k = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!}\, J(u_i, k+n-1, 2).$$

The $J(.,.,.)$ function can be determined from the integral given by

$$J(a, p, 1) = \int_0^a x^p\, \log(x)\, dx = \frac{a^{p+1}}{(p+1)}\left[(p+1)\, \log(a) - 1\right]$$

and

$$J(a, p, 2) = \int_0^a x^p\, \log^2(x)\, dx = \frac{a^{p+1}}{(p+3)}\left\{2 - (p+1)\log(a)\left[2 - \log(a)\,(p+1)\right]\right\}.$$

Appendix D - Implemented functions used throughout the thesis in the statistical software package R.

Appendix D.1 - Plots of the Kummer beta density function

```
rm(list=ls(all=TRUE))
Plot 1
x < -seq(0, 1, 0.001)
a1<-2
b1<-3
c1=-20
f < -function(x)((exp(-(c1 * x))) * (x^(a1 - 1)) * ((1 - x)^(b1 - 1)))
i1 < -integrate(f, 0, 1)
fuc = (gamma(a1 + b1)/(gamma(a1)) * (gamma(b1))) * (i1)
p < -((exp(-(c1 * x))) * (x^(a1 - 1)) * ((1 - x)^(b1 - 1)))/((beta(a1, b1)) * (fuc))
a2<-1
b2<-3
c2=-10
f2 < -function(x)((exp(-(c2 * x))) * (x^(a2 - 1)) * ((1 - x)^(b2 - 1)))
i2 < -integrate(f2, 0, 1)
fuc2 = (gamma(a2 + b2)/(gamma(a2)) * (gamma(b2))) * (i2)
d < -((exp(-(c2 * x))) * (x^(a2 - 1)) * ((1 - x)^(b2 - 1)))/((beta(a2, b2)) * (fuc2))
a3<-4
b3<-3
c3=30
f3 < -function(x)((exp(-(c3 * x))) * (x^(a3 - 1)) * ((1 - x)^(b3 - 1)))
i3 < -integrate(f3, 0, 1)
fuc3 = (gamma(a3 + b3)/(gamma(a3)) * (gamma(b3)) * (i3))
z < -((exp(-(c3 * x))) * (x^(a3 - 1)) * ((1 - x)^(b3 - 1)))/((beta(a3, b3)) * (fuc3))
a4<-5
b4<-3.5
c4=15
f4 < -function(x)((exp(-(c4 * x))) * (x^(a4 - 1)) * ((1 - x)^(b4 - 1)))
i4 < -integrate(f4, 0, 1)
fuc4 = (gamma(a4 + b4)/(gamma(a4)) * (gamma(b4)) * (i4))
s < -((exp(-(c4 * x))) * (x^(a4 - 1)) * ((1 - x)^(b4 - 1)))/((beta(a4, b4)) * (fuc4))
plot(c(0, 1), c(0, 2), type = "n", xlab = "x", ylab = "f(x)", main = "")
lines(x,p,lty=1,lwd=2)
lines(x,d,lty=2,lwd=2)
lines(x,z,lty=3,lwd=2)
```

```
lines(x,s,lty=4,lwd=2)
legend(locator(1),c("a=2,b=3,c=-20","a=1,b=3,c=-
10","a=4,b=3,c=30","a=5,b=3.5,c=15"),lty=1:4,bty="n",cex=1.3)
****************************************
rm(list=ls(all=TRUE))
Plot 2
x <-seq(0,1,0.001)
a1<-4
b1<-1.5
c1=-10
f <-function(x)((exp(-(c1 * x))) * (x^(a1 - 1)) * ((1 - x)^(b1 - 1)))
i1 <-integrate(f,0,1)
fuc = (gamma(a1 + b1)/(gamma(a1)) * (gamma(b1))) * (i1)
p <-((exp(-(c1 * x))) * (x^(a1 - 1)) * ((1 - x)^(b1 - 1)))/((beta(a1,b1)) * (fuc))
a2<-4
b2<-1.5
c2=-5
f2 <-function(x)((exp(-(c2 * x))) * (x^(a2 - 1)) * ((1 - x)^(b2 - 1)))
i2 <-integrate(f2,0,1)
fuc2 = (gamma(a2 + b2)/(gamma(a2)) * (gamma(b2))) * (i2)
d <-((exp(-(c2 * x))) * (x^(a2 - 1)) * ((1 - x)^(b2 - 1)))/((beta(a2,b2)) * (fuc2))
a3<-4
b3<-1.5
c3=-15
f3 <-function(x)((exp(-(c3 * x))) * (x^(a3 - 1)) * ((1 - x)^(b3 - 1)))
i3 <-integrate(f3,0,1)
fuc3 = (gamma(a3 + b3)/(gamma(a3)) * (gamma(b3)) * (i3))
z <-((exp(-(c3 * x))) * (x^(a3 - 1)) * ((1 - x)^(b3 - 1)))/((beta(a3,b3)) * (fuc3))
a4<-4
b4<-1.5
c4=0
f4 <-function(x)((exp(-(c4 * x))) * (x^(a4 - 1)) * ((1 - x)^(b4 - 1)))
i4 <-integrate(f4,0,1)
fuc4 = (gamma(a4 + b4)/(gamma(a4)) * (gamma(b4)) * (i4))
s <-((exp(-(c4 * x))) * (x^(a4 - 1)) * ((1 - x)^(b4 - 1)))/((beta(a4,b4)) * (fuc4))
plot(c(0,1),c(0,12),type = "n",xlab = "x",ylab = "f(x)",main = "")
lines(x,p,lty=1,lwd=2)
lines(x,d,lty=2,lwd=2)
lines(x,z,lty=3,lwd=2)
```

```
lines(x,s,lty=4,lwd=2)
legend(locator(1),c("a=4,b=1.5,c=-10","a=4,b=1.5,c=-5","a=4,b=1.5,c=-
15","a=4,b=1.5,c=0"),lty=1:4,bty="n",cex=1.3)
**************************************
rm(list=ls(all=TRUE))
Plot 3
x < -seq(0,1,0.001)
a1<-4
b1<-1.5
c1=5
f < -function(x)((exp(-(c1 * x))) * (x^(a1 - 1)) * ((1 - x)^(b1 - 1)))
i1 < -integrate(f,0,1)
fuc = (gamma(a1 + b1)/(gamma(a1)) * (gamma(b1))) * (i1)
p < -((exp(-(c1 * x))) * (x^(a1 - 1)) * ((1 - x)^(b1 - 1)))/((beta(a1,b1)) * (fuc))
a2<-4
b2<-1.5
c2=10
f2 < -function(x)((exp(-(c2 * x))) * (x^(a2 - 1)) * ((1 - x)^(b2 - 1)))
i2 < -integrate(f2,0,1)
fuc2=(gamma(a2+b2)/(gamma(a2))*(gamma(b2)))*(i2)
d < -((exp(-(c2 * x))) * (x^(a2 - 1)) * ((1 - x)^(b2 - 1)))/((beta(a2,b2)) * (fuc2))
a3<-4
b3<-1.5
c3=15
f3 < -function(x)((exp(-(c3 * x))) * (x^(a3 - 1)) * ((1 - x)^(b3 - 1)))
i3 < -integrate(f3,0,1)
fuc3 = (gamma(a3 + b3)/(gamma(a3)) * (gamma(b3)) * (i3))
z < -((exp(-(c3 * x))) * (x^(a3 - 1)) * ((1 - x)^(b3 - 1)))/((beta(a3,b3)) * (fuc3))
a4<-4
b4<-1.5
c4=20
f4 < -function(x)((exp(-(c4 * x))) * (x^(a4 - 1)) * ((1 - x)^(b4 - 1)))
i4 < -integrate(f4,0,1)
fuc4 = (gamma(a4 + b4)/(gamma(a4)) * (gamma(b4)) * (i4))
s < -((exp(-(c4 * x))) * (x^(a4 - 1)) * ((1 - x)^(b4 - 1)))/((beta(a4,b4)) * (fuc4))
plot(c(0,1),c(0,6.1),type = "n",xlab = "x",ylab = "f(x)",main = "")
lines(x,p,lty=1,lwd=2)
lines(x,d,lty=2,lwd=2)
lines(x,z,lty=3,lwd=2)
```

```
lines(x,s,lty=4,lwd=2)
legend(locator(1),c("a=4,b=1.5,c=10","a=4,b=1.5,c=5",
"a=4,b=1.5,c=15","a=4,b=1.5,c=0"),lty=1:4,bty="n",cex=1.3)
```

Appendix D.2 - Plots of the KBBS density function

```
rm(list=ls(all=TRUE))
x < -seq(0,100,0.001)
a1<-1
b1<-1
c1=0
alpha1=1
betha1=1
v1 < -((alpha1^( - 1)) * ((((x/betha1)^(1/2)) - ((x/betha1)^( - 1/2)))))
integrand1 < -function(t)(t^(a1 - 1)) * ((1 - t)^(b1 - 1)) * exp(-c1 * t)
km1 < -integrate(integrand1, lower = 0, upper = 1)
p < -((exp(alpha1^( - 2)))/(2 * alpha1 * (sqrt(2 * pi * betha1)) * (km1))) * (x^( - 3/2)) *
(x + betha1) * (exp(-(((x/betha1) + (x/betha1)^( - 1))/(2 * alpha1^(2))))) *
((pnorm(v1))^(a1 - 1)) * ((1 - pnorm(v1))^(b1 - 1)) * (exp(-c1 * ((pnorm(v1)))))
a2<-1
b2<-2.5
c2=-10
alpha2=1
betha2=1
v2 < -((alpha2^( - 1)) * ((((x/betha2)^(1/2)) - ((x/betha2)^( - 1/2)))))
integrand2 < -function(t)(t^(a2 - 1)) * ((1 - t)^(b2 - 1)) * exp(-c2 * t)
km2 < -integrate(integrand2, lower = 0, upper = 1)
g < -((exp(alpha2^( - 2)))/(2 * alpha2 * (sqrt(2 * pi * betha2)) * (km2))) * (x^( - 3/2)) *
(x + betha2) * (exp(-(((x/betha2) + (x/betha2)^( - 1))/(2 * alpha2^(2))))) *
((pnorm(v2))^(a2 - 1)) * ((1 - pnorm(v2))^(b2 - 1)) * (exp(-c2 * ((pnorm(v2)))))
a3<-1
b3<-1
c3=-20
alpha3=1
betha3=1
v3 < -((alpha3^( - 1)) * ((((x/betha3)^(1/2)) - ((x/betha3)^( - 1/2)))))
integrand3 < -function(t)(t^(a3 - 1)) * ((1 - t)^(b3 - 1)) * exp(-c3 * t)
km3 < -integrate(integrand3, lower = 0, upper = 1)
```

$h < -((exp(alpha3^{(} - 2)))/(2 * alpha3 * (sqrt(2 * pi * betha3)) * (km3))) * (x^{(} - 3/2)) *$
$(x + betha3) * (exp(-(((x/betha3) + (x/betha3)^{(} - 1))/(2 * alpha3^{(}2))))) *$
$((pnorm(v3))^{(}a3 - 1)) * ((1 - pnorm(v3))^{(}b3 - 1)) * (exp(-c3 * ((pnorm(v3)))))$
a4<-0.5
b4<-0.5
c4=-5
alpha4=1
betha4=1
$v4 < -((alpha4^{(} - 1)) * ((((x/betha4)^{(}1/2)) - ((x/betha4)^{(} - 1/2)))))$
$integrand4 < -function(t)(t^{(}a4 - 1)) * ((1 - t)^{(}b4 - 1)) * exp(-c4 * t)$
$km4 < -integrate(integrand4, lower = 0, upper = 1)$
$j < -((exp(alpha4^{(} - 2)))/(2 * alpha4 * (sqrt(2 * pi * betha4)) * (km4))) * (x^{(} - 3/2)) *$
$(x + betha4) * (exp(-(((x/betha4) + (x/betha4)^{(} - 1))/(2 * alpha4^{(}2))))) *$
$((pnorm(v4))^{(}a4 - 1)) * ((1 - pnorm(v4))^{(}b4 - 1)) * (exp(-c4 * ((pnorm(v4)))))$
$plot(c(0, 11.2), c(0, 0.72), type = "n", xlab = "x", ylab = "f(x)", main =$
$"KBBS(a, b, c, 1, 1)", cex.lab = 1.3)$
lines(x,p,col='darkgreen',lty=1,lwd=2)
lines(x,g,col='green',lty=1,lwd=2)
lines(x,h,col='red',lty=1,lwd=2)
lines(x,j,col='blue',lty=1,lwd=2)
legend(locator(1),c("BS","a = 1, b = 2.5, c = -10","a = 1, b = 1, c = -20","a = 0.5, b = 0.5, c = -5"), lty=1,bty="n",col=c('darkgreen','green','red','blue'),cex=1.3)

Appendix D.3 - Plots of the KBBS hazard rate function

rm(list=ls(all=TRUE))
$x < -seq(0, 100, 0.1)$
a1<-1.8
b1<-1.5
c1=16
alpha1=2.2
betha1=0.2
$z < -pnorm((alpha1^{(} - 1)) * (((x/betha1)^{(}1/2) - (x/betha1)^{(} - 1/2)))) \; Z < -c()$
$for(iin1 : length(x)) \; a1 < -1.8 \; b1 < -1.5 \; c1 = 164 \; alpha1 = 2.2 \; betha1 = 0.2$
$integrand < -function(t)(t^{(}a1 - 1)) * ((1 - t)^{(}b1 - 1)) * exp(-c1 * t)$
$km < -integrate(integrand, lower = 0, upper = z[i]) \; km1 < -km \; Z < -cbind(Z, km1)$
$Z1 < -c(Z)$
$km2 < -integrate(integrand, lower = 0, upper = 1)$
$F < -Z1/km2$
$v1 < -((alpha1^{(} - 1)) * ((((x/betha1)^{(}1/2)) - ((x/betha1)^{(} - 1/2)))))$

$integrand1 < -function(t)(t^(a1-1)) * ((1-t)^(b1-1)) * exp(-c1 * t)$
$k1 < -integrate(integrand1, lower = 0, upper = 1)$
$p < -((exp(alpha1^(-2)))/(2 * alpha1 * (sqrt(2 * pi * betha1)) * (k1))) * (x^(-3/2)) *$
$(x + betha1) * (exp(-(((x/betha1) + (x/betha1)^(-1))/(2 * alpha1^(2))))) *$
$((pnorm(v1))^(a1-1)) * ((1 - pnorm(v1))^(b1-1)) * (exp(-c1 * ((pnorm(v1)))))$
$p1 = p/(1-F)$
a2<-1.6
b2<-1.5
c2=15
alpha2=2.5
betha2=0.19
$z2 = pnorm((alpha2^(-1)) * (((x/betha2)^(1/2) - (x/betha2)^(-1/2))))$
$Z2 < -c()\ for(iin1 : length(x))\ a2 < -1.6\ b2 < -1.5\ c2 = 15\ alpha2 = 2.5$
$betha2 = 0.19\ integrand22 < -function(t)(t^(a2-1)) * ((1-t)^(b2-1)) * exp(-c2 * t)$
$km3 < -integrate(integrand22, lower = 0, upper = z2[i])\ km4 < -km3$
$Z2 < -cbind(Z2, km4)\ Z3 < -c(Z2)$
$km5 < -integrate(integrand22, lower = 0, upper = 1)$
$F2 < -Z3/km5$
$v2 < -((alpha2^(-1)) * ((((x/betha2)^(1/2)) - ((x/betha2)^(-1/2)))))$
$integrand2 < -function(t)(t^(a2-1)) * ((1-t)^(b2-1)) * exp(-c2 * t)$
$k2 < -integrate(integrand2, lower = 0, upper = 1)$
$g < -((exp(alpha2^(-2)))/(2 * alpha2 * (sqrt(2 * pi * betha2)) * (k2))) * (x^(-3/2)) *$
$(x + betha2) * (exp(-(((x/betha2) + (x/betha2)^(-1))/(2 * alpha2^(2))))) *$
$((pnorm(v2))^(a2-1)) * ((1 - pnorm(v2))^(b2-1)) * (exp(-c2 * ((pnorm(v2)))))$
$g1 = g/(1-F2)$
a3<-1.1
b3<-1
c3=-2
alpha3=0.35
betha3=15
$z5 = pnorm((alpha3^(-1)) * (((x/betha3)^(1/2) - (x/betha3)^(-1/2))))$
$Z5 < -c()\ for(iin1 : length(x))\ a3 < -1.1\ b3 < -1\ c3 = -2\ alpha3 = 0.35\ betha3 = 15$
$integrand33 < -function(t)(t^(a3-1)) * ((1-t)^(b3-1)) * exp(-c3 * t)$
$km6 < -integrate(integrand33, lower = 0, upper = z5[i])\ km7 < -km6$
$Z5 < -cbind(Z5, km7)\ Z6 < -c(Z5)$
$km8 < -integrate(integrand33, lower = 0, upper = 1)$
$F3 < -Z6/km8$
$v3 < -((alpha3^(-1)) * ((((x/betha3)^(1/2)) - ((x/betha3)^(-1/2)))))$
$integrand3 < -function(t)(t^(a3-1)) * ((1-t)^(b3-1)) * exp(-c3 * t)$

$k3 < -integrate(integrand3, lower = 0, upper = 1)$
$h < -((exp(alpha3^{(} - 2)))/(2 * alpha3 * (sqrt(2 * pi * betha3)) * (k3))) * (x^{(} - 3/2)) *$
$(x + betha3) * (exp(-(((x/betha3) + (x/betha3)^{(} - 1))/(2 * alpha3^{(}2))))) *$
$((pnorm(v3))^{(}a3 - 1)) * ((1 - pnorm(v3))^{(}b3 - 1)) * (exp(-c3 * ((pnorm(v3)))))$
$h1 = h/(1 - F3)$
a4<-1.2
b4<-1
c4=-2
alpha4=0.2
betha4=20
$z7 = pnorm((alpha4^{(} - 1)) * (((x/betha4)^{(}1/2) - (x/betha4)^{(} - 1/2))))$
$Z7 < -c()\ for(iin1 : length(x))\ a4 < -1.2\ b4 < -1\ c4 = -2\ alpha4 = 0.2\ betha4 = 20$
$integrand44 < -function(t)(t^{(}a4 - 1)) * ((1 - t)^{(}b4 - 1)) * exp(-c4 * t)$
$km9 < -integrate(integrand44, lower = 0, upper = z7[i])\ km10 < -km9$
$Z7 < -cbind(Z7, km10)\ Z8 < -c(Z7)$
$km11 < -integrate(integrand44, lower = 0, upper = 1)$
$F4 < -Z8/km11$
$v4 < -((alpha4^{(} - 1)) * ((((x/betha4)^{(}1/2)) - ((x/betha4)^{(} - 1/2)))))$
$integrand4 < -function(t)(t^{(}a4 - 1)) * ((1 - t)^{(}b4 - 1)) * exp(-c4 * t)$
$km4 < -integrate(integrand4, lower = 0, upper = 1)$
$j < -((exp(alpha4^{(} - 2)))/(2 * alpha4 * (sqrt(2 * pi * betha4)) * (km4))) * (x^{(} - 3/2)) *$
$(x + betha4) * (exp(-(((x/betha4) + (x/betha4)^{(} - 1))/(2 * alpha4^{(}2))))) *$
$((pnorm(v4))^{(}a4 - 1)) * ((1 - pnorm(v4))^{(}b4 - 1)) * (exp(-c4 * ((pnorm(v4)))))$
$j1 = j/(1 - F4)$
$plot(c(0, 52.0), c(0, 0.09), type = "n", xlab = "x", ylab = "h(x)", main = "")$
lines(x,p1,col='darkgreen',lty=1,lwd=2)
lines(x,g1,col='green',lty=1,lwd=2)
lines(x,h1,col='red',lty=1,lwd=2)
lines(x,j1,col='blue',lty=1,lwd=2)
legend(22,0.09, expression(paste(a,"=1.8; ",b,"=1.5; ",c,"=16; ",alpha,"=2.2;",beta,"=0.2"),
paste(a,"=1.6; ",b,"=1.5; ",c,"=15; ",alpha,"=2.5;",beta,"=0.19"), paste(a,"=1.1; ",b,"=1;
",c,"=-2; ",alpha,"=0.35;",beta,"=15"), paste(a,"=1.2; ",b,"=1; ",c,"=-2;
",alpha,"=0.2;",beta,"=20")), lty=c(1,1,1,1),
lwd=c(2,2,2,2),col=c('darkgreen','green','red','blue'), bty="o", cex=1)
****
rm(list=ls(all=TRUE))
$x < -seq(0, 100, 0.1)$
a1<-2
b1<-1.68

c1=30

alpha1=1.5

betha1=1.4

$z < -pnorm((alpha1^(- 1)) * (((x/betha1)^(1/2) - (x/betha1)^(- 1/2))))$

$Z < -c()\ for(iin1 : length(x))\ a1 < -2\ b1 < -1.68\ c1 = 30\ alpha1 = 1.5\ betha1 = 1.4$

$integrand < -function(t)(t^(a1 - 1)) * ((1 - t)^(b1 - 1)) * exp(-c1 * t)$

$km < -integrate(integrand, lower = 0, upper = z[i])\ km1 < -km\ Z < -cbind(Z, km1)$

$Z1 < -c(Z)$

$km2 < -integrate(integrand, lower = 0, upper = 1)$

$F < -Z1/km2$

$v1 < -((alpha1^(- 1)) * ((((x/betha1)^(1/2)) - ((x/betha1)^(- 1/2)))))$

$integrand1 < -function(t)(t^(a1 - 1)) * ((1 - t)^(b1 - 1)) * exp(-c1 * t)$

$k1 < -integrate(integrand1, lower = 0, upper = 1)$

$p < -((exp(alpha1^(- 2)))/(2 * alpha1 * (sqrt(2 * pi * betha1)) * (k1))) * (x^(- 3/2)) *$
$(x + betha1) * (exp(-(((x/betha1) + (x/betha1)^(- 1))/(2 * alpha1^(2))))) *$
$((pnorm(v1))^(a1 - 1)) * ((1 - pnorm(v1))^(b1 - 1)) * (exp(-c1 * ((pnorm(v1)))))$

$p1 = p/(1 - F)$

a2<-2

b2<-1.55

c2=30

alpha2=1.5

betha2=1.4

$z2 = pnorm((alpha2^(- 1)) * (((x/betha2)^(1/2) - (x/betha2)^(- 1/2))))$

$Z2 < -c()\ for(iin1 : length(x))\ a2 < -2\ b2 < -1.55\ c2 = 30\ alpha2 = 1.5\ betha2 = 1.4$

$integrand22 < -function(t)(t^(a2 - 1)) * ((1 - t)^(b2 - 1)) * exp(-c2 * t)$

$km3 < -integrate(integrand22, lower = 0, upper = z2[i])\ km4 < -km3$

$Z2 < -cbind(Z2, km4)\ Z3 < -c(Z2)$

$km5 < -integrate(integrand22, lower = 0, upper = 1)$

$F2 < -Z3/km5$

$v2 < -((alpha2^(- 1)) * ((((x/betha2)^(1/2)) - ((x/betha2)^(- 1/2)))))$

$integrand2 < -function(t)(t^(a2 - 1)) * ((1 - t)^(b2 - 1)) * exp(-c2 * t)$

$k2 < -integrate(integrand2, lower = 0, upper = 1)$

$g < -((exp(alpha2^(- 2)))/(2 * alpha2 * (sqrt(2 * pi * betha2)) * (k2))) * (x^(- 3/2)) *$
$(x + betha2) * (exp(-(((x/betha2) + (x/betha2)^(- 1))/(2 * alpha2^(2))))) *$
$((pnorm(v2))^(a2 - 1)) * ((1 - pnorm(v2))^(b2 - 1)) * (exp(-c2 * ((pnorm(v2)))))$

$g1 = g/(1 - F2)$

a3<-2

b3<-1.6

c3=30

alpha3=1.5

betha3=1.4

$z5 = pnorm((alpha3^{(} - 1)) * (((x/betha3)^{(}1/2) - (x/betha3)^{(} - 1/2))))$

$Z5 < -c()\ for(iin1 : length(x))\ a3 < -2\ b3 < -1.6\ c3 = 30\ alpha3 = 1.5\ betha3 = 1.4$

$integrand33 < -function(t)(t^{(}a3 - 1)) * ((1 - t)^{(}b3 - 1)) * exp(-c3 * t)$

$km6 < -integrate(integrand33, lower = 0, upper = z5[i])\ km7 < -km6$

$Z5 < -cbind(Z5, km7)\ Z6 < -c(Z5)$

$km8 < -integrate(integrand33, lower = 0, upper = 1)$

$F3 < -Z6/km8$

$v3 < -((alpha3^{(} - 1)) * ((((x/betha3)^{(}1/2)) - ((x/betha3)^{(} - 1/2)))))$

$integrand3 < -function(t)(t^{(}a3 - 1)) * ((1 - t)^{(}b3 - 1)) * exp(-c3 * t)$

$k3 < -integrate(integrand3, lower = 0, upper = 1)$

$h < -((exp(alpha3^{(} - 2)))/(2 * alpha3 * (sqrt(2 * pi * betha3)) * (k3))) * (x^{(} - 3/2)) *$
$(x + betha3) * (exp(-(((x/betha3) + (x/betha3)^{(} - 1))/(2 * alpha3^{(}2)))))) *$
$((pnorm(v3))^{(}a3 - 1)) * ((1 - pnorm(v3))^{(}b3 - 1)) * (exp(-c3 * ((pnorm(v3)))))$

$h1 = h/(1 - F3)$

a4<-2

b4<-1.65

c4=30

alpha4=1.5

betha4=1.4

$z7 = pnorm((alpha4^{(} - 1)) * (((x/betha4)^{(}1/2) - (x/betha4)^{(} - 1/2))))$

$Z7 < -c()\ for(iin1 : length(x))\ a4 < -2\ b4 < -1.65\ c4 = 30\ alpha4 = 1.5\ betha4 = 1.4$

$integrand44 < -function(t)(t^{(}a4 - 1)) * ((1 - t)^{(}b4 - 1)) * exp(-c4 * t)$

$km9 < -integrate(integrand44, lower = 0, upper = z7[i])\ km10 < -km9$

$Z7 < -cbind(Z7, km10)\ Z8 < -c(Z7)$

$km11 < -integrate(integrand44, lower = 0, upper = 1)$

$F4 < -Z8/km11$

$v4 < -((alpha4^{(} - 1)) * ((((x/betha4)^{(}1/2)) - ((x/betha4)^{(} - 1/2)))))$

$integrand4 < -function(t)(t^{(}a4 - 1)) * ((1 - t)^{(}b4 - 1)) * exp(-c4 * t)$

$km4 < -integrate(integrand4, lower = 0, upper = 1)$

$j < -((exp(alpha4^{(} - 2)))/(2 * alpha4 * (sqrt(2 * pi * betha4)) * (km4))) * (x^{(} - 3/2)) *$
$(x + betha4) * (exp(-(((x/betha4) + (x/betha4)^{(} - 1))/(2 * alpha4^{(}2)))))) *$
$((pnorm(v4))^{(}a4 - 1)) * ((1 - pnorm(v4))^{(}b4 - 1)) * (exp(-c4 * ((pnorm(v4)))))$

$j1 = j/(1 - F4)$

$plot(c(6, 13.0), c(0.65, 0.85), type = "n", xlab = "x", ylab = "h(x)", main = "")$

lines(x,p1,col='darkgreen',lty=1,lwd=2)

lines(x,g1,col='green',lty=1,lwd=2)

lines(x,h1,col='red',lty=1,lwd=2)

```
lines(x,j1,col='blue',lty=1,lwd=2)
legend(5.6,0.685, expression(paste(a,"=2; ",b,"=1.68; ",c,"=30; ",alpha,"=1.5;",beta,"=1.4"),
paste(a,"=2; ",b,"=1.55; ",c,"=30; ",alpha,"=1.5;",beta,"=1.4"), paste(a,"=2; ",b,"=1.6;
",c,"=30; ",alpha,"=1.5;",beta,"=1.4"), paste(a,"=2; ",b,"=1.65; ",c,"=30;
",alpha,"=1.5;",beta,"=1.4")), lty=c(1,1,1,1),
lwd=c(2,2,2,2),col=c('darkgreen','green','red','blue'), bty="o", cex=1)
```

Appendix D.4 - Plots of the KBGG density function

```
rm(list=ls(all=TRUE))
x < -seq(0,100,0.001)
a1<-1.5
b1<-2
c1=2
alpha1=1
betha1=1
k1=2
integrand1 < -function(t)(t^(a1-1)) * ((1-t)^(b1-1)) * exp(-c1 * t)
km1 < -integrate(integrand1, lower = 0, upper = 1)
p < -((betha1)/(alpha1 * gamma(k1) * km1)) * ((x/alpha1)^(betha1 * k1) - 1)) *
(exp(-(x/alpha1)^betha1))) * ((pgamma((x/alpha1)^betha1), k1))^(a1-1)) * ((1 -
pgamma((x/alpha1)^betha1), k1))^(b1-1))*exp(-c1*(pgamma((x/alpha1)^betha1), k1)))
a2<-1
b2<-2.5
c2=-5
alpha2=1
betha2=1
k2=2
integrand2 < -function(t)(t^(a2-1)) * ((1-t)^(b2-1)) * exp(-c2 * t)
km2 < -integrate(integrand2, lower = 0, upper = 1)
g < -((betha2)/(alpha2 * gamma(k2) * km2)) * ((x/alpha2)^(betha2 * k2) - 1)) *
(exp(-(x/alpha2)^betha2))) * ((pgamma((x/alpha2)^betha2), k2))^(a2-1)) * ((1 -
pgamma((x/alpha2)^betha2), k2))^(b2-1))*exp(-c2*(pgamma((x/alpha2)^betha2), k2)))
a3<-1
b3<-1
c3=-15
alpha3=1
betha3=1
k3=2
integrand3 < -function(t)(t^(a3-1)) * ((1-t)^(b3-1)) * exp(-c3 * t)
```

```
km3 < -integrate(integrand3, lower = 0, upper = 1)
h < -((betha3)/(alpha3 * gamma(k3) * km3)) * ((x/alpha3)^(betha3 * k3) - 1)) *
(exp(-(x/alpha3)^(betha3))) * ((pgamma((x/alpha3)^(betha3), k3))^(a3 - 1)) * ((1 -
pgamma((x/alpha3)^(betha3), k3))^(b3-1))*exp(-c3*(pgamma((x/alpha3)^(betha2), k3)))
a4<-2.5
b4<-0.5
c4=2.5
alpha4=1
betha4=1
k4=2
integrand4 < -function(t)(t^(a4 - 1)) * ((1 - t)^(b4 - 1)) * exp(-c4 * t)
km4 < -integrate(integrand4, lower = 0, upper = 1)
j < -((betha4)/(alpha4 * gamma(k4) * km4)) * ((x/alpha4)^(betha4 * k4) - 1)) *
(exp(-(x/alpha4)^(betha4))) * ((pgamma((x/alpha4)^(betha4), k4))^(a4 - 1)) * ((1 -
pgamma((x/alpha4)^(betha4), k4))^(b4-1))*exp(-c4*(pgamma((x/alpha4)^(betha4), k4)))
plot(c(0, 11.2), c(0, 0.7), type = "n", xlab = "x", ylab = "f(x)", main =
"KBGG(a, b, c, 1, 1, 2)", cex.lab = 1.3)
lines(x,p,col='red',lty=1,lwd=2)
lines(x,g,col='blue',lty=1,lwd=2)
lines(x,h,col='green',lty=1,lwd=2)
lines(x,j,col='darkgreen',lty=1,lwd=2)
legend(locator(1),c("a = 1.5, b = 2, c = 2","a = 1, b = 2.5, c = -5","a = 1, b = 1, c = -15","a =
2.5, b = 0.5, c = 2.5"), lty=1,bty="n",col=c('red','blue','green','darkgreen'),cex=1.3)
```