UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA
UFSCar-USP

PEDRO LUIZ RAMOS

BAYESIAN AND CLASSICAL INFERENCE FOR THE GENERALIZED GAMMA DISTRIBUTION AND RELATED MODELS

Doctoral dissertation submitted to the Des/UFSCar and to the Institute of Mathematics and Computer Sciences – ICMC-USP, in partial fulfillment of the requirements for the PhD degree Statistics – Interagency Program Graduate in Statistics UFSCar-USP.

Advisor: Prof. Dr. Francisco Louzada

São Carlos

February 2018

UNIVERSIDADE FEDERAL DE SÃO CARLOS
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA
PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA
UFSCar-USP

PEDRO LUIZ RAMOS

ANÁLISE CLÁSSICA E BAYESIANA PARA A DISTRIBUIÇÃO GAMA GENERALIZADA E MODELOS RELACIONADOS

Tese apresentada ao Departamento de Estatística – Des/UFSCar e ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Doutor em Estatística - Programa Interinstitucional de Pós-Graduação em Estatística UFSCar-USP.

Orientador: Prof. Dr. Francisco Louzada

São Carlos

Fevereiro de 2018

UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa Interinstitucional de Pós-Graduação em Estatística

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Tese de Doutorado do candidato Pedro Luiz Ramos, realizada em 22/02/2018:

Prof. Dr. Francisco Louzada Neto
USP

Profa. Dra. Vera Lucia Damasceno Tomazella
UFSCar

Prof. Dr. Francisco Aparecido Rodrigues
USP

Profa. Dra. Teresa Cristina Martins Dias
UFSCar

Prof. Dr. Helio dos Santos Migon
UFRJ

Certifico que a defesa realizou-se com a participação à distância do(s) membro(s) Francisco Aparecido Rodrigues e, depois das arguições e deliberações realizadas, o(s) participante(s) à distância está(ao) de acordo com o conteúdo do parecer da banca examinadora redigido neste relatório de defesa.

Prof. Dr. Francisco Louzada Neto

To my beloved son, Francisco, and my wife and best friend, Maysa,
without whom this thesis would have been
completed one year earlier.

ACKNOWLEDGEMENTS

There are a number of people without whom this thesis might not have been written, and to whom I am greatly indebted.

To my mother, Rita, who has been a source of encouragement and inspiration to me throughout my life. To my meemaw, Carmen, who always made me feel that she had been waiting to see me all day, and when she sees me, her day is completed.

To my beloved and supportive wife, Maysa, who is always by my side making my life worth living. To my little boy, Francisco, the most beautiful chapter of my life.

I would like to thank my brother, Eduardo, an excellent mathematician who supported me and helped in many topics of this thesis. To my father, Luiz, who always gave me the best things in life: his time, his care and his love. To my sisters Lais and Ana Luiza, more precious than gold.

I owe my deepest gratitude to my supervisor Dr. Francisco Louzada. Without his continuous optimism concerning this work, enthusiasm, encouragement and support this study would hardly have been completed.

I express my warmest gratitude to Wally and Eddie for opening their hearts and their house during my sandwich doctorate period. No matter where this life takes us, I hope that I will always hold a special place in all of their hearts because they certainly will in mine.

I would like to express my gratitude for the valuable suggestions from my committee members: Dr. Vera L. D. Tomazella, Dr. Helio S. Migon, Dr. Francisco A. Rodrigues and Dr. Teresa C. M. Dias.

I express my warmest gratitude to Dr. Dipak K. Dey who received me at UCONN and contributed to many discussions that helped to shape this thesis. I also would like to express my gratitude to Professor Dr. Victor H. Lachos for all that he has done. I consider him my friend.

Finally, my thanks go to CAPES and the taxpayers, whose financial support provided me the sandwich doctorate period at the University of Connecticut, USA.

Curiosity. Not good for cats,
great for scientists.
Dr. Larry Fleinhardt, Numb3rs

ABSTRACT

RAMOS, P. L. **Bayesian and classical inference for the generalized gamma distribution and related models**. 2018. 141 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2018.

The generalized gamma (GG) distribution is an important model that has proven to be very flexible in practice for modeling data from several areas. This model has important sub-models, such as the Weibull, gamma, lognormal, Nakagami-m distributions, among others. In this work, our main objective is to develop different estimation procedures for the unknown parameters of the generalized gamma distribution and related models (Nakagami-m and gamma), considering both classical and Bayesian approaches. Under the Bayesian approach, we provide in a simple way necessary and sufficient conditions to check whether or not objective priors lead proper posterior distributions for the Nakagami, gamma, and GG distributions. As a result, one can easily check if the obtained posterior is proper or improper directly looking at the behavior of the improper prior. These theorems are applied to different objective priors such as Jeffreys's rule, Jeffreys prior, maximal data information prior and reference priors. Simulation studies were conducted to investigate the performance of the Bayes estimators. Moreover, maximum a posteriori (MAP) estimators for the Nakagami and gamma distribution that have simple closed-form expressions are proposed Numerical results demonstrate that the MAP estimators outperform the existing estimation procedures and produce almost unbiased estimates for the fading parameter even for a small sample size. Finally, a new lifetime distribution that is expressed as a two-component mixture of the GG distribution is presented.

Keywords: Generalized gamma distribution, Nakagami-m distribution, gamma distribution, Bayesian methods.

RESUMO

RAMOS, P. L. **Análise clássica e Bayesiana para a distribuição gama generalizada e modelos relacionados**. 2018. 141 p. Tese (Doutorado em Estatística – Programa Interinstitucional de Pós-Graduação em Estatística) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2018.

A distribuição gama Generalizada (GG) possui um papel fundamental para modelar dados em diversas áreas. Tal distribuição possui como casos particulares importantes distribuições, tais como, Weibull, Gama, lognormal, Nakagami-m, dentre outras. Nesta tese, tem-se como objetivo principal, considerando as abordagens clássica e Bayesiana, desenvolver diferentes procedimentos de estimação para os parâmetros da distribuição gama generalizada e de alguns dos seus casos particulares dentre eles as distribuições Nakagami-m e Gama. Do ponto de vista Bayesiano, iremos propor de forma simples, condições suficientes e necessárias para verificar se diferentes distribuições a priori não-informativas impróprias conduzem a distribuições posteriori próprias. Tais resultados são apresentados para as distribuições Nakagami-m, gama e gama generalizada. Assim, com a criação de novas prioris não-informativas, para tais modelos, futuros pesquisadores poderão utilizar nossos resultados para verificar se as distribuições a posteriori obtidas são impróprias ou não. Aplicações dos teoremas propostos são apresentados em diferentes prioris objetivas, tais como, a regra de Jeffreys, priori Jeffreys, priori maximal data information e prioris de referência. Iremos também realizar estudos de simulação para investigar a influência destas prioris nas estimativas a posteriori. Além disso, são propostos estimadores de máxima a posteriori em forma fechada para as distribuições Nakagami-m e Gama. Por meio de estudos de simulação verificamos que tais estimadores superam os procedimentos de estimação existentes e produzem estimativas quase não-viciadas para os parâmetros de interesse. Por fim, apresentamos uma nova distribuição obtida considerando um modelo de mistura de distribuições gama generalizada.

Palavras-chave: Distribuição Gama Generalizada, Distribuição Nakagami-m, Distribuição Gama, Métodos Bayesianos.

LIST OF FIGURES

Figure 1 – Different shapes for the hazard function. 28
Figure 2 – TTT-plot shapes, retrieved from Ramos, Moala and Achcar (2014). 29
Figure 3 – Hazard function shapes for NK distribution considering different values of μ and Ω. 40
Figure 4 – Mean residual life function shapes for NK distribution considering different values of μ and Ω. 41
Figure 5 – TTT-plot, empirical reliability function and reliability function of different fitted probability distributions considering the data set related to the time of failure of 18 electronic devices and the cumulative hazard function. 54
Figure 6 – TTT-plot, empirical reliability function and reliability function of different fitted probability distributions for the data set related to the number of 1000s cycles for failure for 60 electrical appliances and the cumulative hazard function. 56
Figure 7 – Survival function fitted by the empirical and by different probability distributions considering the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901-1998 and the hazard function fitted by a GG distribution. 96
Figure 8 – MREs for μ considering $c_3 = (2, 2.1, \cdots, 4)$, $\mu = 6$, $\Omega = 10$, $n = 120$ (left panel) $\mu = 20$, $\Omega = 2$, $n = 20$ (right panel) and for $N = 1,000,000$ simulated samples and $n = 50$. 105
Figure 9 – MREs, RMSEs for μ and Ω considering $\mu = 4, \Omega = 2$ for $N = 1,000,000$ simulated samples and $n = (10, 15, \ldots, 140)$. 105
Figure 10 – MREs, RMSEs for μ considering $\mu = 4$, $\Omega = 80$ for $N = 1,000,000$ simulated samples and $n = (10, 15, \ldots, 140)$. 106
Figure 11 – MREs, RMSEs for different values of $\mu = (0.5, 1.0, \ldots, 20)$ considering $\Omega = 2$ for $N = 1,000,000$ simulated samples and $n = 50$. 106
Figure 12 – MREs, RMSEs for different values of $\mu = (0.5, 1.0, 1.5, \ldots, 20)$ considering $\Omega = 80$ for $N = 1,000,000$ simulated samples and $n = 50$. 106
Figure 13 – MREs for ϕ considering $c_3 = (2, 2.1, \cdots, 4)$, $\phi = 3$, $\lambda = 4$ and $n = 20$ (left panel) $\phi = 10$, $\lambda = 4$ and $n = 20$ (right panel) for $N = 500,000$ simulated samples. 109
Figure 14 – MREs and RMSEs for different values of $\phi = (0.5, 1.0, \ldots, 20)$ for sample size of 8 elements. 109

Figure 15 – MREs and RMSEs for ϕ for samples sizes of $6,7,8,\ldots,20$ elements. Upper panels: considering $\phi=4$, lower panels: considering $\phi=10$. 110
Figure 16 – Density function shapes for GWL distribution considering different values of ϕ,λ and α. 113
Figure 17 – Hazard function shapes for GWL distribution and considering different values of ϕ,λ and α 115
Figure 18 – MRE's, MSE's related from the estimates of $\phi=0.5,\lambda=0.7$ and $\alpha=1.5$ for N simulated samples, considering different values of n obtained using the following estimation method 1-MLE, 2-MPS, 3-ADE, 4-RTADE. 121
Figure 19 – (left panel) the TTT-plot, (middle panel) the fitted survival superimposed to the empirical survival function and (right panels) the hazard function adjusted by GWL distribution 122
Figure 20 – (left panel) the TTT-plot, (middle panel) the fitted survival superimposed to the empirical survival function and (right panels) the hazard function adjusted by GWL distribution 124

LIST OF TABLES

Table 1 – The MRE(MSE) from the estimates of μ considering different values of n with $N = 10,000,000$ simulated samples using the estimation methods: 1 - MM, 2 - MLE, 3- Jeffreys's rule, 4 - Jeffreys prior, 5 - Overall reference prior. 51

Table 2 – The $CP_{95\%}$ from the estimates of μ and Ω considering different values of n with $N = 10,000,000$ simulated samples using the estimation methods: 1 - MM, 2 - MLE, 3- Jeffreys's rule, 4 - Jeffreys prior, 5 - Overall reference prior. 52

Table 3 – Data set related to the lifetime of one hundred observations on breaking stress of carbon fibers 53

Table 4 – The MLEs (Standard Errors) obtained for different distributions considering the lifetime of one hundred observations on breaking stress of carbon fibers. . 53

Table 5 – Posterior mode and 95% credibility intervals for μ and Ω obtained from the data set related to the lifetime of one hundred observations on breaking stress of carbon fibers. 53

Table 6 – Results of AIC, AICc and BIC criteria for different probability distributions considering the data set related to the life time of one hundred observations on breaking stress of carbon fibers. 53

Table 7 – Data set related to the lifetime of 60 (in cycles) electrical devices. 54

Table 8 – The MLEs obtained for different distributions considering the lifetime data set of 60 (in cycles) electrical devices. 55

Table 9 – Posterior mode and 95% CIs for μ and Ω obtained from the data set related to 1000s cycles of failure for 60 electrical appliances. 55

Table 10 – Results of AIC, AICc and BIC criteria for different probability distributions considering the data set related to 1000 cycles of failure for 60 electrical appliances. 55

Table 11 – The MRE(MSE) for for the estimates of ϕ and μ considering different sample sizes. 75

Table 12 – The MRE(MSE) for for the estimates of ϕ and μ considering different sample sizes. 76

Table 13 – The $CP_{95\%}$ from the estimates of μ and Ω considering different values of n with $N = 10,000$ simulated samples. 77

Table 14 – The $CP_{95\%}$ from the estimates of μ and Ω considering different values of n with $N = 10,000$ simulated samples. 77

Table 15 – Bias (RMSE) of the ML estimates and the Bayes estimators (posterior mode) for $10,000$ samples of sizes $n = (50, 100, 200)$ and different values of θ. 94
Table 16 – Coverage probability with a 95% confidence level equals the ML estimates and the Bayes estimators (posterior mode) considering $10,000$ samples of sizes $n = (50, 100, 200)$ and different values of θ. 94
Table 17 – Posterior mean and 95% credibility intervals for ϕ and α from the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901-1998. 95
Table 18 – Posterior mode, standard deviations and 95% credible intervals for ϕ, μ and α from the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901-1998. 95
Table 19 – Results of AIC, AICc and BIC criteria for different probability distributions considering the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901-1998. 95
Table 20 – Posterior mode, MLE, Posterior mean and 90% credibility intervals for ϕ and α from the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901-1998. 96
Table 21 – Lifetimes data (in hours) related to a device put on test. 122
Table 22 – Results of AIC and AICc criteria and the p-value from the KS test for all fitted distributions considering the Aarset dataset. 122
Table 23 – MLE estimates, Standard-error and 95% confidence intervals (CI) for ϕ, λ and α 123
Table 24 – January average flows (m^3/s) of the Cantareira system. 123
Table 25 – Results of AIC and AICc criteria and the p-value from the KS test for all fitted distributions considering the data set related to the january average flows (m^3/s) of the Cantareira system. 123
Table 26 – ML estimates, Standard-error and 95% confidence intervals (CI) for ϕ, λ and α 124

CONTENTS

1 INTRODUCTION . . . 23
1.1 Objectives and Overview . . . 25

2 PRELIMINARIES . . . 27
2.1 Survival Analysis . . . 27
2.1.1 TTT-plot . . . 28
2.2 Frequentist inference . . . 29
2.2.1 Maximum Likelihood Estimation . . . 29
2.3 Bayesian Inference . . . 30
2.3.1 Non-informative priors . . . 30
2.4 Discrimination criterion methods . . . 33
2.5 Some Useful Mathematical Results . . . 33

3 NAKAGAMI-M DISTRIBUTION . . . 37
3.1 Introduction . . . 37
3.2 Nakagami-m distribution . . . 38
3.3 Classical Inference . . . 41
3.3.1 Moment Estimators . . . 41
3.3.2 Maximum Likelihood Estimation . . . 42
3.4 Bayesian Analysis . . . 43
3.4.1 Proper Posterior . . . 43
3.4.2 Jeffreys's rule . . . 45
3.4.3 Jeffreys prior . . . 46
3.4.4 Reference prior . . . 47
3.4.5 Maximal Data Information prior . . . 48
3.4.6 Matching priors . . . 48
3.4.7 Numerical integration . . . 49
3.5 Simulation Study . . . 50
3.6 Applications in Reliability . . . 51
3.6.1 Breaking stress of carbon fibers . . . 52
3.6.2 Cycles up to the failure for electrical appliances . . . 54
3.7 Discussion . . . 56

4 GAMMA DISTRIBUTION . . . 59

4.1 Introduction . . . 59
4.2 Classical Inference . . . 60
4.3 Proper Posterior . . . 61
4.4 Objective priors . . . 65
4.4.1 Uniform Prior . . . 65
4.4.2 Jeffreys Rule . . . 66
4.4.3 Jeffreys prior . . . 66
4.4.4 Miller prior . . . 67
4.4.5 Reference prior . . . 68
4.4.5.1 Reference prior when ϕ is the parameter of interest . . . 68
4.4.5.2 Reference prior when μ is the parameter of interest . . . 69
4.4.5.3 Overall Reference prior . . . 69
4.4.6 Maximal Data Information prior . . . 70
4.4.6.1 Modified MDI prior . . . 70
4.4.7 Matching priors . . . 71
4.4.8 Consensus Prior . . . 72
4.4.8.1 Geometric mean . . . 72
4.4.8.2 Arithmetic mean . . . 73
4.5 Numerical evaluation . . . 74
4.6 Discussion . . . 78

5 GENERALIZED GAMMA DISTRIBUTION . . . 79
5.1 Introduction . . . 79
5.2 Maximum likelihood estimators . . . 80
5.3 Bayesian Analysis . . . 81
5.4 Objective Priors . . . 86
5.4.1 Some common objective priors . . . 86
5.4.2 Priors based on the Fisher information matrix . . . 87
5.4.3 Reference priors . . . 89
5.5 Simulation Analysis . . . 92
5.6 Real data application . . . 94
5.7 Discussion . . . 97

6 CLOSED-FORM ESTIMATORS FOR NAKAGAMI-M AND GAMMA DISTRIBUTIONS . . . 99
6.1 Introduction . . . 99
6.2 Closed-form estimators based on the generalized gamma distribution 100
6.3 Nakagami-m . . . 101
6.3.1 Generalized moment estimators . . . 101
6.3.2 Modified maximum likelihood estimators . . . 101

6.3.3 Maximum a Posteriori estimator 102
6.3.4 Results 104
6.4 Gamma distribution 107
6.4.1 Bias corrected estimators 107
6.4.2 Bias Expression 108
6.4.3 Maximum a Posteriori estimator 108
6.4.4 Simulation and Discussion 108
6.5 Discussion 110

7 MIXTURE OF GENERALIZED GAMMA DISTRIBUTION 111
7.1 Introduction 111
7.2 Generalized Weighted Lindley distribution 112
7.2.1 Moments 112
7.2.2 Survival Properties 114
7.2.3 Entropy 116
7.2.4 Lorenz curves 118
7.3 Maximum likelihood estimators 118
7.4 Simulation Study 120
7.5 Application 120
7.5.1 Lifetimes data 121
7.5.2 Average flows data 123
7.6 Concluding remarks 124

8 COMMENTS AND FURTHER DEVELOPMENT 125
8.1 Comments 125
8.2 Further development 126

BIBLIOGRAPHY 129

A USEFUL PROPORTIONALITIES 137

CHAPTER

1

INTRODUCTION

In recent years, several new extensions of the exponential distribution have been proposed in the literature for describing real problems. Introduced by Stacy (1962), the generalized gamma (GG) distribution is an important distribution that has proven to be very flexible in practice for modeling data from several areas, such as climatology, meteorology, medicine, reliability and image processing data, among others. The GG distribution is a distribution which has several particular cases, such as the exponential, Weibull, gamma, Log-normal, Nakagami-m, Half-normal, Rayleigh, Maxwell-Boltzmann and chi distributions.

Marani, Lavagnini and Buttazzoni (1986) used this distribution to analyze data relating to air quality in Venice, Italy. Tahai and Meyer (1999) proposed new methods for analyzing citations in recent publications to find journals with greater influence using the GG distribution. Aalo, Piboongungon and Iskander (2005) used this distribution to analyze the performance degradation of wireless communication systems. Li *et al.* (2011) used the GG distribution to obtain different techniques for processing SAR (Synthetic aperture radar) images. Other applications of the GG distribution can be seen in Noortwijk (2001), Dadpay, Soofi and Soyer (2007), Balakrishnan and Peng (2006), Raju and Srinivasan (2002) and Ahsanullah, Maswadah and Ali (2013).

The sub-models related to the GG distribution has been widely used in the literature. For instance, considering the Google scholarship, a search with the words *Weibull distribution*, *gamma distribution*, *Log-normal distribution* and *Nakagami-m distribution*, on March 2016, found respectively 217.000, 2.920.000, 535.000 and 28.600 research papers.

Different estimation procedures have been discussed in the literature considering both classical and Bayesian analysis. However, much work still has to be done, in this thesis under the Bayesian approach, we considered different objective priors for the Nakagami-m, gamma and generalized gamma models, such as the Jeffreys Rule (BOX; TIAO, 1973), Jeffreys prior (JEFFREYS, 1946), maximal data information (MDI) prior (ZELLNER, 1977; ZELLNER, 1984) and reference priors (BERNARDO, 1979; BERGER; BERNARDO, 1989; BERGER;

BERNARDO, 1992a; BERGER; BERNARDO, 1992b; BERGER; BERNARDO *et al.*, 1992; BERGER *et al.*, 2015).

These objective priors are usually improper and could lead to improper posteriors. For instance, Noortwijk (2001) considered the non-informative Jeffreys prior to estimating the quartiles of the flood of a given river using the GG distribution. Such prior has the important one-to-one invariant property. However, we proved that such prior leads to an improper posterior and should not be used. Providing a proof that a posterior distribution is proper or improper is not an easy task. Northrop and Attalides (2016) argued that "... there is no general theory providing simple conditions under which improper priors yields proper posteriors for a particular model, so this must be investigated case-by-case". In this study, we overcome this problem by providing in a simple way necessary and sufficient conditions to check whether or not these objective priors lead to proper posteriors distributions for the chosen models. In this way, one can easily check if the obtained posterior is proper or improper considering directly the behavior of the improper prior.

For the Nakagami distribution the main theorem is applied in different objective priors such as Jeffreys's rule, Jeffreys prior, the MDI prior and reference priors. The Jeffreys-rule prior and Jeffreys prior gave proper posterior distribution respectively for $n \geq 1$ and $n \geq 0$, whereas they are matching priors only for one of the parameters. The MDI prior provided improper posterior for any sample sizes and should not be used in Bayesian analysis. The overall reference prior yielded a proper posterior distribution if and only if $n \geq 1$. This prior is the one-at-a-time reference prior for any chosen parameter of interest and any ordering of the nuisance parameters. It is also the only prior that is a matching prior for both parameters. An extensive simulation study showed that the proposed overall reference posterior distribution returns more accurate results, as well as better theoretical properties such as the invariance property under one-to-one transformations of the parameters, consistency under marginalization and consistent sampling properties. The proposed methodology is fully illustrated using two real lifetime data sets, demonstrating that the NK distribution can be used to describe lifetime data.

For the Gamma distribution we investigated the same problem related to the posterior distribution. We proved that among the priors considered in this study the MDI prior was the only that yield an improper posterior for any sample sizes. An extensive simulation study showed that the posterior distribution obtained under Tibshirani prior provided more accurate results in terms of mean relative errors (MREs), mean square errors (MSEs) and coverage probabilities and should be used to obtain inference for this distribution.

Considering the GG distribution, we proved that the uniform prior, the prior obtained from Jeffreys' first rule and the MDI prior lead to improper posteriors. Further, the impropriety of the posterior using the Jeffreys' priors (NOORTWIJK, 2001) led us to consider the scenario where the Jeffreys prior has an independent structure (FONSECA; FERREIRA; MIGON, 2008). However, the four possible objective priors also returned improper posteriors. An alternative was to consider reference priors. Since these priors are sensitive to the ordering of the unknown parameters, from Proposition 2.3.1 we obtained six reference priors, two of them were similar to other reference priors. Among the four distinct reference priors, we proved that only one returned a proper posterior distribution. The obtained posterior has excellent theoretical properties such as invariance property under one-to-one transformations of the parameters, consistency under marginalization and consistent sampling properties and should be used to make inference in the parameters of the GG distribution.

However, under the proposed approaches discussed so far, numerical integration must be used to obtain the posterior estimates and to perform the classical inference. Despite the enormous evolution of computational methods during the last decades, these methods still carry the disadvantage of high computational cost in many applications. Particularly in the case where the parameter estimators need to be obtained in real time, often within devices with embedded technology Song (2008). To overcome this problem we propose a class of maximum a posteriori (MAP) estimators for the parameters of the Nakagami and gamma distributions. They have simple closed-form expressions and can be rewritten as a bias-corrected maximum likelihood estimators (MLEs). Numerical results have shown that the MAP estimation scheme outperforms the existing estimation procedures and produces almost unbiased estimates for the parameters even for small sample size.

Finally, a new lifetime distribution that is expressed as a two-component mixture of the generalized gamma distribution is proposed. This generalization accommodates increasing, decreasing, decreasing-increasing-decreasing, bathtub, or unimodal hazard shapes, making such distribution a flexible model for reliability data. A significant account of mathematical properties of the new distribution is presented as well as two data sets are analyzed for illustrative purposes, proving that the mixture model outperforms several usual three parameters lifetime distributions.

1.1 Objectives and Overview

The main objective of this thesis is to improve the estimation procedures for the GG distribution and some of its related models. In order to achieved that we will:

1. Provide sufficient conditions to check whether or not objective priors lead to proper posteriors distributions for the Nakagami, gamma and GG distributions.

2. Derive different objective priors for the distributions cited above and apply the theorems to check if the obtained priors lead to proper posteriors.

3. Select among the priors that lead to proper posteriors the best ones that return better estimates in terms of MREs, MSEs and coverage probabilities.

4. Propose for the Nakagami and gamma distributions MAP estimators that have simple closed-form expressions for the parameters and can be rewritten as bias-corrected maximum likelihood estimators.

5. Introduce and discuss the properties of a new lifetime distribution that is expressed as a two-component mixture of the generalized gamma distribution.

The remainder of this work is organized as follows. In Chapter 2, we present a literature review of some important topics that will be covered in this work. In Chapter 3, we present a Bayesian inference for the unknown parameters of the Nakagami-m distribution. In Chapter 4, we consider the same approach for the parameters of the gamma distribution. In Chapter 5, we extended the results to the generalized gamma distribution. In Chapter 5, MAP estimators that have simple closed-form expressions for the Nakagami-m and gamma distribution are proposed. In Chapter 6, we proposed a new lifetime distribution expressed as a two-component mixture of the generalized gamma distribution. Finally, in Chapter 7 we present some general comments and possible extensions of this current work.

CHAPTER

2

PRELIMINARIES

In this section, we present a literature review of some important topics that are covered throughout this thesis.

2.1 Survival Analysis

In survival analysis, the responses are usually characterized by the failure times and the occurrence of censorship. These responses are usually measured over time until the occurrence of the event of interest. Therefore, a random variable (RV) T will have only non-negative values and can be expressed by different mathematical functions, such as the probability density function (PDF) $f(t)$, the cumulative distribution function (CDF) $F(t)$, the survival function $S(t)$, the hazard function $h(t)$, among others.

The probability density function of a non-negative RV T, is given by

$$f(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t)}{\Delta t}, \quad f(t) \geq 0. \tag{2.1}$$

The survival function with probability of an observation does not fail until the time t is

$$S(t) = P[T > t] = 1 - P[T \leq t] = 1 - \int_0^t f(t)d(t) = 1 - F(t),\ \ 0 < S(t) < 1,$$

where $F(t)$ is cumulative distribution function.

The hazard function quantify the instantaneous risk of failure at a given time t and is given by

$$h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq +t\Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}, \quad h(t) \geq 0. \tag{2.2}$$

The Figure 1 presents shapes that the hazard function can assume such as constant, decreasing, increasing, unimodal, bathtub shape, among others.

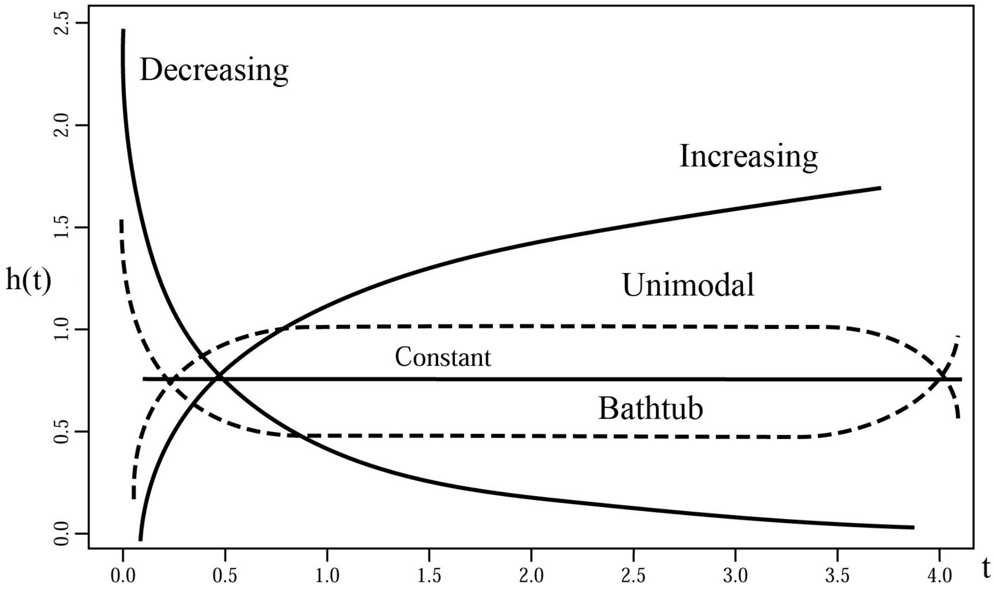

Figure 1 – Different shapes for the hazard function.

Some useful relationships can be obtained from these functions, such as:

$$f(t) = h(t)S(t), \quad h(t) = \frac{S'(t)}{S(t)} = \frac{\partial}{\partial t}\log(S(t)), \quad S(t) = \exp\left[-\int_0^t h(t)dt\right].$$

Glaser (1980) provided a helpful Lemma to study the behavior of the hazard function that is given as follow.

Lemma 2.1.1. Glaser (1980). Let T be a non-negative continuous random variable with twice differentiable PDF, $f(t|\theta)$. Then for $\eta(t|\theta) = -\frac{d}{dt}\log f(t|\theta)$, we have the following results:

1. If $\eta(t|\theta)$ has a decreasing (increasing) shape, then $h(t|\theta)$ has an increasing (decreasing) shape.
2. If $\eta(t|\theta)$ is bathtub (unimodal) shaped, then $h(t|\theta)$ is bathtub (unimodal) shaped.

2.1.1 TTT-plot

The TTT-plot (total time on test) is considered in order to verify the behavior of the empirical hazard function (BARLOW; CAMPO, 1975). The TTT-plot is obtained from the plot of $[r/n, G(r/n)]$ where

$$G(r/n) = \left(\sum_{i=1}^{r} t_i + (n-r)t_{(r)}\right) / \sum_{i=1}^{n} t_i,$$

$r = 1,\ldots,n, i = 1,\ldots,n$ and $t_{(r)}$ is the order statistics. If the curve is concave (convex), the hazard function has an increasing (decreasing) shape. If it starts convex and then becomes concave, or begins as concave and then becomes convex, then the hazard function is bathtub shaped, or inverse bathtub shaped, respectively. For more details see Figure 2.

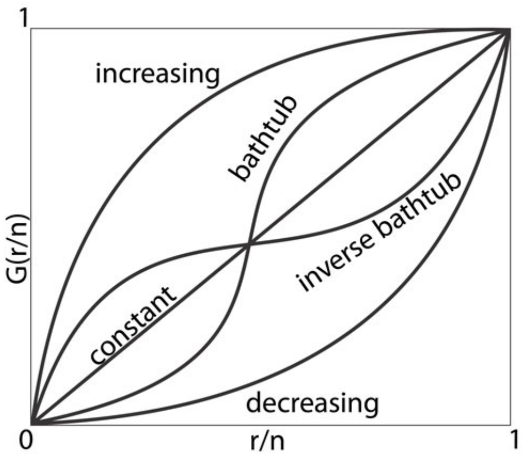

Figure 2 – TTT-plot shapes, retrieved from Ramos, Moala and Achcar (2014).

2.2 Frequentist inference

In a frequentist approach the unknown parameter vector $\theta = (\theta_1, \ldots, \theta_k)$ is considered as having fixed but unknown values. Different classical inferential procedures are available in the literature, such as the maximum likelihood estimators, method of moments, L-moments, ordinary and weighted least-squares, percentile, maximum product of spacings, the maximum goodness-of-fit estimators, among others (LOUZADA; RAMOS; PERDONÁ, 2016; BAKOUCH *et al.*, 2017; DEY *et al.*, 2017; RODRIGUES; LOUZADA; RAMOS, 2018). In this work, the maximum likelihood are considered to obtain the point and interval estimates under the classical approach.

2.2.1 Maximum Likelihood Estimation

The MLEs were chosen due to their good asymptotic properties. These estimators are obtained from maximizing the likelihood function (CASELLA; BERGER, 2002). The likelihood function of θ given t, is

$$L(\theta, t) = \prod_{i=1}^{n} f(t_i|\theta). \tag{2.3}$$

For a model with k parameters, if the likelihood function is differentiable at θ_i, the likelihood equations are obtained by solving the equations

$$\frac{\partial}{\partial \theta_i} \log(L(\theta, t)) = 0, i = 1, 2, \ldots, k. \tag{2.4}$$

Under mild conditions the solutions of (2.4) provide the maximum likelihood estimators. In many cases, numerical methods such as Newton-Raphson are required to find the solution of the nonlinear system. For large samples and under mild conditions they are consistent and efficient with an asymptotically multivariate normal distribution given by

$$\hat{\theta} \sim N_k[\theta, I^{-1}(\theta)] \text{ as } n \to \infty, \tag{2.5}$$

where $I(\theta)$ is the Fisher information matrix, $k \times k$ and $I_{ij}(\theta)$ is the Fisher information element of θ in i and j given by

$$I_{ij}(\theta) = E\left[-\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log(L(\theta,t))^2\right],\ i,j=1,2,\ldots,k. \tag{2.6}$$

For large samples, approximated confidence intervals can be constructed for the individuals parameters θ_i, $i=1,\ldots,k$, with confidence coefficient $100(1-\gamma)\%$, through marginal distributions given by

$$\hat{\theta}_i \sim N[\theta_i, I_{ii}^{-1}(\theta)] \text{ para } n \to \infty. \tag{2.7}$$

2.3 Bayesian Inference

So far, we have presented the estimation procedures using the frequentist approach. Bayesian analysis is an attractive framework in practical problems and became very popular in recent years. Here, we assume that the reader has a basic knowledge about Bayesian procedures. For an overview of Bayesian techniques, the reader is referred to Migon, Gamerman and Louzada (2014).

As the parameters are treated as random variables the distribution associated with such variables are known as prior distribution. The prior distribution is a key part of the Bayesian inference and there are different types of priors distribution available in the literature. Priors can be created using different procedures. For example, a prior distribution could be elicited from the assessment of an experienced expert (O'Hagan *et al.* (2006)). On the other hand, we could be interested in specifying a prior distribution, where the dominant information in the posterior distribution is provided by the data, such priors are known as noninformative prior. In this work, we considered different non-informative priors, such as the Jeffreys Rule (BOX; TIAO, 1973), Jeffreys prior (JEFFREYS, 1946), Maximal Data Information (MDI) prior (ZELLNER, 1977; ZELLNER, 1984) and Reference prior (BERNARDO, 1979). These priors are usually improper and could lead to improper posteriors. Therefore, we investigated if whether these priors lead to proper posterior distributions for the chosen models.

2.3.1 *Non-informative priors*

Jeffreys considered different scenarios for constructing non-informative priors. He considered the cases in which the parameter space was a bounded interval, $(-\infty,\infty)$, or $(0,\infty)$ (KASS; WASSERMAN, 1996). For these two cases, Jeffreys suggested using a constant prior. For $(0,\infty)$, he used the prior $\pi(\theta) = \frac{1}{\theta}$. His main justification for this choice was its invariance under power transformations of the parameters. Therefore, if $\theta \in (0,\infty)^k$, then the Jeffreys rule prior is given by

$$\pi(\theta) \propto \prod_{i=1}^{k} \frac{1}{\theta_i}. \tag{2.8}$$

In a further study, Jeffreys (1946) proposed his "general rule" in which the non-informative prior is obtained from the square root of the determinant of the Fisher information matrix $I(\theta)$ and θ is the vector of parameters. This prior has been widely used due to its invariance property under one-to-one transformations of parameters. For example, for any one-to-one function $\Phi = \Phi(\theta)$, the posterior $p(\Phi|t)$ obtained from the reparametrized distribution $f(t|\Phi)$ must be coherent with the posterior $p(\theta|t)$ obtained from the original distribution $f(t|\theta)$, in the sense that, $p(\Phi|t) = p(\theta|t)|\frac{d\theta}{d\Phi}|$.

The Jeffreys prior is obtained through the square root of the determinant of the Fisher information matrix $I(\theta)$ given by

$$\pi(\theta) \propto \sqrt{\det I(\theta)}. \tag{2.9}$$

Zellner (1977) introduced another procedure to derive a noninformative prior $\pi(\theta)$ in which the gain in the information supplied by the data is the largest as possible relative to the prior information, maximizing the information provided by the data. The resulting noninformative prior distribution is known as Maximal Data Information (MDI) prior and is defined as

$$\pi(\theta) \propto \exp\left(Q(\theta)\right), \tag{2.10}$$

where $Q(\theta)$ is the negative Shanon Entropy of $f(t\,|\theta)$ given by

$$Q(\theta) = \int f(t\,|\theta) \log f(t\,|\theta) dt, \tag{2.11}$$

i.e, one measure of the information of $f(t\,|\theta)$. The MDI prior has invariant limitations, in this case, is only invariant for linear transformations of T or θ.

Another important noninformative prior was introduced by Bernardo (1979) with further developments (BERGER; BERNARDO, 1989; BERGER; BERNARDO, 1992a; BERGER; BERNARDO, 1992b; BERGER; BERNARDO *et al.*, 1992; BERGER *et al.*, 2015). The proposed reference prior is minimally informative in a precise information-theoretic sense. Moreover, the information provided by the data dominate the prior information, reflecting the vague nature of the prior knowledge. To achieve such prior the authors maximize the expected Kullback-Leibler divergence between the posterior distribution and the prior. The obtained reference prior provides a posterior distribution with interesting properties, such as:

- Consistent marginalization: For all data t, if the posterior $p_1(\theta|t)$ obtained from original model $f(t|\theta,\lambda)$ is of the form $p_1(\theta|t) = p_1(\theta|x)$ for some statistic $x = x(t)$ whose sampling distribution $p(x|\theta,\lambda) = p(x|\theta)$ only depends on θ, then the posterior $p_2(\theta|x)$ obtained from the marginal distribution $f(x|\theta)$ must be the same as the posterior $p_1(\theta|x)$ obtained from $f(t|\theta,\lambda)$.

- Consistent sampling properties: The properties under repeated sampling of the posterior distribution must be consistent with the model $f(t|\theta)$ (COX; HINKLEY, 1979).
- Invariant under one-to-one transformations: The same properties presented in the Jeffreys prior section.

Bernardo (2005) presented different procedures to derive reference priors in the presence of nuisance parameters. The following propositions are useful to obtain the reference priors for the chosen models.

Proposition 2.3.1. Let $f(x|\theta,\lambda)$ be a parametric model, where θ is the parameter of interest, $\lambda=(\lambda_1,\ldots,\lambda_m)$ is a vector of nuisance parameters and $I(\theta,\lambda)$ Fisher's information matrix $(m+1)\times(m+1)$. It is assumed that the joint distribution of (θ,λ) is asymptotically normal with mean and covariance matrix $S(\hat{\theta},\hat{\lambda})=I^{-1}(\hat{\theta},\hat{\lambda})$, where $(\hat{\theta},\hat{\lambda})$ correspondents of MLEs. Moreover, S_j is a $j\times j$ upper left submatrix of S and $\imath_{i,j}(\theta,\lambda)$ is an element of I_j. If the nuisance parameter spaces $\wedge_i(\theta,\lambda_1,\ldots,\lambda_{j-1})=\wedge_i$ are independent of θ and λ_i's and the functions $\imath_{i,i},\ldots,\imath_{m,m},\;\; i=1,\ldots,m$ factorize in the form

$$s_{1,1}^{-\frac{1}{2}}(\theta,\lambda)=f_0(\theta)g_0(\lambda) \quad \text{and} \quad \imath_{i+1,i+1}^{\frac{1}{2}}(\theta,\lambda)=f_i(\lambda_i)g_i(\theta,\lambda_{-i}).$$

where $\lambda_{-i}=(\lambda_1,\ldots,\lambda_{i-1},\lambda_{i+1},\ldots,\lambda_m)$. Then

$$\pi(\theta)\propto f_0(\theta), \quad \pi(\lambda_i|\theta,\lambda_{-i})\propto f_i(\lambda_i), i=1,\ldots,m, \tag{2.12}$$

and the reference prior when θ is the parameter of interest and λ is the vector of nuisance parameters is given by $\pi_\theta(\theta,\lambda)=f_0(\theta)\prod_{i=1}^m f_j(\lambda_i)$.

Proposition 2.3.2. (BERGER *et al.*, 2015, p.196) Consider the unknown vector of parameters $\theta=(\theta_1,\ldots,\theta_m)$ and the posterior distribution $\mathrm{p}(\theta|t)$ with asymptotically normal distribution and dispersion matrix $S(\theta)=I^{-1}(\theta)$. If $I(\theta)$ is of the form

$$I(\theta)=\operatorname{diag}(f_1(\theta_1)g_1(\theta_{-1}),\ldots,f_m(\theta_m)g_m(\theta_{-m})),$$

where $f_i(\cdot)$ and $g_i(\cdot)$ are positive functions of θ_i for $i=1,\ldots,m$, then the one-at-a-time reference prior, for any chosen parameter of interest and any ordering of the nuisance parameters in the derivation, hereafter, the overall reference prior is given by

$$\pi(\theta)=\sqrt{f_1(\theta_1)\ldots f_m(\theta_m)}. \tag{2.13}$$

Tibshirani (1989) proposes an alternative method to derive a class of non-informative priors $\pi(\theta_1,\theta_2)$ where θ_1 is the parameter of interest so that the credible interval for θ_1 has a coverage error $O(n^{-1})$ in the frequentist sense, i.e.,

$$P\left[\theta_1\leq\theta_1^{1-\alpha}(\pi;X)|(\theta_1,\theta_2)\right]=1-\alpha-O(n^{-1}), \tag{2.14}$$

where $\theta_1^{1-\alpha}(\pi;X)|(\theta_1,\theta_2)$ denote the $(1-\alpha)$th quantile of the posterior distribution of θ_1. The class of priors satisfying (2.14) are known as matching priors.

To achieve this, Tibshirani (1989) proposed to reparametrize the model in terms of the orthogonal parameters (δ,λ) in the sense discussed by Cox and Reid (1987). That is, $I_{\delta,\lambda}(\delta,\lambda)=0$ for all (δ,λ), where δ is the parameter of interest and λ is the orthogonal nuisance parameter. Thus, the matching priors are all priors of the form

$$\pi(\delta,\lambda)=g(\lambda)\sqrt{I_{\delta\delta}(\delta,\lambda)}, \tag{2.15}$$

where $g(\lambda)>0$ is an arbitrary function and $I_{\delta\delta}(\delta,\,\lambda)$ is the δ entry of the Fisher information matrix. Further, Mukerjee and Dey (1993) discussed sufficiency and necessary conditions for a class of Tibshirani priors be matching prior up to $o(n^{-1})$.

2.4 Discrimination criterion methods

In situations that involve uncertainty measures, discrimination criterion methods are of great importance in statistical analysis as a goodness of fit for model selection. Let k be the number of parameters to be fitted and $\hat{\theta}$ is the estimate of θ some discrimination criterion methods based on log-likelihood function are given by

- Akaike information criterion: $AIC=-2\log(L(\hat{\theta};t))+2k$.
- Corrected Akaike information criterion: $AICC=AIC+\frac{2k(k+1)}{(n-k-1)}$.
- Bayesian information criterion: $BIC=-2\log(L(\hat{\theta};t))+k\log(n)$.

Given observed data and a set of candidate models the best model is the one which provides the minimum values. These procedures includes penalty discourages overfitting, i.e, increasing the number of parameters with poor predictive results. For an overview of discrimination criterion methods, the reader is referred to Burnham and Anderson (2004).

2.5 Some Useful Mathematical Results

In this section, we present some useful propositions that are used to prove some posterior properties.

Let $\mathbb{R}^+$ denote the strictly positive real numbers.

Definition 2.5.1. Let $\mathrm{g}:\mathscr{U}\to\mathbb{R}^+$ and $\mathrm{h}:\mathscr{U}\to\mathbb{R}^+$, where $\mathscr{U}\subset\mathbb{R}$, and let $a\in\overline{\mathbb{R}}$. We can say that $\mathrm{g}(x)\underset{x\to a}{\propto}\mathrm{h}(x)$ if

$$\liminf_{x\to a}\frac{\mathrm{g}(x)}{\mathrm{h}(x)}>0 \ \text{ and } \ \limsup_{x\to a}\frac{\mathrm{g}(x)}{\mathrm{h}(x)}<\infty.$$

We define the meaning of the relations $\mathrm{g}(x)\underset{x\to a^+}{\propto}\mathrm{h}(x)$ and $\mathrm{g}(x)\underset{x\to a^-}{\propto}\mathrm{h}(x)$ for $a\in\mathbb{R}$ analogously.

The following propositions and definitions are useful to prove the results related to the posterior distribution.

We denoted by $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ the *extended real number line* with its usual order ($\leq$), while $\mathbb{R}^+$ denotes the strictly positive real numbers, $\overline{\mathbb{R}^+} = \mathbb{R}^+ \cup \{\infty\}$ and $\mathbb{R}_0^+$ denotes the non-negative real numbers.

Definition 2.5.2. Let $\mathrm{g} : \mathscr{U} \to \overline{\mathbb{R}^+}$ and $\mathrm{h} : \mathscr{U} \to \overline{\mathbb{R}^+}$, where $\mathscr{U} \subset \mathbb{R}$. We will say that $\mathrm{g}(x) \propto \mathrm{h}(x)$ if there is $c_0 \in \mathbb{R}^+$ and $c_1 \in \mathbb{R}^+$ such that $c_0\,\mathrm{h}(x) \leq \mathrm{g}(x) \leq c_1\,\mathrm{h}(x)$ for every $x \in \mathscr{U}$.

Note that if for some $c \in \mathbb{R}^+$ we have $\lim_{x\to a} \frac{\mathrm{g}(x)}{\mathrm{h}(x)} = c$, then $\mathrm{g}(x) \underset{x\to a}{\propto} \mathrm{h}(x)$. The following proposition relates Definition 2.5.2 and Definition 2.5.1 for $\mathscr{U} = (a,b)$.

Proposition 2.5.3. Let $\mathrm{g} : (a,b) \to \mathbb{R}^+$ and $\mathrm{h} : (a,b) \to \mathbb{R}^+$ be continuous functions on $(a,b) \subset \mathbb{R}$, where $a \in \overline{\mathbb{R}}$ and $b \in \overline{\mathbb{R}}$. Then $\mathrm{g}(x) \propto \mathrm{h}(x)$ if and only if $\mathrm{g}(x) \underset{x\to a}{\propto} \mathrm{h}(x)$ and $\mathrm{g}(x) \underset{x\to b}{\propto} \mathrm{h}(x)$.

Proof. Suppose $\mathrm{g}(x) \underset{x\to a}{\propto} \mathrm{h}(x)$ and $\mathrm{g}(x) \underset{x\to b}{\propto} \mathrm{h}(x)$. Then, by Definition 2.5.1, $\liminf_{x\to a} \dfrac{\mathrm{g}(x)}{\mathrm{h}(x)} = w_1$ and $\limsup_{x\to a} \dfrac{\mathrm{g}(x)}{\mathrm{h}(x)} = w_2$ for some w_1 and w_2 both in $(0,\infty)$. Therefore, from the definition of liminf and limsup there is some $a' \in (a,b)$ such that $\dfrac{w_1}{2} \leq \dfrac{\mathrm{g}(x)}{\mathrm{h}(x)} \leq \dfrac{3w_2}{2}$ for every $x \in (a,a']$. Analogously, there is some v_1 and v_2, both in $(0,\infty)$, and $b' \in (0,\infty)$ such that $\dfrac{v_1}{2} \leq \dfrac{\mathrm{g}(x)}{\mathrm{h}(x)} \leq \dfrac{3v_2}{2}$ for every $v \in [b',b)$. On the other hand, since $\dfrac{\mathrm{g}(x)}{\mathrm{h}(x)}$ is continuous in $[a',b']$, the Weierstrass extreme value Theorem (RUDIN *et al.*, 1964) states that there is some x_0 and $x_1 \in [a',b']$ such that $\dfrac{\mathrm{g}(x_1)}{\mathrm{h}(x_1)} \leq \dfrac{\mathrm{g}(x)}{\mathrm{h}(x)} \leq \dfrac{\mathrm{g}(x_2)}{\mathrm{h}(x_2)}$ for every $x \in [a',b']$. Finally, choosing $m = \min\left(\dfrac{w_1}{2}, \dfrac{v_1}{2}, \dfrac{\mathrm{g}(x_1)}{\mathrm{h}(x_1)}\right) > 0$ and $M = \max\left(\dfrac{3w_2}{2}, \dfrac{3v_2}{2}, \dfrac{\mathrm{g}(x_2)}{\mathrm{h}(x_2)}\right) < \infty$, it follows from the above considerations that $m \leq \dfrac{\mathrm{g}(x)}{\mathrm{h}(x)} \leq M$ for every $x \in (a,b)$, which by Definition 2.5.2 means that $g(x) \propto h(x)$.

Suppose $g(x) \propto h(x)$. By Definition 2.5.2, there are some $m > 0$ and $M < 0$ such that $m \leq \dfrac{\mathrm{g}(x)}{\mathrm{h}(x)} \leq M$ for every $x \in (a,b)$. This implies that

$$\liminf_{x\to a} \frac{\mathrm{g}(x)}{\mathrm{h}(x)} \geq m > 0 \text{ and } \limsup_{x\to a} \frac{\mathrm{g}(x)}{\mathrm{h}(x)} \leq M < \infty,$$

which by Definition 2.5.1 means that $\mathrm{g}(x) \underset{x\to a}{\propto} \mathrm{h}(x)$. The proof that $\mathrm{g}(x) \underset{x\to b}{\propto} \mathrm{h}(x)$ is analogous to the previous case. □

As a direct consequence of Proposition 2.5.3, we have the following Proposition. This is the main proposition that will be used in the next subsection.

Proposition 2.5.4. Let $\mathrm{g}:(a,b)\to\mathbb{R}^+$ and $\mathrm{h}:(a,b)\to\mathbb{R}^+$ be continuous functions in $(a,b)\subset\mathbb{R}$, where $a\in\overline{\mathbb{R}}$ and $b\in\overline{\mathbb{R}}$, and let $c\in(a,b)$. Then if $\mathrm{g}(x)\underset{x\to a}{\propto}\mathrm{h}(x)$ or $\mathrm{g}(x)\underset{x\to b}{\propto}\mathrm{h}(x)$ we have

$$\int_a^c g(t)\,dt \propto \int_a^c h(t)\,dt \quad\text{or}\quad \int_c^b g(t)\,dt \propto \int_c^b h(t)\,dt\,.$$

Proof. Suppose $\mathrm{g}(x)\underset{x\to a}{\propto}\mathrm{h}(x)$. By continuity and non-nullity of $g(x)$ and $h(x)$ in c we have $\mathrm{g}(x)\underset{x\to c}{\propto}\mathrm{h}(x)$. Therefore, by Proposition 2.5.3, we have that $g(x)\propto h(x)$ in (a,c). This implies that

$$\int_a^c g(t)\,dt \propto \int_a^c h(t)\,dt\,.$$

The proof of the case $\mathrm{g}(x)\underset{x\to b}{\propto}\mathrm{h}(x)$ is analogous. □

The following propositions are useful to prove the results related to the posterior distribution. Let $\mathbb{R}^+$ denote the positive real numbers and $\mathbb{R}_0^+$ denote the positive real numbers including 0.

Definition 2.5.5. Let $\mathrm{g}:\mathscr{U}\to\overline{\mathbb{R}}_0^+$ and $\mathrm{h}:\mathscr{U}\to\overline{\mathbb{R}}_0^+$, where $\mathscr{U}\subset\mathbb{R}$. We say that $\mathrm{g}(x)\lesssim\mathrm{h}(x)$ if there exist $M\in\mathbb{R}^+$ such that $\mathrm{g}(x)\le M\,\mathrm{h}(x)$ for every $x\in\mathscr{U}$. If $\mathrm{g}(x)\lesssim\mathrm{h}(x)$ and $\mathrm{h}(x)\lesssim\mathrm{g}(x)$ then we say that $\mathrm{g}(x)\propto\mathrm{h}(x)$.

Definition 2.5.6. Let $a\in\overline{\mathbb{R}}$, $\mathrm{g}:\mathscr{U}\to\mathbb{R}^+$ and $\mathrm{h}:\mathscr{U}\to\mathbb{R}^+$, where $\mathscr{U}\subset\mathbb{R}$. We say that $\mathrm{g}(x)\underset{x\to a}{\lesssim}\mathrm{h}(x)$ if $\limsup_{x\to a}\dfrac{\mathrm{g}(x)}{\mathrm{h}(x)}<\infty$. If $\mathrm{g}(x)\underset{x\to a}{\lesssim}\mathrm{h}(x)$ and $\mathrm{h}(x)\underset{x\to a}{\lesssim}\mathrm{g}(x)$ then we say that $\mathrm{g}(x)\underset{x\to a}{\propto}\mathrm{h}(x)$.

The meaning of the relations $\mathrm{g}(x)\underset{x\to a^+}{\lesssim}\mathrm{h}(x)$ and $\mathrm{g}(x)\underset{x\to a^-}{\lesssim}\mathrm{h}(x)$ for $a\in\mathbb{R}$ are defined analogously.

Note that, if for some $c\in\mathbb{R}^+$ we have $\lim_{x\to a}\dfrac{\mathrm{g}(x)}{\mathrm{h}(x)}=c$, then $\mathrm{g}(x)\underset{x\to a}{\propto}\mathrm{h}(x)$. The following proposition is a direct consequence of the above definition.

Proposition 2.5.7. For $a\in\overline{\mathbb{R}}$ and $r\in\mathbb{R}^+$, let $f_1(x)\underset{x\to a}{\lesssim}f_2(x)$ and $g_1(x)\underset{x\to a}{\lesssim}g_2(x)$ then the following hold

$$f_1(x)g_1(x)\underset{x\to a}{\lesssim}f_2(x)g_2(x) \quad\text{and}\quad f_1(x)^r\underset{x\to a}{\lesssim}f_2(x)^r.$$

The following proposition relates Definition 2.5.5 and Definition 2.5.6.

Proposition 2.5.8. Let $\mathrm{g}:(a,b)\to\mathbb{R}^+$ and $\mathrm{h}:(a,b)\to\mathbb{R}^+$ be continuous functions on $(a,b)\subset\mathbb{R}$, where $a\in\overline{\mathbb{R}}$ and $b\in\overline{\mathbb{R}}$. Then $\mathrm{g}(x)\lesssim\mathrm{h}(x)$ if and only if $\mathrm{g}(x)\underset{x\to a}{\lesssim}\mathrm{h}(x)$ and $\mathrm{g}(x)\underset{x\to b}{\lesssim}\mathrm{h}(x)$.

Proof. Suppose that $\mathrm{g}(x)\underset{x\to a}{\lesssim}\mathrm{h}(x)$ and $\mathrm{g}(x)\underset{x\to b}{\lesssim}\mathrm{h}(x)$. Then, by Definition 2.5.6, $\limsup_{x\to a}\dfrac{\mathrm{g}(x)}{\mathrm{h}(x)}=w$ for some $w\in\mathbb{R}^+$. Therefore, from the definition of lim sup there exist some $a' \in (a,b)$ such that $\dfrac{\mathrm{g}(x)}{\mathrm{h}(x)} \leq \dfrac{3w}{2}$ for every $x \in (a,a']$. Proceeding analogously, there must exist some $v \in \mathbb{R}^+$ and $b' \in (a',b)$ such that $\dfrac{\mathrm{g}(x)}{\mathrm{h}(x)} \leq \dfrac{3v}{2}$ for every $x \in [b',b)$. On the other hand, since $\dfrac{\mathrm{g}(x)}{\mathrm{h}(x)}$ is continuous in $[a',b']$, the Weierstrass Extreme Value Theorem states that there exist some $x_1 \in [a',b']$ such that $\dfrac{\mathrm{g}(x)}{\mathrm{h}(x)} \leq \dfrac{\mathrm{g}(x_1)}{\mathrm{h}(x_1)}$ for every $x \in [a',b']$. Finally, choosing $M = \max\left(\dfrac{3w}{2}, \dfrac{3v}{2}, \dfrac{\mathrm{g}(x_1)}{\mathrm{h}(x_1)}\right) < \infty$, it follows that $\dfrac{\mathrm{g}(x)}{\mathrm{h}(x)} \leq M$ for every $x \in (a,b)$, which by Definition 2.5.5 means that $g(x) \lesssim h(x)$.

Now suppose $g(x) \lesssim h(x)$. By Definition 2.5.5, there exist some $M < 0$ such that $\dfrac{\mathrm{g}(x)}{\mathrm{h}(x)} \leq M$ for every $x \in (a,b)$. This implies that $\limsup_{x\to a} \frac{\mathrm{g}(x)}{\mathrm{h}(x)} \leq M < \infty$ which by Definition 2.5.6 means that $\mathrm{g}(x) \underset{x\to a}{\lesssim} \mathrm{h}(x)$. The proof that $\mathrm{g}(x) \underset{x\to b}{\lesssim} \mathrm{h}(x)$ must also be satisfied is analogous to the previous case. Therefore the theorem is proved. $\square$

Note that if $\mathrm{g} : (a,b) \to \mathbb{R}^+$ and $\mathrm{h} : (a,b) \to \mathbb{R}^+$ are continuous functions on $(a,b) \subset \mathbb{R}$, then by continuity it follows directly that $\lim_{x\to c} \dfrac{\mathrm{g}(x)}{\mathrm{h}(x)} = \dfrac{\mathrm{g}(c)}{\mathrm{h}(c)} > 0$ and therefore $\mathrm{g}(x) \underset{x\to c}{\propto} \mathrm{h}(x)$ for every $c \in (a,b)$. This fact and the Proposition 2.5.8 imply directly the following.

Proposition 2.5.9. Let $\mathrm{g} : (a,b) \to \mathbb{R}^+$ and $\mathrm{h} : (a,b) \to \mathbb{R}^+$ be continuous functions in $(a,b) \subset \mathbb{R}$, where $a \in \overline{\mathbb{R}}$ and $b \in \overline{\mathbb{R}}$, and let $c \in (a,b)$. Then if $\mathrm{g}(x) \underset{x\to a}{\lesssim} \mathrm{h}(x)$ (or $\mathrm{g}(x) \underset{x\to b}{\lesssim} \mathrm{h}(x)$) we have that $\int_a^c g(t)\,dt \lesssim \int_a^c h(t)\,dt$ (respectively $\int_c^b g(t)\,dt \lesssim \int_c^b h(t)\,dt$).

CHAPTER 3

NAKAGAMI-M DISTRIBUTION

3.1 Introduction

The Nakagami-m (NK) distribution is a powerful statistical tool for modeling fading radio signals. Proposed by Nakagami (1960), this model has received considerable attention due to its flexibility to describe a wide range of communication engineering problems. For instance, considering the IEEE Xplore digital Library, a search carried out using the word ”Nakagami” on April 2017 found 3,660 research papers.

The NK distribution has been used successfully in other fields such as medical imaging processing (SHANKAR *et al.*, 2001; TSUI; HUANG; WANG, 2006), hydrologic engineering (SARKAR; GOEL; MATHUR, 2009; SARKAR; GOEL; MATHUR, 2010), seismological analysis (CARCOLE; SATO, 2009; NAKAHARA; CARCOLÉ, 2010) and traffic modeling of multimedia data (KIM; LATCHMAN, 2009). However, there are no comprehensive references in the literature which consider the NK distribution as a reliability model. In this chapter, we present the reliability properties for this model and also prove that its hazard rate (mean residual life) function presents increasing (decreasing) or bathtub (unimodal) shapes.

The parameter estimations of the NK distribution were discussed earlier. An unbiased estimator for parameter Ω is easily obtained using the method of moments (NAKAGAMI, 1960). However, considerable effort has been made to derive efficient estimators for the fading parameter. Cheng and Beaulieu (2001) considered the maximum likelihood (ML) method. Cheng and Beaulieu (2002) further suggested an estimator based on generalized moments (GM). Zhang (2002) numerically compared the accuracy of various estimators and suggested using an approximation to the ML estimators. Gaeddert and Annamalai (2004) developed estimators based on approximations of transcendental equations that arose when computing the ML and GM estimators. However, these estimators are approximations to the natural procedures and are motivated by fast computation, avoiding solving nonlinear equations. Wang, Song and

Cheng (2012) proposed a closed-form estimator for the fading parameter obtained as a limiting procedure of the traditional GM estimators. However, these estimators depend on the asymptotic properties to construct the confidence intervals.

Considering a Bayesian approach, Son and Oh (2007) discussed Bayes estimation using independent gamma prior distributions for the parameters of the NK distribution. However, Bernardo (2005) argued that using simple proper priors, presumed to be non-informative, often hides important unwarranted assumptions which may easily dominate, or even invalidate, the statistical analysis and should be strongly discouraged. Beaulieu and Chen (2007) discussed MAP estimators using informative priors. However, in applications, it is difficult to obtain prior information for the unknown parameters. To overcome this problem, a Bayesian analysis can be performed with non-informative priors, i.e., priors constructed by formal rules.

In this chapter, different objective priors for the NK distribution are presented such as Jeffreys's rule, Jeffreys prior, the MDI prior and the reference prior. These priors are improper and could lead to improper posteriors. We propose a theorem that provides sufficient and necessary conditions for a general class of posterior to be proper posterior distributions. The proposed theorem is used to investigate if these priors lead to proper or improper posterior distributions. Later, a posterior distribution based on the reference prior is obtained. This proper posterior returns better numerical results and also excellent theoretical properties such as invariance property under one-to-one transformations of the parameters, consistency under marginalization and consistent sampling properties. The proposed posterior distribution also satisfies the matching prior properties for both Ω and μ. Finally, our methodology is illustrated using two real lifetime data sets, proving that the NK distribution can be used to describe lifetime data.

The remainder of this chapter is organized as follows. Section 2 presents mathematical properties for the NK and reviews two common classical approaches. Section 3 presents the main theorem that provides sufficient and necessary conditions for a general class of posterior to be proper with applications in non-informative priors. In Section 4, a simulation study is presented in order to identify the most efficient estimation procedure. Section 5 presents an analysis of two lifetime data sets. Finally, Section 6 summarizes the study.

3.2 Nakagami-m distribution

Let T be a random variable with NK distribution, the PDF is given by

$$f(t|\theta) = \frac{2}{\Gamma(\mu)}\left(\frac{\mu}{\Omega}\right)^{\mu} t^{2\mu-1}\exp\left(-\frac{\mu}{\Omega}t^2\right), \tag{3.1}$$

for all $t>0$, where $\theta=(\mu,\Omega)$, $\mu \geq 0.5$ and $\Omega>0$ are, respectively, the shape (also known as a fading parameter) and scale parameters and $\Gamma(\phi)=\int_0^\infty e^{-x}x^{\phi-1}dx$ is the gamma function. We use Ω and μ to represent the parameters since they are commonly used in signal processing.

Important probability distributions can be obtained from the NK distribution such as the Rayleigh distribution ($\mu = 1$) and the half-normal distribution ($\mu = 0.5$). Moreover, this model is also related to the gamma distribution. For instance, if $Y \sim \text{Gamma}(a,b)$, then $T = \sqrt{Y}$ has a NK distribution with $\mu = a$ and $\Omega = ab$. Due to this relationship, the μ parameter can also take on values between $0 < \mu < 0.5$.

Let T be a continuous lifetime (non-negative) random variable with NK distribution. The raw moments are given as

$$E(T^r) = \frac{\Gamma(\mu + r/2)}{\Gamma(\mu)} \left(\frac{\Omega}{\mu}\right)^{r/2}, \tag{3.2}$$

for $r \in \mathbb{N}$. After some algebraic manipulation, the mean and variance of (3.1) are respectively given by

$$E(T) = \frac{\Gamma(\mu + 1/2)}{\Gamma(\mu)} \left(\frac{\Omega}{\mu}\right)^{\frac{1}{2}} \text{ and} \tag{3.3}$$

$$Var(T) = \Omega \left(1 - \left(\frac{\Gamma(\mu + 1/2)}{\Gamma(\mu)}\right)^2\right). \tag{3.4}$$

The median and the mode of the NK distribution are

$$\text{Med}(T) = \sqrt{\Omega} \quad \text{and} \quad \text{Mode}(T|\theta) = \frac{\sqrt{2}}{2}\left(\frac{(2\mu - 1)\Omega}{\mu}\right)^{\frac{1}{2}}.$$

The reliability function that represents the probability that an observation does not fail until t is

$$S(t|\theta) = \frac{1}{\Gamma(\mu)} \Gamma\left(\mu, \frac{\mu}{\Omega} t^2\right),$$

where $\Gamma(y,x) = \int_x^\infty w^{y-1} e^{-w} dw$ is the upper incomplete gamma function. For the NK distribution, the hazard function is given by

$$h(t|\theta) = 2\left(\frac{\mu}{\Omega}\right)^{\mu} t^{2\mu - 1} \exp\left(-\frac{\mu}{\Omega} t^2\right) \Gamma\left(\mu, \frac{\mu}{\Omega} t^2\right)^{-1}. \tag{3.5}$$

Theorem 3.2.1. The hazard rate function $h(t|\theta)$ of the NK distribution is bathtub (increasing) shaped for $0 < \mu < 0.5$ ($\mu \geq 0.5$), for all $\Omega > 0$.

Proof. Firstly

$$\eta(t|\theta) = -\frac{d}{dt} \log f(t|\theta) = -\frac{(2\mu - 1)}{t} - \frac{2\mu t}{\Omega}. \tag{3.6}$$

Following Lemma 2.1.1, for $\mu \geq 0.5$ and $\Omega > 0$, $\eta(t|\theta)$ has a decreasing shape, i.e., $h(t|\theta)$ has an increasing shape. For $0 < \mu < 0.5$ and $\Omega > 0$, $\eta(t|\theta)$ is bathtub shaped with a global minimum at $t^* = \sqrt{\Omega - \frac{\Omega}{2\mu}}$. Therefore, $h(t|\theta)$ is also bathtub shaped. $\square$

The behavior of the hazard function (3.5) when $t \to 0$ and $t \to \infty$ is given by

$$h(0|\theta) = \begin{cases} \infty, & \text{if } \mu < 0.5 \\ \sqrt{\dfrac{2}{\Omega}}, & \text{if } \mu = 0.5 \\ 0, & \text{if } \mu > 0.5 \end{cases} \quad \text{and} \quad h(\infty|\theta) = \infty.$$

Figure 3 presents examples for the shapes of the hazard function for different values of μ and Ω.

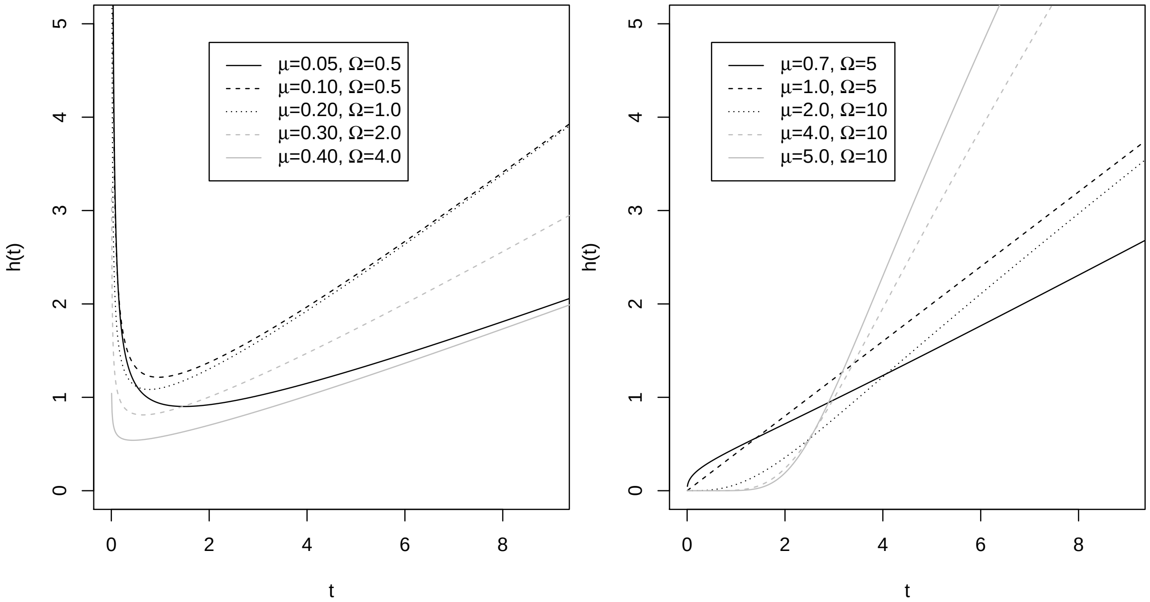

Figure 3 – Hazard function shapes for NK distribution considering different values of μ and Ω.

The mean residual life (MRL) represents the expected additional lifetime given that a component has survived until time t.

Proposition 3.2.2. The mean residual life function $r(t|\theta)$ of the NK distribution is given by

$$r(t|\theta) = \frac{1}{S(t|\theta)} \int_t^{\infty} y f(y|\theta)) dy - t = \sqrt{\frac{\Omega}{\mu}} \left(\frac{\Gamma\left(\mu + \frac{1}{2}, \frac{\mu}{\Omega} t^2\right)}{\Gamma\left(\mu, \frac{\mu}{\Omega} t^2\right)} \right) - t. \tag{3.7}$$

The behaviors of the MRL function (3.5) when $t \to 0$ and $t \to \infty$ are, respectively

$$r(0|\theta) = \sqrt{\frac{\Omega}{\mu}} \left(\frac{\Gamma\left(\mu + \frac{1}{2}\right)}{\Gamma(\mu)} \right) \quad \text{and} \quad r(\infty|\theta) = \frac{1}{h(\infty|\theta)} = 0.$$

The following Lemma is useful to obtain the shapes of the MRL function.

Lemma 3.2.3. Let T be a continuous lifetime random variable with hazard function $h(t|\theta)$ and mean residual life function $r(t|\theta)$.

1. If $h(t|\theta)$ has a decreasing (increasing) shape, then $r(t|\theta)$ has an increasing (decreasing) shape (BRYSON; SIDDIQUI, 1969) .

2. If $h(t|\theta)$ is bathtub shaped and $h(0)r(0) > 1$, then $r(t|\theta)$ is unimodal shaped (OLCAY, 1995).

Theorem 3.2.4. The mean residual life function $r(t|\theta)$ of the NK distribution has a unimodal (decreasing) shape for $0 < \mu < 0.5\,(\mu \geq 0.5)$, for all $\Omega > 0$.

Proof. For $\mu \geq 0.5$ and $\Omega > 0$, $h(t|\theta)$ has an increasing shape. Then, by Lemma 3.2.3, $r(t|\theta)$ has a decreasing shape. For $\Omega > 0$ and $0 < \mu < 0.5$, $h(t|\theta)$ has a bathtub shape and $h(0)r(0) > 1$. Therefore, based on Lemma 3.2.3, $r(t|\theta)$ has a unimodal shape. □

Figure 4 presents examples of the shapes of the mean residual life function for different values of μ and Ω.

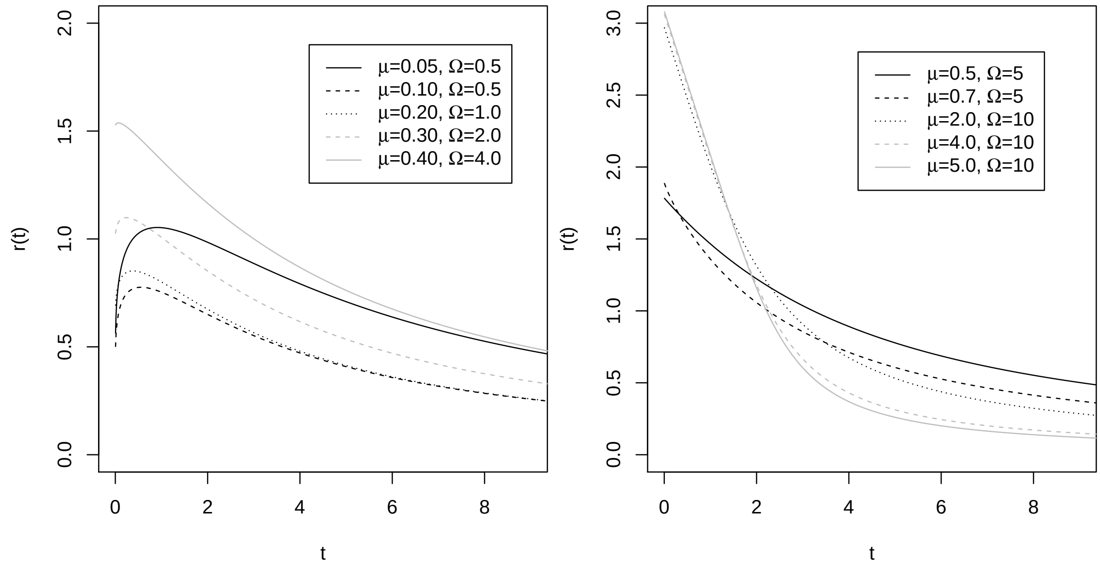

Figure 4 – Mean residual life function shapes for NK distribution considering different values of μ and Ω.

3.3 Classical Inference

3.3.1 Moment Estimators

The method of moments (MM) is one of the oldest methods used for estimating parameters in statistical models. Nakagami (1960) proposed the following moment estimators

$$\hat{\mu} = \frac{\left(\sum_{i=1}^{n} t_i^2\right)^2}{\left(n\sum_{i=1}^{n} t_i^4\right) - \left(\sum_{i=1}^{n} t_i^2\right)^2} \quad \text{and} \quad \hat{\Omega} = \frac{1}{n}\sum_{i=1}^{n} t_i^2. \tag{3.8}$$

Note that the second moment of the NK distribution is Ω. Therefore, $\hat{\Omega} = \sum_{i=1}^{n} t_i^2/n$ is an unbiased estimator for Ω.

3.3.2 Maximum Likelihood Estimation

Let $T_1, \ldots, T_n$ be a random sample such that $T \sim \mathrm{NK}(\Omega, \mu)$. The likelihood function from (3.1) is given by

$$L(\theta|t) = \frac{2^n}{\Gamma(\mu)^n}\left(\frac{\mu}{\Omega}\right)^{n\mu}\left\{\prod_{i=1}^{n} t_i^{2\mu-1}\right\}\exp\left(-\frac{\mu}{\Omega}\sum_{i=1}^{n} t_i^2\right). \tag{3.9}$$

The log-likelihood function is

$$\log L(\theta|t) = n\log(2) - n\log(\Gamma(\mu)) + n\mu\log\left(\frac{\mu}{\Omega}\right) + (2\mu-1)\sum_{i=1}^{n}\log(t_i) - \frac{\mu}{\Omega}\sum_{i=1}^{n} t_i^2. \tag{3.10}$$

The estimates are obtained by maximizing the likelihood function. From the expressions $\frac{\partial}{\partial\Omega}\log L(\theta|t) = 0$, $\frac{\partial}{\partial\mu}\log L(\theta|t) = 0$, the likelihood equations are given as

$$n\left(1+\log\left(\frac{\mu}{\Omega}\right)\right) - n\psi(\mu) + 2\sum_{i=1}^{n}\log(t_i) = \frac{1}{\Omega}\sum_{i=1}^{n} t_i^2 \tag{3.11}$$

and

$$-\frac{n\mu}{\Omega} + \frac{\mu}{\Omega^2}\sum_{i=1}^{n} t_i^2 = 0, \tag{3.12}$$

where $\psi(k) = \frac{\partial}{\partial k}\log\Gamma(k) = \frac{\Gamma'(k)}{\Gamma(k)}$ is the digamma function. The MLE for $\hat{\Omega}$ is

$$\hat{\Omega} = \frac{1}{n}\sum_{i=1}^{n} t_i^2. \tag{3.13}$$

Substituting $\hat{\Omega}$ in (3.11), the estimate for $\hat{\mu}$ can be obtained solving

$$\log(\mu) - \psi(\mu) = \log\left(\frac{1}{n}\sum_{i=1}^{n} t_i^2\right) - \frac{2}{n}\sum_{i=1}^{n}\log(t_i). \tag{3.14}$$

Under mild conditions, the MLE are asymptotically normally distributed with a joint bivariate normal distribution given by

$$(\hat{\mu}_{MLE}, \hat{\Omega}_{MLE}) \sim N_2[(\mu,\Omega), I^{-1}(\mu,\Omega))] \text{ for } n \to \infty,$$

where $I(\mu,\Omega)$ is the Fisher information matrix

$$I(\mu,\Omega) = n\begin{bmatrix} \dfrac{(\mu\psi'(\mu)-1)}{\mu} & 0 \\ 0 & \dfrac{\mu}{\Omega^2} \end{bmatrix}, \tag{3.15}$$

and $\psi'(k) = \frac{\partial}{\partial k}\psi(k)$ is the trigamma function. These results are useful to construct asymptotic confidence intervals for the parameters of the NK distribution.

3.4 Bayesian Analysis

In this section, sufficient and necessary conditions are presented for a general class of posterior to obtain proper posterior distributions. Our proposed methodology is illustrated in different non-informative priors for parameters μ and Ω of the NK distribution.

3.4.1 Proper Posterior

The joint posterior distribution for θ is equal to the product of the likelihood function (3.9) and the joint prior distribution $\pi(\theta)$ divided by a normalizing constant $d(t)$, resulting in

$$p(\theta|t) = \frac{1}{d(t)}\frac{\pi(\theta)}{\Gamma(\mu)^n}\left(\frac{\mu}{\Omega}\right)^{n\mu}\left\{\prod_{i=1}^{n} t_i^{2\mu-1}\right\}\exp\left(-\frac{\mu}{\Omega}\sum_{i=1}^{n} t_i^2\right), \tag{3.16}$$

where

$$d(t) = \int_{\mathscr{A}} \frac{\pi(\theta)}{\Gamma(\mu)^n}\left(\frac{\mu}{\Omega}\right)^{n\mu}\left\{\prod_{i=1}^{n} t_i^{2\mu-1}\right\}\exp\left(-\frac{\mu}{\Omega}\sum_{i=1}^{n} t_i^2\right)d\theta, \tag{3.17}$$

and $\mathscr{A} = \{(0,\infty)\times(0,\infty)\}$ is the parameter space of θ. For any prior distribution in the form $\pi(\theta) \propto \pi(\mu)\pi(\Omega)$, our purpose is to find sufficient and necessary conditions for the posterior to be proper, i.e., $d(t) < \infty$.

Theorem 3.4.1. Suppose the behavior of $\pi(\Omega)$ is given by $\pi(\Omega) \propto \Omega^k$, for $k \in \mathbb{R}$ with $k > -1$ and $\pi(\mu)$ is strictly positive. Then, the posterior distribution (3.16) is improper. On the other hand, suppose that the behavior of $\pi(\Omega)$ and $\pi(\mu)$ is given by

$$\pi(\Omega) \propto \Omega^k, \quad \pi(\mu) \underset{\mu\to 0^+}{\propto} \mu^{r_0} \quad \text{and} \quad \pi(\mu) \underset{\mu\to\infty}{\propto} \mu^{r_\infty},$$

for $k \in \mathbb{R}$ with $k \leq -1$, $r_0 \in \mathbb{R}$ and $r_\infty \in \mathbb{R}$. Then, the posterior distribution (3.17) is proper if and only if $n > -r_0$ in case $k = -1$, and is proper if and only if $n > -r_0 - k - 2$ in case $k < -1$.

Proof. Let $\mathscr{B} = \{(0,\infty)\times(0,\infty)\}$ and consider the change of coordinates through the transformation $\theta : \mathscr{B} \to \mathscr{A}$

$$\theta(\phi,\lambda) = (\mu(\phi,\lambda),\Omega(\phi,\lambda)) = \left(\phi, \frac{\phi}{\lambda}\right). \tag{3.18}$$

Note that $\mathscr{A} = \theta(\mathscr{B})$. Since $|\det(D\theta(\phi,\lambda))| = \phi\lambda^{-2}$ (where $D\theta(\phi,\lambda)$ is the Jacobian matrix of the function $\theta(\phi,\lambda)$), denoting $\Theta = (\phi,\lambda)$ and applying the Theorem of Change of Variables on the Lebesgue integral (FOLLAND, 1999), we have that

$$\begin{aligned} d(t) &\propto \int_{\mathscr{A}} \frac{\Omega^k \pi(\mu)}{\Gamma(\mu)^n}\left(\frac{\mu}{\Omega}\right)^{n\mu}\left\{\prod_{i=1}^{n} t_i^{2\mu}\right\}\exp\left(-\frac{\mu}{\Omega}\sum_{i=1}^{n} t_i^2\right)d\theta \\ &= \int_{\mathscr{B}} \frac{\phi^{k+1}\pi(\phi)\lambda^{n\phi-k-2}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n} t_i^{2\phi}\right\}\exp\left(-\lambda\sum_{i=1}^{n} t_i^2\right)d\Theta. \end{aligned}$$

Since $\phi^{k+1}\pi(\phi)\lambda^{n\phi-k-2}(\Gamma(\phi)^n)^{-1}\prod_{i=1}^n t_i^{2\phi}\exp\left(-\lambda\sum_{i=1}^n t_i^2\right) \geq 0$, by the Fubini-Tonelli Theorem (FOLLAND, 1999), we have

$$\begin{aligned} d(t) &\propto \int_{\mathscr{B}} \frac{\phi^{k+1}\pi(\phi)\lambda^{n\phi-k-2}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n t_i^{2\phi}\right\}\exp\left\{-\lambda\sum_{i=1}^n t_i^2\right\}d\Theta \\ &= \int_0^\infty \frac{\phi^{k+1}\pi(\phi)}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n t_i^{2\phi}\right\}\int_0^\infty \lambda^{n\phi-k-2}\exp\left\{-\lambda\sum_{i=1}^n t_i^2\right\}d\lambda d\phi\,. \end{aligned} \tag{3.19}$$

The rest of the proof is divided into two cases:

- *Case* $k > -1$.

Since $e^{-bx} \underset{x\to 0^+}{\propto} e^0 = 1$, by Proposition 2.5.4 we have $\int_0^\infty x^{a-1}e^{-bx}dx = \infty$ for any $a \leq 0$ and $b \in \mathbb{R}$. Moreover, for $0 < \phi < \frac{k+1}{n}$, we have $n\phi - k - 2 < n\frac{(k+1)}{n} - k - 2 = -1$ and

$$\begin{aligned} d(t) &\propto \int_0^\infty \frac{\phi^{k+1}\pi(\phi)}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n t_i^{2\phi}\right\}\int_0^\infty \lambda^{n\phi-k-2}\exp\left\{-\lambda\sum_{i=1}^n t_i^2\right\}d\lambda d\phi \\ &\geq \int_0^{\frac{1+k}{n}} \frac{\phi^{k+1}\pi(\phi)}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n t_i^{2\phi}\right\}\int_0^\infty \lambda^{n\phi-k-2}\exp\left\{-\lambda\sum_{i=1}^n t_i^2\right\}d\lambda d\phi \\ &= \int_0^{\frac{1+k}{n}} \frac{\phi^{k+1}\pi(\phi)}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n t_i^{2\phi}\right\}\times\infty\, d\phi = \int_0^{\frac{1+k}{n}} \infty\, d\phi = \infty. \end{aligned}$$

- *Case* $k \leq -1$.

Let $v(\phi) = \dfrac{\phi^{k+1}\pi(\phi)\Gamma(n\phi-k-1)}{\Gamma(\phi)^n}$. By equation (3.19),

$$\begin{aligned} d(t) &\propto \int_0^\infty \frac{\phi^{k+1}\pi(\phi)}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n t_i^{2\phi}\right\}\int_0^\infty \lambda^{n\phi-k-2}\exp\left\{-\lambda\sum_{i=1}^n t_i^2\right\}d\lambda d\phi \\ &= \int_0^\infty v(\phi)\frac{\left(\prod_{i=1}^n t_i^{2\phi}\right)}{\left(\sum_{i=1}^n t_i^2\right)^{n\phi-k-1}}d\phi \propto \int_0^\infty v(\phi)n^{-n\phi}\frac{\left(\sqrt[n]{\prod_{i=1}^n t_i^2}\right)^{n\phi}}{\left(\frac{1}{n}\sum_{i=1}^n t_i^2\right)^{n\phi}}d\phi \\ &= \int_0^\infty v(\phi)n^{-n\phi}e^{-nq(t)\phi}d\phi = \int_0^1 v(\phi)n^{-n\phi}e^{-nq(t)\phi}d\phi + \int_1^\infty v(\phi)n^{-n\phi}e^{-nq(t)\phi}d\phi \\ &= s_1 + s_2, \end{aligned}$$

where $q(t) = \log\left(\dfrac{\frac{1}{n}\sum_{i=1}^n t_i^2}{\sqrt[n]{\prod_{i=1}^n t_i^2}}\right) > 0$ by the inequality of the arithmetic and geometric means. Therefore $d(t) < \infty$ if and only if $s_1 < \infty$ and $s_2 < \infty$.

Using the Stirling approximation of gamma function (ABRAMOWITZ; STEGUN, 1972), we have

$$\Gamma(z) \underset{z\to\infty}{\propto} z^{z-\frac{1}{2}}e^{-z}.$$

It is also valid that $\Gamma(z+a) \underset{z\to\infty}{\propto} \Gamma(z)z^a$ for $a \geq 0$ and $z \in \mathbb{R}$ (ABRAMOWITZ; STEGUN, 1972). Therefore $\Gamma(n\phi-k-1) \underset{\phi\to\infty}{\propto} \Gamma(n\phi)(n\phi)^{-k-1}$ and

$$\begin{aligned}
v(\phi) = \frac{\phi^{k+1}\pi(\phi)\Gamma(n\phi-k-1)}{\Gamma(\phi)^n} &\underset{\phi\to\infty}{\propto} \frac{\phi^{k+1}\phi^{r_\infty}(n\phi)^{n\phi-\frac{1}{2}}e^{-n\phi}(n\phi)^{-k-1}}{\phi^{n\phi-\frac{n}{2}}e^{-n\phi}} \\
&\propto \frac{\phi^{r_\infty}(n\phi)^{n\phi-\frac{1}{2}}}{\phi^{n\phi-\frac{n}{2}}} \propto \phi^{r_\infty+\frac{n-1}{2}}n^{n\phi}.
\end{aligned}$$

Therefore, by Proposition 2.5.4

$$s_2 = \int_1^\infty v(\phi)n^{-n\phi}e^{-nq(t)\phi}d\phi \propto \int_1^\infty \phi^{r_\infty+\frac{n-1}{2}}e^{-nq(t)\phi}d\phi = \frac{\Gamma(r_\infty+\frac{n-1}{2}, nq(t))}{(nq(t))^{r_\infty+\frac{n-1}{2}}} < \infty.$$

Suppose $k = -1$. Given that $\Gamma(z) \underset{z\to 0^+}{\propto} \frac{1}{z}$ (ABRAMOWITZ; STEGUN, 1972) and $n^{-n\phi}e^{-nq\phi} \underset{\phi\to 0^+}{\propto} n^0e^0 = 1$, we have

$$s_1 = \int_0^1 \frac{\pi(\phi)\Gamma(n\phi)}{\Gamma(\phi)^n}n^{-n\phi}e^{-nq\phi}\,d\phi \propto \int_0^1 \frac{\phi^{r_0}\frac{1}{n\phi}}{\frac{1}{\phi^n}} \times 1\,d\phi \propto \int_0^1 \phi^{r_0+n-1}d\phi$$

which implies $s_1 < \infty$ if and only if $n > -r_0$.

On the other hand, if suppose $k < -1$, we have $-k-1 > 0$, which implies $\Gamma(n\phi-k-1) \underset{\phi\to 0^+}{\propto} \Gamma(0-k-1) \propto 1$. Then

$$s_1 = \int_0^1 \frac{\phi^{k+1}\pi(\phi)\Gamma(n\phi-k-1)}{\Gamma(\phi)^n}n^{-n\phi}e^{-nq\phi}\,d\phi \propto \int_0^1 \frac{\phi^{k+1}\phi^{r_0}}{\frac{1}{\phi^n}} \times 1\,d\phi = \int_0^1 \phi^{r_0+n+k+1}d\phi$$

which implies $s_1 < \infty$ if and only if $n > -r_0-k-2$ and the proof is completed.

□

3.4.2 Jeffreys's rule

As the parameters of the NK distribution are included in the interval $(0,\infty)$, the prior distribution using Jeffreys's rule is given by

$$\pi_1(\mu,\Omega) \propto \frac{1}{\mu\Omega}. \tag{3.20}$$

The joint posterior distribution for μ and Ω, produced by the Jeffreys's rule, is proportional to the product of the likelihood function (3.9) and the prior (3.20) resulting in

$$p_1(\mu,\Omega|t) \propto \frac{\mu^{n\mu-1}}{\Omega^{n\mu+1}\Gamma(\mu)^n}\left\{\prod_{i=1}^{n} t_i^{2\mu-1}\right\}\exp\left(-\frac{\mu}{\Omega}\sum_{i=1}^{n} t_i^2\right). \tag{3.21}$$

Theorem 3.4.2. The posterior density (PD) (3.21) is proper if and only if $n \geq 1$.

Proof. Here we have $\pi(\Omega)=\Omega^{-1}$ and $\pi(\mu)=\mu^{-1}$, then $k=-1$ and $r_0=r_\infty=-1$. Therefore, the result follows directly from Theorem 3.4.1. □

The marginal posterior distribution for μ is given by

$$p_1(\mu|t) \propto \frac{\mu^{n\mu-1}}{\Gamma(\mu)^n}\int_0^\infty \Omega^{-n\mu-1}\prod_{i=1}^{n} t_i^{2\mu}\exp\left(-\frac{\mu}{\Omega}\sum_{i=1}^{n} t_i^2\right)d\Omega \propto \frac{\Gamma(n\mu)}{\mu\Gamma(\mu)^n}\frac{\left(\prod_{i=1}^{n} t_i^2\right)^{\mu}}{\left(\sum_{i=1}^{n} t_i^2\right)^{n\mu}}. \tag{3.22}$$

The conditional posterior distribution for Ω is

$$p_1(\Omega|\mu,t) \sim \mathrm{IG}\left(n\mu,\mu\sum_{i=1}^{n} t_i^2\right), \tag{3.23}$$

where $\mathrm{IG}(\cdot)$ is an inverse gamma distribution with PDF given by $f(x,a,b)=b^a x^{-a-1}\exp(-bx^{-1})/\Gamma(a)$.

3.4.3 Jeffreys prior

For the NK distribution, calculating the square root of the determinant of $I(\mu,\Omega)$ given in (3.15), we have

$$\pi_2(\mu,\Omega) \propto \frac{\sqrt{\mu\psi'(\mu)-1}}{\Omega}. \tag{3.24}$$

The joint posterior distribution for μ and Ω, produced by the Jeffreys prior, is proportional to the product of the likelihood function (3.9) and the prior (3.24) resulting in,

$$p_2(\mu,\Omega|t) \propto \frac{\mu^{n\mu}\sqrt{\mu\psi'(\mu)-1}}{\Omega^{n\mu+1}\Gamma(\mu)^n}\left\{\prod_{i=1}^{n} t_i^{2\mu-1}\right\}\exp\left(-\frac{\mu}{\Omega}\sum_{i=1}^{n} t_i^2\right). \tag{3.25}$$

Theorem 3.4.3. The PD (3.25) is proper if and only if $n \geq 0$.

Proof. Here we have $\pi(\Omega)=\Omega^{-1}$. Given that $\lim_{z\to 0^+}\frac{\psi'(z)}{z^{-2}}=1$ (ABRAMOWITZ; STEGUN, 1972), it follows that

$$\lim_{z\to 0^+}\frac{\sqrt{z\psi'(z)-1}}{z^{-\frac{1}{2}}}=\lim_{z\to 0^+}\sqrt{\frac{\psi'(z)}{z^{-2}}-z}=1,$$

which implies

$$\sqrt{z\psi'(z)-1} \underset{z\to 0^+}{\propto} z^{-\frac{1}{2}}. \tag{3.26}$$

Moreover, following Abramowitz and Stegun (1972), we obtain

$$\psi'(z) = \frac{1}{z} + \frac{1}{2z^2} + o\left(\frac{1}{z^3}\right).$$

Then

$$\frac{z\psi'(z)-1}{z^{-1}} = \frac{1}{2} + o\left(\frac{1}{z}\right) \Rightarrow \lim_{z\to\infty} \frac{\sqrt{z\psi'(z)-1}}{z^{-\frac{1}{2}}} = \frac{1}{\sqrt{2}},$$

which implies

$$\sqrt{z\psi'(z)-1} \underset{z\to\infty}{\propto} z^{-\frac{1}{2}}. \tag{3.27}$$

Therefore $k=-1$ and $r_0 = r_\infty = -\frac{1}{2}$, which, by Theorem 3.4.1 implies that the posterior is proper for $n \geq 0$. □

The marginal posterior distribution for μ is given by

$$p_2(\mu|t) \propto \frac{\sqrt{\mu\psi'(\mu)-1}\,\Gamma(n\mu)}{\Gamma(\mu)^n} \frac{\left(\prod_{i=1}^n t_i^2\right)^{\mu}}{\left(\sum_{i=1}^n t_i^2\right)^{n\mu}}. \tag{3.28}$$

The conditional posterior distribution for Ω is given by

$$p_2(\Omega|\mu,t) \sim \mathrm{IG}\left(n\mu, \mu\sum_{i=1}^n t_i^2\right). \tag{3.29}$$

3.4.4 Reference prior

As the Fisher information $I(\mu,\Omega)$ has a special form, considering Proposition 2.3.2 the overall reference prior for the NK distribution is given by

$$\pi_3(\mu,\Omega) \propto \frac{1}{\Omega}\sqrt{\frac{\mu\psi'(\mu)-1}{\mu}}. \tag{3.30}$$

The joint posterior distribution for μ and Ω, produced by the overall reference prior is proportional to the product of the likelihood function (3.9) and the prior distribution (3.30), resulting in

$$p_3(\mu,\Omega|t) \propto \frac{\sqrt{\mu\psi'(\mu)-1}}{\mu^{\frac{1}{2}}\Omega\Gamma(\mu)^n}\left(\frac{\mu}{\Omega}\right)^{n\mu}\left\{\prod_{i=1}^n t_i^{2\mu-1}\right\}\exp\left(-\frac{\mu}{\Omega}\sum_{i=1}^n t_i^2\right). \tag{3.31}$$

Theorem 3.4.4. The PD (3.31) is proper if and only if $n \geq 1$.

Proof. We have proved that $\sqrt{\mu\psi'(\mu)-1} \underset{z\to 0}{\propto} \mu^{-\frac{1}{2}}$ and that $\sqrt{\mu\psi'(\mu)-1} \underset{z\to\infty}{\propto} \mu^{-\frac{1}{2}}$. From that, it follows that $\sqrt{\frac{\mu\psi'(\mu)-1}{\mu}} \underset{z\to 0}{\propto} \mu^{-1}$ and $\sqrt{\frac{\mu\psi'(\mu)-1}{\mu}} \underset{z\to\infty}{\propto} \mu^{-1}$. Therefore $k=-1$ and $r_0 = r_\infty = -1$, therefore the result follows directly from Theorem 3.4.1. □

The marginal posterior distribution for μ is given by

$$p_3(\mu|t) \propto \sqrt{\frac{\mu\psi'(\mu)-1}{\mu}}\frac{\Gamma(n\mu)}{\Gamma(\mu)^n}\frac{\left(\prod_{i=1}^n t_i\right)^{2\mu-1}}{\left(\sum_{i=1}^n t_i^2\right)^{n\mu}}. \tag{3.32}$$

The conditional posterior distribution for Ω is given by

$$p_3(\Omega|\mu,t) \sim \mathrm{IG}\left(n\mu, \mu\sum_{i=1}^n t_i^2\right). \tag{3.33}$$

3.4.5 Maximal Data Information prior

For the NK distribution, $H(\mu,\Omega)$ can be written as

$$H(\mu,\Omega) = \log\left(\frac{2}{\Gamma(\mu)}\right) + \mu\psi(\mu) + \log\left(\frac{\mu^{\frac{1}{2}}}{\Omega}\right) - \frac{\psi(\mu)}{2} - \mu.$$

Therefore, the MDI prior (2.10) for the NK distribution (3.1) is given by

$$\pi_Z(\mu,\Omega) \propto \frac{1}{\Gamma(\mu)}\left(\frac{\mu}{\Omega}\right)^{\frac{1}{2}}\exp\left\{\psi(\mu)\left(\mu-\frac{1}{2}\right)-\mu\right\}. \tag{3.34}$$

The joint posterior distribution for μ and Ω, produced by the MDI prior, is proportional to the product of the likelihood function (3.9) and the prior distribution (3.34), resulting in

$$p_Z(\mu,\Omega|t) \propto \frac{1}{\Gamma(\mu)^{n+1}}\left(\frac{\mu}{\Omega}\right)^{n\mu+\frac{1}{2}}\prod_{i=1}^n t_i^{2\mu}\exp\left(-\frac{\mu}{\Omega}\sum_{i=1}^n t_i^2+\psi(\mu)\left(\mu-\frac{1}{2}\right)-\mu\right). \tag{3.35}$$

Theorem 3.4.5. The joint PD (3.35) is improper for any $n \in \mathbb{N}$.

Proof. Since $\pi(\Omega) = \Omega^{-\frac{1}{2}}$, we have $k = -\frac{1}{2}$ and the result follows from Theorem 3.4.1. □

3.4.6 Matching priors

Proposition 3.4.6. The overall reference prior (3.30) is a matching prior when μ is the parameter of interest and Ω is the orthogonal nuisance parameter and also when Ω is the parameter of interest and μ is the orthogonal nuisance parameter.

Proof. Note that $I_{\Omega,\mu}(\Omega,\mu) = 0$ for all (Ω,μ). Since $I_{\Omega,\Omega}(\Omega,\mu) = \dfrac{\mu}{\Omega^2}$, if $g(\mu) = \frac{\sqrt{\mu\psi'(\mu)-1}}{\mu}$ then the overall reference prior (3.30) can be written in the form (2.15).

On the other hand, $I_{\mu,\mu}(\mu,\Omega) = \frac{(\mu\psi'(\mu)-1)}{\mu}$, if $g(\Omega) = \frac{1}{\Omega}$ then the overall reference prior (3.30) can be written in the form (2.15) and the proof is complete. □

Proposition 3.4.7. The Jeffreys prior (3.24) is a matching prior when Ω is the parameter of interest and μ is the orthogonal nuisance parameter.

Proof. Firstly, $I_{\Omega,\mu}(\Omega,\mu) = 0$ for all (Ω,μ). Since $I_{\Omega,\Omega}(\Omega,\mu) = \frac{\mu}{\Omega^2}$, if $g(\mu) = \sqrt{\frac{\mu\psi'(\mu)-1}{\mu}}$ then the Jeffreys prior (3.24) can be written in the form (2.15). $\square$

Proposition 3.4.8. The Jeffreys rule prior (3.20) is a matching prior when Ω is the parameter of interest and μ is the orthogonal nuisance parameter.

Proof. Firstly, $I_{\Omega,\mu}(\Omega,\mu) = 0$ for all (Ω,μ). Since $I_{\Omega,\Omega}(\Omega,\mu) = \frac{\mu}{\Omega^2}$, if $g(\mu) = \mu^{-\frac{3}{2}}$ then the Jeffreys-rule prior (3.20) can be written in the form (2.15). $\square$

These results showed that for both Ω and μ, only the overall reference prior is a matching prior.

3.4.7 Numerical integration

Since the marginal distributions of μ do not belong to any known parametric family, we have to resort to numerical integration. To obtain the posterior mode of μ and its respective credibility intervals, the following procedure was adopted:

1. Find the value that maximizes the marginal distributions of μ given in (3.22), (3.28) or (3.32).

2. Compute the normalizing constant
$$d_i(t) = \int_0^\infty p_i(\mu|t)d\mu \text{ for } i = 1,2,3.$$

3. For $i = 1,2,3$ find the values of credibility intervals
$$\frac{1}{d_i(t)}\int_0^{\mu_{L_i}} p_i(\mu|t)d\mu = \frac{\alpha}{2} \text{ and } \frac{1}{d_i(t)}\int_{\mu_{S_i}}^\infty p_i(\mu|t)d\mu = \frac{\alpha}{2},$$
where μ_{L_i} is the $\alpha/2$ posterior quantile for μ and μ_{S_i} is the $1-\alpha/2$ posterior quantile. Then (μ_{L_i}, μ_{S_i}) is the credibility interval for μ.

Note that the Metropolis-Hastings algorithm (see Gamerman and Lopes (2006) for a detailed discussion) could be considered in order to generate samples from marginal posteriors. However, direct integration is faster and avoids the need for initial values, as well as convergence for the Markov chain Monte Carlo methods. Considering the conditional posterior distribution (3.29), two Bayes estimators can be obtained in closed-form using the posterior mean and the posterior mode. These estimators are given by

$$\hat{\Omega}_{Mean} = \frac{\mu\sum_{i=1}^n t_i^2}{n\mu - 1} \quad \text{and} \quad \hat{\Omega}_{Mode} = \frac{\mu\sum_{i=1}^n t_i^2}{n\mu + 1}. \tag{3.36}$$

On the other hand, for sample small sizes, these estimators are biased specially when μ is small. Note that

$$\frac{\mu\sum_{i=1}^{n} t_i^2}{n\mu+1} \leq \frac{\sum_{i=1}^{n} t_i^2}{n} \leq \frac{\mu\sum_{i=1}^{n} t_i^2}{n\mu-1}. \tag{3.37}$$

This implies that for $n \to \infty$, $\hat{\Omega}_{Mean} \approx \hat{\Omega}_{Mode} \approx \sum_{i=1}^{n} t_i^2/n$, i.e., the posterior mean and mode of Ω are asymptotically unbiased. Therefore, as a Bayes estimator, we choose the unbiased estimator $\hat{\Omega}_{Bayes} = \sum_{i=1}^{n} t_i^2/n$.

It is noteworthy that only $\hat{\Omega}_{Bayes}$ does not depend on μ. However, the credibility interval must be evaluated considering the quantile function of the IG$\left(n\mu, \mu\sum_{i=1}^{n} t_i^2\right)$ to satisfy the matching prior properties.

3.5 Simulation Study

In this section, a simulation study via Monte Carlo methods is used to compare the influence of different non-informative priors in the posterior distributions to find the most efficient estimation method by computing the mean relative errors (MRE) and the mean square errors (MSE), given by

$$\text{MRE}_{\mu} = \frac{1}{N}\sum_{i=1}^{N}\frac{\hat{\mu}_i}{\mu} \quad \text{and} \quad \text{MSE}_{\mu} = \sum_{i=1}^{N}\frac{(\hat{\mu}_i-\mu)^2}{N}, \tag{3.38}$$

where $N = 10,000,000$ is the number of estimates obtained through the posterior mode of μ. The MRE and the MSE of the Ω were not presented as they are the same for all methods and also Ω_{Bayes} is an unbiased estimator. The 95% coverage probability ($CP_{95\%}$) of the credibility intervals (CI) and the asymptotic confidence intervals for Ω and μ are evaluated. Considering this approach, the best estimators will show MRE closer to one and MSE closer to zero. In addition, for a large number of experiments considering a 95% confidence level, the frequencies of intervals that covered the true values of θ should be closer to 95%.

The results were computed using the software R (R Core Development Team). The seed used to generate the pseudo-random samples from the NK distribution was 2017. Considering $n = (20, \ldots, 120)$, the results were presented only for $\theta = ((4,2),(0.5,5))$ due to the lack of space. However, the obtained results are similar for other choices of μ and Ω. For each simulated sample, we computed the posterior modes for μ, Ω and the credibility (confidence) intervals for both parameters. Tables 1 and 2 present the MREs, MSEs and $CP_{95\%}$ from different estimators of μ and the $CP_{95\%}$ from Ω.

From these results, we observe that

1. The MRE (MSE) for all estimators of the parameters tend to one (zero) for large n, i.e., the estimators are asymptotically unbiased for the parameters.

Table 1 – The MRE(MSE) from the estimates of μ considering different values of n with $N = 10,000,000$ simulated samples using the estimation methods: 1 - MM, 2 - MLE, 3- Jeffreys's rule, 4 - Jeffreys prior, 5 - Overall reference prior.

θ	n	MM	MLE	Jeffreys's rule	Jeffreys Prior	Reference Prior
	20	1.16733(2.74804)	1.13136(2.12334)	0.97383(1.34895)	1.02236(1.50408)	0.96993(1.34994)
	30	1.12631(1.91607)	1.10068(1.50871)	0.99855(1.09233)	1.03007(1.18738)	0.99606(1.09138)
	40	1.09602(1.35709)	1.07609(1.06296)	1.00127(0.83116)	1.02436(0.88519)	0.99944(0.83050)
	50	1.07651(1.02625)	1.06020(0.79743)	1.00126(0.65397)	1.01945(0.68752)	0.99982(0.65355)
$\mu = 4$,	60	1.06368(0.82170)	1.04989(0.63405)	1.00127(0.53671)	1.01628(0.55951)	1.00008(0.53641)
	70	1.05435(0.68307)	1.04236(0.52448)	1.00101(0.45449)	1.01377(0.47088)	0.99999(0.45428)
$\Omega = 2$	80	1.04747(0.58415)	1.03693(0.44672)	1.00094(0.39386)	1.01205(0.40624)	1.00005(0.39370)
	90	1.04212(0.50990)	1.03267(0.38857)	1.00082(0.34733)	1.01065(0.35698)	1.00003(0.34720)
	100	1.03780(0.45237)	1.02925(0.34414)	1.00068(0.31107)	1.00950(0.31881)	0.99998(0.31097)
	110	1.03448(0.40707)	1.02663(0.30870)	1.00073(0.28148)	1.00872(0.28786)	1.00009(0.28139)
	120	1.03152(0.36939)	1.02431(0.27967)	1.00062(0.25698)	1.00793(0.26229)	1.00004(0.25691)
	20	1.36875(0.11215)	1.12627(0.03256)	1.00804(0.02185)	1.03982(0.02395)	1.00096(0.02144)
	30	1.25173(0.06330)	1.07985(0.01740)	1.00498(0.01330)	1.02517(0.01410)	1.00045(0.01314)
	40	1.19214(0.04367)	1.05813(0.01169)	1.00336(0.00955)	1.01815(0.00996)	1.00002(0.00946)
	50	1.15631(0.03331)	1.04588(0.00878)	1.00269(0.00746)	1.01437(0.00772)	1.00005(0.00741)
$\mu = 0.5$,	60	1.13211(0.02690)	1.03800(0.00701)	1.00234(0.00612)	1.01199(0.00629)	1.00015(0.00608)
	70	1.11433(0.02255)	1.03226(0.00583)	1.00190(0.00518)	1.01012(0.00531)	1.00004(0.00516)
$\Omega = 5$	80	1.10097(0.01944)	1.02803(0.00498)	1.00160(0.00450)	1.00876(0.00459)	0.99998(0.00448)
	90	1.09045(0.01711)	1.02486(0.00436)	1.00146(0.00398)	1.00780(0.00405)	1.00002(0.00396)
	100	1.08193(0.01525)	1.02233(0.00386)	1.00133(0.00356)	1.00702(0.00362)	1.00004(0.00355)
	110	1.07493(0.01378)	1.02020(0.00347)	1.00116(0.00322)	1.00632(0.00327)	0.99999(0.00321)
	120	1.06895(0.01256)	1.01853(0.00316)	1.00111(0.00295)	1.00584(0.00299)	1.00004(0.00294)

2. For Ω, the same MREs and MSEs were obtained from both approaches. However, considering the Bayesian approach, we obtain accurate coverage probability through the CIs, specially for $n < 100$.

3. In the case of μ, the posterior mode using the Jeffreys's rule and the overall reference prior returns nearly an unbiased estimator for μ, which indicates a better performance than the classical approaches. The better performance of this approach is also confirmed through the coverage probability obtained from the CIs.

Combining all the numerical results, both posterior distributions with the Jeffreys's rule and the overall reference prior produced better results. However, the overall reference posterior distribution has better theoretical properties such as invariance property under one-to-one transformations of the parameters, consistency under marginalization, consistent sampling properties and the fact that it is a matching prior for both Ω and μ. Therefore, we conclude that the overall reference posterior distribution should be used to make an inference on the unknown parameters of the NK distribution.

3.6 Applications in Reliability

In this section, our proposed methodology is adopted in two data sets. The NK distribution is compared with other usual two parameter lifetime distributions. The following lifetime

Table 2 – The $CP_{95\%}$ from the estimates of μ and Ω considering different values of n with $N = 10,000,000$ simulated samples using the estimation methods: 1 - MM, 2 - MLE, 3- Jeffreys's rule, 4 - Jeffreys prior, 5 - Overall reference prior.

θ	n	MLE		Jeffreys's rule		Jeffreys Prior		Reference Prior	
		μ	θ	μ	θ	μ	θ	μ	θ
$\mu = 4, \Omega = 2$	20	0.9710	0.9280	0.9730	0.9510	0.9680	0.9460	0.9730	0.9520
	30	0.9610	0.9340	0.9520	0.9490	0.9490	0.9450	0.9520	0.9490
	40	0.9560	0.9380	0.9510	0.9490	0.9480	0.9460	0.9510	0.9490
	50	0.9550	0.9400	0.9510	0.9490	0.9490	0.9470	0.9510	0.9490
	60	0.9540	0.9420	0.9500	0.9490	0.9480	0.9480	0.9500	0.9490
	70	0.9530	0.9430	0.9490	0.9490	0.9470	0.9480	0.9490	0.9500
	80	0.9530	0.9440	0.9490	0.9490	0.9470	0.9480	0.9490	0.9490
	90	0.9530	0.9440	0.9480	0.9490	0.9470	0.9480	0.9480	0.9500
	100	0.9520	0.9450	0.9480	0.9500	0.9470	0.9490	0.9480	0.9500
	110	0.9520	0.9450	0.9490	0.9500	0.9480	0.9490	0.9490	0.9500
	120	0.9520	0.9460	0.9500	0.9500	0.9490	0.9490	0.9500	0.9500
$\mu = 0.5, \Omega = 5$	20	0.9610	0.8930	0.9490	0.9470	0.9450	0.9440	0.9500	0.9480
	30	0.9570	0.9100	0.9500	0.9480	0.9470	0.9460	0.9500	0.9490
	40	0.9550	0.9200	0.9500	0.9490	0.9480	0.9470	0.9500	0.9490
	50	0.9540	0.9250	0.9500	0.9490	0.9480	0.9480	0.9500	0.9490
	60	0.9540	0.9290	0.9500	0.9490	0.9490	0.9480	0.9500	0.9490
	70	0.9530	0.9320	0.9500	0.9490	0.9490	0.9480	0.9500	0.9500
	80	0.9530	0.9340	0.9500	0.9490	0.9490	0.9490	0.9500	0.9500
	90	0.9520	0.9360	0.9500	0.9490	0.9500	0.9490	0.9500	0.9500
	100	0.9520	0.9370	0.9500	0.9490	0.9500	0.9490	0.9510	0.9500
	110	0.9520	0.9390	0.9500	0.9500	0.9500	0.9490	0.9500	0.9500
	120	0.9520	0.9390	0.9500	0.9500	0.9500	0.9490	0.9500	0.9500

distributions were considered. Let $\beta > 0$ and $\alpha > 0$, the Weibull distribution with PDF given by

$$f_W(t|\beta,\alpha) = \frac{\alpha}{\beta}\left(\frac{\alpha}{\beta}\right)^{\alpha-1} e^{-\left(\frac{\alpha}{\beta}\right)^{\alpha}}.$$

The lognormal distribution with PDF given by

$$f_L(t|\beta,\alpha) = \frac{1}{\alpha\sqrt{2\pi}} e^{-\frac{(\ln x-\beta)^2}{2\alpha^2}}.$$

The gamma distribution with PDF given by

$$f_G(t|\beta,\alpha) = \frac{\beta^{\alpha}}{\Gamma(\beta)} t^{\alpha-1} e^{-\beta t}.$$

The model selection was carried out by considering the negative log-likelihood value. The best model is the one which provides the minimum value.

3.6.1 Breaking stress of carbon fibers

The following data set (see Table 3) was extracted from Pal, Ali and Woo (2006), which is related to the lifetime of one hundred observations on breaking stress of carbon fibers (in Gba). These carbon fibers are used to construct fibrous composite materials. Table 4 presents the MLEs for the parameters of the Weibull, gamma and lognormal distribution. The posterior summaries

Table 3 – Data set related to the lifetime of one hundred observations on breaking stress of carbon fibers

3.70	2.74	2.73	2.50	3.60	3.11	3.27	2.87
1.47	3.11	4.42	2.41	3.19	3.22	1.69	3.28
3.09	1.87	3.15	4.90	3.75	2.43	2.95	2.97
3.39	2.96	2.53	2.67	2.93	3.22	3.39	2.81
4.20	3.33	2.55	3.31	3.31	2.85	2.56	3.56
3.15	2.35	2.55	2.59	2.38	2.81	2.77	2.17
2.83	1.92	1.41	3.68	2.97	1.36	0.98	2.76
4.91	3.68	1.84	1.59	3.19	1.57	0.81	5.56
1.73	1.59	2.00	1.22	1.12	1.71	2.17	1.17
5.08	2.48	1.18	3.51	2.17	1.69	1.25	4.38
1.84	0.39	3.68	2.48	0.85	1.61	2.79	4.70
2.03	1.80	1.57	1.08	2.03	1.61	2.12	1.89
2.88	2.82	2.05	3.65				

Table 4 – The MLEs (Standard Errors) obtained for different distributions considering the lifetime of one hundred observations on breaking stress of carbon fibers.

$\hat{\theta}$	Weibull	Gamma	Lognormal
$\hat{\alpha}$	2.7928(0.2141)	5.9526(0.8193)	0.4439(0.0314)
$\hat{\beta}$	2.9436(0.1111)	2.2708(0.3261)	0.8773(0.0443)

Table 5 – Posterior mode and 95% credibility intervals for μ and Ω obtained from the data set related to the lifetime of one hundred observations on breaking stress of carbon fibers.

$\hat{\theta}$	Mode	$\mathrm{CI}_{95\%}(\theta)$
$\hat{\mu}$	1.7104	(1.3269; 2.2176)
$\hat{\Omega}$	7.8894	(6.8300; 9.2174)

Table 6 – Results of AIC, AICc and BIC criteria for different probability distributions considering the data set related to the life time of one hundred observations on breaking stress of carbon fibers.

Criteria	Nakagami	Weibull	Gamma	Lognormal
AIC	**286.92**	287.06	290.47	300.84
AICc	**283.04**	283.18	286.59	296.96
HQIC	**289.02**	289.17	292.58	302.95
BIC	**292.13**	292.27	295.68	306.05

are given in Table 5. Table 6 presents the AIC, AICc and BIC values for different probability distributions considering the proposed data set.

We conclude that from the results of Table 6, the NK distribution has the best fit among the chosen models for describing the proposed data set. The quality of the fit can also be observed in Figure 5 by the reliability function fitted, considering different probability distributions and the empirical reliability function. Moreover, based on the TTT-plot, there is an indication that the hazard function has an increasing shape. This result is confirmed by the posterior estimates since $\hat{\mu} = 1.7148$ (increasing shape when $\mu \geq 0.5$). Therefore, considering our proposed methodology, the data related to the lifetime of the electrical devices can be described by the NK distribution.

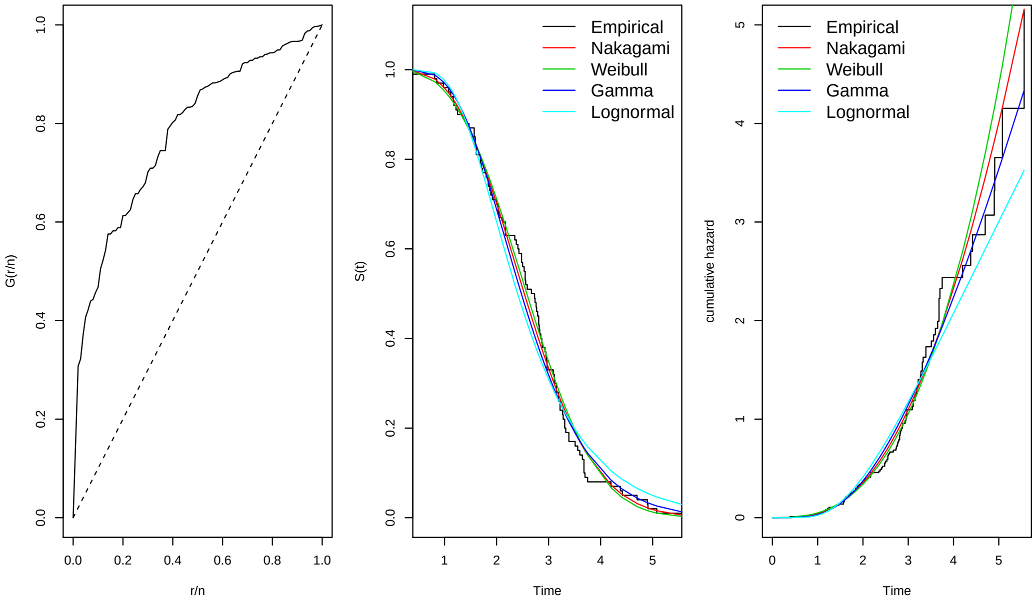

Figure 5 – TTT-plot, empirical reliability function and reliability function of different fitted probability distributions considering the data set related to the time of failure of 18 electronic devices and the cumulative hazard function.

3.6.2 Cycles up to the failure for electrical appliances

We reanalyzed the data extracted from (LAWLESS, 2011, p.112) which consists of a number of cycles divided by 1000 up to the failure for 60 electrical appliances in a life test (available in Table 7).

Table 7 – Data set related to the lifetime of 60 (in cycles) electrical devices.

0.014	0.034	0.059	0.061	0.069	0.080	0.123	0.142
0.165	0.210	0.381	0.464	0.479	0.556	0.574	0.839
0.917	0.969	0.991	1.064	1.088	1.091	1.174	1.270
1.275	1.355	1.397	1.477	1.578	1.649	1.702	1.893
1.932	2.001	2.161	2.292	2.326	2.337	2.628	2.785
2.811	2.886	2.993	3.122	3.248	3.715	3.790	3.857
3.912	4.100	4.106	4.116	4.315	4.510	4.580	5.267
5.299	5.583	6.065	9.701				

Based on the TTT-plot (see Figure 6), there is an indication that the hazard function is bathtub shaped. Therefore, one reviewer suggested extending the analysis by adding further competing models appropriate for bathtub shapes. The lifetime distributions considered with this property are the weighted Lindley distribution (GHITANY *et al.*, 2011a) with PDF given by

$$f_{WL}(t|\beta,\alpha) = \frac{\beta^{\alpha+1}}{(\beta+\alpha)\Gamma(\alpha)} t^{\alpha-1}(1+t)e^{-\beta t},$$

and the Chen distribution (CHEN, 2000) with PDF given by

$$f_C(t|\beta,\alpha) = \alpha\beta t^{\beta-1} e^{t^{\beta}} e^{\alpha(1-e^{t^{\beta}})},$$

where $\beta > 0$ and $\alpha > 0$.

Table 8 presents the MLEs for the parameters of the Weibull, gamma, lognormal, weighted Lindley and Chen distribution. Table 9 displays the posterior estimates and the 95% CIs for μ and Ω using the overall reference prior. Table 10 presents the AIC, AICc and BIC values for different probability distribution considering the proposed data set.

Table 8 – The MLEs obtained for different distributions considering the lifetime data set of 60 (in cycles) electrical devices.

$\hat{\theta}$	Weibull	Gamma	Lognormal	W. Lindley	Chen
$\hat{\alpha}$	1.0008(0.1066)	0.9307(0.4244)	0.1597(0.1858)	0.7332(0.1362)	0.2453(0.0443)
$\hat{\beta}$	2.1937(0.2961)	0.1486(0.08843)	1.4392(0.1313)	0.5874(0.0927)	0.5317(0.0392)

Table 9 – Posterior mode and 95% CIs for μ and Ω obtained from the data set related to 1000s cycles of failure for 60 electrical appliances.

$\hat{\theta}$	Mode	$CI_{95\%}(\theta)$
$\hat{\mu}$	0.35539	(0.2672 0.4768)
$\hat{\Omega}$	8.43430	(5.6862; 13.8043)

Table 10 – Results of AIC, AICc and BIC criteria for different probability distributions considering the data set related to 1000 cycles of failure for 60 electrical appliances.

Criteria	NK	Weibull	Gamma
AIC	**214.32**	218.22	218.02
AICc	**214.53**	218.43	218.23
BIC	**218.50**	222.41	222.21
Criteria	Lognormal	W. Lindley	Chen
AIC	**237.12**	215.54	214.84
AICc	**237.33**	215.75	215.05
BIC	**241.31**	219.73	219.03

From the results of Table 10, it can be concluded that among the chosen models, the NK distribution has the best fit for describing the number of 1000s cycles per failure for 60 electrical appliances.

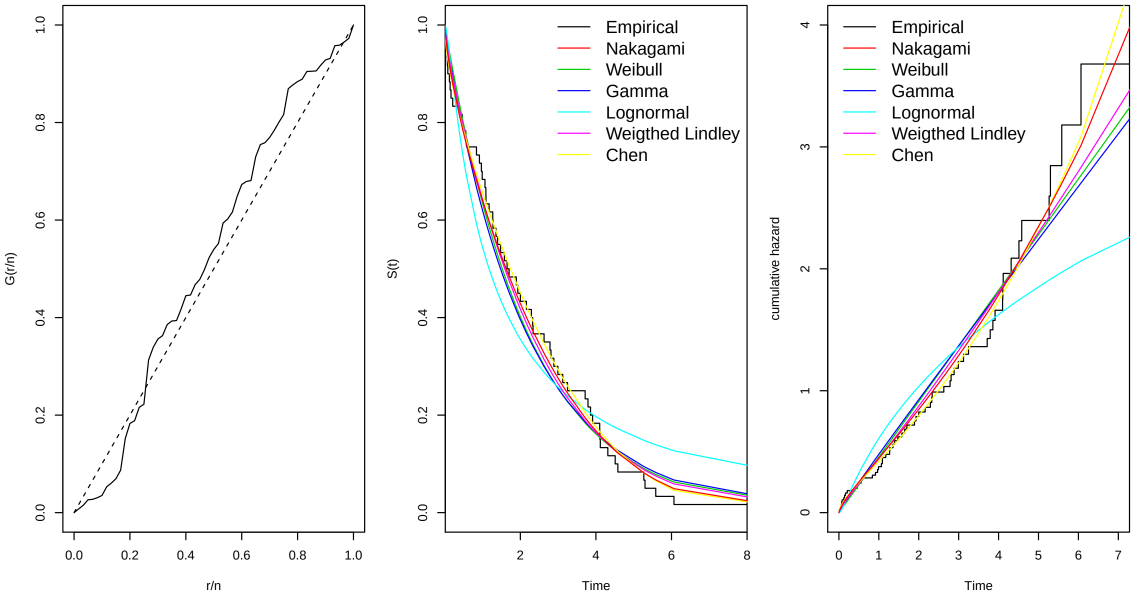

Figure 6 – TTT-plot, empirical reliability function and reliability function of different fitted probability distributions for the data set related to the number of 1000s cycles for failure for 60 electrical appliances and the cumulative hazard function.

In Figure 6, the goodness of fit can be observed by the reliability function fitted by the different probability distributions and the empirical reliability function. The bathtub shape in the hazard function is confirmed by the posterior estimates since $\hat{\mu} = 0.33365$ (bathtub shaped when $0 < \mu < 0.5$). Therefore, considering our proposed methodology, the data related to the lifetime of the electrical devices can be described by the NK distribution.

3.7 Discussion

We presented a theorem that provides sufficient and necessary conditions for a general class of posterior distribution to be proper. An interesting aspect of our findings is that it can be observed that the posterior is proper or improper considering the behavior of the proposed objective prior. The main theorem is applied in different objective priors for the NK distribution such as Jeffreys's rule, Jeffreys prior, the MDI prior and reference priors. The Jeffreys-rule prior and Jeffreys prior gave proper posterior distribution respectively for $n \geq 1$ and $n \geq 0$, whereas they are matching priors only for Ω. The MDI prior provided improper posterior for any sample sizes and should not be used in Bayesian analysis.

The overall reference prior yielded a proper posterior distribution if and only if $n \geq 1$. This prior is the one-at-a-time reference prior for any chosen parameter of interest and any ordering of the nuisance parameters. It is also the only prior that is a matching prior for both Ω and μ.

An extensive simulation study showed that the proposed overall reference posterior distribution returns more accurate results, as well as better theoretical properties such as the invariance property under one-to-one transformations of the parameters, consistency under marginalization and consistent sampling properties. Therefore, the overall reference posterior distribution should be used to make inference in the unknown parameters of the NK distribution. Finally, our proposed methodology is fully illustrated using two real lifetime data sets, demonstrating that the NK distribution can be used to describe lifetime data.

CHAPTER 4

GAMMA DISTRIBUTION

4.1 Introduction

The gamma distribution is one of the most well-known distributions used in statistical analysis. Such distribution arises naturally in many areas such as environmental analysis, reliability analysis, clinical trials, signal processing and other physical situations. Let X be a non-negative random variable with the gamma distribution given by

$$f(x|\phi,\mu) = \frac{\mu^{\phi}}{\Gamma(\phi)}x^{\phi-1}e^{-\mu x}, \tag{4.1}$$

where $\phi > 0$ and $\mu > 0$ are unknown shape and scale parameters, respectively, and $\Gamma(\phi) = \int_0^\infty e^{-x}x^{\phi-1}dx$ is the gamma function.

Commonly-used frequentist methods of inference for gamma distribution are standard in the statistical literature. Considering the Bayesian approach, where a prior distribution must be assigned, different objective priors for the gamma distribution have been discussed earlier by Miller (1980), Sun and Ye (1996) and Berger *et al.* (2015). Although these priors are constructed by formal rules, they are improper, i.e., do not correspond to proper probability distribution and could lead to improper posteriors, which is undesirable. Northrop and Attalides (2016) argued that "… there is no general theory providing simple conditions under which an improper prior yields a proper posterior for a particular model, so this must be investigated case-by-case". In this study, we overcome this problem by providing in a simple way necessary and sufficient conditions to check whether or not these objective priors lead proper posterior distributions. Even if the posterior distribution is proper the posterior moments for the parameters can be infinite. Further, we also provided sufficient conditions to verify if the posterior moments are finite. Therefore, one can easily check if the obtained posterior is proper or improper and also if its posterior moments are finite considering directly the behavior of the improper prior. Our proposed methodology is fully illustrated in more than ten objective priors such as independent uniform priors, Jeffreys' rule, Jeffreys' prior, MDI prior, reference priors and matching priors, to list a few. Finally, the effect of these priors in the posterior distribution are compared via numerical simulation.

The remainder of this chapter is organized as follows. Section 2 presents a theorem that provides necessary and sufficient conditions for the posterior distributions to be proper and also sufficient conditions to check if the posterior moments of the parameters are finite. Section 3 presents the applications of our main theorem in different objective priors. In Section 4, a simulation study is conducted in order to identify the most efficient estimation procedure. Finally, Section 5 summarizes the study.

4.2 Classical Inference

The MLEs for the gamma distribution are obtained by maximizing the likelihood function, given by

$$L(\mu,\phi;t)=\frac{\mu^{n\phi}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n}t_i^{\phi-1}\right\}\exp\left(-\mu\sum_{i=1}^{n}t_i\right). \tag{4.2}$$

The log-likelihood function is

$$\ell(\mu,\phi;t)=-n\log\Gamma(\phi)+n\phi\log(\mu)+(\phi-1)\sum_{i=1}^{n}\log(t_i)-\mu\sum_{i=1}^{n}t_i. \tag{4.3}$$

The estimates are obtained from maximizing the likelihood function. From the expressions $\frac{\partial}{\partial\phi}\ell(\mu,\phi;t)=0$, $\frac{\partial}{\partial\mu}\ell(\mu,\phi;t)=0$, the likelihood equations are given as

$$-n\psi(\phi)+n\log(\mu)+\sum_{i=1}^{n}\log(t_i)=0 \tag{4.4}$$

and

$$-\frac{n\phi}{\mu}+\sum_{i=1}^{n}t_i=0, \tag{4.5}$$

where $\psi(k)=\frac{\partial}{\partial k}\log\Gamma(k)=\frac{\Gamma'(k)}{\Gamma(k)}$ is the digamma function. After some algebraic manipulations the MLE for $\hat{\mu}$ is given by

$$\hat{\mu}=\frac{n\phi}{\sum_{i=1}^{n}t_i}. \tag{4.6}$$

Note that, substituting $\hat{\mu}_{MLE}$ in (4.4), the estimate for $\hat{\phi}_{MLE}$ can be obtained solving

$$\log(\phi)-\psi(\phi)=\log\left(\frac{1}{n}\sum_{i=1}^{n}t_i\right)-\frac{1}{n}\sum_{i=1}^{n}\log(t_i). \tag{4.7}$$

The MLE estimates are asymptotically normally distributed with a joint bivariate normal distribution given by

$$(\hat{\phi}_{MLE},\hat{\mu}_{MLE})\sim N_2[(\phi,\mu),I^{-1}(\phi,\mu))] \text{ for } n\to\infty,$$

where $I(\phi,\mu)$ is the Fisher information matrix

$$I(\phi,\mu) = n\begin{bmatrix} \psi'(\phi) & -\dfrac{1}{\mu} \\ -\dfrac{1}{\mu} & \dfrac{\phi}{\mu^2} \end{bmatrix}, \tag{4.8}$$

and $\psi'(k) = \frac{\partial}{\partial k}\psi(k)$ is the trigamma function.

4.3 Proper Posterior

The joint posterior distribution for $\theta = (\phi,\mu)$ is given by the product of the likelihood function and the prior distribution $\pi(\theta)$ divided by a normalizing constant $d(x)$, resulting in

$$p(\theta|x) = \frac{\pi(\theta)}{d(x)}\frac{\mu^{n\phi}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n} x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^{n} x_i\right\}, \tag{4.9}$$

where

$$d(x) = \int\limits_{\mathscr{A}} \frac{\pi(\theta)\mu^{n\phi}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n} x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^{n} x_i\right\}d\theta \tag{4.10}$$

and $\mathscr{A} = \{(0,\infty)\times(0,\infty)\}$ is the parameter space of θ. For any prior distribution in the form: $\pi(\theta) \propto \pi_1(\mu)\pi_2(\phi)$, our purpose is to find necessary and sufficient conditions for these class of posterior be proper, i.e., $d(x) < \infty$. The following propositions are useful.

Theorem 4.3.1. Let the behavior of $\pi(\mu)$ be given by $\pi(\mu) \propto \mu^c$, for $c \in \mathbb{R}$. Then

i) If $c < -1$, then the posterior distribution (4.10) is improper.

ii) If $c \geq -1$ and $\lim_{\phi\to 0^+}\pi(\phi)\phi^s = \infty$ $\forall s \in \mathbb{N}$ then the posterior distribution (4.10) is improper.

iii) If $c \geq -1$ and the behavior of $\pi(\phi)$ is given by

$$\pi(\phi) \underset{\phi\to 0^+}{\propto} \phi^{s_0} \quad \text{and} \quad \pi(\phi) \underset{\phi\to\infty}{\propto} \phi^{s_\infty},$$

where $s_0 \in \mathbb{R}$ and $s_\infty \in \mathbb{R}$, then the posterior distribution (4.10) is proper if and only if $n > -s_0$ in case $c = -1$, and is proper if and only if $n > -s_0 - 1$ in case $c > -1$.

Proof. Let

$$d(x) \propto \int\limits_{\mathscr{B}} \frac{\pi(\phi)\mu^{n\phi+c}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n} x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^{n} x_i\right\}d\Theta \tag{4.11}$$

Since $\pi(\phi)\mu^{n\phi+c}\Gamma(\phi)^{-n}\prod_{i=1}^{n}x_i^{\phi}\exp\left(-\mu\sum_{i=1}^{n}x_i\right)\geq 0$, by the Fubini-Tonelli Theorem we have

$$\begin{aligned} d(x) &\propto \int_{\mathscr{B}} \frac{\pi(\phi)\mu^{n\phi+c}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n}x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^{n}x_i\right\}d\Theta \\ &= \int_0^{\infty}\frac{\pi(\phi)}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n}x_i^{\phi}\right\}\int_0^{\infty}\mu^{n\phi+c}\exp\left\{-\mu\sum_{i=1}^{n}x_i\right\}d\mu d\phi\,. \end{aligned} \tag{4.12}$$

The rest of the proof is divided in three items which are given bellow:

Case i): Suppose $c<-1$. Notice that $\int_0^{\infty}x^{k-1}e^{-hx}dx=\infty$ for any $k\leq 0$ and $h\in\mathbb{R}$. Then, for $0<\phi<\frac{-c-1}{n}$ we have $n\phi+c<n\frac{(-c-1)}{n}+c=-1$, and it follows that

$$\begin{aligned} d(x) &\propto \int_0^{\infty}\frac{\pi(\phi)}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n}x_i^{\phi}\right\}\int_0^{\infty}\mu^{n\phi+c}\exp\left\{-\mu\sum_{i=1}^{n}x_i\right\}d\mu d\phi \\ &\geq \int_0^{\frac{-c-1}{n}}\frac{\pi(\phi)}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n}x_i^{\phi}\right\}\int_0^{\infty}\mu^{n\phi+c}\exp\left\{-\mu\sum_{i=1}^{n}x_i\right\}d\mu d\phi \\ &= \int_0^{\frac{-c-1}{n}}\frac{\pi(\phi)}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n}x_i^{\phi}\right\}\times\infty\, d\phi = \int_0^{\frac{-c-1}{n}}\infty\, d\phi=\infty \end{aligned}$$

and the case i) is proved.

Now suppose $c\geq -1$. Denoting

$$v(\phi)=\frac{\pi(\phi)\Gamma(n\phi+c+1)}{\Gamma(\phi)^n}\ \text{ and }\ q(x)=\log\left(\frac{\frac{1}{n}\sum_{i=1}^{n}x_i}{\sqrt[n]{\prod_{i=1}^{n}x_i}}\right)>0,$$

we have that $q(x)>0$ by the inequality of the arithmetic and geometric means, and

$$\begin{aligned} d(x) &= \int_0^{\infty}v(\phi)\frac{\left(\prod_{i=1}^{n}x_i\right)^{\phi}}{\left(\sum_{i=1}^{n}x_i\right)^{n\phi+c+1}}d\phi \propto \int_0^{\infty}v(\phi)\frac{1}{n^{n\phi}}\frac{\left(\sqrt[n]{\prod_{i=1}^{n}x_i}\right)^{n\phi}}{\left(\frac{1}{n}\sum_{i=1}^{n}x_i\right)^{n\phi}}d\phi=\int_0^{\infty}v(\phi)n^{-n\phi}e^{-nq(x)\phi}d\phi \\ &= \int_0^{1}v(\phi)n^{-n\phi}e^{-nq(x)\phi}d\phi+\int_1^{\infty}v(\phi)n^{-n\phi}e^{-nq(x)\phi}d\phi=d_0(x)+d_{\infty}(x), \end{aligned}$$

where $d_0(x)=\int_0^1 v(\phi)n^{-n\phi}e^{-nq(x)\phi}d\phi$ and $d_{\infty}(x)=\int_1^{\infty}v(\phi)n^{-n\phi}e^{-nq(x)\phi}d\phi$.

Then $d(x)<\infty$ if and only if $d_0(x)<\infty$ and $d_{\infty}(x)<\infty$. These results lead us to the two remaining cases.

Case ii): Suppose $c \geq -1$ and $\lim_{\phi\to 0^+} \pi(\phi)\phi^s = \infty\ \forall s \in \mathbb{N}$. From Abramowitz and Stegun (1972), we have $\Gamma(z) \underset{z\to 0^+}{\propto} \frac{1}{z}$. Then, if $c = -1$

$$\begin{aligned} d_0(x) &= \int_0^1 \frac{\pi(\phi)\Gamma(n\phi)}{\Gamma(\phi)^n} n^{-n\phi} e^{-nq(x)\phi}\, d\phi \propto \int_0^1 \frac{\pi(\phi)\frac{1}{n\phi}}{\frac{1}{\phi^n}} \times 1 \times 1\, d\phi \\ &\propto \int_0^1 \pi(\phi)\phi^{n-1} d\phi = \int_1^\infty \pi\left(u^{-1}\right) u^{-n-1} du = \infty, \end{aligned}$$

where the last equality comes from the fact that

$$\lim_{u\to\infty} \pi\left(u^{-1}\right) u^{-n-1} = \lim_{\phi\to 0^+} \pi(\phi)\phi^{n+1} = \infty.$$

Therefore, $d(x) = \infty$ if $c = -1$. On the other hand, if $c > -1$ then $n\phi + c + 1 > 0$ for $\phi > 0$, which implies $\Gamma(n\phi + c + 1) \underset{\phi\to 0^+}{\propto} 1$ and

$$\begin{aligned} d_0(x) &= \int_0^1 \frac{\pi(\phi)\Gamma(n\phi + c + 1)}{\Gamma(\phi)^n} n^{-n\phi} e^{-nq\phi}\, d\phi \propto \int_0^1 \frac{\pi(\phi)}{\frac{1}{\phi^n}} \times 1 \times 1\, d\phi \\ &= \int_0^1 \pi(\phi)\phi^n d\phi = \int_1^\infty \pi\left(u^{-1}\right) u^{-n-2} du = \infty. \end{aligned}$$

Therefore, $d(x) = \infty$ if $c > -1$ and the case ii) is proved.

Case iii): Suppose that $c \geq -1$ and the behavior of $\pi(\phi)$ is given by

$$\pi(\phi) \underset{\phi\to 0^+}{\propto} \phi^{s_0} \quad \text{and} \quad \pi(\phi) \underset{\phi\to\infty}{\propto} \phi^{s_\infty},$$

where $s_0 \in \mathbb{R}$ and $s_\infty \in \mathbb{R}$. Following Abramowitz and Stegun (1972), we obtain that $\Gamma(z) \underset{z\to\infty}{\propto} z^{z-\frac{1}{2}} e^{-z}$ and $\Gamma(z + a) \underset{z\to\infty}{\propto} \Gamma(z) z^a$ for $a \in \mathbb{R}^+$. Then $\Gamma(n\phi + c + 1) \underset{\phi\to\infty}{\propto} \Gamma(n\phi)(n\phi)^{c+1}$ and

$$\begin{aligned} v(\phi) &= \frac{\pi(\phi)\Gamma(n\phi + c + 1)}{\Gamma(\phi)^n} \underset{\phi\to\infty}{\propto} \frac{\phi^{s_\infty}(n\phi)^{n\phi - \frac{1}{2}} e^{-n\phi}(n\phi)^{c+1}}{\phi^{n\phi - \frac{n}{2}} e^{-n\phi}} \\ &\propto \frac{\phi^{s_\infty + c + 1}(n\phi)^{n\phi - \frac{1}{2}}}{\phi^{n\phi - \frac{n}{2}}} \propto \phi^{s_\infty + c + \frac{n+1}{2}} n^{n\phi}. \end{aligned}$$

Therefore

$$\begin{aligned} d_\infty(x) &= \int_1^\infty v(\phi) n^{-n\phi} e^{-nq(x)\phi} d\phi \propto \int_1^\infty \phi^{s_\infty + c + \frac{n+1}{2}} e^{-nq(x)\phi} d\phi \\ &= \frac{\Gamma(s_\infty + c + \frac{n+1}{2}, nq(x))}{(nq(x))^{s_\infty + c + \frac{n+1}{2}}} < \infty, \end{aligned}$$

i.e., $d_\infty(x) < \infty$ for all $s_\infty \in \mathbb{R}$. Therefore $d(x) < \infty \Leftrightarrow d_0(x) < \infty$.

Now, following the same from case *ii*), if $c = -1$ we have

$$d_0(x) = \int_0^1 \frac{\pi(\phi)\Gamma(n\phi)}{\Gamma(\phi)^n} n^{-n\phi} e^{-nq\phi}\, d\phi \propto \int_0^1 \frac{\phi^{s_0}\frac{1}{n\phi}}{\frac{1}{\phi^n}}\, d\phi \propto \int_0^1 \phi^{s_0+n-1} d\phi,$$

i.e., $d(x) < \infty$ if and only if $n > -s_0$ when $c = -1$. On the other hand, if $c > -1$

$$d_0(x) = \int_0^1 \frac{\pi(\phi)\Gamma(n\phi+c+1)}{\Gamma(\phi)^n} n^{-n\phi} e^{-nq\phi}\, d\phi \propto \int_0^1 \frac{\phi^{s_0}}{\frac{1}{\phi^n}}\, d\phi = \int_0^1 \phi^{s_0+n} d\phi,$$

i.e., $d(x) < \infty$ if and only if $n > -s_0 - 1$ when $c > -1$ and the proof is completed.

□

Theorem 4.3.2. Let $\pi(\phi,\mu) = \pi(\phi)\pi(\mu)$ and the behavior of $\pi(\mu)$, $\pi(\phi)$ be given by

$$\pi(\mu) \propto \mu^c, \quad \pi(\phi) \underset{\mu \to 0^+}{\propto} \phi^{s_0} \quad \text{and} \quad \pi(\phi) \underset{\phi\to\infty}{\propto} \phi^{s_\infty},$$

for $c \in \mathbb{R}$ where $s_0 \in \mathbb{R}$ and $s_\infty \in \mathbb{R}$. If the posterior of $\pi(\phi,\mu)$ is proper then the posterior mean of ϕ and μ are finite for this prior, as well as all moments.

Proof. Since the posterior is proper, by Theorem 4.3.1 we have that $c \geq -1$, $n > -s_0 - 1$ if $c > -1$ and $n > -s_0$ if $c > -1$.

Let $\pi^*(\phi,\mu) = \phi\pi(\phi,\mu)$. Then $\pi^*(\phi,\mu) = \pi^*(\phi)\pi^*(\mu)$, where $\pi^*(\phi) = \phi\pi(\phi)$ and $\pi^*(\mu) = \pi(\mu)$, and we have

$$\pi^*(\mu) \propto \mu^c, \quad \pi^*(\phi) \underset{\mu\to 0^+}{\propto} \phi^{s_0+1} \quad \text{and} \quad \pi^*(\phi) \underset{\phi\to\infty}{\propto} \phi^{s_\infty+1}.$$

Since $c \geq -1$, $n > -s_0 - 1 > -(s_0+1) - 1$ if $c > -1$ and $n > -s_0 > -(s_0+1)$ if $c > -1$, it follows from Theorem 4.3.1 that the posterior

$$\pi^*(\phi,\mu)\frac{\mu^{n\phi}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n x_i^\phi\right\}\exp\left\{-\mu\sum_{i=1}^n x_i\right\}$$

relative to the prior $\pi^*(\phi,\mu)$ is proper. Therefore

$$E[\phi|x] = \int_0^\infty\int_0^\infty \phi\pi(\phi,\mu)\frac{\mu^{n\phi}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n x_i^\phi\right\}\exp\left\{-\mu\sum_{i=1}^n x_i\right\} d\mu d\phi < \infty.$$

Analogously one can prove that

$$E[\mu|x] = \int_0^\infty\int_0^\infty \mu\pi(\phi,\mu)\frac{\mu^{n\phi}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n x_i^\phi\right\}\exp\left\{-\mu\sum_{i=1}^n x_i\right\} d\mu d\phi < \infty.$$

Therefore we have proved that if a prior $\pi(\phi,\mu)$ satisfying the assumptions of the theorem leads to a proper posterior, then the priors $\phi\pi(\phi,\mu)$ and $\mu\pi(\phi,\mu)$ also leads to proper posteriors, and it follows by induction that $\phi^r\mu^s\pi(\phi,\mu)$ also leads to proper posteriors for any r and s in $\mathbb{N}$, which concludes the proof. □

Proposition 4.3.3. Suppose $\pi_i(\phi,\mu)$, $i=1,\cdots,m$ lead to proper posteriors for $n\in\mathbb{N}$, and consider the constants $k_i\geq 0$, $i=1,\cdots,m$. Then

i) $\sum_{i=1}^m k_i\pi_i(\phi,\mu)$ leads to a proper posterior

ii) $\prod_{i=1}^m \pi_i(\phi,\mu)^{k_i}$ leads to a proper posterior if $\sum_{i=1}^m k_i=1$.

Proof. The item $i)$ is a direct of consequence of the linearity of the Lebesgue integral while $ii)$ is a direct consequence of the Holder's inequality. □

4.4 Objective priors

In this section, we applied the proposed theorems in different objective priors.

4.4.1 *Uniform Prior*

A simple noninformative prior can be obtained considering uniform priors contained in the interval $(0,\infty)$. This prior usually is not attractive due its lack of invariance to reparameterisation. The uniform prior is given by $\pi_1(\phi,\mu)\propto 1$. The joint posterior distribution for ϕ and μ, produced by the uniform prior is

$$\pi_1(\phi,\mu|x)\propto\frac{\mu^{n\phi}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^n x_i\right\}. \tag{4.13}$$

Theorem 4.4.1. The posterior distribution (4.13) is proper if and only if $n\geq 0$, in which case the posterior moments for ϕ and μ are finite.

Proof. Since $\pi(\mu)=\mu^0$ and $\pi(\phi)=\phi^0$, then $c=0$ and $s_0=s_\infty=0$. Therefore, the result follows directly from the Theorem 4.3.1 and by Theorem 4.3.2. □

The marginal posterior distribution for ϕ is

$$\pi_1(\phi|x)\propto\frac{1}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n x_i^{\phi}\right\}\int_0^{\infty}\mu^{n\phi}\exp\left\{-\mu\sum_{i=1}^n x_i\right\}d\mu\propto\frac{\phi\Gamma(n\phi)}{\Gamma(\phi)^n}\left(\frac{\sqrt[n]{\prod_{i=1}^n x_i}}{\sum_{i=1}^n x_i}\right)^{n\phi}.$$

The conditional posterior distribution for μ is given by

$$\pi_1(\mu|\phi,x)\sim\mathrm{Gamma}\left(n\phi+1,\sum_{i=1}^n x_i\right). \tag{4.14}$$

4.4.2 Jeffreys Rule

Since the parameters of the gamma distribution are contained in the interval $(0,\infty)$, the prior using the Jeffreys rule (MILLER, 1980) is

$$\pi_2\left(\phi,\mu\right) \propto \frac{1}{\phi\mu} \ . \tag{4.15}$$

The joint posterior distribution for ϕ and μ produced by the Jeffreys rule prior is given by

$$\pi_2(\phi,\mu|x) \propto \frac{\mu^{n\phi-1}}{\phi\Gamma(\phi)^n}\left\{\prod_{i=1}^{n} x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^{n} x_i\right\}. \tag{4.16}$$

Theorem 4.4.2. The posterior density (4.16) is proper if and only if $n \geq 2$, in which case the posterior moments for ϕ and μ are finite.

Proof. Since $\pi(\mu) = \mu^{-1}$ and $\pi(\phi) = \phi^{-1}$, then $c = -1$ and $s_0 = s_\infty = -1$. Therefore the result follows directly from the Theorem 4.3.1 and Theorem 4.3.2. □

The marginal posterior distribution for ϕ is given by

$$\pi_2(\phi|x) \propto \frac{\Gamma(n\phi)}{\phi\Gamma(\phi)^n}\left(\frac{\sqrt[n]{\prod_{i=1}^{n} x_i}}{\sum_{i=1}^{n} x_i}\right)^{n\phi} .$$

The conditional posterior distribution for μ is

$$\pi_2(\mu|\phi,x) \sim \text{Gamma}\left(n\phi, \sum_{i=1}^{n} x_i\right). \tag{4.17}$$

4.4.3 Jeffreys prior

For the gamma distribution, the Jeffreys prior (MILLER, 1980) is given by

$$\pi_3\left(\phi,\mu\right) \propto \frac{\sqrt{\phi\psi'(\phi)-1}}{\mu}. \tag{4.18}$$

The joint posterior distribution for ϕ and μ produced by the Jeffreys prior is

$$\pi_3(\phi,\mu|x) \propto \frac{\mu^{n\phi-1}\sqrt{\phi\psi'(\phi)-1}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n} x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^{n} x_i\right\}. \tag{4.19}$$

Theorem 4.4.3. The posterior density (4.19) is proper if and only if $n \geq 1$, in which case the posterior moments for ϕ and μ are finite.

Proof. Here, we have $\pi(\mu)=\mu^{-1}$. Following Abramowitz and Stegun (1972) it can easily prove that $\lim_{z\to 0^+}\dfrac{\psi'(z)}{z^{-2}}=1$, then

$$\lim_{\phi\to 0^+}\frac{\sqrt{\phi\psi'(\phi)-1}}{\phi^{-\frac{1}{2}}}=\lim_{\phi\to 0^+}\sqrt{\frac{\psi'(\phi)}{\phi^{-2}}-\phi}=1,$$

which implies

$$\sqrt{\phi\psi'(\phi)-1}\underset{\phi\to 0^+}{\propto}\phi^{-\frac{1}{2}}.$$

Moreover, following Abramowitz and Stegun (1972), we also obtain that $\psi'(z)=\dfrac{1}{z}+\dfrac{1}{2z^2}+o\left(\dfrac{1}{z^3}\right)$, then

$$\frac{\phi\psi'(\phi)-1}{\phi^{-1}}=\frac{1}{2}+o\left(\frac{1}{\phi}\right)\Rightarrow\lim_{\phi\to\infty}\frac{\sqrt{\phi\psi'(\phi)-1}}{\phi^{-\frac{1}{2}}}=\frac{1}{\sqrt{2}},$$

which implies

$$\sqrt{\phi\psi'(\phi)-1}\underset{\phi\to\infty}{\propto}\phi^{-\frac{1}{2}}.$$

Therefore, $c=-1$ and $s_0=s_\infty=-\frac{1}{2}$, using the Theorem 4.3.1, the posterior is proper if and only if $n\geq 1$ and the posterior moments are finite using Theorem 4.3.2. □

The conditional posterior distribution for μ is (4.17). The marginal posterior distribution for ϕ is given by

$$\pi_3(\phi|x)\propto\frac{\Gamma(n\phi)\sqrt{\phi\psi'(\phi)-1}}{\Gamma(\phi)^n}\left(\frac{\sqrt[n]{\prod_{i=1}^n x_i}}{\sum_{i=1}^n x_i}\right)^{n\phi}.$$

4.4.4 Miller prior

Miller (1980) discussed three objective priors for the parameters of the gamma distribution, where the first two were the Jeffreys Rule and the Jeffreys prior. However, the author chose a prior using the justification that such approach involves less computational subroutines. This prior is given by

$$\pi_4(\phi,\mu)\propto\frac{1}{\mu}\cdot \tag{4.20}$$

Note that much progress has been made in computational analysis and many of these computational limitations have been overcome specially after Gelfand and Smith (1990) successfully applied the Gibbs sampling in Bayesian Analysis.

The joint posterior distribution for ϕ and μ produced by the Miller's prior is

$$\pi_4(\phi,\mu|x)\propto\frac{\mu^{n\phi-1}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^n x_i\right\}. \tag{4.21}$$

Theorem 4.4.4. The posterior density (4.21) is proper if and only if $n \geq 1$, in which case the posterior moments for ϕ and μ are finite.

Proof. Since $\pi(\mu) = \mu^{-1}$ and $\pi(\phi) = \phi^0$, then $c = -1$ and $s_0 = s_\infty = 0$. Therefore, the result follows directly from the Theorem 4.3.1 and Theorem 4.3.2. □

The conditional posterior distribution for μ is (4.17). The marginal posterior distribution for ϕ is given by

$$\pi_4(\phi|x) \propto \frac{\Gamma(n\phi)}{\Gamma(\phi)^n}\left(\frac{\sqrt[n]{\prod_{i=1}^n x_i}}{\sum_{i=1}^n x_i}\right)^{n\phi}.$$

4.4.5 Reference prior

4.4.5.1 Reference prior when ϕ is the parameter of interest

From Proposition 2.3.1 the reference prior when ϕ is the parameter of interest and μ is the nuisance parameter is given by

$$\pi_5(\phi, \mu) \propto \frac{1}{\mu}\sqrt{\frac{\phi\psi'(\phi) - 1}{\phi}}. \tag{4.22}$$

The joint posterior distribution for ϕ and μ, produced by the reference prior (4.22) is given by

$$\pi_5(\phi, \mu|x) \propto \sqrt{\frac{\phi\psi'(\phi) - 1}{\phi}}\frac{\mu^{n\phi-1}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^n x_i\right\}. \tag{4.23}$$

Theorem 4.4.5. The posterior density (4.23) is proper if and only if $n \geq 2$, in which case the posterior moments for ϕ and μ are finite.

Proof. We proved in Theorem 4.4.3 that $\sqrt{\phi\psi'(\phi) - 1} \underset{\phi\to 0^+}{\propto} \phi^{-\frac{1}{2}}$ and $\sqrt{\phi\psi'(\phi) - 1} \underset{\phi\to\infty}{\propto} \phi^{-\frac{1}{2}}$. It follows that

$$\sqrt{\frac{\phi\psi'(\phi) - 1}{\phi}} \underset{\phi\to 0^+}{\propto} \phi^{-1} \quad \text{and} \quad \sqrt{\frac{\phi\psi'(\phi) - 1}{\phi}} \underset{\phi\to\infty}{\propto} \phi^{-1}.$$

Then $c = -1$ and $s_0 = s_\infty = -1$, therefore the result follows directly from the Theorem 4.3.1 and Theorem 4.3.2. □

The conditional posterior distribution for μ is (4.17). The marginal posterior distribution for ϕ is given by

$$\pi_5(\phi|x) \propto \sqrt{\frac{\phi\psi'(\phi) - 1}{\phi}}\frac{\Gamma(n\phi)}{\Gamma(\phi)^n}\left(\frac{\sqrt[n]{\prod_{i=1}^n x_i}}{\sum_{i=1}^n x_i}\right)^{n\phi}.$$

4.4.5.2 Reference prior when μ is the parameter of interest

The reference prior when μ is the parameter of interest and ϕ is the nuisance parameter is given by

$$\pi_6(\phi,\mu) \propto \frac{\sqrt{\psi'(\phi)}}{\mu}. \tag{4.24}$$

The joint posterior distribution for ϕ and μ, produced by the reference prior (4.24) is given by

$$\pi_6(\phi,\mu|x) \propto \mu^{n\phi-1}\frac{\sqrt{\psi'(\phi)}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n} x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^{n} x_i\right\}. \tag{4.25}$$

Theorem 4.4.6. The posterior density (4.25) is proper if and only if $n \geq 2$, in which case the posterior moments for ϕ and μ are finite.

Proof. Following Abramowitz and Stegun (1972), $\lim_{\phi\to 0^+}\dfrac{\psi'(\phi)}{\phi^{-2}} = 1$ and $\lim_{\phi\to\infty}\dfrac{\psi'(\phi)}{\phi^{-1}} = 1$. Then, $\sqrt{\psi'(\phi)} \underset{\phi\to 0^+}{\propto} \phi^{-1}$ and $\sqrt{\psi'(\phi)} \underset{\phi\to\infty}{\propto} \phi^{-\frac{1}{2}}$. Thus, $c=-1$, $s_0=-1$ and $s_\infty = -\frac{1}{2}$. Therefore, the result follows from the Theorem 4.3.1 and Theorem 4.3.2. □

The conditional posterior distribution for μ is (4.17). The marginal posterior distribution for ϕ is given by

$$\pi_6(\phi|x) \propto \sqrt{\psi'(\phi)}\frac{\Gamma(n\phi)}{\Gamma(\phi)^n}\left(\frac{\sqrt[n]{\prod_{i=1}^{n} x_i}}{\sum_{i=1}^{n} x_i}\right)^{n\phi}.$$

There are different ways to derive the same reference priors in the presence of nuisance parameters, e.g, Liseo (1993), Sun and Ye (1996) and Moala, Ramos and Achcar (2013).

4.4.5.3 Overall Reference prior

The reference priors presented so far consider the presence of nuisance parameters. However, in many situation we are simultaneously interested in all parameters of the model. Sun and Ye (1996) considered the Bar-Lev and Reiser (1982) two parameter exponential family and presented a straightforward procedure to derive overall reference priors. Since the gamma distribution can be expressed as Bar-Lev and Reiser's two parameter exponential distribution, the overall reference (BERGER *et al.*, 2015) is given by

$$\pi_7(\phi,\mu) \propto \frac{1}{\mu}\sqrt{\frac{\phi\psi'(\phi)-1}{\phi}} \tag{4.26}$$

which is the same as the reference prior when ϕ is the parameter of interest and μ is the nuisance parameter.

4.4.6 Maximal Data Information prior

The MDI prior for the gamma distribution (4.1) is given by

$$\pi_8(\phi,\mu) \propto \frac{\mu}{\Gamma(\phi)}\exp\left\{(\phi-1)\psi(\phi)-\phi\right\}. \tag{4.27}$$

The joint posterior distribution for ϕ and μ, produced by the MDI prior, is

$$\pi_8(\phi,\mu|x) \propto \frac{\mu^{n\phi+1}}{\Gamma(\phi)^{n+1}}\left\{\prod_{i=1}^{n} x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^{n}x_i+(\phi-1)\psi(\phi)-\phi\right\}. \tag{4.28}$$

Moala, Ramos and Achcar (2013) argued that the posterior distribution (4.28) is improper. However, the authors did not present a proof of such result. The following theorem present a formally rigorous proof in which confirmed such conjecture.

Theorem 4.4.7. The joint posterior density (4.28) is improper for any $n \in \mathbb{N}$.

Proof. Following Abramowitz and Stegun (1972), $\lim_{\phi\to 0^+}\dfrac{\Gamma(\phi)}{\phi^{-1}} = 1$ and $\lim_{\phi\to 0^+}\dfrac{\psi(\phi)}{\phi^{-1}} = -1$. Thus,

$$\begin{aligned}
\lim_{\phi\to 0^+}\frac{\pi(\phi)}{\phi^{s_0}} &= \lim_{\phi\to 0^+}\frac{\frac{1}{\Gamma(\phi)}e^{(\phi-1)\psi(\phi)-\phi}}{\phi^{s_0}} = \lim_{\phi\to 0^+}\frac{\phi^{-1}}{\Gamma(\phi)}\frac{e^{(\phi-1)\psi(\phi)-\phi}}{e^{\phi^{-1}}}\frac{e^{\phi^{-1}}}{\phi^{s_0-1}} \\
&= \lim_{\phi\to 0^+} 1\times e^{\phi\psi(\phi)-\phi}e^{-\psi(\phi)-\phi^{-1}}\frac{e^{\phi^{-1}}}{\phi^{s_0-1}} = \lim_{\phi\to 0^+} e^{\frac{\psi(\phi)}{\phi^{-1}}-\phi}e^{-\psi(\phi+1)}\frac{e^{\phi^{-1}}}{\phi^{s_0-1}} \\
&= e^{-1}e^{-\psi(1)}\lim_{\phi\to 0^+}\frac{e^{\phi^{-1}}}{\phi^{s_0-1}} = e^{-1}e^{-\psi(1)}\lim_{u\to\infty}\frac{e^{u}}{u^{-s_0+1}} = \infty.
\end{aligned} \tag{4.29}$$

Since $c=-1$ and $\lim_{\phi\to 0^+}\dfrac{\pi(\phi)}{\phi^{s_0}} = \infty\ \forall s_0 \in \mathbb{N}$, the result follows from the Theorem 4.3.1.

□

4.4.6.1 Modified MDI prior

Moala et al. (2013) introduces a modified maximal data information (MMDI) prior given by

$$\pi_9(\phi,\mu) \propto \frac{\mu}{\Gamma(\phi)}\exp\left\{(\phi-1)\frac{\psi(\phi)}{\Gamma(\phi)}-\phi\right\}. \tag{4.30}$$

The joint posterior distribution for ϕ and μ, produced by the MMDI prior, is

$$\pi_9(\phi,\mu|x) \propto \frac{\mu^{n\phi+1}}{\Gamma(\phi)^{n+1}}\left\{\prod_{i=1}^{n} x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^{n}x_i+(\phi-1)\frac{\psi(\phi)}{\Gamma(\phi)}-\phi\right\}. \tag{4.31}$$

Theorem 4.4.8. The posterior density (4.31) is proper for every $n \in \mathbb{N}$, in which case the posterior moments for ϕ and μ are finite.

Proof. From, $\lim_{\phi\to 0^+}\dfrac{\Gamma(\phi)}{\phi^{-1}}=1$ and $\lim_{\phi\to 0^+}\dfrac{\psi(\phi)}{\phi^{-1}}=-1$. Thus $\lim_{\phi\to 0^+}\dfrac{\psi(\phi)}{\Gamma(\phi)}=-1$ and

$$\begin{aligned}\lim_{\phi\to 0^+}\frac{\pi_9(\phi)}{\phi}&=\lim_{\phi\to 0^+}\frac{\frac{1}{\Gamma(\phi)}e^{(\phi-1)\frac{\psi(\phi)}{\Gamma(\phi)}-\phi}}{\phi}=\lim_{\phi\to 0^+}\frac{\phi^{-1}}{\Gamma(\phi)}e^{(\phi-1)\frac{\psi(\phi)}{\Gamma(\phi)}-\phi}\\&=1\times e^{(-1)(-1)-0}=e>0.\end{aligned}\tag{4.32}$$

On the other hand, $\lim_{\phi\to\infty}\frac{\psi(\phi)}{\log(\phi)}=1$ and by the Stirling approximation Abramowitz and Stegun (1972) we have $\lim_{\phi\to 0^+}\frac{\Gamma(\phi)}{\phi^{\phi-\frac{1}{2}}e^{-\phi}}=\sqrt{2\pi}$ and $\lim_{\phi\to\infty}\frac{\Gamma(\phi)}{\phi^2}=\infty$. Then

$$\begin{aligned}\lim_{\phi\to\infty}\frac{\pi_9(\phi)}{\phi^{\frac{1}{2}-\phi}}&=\lim_{\phi\to 0^+}\frac{\frac{1}{\Gamma(\phi)}e^{(\phi-1)\frac{\psi(\phi)}{\Gamma(\phi)}-\phi}}{\phi^{\frac{1}{2}-\phi}}=\lim_{\phi\to 0^+}\frac{\phi^{\phi-\frac{1}{2}}e^{-\phi}}{\Gamma(\phi)}e^{(\phi-1)\frac{\psi(\phi)}{\Gamma(\phi)}}\\&=\frac{1}{\sqrt{2\pi}}\lim_{\phi\to 0^+}e^{\left(1-\frac{1}{\phi}\right)\frac{\psi(\phi)}{\log(\phi)}\frac{\log(\phi)}{\phi}\frac{\phi^2}{\Gamma(\phi)}}=\frac{1}{\sqrt{2\pi}}e^{1\times 1\times 0\times 0}=\frac{1}{\sqrt{2\pi}}>0.\end{aligned}\tag{4.33}$$

Now, define

$$\pi_9^*(\phi)=\begin{cases}\phi, & \text{if } \phi\le 1\\ \phi^{\frac{1}{2}-\phi} & \text{if } \phi>1\end{cases}\quad\text{and}\quad \chi(\phi)=\begin{cases}\phi, & \text{if } \phi\le 1\\ \phi^{-\frac{1}{2}} & \text{if } \phi>1.\end{cases}\tag{4.34}$$

Then, from (4.32) and (4.33) we have $\pi_9(\phi)\underset{\phi\to 0^+}{\propto}\pi_9^*(\phi)$ and $\pi_9(\phi)\underset{\phi\to\infty}{\propto}\pi_9^*(\phi)$, which implies that $\pi_9(\phi)\propto\pi_9^*(\phi)$ from Proposition 2.5.3. However, $\pi_9^*(\phi)\le\chi(\phi)$ and the prior $\pi_9(\mu)\chi(\phi)=\mu\chi(\phi)$ leads to a proper posterior as well as posterior moments for every $n\in\mathbb{N}$ by Theorem 4.3.1 and Theorem 4.3.2. Therefore $\phi^r\mu^s\pi_9(\phi,\mu)\propto\phi^r\mu^s\pi_9(\mu)\pi_9^*(\phi)\le\phi^r\mu^s\pi_9(\mu)\chi(\phi)$ also leads to a proper posterior for every $n\in\mathbb{N}$, $s\in\mathbb{N}$ and $r\in\mathbb{N}$ which proves the result.

□

The marginal posterior distribution for ϕ is given by

$$\pi_9(\phi|x)\propto\frac{(\phi\psi'(\phi)-1)}{\sqrt{\phi}}\frac{\Gamma(n\phi+2)}{\Gamma(\phi)^n}\exp\left\{(\phi-1)\frac{\psi(\phi)}{\Gamma(\phi)}-\phi\right\}\left(\frac{\sqrt[n]{\prod_{i=1}^n x_i}}{\sum_{i=1}^n x_i}\right)^{n\phi}.$$

The conditional posterior distribution for μ is given by

$$\pi_9(\mu|\phi,x)\sim\text{Gamma}\left(n\phi+2,\sum_{i=1}^n x_i\right).$$

4.4.7 Matching priors

Sun and Ye (1996) prove that the reference prior (4.22) is also a Tibshirani prior when ϕ is the parameter of interest and μ is the nuisance parameter and the Tibshirani prior when μ is the parameter of interest and ϕ is the nuisance parameter with order $O(n^{-1})$. They also proved that when ϕ is the parameter of interest, there is no matching prior up to order $o(n^{-1})$. Finally, they present a Tibshirani prior when μ is the parameter of interest that is matching prior up to order $o(n^{-1})$, such prior is given as follows

$$\pi_{10}(\phi,\mu) \propto \frac{\phi\psi'(\phi)-1}{\mu\sqrt{\phi}}. \tag{4.35}$$

The joint posterior distribution for ϕ and μ, produced by the Tibshirani prior (4.35) is given by

$$\pi_{10}(\phi,\mu|x) \propto \frac{(\phi\psi'(\phi)-1)}{\sqrt{\phi}}\frac{\mu^{n\phi-1}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n}x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^{n}x_i\right\}. \tag{4.36}$$

Theorem 4.4.9. The posterior density (4.36) is proper if and only if $n \geq 2$, in which case the posterior moments for ϕ and μ are finite.

Proof. We proved in Theorem 4.4.3 that $\sqrt{\phi\psi'(\phi)-1} \underset{\phi\to 0^+}{\propto} \phi^{-\frac{1}{2}}$ and that $\sqrt{\phi\psi'(\phi)-1} \underset{\phi\to\infty}{\propto} \phi^{-\frac{1}{2}}$. From that, it follows that

$$\frac{\phi\psi'(\phi)-1}{\sqrt{\phi}} \underset{\phi\to 0^+}{\propto} \frac{\phi^{-1}}{\phi^{\frac{1}{2}}} = \phi^{-\frac{3}{2}} \text{ and } \frac{\phi\psi'(\phi)-1}{\sqrt{\phi}} \underset{\phi\to\infty}{\propto} \frac{\phi^{-1}}{\phi^{\frac{1}{2}}} = \phi^{-\frac{3}{2}}.$$

Thus $c=-1$ and $s_0 = s_\infty = -\frac{3}{2}$, therefore the result follows directly from the Theorem 4.3.1 and Theorem 4.3.2. □

The conditional posterior distribution for μ is (4.17). The marginal posterior distribution for ϕ is given by

$$\pi_{10}(\phi|x) \propto \frac{(\phi\psi'(\phi)-1)}{\sqrt{\phi}}\frac{\Gamma(n\phi)}{\Gamma(\phi)^n}\left(\frac{\sqrt[n]{\prod_{i=1}^{n}x_i}}{\sum_{i=1}^{n}x_i}\right)^{n\phi}.$$

4.4.8 Consensus Prior

A rather natural approach to find an objective prior is to start with a collection of objective priors and take its average. Berger *et al.* (2015) discussed this prior averaging approach under the two most natural averages, the geometric mean and the arithmetic mean.

4.4.8.1 Geometric mean

Let $\pi_i(\phi,\mu), i=3,5,6,7,10$ be a collection of objective priors. Such priors were selected conveniently due its invariance property under one-to-one transformations. Then, our geometric mean (GM) prior is given by

$$\pi_{11}(\phi,\mu) \propto \frac{1}{\mu}\sqrt[5]{\frac{(\phi\psi'(\phi)-1)^{\frac{5}{2}}\psi'(\phi)^{\frac{1}{2}}}{\phi^{\frac{3}{2}}}} \propto \frac{1}{\mu}\frac{\sqrt{\phi\psi'(\phi)-1}\,\sqrt[10]{\psi'(\phi)}}{\phi^{\frac{3}{10}}}. \tag{4.37}$$

Note that, since our prior was constructed as a geometric mean of one-to-one invariant priors then such prior has also invariance property under one-to-one transformations.

The joint posterior distribution for ϕ and μ, produced by the consensus prior, is

$$\pi_{11}(\phi,\mu|x) \propto \frac{\psi'(\phi)^{\frac{1}{10}}\sqrt{\phi\psi'(\phi)-1}}{\phi^{\frac{3}{10}}}\frac{\mu^{n\phi-1}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n}x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^{n}x_i\right\}. \tag{4.38}$$

Theorem 4.4.10. The posterior density (4.38) is proper if and only if $n \geq 2$, in which case the posterior moments for ϕ and μ are finite.

Proof. The result follows directly from the Theorem 4.3.3 and by Theorem 4.3.2. □

The conditional posterior distribution for μ is (4.17). The marginal posterior distribution for ϕ is given by

$$\pi_{11}(\phi|x) \propto \frac{\psi'(\phi)^{\frac{1}{10}}\sqrt{\phi\psi'(\phi)-1}}{\phi^{\frac{3}{10}}}\frac{\Gamma(n\phi)}{\Gamma(\phi)^n}\left(\frac{\sqrt[n]{\prod_{i=1}^{n}x_i}}{\sum_{i=1}^{n}x_i}\right)^{n\phi}.$$

4.4.8.2 Arithmetic mean

Let $\pi_i(\phi,\mu), i=3,5,6,7,10$ be a collection of objective priors. Then, our arithmetic mean (AM) prior is given by

$$\pi_{12}(\phi,\mu) \propto \frac{\pi_{12}(\phi)}{\mu}$$

where

$$\pi_{12}(\phi) = \left(\frac{2\sqrt{\phi\psi'(\phi)-1}+\sqrt{\phi\psi'(\phi)}+\sqrt{\phi^2\psi'(\phi)-\phi+\phi\psi'(\phi)-1}}{\sqrt{\phi}}\right).$$

The joint posterior distribution for ϕ and μ, produced by the consensus prior, is

$$\pi_{12}(\phi,\mu|x) \propto \pi_{12}(\phi)\frac{\mu^{n\phi-1}}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n}x_i^{\phi}\right\}\exp\left\{-\mu\sum_{i=1}^{n}x_i\right\}. \tag{4.39}$$

Theorem 4.4.11. The posterior density (4.39) is proper if and only if $n \geq 2$, in which case the posterior moments for ϕ and μ are finite.

Proof. The result follows directly from the Theorem 4.3.3 and by Theorem 4.3.2. □

The conditional posterior distribution for μ is (4.17). The marginal posterior distribution for ϕ is given by

$$\pi_{12}(\phi|x) \propto \pi_{12}(\phi)\frac{\Gamma(n\phi)}{\Gamma(\phi)^n}\left(\frac{\sqrt[n]{\prod_{i=1}^{n}x_i}}{\sum_{i=1}^{n}x_i}\right)^{n\phi}.$$

4.5 Numerical evaluation

A simulation study is presented to compare the influence of different objective priors in the posterior distributions and find the most efficient estimation method by computing the mean relative errors (MRE) and the mean square errors (MSE), given by

$$\mathrm{MRE}_i \frac{1}{N}\sum_{j=1}^{N}\frac{\hat{\theta}_{i,j}}{\theta_i} \quad \text{and} \quad \mathrm{MSE}_i=\sum_{j=1}^{N}\frac{(\hat{\theta}_{i,j}-\theta_i)^2}{N}, \quad i=1,2$$

where $\theta=(\phi,\mu)$ and $N=10,000$ is the number of estimates obtained through the posterior means of ϕ and μ. The 95% coverage probability ($CP_{95\%}$) of the credibility intervals for ϕ and μ are evaluated. Considering this approach, the best estimators will show MRE closer to one and MSE closer to zero. In addition, for a large number of experiments considering a 95% confidence level, the frequencies of intervals that covered the true values of θ should be closer to 95%.

The results were computed using the software R. Considering $n=(10,20,\ldots,120)$ the results were presented only for $\theta=((4,2),(0.5,5))$ for reasons of space. However, the following results were similar for other choices of ϕ and μ.

The Markov chain Monte Carlo (MCMC) algorithm is considered in order to generate samples from marginal posteriors. Since the marginal posterior of ϕ do not belong to any known parametric family and the marginal posterior of μ has a gamma distribution. Then, we considered the Metropolis-Hastings with Gibbs Sampling. The gamma distribution was used as transition kernel $q\left(\phi^{(j)}|\phi^{(*)},b\right)$ for sampling values of ϕ, in this case b is fixed value that control the rate of acceptance. While we choose b to be equal to one, other higher values can also be considered. Moreover, to increase the time of convergence of the algorithm, we consider the method of moments as a good initial value for ϕ and μ given by

$$\phi^{(1)}=\frac{\bar{x}^2}{s^2} \quad \text{and} \quad \mu^{(1)}=\frac{\bar{x}}{s^2}. \tag{4.40}$$

The Metropolis-Hastings algorithm operates as follows:

1. Start with an initial value $\phi^{(1)}$ and set the iteration counter $j=1$;
2. Generate a random value $\phi^{(*)}$ from the proposal Gamma$(\phi^{(j)},1)$;
3. Evaluate the acceptance probability

$$\lambda\left(\phi^{(j)},\phi^{(*)}\right)=\min\left(1,\frac{p\left(\phi^{(*)}|x\right)}{p\left(\phi^{(j)}|x\right)}\frac{\mathrm{q}\left(\phi^{(j)},\phi^{(*)},1\right)}{\mathrm{q}\left(\phi^{(*)},\phi^{(j)},1\right)}\right),$$

where $p(\cdot)$ is one of the marginal posterior distributions.

4. Generate a random value u from an independent uniform in $(0,1)$;

5. If $\lambda\left(\phi^{(j)}, \phi^{(*)}\right) \geq u(0,1)$ then $\phi^{(j+1)} = \phi^{(*)}$. Otherwise, $\phi^{(j+1)} = \phi^{(j)}$;
6. Generate a random value $\phi^{(j+1)}$ from the conditional posterior Gamma$(\cdot,\cdot)$ related with the marginal posterior distribution chosen in the step 3;
7. Change the counter from j to $j+1$ and return to step 2 until convergence is reached.

Using the MCMC methods, we computed the posterior mean for ϕ, μ and the credibility (confidence) intervals for both parameters. Tables 13 and 12 present the MREs, MSEs and $CP_{95\%}$ from the different estimators of ϕ and μ .

Table 11 – The MRE(MSE) for for the estimates of ϕ and μ considering different sample sizes.

θ	n	MLE	Uniform	Jeffreys' Rule	Jeffreys' Prior	Miller	Reference ϕ
$\phi=2$	10	1.398(3.356)	1.336(1.305)	1.130(0.655)	1.175(0.768)	1.232(0.928)	1.124(0.646)
	20	1.161(0.733)	1.209(0.598)	1.080(0.354)	1.109(0.397)	1.144(0.457)	1.076(0.351)
	30	1.101(0.384)	1.153(0.360)	1.059(0.234)	1.080(0.256)	1.106(0.288)	1.056(0.232)
	40	1.073(0.248)	1.122(0.256)	1.048(0.177)	1.064(0.191)	1.084(0.211)	1.045(0.176)
	50	1.058(0.184)	1.100(0.195)	1.039(0.142)	1.053(0.151)	1.070(0.164)	1.038(0.141)
	60	1.048(0.145)	1.084(0.157)	1.032(0.120)	1.044(0.126)	1.058(0.136)	1.031(0.119)
	70	1.040(0.117)	1.071(0.127)	1.026(0.100)	1.036(0.105)	1.049(0.111)	1.025(0.099)
	80	1.034(0.101)	1.067(0.110)	1.027(0.088)	1.036(0.092)	1.047(0.097)	1.026(0.087)
	90	1.031(0.089)	1.057(0.094)	1.022(0.077)	1.030(0.080)	1.040(0.084)	1.021(0.077)
	100	1.028(0.079)	1.055(0.086)	1.022(0.071)	1.030(0.074)	1.039(0.078)	1.021(0.071)
	110	1.025(0.072)	1.047(0.074)	1.018(0.063)	1.024(0.065)	1.033(0.068)	1.017(0.062)
	120	1.024(0.064)	1.045(0.069)	1.018(0.059)	1.024(0.060)	1.031(0.063)	1.017(0.058)
$\mu=0.5$	10	1.467(0.276)	1.395(0.107)	1.157(0.053)	1.204(0.062)	1.262(0.074)	1.151(0.053)
	20	1.188(0.059)	1.246(0.049)	1.098(0.029)	1.127(0.032)	1.163(0.037)	1.094(0.028)
	30	1.117(0.031)	1.181(0.030)	1.073(0.019)	1.094(0.021)	1.121(0.023)	1.070(0.019)
	40	1.084(0.020)	1.143(0.022)	1.059(0.015)	1.075(0.016)	1.096(0.018)	1.056(0.015)
	50	1.067(0.015)	1.118(0.016)	1.048(0.011)	1.061(0.012)	1.079(0.013)	1.046(0.011)
	60	1.056(0.012)	1.100(0.013)	1.041(0.010)	1.052(0.010)	1.067(0.011)	1.039(0.010)
	70	1.046(0.010)	1.085(0.011)	1.034(0.008)	1.044(0.009)	1.056(0.009)	1.032(0.008)
	80	1.040(0.008)	1.080(0.009)	1.034(0.007)	1.043(0.008)	1.054(0.008)	1.033(0.007)
	90	1.036(0.007)	1.068(0.008)	1.027(0.006)	1.035(0.007)	1.045(0.007)	1.026(0.006)
	100	1.032(0.006)	1.064(0.007)	1.027(0.006)	1.034(0.006)	1.043(0.006)	1.026(0.006)
	110	1.029(0.006)	1.055(0.006)	1.022(0.005)	1.028(0.005)	1.036(0.005)	1.021(0.005)
	120	1.027(0.005)	1.053(0.006)	1.022(0.005)	1.028(0.005)	1.036(0.005)	1.021(0.005)
$\phi=4$	10	1.398(12.062)	1.348(5.558)	1.128(2.743)	1.179(3.256)	1.238(3.930)	1.124(2.724)
	20	1.174(3.233)	1.217(2.522)	1.079(1.467)	1.111(1.659)	1.148(1.909)	1.077(1.458)
	30	1.108(1.619)	1.159(1.535)	1.059(0.986)	1.082(1.084)	1.109(1.217)	1.057(0.982)
	40	1.078(1.049)	1.127(1.136)	1.048(0.787)	1.067(0.850)	1.088(0.934)	1.047(0.784)
	50	1.062(0.774)	1.105(0.832)	1.040(0.601)	1.055(0.642)	1.072(0.698)	1.039(0.599)
	60	1.051(0.613)	1.088(0.673)	1.033(0.508)	1.045(0.537)	1.060(0.577)	1.032(0.506)
	70	1.044(0.506)	1.077(0.562)	1.029(0.436)	1.040(0.459)	1.053(0.489)	1.028(0.435)
	80	1.038(0.435)	1.069(0.475)	1.027(0.376)	1.037(0.394)	1.048(0.418)	1.026(0.375)
	90	1.033(0.377)	1.061(0.410)	1.023(0.333)	1.031(0.346)	1.041(0.365)	1.022(0.332)
	100	1.031(0.336)	1.053(0.350)	1.018(0.290)	1.027(0.301)	1.036(0.315)	1.018(0.290)
	110	1.027(0.299)	1.049(0.320)	1.018(0.269)	1.025(0.278)	1.034(0.291)	1.017(0.268)
	120	1.025(0.273)	1.046(0.292)	1.017(0.247)	1.024(0.256)	1.032(0.267)	1.017(0.247)
$\mu=2$	10	1.429(3.440)	1.377(1.616)	1.142(0.798)	1.194(0.942)	1.253(1.130)	1.138(0.793)
	20	1.187(0.907)	1.235(0.728)	1.088(0.423)	1.120(0.476)	1.157(0.545)	1.086(0.420)
	30	1.115(0.455)	1.174(0.446)	1.066(0.286)	1.090(0.313)	1.117(0.350)	1.064(0.285)
	40	1.084(0.299)	1.138(0.327)	1.053(0.226)	1.072(0.243)	1.093(0.266)	1.052(0.225)
	50	1.067(0.223)	1.115(0.240)	1.045(0.172)	1.060(0.184)	1.078(0.200)	1.044(0.172)
	60	1.055(0.175)	1.096(0.193)	1.037(0.145)	1.050(0.153)	1.065(0.164)	1.036(0.145)
	70	1.047(0.145)	1.083(0.160)	1.032(0.124)	1.043(0.130)	1.056(0.138)	1.031(0.123)
	80	1.041(0.123)	1.076(0.136)	1.030(0.108)	1.040(0.113)	1.051(0.119)	1.029(0.107)
	90	1.036(0.108)	1.067(0.116)	1.026(0.094)	1.035(0.098)	1.045(0.103)	1.025(0.094)
	100	1.033(0.095)	1.058(0.100)	1.021(0.083)	1.029(0.086)	1.038(0.090)	1.020(0.083)
	110	1.029(0.086)	1.053(0.091)	1.020(0.076)	1.027(0.079)	1.035(0.082)	1.019(0.076)
	120	1.027(0.078)	1.051(0.085)	1.020(0.071)	1.026(0.074)	1.034(0.077)	1.019(0.071)

Table 12 – The MRE(MSE) for for the estimates of ϕ and μ considering different sample sizes.

θ	n	Reference μ	MDIP	Tibshirani	Consensus GM	Consensus AM
	10	1.169(0.762)	1.082(0.304)	1.067(0.542)	1.131(0.664)	1.252(1.014)
	20	1.105(0.394)	1.070(0.230)	1.041(0.312)	1.081(0.358)	1.156(0.487)
	30	1.077(0.254)	1.061(0.179)	1.030(0.211)	1.060(0.236)	1.114(0.302)
	40	1.061(0.190)	1.054(0.149)	1.025(0.164)	1.048(0.179)	1.091(0.220)
$\phi = 2$	50	1.051(0.151)	1.049(0.124)	1.021(0.133)	1.040(0.143)	1.075(0.170)
	60	1.042(0.125)	1.042(0.109)	1.016(0.113)	1.032(0.120)	1.062(0.140)
	70	1.034(0.104)	1.036(0.093)	1.012(0.095)	1.026(0.100)	1.052(0.114)
	80	1.035(0.092)	1.038(0.083)	1.015(0.083)	1.028(0.088)	1.050(0.100)
	90	1.029(0.080)	1.032(0.074)	1.011(0.074)	1.022(0.077)	1.043(0.086)
	100	1.028(0.074)	1.032(0.069)	1.012(0.068)	1.023(0.071)	1.041(0.079)
	110	1.023(0.065)	1.027(0.061)	1.009(0.061)	1.018(0.063)	1.035(0.069)
	120	1.023(0.060)	1.027(0.058)	1.009(0.057)	1.018(0.059)	1.033(0.064)
	10	1.198(0.062)	1.160(0.033)	1.093(0.044)	1.159(0.054)	1.283(0.081)
	20	1.123(0.032)	1.121(0.023)	1.058(0.025)	1.099(0.029)	1.175(0.039)
	30	1.091(0.021)	1.101(0.017)	1.044(0.018)	1.074(0.020)	1.129(0.025)
	40	1.072(0.016)	1.085(0.014)	1.035(0.014)	1.059(0.015)	1.102(0.018)
$\mu = 0.5$	50	1.059(0.012)	1.074(0.011)	1.029(0.011)	1.048(0.012)	1.084(0.014)
	60	1.050(0.010)	1.065(0.010)	1.025(0.009)	1.041(0.010)	1.071(0.011)
	70	1.042(0.009)	1.057(0.008)	1.020(0.008)	1.034(0.008)	1.060(0.009)
	80	1.042(0.008)	1.055(0.007)	1.021(0.007)	1.034(0.007)	1.057(0.008)
	90	1.034(0.007)	1.047(0.006)	1.016(0.006)	1.027(0.006)	1.048(0.007)
	100	1.033(0.006)	1.046(0.006)	1.017(0.006)	1.027(0.006)	1.046(0.006)
	110	1.027(0.005)	1.039(0.005)	1.012(0.005)	1.022(0.005)	1.039(0.006)
	120	1.027(0.005)	1.039(0.005)	1.014(0.005)	1.022(0.005)	1.038(0.005)
	10	1.177(3.243)	0.841(0.779)	1.066(2.290)	1.134(2.810)	1.270(4.420)
	20	1.109(1.651)	0.879(0.616)	1.040(1.302)	1.083(1.493)	1.167(2.086)
	30	1.081(1.081)	0.902(0.507)	1.030(0.897)	1.062(1.000)	1.123(1.308)
	40	1.066(0.849)	0.919(0.452)	1.026(0.730)	1.050(0.796)	1.099(0.993)
	50	1.054(0.642)	0.929(0.380)	1.021(0.563)	1.042(0.606)	1.081(0.736)
$\phi = 4$	60	1.045(0.536)	0.936(0.346)	1.017(0.480)	1.034(0.511)	1.068(0.603)
	70	1.040(0.458)	0.943(0.311)	1.015(0.416)	1.030(0.440)	1.060(0.510)
	80	1.036(0.393)	0.949(0.277)	1.015(0.359)	1.028(0.379)	1.054(0.433)
	90	1.031(0.345)	0.952(0.255)	1.011(0.319)	1.023(0.334)	1.047(0.377)
	100	1.026(0.300)	0.954(0.232)	1.009(0.280)	1.019(0.291)	1.040(0.325)
	110	1.025(0.278)	0.958(0.218)	1.009(0.261)	1.019(0.270)	1.038(0.299)
	120	1.023(0.255)	0.962(0.203)	1.009(0.240)	1.018(0.249)	1.035(0.273)
	10	1.191(0.939)	0.877(0.198)	1.079(0.667)	1.148(0.817)	1.285(1.265)
	20	1.118(0.474)	0.903(0.164)	1.049(0.375)	1.092(0.430)	1.177(0.593)
	30	1.088(0.313)	0.921(0.138)	1.037(0.260)	1.069(0.290)	1.131(0.375)
	40	1.071(0.243)	0.933(0.124)	1.030(0.210)	1.055(0.228)	1.104(0.282)
	50	1.059(0.184)	0.942(0.104)	1.026(0.161)	1.047(0.174)	1.087(0.210)
$\mu = 2$	60	1.049(0.153)	0.947(0.095)	1.021(0.137)	1.038(0.146)	1.072(0.171)
	70	1.042(0.130)	0.952(0.086)	1.018(0.118)	1.033(0.125)	1.063(0.144)
	80	1.039(0.112)	0.958(0.077)	1.018(0.103)	1.031(0.108)	1.057(0.123)
	90	1.034(0.098)	0.961(0.070)	1.015(0.090)	1.027(0.094)	1.050(0.106)
	100	1.028(0.086)	0.961(0.064)	1.011(0.080)	1.022(0.083)	1.043(0.092)
	110	1.027(0.079)	0.964(0.060)	1.011(0.074)	1.020(0.077)	1.040(0.084)
	120	1.026(0.074)	0.968(0.057)	1.011(0.069)	1.020(0.072)	1.038(0.079)

Table 13 – The $CP_{95\%}$ from the estimates of μ and Ω considering different values of n with $N = 10,000$ simulated samples.

θ	n	MLE		Uniform		Jeffreys' Rule		Jeffreys' Prior		Miller		Reference ϕ	
		μ	θ	μ	θ	μ	θ	μ	θ	μ	θ	μ	θ
	10	0.973	0.975	0.892	0.891	0.950	0.953	0.942	0.948	0.928	0.936	0.948	0.953
	20	0.964	0.966	0.907	0.908	0.946	0.949	0.941	0.945	0.938	0.937	0.944	0.949
$\phi = 2$	30	0.958	0.959	0.918	0.916	0.948	0.950	0.944	0.945	0.937	0.939	0.948	0.951
	40	0.957	0.956	0.923	0.921	0.948	0.946	0.944	0.944	0.938	0.942	0.947	0.947
	50	0.955	0.955	0.929	0.928	0.950	0.950	0.948	0.947	0.942	0.944	0.951	0.949
	60	0.955	0.955	0.929	0.928	0.947	0.946	0.943	0.944	0.941	0.940	0.944	0.946
$\mu = 0.5$	70	0.955	0.957	0.934	0.931	0.946	0.948	0.943	0.947	0.942	0.943	0.946	0.948
	80	0.953	0.953	0.934	0.934	0.948	0.947	0.947	0.948	0.943	0.944	0.946	0.949
	90	0.954	0.955	0.938	0.936	0.949	0.949	0.948	0.948	0.944	0.943	0.948	0.948
	100	0.955	0.952	0.938	0.933	0.945	0.945	0.945	0.944	0.943	0.940	0.944	0.944
	110	0.949	0.953	0.941	0.939	0.948	0.951	0.950	0.950	0.945	0.945	0.946	0.948
	120	0.953	0.953	0.942	0.943	0.950	0.951	0.949	0.950	0.948	0.950	0.948	0.953
	10	0.974	0.972	0.886	0.885	0.948	0.951	0.942	0.942	0.928	0.931	0.949	0.952
$\phi = 4$	20	0.963	0.965	0.907	0.905	0.947	0.946	0.942	0.941	0.935	0.933	0.947	0.944
	30	0.962	0.961	0.917	0.914	0.952	0.950	0.948	0.947	0.939	0.939	0.952	0.949
	40	0.958	0.962	0.923	0.920	0.947	0.946	0.946	0.943	0.939	0.939	0.947	0.948
	50	0.961	0.959	0.924	0.926	0.947	0.947	0.944	0.945	0.939	0.940	0.945	0.947
	60	0.956	0.957	0.932	0.930	0.950	0.950	0.948	0.946	0.942	0.943	0.951	0.949
	70	0.959	0.961	0.929	0.929	0.945	0.946	0.945	0.944	0.940	0.938	0.946	0.946
$\mu = 2$	80	0.955	0.956	0.936	0.936	0.950	0.952	0.949	0.949	0.944	0.946	0.950	0.951
	90	0.956	0.956	0.934	0.938	0.946	0.948	0.945	0.948	0.941	0.945	0.944	0.949
	100	0.955	0.956	0.941	0.939	0.949	0.952	0.949	0.949	0.947	0.948	0.951	0.951
	110	0.956	0.954	0.940	0.939	0.949	0.949	0.949	0.949	0.947	0.947	0.949	0.950
	120	0.953	0.953	0.940	0.937	0.949	0.945	0.947	0.944	0.944	0.943	0.947	0.945

Table 14 – The $CP_{95\%}$ from the estimates of μ and Ω considering different values of n with $N = 10,000$ simulated samples.

θ		Reference μ		MDIP		Tibshirani		Consensus GM		Consensus AM	
		μ	θ	μ	θ	μ	θ	μ	θ	μ	θ
	10	0.943	0.948	0.969	0.964	0.950	0.957	0.949	0.953	0.922	0.928
	20	0.942	0.946	0.960	0.956	0.947	0.951	0.945	0.949	0.931	0.932
$\phi = 2$	30	0.946	0.946	0.958	0.949	0.950	0.950	0.949	0.949	0.933	0.936
	40	0.943	0.944	0.953	0.948	0.949	0.948	0.948	0.947	0.937	0.938
	50	0.948	0.948	0.955	0.948	0.950	0.951	0.949	0.949	0.941	0.941
	60	0.944	0.943	0.948	0.945	0.945	0.946	0.946	0.946	0.938	0.938
$\mu = 0.5$	70	0.944	0.947	0.948	0.948	0.944	0.948	0.946	0.947	0.938	0.938
	80	0.946	0.945	0.948	0.946	0.949	0.948	0.946	0.948	0.940	0.942
	90	0.948	0.947	0.950	0.945	0.948	0.949	0.949	0.947	0.944	0.943
	100	0.944	0.943	0.946	0.942	0.946	0.945	0.944	0.944	0.943	0.941
	110	0.947	0.950	0.949	0.948	0.948	0.948	0.948	0.947	0.944	0.944
	120	0.948	0.950	0.950	0.950	0.950	0.952	0.949	0.951	0.946	0.947
	10	0.942	0.942	0.920	0.942	0.951	0.954	0.949	0.949	0.918	0.920
$\phi = 4$	20	0.943	0.940	0.923	0.940	0.950	0.946	0.947	0.944	0.928	0.928
	30	0.949	0.947	0.921	0.938	0.950	0.949	0.951	0.948	0.934	0.933
	40	0.945	0.944	0.925	0.939	0.946	0.948	0.948	0.947	0.935	0.935
	50	0.946	0.944	0.929	0.943	0.948	0.946	0.945	0.947	0.936	0.937
	60	0.947	0.947	0.933	0.940	0.952	0.950	0.950	0.948	0.940	0.941
	70	0.943	0.946	0.931	0.939	0.946	0.947	0.945	0.944	0.938	0.938
$\mu = 2$	80	0.949	0.949	0.937	0.942	0.950	0.952	0.948	0.951	0.943	0.945
	90	0.945	0.947	0.931	0.943	0.944	0.949	0.944	0.948	0.940	0.943
	100	0.951	0.951	0.936	0.942	0.949	0.950	0.950	0.952	0.946	0.945
	110	0.947	0.948	0.937	0.943	0.949	0.950	0.950	0.951	0.944	0.946
	120	0.947	0.944	0.936	0.941	0.946	0.947	0.947	0.945	0.944	0.942

From these results, we observed that the MREs (MSEs) for all estimators of the parameters tend to one (zero) for large n, i.e., the estimators are asymptotically unbiased for the parameters. Moreover, for both parameters the posterior mean using the Tibshirani prior indicates better performance than the obtained with other priors and the MLE. The better performance of this approach is also confirmed through the coverage probability obtained from the credibility intervals. Therefore, we conclude that the posterior distribution obtained with Tibshirani prior should be used to make inference on the parameters of the gamma distribution.

4.6 Discussion

In this study, we presented a theorem that provides simple conditions under which an improper prior yields a proper posterior for the gamma distribution. Further, we provided sufficient conditions to verify if the posterior moments of the parameters are finite. An interesting aspect of our findings is that one can check if the posterior is proper or improper and also if its posterior moments are finite looking directly the behavior of the proposed improper prior.

The proposed methodology is applied in different objective priors. The MDI prior was the only that yield an improper posterior for any sample sizes. An extensive simulation study showed that the posterior distribution obtained under Tibshirani prior provided more accurate results in terms of MRE, MSE and coverage probabilities. Therefore, this posterior distribution should be used to make inference in the unknown parameters of the gamma distribution. This study can be easily extended for other distributions such as Weibull, generalized gamma and the generalized extreme value distribution providing simple conditions to check under which an improper prior yields a proper posterior.

CHAPTER

5

GENERALIZED GAMMA DISTRIBUTION

5.1 Introduction

The generalized gamma (GG) distribution plays an important role in statistics and has proven to be very flexible in practice for modeling data from several areas, such as climatology, meteorology medicine, reliability and image processing data, among others. Introduced by Stacy and Mihram (1965) the GG distribution unify many important models such as the exponential, Weibull, gamma, lognormal, generalized normal, Nakagami-m, half-normal, Rayleigh, Maxwell-Boltzmann and chi distribution, to list a few. A random variable X follows a GG distribution if its probability density function (PDF) is given by

$$f(x|\theta) = \frac{\alpha}{\Gamma(\phi)}\mu^{\alpha\phi}x^{\alpha\phi-1}\exp\left(-(\mu x)^{\alpha}\right), \quad x > 0 \tag{5.1}$$

where $\Gamma(\phi) = \int_0^\infty e^{-x}x^{\phi-1}dx$ is the gamma function, $\theta = (\phi, \mu, \alpha)$, $\alpha > 0$ and $\phi > 0$ are the shape parameters and $\mu > 0$ is a scale parameter.

The parameter estimators for the GG distribution have been discussed earlier considering the maximum likelihood (ML) method (STACY; MIHRAM, 1965). However, the ML estimators are not well-behaved (HAGER; BAIN, 1970) and its asymptotic properties may not be achieved even for samples greater than 400 (PRENTICE, 1974). From a Bayesian point of view, a subjective analysis can be considered where the prior distribution supplies information from an expert (O'HAGAN *et al.*, 2006). On the other hand, in many situations, we are interested in obtaining a prior distribution which guarantees that the information provided by the data will not be overshadowed by subjective information. In this case, an objective analysis is recommended by considering non-informative priors that are derived by formal rules (BERNARDO, 2005). Although several studies have considered weakly informative priors (flat priors) as presumed non-informative priors, Bernardo (2005) argued that using simple proper priors presumed to be non-informative, often hides important unwarranted assumptions which may easily dominate, or even invalidate the statistical analysis and should be strongly discouraged.

Objective priors have been discussed for the generalized gamma distribution (NOORTWIJK, 2001). The obtained priors are constructed by formal rules (KASS; WASSERMAN, 1996) and are usually improper, i.e., do not correspond to proper probability distribution and could lead to improper posteriors, which is undesirable. According to Northrop and Attalides (2016), there are no simple conditions that can be used to prove that an improper prior yields a proper posterior for a particular distribution, therefore a case-by-case investigation is needed to check the propriety of posterior distribution. This study overcomes this problem by providing in a simple way necessary and sufficient conditions to check whether or not objective priors lead proper posterior distributions for the generalized gamma distribution. As a result, one can easily check if the obtained posterior is proper or improper directly looking at the behavior of the improper prior.

The proposed methodology is fully illustrated in twelve improper priors such as independent uniform priors, Jeffreys' rule (KASS; WASSERMAN, 1996), Jeffreys' prior (JEFFREYS, 1946), maximal data information (MDI) prior (ZELLNER, 1977; ZELLNER, 1984), reference priors (BERNARDO, 1979; BERNARDO, 2005; BERGER *et al.*, 2015), to list a few. We proved that among the priors considered only one reference prior returned a proper posterior distribution. The proper reference posterior has excellent theoretical properties such as invariance property under one-to-one transformations of the parameters, consistency under marginalization and consistent sampling properties. Despite the fact that the posterior distribution may be proper the posterior moments can be infinite. Therefore, we also provided sufficient conditions to verify if the posterior moments are finite.

The remainder of this chapter is organized as follows. Section 2 presents a theorem that provides necessary and sufficient conditions for the posterior distributions to be proper and also sufficient conditions to check if the posterior moments of the parameters are finite. Section 3 presents the applications of our main theorem in different objective priors. Finally, Section 4 summarizes the study.

5.2 Maximum likelihood estimators

Among the classical statistical inference methods, the ML method is usually preferred due to its better asymptotic properties. The ML estimators are obtained by maximizing the likelihood function. Let $T_1,\ldots,T_n$ be a random sample where $T \sim \mathrm{GG}(\alpha,\mu,\phi)$, the likelihood function for the parameter vector $\theta=(\alpha,\mu,\phi)$ is given by

$$L(\theta;t)=\frac{\alpha^n}{\Gamma(\phi)^n}\mu^{n\alpha\phi}\left\{\prod_{i=1}^{n}t_i^{\alpha\phi-1}\right\}\exp\left\{-\mu^{\alpha}\sum_{i=1}^{n}t_i^{\alpha}\right\}. \tag{5.2}$$

The maximum likelihood estimates of the parameters are obtained by solving the likelihood equations $\frac{\partial}{\partial\alpha}\log(L(\theta;t))=0$, $\frac{\partial}{\partial\mu}\log(L(\theta;t))=0$, $\frac{\partial}{\partial\phi}\log(L(\theta;t))=0$. Therefore, from

(5.2), we have

$$n\,\psi(\hat{\phi}) = n\hat{\alpha}\log(\hat{\mu}) + \hat{\alpha}\sum_{i=1}^{n}\log(t_i), \tag{5.3}$$

$$n\hat{\phi} = \hat{\mu}^{\hat{\alpha}}\sum_{i=1}^{n} t_i^{\hat{\alpha}}, \tag{5.4}$$

$$\frac{n}{\hat{\alpha}} + n\hat{\phi}\log(\mu) + \phi\sum_{i=1}^{n}\log(t_i) = \hat{\mu}^{\hat{\alpha}}\sum_{i=1}^{n} t_i^{\hat{\alpha}}\log(\hat{\mu}t_i), \tag{5.5}$$

where $\psi(k) = \frac{\partial}{\partial k}\log\Gamma(k) = \frac{\Gamma'(k)}{\Gamma(k)}$ is the digamma function. The solutions of (5.3-5.5) provide the maximum likelihood estimators (STACY; MIHRAM, 1965; HAGER; BAIN, 1970). Numerical methods such as Newton-Rapshon are required to find the solution of the nonlinear system.

Under mild conditions that in some cases are not fulfill (PRENTICE, 1974), the MLE estimates are asymptotically normal distributed with a trivariate normal distribution given by,

$$\hat{\theta} \sim N_3[\theta, I^{-1}(\theta)] \text{ for } n \to \infty, \tag{5.6}$$

where $I(\theta)$ is the Fisher information matrix given by,

$$I(\alpha,\mu,\phi) = \begin{bmatrix} \dfrac{1+2\psi(\phi)+\phi\psi'(\phi)+\phi\psi(\phi)^2}{\alpha^2} & -\dfrac{1+\phi\psi(\phi)}{\mu} & -\dfrac{\psi(\phi)}{\alpha} \\ -\dfrac{1+\phi\psi(\phi)}{\mu} & \dfrac{\phi\alpha^2}{\mu^2} & \dfrac{\alpha}{\mu} \\ -\dfrac{\psi(\phi)}{\alpha} & \dfrac{\alpha}{\mu} & \psi'(\phi) \end{bmatrix}. \tag{5.7}$$

5.3 Bayesian Analysis

The joint posterior distribution for θ is given by the product of the likelihood function and the prior distribution $\pi(\theta)$ divided by a normalizing constant $d(x)$, resulting in

$$p(\theta|x) = \frac{\pi(\theta)}{d(x)}\frac{\alpha^n}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n} x_i^{\alpha\phi-1}\right\}\mu^{n\alpha\phi}\exp\left\{-\mu^{\alpha}\sum_{i=1}^{n} x_i^{\alpha}\right\}, \tag{5.8}$$

where

$$d(x) = \int_{\mathscr{A}} \pi(\theta)\frac{\alpha^n}{\Gamma(\phi)^n}\left\{\prod_{i=1}^{n} x_i^{\alpha\phi-1}\right\}\mu^{n\alpha\phi}\exp\left\{-\mu^{\alpha}\sum_{i=1}^{n} x_i^{\alpha}\right\}d\theta \tag{5.9}$$

and $\mathscr{A} = \{(0,\infty)\times(0,\infty)\times(0,\infty)\}$ is the parameter space of θ. Consider any prior in the form $\pi(\theta) \propto \pi(\mu)\pi(\alpha)\pi(\phi)$, the main aim is to find necessary and sufficient conditions for this class of posterior to be proper, i.e., $d(x) < \infty$.

Theorem 5.3.1. Suppose that $\pi(\alpha,\beta,\mu) < \infty$ for all $(\alpha,\beta,\mu) \in \mathbb{R}_+^3$, that $n \in \mathbb{N}^+$, and suppose that $\pi(\mu,\alpha,\phi) = \pi(\mu)\pi(\alpha)\pi(\mu)$ where

$$\pi(\mu) \underset{\mu\to 0^+}{\lesssim} \mu^{k_0}, \quad \pi(\mu) \underset{\alpha\to\infty}{\lesssim} \mu^{k_\infty}, \quad \pi(\alpha) \underset{\alpha\to 0^+}{\lesssim} \alpha^{q_0},$$

$$\pi(\alpha) \underset{\alpha\to\infty}{\lesssim} \alpha^{q_\infty}, \quad \pi(\phi) \underset{\phi\to 0^+}{\lesssim} \phi^{r_0} \quad \text{and} \quad \pi(\phi) \underset{\phi\to\infty}{\lesssim} \phi^{r_\infty},$$

such that $k_0 \geq -1$, $k_\infty \leq -1$, $q_\infty < r_0$, $2r_\infty + 1 < q_0$ and $n > -q_0$, then $p(\theta|x)$ is proper.

Proof. Since $\pi(\alpha)\alpha^n \frac{\pi(\phi)}{\Gamma(\phi)^n}\prod_{i=1}^n x_i^{\alpha\phi}\pi(\mu)\mu^{n\alpha\phi-1}\exp\left(-\mu^\alpha\sum_{i=1}^n x_i^\alpha\right) \geq 0$ always, by Tonelli's theorem we have:

$$\begin{aligned} d(x) &= \int_{\mathscr{A}} \pi(\alpha)\alpha^n \frac{\pi(\phi)}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n x_i^{\alpha\phi-1}\right\}\pi(\mu)\mu^{n\alpha\phi}\exp\left\{-\mu^\alpha\sum_{i=1}^n x_i^\alpha\right\}d\theta \\ &= \int_0^\infty\int_0^\infty\int_0^\infty \pi(\alpha)\alpha^n \frac{\pi(\phi)}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n x_i^{\alpha\phi-1}\right\}\pi(\mu)\mu^{n\alpha\phi}\exp\left\{-\mu^\alpha\sum_{i=1}^n x_i^\alpha\right\}d\mu\, d\phi\, d\alpha. \end{aligned}$$

Now, since $k_0 \geq -1$ and $k_\infty \leq -1$ by hypothesis we have that $\pi(\mu) \underset{\mu\to 0^+}{\lesssim} \mu^{-1}$, $\pi(\mu) \underset{\mu\to\infty}{\lesssim} \mu^{-1}$ and therefore $\pi(\mu) \lesssim \mu^{-1}$, then

$$\begin{aligned} d(x) &\lesssim \int_0^\infty\int_0^\infty\int_0^\infty \pi(\alpha)\alpha^n \frac{\pi(\phi)}{\Gamma(\phi)^n}\left(\prod_{i=1}^n x_i^\alpha\right)^\phi \mu^{n\alpha\phi-1}\exp\left\{-\mu^\alpha\sum_{i=1}^n x_i^\alpha\right\}d\mu\, d\phi\, d\alpha \\ &= \int_0^\infty\int_0^\infty \pi(\alpha)\alpha^{n-1}\frac{\pi(\phi)}{\Gamma(\phi)^n}\left(\prod_{i=1}^n x_i^\alpha\right)^\phi \frac{\Gamma(n\phi)}{\left(\sum_{i=1}^n x_i^\alpha\right)^{n\phi}}\, d\phi\, d\alpha \\ &= \int_0^\infty\int_0^\infty \pi(\alpha)\alpha^{n-1}\pi(\phi)\frac{\Gamma(n\phi)}{\Gamma(\phi)^n}e^{-n\,\mathrm{q}(\alpha)\phi}\, d\phi\, d\alpha \end{aligned}$$

where $\mathrm{q}(\alpha)$ is given in Proposition A.0.4. Therefore, from the proportionalities in Appendix A it follows that

$$\begin{aligned} d(x) &\lesssim \int_0^\infty\int_0^\infty \pi(\alpha)\alpha^{n-1}\pi(\phi)\frac{\Gamma(n\phi)}{\Gamma(\phi)^n}e^{-n\,\mathrm{q}(\alpha)\phi}\, d\phi\, d\alpha \\ &\propto \int_0^1\int_0^1 f(\alpha,\phi)\, d\phi\, d\alpha + \int_1^\infty\int_0^1 f(\alpha,\phi)\, d\phi\, d\alpha + \int_0^1\int_1^\infty g(\alpha,\phi)\, d\phi\, d\alpha + \int_1^\infty\int_1^\infty g(\alpha,\phi)\, d\phi\, d\alpha \\ &= s_1 + s_2 + s_3 + s_4, \end{aligned} \tag{5.10}$$

where $f(\alpha,\phi) = \pi(\alpha)\alpha^{n-1}\pi(\phi)\phi^{n-1}e^{-n\,\mathrm{q}(\alpha)\phi}$, $g(\alpha,\phi) = \pi(\alpha)\alpha^{n-1}\pi(\phi)\phi^{\frac{n-1}{2}}e^{-n\,\mathrm{p}(\alpha)\phi}$ and s_1, s_2, s_3 and s_4 denote the respective four real numbers in the sum that precedes it. It follows that $d(x) < \infty$, if and only if $s_1 < \infty$, $s_2 < \infty$, $s_3 < \infty$ and $s_4 < \infty$. Now, it follows that

$$\begin{aligned} s_1 &\lesssim \int_0^1 \alpha^{q_0+n-1}\int_0^1 \phi^{n+r_0-1}e^{-n\,\mathrm{q}(\alpha)\phi}\, d\phi\, d\alpha \\ &= \int_0^1 \alpha^{q_0+n-1}\frac{\gamma(n+r_0, n\,\mathrm{q}(\alpha))}{(n\,\mathrm{q}(\alpha))^{n+r_0}}\, d\alpha \propto \int_0^1 \alpha^{q_0+n-1}\, d\alpha < \infty, \end{aligned}$$

$$
\begin{aligned}
s_2 &\lesssim \int_1^{\infty} \alpha^{q_\infty+n-1} \int_0^1 \phi^{n+r_0-1} e^{-n\mathrm{q}(\alpha)\phi}\, d\phi\, d\alpha \\
&= \int_1^{\infty} \alpha^{q_\infty+n-1} \frac{\gamma(n+r_0, n\mathrm{q}(\alpha))}{(n\mathrm{q}(\alpha))^{n+r_0}}\, d\alpha \propto \int_1^{\infty} \alpha^{q_\infty - r_0 - 1}\, d\alpha < \infty,
\end{aligned}
$$

$$
\begin{aligned}
s_3 &\lesssim \int_0^1 \alpha^{q_0+n-1} \int_1^{\infty} \phi^{\frac{n+1+2r_\infty}{2}-1} e^{-n\mathrm{p}(\alpha)\phi}\, d\phi\, d\alpha \\
&= \int_0^1 \alpha^{q_0+n-1} \frac{\Gamma(\frac{n+1+2r_\infty}{2}, n\mathrm{p}(\alpha))}{(n\mathrm{p}(\alpha))^{\frac{n+1+2r_\infty}{2}}}\, d\alpha \propto \int_0^1 \alpha^{(q_0-2r_\infty-1)-1}\, d\alpha < \infty, \text{ and}
\end{aligned}
$$

$$
\begin{aligned}
s_4 &\lesssim \int_1^{\infty} \alpha^{q_\infty+n-1} \int_1^{\infty} \phi^{\frac{n+1+2r_\infty}{2}-1} e^{-n\mathrm{p}(\alpha)\phi}\, d\phi\, d\alpha \\
&= \int_1^{\infty} \alpha^{q_\infty+n-1} \frac{\Gamma(\frac{n+1+2r_\infty}{2}, n\mathrm{p}(\alpha))}{(n\mathrm{p}(\alpha))^{\frac{n+1+2r_\infty}{2}}}\, d\alpha \propto \int_1^{\infty} \alpha^{q_\infty+n-2} e^{-nk\alpha}\, d\alpha < \infty,
\end{aligned}
$$

where in the last line $k \in \mathbb{R}^+$ is given in Proposition A.0.6. Therefore, from $s_i < \infty, i = 1, \ldots, 4$, we have that $d = s_1 + s_2 + s_3 + s_4 < \infty$. □

Theorem 5.3.2. Suppose that $\pi(\alpha, \beta, \mu) > 0\ \forall(\alpha, \beta, \mu) \in \mathbb{R}^3_+$ and that $n \in \mathbb{N}^+$, then the following items are valid

i) $\pi(\mu, \alpha, \beta) \gtrsim \pi(\mu)\pi(\alpha)\pi(\beta)$ for all $\beta \in [b_0, b_1]$ where $0 \leq b_0 < b_1$ and one of the following hold

  - $\pi(\mu) \underset{\mu \to 0^+}{\gtrsim} \mu^{k_0}$ where $k_0 < -1$, or
  - $\pi(\mu) \underset{\mu \to \infty}{\gtrsim} \mu^{k_\infty}$ and $\pi(\alpha) \underset{\alpha \to 0^+}{\gtrsim} \alpha^{q_0}$, where $k_\infty > -1$ and $q_0 \in \mathbb{R}$,

  then $p(\theta|x)$ is improper.

ii) $\pi(\mu, \alpha, \beta) \gtrsim \pi(\mu)\pi(\alpha)\pi(\beta)$ in which

$$\pi(\mu) \underset{\mu \to 0^+}{\gtrsim} \mu^{k_0} \text{ and } \pi(\mu) \underset{\mu \to \infty}{\gtrsim} \mu^{k_\infty},$$

  where $k_0 \geq -1$ and $k_\infty \leq -1$, and one of the following occur

  - $\pi(\phi) \underset{\phi \to 0^+}{\gtrsim} \phi^{r_0}$ and $\pi(\alpha) \underset{\alpha \to \infty}{\gtrsim} \alpha^{q_\infty}$ where $q_\infty \geq r_0$, or
  - $\pi(\alpha) \underset{\alpha \to 0^+}{\gtrsim} \alpha^{q_0}$ and $\pi(\phi) \underset{\phi \to \infty}{\gtrsim} \phi^{r_\infty}$ where $2r_\infty + 1 \geq q_0$, or $n \leq -q_0$,

  then $p(\theta|x)$ is improper.

Proof. Suppose that hypothesis of item *i*) hold.

First suppose that $\pi(\mu) \underset{\mu\to 0^+}{\gtrsim} \mu^{k_\infty}$ with $k_0 < -1$. For $h = \sqrt{\frac{-k_0-1}{2n}} > 0$, fixing $0 < \alpha \leq h$ and $0 < \phi \leq h$ we have that $n\alpha\phi + (k_0+1) - 1 \leq nh^2 + (k_0+1) - 1 = \frac{(k_0+1)}{2} - 1 < -1$. Moreover, for every $\alpha > 0$ fixed we have that $\exp\left\{-\mu^\alpha \sum_{i=1}^n x_i^\alpha\right\} \underset{\mu\to 0^+}{\propto} 1$, hence, from Proposition 2.5.9 we have that

$$\int_0^\infty \pi(\mu)\mu^{n\alpha\phi} \exp\left\{-\mu^\alpha \sum_{i=1}^n x_i^\alpha\right\} d\mu \gtrsim \int_0^1 \mu^{n\alpha\phi+(k_0+1)-1} = \infty,$$

for all $\alpha \in (0,h]$ and $\phi \in (0,h]$. Therefore

$$\begin{aligned} d(x) &\gtrsim \int_{h/2}^{h}\int_{h/2}^{h} \pi(\alpha)\alpha^n \frac{\pi(\phi)}{\Gamma(\phi)^n}\left(\prod_{i=1}^n x_i^\alpha\right)^\phi \int_0^\infty \mu^{n\alpha\phi+(k+1)-1} \exp\left\{-\mu^\alpha \sum_{i=1}^n x_i^\alpha\right\} d\mu\, d\phi\, d\alpha \\ &\propto \int_{h/2}^{h}\int_{h/2}^{h} \infty \, d\phi\; d\alpha = \infty, \end{aligned}$$

that is, $d(x) = \infty$. Now suppose that $\pi(\mu) \underset{\mu\to\infty}{\gtrsim} \mu^{k_\infty}$ and $\pi(\alpha) \underset{\alpha\to 0^+}{\gtrsim} \alpha^{q_0}$, where $k_\infty > -1$ and $q_0 \in \mathbb{R}$. Then, from the proportionalities in Appendix A we have that

$$\begin{aligned} d(x) &\propto \int_{\mathscr{A}} \pi(\alpha)\alpha^n \frac{\pi(\phi)}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n x_i^{\alpha\phi-1}\right\}\pi(\mu)\mu^{n\alpha\phi}\exp\left\{-\mu^\alpha\sum_{i=1}^n x_i^\alpha\right\} d\theta \\ &= \int_0^\infty\int_0^\infty\int_0^\infty \pi(\alpha)\alpha^n \frac{\pi(\phi)}{\Gamma(\phi)^n}\left\{\prod_{i=1}^n x_i^{\alpha\phi-1}\right\}\pi(\mu)\mu^{n\alpha\phi}\exp\left\{-\mu^\alpha\sum_{i=1}^n x_i^\alpha\right\} d\mu\, d\phi\, d\alpha \\ &\gtrsim \int_0^\infty\int_0^\infty\int_1^\infty \alpha^{n+q_0}\frac{\pi(\phi)}{\Gamma(\phi)^n}\left(\prod_{i=1}^n x_i^\alpha\right)^\phi \mu^{n\alpha\phi+(k_\infty+1)-1}\exp\left\{-\mu^\alpha\sum_{i=1}^n x_i^\alpha\right\} d\mu\, d\phi\, d\alpha \\ &= \int_0^\infty\int_0^\infty\int_{\sum_{i=1}^n x_i^\alpha}^\infty \alpha^{n+q_0}\frac{\pi(\phi)}{\Gamma(\phi)^n}\frac{\left(\prod_{i=1}^n x_i^\alpha\right)^\phi}{\left(\sum_{i=1}^n x_i^\alpha\right)^{n\phi+\frac{k_\infty+1}{\alpha}}} u^{n\phi+\frac{k_\infty+1}{\alpha}-1}e^{-u}\, du\, d\phi\, d\alpha \\ &\geq \int_0^\infty\int_1^\infty \frac{\pi(\phi)}{\Gamma(\phi)^n} n^{-n\phi}u^{n\phi-1}e^{-u}\int_0^\infty \alpha^{n+q_0}e^{-\mathrm{p}(\alpha)\left(n\phi+\frac{k_\infty+1}{\alpha}\right)+(\log u-\log n)\frac{k+1}{\alpha}}\, d\alpha\, du\, d\phi \end{aligned}$$

where in the above we used the change of variables $u = \mu^\alpha \sum_{i=1}^n x_i^\alpha$ in the integral, in the last inequality we used the fact that $\sum_{i=1}^n x_i^\alpha \geq 1$ for $\alpha \geq 0$, and $\mathrm{p}(\alpha)$ is given as in Proposition A.0.3. Now, since $\mathrm{p}(\alpha) \underset{\alpha\to 0^+}{\propto} \alpha^2$ from Proposition A.0.3 it follows due to the Proposition 2.5.8 that $\mathrm{p}(\alpha) \propto \alpha^2$ for $\alpha \in [0,1]$ and therefore $\lim_{\alpha\to 0^+} e^{-\mathrm{p}(\alpha)\left(n\phi+\frac{k_\infty+1}{\alpha}\right)} = \lim_{\alpha\to 0^+} e^{-\frac{p(\alpha)}{\alpha^2}(n\phi\alpha+k_\infty+1)\alpha} = e^0 = 1$. Thus, since $n \geq 1$ and $\log u - \log n > 0$ for $u \geq 3n > e\cdot n$, and since $\int_0^1 \alpha^H e^{\frac{L}{\alpha}} = \infty$ for every $H \in \mathbb{R}$ and $L \in R^+$ (which can be easily checked via the change of variable $\beta = \frac{1}{\alpha}$ in the integral), it follows that

$$d(x) \gtrsim \int_0^\infty \int_{3n}^\infty \pi(\phi) \frac{1}{\Gamma(\phi)^n} n^{-n\phi} u^{n\phi-1} e^{-u} \int_0^1 \alpha^{n+q_0} e^{(\log u - \log n)\frac{k+1}{\alpha}} \, d\alpha \, du \, d\phi$$
$$= \int_0^\infty \int_{3n}^\infty \infty \, du \, d\phi = \infty,$$

and therefore $d(x) = \infty$.

Now suppose the hypothesys of ii) hold. First suppose that

$$\pi(\phi) \underset{\phi \to 0^+}{\gtrsim} \phi^{r_0} \quad \text{and } \pi(\alpha) \underset{\alpha \to \infty}{\gtrsim} \alpha^{q_\infty}$$

where $q_\infty \geq r_0$. Then, following the same steps that resulted in (5.10) and the same expressions for s_i, where $i = 1, \cdots, 4$, we have that $d(x) \gtrsim s_1 + s_2 + s_3 + s_4$ where

$$s_2 \gtrsim \int_1^\infty \alpha^{q_\infty + n - 1} \int_0^1 \phi^{n + r_0 - 1} e^{-n\mathrm{q}(\alpha)\phi} \, d\phi \, d\alpha$$
$$= \int_1^\infty \alpha^{q_\infty + n - 1} \frac{\gamma(n + r_0, n\mathrm{q}(\alpha))}{(n\mathrm{q}(\alpha))^{n + r_0}} \, d\alpha \propto \int_1^\infty \alpha^{q_\infty - r_0 - 1} \, d\alpha = \infty$$

and therefore $d(x) = \infty$.

Now suppose that

$$\pi(\alpha) \underset{\alpha \to 0^+}{\gtrsim} \alpha^{q_0} \quad \text{and} \quad \pi(\phi) \underset{\phi \to \infty}{\gtrsim} \phi^{r_\infty}$$

where $2r_\infty + 1 \geq q_0$ or $n \leq -q_0$. Then, following the same steps that resulted in (5.10) and the same expressions for s_i, where $i = 1, \ldots, 4$, we have that $d(x) \gtrsim s_1 + s_2 + s_3 + s_4$ where

$$s_3 \gtrsim \int_0^1 \alpha^{q_0 + n - 1} \int_1^\infty \phi^{\frac{n + 1 + 2r_\infty}{2} - 1} e^{-n\mathrm{p}(\alpha)\phi} \, d\phi \, d\alpha$$
$$= \int_0^1 \alpha^{q_0 + n - 1} \frac{\Gamma(\frac{n+1+2r_\infty}{2}, n\mathrm{p}(\alpha))}{(n\mathrm{p}(\alpha))^{\frac{n+1+2r_\infty}{2}}} \, d\alpha \propto \int_0^1 \alpha^{(q_0 - 2r_\infty - 1) - 1} \, d\alpha = \infty$$

which implies $d(x) = \infty$. □

Theorem 5.3.3. Suppose that $0 < \pi(\alpha, \beta, \mu) < \infty$ for all $(\alpha, \beta, \mu) \in \mathbb{R}_+^3$, and suppose that $\pi(\mu, \alpha, \phi) = \pi(\mu)\pi(\alpha)\pi(\phi)$ where

$$\pi(\mu) \underset{\mu \to 0^+}{\propto} \mu^{k_0}, \quad \pi(\mu) \underset{\alpha \to \infty}{\propto} \mu^{k_\infty}, \quad \pi(\alpha) \underset{\alpha \to 0^+}{\propto} \alpha^{q_0},$$
$$\pi(\alpha) \underset{\alpha \to \infty}{\propto} \alpha^{q_\infty}, \quad \pi(\phi) \underset{\phi \to 0^+}{\propto} \phi^{r_0} \quad \text{and} \quad \pi(\phi) \underset{\phi \to \infty}{\propto} \phi^{r_\infty},$$

then $\alpha^q \phi^r \mu^k \pi(\alpha, \phi, \mu)$ leads to a proper posterior if and only if $-1 - k_0 \leq k \leq -1 - k_\infty$, $2r + (2r_\infty + 1 - q_0) < q < r + (r_0 - q_\infty)$ and $n \geq -q_0$.

Proof. By Theorems 5.3.2 and 5.3.3 we have that $\alpha^q\beta^r\mu^k\pi(\alpha,\beta,\mu)$ leads to a proper posterior if and only if $k+k_0\leq -1$, $k+k_\infty\geq -1$, $q+q_\infty<r+r_0$, $2(r+r_\infty)\leq q+q_0$ and $n>-q_0$. Combining theses inequalities the proof is completed. □

5.4 Objective Priors

5.4.1 Some common objective priors

A naive approach to obtain objective priors is to consider uniform priors contained in the interval $(0,\infty)$. However, uniform priors are usually not attractive due to its lack of invariance over reparametrizations. The uniform prior for GG distribution is given by $\pi_1(\phi,\mu,\alpha)\propto 1$.

Corollary 5.4.1. The posterior distribution obtained using a joint uniform prior is improper for all $n\in\mathbb{N}^+$.

Proof. Since $\pi_1(\phi,\mu,\alpha)=\mu^0\alpha^0\phi^0$ we apply Theorem 5.3.2 ii) with $k_0=k_\infty=q_\infty=r_0=0$ and since $q_\infty\geq r_0$ we have that $\pi(\alpha,\beta,\mu)$ leads to an improper posterior for all $n\in\mathbb{N}^+$. □

Another common approach was suggested by Jeffreys' that considered different procedures for constructing objective priors. As the parameters of the GG distribution are contained in the interval $(0,\infty)$, the prior using Jeffreys' rule is $\pi_2(\phi,\mu,\alpha)\propto(\phi\mu\alpha)^{-1}$.

Corollary 5.4.2. The posterior distribution obtained using Jeffreys' rule is improper for all $n\in\mathbb{N}^+$.

Proof. Since $\pi_1(\phi,\mu,\alpha)=\mu^{-1}\alpha^{-1}\phi^{-1}$ we can apply Theorem 5.3.2 ii) with $k_0=k_\infty=q_\infty=r_0=-1$, where $q_\infty\geq r_0$, and therefore we have that $\pi(\alpha,\beta,\mu)$ leads to an improper posterior for all $n\in\mathbb{N}^+$. □

Zellner (1984) discussed another procedure to obtain an objective prior, the MDI prior for the GG distribution is given by

$$\pi_3(\theta)\propto\frac{\alpha\mu}{\Gamma(\phi)}\exp\left\{\psi(\phi)\left(\phi-\frac{1}{\alpha}\right)-\phi\right\}. \tag{5.11}$$

Corollary 5.4.3. The joint posterior density using the MDI prior (5.11) is improper for any $n\in\mathbb{N}^+$.

Proof. Since $\psi(\phi)<0$ for all $\phi\in(0,1]$ Abramowitz and Stegun (1972), we have that $\exp\left(-\psi(\phi)\frac{1}{\alpha}\right)\geq$ 1 for all $\phi\in[0.5,1]$ and therefore

$$\pi_3(\theta)\gtrsim\alpha\mu\frac{\exp\left(\psi(\phi)\phi-\phi\right)}{\Gamma(\phi)}.$$

in the interval $[0.5, 1]$. It follows that the hypothesis in Theorem 5.3.2 i) is satisfied with $b_0 = 0.5$, $b_1 = 1$, $k_\infty = 1 > -1$ and $q_0 = 1$, and therefore we have that $\pi_3(\theta)$ leads to an improper posterior for all $n \in \mathbb{N}^+$. □

5.4.2 Priors based on the Fisher information matrix

One important objective prior is based on Jeffreys' general rule (JEFFREYS, 1946) and known as Jeffreys' prior. Noortwijk (2001) provided the Jeffreys prior for the GG distribution, which is given by

$$\pi_4(\theta) \propto \frac{\sqrt{\phi^2\psi'(\phi)^2 - \psi'(\phi) - 1}}{\mu}. \tag{5.12}$$

Corollary 5.4.4. The posterior distribution using the Jeffreys prior (5.12) is improper for all $n \in \mathbb{N}^+$.

Proof. From Proposition A.0.1, we have that

$$\sqrt{\phi^2\psi'(\phi)^2 - \psi'(\phi) - 1} \underset{\phi \to 0^+}{\propto} 1 \text{ and } \sqrt{\phi^2\psi'(\phi)^2 - \psi'(\phi) - 1} \underset{\phi \to \infty}{\propto} \frac{1}{\phi}. \tag{5.13}$$

Since $\pi_4(\phi) \propto \sqrt{\phi^2\psi'(\phi)^2 - \psi'(\phi) - 1} \underset{\phi \to 0^+}{\propto} 1$, the hypotheses of Theorem 5.3.2 ii) hold with $k_0 = k_\infty = -1$ and $r_0 = q_\infty = 0$, where $q_\infty \geq r_0$, and therefore $\pi_4(\theta)$ leads to an improper posterior for all $n \in \mathbb{N}^+$. □

Fonseca, Ferreira and Migon (2008) considered the scenario where the Jeffreys prior has an independent structure, i.e., the prior has the form $\pi_{J2}(\theta) \propto \sqrt{|\operatorname{diag} I(\theta)|}$, where $\operatorname{diag} I(\cdot)$ is the diagonal matrix of $I(\cdot)$. For the GG distribution the prior is given by

$$\pi_5(\theta) \propto \frac{\sqrt{\phi\psi'(\phi)\left(1 + 2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2\right)}}{\mu}. \tag{5.14}$$

Corollary 5.4.5. The posterior distribution using the independent Jeffreys' prior (5.14) is improper for all $n \in \mathbb{N}^+$.

Proof. By Abramowitz and Stegun (1972), we have the recurrence relations

$$\psi(\phi) = -\frac{1}{\phi} + \psi(\phi+1) \quad \text{and} \quad \psi'(\phi) = \frac{1}{\phi^2} + \psi'(\phi+1). \tag{5.15}$$

It follows that

$$\begin{aligned}
&2\psi(\phi) + \phi\psi'(\phi) + \phi\psi(\phi)^2 + 1 = \\
&2\left(-\frac{1}{\phi} + \psi(\phi+1)\right) + \phi\left(\frac{1}{\phi^2} + \psi'(\phi+1)\right) + \phi\left(\frac{1}{\phi^2} - \frac{2}{\phi}\psi(\phi+1) + \psi(\phi+1)^2\right) + 1 = \\
&1 + \phi\left(\psi(\phi+1)^2 + \psi'(\phi+1)\right).
\end{aligned}$$

Hence, $2\psi(\phi)+\phi\psi'(\phi)+\phi\psi(\phi)^2+1 \underset{\phi\to 0^+}{\propto} 1$, which implies that

$$\pi_5(\phi) \propto \sqrt{\phi\psi'(\phi)\left(1+2\psi(\phi)+\phi\psi'(\phi)+\phi\psi(\phi)^2\right)} \underset{\phi\to 0^+}{\propto} \frac{1}{\sqrt{\phi}}, \tag{5.16}$$

then, Theorem 5.3.2 ii) can be applied with $k_0=k_\infty=-1$, $r_0=-\frac{1}{2}$ and $q_\infty=0$ where $q_\infty \geq r_0$ and therefore $\pi_5(\theta)$ leads to an improper posterior. $\square$

This approach can be further extended considering that only one parameter is independent. For instance, let (θ_1,θ_2) be dependent parameters and θ_3 be independent then under the partition the $((\theta_1,\theta_2),\theta_3)$-Jeffreys prior is given by

$$\pi(\theta) \propto \sqrt{\left(I_{11}(\theta)I_{22}(\theta)-I_{12}^2(\theta)\right)I_{33}(\theta)}. \tag{5.17}$$

For the GG distribution the partition $((\phi,\mu),\alpha)$-Jeffreys' prior is given by

$$\pi_6(\theta) \propto \frac{\sqrt{(\phi\psi'(\phi)-1)\left(1+2\psi(\phi)+\phi\psi'(\phi)+\phi\psi(\phi)^2\right)}}{\mu}. \tag{5.18}$$

Corollary 5.4.6. The posterior distribution using the $((\phi,\mu),\alpha)$-Jeffreys' prior is improper for all $n\in\mathbb{N}^+$.

Proof. From the recurrence relations (5.15) we have that

$$\phi\psi'(\phi)-1=\frac{1}{\phi}+\phi\psi'(\phi+1)-1 \Rightarrow \phi\psi'(\phi)-1 \underset{\phi\to 0^+}{\propto} \frac{1}{\phi}. \tag{5.19}$$

Together with the relation (5.16) this implies that

$$\pi_6(\phi) \propto \sqrt{(\phi\psi'(\phi)-1)\left(1+2\psi(\phi)+\phi\psi'(\phi)+\phi\psi(\phi)^2\right)} \underset{\phi\to 0^+}{\propto} \frac{1}{\sqrt{\phi}}. \tag{5.20}$$

Therefore, Theorem 5.3.2 ii) can be applied with $k_0=k_\infty=-1$, $r_0=-\frac{1}{2}$ and $q_\infty=0$ where $q_\infty \geq r_0$ and therefore $\pi_6(\theta)$ leads to an improper posterior. $\square$

On the other hand, the $((\phi,\alpha),\mu)$-Jeffreys prior is given by

$$\pi_7(\theta) \propto \frac{\sqrt{\phi\psi'(\phi)\left(1+2\psi(\phi)+\phi\psi'(\phi)+\phi\psi(\phi)^2\right)-\phi\psi(\phi)^2}}{\mu}. \tag{5.21}$$

Corollary 5.4.7. The posterior distribution using the independent Jeffreys' prior (5.21) is improper for all $n\in\mathbb{N}^+$.

Proof. From (5.15) we have that

$$
\begin{aligned}
\pi_7^{\frac{1}{2}}(\phi) \propto\; & \phi\psi^{'}(\phi)\left(1+2\psi(\phi)+\phi\psi^{'}(\phi)+\phi\psi(\phi)^2\right)-\phi\psi(\phi)^2 \\
& = \left(\phi^{-1}+\phi\psi'(\phi+1)\right)\left(1+\phi\left(\psi(\phi+1)^2+\psi'(\phi+1)\right)\right)-\phi\left(-\phi^{-1}+\psi(\phi+1)\right)^2 \\
& = \phi\left(\psi'(\phi+1)-\psi(\phi+1)^2+\phi\psi'(\phi+1)(\psi(\phi+1)^2+\psi'(\phi+1))\right)+\psi(\phi+1)^2 \\
& \quad +2\psi(\phi+1)+\psi'(\phi+1) \\
& \propto \phi\psi^{'}(\phi)\left(1+2\psi(\phi)+\phi\psi^{'}(\phi)+\phi\psi(\phi)^2\right)-\phi\psi(\phi)^2 \\
& \underset{\phi\to 0^+}{\propto} \psi(1)^2+2\psi(1)+\psi'(1)=\gamma^2-2\gamma+\frac{\pi}{6}>0,
\end{aligned}
$$

then, Theorem 5.3.2 ii) can be applied with $k_0 = k_\infty = -1$, $r_0 = 0$ and $q_\infty = 0$ where $q_\infty \geq r_0$ and therefore $\pi_7(\theta)$ leads to an improper posterior. □

Finally, the $((\alpha,\mu),\phi)$-Jeffreys prior is

$$
\pi_8(\theta) \propto \frac{\sqrt{\psi'(\phi)(\phi^2\psi'(\phi)+\phi-1)}}{\mu}. \tag{5.22}
$$

Corollary 5.4.8. The posterior distribution using the independent Jeffreys' prior (5.22) is improper for all $n \in \mathbb{N}^+$.

Proof. From the recurrence relations (5.15) we have that

$$
\phi^2\psi^{'}(\phi)+\phi-1=\phi\left(1+\phi\psi^{'}(\phi+1)\right) \Rightarrow \phi^2\psi^{'}(\phi)+\phi-1 \underset{\phi\to 0^+}{\propto} \phi \tag{5.23}
$$

as $\psi'(\phi) \propto \frac{1}{\phi^2}$ it follows that

$$
\pi_8(\phi) \propto \sqrt{\psi'(\phi)(\phi^2\psi'(\phi)+\phi-1)} \underset{\phi\to 0^+}{\propto} \frac{1}{\sqrt{\phi}},
$$

and Theorem 5.3.2 ii) can be applied with $k_0 = k_\infty = -1$, $r_0 = -\frac{1}{2}$ and $q_\infty = 0$ where $q_\infty \geq r_0$. Therefore $\pi_8(\theta)$ leads to an improper posterior. □

5.4.3 Reference priors

Let (α,ϕ,μ) be the ordered parameters of interest, then conditional priors of the (α,ϕ,μ)-reference prior are given by

$$
\pi(\alpha) \propto \frac{1}{\alpha}, \quad \pi(\phi|\alpha) \propto \sqrt{\frac{\phi\psi'(\phi)-1}{\phi}}, \quad \pi(\mu|\alpha,\phi) \propto \frac{1}{\mu}.
$$

Therefore, (α,ϕ,μ)-reference prior is given by

$$
\pi_9(\theta) \propto \frac{1}{\alpha\mu}\sqrt{\frac{\phi\psi'(\phi)-1}{\phi}}.
$$

Corollary 5.4.9. The posterior density using the (α,ϕ,μ)-reference prior is improper for all $n\in\mathbb{N}^+$.

Proof. By equation (5.19) we have that

$$\pi_9(\phi)\propto\sqrt{\frac{\phi\psi'(\phi)-1}{\phi}}\underset{\phi\to 0^+}{\propto}\frac{1}{\phi},$$

therefore, item (ii) of Theorem 5.3.2 can be applied with $k_0=k_\infty=r_0=q_\infty=-1$ where $q_\infty\geq r_0$ which implies that $\pi_9(\theta)$ leads to an improper posterior.

□

On the other hand, if (α,μ,ϕ) are the ordered parameters, then the conditional reference priors are

$$\pi(\alpha)\propto\frac{1}{\alpha},\quad \pi(\mu|\alpha)\propto\frac{1}{\mu},\quad \pi(\phi|\alpha,\mu)\propto\sqrt{\psi'(\phi)},$$

and the (α,μ,ϕ)-reference prior is

$$\pi_{10}(\theta)\propto\frac{\sqrt{\psi'(\phi)}}{\mu\alpha}.$$

Corollary 5.4.10. The posterior density using the (α,μ,ϕ)-reference prior is improper for all $n\in\mathbb{N}^+$.

Proof. By the equation (5.15) we have that $\sqrt{\psi'(\phi)}\underset{\phi\to 0^+}{\propto}\frac{1}{\phi}$. Thus, as in Corollary 5.4.9 we have that $\pi_{10}(\theta)$ leads to an improper posterior for all $n\in\mathbb{N}^+$. □

In the case of (μ,ϕ,α) be the vector of ordered parameters, we have that the conditional priors are

$$\pi(\mu)\propto\frac{1}{\mu},\quad \pi(\phi|\mu)\propto\sqrt{\psi'(\phi)-\frac{\psi(\phi)^2}{2\psi(\phi)+\phi\psi'(\phi)+\phi\psi(\phi^2)+1}},\quad \pi(\alpha|\phi,\mu)\propto\frac{1}{\alpha}.$$

and the (μ,ϕ,α)-reference prior is given by

$$\pi_{11}(\theta)\propto\frac{1}{\mu\alpha}\sqrt{\psi'(\phi)-\frac{\psi(\phi)^2}{2\psi(\phi)+\phi\psi'(\phi)+\phi\psi(\phi^2)+1}}.$$

Corollary 5.4.11. The posterior density using the (μ,ϕ,α)-reference prior is improper for all $n\in\mathbb{N}^+$.

Proof. From Abramowitz and Stegun (1972), we have

$$\psi(\phi)=\log(\phi)-\frac{1}{2\phi}-\frac{1}{12\phi^2}+o\left(\frac{1}{\phi^2}\right)\quad\text{and}\quad\psi'(\phi)=\frac{1}{\phi}+\frac{1}{2\phi^2}+o\left(\frac{1}{\phi^2}\right),\tag{5.24}$$

where it follows directly that

$$\psi(\phi)^2 = \log(\phi)^2 - \frac{\log(\phi)}{\phi} + o\left(\frac{1}{\phi}\right).$$

Therefore $2\psi(\phi)+\phi\psi'(\phi)+\phi\psi(\phi)^2+1 = \phi\log(\phi)^2+\log(\phi)+2+o(1)$ and

$$\begin{aligned}
\pi_{11}(\phi) &\propto \sqrt{\psi'(\phi) - \frac{\psi(\phi)^2}{2\psi(\phi)+\phi\psi'(\phi)+\phi\psi(\phi)^2+1}} \\
&= \sqrt{\frac{\left(\frac{1}{\phi}+\frac{1}{2\phi^2}+o\left(\frac{1}{\phi^2}\right)\right)\left(\phi\log(\phi)^2+\log(\phi)+2+o(1)\right)-\log(\phi)^2+\frac{\log(\phi)}{\phi}+o\left(\frac{1}{\phi}\right)}{\phi\log(\phi)^2+\log(\phi)+2+o(1)}} \\
&= \sqrt{\frac{\frac{1}{\phi}\left(\log(\phi)^2+o(\log(\phi)^2)\right)}{\phi\left(\log(\phi)^2+o(\log(\phi)^2)\right)}} = \frac{1}{\phi}\sqrt{\frac{1+o(1)}{1+o(1)}}.
\end{aligned}$$

Thus

$$\pi_{11}(\phi) \propto \sqrt{\psi'(\phi) - \frac{\psi(\phi)^2}{2\psi(\phi)+\phi\psi'(\phi)+\phi\psi(\phi)^2+1}} \underset{\phi\to 0^+}{\propto} \frac{1}{\phi},$$

and therefore Theorem 5.3.2 ii) can be applied with $k_0 = k_\infty = q_0 = r_\infty = -1$ where $2r_\infty+1 \geq q_0$. Therefore $\pi_{11}(\theta)$ leads to an improper posterior.

□

If (μ,α,ϕ) are the ordered parameters then the conditional priors are given by

$$\pi(\mu) \propto \frac{1}{\mu}, \quad \pi(\alpha|\mu) \propto \frac{1}{\alpha}, \quad \pi(\phi|\alpha,\mu) \propto \sqrt{\psi'(\phi)}$$

and the joint (μ,α,ϕ)-reference prior has the same form of $\pi_{10}(\theta)$ and its posterior is improper from Corollary 5.4.10.

Finally, let (ϕ,α,μ) be the ordered parameters, then the conditional priors are

$$\pi(\phi) \propto \sqrt{\frac{\phi^2\psi'(\phi)^2-\psi'(\phi)-1}{\phi^2\psi'(\phi)+\phi-1}}, \quad \pi(\alpha|\phi) \propto \frac{1}{\alpha}, \quad \pi(\mu|\alpha,\phi) \propto \frac{1}{\mu}$$

and the (ϕ,α,μ)-reference prior is given by

$$\pi_{12}(\theta) \propto \frac{1}{\mu\alpha}\sqrt{\frac{\phi^2\psi'(\phi)^2-\psi'(\phi)-1}{\phi^2\psi'(\phi)+\phi-1}}. \tag{5.25}$$

It is important to point out that (ϕ,μ,α)-reference prior is the same as the (ϕ,α,μ)-reference prior, which completes all possible reference priors obtained from Proposition 2.3.1.

Corollary 5.4.12. The posterior distribution using the (ϕ,α,μ)-reference prior (5.25) is proper for $n \geq 2$ and its higher moments are improper for all $n \in \mathbb{N}^+$.

Proof. From (5.13) and by the asymptotic relations (5.24) we have that

$$\phi^2\psi'(\phi)+\phi-1 = 2\phi-\frac{1}{2}+o(1) \underset{\phi\to\infty}{\propto} \phi$$

which together with equation (5.23) implies that

$$\sqrt{\phi^2\psi'(\phi)+\phi-1} \underset{\phi\to 0^+}{\propto} \sqrt{\phi} \text{ and } \sqrt{\phi^2\psi'(\phi)+\phi-1} \underset{\phi\to\infty}{\propto} \sqrt{\phi}.$$

Hence, from the above proportionalities we have that

$$\sqrt{\frac{\phi^2\psi'(\phi)^2-\psi'(\phi)-1}{\phi^2\psi'(\phi)+\phi-1}} \underset{\phi\to 0^+}{\propto} \frac{1}{\sqrt{\phi}} \text{ and } \sqrt{\frac{\phi^2\psi'(\phi)^2-\psi'(\phi)-1}{\phi^2\psi'(\phi)+\phi-1}} \underset{\phi\to\infty}{\propto} \frac{1}{\sqrt{\phi^3}}.$$

Therefore, Theorem 5.3.1 can be applied with $k_0 = k_\infty = q_0 = q_\infty = -1$, $r_0 = -\frac{1}{2}$ and $r_\infty = -\frac{3}{2}$ where $k_0 \geq -1$, $k_\infty \leq -1$, $q_\infty < r_0$ and $2r_\infty + 1 < q_0$, and therefore $\pi_{12}(\theta)$ leads to an proper posterior for every $n > -q_0 = 1$.

In order to prove that the higher moments are improper suppose $\alpha^q\phi^r\mu^k\pi(\theta)$ leads to a proper posterior for $r\in\mathbb{N}$, $q\in\mathbb{N}$ and $r\in\mathbb{N}$. By Theorem 5.3.3 we have $k+k_0\leq -1$, $k+k_\infty \geq -1$, $q+q_\infty < r+r_0$, $2(r+r_\infty)\leq q+q_0$ and $n\geq -q_0$, i.e., $k=0$ and $2r-1<q<r+\frac{1}{2}$. The inequality $2r-1<r+\frac{1}{2}$ leads to $r<\frac{3}{2}$, i.e., $r=0$ or $r=1$. By the previous inequality, the case where $r=0$ leads to $-1<q<\frac{1}{2}$, that is, $q=0$. Now, for $r=1$ we have the inequality $1<q<\frac{3}{2}$ which do not have integer solution. Therefore, the only possible solution is $q=r=k=0$ which implies that the higher moments are improper. □

Due to the consistent marginalization property of the reference prior the reference marginal posterior distribution of ϕ and α is

$$p_{12}(\phi,\alpha|x) \propto \alpha^{n-2}\frac{\Gamma(n\phi)}{\Gamma(\phi)^n}\sqrt{\frac{\phi^2\psi'(\phi)^2-\psi'(\phi)-1}{\phi^2\psi'(\phi)+\phi-1}}\left(\frac{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}{\sum_{i=1}^n t_i^\alpha}\right)^{n\phi},$$

while the conditional posterior distributions for μ given ϕ and α is given by

$$p_{12}(\mu|\phi,\alpha,x) \sim \mathrm{GG}\left(n\phi, \left(\sum_{i=1}^n t_i^\alpha\right)^{\frac{1}{\alpha}}, \alpha\right).$$

These results are useful to obtain posterior estimates using Markov chain Monte Carlo methods. Since we proved that the posterior mean for the parameter does not return finite values, the posterior median or mode can be an alternatives as a posterior estimate.

5.5 Simulation Analysis

In this section, a simulation study using Monte Carlo methods is presented to compare the efficiency of ML method with our proposed Bayesian approach by computing the Bias and the root-mean-square error (RMSE), given by

$$\text{Bias}_i = \sum_{j=1}^{N} \frac{\hat{\theta}_{i,j}}{N} - \theta_i \ , \quad \text{RMSE}_i = \sqrt{\sum_{j=1}^{N} \frac{(\hat{\theta}_{i,j} - \theta_i)^2}{N}}, \quad \text{for}\, i = 1,2,3 \tag{5.26}$$

where $N = 10,000$ is the number of estimates obtained throughout the MLE and the posterior modes. The 95% coverage probability of the asymptotic confidence intervals and the Credible Intervals ($\text{CI}_{95\%}$) were also evaluated. Considering this approach, the best estimators will show both Bias and RMSE closer to zero. In addition, for a large number of experiments considering a 95% confidence level, the frequencies of intervals that covered the true values of θ should be closer to 95%.

To find the ML estimators, the Newton-Raphson method was adopted. In this case, the initial values to start the iterative procedure must be assigned. To ensure a fair comparison, both procedures were under the same conditions (same initial values and samples). The initial values considered were the same values used to generate the samples.

Clearly, the normalizing constant for the marginal posterior densities require two-dimensional integration. Therefore, the MCMC method was considered to obtain the posterior estimates. Since the conditional distributions of α and ϕ were not easily identified, the Metropolis-Hastings algorithm (GAMERMAN; LOPES, 2006) was considered to simulate the posterior quantities. For each simulated data set, $15,500$ iterations were performed using MCMC methods. As a burn-in, the first $1,000$ initial values were discarded, the considered thin was 30 to reduce the correlation among the chains. The Geweke criterion (GEWEKE *et al.*, 1991) was used to check the convergence of the obtained chains under a 95% confidence level. These values were used to compute the posterior mode estimates, yielding $10,000$ estimates for ϕ, μ and α.

The chosen values to perform this procedure were $\theta = ((0.5, 0.5, 3), (2, 1, 0.5), (4, 2, 2), (0.4, 1.5, 5))$ and $n = (50, 100, 200)$. The seed used to generate the random values in the R software was 2016. Table 15 presents the Bias and the RMSE of the estimates obtained through the MLE and the Bayes estimators for $10,000$ simulated samples under different values of θ and n. Table 16 shows the coverage probability with a 95% confidence level.

The Bayes estimators returned estimates with smaller Bias and RMSE than the MLEs, specially for small and moderate sample sizes. For large samples, both estimators returned similar values, i.e., as there is an increase in n, both methodologies behave similarly. Both bias and RMSE have shown to be consistent and asymptotically unbiased for the parameters. However, the CIs of the MLEs using the asymptotic method does not have good coverage probabilities. These results correlate with Prentice (1974), i.e., even for large sample sizes, the approximate normal distribution for the parameters using the ML theory could not be achieved. On the other hand, the credible interval based on the Bayes estimators provided excellent coverage probabilities even for small sample sizes. For these reasons, our Bayes estimators should be considered to achieve the parameter estimators of GG distribution.

Table 15 – Bias (RMSE) of the ML estimates and the Bayes estimators (posterior mode) for 10,000 samples of sizes $n = (50, 100, 200)$ and different values of θ.

	Classical Inference			Bayesian Inference		
θ	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\phi = 0.5$	0.179(0.73)	0.046(0.26)	0.024(0.18)	-0.015(0.23)	0.001(0.19)	0.014(0.16)
$\mu = 0.5$	0.156(0.91)	0.026(0.11)	0.014(0.07)	-0.002(0.08)	0.001(0.07)	0.006(0.05)
$\alpha = 3$	0.487(1.81)	0.240(1.04)	0.145(0.74)	-0.403(0.97)	-0.248(0.79)	-0.209(0.63)
$\phi = 2$	-0.329(1.08)	-0.102(0.93)	-0.115(0.77)	-0.508(0.62)	-0.385(0.53)	0.037(0.63)
$\mu = 1$	0.870(3.71)	1.166(3.85)	0.705(2.96)	-0.431(0.47)	-0.117(0.27)	0.818(2.64)
$\alpha = 0.5$	0.255(0.68)	0.096(0.22)	0.064(0.15)	0.220(0.27)	0.129(0.16)	0.105(0.15)
$\phi = 4$	1.616(4.92)	1.115(4.03)	0.335(2.95)	-1.811(2.07)	-1.621(1.94)	-0.755(1.46)
$\mu = 2$	1.890(4.34)	1.282(3.34)	0.557(2.23)	-0.791(0.86)	-0.562(0.75)	-0.114(0.76)
$\alpha = 2$	0.531(1.64)	0.306(1.08)	0.279(0.79)	0.993(1.33)	0.648(0.96)	0.496(0.75)
$\phi = 0.4$	0.234(0.68)	0.128(0.25)	0.035(0.14)	0.042(0.21)	0.078(0.17)	0.026(0.13)
$\mu = 1.5$	0.249(0.98)	0.105(0.21)	0.032(0.11)	0.031(0.17)	0.055(0.14)	0.022(0.10)
$\alpha = 5$	-0.013(2.07)	-0.468(1.12)	0.088(1.13)	0.502(1.89)	-0.317(1.04)	0.050(1.12)

Table 16 – Coverage probability with a 95% confidence level equals the ML estimates and the Bayes estimators (posterior mode) considering 10,000 samples of sizes $n = (50, 100, 200)$ and different values of θ.

	Classical Inference			Bayesian Inference		
θ	$n = 50$	$n = 100$	$n = 200$	$n = 50$	$n = 100$	$n = 200$
$\phi = 0.5$	89.30%	92.36%	92.39%	98.68%	97.91%	95.81%
$\mu = 0.5$	90.03%	92.31%	92.53%	99.34%	98.95%	97.61%
$\alpha = 3$	93.89%	96.01%	95.44%	98.64%	97.74%	95.62%
$\phi = 2$	75.62%	84.34%	85.74%	92.78%	94.02%	89.66%
$\mu = 1$	61.48%	71.16%	72.63%	81.16%	87.55%	89.38%
$\alpha = 0.5$	100.00%	99.01%	97.92%	91.97%	93.21%	89.35%
$\phi = 4$	80.92%	82.53%	81.37%	92.25%	92.60%	87.87%
$\mu = 2$	77.12%	78.90%	77.65%	91.79%	92.37%	88.04%
$\alpha = 2$	100.00%	98.82%	97.02%	92.06%	92.38%	87.86%
$\phi = 0.4$	97.66%	99.97%	95.76%	97.40%	95.72%	95.66%
$\mu = 1.5$	97.27%	99.83%	96.09%	98.50%	97.98%	97.85%
$\alpha = 5$	90.11%	91.88%	94.76%	97.45%	95.94%	95.82%

5.6 Real data application

As the Bayesian analysis was used improperly by Noortwijk (2001), the data set related to the annual maximum discharges of the river Rhine at Lobith, Netherlands from 1901 to 1998 was reanalyzed. The results presented by (NOORTWIJK, 2001, table 1, pg 6) can be seen in Table 17.

From the credibility intervals of ϕ and α available in Table 17, there is a good indication that the numerical techniques did not provide good results. These large credibility intervals were probably influenced due to the improper posterior distribution obtained from the Jeffreys prior. In this study, considering a proper posterior distribution, the GG distribution can be used to analyze

Table 17 – Posterior mean and 95% credibility intervals for ϕ and α from the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901-1998.

θ	Mean	$CI_{95\%}(\theta)$
ϕ	1.380	(0.01; 6.00)
$1/\mu$	4936	$(\cdots;\cdots)$
α	2.310	(0.01; 6.00)

$\cdots$ Not presented

this data under the same assumptions as Section 4. The posterior summaries obtained using the MCMC methods and the reference prior (5.25 are given in Table 18.

Table 18 – Posterior mode, standard deviations and 95% credible intervals for ϕ, μ and α from the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901-1998.

θ	Mode	SD	$CI_{95\%}(\theta)$
ϕ	3.5449	1.8416	(1.0681; 7.6216)
$1/\mu$	2,041.4	2,118.5	(1,029.1; 8,325.6)
α	1.6593	0.5730	(1.0745; 3.2432)

Note that the $CI_{95\%}(\theta)$, are very closer to $\alpha = 1$ or $\phi = 1$, i.e., the GG distribution may reduce to gamma or the Weibull distribution. The obtained results were compared with the sub-models such as Weibull, gamma and lognormal distributions using the AIC, CAIC and BIC.

Table 19 – Results of AIC, AICc and BIC criteria for different probability distributions considering the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901-1998.

Criteria	G. Gamma	Weibull	Gamma	Lognormal
AIC	428.20	430.11	**426.52**	429.81
AICc	428.45	430.23	**426.65**	429.93
BIC	435.95	435.28	**431.69**	434.98

Considering any criteria, it can be concluded from the results in Table 19 that among the chosen models, the gamma distribution fit best considering the annual maximum discharges of the river Rhine at Lobith from 1901 to 1998. Moreover, to verify the goodness of fit, Figure 7 shows the survival function adjusted for different distributions of overlapping probability in the empirical function.

Noortwijk (2001) argued that "... the Dutch river dikes have to withstand water levels and discharges with an average return period of up to 1250 years, where a downstream water level can be determined on the basis of the upstream discharge by using a river flow simulation model".

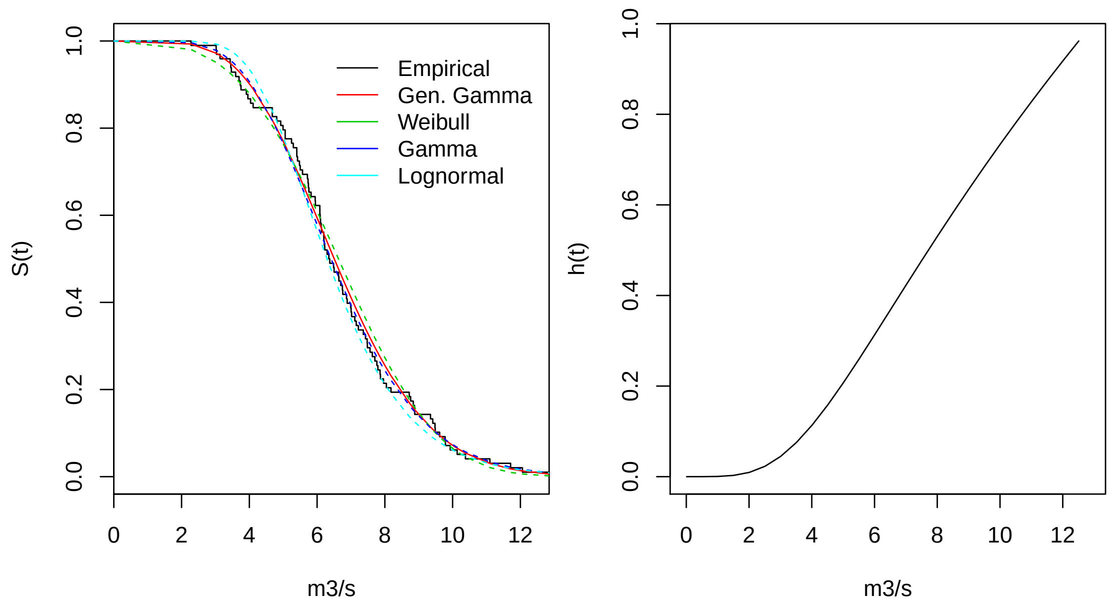

Figure 7 – Survival function fitted by the empirical and by different probability distributions considering the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901-1998 and the hazard function fitted by a GG distribution.

The main aim was to find the annual maximum river discharge in which the probability of exceedance is $1/1250$ per year. Table 20 presents the discharge with a probability of exceedance of $1/1250$ and the 90% uncertainty interval for the GG distribution (Van Noortwijk, our results and the classical inference) and the two-parameter gamma distribution. To evaluate the Bayes estimators (v) of the two-parameter gamma distribution, we considered a posterior distribution obtained with the Tibshirani prior.

Table 20 – Posterior mode, MLE, Posterior mean and 90% credibility intervals for ϕ and α from the data set related to the annual maximum discharges of the river Rhine at Lobith during 1901-1998.

Distribution	River discharge	$CI_{90\%}$
G. gamma (Van Noortwijk)	15,150	(12,950; 16,950)
G. gamma (Our Approach)	15,515	(10,561; 22,212)
G. gamma (Classical Inference)	14,780	(12,699; 16,542)
Two-parameter gamma	15,690	(14,342; 17,538)

The improper posterior produced an underestimated annual maximum discharge. The difference between the Van Noortwijk estimate and ours was 365 $m3/s$. Hence, the Dutch river dikes will have to withstand water levels and discharges of up to $15,690$ m3/s. The ML estimators of the GG distribution also returned an underestimated value for the River maximum discharge. On the other hand, the results obtained from the gamma distribution are similar to those obtained from the GG distribution in our approach. Our results clearly showed that the gamma distribution should be used to estimate the annual maximum discharges of the river Rhine at Lobith.

5.7 Discussion

We have provided in a simple way necessary and sufficient conditions to check whether or not improper priors lead to proper posterior distributions for the GG distribution. In this case, one can easily check if the obtained posterior is proper or improper directly looking at the behavior of the improper prior. From the main theorem, we proved that the uniform prior, the prior obtained from Jeffreys' first rule and the MDI prior lead to improper posteriors.

The impropriety of the posterior using the Jeffreys' priors (NOORTWIJK, 2001) led us to consider the scenario where the Jeffreys prior has an independent structure (FONSECA; FERREIRA; MIGON, 2008). However, the four possible objective priors also returned improper posteriors. An alternative was to consider the reference priors. Since these priors are sensitive to the ordering of the unknown parameters, from Proposition 2.3.1 we obtained six reference priors, two of them were similar to other reference priors. Among the four distinct reference priors, we proved that only one leads to a proper posterior distribution without the need of compact approximations or truncate possible values of the parameters. The proper reference posterior has excellent theoretical properties such as invariance property under one-to-one transformations of the parameters, consistency under marginalization and consistent sampling properties and should be used to make inference in the parameters of the GG distribution.

CHAPTER

6

CLOSED-FORM ESTIMATORS FOR NAKAGAMI-M AND GAMMA DISTRIBUTIONS

6.1 Introduction

Despite the enormous evolution of computational methods during the last decades, the methods discussed up to here still carry the disadvantage of high computational cost in many applications. Particularly in the case where the parameter estimators need to be obtained in real time, often within devices with embedded technology (SONG, 2008). To overcome this problem closed-form estimators are preferred.

For Nakagami-m distribution an unbiased closed-form estimator for the scale parameter can be easily obtained through the method of moments Nakagami (1960). However, considerable effort has been made to derive efficient estimators for the fading parameter. The maximum likelihood (ML) estimators have been discussed earlier Cheng and Beaulieu (2001). The accuracy of different procedures have been compared numerically and the use of a closed-form approximation to the ML estimator has been suggested Zhang (2002). A generalized moments estimator has been presented Cheng and Beaulieu (2002). Some estimators based on approximations of the transcendental equations that arise in the computation of ML and the generalized moment (GM) estimators have been discussed Gaeddert and Annamalai (2004). However, such estimators are only approximations to the natural methods motivated by fast computation avoiding solving nonlinear equations. A closed-form estimator for the fading parameter obtained as a limiting procedure of the traditional GM estimators have been proposed Wang, Song and Cheng (2012). For gamma distribution the estimation procedure based on the method of moments produce closed-form estimator. Hwang and Huang (2002) presented a more efficient moment estimator based on its characterization that outperforms the MLE for some cases. However, for both models the estimators presented so far have considerable bias, specially for small samples sizes.

In this chapter, we discuss maximum likelihood estimators that have closed-form expressions based on the (GG) distribution. Using the same idea we propose a class of maximum a posteriori (MAP) estimators for the parameters of the Nakagami-m and gamma distributions. They have simple closed-form expressions and can be rewritten as a bias corrected MLEs. Finally, numerical results have shown that the MAP estimation scheme outperforms the existing estimation procedures and produces almost unbiased estimates for the parameters even for small sample size.

The remainder of this chapter is organized as follows: Section 3 presents the new estimators based on the GG distribution. Section 4 displays the proposed approach on the Nakagami-m distribution. Section 5 presents the same methodology on the gamma distribution. Finally, Section 6 summarizes the study.

6.2 Closed-form estimators based on the generalized gamma distribution

Let T be a random variable the generalized gamma (GG) distribution with a PDF given by (5.1). Here, we considered a simple reparametrization where $\mu^\alpha = \frac{\phi}{\lambda}$, i.e., the PDF is

$$f(t|\theta) = \frac{\alpha}{\Gamma(\phi)}\left(\frac{\phi}{\lambda}\right)^{\phi} t^{\alpha\phi-1}\exp\left(\frac{\phi}{\lambda}t^{\alpha}\right). \tag{6.1}$$

Although a non common parametrization is considered, when α is known and $\alpha \in \mathbb{N}$ the obtained models have orthogonal parameters (ϕ, λ) in the sense of Cox and Reid (1987). The likelihood function from (6.1) is given by

$$L(\theta;t) = \frac{\alpha^n}{\Gamma(\phi)^n}\left(\frac{\phi}{\lambda}\right)^{n\phi}\left\{\prod_{i=1}^{n} t_i^{\alpha\phi-1}\right\}\exp\left(-\frac{\phi}{\lambda}\sum_{i=1}^{n} t_i^{\alpha}\right). \tag{6.2}$$

The maximum likelihood estimates of the parameters are obtained by solving the following likelihood equations

$$\hat\phi = \frac{n}{\left(\frac{1}{\lambda}\sum_{i=1}^{n} t_i^{\alpha}\log\left(t_i^{\alpha}\right) - \sum_{i=1}^{n}\log\left(t_i^{\alpha}\right)\right)}, \tag{6.3}$$

$$\hat\lambda = \frac{1}{n}\sum_{i=1}^{n} t_i^{\hat\alpha} \tag{6.4}$$

and the MLE for α is obtained solving the non-linear equation

$$\log(\hat\phi) - \psi^{(0)}(\hat\phi) = \log(\hat\lambda) - \frac{1}{n}\sum_{i=1}^{n}\log\left(t_i^{\hat\alpha}\right). \tag{6.5}$$

Considering that α is know we can remove equation (6.5) and focus on the first two equations to obtain closed-form estimators for different distributions such as Nakagami, gamma, to list a few. The applications of this idea are discussed in the following sections.

6.3 Nakagami-m

In this section, we discuss closed-form estimators for Nakagami-m parameters. Firstly, we review an estimator based on the generalized moments, then we discuss the new method.

6.3.1 Generalized moment estimators

A useful estimation procedure for the fading parameter of the NK distribution was proposed earlier (CHENG; BEAULIEU, 2002) and is given through the fractional moment estimator

$$\hat{\mu}_{1/p} = \frac{\hat{m}_{1/p}\hat{m}_2}{2p\left(\hat{m}_{2+1/p} - \hat{m}_{1/p}\hat{m}_2\right)}, \tag{6.6}$$

where the kth-order moment is given by

$$m_k = \frac{\Gamma(\mu+k/2)}{\Gamma(\mu)}\left(\frac{\Omega}{\mu}\right)^{k/2}$$

and p is a positive real number. Further, the limiting estimator, $\hat{\mu}_0$, that combined with the fractional moment estimator (WANG; SONG; CHENG, 2012), is given by

$$\hat{\mu}_k = \begin{cases} \dfrac{\hat{m}_{1/p}\hat{m}_2}{2p\left(\hat{m}_{2+1/p} - \hat{m}_{1/p}\hat{m}_2\right)}, & k > 0 \\ \dfrac{\hat{m}_2}{\left(\frac{1}{n}\sum_{i=1}^n t_i^2 \log\left(t_i^2\right) - \frac{1}{n}\hat{m}_2 \sum_{i=1}^n \log\left(t_i^2\right)\right)}, & k = 0. \end{cases} \tag{6.7}$$

The authors showed that in the limiting case, $\hat{\mu}_0$ is expected to achieve the best performance among this fractional moment-based estimator family. Hence, we considered only the case when $k = 0$, namely $\hat{\mu}_{GME}$. It was also presented for $k = 0$ the asymptotic variance, $\mathrm{Var}(\mu_{MAP})$, given by

$$Var(\mu_{\mathrm{GME}}) = \mu^2 + \mu^3\psi^{(1)}(\mu+1). \tag{6.8}$$

6.3.2 Modified maximum likelihood estimators

Consider the maximum likelihood estimators (6.3) and (6.4). By substituting $\mu = \phi$, $\Omega = \lambda$ and $\alpha = 2$ we have the following modified maximum likelihood estimators

$$\hat{\mu}_{MMLE} = \frac{n}{\left(\frac{1}{\hat{\Omega}}\sum_{i=1}^n t_i^2 \log\left(t_i^2\right) - \sum_{i=1}^n \log\left(t_i^2\right)\right)} \quad \text{and} \quad \hat{\Omega}_{MMLE} = \frac{1}{n}\sum_{i=1}^n t_i^2, \tag{6.9}$$

which is the same of (6.7) when $k = 0$. Hereafter, we are going to call the MMLEs as GMEs since they have the same structure of the estimators presented by Wang, Song and Cheng (2012). The derivation of the variance of the MMLEs for Ω can be easily obtained considering

$$Var(\Omega_{\text{MMLE}}) = \frac{1}{n}\left(E[X^4] - E[X^2]^2\right) = \frac{\Omega^2}{n\mu}$$

while the variance of μ is the same as (6.8).

6.3.3 Maximum a Posteriori estimator

In this section, we consider the Bayesian inference to derive Bayes estimators with smaller bias. Under this approach, the most common objective priors were considered such as Jeffreys Prior, Reference Prior, MDI prior (see Chapter 5 for more details). However, such priors depend on polygamma function which did not allow to obtain MAP estimators in closed form. The chosen objective prior for the parameters is given by

$$\pi(\theta) \propto \frac{1}{\Omega^{c_1}\mu^{c_2}\alpha^{c_3}} \tag{6.10}$$

where $\theta = (\Omega, \mu, \alpha)$ and $c_i \geq 0, i = 1,2,3$ are known hyperparameteres. From the product of the likelihood function (6.2) and the prior distribution (6.10), the joint posterior distribution for θ is given by

$$\pi(\theta|t) = \frac{1}{d(t)}\frac{\alpha^{n-c_3}}{\mu^{c_2}\Omega^{c_1}\Gamma(\mu)^n}\left(\frac{\mu}{\Omega}\right)^{n\mu}\left\{\prod_{i=1}^{n} t_i^{\alpha\mu-1}\right\}\exp\left(-\frac{\mu}{\Omega}\sum_{i=1}^{n} t_i^{\alpha}\right), \tag{6.11}$$

where

$$d(t) = \int\limits_{\mathscr{A}} \frac{\alpha^{n-c_3}}{\mu^{c_2}\Omega^{c_1}\Gamma(\mu)^n}\left(\frac{\mu}{\Omega}\right)^{n\mu}\left\{\prod_{i=1}^{n} t_i^{\alpha\mu-1}\right\}\exp\left(-\frac{\mu}{\Omega}\sum_{i=1}^{n} t_i^{\alpha}\right) d\theta$$

and $\mathscr{A} = \{(0,\infty)\times(0.5,\infty)\times(\varepsilon, M)\}$ is the parameter space of θ, where $0 < \varepsilon < 2$ is a small constant and $M > 2$ is a large constant. We chose (ε, M) for the interval of α since the only interest is in the case where $\alpha = 2$. Therefore, any interval (ε, M) containing $\alpha = 2$ will be satisfactory for our purposes.

The MAP of θ is computed through $\hat{\theta}_{MAP} = \arg\max_{\theta} \log(\pi(\theta|t))$. After some algebraic manipulation we have

$$\hat{\Omega} = \frac{\mu\sum_{i=1}^{n} t_i^{\hat{\alpha}}}{n\mu + c_1}. \tag{6.12}$$

It is important to point out that (6.12) will be equal (6.4), if and only if $c_1 = 0$, i.e, Ω is unbiased when $\alpha = 2$. Therefore, we consider only that $c_1 = 0$. To obtain reliable inference results, we have to check if (6.11) is a proper posterior distribution, i.e, $d(t) < \infty$.

Theorem 6.3.1. The posterior distribution (6.11) is proper.

Proof. Let $\mathscr{B} = \{(\varepsilon, M) \times [0.5, \infty) \times (0, \infty)\}$ and consider the change in the coordinates through the transformation $\theta : \mathscr{B} \to \mathscr{A}$, where $(\alpha, \mu, \Omega) = \theta(\beta, \phi, \lambda) = \left(\beta, \phi, \frac{\phi}{\lambda}\right)$ and $\mathscr{A} = \theta(\mathscr{B})$

Noticing that $|\det(D\theta(\phi, \lambda))| = \phi\lambda^{-2}$, denoting $\Theta = (\beta, \phi, \lambda)$ and applying the change of variables on the Lebesgue integral and the Fubini-Tonelli Theorem (FOLLAND, 1999) we have that

$$\begin{aligned}
d(t) &\propto \int\limits_{\mathscr{A}} \frac{\alpha^{n-c_3}}{\mu^{c_2}\Gamma(\mu)^n} \left(\frac{\mu}{\Omega}\right)^{n\mu} \left\{\prod_{i=1}^n t_i^{\alpha\mu}\right\} \exp\left(-\frac{\mu}{\Omega}\sum_{i=1}^n t_i^{\alpha}\right) d\theta \\
&= \int\limits_{\mathscr{B}} \frac{\beta^{n-c_3}\lambda^{n\phi-2}}{\phi^{c_2-1}\Gamma(\phi)^n} \left\{\prod_{i=1}^n t_i^{\beta\phi}\right\} \exp\left(-\lambda\sum_{i=1}^n t_i^{\beta}\right) d\Theta \\
&\propto \int\limits_{\varepsilon}^{M}\int\limits_{0.5}^{\infty}\int\limits_{0}^{\infty} \frac{\beta^{n-c_3}}{\phi^{c_2-1}\Gamma(\phi)^n} \left\{\prod_{i=1}^n t_i^{\beta\phi}\right\} \lambda^{n\phi-2} e^{-\lambda\sum_{i=1}^n t_i^{\beta}}\, d\lambda d\phi d\beta \\
&= \int\limits_{\varepsilon}^{M} \beta^{n-c_3} \int\limits_{0.5}^{\infty} \frac{\Gamma(n\phi-1)}{\phi^{c_2-1}\Gamma(\phi)^n} \frac{\left(\prod_{i=1}^n t_i^{\beta}\right)^{\phi}}{\left(\sum_{i=1}^n t_i^{\beta}\right)^{n\phi-1}} d\phi d\beta \\
&\propto \int\limits_{\varepsilon}^{M} \beta^{n-c_3} \sum_{i=1}^n t_i^{\beta} \int\limits_{0.5}^{\infty} \phi^{\frac{n-1-2c_2}{2}} e^{np(\beta)\phi} d\phi d\beta \\
&\propto \int\limits_{\varepsilon}^{M} \beta^{n-c_3} \sum_{i=1}^n t_i^{\beta} \frac{\Gamma(\frac{n+1-2c_2}{2}, 0.5n\,\mathrm{p}(\beta))}{(n\,\mathrm{p}(\beta))^{\frac{n+1-2c_2}{2}}} d\beta < \infty
\end{aligned}$$

where $p(\beta) = \left(\sqrt[n]{\prod_{i=1}^n t_i^{\beta}}\right)\left(\frac{1}{n}\sum_{i=1}^n t_i^{\beta}\right)^{-1} > 0$ by the inequality of the arithmetic and geometric mean. □

The other MAP estimators are given by

$$\hat{\mu} = \frac{(n-c_3)}{\left(\frac{1}{\hat{\Omega}}\sum_{i=1}^n t_i^{\alpha}\log\left(t_i^{\alpha}\right) - \sum_{i=1}^n \log\left(t_i^{\alpha}\right)\right)}\,, \tag{6.13}$$

and the MAP for α is obtained by solving the non-linear equation

$$\log\left(\hat{\mu}\right) - \psi^{(0)}(\hat{\mu}) = \log\left(\hat{\Omega}\right) - \frac{1}{n}\sum_{i=1}^n \log(t_i^{\hat{\alpha}}) + \frac{c_2}{n\hat{\mu}}.$$

Therefore, for $\alpha = 2$, a hybrid MAP estimator of Ω is given by $\hat{\Omega}_{MAP} = \frac{1}{n}\sum_{i=1}^n t_i^2$ and the Nakagami-m fading parameter can be estimated by

$$\hat{\mu}_{\text{MAP}} = \frac{(n-c_3)\frac{1}{n}\sum_{i=1}^n t_i^2}{\left(\sum_{i=1}^n t_i^2\log\left(t_i^2\right) - \frac{1}{n}\sum_{i=1}^n t_i^2\sum_{i=1}^n \log\left(t_i^2\right)\right)}. \tag{6.14}$$

Theorem 6.3.2. Let $\hat{\mu}_{MAP}$ be an estimator of μ , and let $t = (t_1, \ldots, t_n)$, with $t_n \geq \ldots \geq t_1$, not all equal, then for $n > c_3$ we have:

$$\hat{\mu}_{\text{MAP}} = \frac{(n-c_3)\frac{1}{n}\sum_{i=1}^n t_i^2}{\sum_{k=1}^{n-1} k\left(\frac{1}{n}\sum_{i=1}^n t_i^2 - \frac{1}{k}\sum_{i=1}^k t_i^2\right)\log(t_{k+1}^2/t_k^2)} > 0.$$

Proof. We have that

$$
\begin{aligned}
\hat{\mu} =& \frac{(n-c_3)\frac{1}{n}\sum_{i=1}^{n} t_i^2}{\sum_{k=1}^{n}\left(t_k^2 - \frac{1}{n}\sum_{i=1}^{n} t_i^2\right)\log(t_k^2)} \\
=& \frac{(n-c_3)\frac{1}{n}\sum_{i=1}^{n} t_i^2}{\sum_{k=1}^{n}\left(\frac{(k-1)}{n}\sum_{i=1}^{n} t_i^2 - \sum_{i=1}^{k-1} t_i^2 - \frac{k}{n}\sum_{i=1}^{n} t_i^2 + \sum_{i=1}^{k} t_i^2\right)\log(t_k^2)} \\
=& \frac{(n-c_3)\frac{1}{n}\sum_{i=1}^{n} t_i^2}{\sum_{k=1}^{n-1} k\left(\frac{1}{n}\sum_{i=1}^{n} t_i^2 - \frac{1}{k}\sum_{i=1}^{k} t_i^2\right)\log(t_{k+1}^2/t_k^2)}.
\end{aligned}
$$

Now, since $t_n \geq t_{n-1} \geq \cdots \geq t_1 > 0$, we have that $\log(t_{k+1}^2/t_k^2) \geq 0$ and $\frac{1}{n}\sum_{i=1}^{n} t_i^2 > \frac{1}{k}\sum_{i=1}^{k} t_i^2$ for every $1 \leq k < n$. Moreover, since $t_1, \cdots, t_n$ are not all equal, we have that $\log(t_{j+1}^2/t_j^2) > 0$ for some $0 \leq j < n$. Then $\sum_{k=1}^{n-1} k\left(\frac{1}{n}\sum_{i=1}^{n} t_i^2 - \frac{1}{k}\sum_{i=1}^{k} t_i^2\right)\log(t_{k+1}^2/t_k^2) > 0$ which, for $n > c_3$, implies that $\hat{\mu} > 0$.

□

Note that the MAP estimator does not depend on c_2. Moreover, the $\hat{\mu}_{\text{MAP}}$ can be rewritten as a bias corrected generalized moment estimator

$$
\hat{\mu}_{\text{MAP}} = \frac{(n-c_3)}{n}\hat{\mu}_{\text{GME}}. \tag{6.15}
$$

Due to this relationship, the asymptotic variance, $\text{Var}(\mu_{MAP})$, can be obtained by

$$
Var(\mu_{\text{MAP}}) = Var\left(\frac{(n-c_3)}{n}\hat{\mu}_{\text{GME}}\right) = \frac{(n-c_3)^2}{n^2}\left(\mu^2 + \mu^3\psi^{(1)}(\mu+1)\right).
$$

6.3.4 Results

In this section, we present Monte Carlo simulation studies to compare the efficiency of our proposed estimation method. The comparison between such procedures is carried out by computing the mean relative error (MRE) and the root-mean-square error (RMSE), given by

$$
\text{MRE}_{\theta_i} = \frac{1}{N}\sum_{j=1}^{N}\frac{\hat{\theta}_{i,j}}{\theta_i} \quad \text{and} \quad \text{RMSE}_{\theta_i} = \sqrt{\sum_{j=1}^{N}\frac{(\hat{\theta}_{i,j}-\theta_j)^2}{N}},
$$

for $i = 1, 2$, where $N = 1{,}000{,}000$ is the number of estimates obtained through the ME, ML, GM and MAP estimators. The MRE and the RMSE of the Ω are the same for the different estimation procedures.

Considering this approach, we expect that the most efficient estimators would yield the MREs closer to one with smaller RMSEs. These results were computed using the software R (R Core Development Team). The seed used to generate the pseudo-random samples from the NK distribution was 2016.

The MAP estimator depends on c_3. Therefore, we have to find a value for c_3 in which the MRE is closer to one. Figure 8 presents the MREs for μ_{MAP} considering different values of c_3, for $\mu = 20$, $\Omega = 2$, $n = 20$ and $\mu = 20$, $\Omega = 2$, $n = 20$. We omitted the results of the simulation study for different values of μ, Ω and n since they are similar to the one presented here.

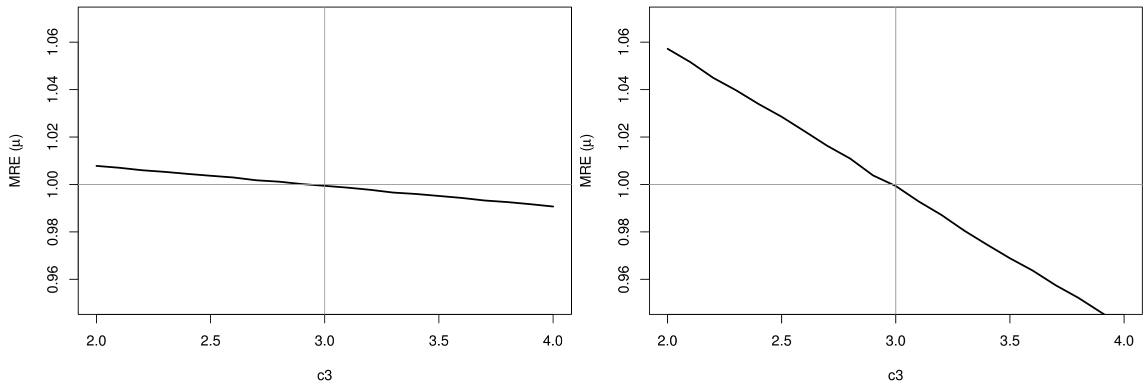

Figure 8 – MREs for μ considering $c_3 = (2, 2.1, \cdots, 4)$, $\mu = 6$, $\Omega = 10$, $n = 120$ (left panel) $\mu = 20$, $\Omega = 2$, $n = 20$ (right panel) and for $N = 1,000,000$ simulated samples and $n = 50$.

From the Figure 8, we observed that a good choice is $c_3 = 3$. Therefore, we considered that $c_3 = 3$ in (6.14). Figures 9-12 show the MREs, RMSEs from the estimates of μ obtained using the MC method. Figure 9 also presents the MREs, RMSEs from the estimates of Ω, we omitted the other graphics since they were similar. The horizontal lines in the figures correspond to MREs and RMSEs being one and zero respectively.

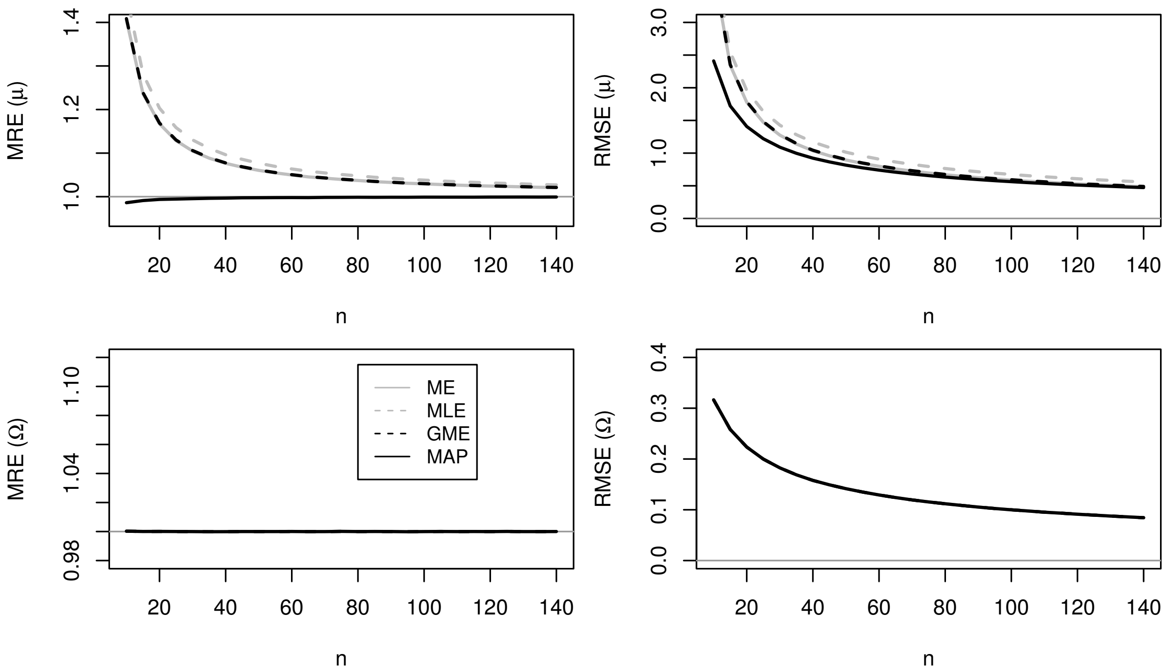

Figure 9 – MREs, RMSEs for μ and Ω considering $\mu = 4, \Omega = 2$ for $N = 1,000,000$ simulated samples and $n = (10, 15, \ldots, 140)$.

From the Figures 9 and 10, we observed that the estimates of the fading parameter are asymptotically unbiased, i.e., the MREs tend to one when n increases and the RMSEs decrease

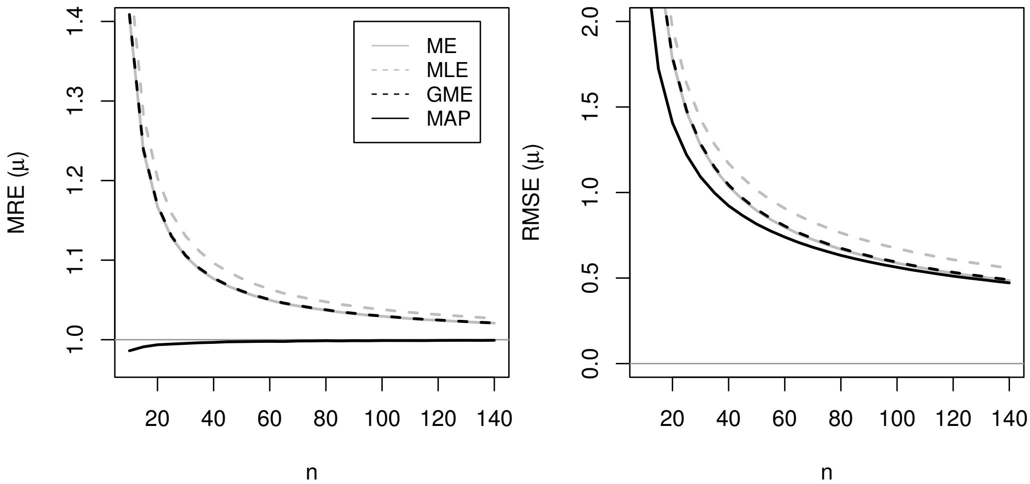

Figure 10 – MREs, RMSEs for μ considering $\mu = 4$, $\Omega = 80$ for $N = 1,000,000$ simulated samples and $n = (10, 15, \ldots, 140)$.

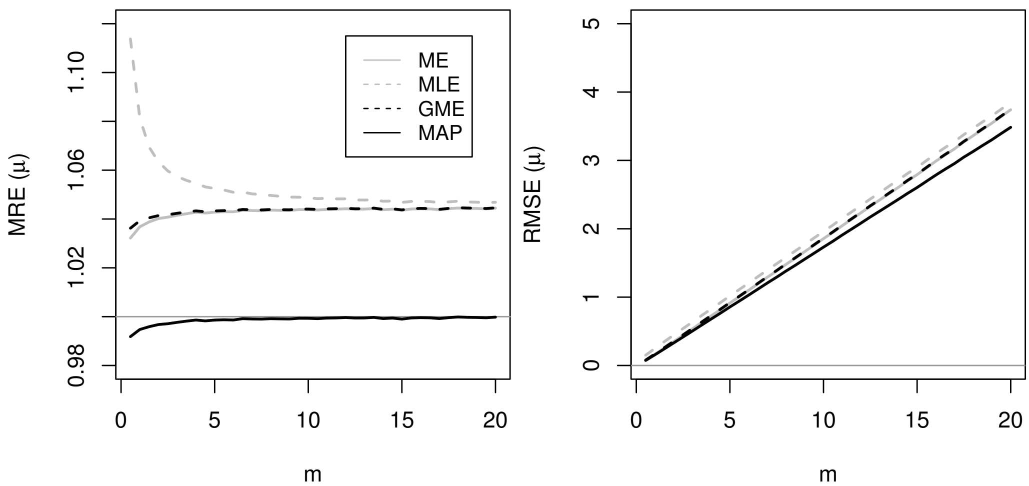

Figure 11 – MREs, RMSEs for different values of $\mu = (0.5, 1.0, \ldots, 20)$ considering $\Omega = 2$ for $N = 1,000,000$ simulated samples and $n = 50$.

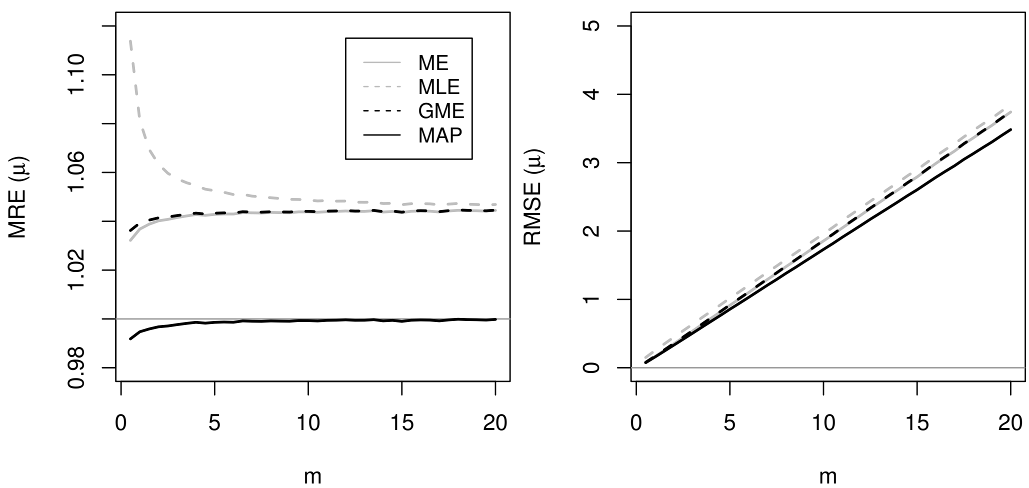

Figure 12 – MREs, RMSEs for different values of $\mu = (0.5, 1.0, 1.5, \ldots, 20)$ considering $\Omega = 80$ for $N = 1,000,000$ simulated samples and $n = 50$.

to zero. Moreover, the MAP estimators present extremely efficient estimates for μ even for small sample sizes, for instance, considering $n = 15$, the errors related to the MAP are of the order 10^{-3} while for the GME are 10^{-1}, i.e, the MAP estimator is almost unbiased for small samples. From Figures 11 and 12, we obtained similar results considering different values of μ. Taking into account the results of the simulation studies, the MAP estimators should be considered for estimating the fading parameter of the NK distribution.

6.4 Gamma distribution

The same approach can be considered for the gamma distribution. Let X be a non negative random variable with a gamma PDF given by

$$f(t|\phi,\mu) = \frac{\mu^{\phi}}{\Gamma(\phi)} t^{\phi-1} \exp\left(-\mu t\right),$$

where $\phi > 0$ and $\mu > 0$ are the shape and scale parameters. Considering $\alpha = 1$ in (6.3) and (6.4) the hybrid maximum likelihood estimators for the gamma distribution are given by

$$\hat{\phi}_{YC} = \frac{n\sum_{i=1}^{n} t_i}{\left(n\sum_{i=1}^{n} t_i \log\left(t_i\right) - \sum_{i=1}^{n} t_i \sum_{i=1}^{n} \log\left(t_i\right)\right)} \tag{6.16}$$

and

$$\hat{\mu}_{YC} = \frac{1}{n^2}\left(n\sum_{i=1}^{n} t_i \log\left(t_i\right) - \sum_{i=1}^{n} t_i \sum_{i=1}^{n} \log\left(t_i\right)\right). \tag{6.17}$$

For gamma distribution, the proposed estimators were firstly presented by Ye and Chen (2017). Further, the authors also discussed a bias corrective approach that is presented as follow.

6.4.1 Bias corrected estimators

A useful bias corrections for (6.16) and (6.17) were presented by Ye and Chen (2017). The modified maximum likelihood estimators for ϕ and μ are given by

$$\hat{\phi}_{BC_1} = \frac{(n-1)}{(n+2)} \frac{n\sum_{i=1}^{n} t_i}{\left(n\sum_{i=1}^{n} t_i \log\left(t_i\right) - \sum_{i=1}^{n} t_i \sum_{i=1}^{n} \log\left(t_i\right)\right)} \tag{6.18}$$

and

$$\hat{\mu}_{BC_1} = \frac{1}{n(n-1)}\left(n\sum_{i=1}^{n} t_i \log\left(t_i\right) - \sum_{i=1}^{n} t_i \sum_{i=1}^{n} \log\left(t_i\right)\right), \tag{6.19}$$

hereafter, BC_1 estimators. Although $\hat{\mu}_{BC_1}$ is an unbiased estimators for μ, the $\hat{\phi}_{BC_1}$ has a systematic bias as it shall be shown in the next section. Therefore, our effort is to present improved bias corrected estimators for ϕ.

6.4.2 Bias Expression

Cox and Reid (1987) presented elegant expressions to derive the bias for the parameters of parametric models. However, to derive the bias correction for (6.16) we would have to calculate the bias of the generalized gamma distribution parameter and use the delta method. Unfortunately, these bias are very complex to be calculated. On the order hand, Ye and Chen (2017) showed that both, the MLE and the closed form estimator of ϕ returned similar results. Therefore, we may consider the bias correction presented by Cox and Reid (1987) for the MLE of ϕ that after tedious calculations is given by

$$Bias(\hat{\phi}) = \frac{\hat{\phi}\psi^{(1)}(\hat{\phi}) - \hat{\phi}^2\psi^{(2)}(\hat{\phi}) - 2}{2n\left(\hat{\phi}\psi^{(1)}(\hat{\phi}) - 1\right)^2} + O(n^{-2}). \tag{6.20}$$

It follows that

$$\hat{\phi}_{\mathrm{BC}_2} = \hat{\phi}_{\mathrm{YC}} - Bias(\hat{\phi}_{\mathrm{YC}}), \tag{6.21}$$

hereafter, BC_2 estimator. This expression is easily obtained by using the orthogonal reparametrization of the gamma distribution and considering the same steps as described by Schwartz, Godwin and Giles (2013). However, the solution of this estimator involves the computation of transcendental functions, increasing considerably the computational time. In the following, we describe a different approach.

6.4.3 Maximum a Posteriori estimator

Note that from (6.13) and considering the parametrization given in taking ϕ

$$\hat{\phi}_{\mathrm{MAP}} = \frac{(n-c)\sum_{i=1}^n t_i^\alpha}{\left(n\sum_{i=1}^n t_i^\alpha \log\left(t_i^\phi\right) - \sum_{i=1}^n t_i^\alpha \sum_{i=1}^n \log\left(t_i^\alpha\right)\right)},$$

where α is the parameter that select the chosen distribution and c is a constant that calibrates the estimator in order to decrease the bias. For $c = 0$ and $\alpha = 1$ we have (6.16). As the value of c is unknown, we performed a similar study considering $\alpha = 1$ and obtained $c = 2.9$ as the optimal value. This study is available in the next section. Therefore, we have

$$\hat{\phi}_{\mathrm{BC}_3} = \frac{(n-2.9)\sum_{i=1}^n t_i}{\left(n\sum_{i=1}^n t_i \log(t_i) - \sum_{i=1}^n t_i \sum_{i=1}^n \log(t_i)\right)}, \tag{6.22}$$

hereafter, BC_3 estimator. The asymptotic variance of (9) is given by

$$Var(\hat{\phi}_{\mathrm{BC}_3}) = Var\left(\frac{(n-2.9)}{n}\hat{\phi}_{\mathrm{YC}}\right) = \frac{(n-2.9)^2}{n^2}\phi^2\left(1+\phi\psi^{(1)}(\phi+1)\right).$$

6.4.4 Simulation and Discussion

Since the MAP estimator depends on c_3. We have to find a value for c_3 in which the MRE is closer to one. Figure 8 presents the MREs for ϕ_{MAP} considering different values of c_3,

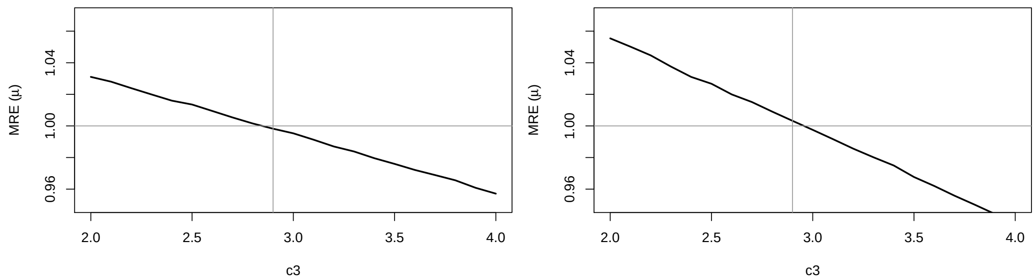

Figure 13 – MREs for ϕ considering $c_3 = (2, 2.1, \cdots, 4)$, $\phi = 3$, $\lambda = 4$ and $n = 20$ (left panel) $\phi = 10$, $\lambda = 4$ and $n = 20$ (right panel) for $N = 500,000$ simulated samples.

for $\phi = (3, 10)$, $\lambda = 4$ and $n = 20$. We omitted the results of the simulation study for different values of ϕ, λ and n once they are similar to the one presented here.

From the Figure 13, we observed that a good choice is $c_3 = 2.9$. Therefore, we considered that $c_3 = 2.9$ in (6.22) and $n \geq 3$. A simulation is performed in order to compare the performance of the proposed estimators. We follow Ye and Chen (2017) and consider $\beta = 1$. The mean relative error (MRE) and the root-mean-square error (RMSE) are considered as comparative measures based on 50,000,000 simulated samples. Therefore, we expect that the most efficient estimators would yield MREs closer to one with smaller RMSEs. The results are shown in Figures 14-15. We observe that both BC2 and BC3 provided much smallest bias than BC1, with preference to BC3, which is straightforwardly obtained.

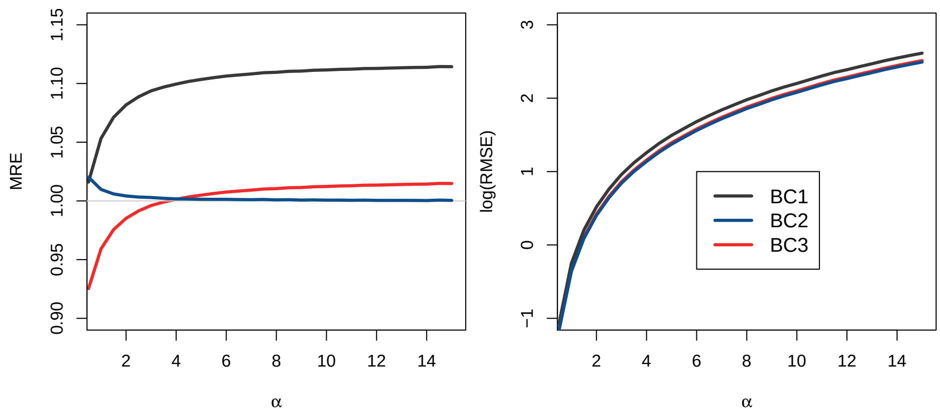

Figure 14 – MREs and RMSEs for different values of $\phi = (0.5, 1.0, \ldots, 20)$ for sample size of 8 elements.

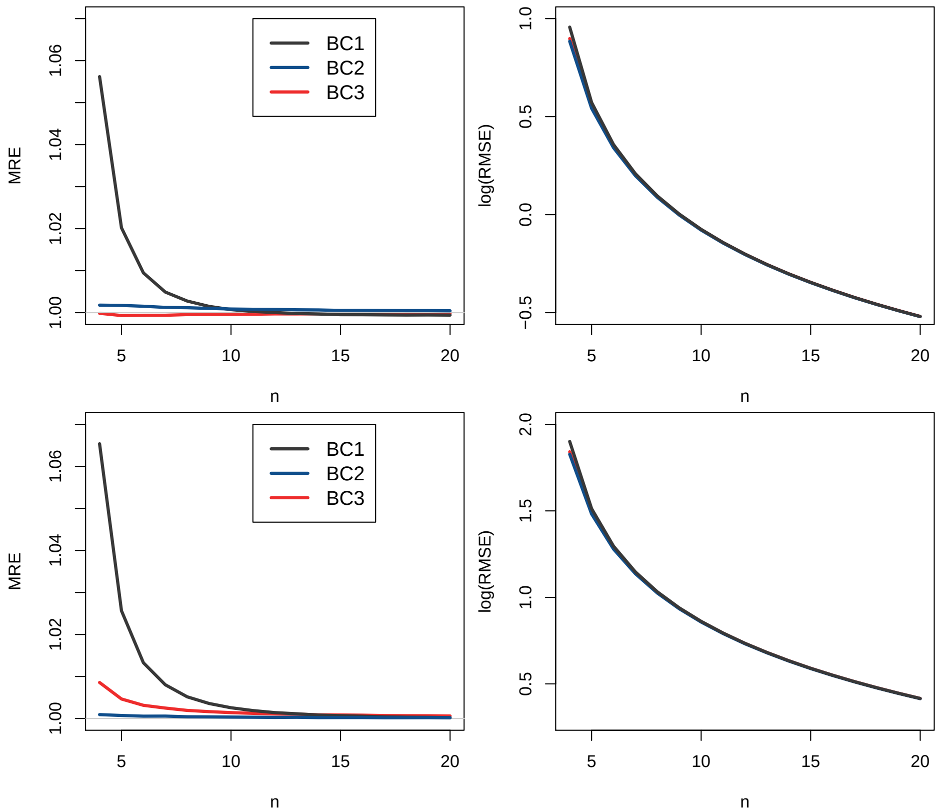

Figure 15 – MREs and RMSEs for ϕ for samples sizes of $6, 7, 8, \ldots, 20$ elements. Upper panels: considering $\phi = 4$, lower panels: considering $\phi = 10$.

6.5 Discussion

In this work, we have introduced MAP estimators for the Nakagami-m and gamma distribution. Some mathematical properties for this new estimator are presented. We show that such estimator can be rewritten as a bias corrected modified maximum likelihood estimators. Numerical results have shown that the MAP estimators outperforms the existing estimator procedures for the parameters of both distribution. In addition, we conclude that the MAP estimators present almost unbiased estimates for the parameter even for small sample sizes.

CHAPTER

7

MIXTURE OF GENERALIZED GAMMA DISTRIBUTION

7.1 Introduction

In recent years, several new extensions of the exponential distribution have been introduced in the literature for describing real problems. Ghitany, Atieh and Nadarajah (2008) investigated different properties of the Lindley distribution and outlined that in many cases the Lindley distribution is a better model than one based on the exponential distribution. Since then, many generalizations of the Lindley distribution have been introduced, such as generalized Lindley (ZAKERZADEH; DOLATI, 2009), weighted Lindley (GHITANY *et al.*, 2011b), extended Lindley (BAKOUCH *et al.*, 2012), exponential Poisson Lindley (BARRETO-SOUZA; BAKOUCH, 2013), Power Lindley (GHITANY *et al.*, 2013) distribution, among others.

Here, a new lifetime distribution family is proposed by considering a PDF expressed as a two-component mixture

$$f(t|\phi,\lambda,\alpha) = pf_1(t|\phi,\lambda,\alpha)+(1-p)f_2(t|\phi,\lambda,\alpha)$$

where $p=\lambda/(\lambda+\phi)$ and $T_j \sim \text{GG}(\phi+j-1,\lambda,\alpha)$, for $j=1,2$, i.e, $f_j(t|\lambda,\phi)$ has generalized gamma distribution. The PDF is given by

$$f(t|\phi,\lambda,\alpha) = \frac{\alpha\lambda^{\alpha\phi}}{(\lambda+\phi)\Gamma(\phi)}t^{\alpha\phi-1}(\lambda+(\lambda t)^{\alpha})e^{-(\lambda t)^{\alpha}}, \tag{7.1}$$

for all $t>0$, $\phi>0, \lambda>0$ and $\alpha>0$. The proposed distribution can be referred as generalized weighted Lindley (GWL) distribution. Important probability distributions can be obtained from the GWL distribution as the weighted Lindley distribution $(\alpha=1)$, Power Lindley distribution $(\phi=1)$ and the Lindley distribution ($\phi=1$ and $\alpha=1$). Due to this relationship, such model could also be named as weighted power Lindley or generalized power Lindley distribution. The proposed model has different forms of the hazard function, such as: increasing, decreasing, bathtub, unimodal or decreasing-increasing-decreasing shape, making the GWL distribution a flexible model for reliability data. Moreover, a significant account of mathematical properties of the new distribution is provided.

The inferential procedures of the parameters of the GLW distribution are presented considering the maximum likelihood estimators (MLE). Finally, we analyze two data sets for illustrative purposes, proving that the GWL outperform several usual three parameters lifetime distributions such as the generalized gamma distribution, the generalized Weibull (GW) distribution (MUDHOLKAR; SRIVASTAVA; KOLLIA, 1996), the generalized exponential-Poisson (GEP) distribution (BARRETO-SOUZA; CRIBARI-NETO, 2009) and the exponentiated Weibull (EP) distribution (MUDHOLKAR; SRIVASTAVA; FREIMER, 1995).

The chapter is organized as follows. In Section 2, we provide a significant account of mathematical properties of the new distribution. In Section 3, we present the maximum likelihood estimators. In Section 4 a simulation study is presented in order to verify the effiency of the MLES. In Section 5 the methodology is illustrated in two real data sets. Some final comments are presented in Section 6.

7.2 Generalized Weighted Lindley distribution

The cumulative distribution function from the GWL distribution is given by

$$F(t|\phi,\lambda,\alpha) = \frac{\gamma[\phi,(\lambda t)^{\alpha}](\lambda+\phi) - (\lambda t)^{\alpha\phi} e^{-(\lambda t)^{\alpha}}}{(\lambda+\phi)\Gamma(\phi)}, \tag{7.2}$$

The behaviors of the PDF (7.1) when $t \to 0$ and $t \to \infty$ are, respectively, given by

$$f(0) = \begin{cases} \infty, & \text{if } \alpha\phi < 1 \\ \dfrac{\alpha\lambda^2}{(\lambda+\phi)\Gamma(\phi)}, & \text{if } \alpha\phi = 1\,, \\ 0, & \text{if } \alpha\phi > 1 \end{cases} \qquad f(\infty) = 0.$$

Figure 16 gives examples of the shapes of the density function for different values of ϕ, λ and α.

7.2.1 Moments

Many important features and properties related to a distribution can be obtained through its moments, such as mean, variance, kurtosis and skewness. In this section, we present some important moments, such as the moment generating function, r-th moment, r-th central moment among others.

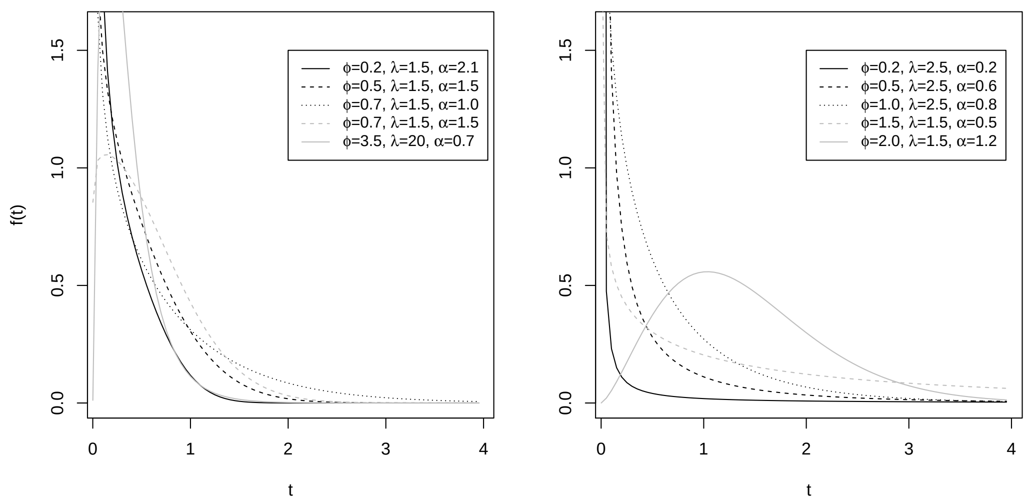

Figure 16 – Density function shapes for GWL distribution considering different values of ϕ, λ and α.

Theorem 7.2.1. For the random variable T with GWL distribution, the moment generating function is given by

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{\lambda^r r!} \frac{\left(\frac{r}{\alpha} + \phi + \lambda\right) \Gamma(\frac{r}{\alpha} + \phi)}{(\lambda + \phi)\Gamma(\phi)}. \tag{7.3}$$

Proof. Note that, the moment generating function from GG distribution is given by

$$M_{X,j}(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \frac{\Gamma(\frac{r}{\alpha} + \phi + j - 1)}{\lambda^r \Gamma(\phi + j - 1)}.$$

Therefore, as the GWL (7.1) distribution can be expressed as a two-component mixture, we have

$$\begin{aligned}
M_X(t) = E\left[e^{tX}\right] &= \int_0^{\infty} e^{tx} f(x|\phi,\lambda,\alpha)dx = pM_{X,1}(t) + (1-p)M_{X,2}(t) \\
&= \frac{\lambda}{(\lambda+\phi)} \sum_{r=0}^{\infty} \frac{t^r}{r!} \frac{\Gamma(\frac{r}{\alpha}+\phi)}{\lambda^r\Gamma(\phi)} + \frac{\phi}{(\lambda+\phi)} \sum_{r=0}^{\infty} \frac{t^r}{r!} \frac{\Gamma(\frac{r}{\alpha}+\phi+1)}{\lambda^r\Gamma(\phi+1)} \\
&= \frac{1}{(\lambda+\phi)} \sum_{r=0}^{\infty} \frac{t^r}{r!} \frac{\lambda\Gamma(\frac{r}{\alpha}+\phi)}{\lambda^r\Gamma(\phi)} + \frac{1}{(\lambda+\phi)} \sum_{r=0}^{\infty} \frac{t^r}{r!} \frac{\left(\frac{r}{\alpha}+\phi\right)\Gamma(\frac{r}{\alpha}+\phi)}{\lambda^r\Gamma(\phi)} \\
&= \sum_{r=0}^{\infty} \frac{t^r}{\lambda^r r!} \frac{\left(\frac{r}{\alpha}+\phi+\lambda\right)\Gamma(\frac{r}{\alpha}+\phi)}{(\lambda+\phi)\Gamma(\phi)}.
\end{aligned}$$

□

Corollary 7.2.2. For the random variable T with GWL distribution, the r-th moment is given by

$$\mu_r = E[T^r] = \frac{\left(\frac{r}{\alpha} + \phi + \lambda\right) \Gamma(\frac{r}{\alpha} + \phi)}{(\lambda + \phi)\lambda^r \Gamma(\phi)}. \tag{7.4}$$

Proof. From the literature, $\mu_r = M_X^{(r)}(0) = \frac{d^n M_X(0)}{dt^n}$ and the result follows. □

Corollary 7.2.3. For the random variable T with GWL distribution, the r-th central moment is given by

$$\begin{aligned} M_r = E[T-\mu]^r &= \sum_{i=0}^{r}\binom{r}{i}(-\mu)^{r-i}E[T^i] \\ &= \sum_{i=0}^{r}\binom{r}{i}\left(-\frac{\left(\frac{1}{\alpha}+\phi+\lambda\right)\Gamma\left(\frac{1}{\alpha}+\phi\right)}{\lambda(\lambda+\phi)\Gamma(\phi)}\right)^{r-i}\frac{\left(\frac{i}{\alpha}+\phi+\lambda\right)\Gamma(\frac{i}{\alpha}+\phi)}{(\lambda+\phi)\lambda^i\Gamma(\phi)}. \end{aligned} \tag{7.5}$$

Corollary 7.2.4. A random variable T with GWL distribution, has the mean and variance given by

$$\mu = \frac{\left(\frac{1}{\alpha}+\phi+\lambda\right)\Gamma\left(\frac{1}{\alpha}+\phi\right)}{\lambda(\lambda+\phi)\Gamma(\phi)}, \tag{7.6}$$

$$\sigma^2 = \frac{\lambda(\lambda+\phi)\left(\frac{2}{\alpha}+\phi+\lambda\right)\Gamma\left(\frac{2}{\alpha}+\phi\right)-\left(\frac{1}{\alpha}+\phi+\lambda\right)^2\Gamma\left(\frac{1}{\alpha}+\phi\right)^2}{\lambda^2(\lambda+\phi)^2\Gamma(\phi)^2}. \tag{7.7}$$

Proof. From (7.4) and considering $r=1$ follows $\mu_1=\mu$. The second result follows from (7.5) considering $r=2$ and with some algebra follow the results. □

Different type of moments can be easily achieved for GWL distribution, one in particular, that has play a important role in information theory is given by

$$E[\log(T)] = \frac{\left(\psi(\phi)-\alpha\log\lambda+(\lambda+\phi)^{-1}\right)}{\alpha}. \tag{7.8}$$

7.2.2 Survival Properties

In this section, we present the survival, the hazard and mean residual life function for the GWL distribution. The survival function of $T\sim \mathrm{GWL}(\phi,\lambda,\alpha)$ with the probability of an observation does not fail until a specified time t is

$$S(t|\phi,\lambda,\alpha) = \frac{\Gamma[\phi,(\lambda t)^\alpha](\lambda+\phi)+(\lambda t)^{\alpha\phi}e^{-(\lambda t)^\alpha}}{(\lambda+\phi)\Gamma(\phi)}\,. \tag{7.9}$$

The hazard function is given by

$$h(t|\phi,\lambda,\alpha) = \frac{f(t|\phi,\lambda,\alpha)}{S(t|\phi,\lambda,\alpha)} = \frac{\alpha\lambda^{\alpha\phi}t^{\alpha\phi-1}(\lambda+(\lambda t)^\alpha)e^{-(\lambda t)^\alpha}}{\Gamma[\phi,(\lambda t)^\alpha](\lambda+\phi)+(\lambda t)^{\alpha\phi}e^{-(\lambda t)^\alpha}}. \tag{7.10}$$

The behaviors of the hazard function (7.10) when $t\to 0$ and $t\to\infty$are

$$h(0) = \begin{cases}\infty, & \text{if } \alpha\phi<1 \\ \dfrac{\alpha\lambda^2}{(\lambda+\phi)\Gamma(\phi)}, & \text{if } \alpha\phi=1 \\ 0, & \text{if } \alpha\phi>1\end{cases} \quad \text{and} \quad h(\infty) = \begin{cases}0, & \text{if } \alpha\phi<1 \\ \lambda, & \text{if } \alpha\phi=1 \\ \infty, & \text{if } \alpha\phi>1.\end{cases}$$

Theorem 7.2.5. The hazard rate function $h(t)$ of the generalized weighted Lindley distribution has increasing, decreasing, bathtub, unimodal or decreasing-increasing-decreasing shape.

Proof. It is not simple to apply the Glaser's theorem in the GLW distribution. As the hazard rate function (7.10) is complex, we consider the following cases:

1. Let $\alpha = 1$, then GWL distribution reduces to the WL distribution. In this case, Ghitany et al. (2008) proved that the hazard function has bathtub (increasing) shape if $0 < \phi < 1$ $(\phi > 0)$, for all $\lambda > 0$.

2. Let $\phi = 1$, then GWL distribution reduces to the PL distribution. In this case, considering $\beta = \lambda^{\alpha}$, Ghitany *et al.* (2013) proved that the hazard function is

   - increasing if $\{0 < \alpha \geq 1, \beta > 0\}$;
   - decreasing if $\left\{0 < \alpha \leq \frac{1}{2}, \beta > 0\right\}$ or $\left\{\frac{1}{2} < \alpha < 1, \beta \geq (2\alpha-1)^2(4\alpha(1-\alpha))^{-1}\right\}$;
   - decreasing-increasing-decreasing if $\left\{\frac{1}{2} < \alpha < 1, 0 < \beta < (2\alpha-1)^2(4\alpha(1-\alpha))^{-1}\right\}$.

3. Let $\alpha = 2$ and $\lambda = 1$, from Glaser's theorem, the hazard rate function has decreasing (unimodal) shape if $0 < \phi < 1$ $(\phi > 1)$.

□

These properties make the GWL distribution a flexible model for reliability data. Figure 17 gives examples from the shapes of the hazard function for different values of ϕ, λ and α.

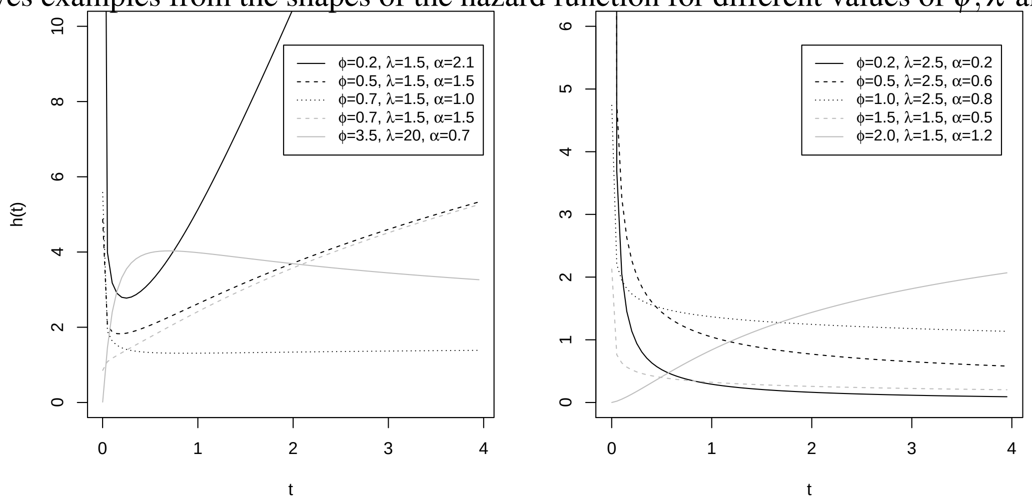

Figure 17 – Hazard function shapes for GWL distribution and considering different values of ϕ, λ and α

The mean residual life (MRL) has been used widely in survival analysis and represents the expected additional lifetime given that a component has survived until time t, the following result presents the MRL function of the GWL distribution

Proposition 7.2.6. The mean residual life function $r(t|\phi,\lambda,\alpha)$ of the GWL distribution is given by

$$r(t|\phi,\lambda,\alpha)=\frac{\left(\phi+\frac{1}{\alpha}+\lambda\right)\Gamma\left(\phi+\frac{1}{\alpha},(\lambda t)^{\alpha}\right)-\lambda t(\lambda+\phi)\Gamma\left(\phi,(\lambda t)^{\alpha}\right)}{\lambda[(\lambda+\phi)\Gamma(\phi,(\lambda t)^{\alpha})+(\lambda t)^{\alpha\phi}e^{-(\lambda t)^{\alpha}}]}. \tag{7.11}$$

Proof. Note that

$$\begin{aligned}
r(t|\phi,\lambda,\alpha) &= \frac{1}{S(t)}\int_t^{\infty} yf(y|\lambda,\phi)dy-t \\
&= \frac{1}{S(t)}\left[p\int_t^{\infty} yf_1(y|\lambda,\phi)dy+(1-p)\int_x^{\infty} yf_2(y|\lambda,\phi)dy\right]-t \\
&= \frac{\left(\phi+\frac{1}{\alpha}+\lambda\right)\Gamma\left(\phi+\frac{1}{\alpha},(\lambda t)^{\alpha}\right)-\lambda t(\lambda+\phi)\Gamma\left(\phi,(\lambda t)^{\alpha}\right)}{\lambda[(\lambda+\phi)\Gamma(\phi,(\lambda t)^{\alpha})+(\lambda t)^{\alpha\phi}e^{-(\lambda t)^{\alpha}}]}.
\end{aligned}$$

□

The behaviors of the mean residual life function when $t\to 0$ and $t\to\infty$, respectively, are given by

$$r(0)=\frac{1}{\lambda\left((\lambda+\phi)\Gamma(\phi)\right)} \quad \text{and} \quad r(\infty)\begin{cases}\infty, & \text{if } \alpha<1\\ \dfrac{1}{\lambda}, & \text{if } \alpha=1\\ 0, & \text{if } \alpha>1\end{cases}.$$

7.2.3 Entropy

In information theory, entropy has played a central role as a measure of the uncertainty associated with a random variable. Proposed by Shannon (2001), Shannon's entropy is one of the most important metrics in information theory. Shannon's entropy for the GWL distribution can be obtained by solving the following equation

$$H_S(\phi,\lambda,\alpha)=-\int_0^{\infty}\log\left(\frac{\alpha\lambda^{\alpha\phi}t^{\alpha\phi-1}(\lambda+(\lambda t)^{\alpha})e^{-(\lambda t)^{\alpha}}}{(\lambda+\phi)\Gamma(\phi)}\right)f(t|\phi,\lambda,\alpha)dt. \tag{7.12}$$

Proposition 7.2.7. A random variable T with GWL distribution, has Shannon's Entropy given by

$$\begin{aligned}
H_S(\phi,\lambda,\alpha)= &\log(\lambda+\phi)+\log\Gamma(\phi)-\log\alpha-\log\lambda-\frac{\phi(1+\phi+\lambda)}{(\lambda+\phi)} \\
&-\frac{\psi(\phi)(\alpha\phi-1)}{\alpha}-\frac{(\alpha\phi-1)}{\alpha(\lambda+\phi)}-\frac{\eta(\phi,\lambda)}{(\lambda+\phi)\Gamma(\phi)}.
\end{aligned} \tag{7.13}$$

where

$$\eta(\phi,\lambda)=\int_0^{\infty}(\lambda+y)\log(\lambda+y)y^{\phi-1}e^{-y}dy=\int_0^1(\lambda-\log u)\log(\lambda-\log u)(-\log u)^{\phi-1}du.$$

Proof. From the equation (7.12) we have

$$\begin{aligned} H_S(\phi,\lambda,\alpha) = & -\log\alpha - \alpha\phi\log\lambda + \log(\lambda+\phi) + \log(\Gamma(\phi)) + \lambda^{\alpha} E[T^{\alpha}] \\ & -(\alpha\phi-1)E[\log T] - E\left[\log(\lambda+(\lambda T)^{\alpha})\right] \end{aligned} \tag{7.14}$$

Note that

$$E\left[\log(\lambda+(\lambda T)^{\alpha})\right] = \int_0^{\infty} \log(\lambda+(\lambda T)^{\alpha} \frac{\alpha\lambda^{\alpha\phi} t^{\alpha\phi-1}(\lambda+(\lambda t)^{\alpha})e^{-(\lambda t)^{\alpha}}}{(\lambda+\phi)\Gamma(\phi)} dt,$$

using the change of variable $y=(\lambda t)^{\alpha}$ and after some algebra

$$\begin{aligned} E\left[\log(\lambda+(\lambda T)^{\alpha})\right] &= \frac{1}{(\lambda+\phi)\Gamma(\phi)} \int_0^{\infty} (\lambda+y)\log(\lambda+y) y^{\phi-1} e^{-y} dy \\ &= \frac{\eta(\phi,\lambda)}{(\lambda+\phi)\Gamma(\phi)}. \end{aligned}$$

From equations (7.4) and (7.8), we can easily find the solution of $E[T^{\alpha}]$ and $E[\log T]$ and the result follows. □

Other popular entropy measure is proposed by Renyi (1961). Some recent applications of the Renyi entropy can be seen in Popescu and Aiordachioaie (2013). If T has the probability density function (1) then Renyi entropy is defined by

$$\frac{1}{1-\rho} \log \int_0^{\infty} f^{\rho}(x) dx. \tag{7.15}$$

Proposition 7.2.8. A random variable T with GWL distribution, has the Renyi entropy given by

$$H_R(\rho) = \frac{(\rho-1)(\log\alpha+\log\lambda) - \rho\left(\log(\lambda+\phi)+\log\Gamma(\phi)\right) - \log(\delta(\rho,\phi,\lambda,\alpha))}{1-\rho} \tag{7.16}$$

where $\delta(\rho,\phi,\lambda,\alpha) = \int_0^{\infty} y^{\frac{\rho\phi-\rho+1-\alpha}{\alpha}} (\lambda+y)^{\rho} e^{-\rho y} dy$.

Proof. The Renyi entropy is given by

$$\begin{aligned} H_R(\rho) &= \frac{1}{1-\rho} \log\left(\frac{\alpha^{\rho}\lambda^{\rho}}{(\lambda+\phi)^{\rho}\Gamma(\phi)^{\rho}} \int_0^{\infty} (\lambda t)^{\alpha\rho\left(\phi-\frac{1}{\alpha}\right)} \left(\lambda+(\lambda t)^{\alpha}\right)^{\rho} e^{-\rho(\lambda t)^{\alpha}} dt \right) \\ &= \frac{1}{1-\rho} \log\left(\frac{\alpha^{\rho}\lambda^{\rho}}{(\lambda+\phi)^{\rho}\Gamma(\phi)^{\rho}} \int_0^{\infty} y^{\frac{\rho\phi-\rho+1-\alpha}{\alpha}} (\lambda+y)^{\rho} e^{-\rho y} dy \right) \\ &= \frac{1}{1-\rho} \log\left(\frac{\alpha^{\rho}\lambda^{\rho}}{(\lambda+\phi)^{\rho}\Gamma(\phi)^{\rho}} \delta(\rho,\phi,\lambda,\alpha) \right) \end{aligned}$$

and with some algebra the proof is completed. □

7.2.4 Lorenz curves

The Lorenz curve (LORENZ, 1905) is a well-known measure used in reliability, income inequality, life testing and renewal theory. The Lorenz curve for a non-negative T random variable is given by the consecutive plot of

$$L(F(t)) = \frac{\int_0^t xf(x)dx}{\int_0^\infty xf(x)dx} = \frac{1}{\mu}\int_0^t xf(x)dx.$$

The Lorenz curve of the GWL distribution is

$$L(p) = \frac{\left(\frac{1}{\alpha}+\phi+\lambda\right)\gamma\left(\phi+\frac{1}{\alpha},\left(\lambda F^{-1}(p)\right)^{\alpha}\right) - \left(\lambda F^{-1}(p)\right)^{\alpha\phi-1} e^{-\left(\lambda F^{-1}(p)\right)^{\alpha}}}{\left(\frac{1}{\alpha}+\phi+\lambda\right)\Gamma\left[\frac{1}{\alpha}+\phi\right]}$$

where $F^{-1}(p) = t_p$.

7.3 Maximum likelihood estimators

Among the statistical inference methods, the maximum likelihood method is widely used due its better asymptotic properties. Let $T_1,\ldots,T_n$ be a random sample such that $T \sim$ GWL(ϕ,λ,α). In this case, the likelihood function from (7.1) is given by,

$$L(\phi,\lambda,\alpha;t) = \frac{\alpha^n\lambda^{n\alpha\phi}}{(\lambda+\phi)\Gamma(\phi)^n}\left\{\prod_{i=1}^n t_i^{\alpha\phi-1}\right\}\prod_{i=1}^n(\lambda+(\lambda t_i)^{\alpha})\exp\left\{-\lambda^{\alpha}\sum_{i=1}^n t_i^{\alpha}\right\}. \tag{7.17}$$

The log-likelihood function $l(\phi,\lambda,\alpha;t) = \log L(\phi,\lambda,\alpha;t)$ is given by,

$$\begin{aligned} l(\phi,\lambda,\alpha;t) = {} & n\log\alpha + n\alpha\phi\log\lambda - n\log(\lambda+\phi) - n\log\Gamma(\phi) + (\alpha\phi-1)\sum_{i=1}^n \log(t_i) \\ & + \sum_{i=1}^n \log\left(\lambda+(\lambda t_i)^{\alpha}\right) - \lambda^{\alpha}\sum_{i=1}^n t_i^{\alpha}. \end{aligned} \tag{7.18}$$

From the expressions $\frac{\partial}{\partial\phi}l(\phi,\lambda,\alpha;t) = 0$, $\frac{\partial}{\partial\lambda}l(\phi,\lambda,\alpha;t) = 0$, $\frac{\partial}{\partial\alpha}l(\phi,\lambda,\alpha;t) = 0$, we get the likelihood equations

$$n\hat{\alpha}\log(\hat{\lambda}) + \hat{\alpha}\sum_{i=1}^n \log(t_i) = \frac{n}{\hat{\lambda}+\hat{\phi}} + n\psi(\hat{\phi}) \tag{7.19}$$

$$\frac{n\hat{\alpha}\hat{\phi}}{\hat{\lambda}} + \sum_{i=1}^n \frac{1+\hat{\alpha}\hat{\lambda}^{\hat{\alpha}-1}t_i^{\hat{\alpha}}}{\hat{\lambda}+(t_i)^{\hat{\alpha}}} = \hat{\alpha}\hat{\lambda}^{\hat{\alpha}-1}\sum_{i=1}^n t_i^{\hat{\alpha}} + \frac{n}{\hat{\lambda}+\hat{\phi}} \tag{7.20}$$

$$\frac{n}{\hat{\alpha}} + n\hat{\phi}\log(\hat{\lambda}) + \hat{\phi}\sum_{i=1}^n \log(t_i) + \sum_{i=1}^n \frac{(\hat{\lambda}t_i)^{\hat{\alpha}}\log(\hat{\lambda}t_i)}{\hat{\lambda}+(\hat{\lambda}t_i)^{\hat{\alpha}}} = \hat{\lambda}^{\hat{\alpha}}\sum_{i=1}^n t_i^{\hat{\alpha}}\log(\hat{\lambda}t_i). \tag{7.21}$$

The solutions of such non-linear system provide the maximum likelihood estimates. Numerical methods such as Newton-Rapshon are required to find the solution of the nonlinear system. Note that from (7.19) and (7.21) and after some algebra we have

$$\hat{\alpha} = \frac{1}{\left(n\log(\hat{\lambda}) + \sum_{i=1}^{n}\log(t_i)\right)}\left(\frac{n}{\hat{\lambda}+\hat{\phi}} + n\psi(\hat{\phi})\right) \tag{7.22}$$

$$\hat{\phi} = \frac{\left(\hat{\lambda}^{\hat{\alpha}}\sum_{i=1}^{n} t_i{}^{\hat{\alpha}}\log(\hat{\lambda}t_i) - \sum_{i=1}^{n}\frac{(\hat{\lambda}t_i)^{\hat{\alpha}}\log(\hat{\lambda}t_i)}{\hat{\lambda}+(\hat{\lambda}t_i)^{\hat{\alpha}}} - \frac{n}{\hat{\alpha}}\right)}{\left(n\log(\hat{\lambda}) + \sum_{i=1}^{n}\log(t_i)\right)}. \tag{7.23}$$

Under mild conditions, for large sample sizes, the obtained estimators are not biased and they are asymptotically efficient with an asymptotically joint multivariate normal distribution given by

$$(\hat{\phi},\hat{\lambda},\hat{\alpha}) \sim N_3[(\phi,\lambda,\alpha), I^{-1}(\phi,\lambda,\alpha)] \text{ for } n\to\infty, \tag{7.24}$$

where $I(\phi,\lambda,\alpha)$ is the Fisher information matrix given by,

$$I(\phi,\lambda,\alpha) = \begin{bmatrix} I_{\phi,\phi}(\phi,\lambda,\alpha) & I_{\phi,\lambda}(\phi,\lambda,\alpha) & I_{\phi,\alpha}(\phi,\lambda,\alpha) \\ I_{\phi,\lambda}(\phi,\lambda,\alpha) & I_{\lambda,\lambda}(\phi,\lambda,\alpha) & I_{\lambda,\alpha}(\phi,\lambda,\alpha) \\ I_{\phi,\alpha}(\phi,\lambda,\alpha) & I_{\lambda,\alpha}(\phi,\lambda,\alpha) & I_{\alpha,\alpha}(\phi,\lambda,\alpha) \end{bmatrix}, \tag{7.25}$$

where the elements of the matrix are given by

$$\begin{aligned}
I_{\phi,\phi}(\phi,\lambda,\alpha) &= -E\left[\frac{\partial l(\theta;t)}{\partial\phi^2}\right] = -\frac{1}{(\lambda+\phi)^2} + \psi'(\theta) \\
I_{\phi,\lambda}(\phi,\lambda,\alpha) &= -E\left[\frac{\partial l(\theta;t)}{\partial\phi\partial\lambda}\right] = -\frac{\alpha}{\lambda} + \frac{1}{(\lambda+\phi)^2} \\
I_{\phi,\alpha}(\phi,\lambda,\alpha) &= -E\left[\frac{\partial l(\theta;t)}{\partial\phi\partial\alpha}\right] = \frac{-\alpha\log(\lambda) - \psi(\phi) + \alpha\log(\lambda) - (\lambda+\phi)^{-1}}{\alpha} \\
I_{\lambda,\lambda}(\phi,\lambda,\alpha) &= -E\left[\frac{\partial l(\theta;t)}{\partial\lambda^2}\right] = \frac{\alpha\phi}{\lambda^2} + (\alpha-1)\lambda^{\alpha-2}(\psi(\phi) - \alpha\log(\lambda) + (\lambda+\phi)^{-1}) \\
&\quad + E\left[\frac{\alpha T^{\alpha}\lambda^{\alpha-2}((\alpha-2)\lambda - (\lambda T)^{\alpha})}{(\lambda+(\lambda T)^{\alpha})}\right] - \frac{1}{(\lambda+\phi)^2} \\
I_{\alpha,\alpha}(\phi,\lambda,\alpha) &= -E\left[\frac{\partial l(\theta;t)}{\partial\alpha^2}\right] = \frac{\phi(\lambda+\phi+1)\left(\psi(\phi)^2 + \psi(\phi)\right)}{\alpha^2(\lambda+\phi)} + \frac{1}{\alpha^2} \\
&\quad + \frac{2(\lambda+2\phi+1)\psi(\phi)+2}{\alpha^2(\lambda+\phi)} - E\left[\frac{\lambda(\lambda T)^{\alpha}\log(\lambda T)^2}{(\lambda+(\lambda T)^{\alpha})}\right] \\
I_{\alpha,\lambda}(\phi,\lambda,\alpha) &= -E\left[\frac{\partial l(\theta;t)}{\partial\alpha\partial\lambda}\right] = -\frac{\phi}{\lambda} + \frac{\lambda\left(1+\phi\psi(\phi)\right) + \phi\left(1+(\phi+1)\psi(\phi+1)\right)}{\lambda(\lambda+\phi)} \\
&\quad - E\left[\frac{\left(1+\alpha\lambda^{\alpha-1}T^{\alpha}\right)(\lambda T)^{\alpha}\log(\lambda T)}{(\lambda+(\lambda T)^{\alpha})^2}\right] + \frac{\left(\phi+\lambda+1-\frac{1}{\alpha}\right)\Gamma\left(\phi+1-\frac{1}{\alpha}\right)}{(\lambda+\phi)\Gamma(\phi)} \\
&\quad - E\left[\frac{\left(\alpha\lambda^{\alpha-1}T^{\alpha}\log(\lambda T) + (\lambda T)^{\alpha-1}\right)}{(\lambda+(\lambda T)^{\alpha})}\right].
\end{aligned}$$

7.4 Simulation Study

In this section, we present an intensive simulation study to compare the efficiency of our estimation procedure. The mean relative estimates (MRE) and the mean square errors (MSE) were computed in which are given by

$$\frac{1}{N}\sum_{j=1}^{N}\frac{\hat{\theta}_{i,j}}{\theta_i}, \quad \frac{1}{N}\sum_{j=1}^{N}(\hat{\theta}_{i,j}-\theta_i)^2, \quad \text{for} \quad i=1,2,3.$$

The results were computed using the software R using the seed 2015 to generate the pseudo-random values. The chosen values to perform this procedure were $N = 10,000$ and $n = (50,60,\ldots,300)$. We presented the results only for $\theta = (2,0.5,0.1)$ for reasons of space. However, the following results were similar for other choices of θ.

Figures 18 presents MREs, MSEs from the estimates of ϕ,λ and α obtained using the MLE for N simulated samples and considering different values of $\theta = (2,0.5,0.1)$ and n. The horizontal lines in both figures corresponds to MREs and MSEs being respectively one and zero.

Based on these results, we observe that the MSE of the MLEs tend to zero for large n and also, as expected, the values of MREs tend to one, i.e. the estimates are consistent and asymptotically unbiased for the parameters. Therefore, the MLEs can be easily used for estimating the parameters of the GWL distribution.

7.5 Application

In this section, we compare the GWL distribution fit with several usual three parameters lifetime distributions considering two data sets one with bathtub hazard rate and one with the increasing hazard function. For sake of comparison the following lifetime distributions were considered, the generalized gamma (GG) distribution, the generalized Weibull (GW) distribution with PDF given by $f(t) = (\alpha\phi)^{-1}(t/\phi)^{1/\alpha-1}\left(1-\lambda(t/\phi)^{1/\alpha}\right)^{1/\lambda-1}$, where $\lambda\in\mathbb{R}$ the generalized exponential-Poisson (GEP) distribution with PDF given by $f(t) = \left(\alpha\beta\phi/(1-e^{-\phi})^{\alpha}\right)e^{-\phi-\beta t+\phi\exp(-\beta t)}\left(1-e^{-\phi+\phi\exp(-\beta t)}\right)^{\alpha-1}$ and the exponentiated Weibull (EW) distribution with PDF $f(t) = \alpha\phi(t/\beta)^{\alpha-1}\exp\left(-(t/\beta)^{\alpha}\right)\left(1-\exp\left(-(t/\beta)^{\alpha}\right)\right)^{\phi-1}/\beta$.

Firstly, we considered the TTT-plot (total time on test) in order to verify the behavior of the empirical hazard function. To check the goodness of fit we also considered the Kolmogorov-Smirnov (KS) test . This procedure is based on the KS statistic $D_n = \sup|F_n(t)-F(t;\phi,\lambda,\alpha)|$, where $\sup t$ is the supremum of the set of distances, $F_n(t)$ is the empirical distribution function and $F(t;\alpha,\beta,\lambda)$ is c.d.f. In this case, we test the null hypothesis that the data comes from $F(t;\alpha,\beta,\lambda)$ with significance level of 5%, the null hypothesis we will be rejected if the returned p-value is smaller than 0.05.

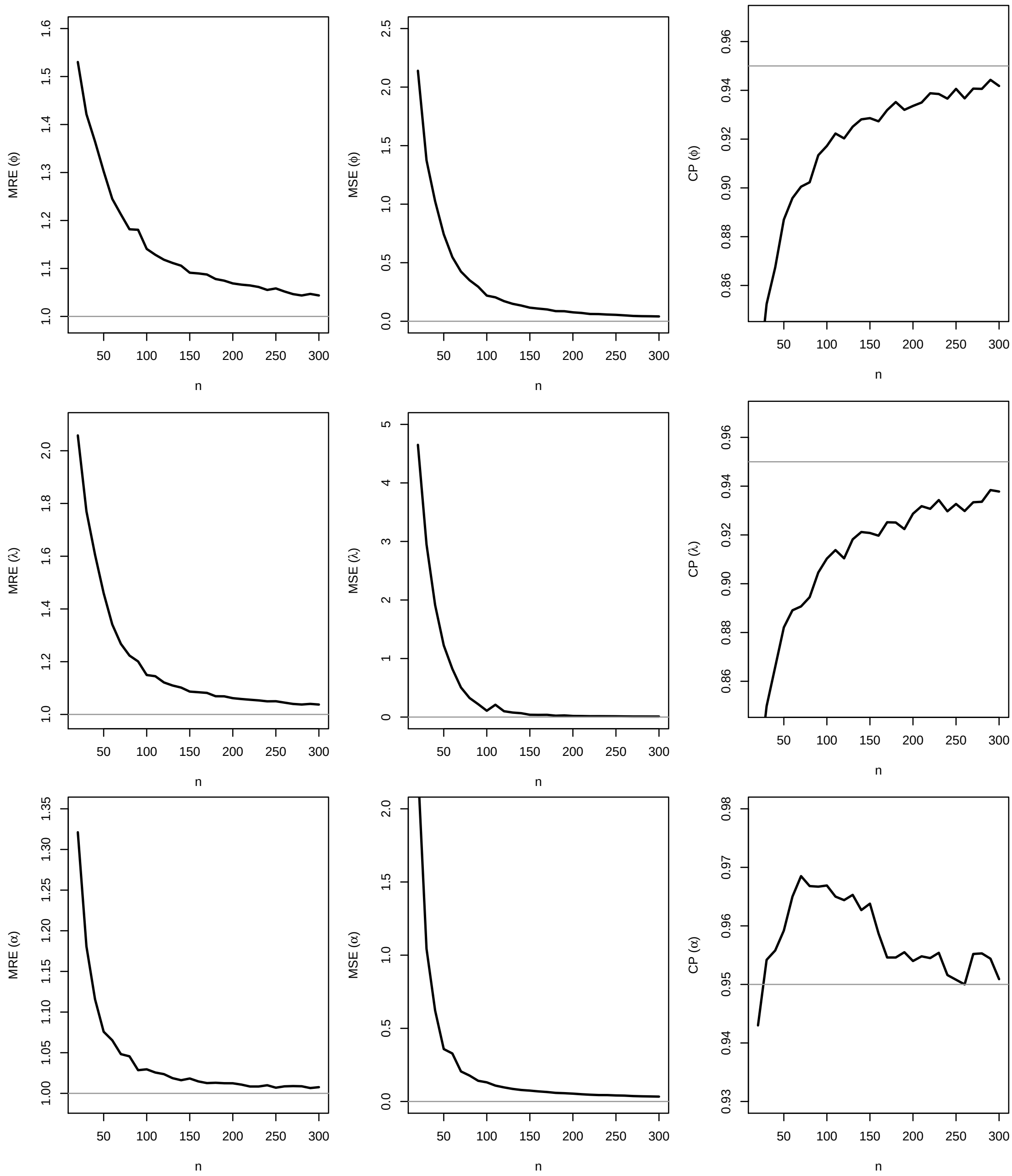

Figure 18 – MRE's, MSE's related from the estimates of $\phi = 0.5, \lambda = 0.7$ and $\alpha = 1.5$ for N simulated samples, considering different values of n obtained using the following estimation method 1-MLE, 2-MPS, 3-ADE, 4-RTADE.

7.5.1 Lifetimes data

Presented by Aarset (1987) the dataset is related to the lifetime in hours of 50 devices put on test (see Table 21).

Figure 20 presets (left panel) the TTT-plot, (middle panel) the fitted survival superimposed to the empirical survival function and (right panels) the hazard function adjusted by GWL distribution. Table 22 presents the AIC and AICc criteria and the p-value from the KS test for all fitted distributions considering the Aarset dataset.

Table 21 – Lifetimes data (in hours) related to a device put on test.

0.1	0.2	1	1	1	1	1	2	3	6	7	11	12
18	18	18	18	18	21	32	36	40	45	46	47	50
55	60	63	63	67	67	67	67	72	15	79	82	82
83	84	84	84	85	85	85	85	85	86	86		

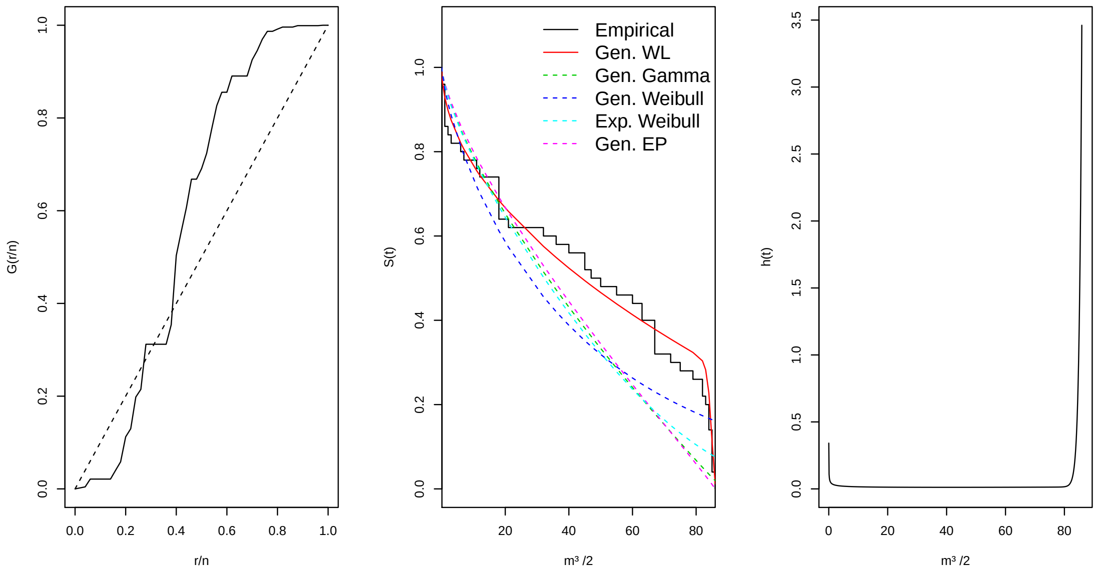

Figure 19 – (left panel) the TTT-plot, (middle panel) the fitted survival superimposed to the empirical survival function and (right panels) the hazard function adjusted by GWL distribution

Table 22 – Results of AIC and AICc criteria and the p-value from the KS test for all fitted distributions considering the Aarset dataset.

Criteria	Gen. WL	Gen. Gamma	Gen. Weibull	Exp. Weibull	Gen. EP
AIC	**415.400**	448.294	430.055	463.674	486.255
AICc	**409.922**	442.816	424.576	458.196	480.777
KS	**0.8343**	0.0115	0.0453	0.0222	0.0302

Comparing the empirical survival function with the adjusted distributions it can be observed a better fit for the GWL distribution among the chosen models. These result is confirmed from AIC and AICC since GWL distribution has the minimum values and the p-values returned from the KS test are greater than 0.05. It is worth mentioning that, considering a significance level of 5%, the others models are not able to fit the proposed data.

Table 23 displays the MLE estimates, standard errors and the confidence intervals for ϕ, λ and α of the GWL distribution.

Table 23 – MLE estimates, Standard-error and 95% confidence intervals (CI) for ϕ, λ and α

θ	$\hat{\theta}_{MLE}$	S.E($\hat{\theta}$)	CI$_{95\%}(\theta)$
ϕ	0.0050	5.69145e-07	(0.005046; 0.005048)
λ	0.0118	1.13547e-09	(0.011762; 0.011762)
α	102.0427	5.97761	(90.326; 113.758)

7.5.2 Average flows data

The study of average flows has been proved of high importance to protect and maintain aquatic resources in streams and rivers (Reiser et al., 1989). In this section, we consider a real data set related to the average flows (m^3/s) of the Cantareira system during January at São Paulo city in Brazil. Its worth mentioning that the Cantareira system provides water to 9 million people in the São Paulo metropolitan area. The data set available in Table 24 was obtained from the website of the National Water Agency including a period from 1930 to 2012.

Table 24 – January average flows (m^3/s) of the Cantareira system.

82.0	80.9	102.5	65.3	65.5	47.1	53.0	139.4	82.4	80.2	92.5
50.0	50.4	50.2	36.2	35.9	100.0	94.2	78.1	54.8	86.9	80.1
60.3	26.9	48.5	51.0	51.1	84.5	76.9	69.4	77.3	109.2	55.3
106.3	30.5	94.2	87.3	115.0	70.0	31.3	87.1	35.9	67.7	55.1
89.9	50.1	52.6	82.0	54.1	44.3	69.2	94.4	83.4	122.7	88.1
73.3	35.9	82.4	64.9	90.8	80.4	55.3	31.4	45.7	43.6	45.8
96.8	85.8	43.6	122.3	66.5	41.0	75.4	79.4	34.8	78.8	52.4
77.1	47.0	67.4	132.8	144.9	64.1					

Figure 20 shows (left panel) the TTT-plot, (middle panel) the fitted survival superimposed to the empirical survival function and (right panels) the hazard function adjusted by GWL distribution. Table 25 presents the AIC and AICc criteria and the p-value from the KS test for all fitted distributions considering the data set related to the January average flows (m^3/s) of the Cantareira system.

Table 25 – Results of AIC and AICc criteria and the p-value from the KS test for all fitted distributions considering the data set related to the january average flows (m^3/s) of the Cantareira system.

Criteria	Gen. WL	Gen. Gamma	Gen. Weibull	Exp. Weibull	Gen. EP
AIC	**775.431**	775.461	777.280	780.304	778.873
AICc	**769.735**	769.765	771.584	774.608	773.176
KS	**0.4683**	0.4223	0.3935	0.1654	0.4599

Comparing the empirical survival function with the adjusted distributions it can be observed a better fit for the GWL distribution among the chosen models. This result is confirmed

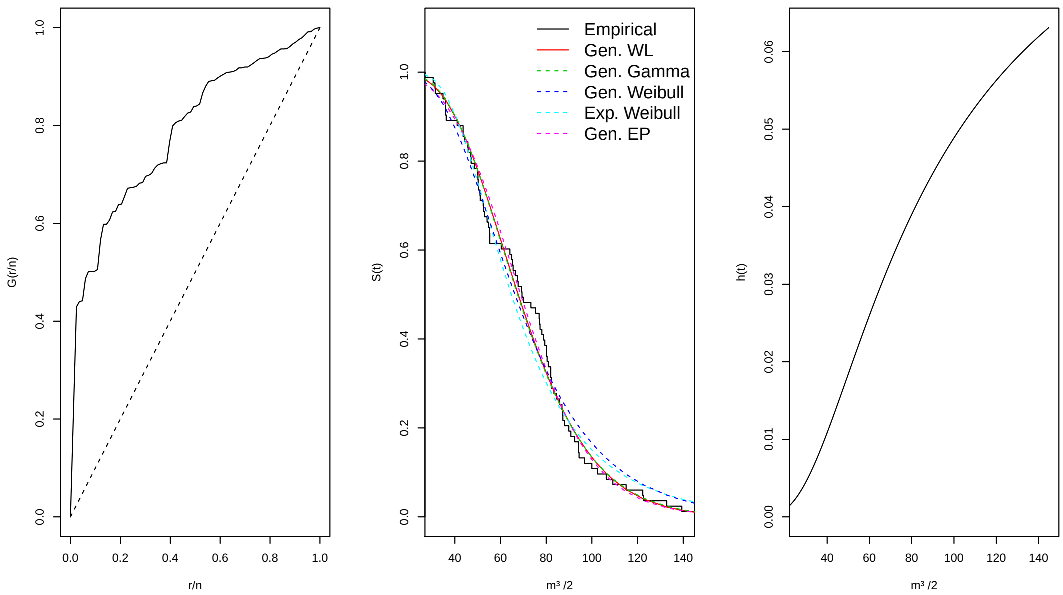

Figure 20 – (left panel) the TTT-plot, (middle panel) the fitted survival superimposed to the empirical survival function and (right panels) the hazard function adjusted by GWL distribution

from AIC and AICC since GWL distribution has the minimum values and also the p-values returned from the KS test are greater than 0.05. Table 26 displays the ML estimates, standard errors and the confidence intervals for ϕ, λ and α of the GWL distribution.

Table 26 – ML estimates, Standard-error and 95% confidence intervals (CI) for ϕ, λ and α

θ	$\hat{\theta}_{MLE}$	S.E($\hat{\theta}$)	CI$_{95\%}(\theta)$
ϕ	7.0485	1.5425	(2.3847; 11.7124)
λ	0.1244	0.0557	(0.1183; 0.1305)
α	0.9579	0.1173	(0.9310; 0.9849)

7.6 Concluding remarks

In this chapter, we proposed a lifetime distribution that can be written as two-mixture generalized gamma distribution. The new model named as GLW distribution is a simple generalization of the weighted Lindley distribution proposed by Ghitany *et al.* (2011b), which accommodates increasing, decreasing, decreasing-increasing-decreasing, bathtub, or unimodal hazard functions, making the GWL distribution a flexible model for reliability data. The mathematical properties and the estimation procedure of the new distribution are discussed. Finally, we analyze two data sets for illustrative purposes, proving that the GWL outperform several usual three parameters lifetime distributions.

CHAPTER

8

COMMENTS AND FURTHER DEVELOPMENT

8.1 Comments

In this monograph, considering the Nakagami, gamma and generalized gamma distributions we presented main theorems that provide sufficient and necessary conditions for a wide class of posterior distribution to be proper. An interesting aspect of our findings is that one can easily check that a posterior is improper from the behavior of the proposed prior. For the NK distribution, we derived various objective priors such as Jeffreys Rule, Jeffreys prior, MDI prior and reference priors. We presented an overall reference prior that yields a proper posterior distribution if and only if $n \geq 1$. The proposed overall reference posterior distribution returned more accurate results as well as better theoretical properties such as invariance property under one-to-one transformations of the parameters, consistency under marginalization, consistent sampling properties.

For the gamma distribution we applied our proposed methodology in twelve objective priors, we observed that the MDI prior was the only prior that returned an improper posterior. An extensive simulation study showed that the posterior distribution obtained under Tibshirani prior provided more accurate results in terms of MRE, MSE and coverage probabilities.

Considering the generalized gamma distribution, we applied the proposed theorem in twelve different objective priors for the unknown parameters. We proved that only the (ϕ, α, μ)-reference prior leaded to a proper posterior distribution, such reference prior finally enable the use of the GG distribution in practice from the Bayesian point of view, in a methodologically correct way, with a objective prior that is one-to-one invariant, consistent marginalization and with consistent sampling properties, breaking with the problem of estimating the parameters of this important distribution.

Further, we proposed a new class of maximum a posteriori estimators for the parameters of the Nakagami-m and gamma distributions. These estimators have simple closed-form expressions and were rewritten as a bias corrected modified maximum likelihood estimators. Simulation studies were carried out to compare different estimation procedures. Numerical results revels that our new estimation schemes outperforms the existing closed-form estimators for the proposed distributions and produces extremely efficient estimates for both parameters, even for small sample sizes.

Finally, a new lifetime distribution that is expressed as a two-component mixture of the GG distribution is presented. The GLW distribution is a simple generalization of the weighted Lindley distribution proposed by Ghitany *et al.* (2011b), which accommodates increasing, decreasing, decreasing-increasing-decreasing, bathtub, or unimodal hazard functions, making the GWL distribution a flexible model for reliability data. The mathematical properties and the estimation procedure of the new distribution were discussed. Two data sets were analyzed for illustrative purposes, proving that the GWL outperform several usual three parameters lifetime distributions.

This thesis is based on nine papers developed during my doctoral. Five papers have been already published (RAMOS; LOUZADA, 2016; RAMOS; LOUZADA; RAMOS, 2016; RAMOS *et al.*, 2017; RAMOS; LOUZADA; RAMOS, 2018; LOUZADA; RAMOS, 2018a), and four are under review (RAMOS *et al.*, 2018; RAMOS *et al.*, 2018; RAMOS; LOUZADA, 2018; LOUZADA; RAMOS, 2018b)

8.2 Further development

There are a large number of possible extensions of this current work. We would like to present main theorems that will provide sufficient and necessary conditions for a wide class of posterior distribution be proper for the Weibull and Lognormal distributions. Another extension could be to follow Roy and Dey Roy and Dey (2014) that considered an objective Bayesian analysis for the generalized extreme value regression distribution. The proposed results can be further extended for the proposed models.

Reliability analysis is as much about accurately representing the past, as it is about perfectly modeling the future. We can consider the Bayes prediction for the proposed models. For instance for the Nakagami-m distribution using the observed order statistics. The Bayesian approach can be considered due to its facility in obtaining the predictive density of the future observation. The notation and the steps assumed follows Kundu and Raqab (2012). Let $x_{(m)}$ denote the m-th order statistic, $X_{(1)} < \ldots < X_{(m)}$ be the observed sample and $X_{(m+1)} < \ldots < X_{(n)}$ be the unobserved future sample.

From the Markov property of the conditional order statistics, we have

$$
\begin{aligned}
f_{X_{m+k}}(y|x) = & f_{X_{m+k}|X_m}(y|x) = \frac{(n-m)!}{(k-1)!(n-m-k)!} \\
& \times \frac{f(y)\left(F(y)-F(x_m)\right)^{k-1}\left(1-F(y)\right)^{n-m-k}}{\left(1-F(x_m)\right)^{n-m}}
\end{aligned}
\tag{8.1}
$$

for $y > x_{(m)}$. After some algebra we have

$$
\begin{aligned}
f_{X_{m+k}}(y|x) = & \frac{2(n-m)!}{(k-1)!(n-m-k)!} y^{-2\mu-1}\left(\frac{\mu}{\Omega}\right)^{\mu}\exp\left(\frac{\mu}{\Omega}y^2\right) \\
& \times \frac{\left(\gamma\left(\mu,\frac{\mu}{\Omega}y^2\right)-\gamma\left(\mu,\frac{\mu}{\Omega}x_m^2\right)\right)^{k-1}\left(\Gamma\left(\mu,\frac{\mu}{\Omega}y^2\right)\right)^{n-m-k}}{\left(\Gamma\left(\mu,\frac{\mu}{\Omega}x_{(m)}^2\right)\right)^{n-m}}.
\end{aligned}
\tag{8.2}
$$

The posterior predictive density of X_{m+k} given x is

$$
p_{X_{m+k}}(y|x) = \int_0^\infty \int_0^\infty f_{X_{m+k}}(y|x)\pi(\Omega,\mu|x)d\Omega d\mu.
$$

Therefore, the predictive density of $X_{(m+k)}$ under the assumption of $y > x_{(m)}$ is

$$
f^*_{X_{m+k}}(y|x) = \int_0^\infty \int_0^\infty f_{X_{m+k}|X_m}(y|x)\pi(\Omega,\mu|x)d\Omega d\mu. \tag{8.3}
$$

Simulation studies can be considered for the proposed predictive distribution. We will also explore such approach for different probability density functions.

BIBLIOGRAPHY

AALO, V. A.; PIBOONGUNGON, T.; ISKANDER, C.-D. Bit-error rate of binary digital modulation schemes in generalized gamma fading channels. **IEEE Communications Letters**, IEEE, v. 9, n. 2, p. 139–141, 2005. Citation on page 23.

AARSET, M. V. How to identify a bathtub hazard rate. **IEEE Transactions on Reliability**, v. 36, n. 1, p. 106–108, 1987. Citation on page 121.

ABRAMOWITZ, M.; STEGUN, I. A. **Handbook of Mathematical Functions**. 10. ed. Washington, D.C.: NBS, 1972. Citations on pages 45, 46, 47, 63, 67, 69, 70, 71, 86, 87, 90, and 137.

AHSANULLAH, M.; MASWADAH, M.; ALI, M. S. Kernel inference on the generalized gamma distribution based on generalized order statistics. **Journal of Statistical Theory and Applications**, Atlantis Press, v. 12, n. 2, p. 152–172, 2013. Citation on page 23.

BAKOUCH, H.; DEY, S.; RAMOS, P.; LOUZADA, F. Binomial-exponential 2 distribution: Different estimation methods with weather applications. **TEMA**, v. 18, n. 2, p. 233–251, 2017. Citation on page 29.

BAKOUCH, H. S.; AL-ZAHRANI, B. M.; AL-SHOMRANI, A. A.; MARCHI, V. A.; LOUZADA, F. An extended lindley distribution. **Journal of the Korean Statistical Society**, Elsevier, v. 41, n. 1, p. 75–85, 2012. Citation on page 111.

BALAKRISHNAN, N.; PENG, Y. Generalized gamma frailty model. **Statistics in medicine**, Wiley Online Library, v. 25, n. 16, p. 2797–2816, 2006. Citation on page 23.

BAR-LEV, S. K.; REISER, B. An exponential subfamily which admits umpu tests based on a single test statistic. **The Annals of Statistics**, JSTOR, p. 979–989, 1982. Citation on page 69.

BARLOW, R. E.; CAMPO, R. A. **Total Time on Test Processes and Applications to Failure Data Analysis**. Pennsylvania: [s.n.], 1975. Citation on page 28.

BARRETO-SOUZA, W.; BAKOUCH, H. S. A new lifetime model with decreasing failure rate. **Statistics**, Taylor & Francis, v. 47, n. 2, p. 465–476, 2013. Citation on page 111.

BARRETO-SOUZA, W.; CRIBARI-NETO, F. A generalization of the exponential-poisson distribution. **Statistics & Probability Letters**, Elsevier, v. 79, n. 24, p. 2493–2500, 2009. Citation on page 112.

BEAULIEU, N. C.; CHEN, Y. A map estimator for the m parameter in nakagami fading ultra-wide bandwidth indoor channels. **IEEE transactions on wireless communications**, IEEE, v. 6, n. 3, p. 840–844, 2007. Citation on page 38.

BERGER, J. O.; BERNARDO, J. M. Estimating a product of means: Bayesian analysis with reference priors. **Journal of the American Statistical Association**, Taylor & Francis, v. 84, n. 405, p. 200–207, 1989. Citations on pages 23, 24, and 31.

____. Ordered group reference priors with application to the multinomial problem. **Biometrika**, Biometrika Trust, v. 79, n. 1, p. 25–37, 1992. Citations on pages 23, 24, and 31.

____. Reference priors in a variance components problem. In: **Bayesian Analysis in Statistics and Econometrics**. [S.l.]: Springer, 1992. p. 177–194. Citations on pages 23, 24, and 31.

BERGER, J. O.; BERNARDO, J. M. *et al.* On the development of reference priors. **Bayesian statistics**, v. 4, n. 4, p. 35–60, 1992. Citations on pages 23, 24, and 31.

BERGER, J. O.; BERNARDO, J. M.; SUN, D. *et al.* Overall objective priors. **Bayesian Analysis**, International Society for Bayesian Analysis, v. 10, n. 1, p. 189–221, 2015. Citations on pages 23, 24, 31, 32, 59, 69, 72, and 80.

BERNARDO, J. M. Reference posterior distributions for bayesian inference. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 113–147, 1979. Citations on pages 23, 24, 30, 31, and 80.

____. Reference analysis. **Handbook of statistics**, Elsevier, v. 25, p. 17–90, 2005. Citations on pages 32, 38, 79, and 80.

BOX, G. E.; TIAO, G. C. **Bayesian inference in statistical analysis**. [S.l.]: John Wiley & Sons, 1973. Citations on pages 23 and 30.

BRYSON, M. C.; SIDDIQUI, M. Some criteria for aging. **Journal of the American Statistical Association**, Taylor & Francis, v. 64, n. 328, p. 1472–1483, 1969. Citation on page 41.

BURNHAM, K. P.; ANDERSON, D. R. Multimodel inference: understanding aic and bic in model selection. **Sociological methods & research**, Sage Publications Sage CA: Thousand Oaks, CA, v. 33, n. 2, p. 261–304, 2004. Citation on page 33.

CARCOLE, E.; SATO, H. Statistics of the fluctuations of the amplitude of coda waves of local earthquakes. In: **Seismological Society of Japan, 2009 Fall Meeting, C31-13, Kyoto, Japan**. [S.l.: s.n.], 2009. Citation on page 37.

CASELLA, G.; BERGER, R. L. **Statistical inference**. [S.l.]: Duxbury Pacific Grove, CA, 2002. Citation on page 29.

CHEN, Z. A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. **Statistics & Probability Letters**, Elsevier, v. 49, n. 2, p. 155–161, 2000. Citation on page 55.

CHENG, J.; BEAULIEU, N. C. Maximum-likelihood based estimation of the nakagami m parameter. **IEEE Communications letters**, v. 5, n. 3, p. 101–103, 2001. Citations on pages 37 and 99.

____. Generalized moment estimators for the nakagami fading parameter. **IEEE communications letters**, Institute of Electrical and Electronics Engineers, v. 6, n. 4, p. 144–146, 2002. Citations on pages 37, 99, and 101.

COX, D. R.; HINKLEY, D. V. **Theoretical statistics**. [S.l.]: CRC Press, 1979. Citation on page 32.

COX, D. R.; REID, N. Parameter orthogonality and approximate conditional inference. **Journal of the Royal Statistical Society. Series B (Methodological)**, JSTOR, p. 1–39, 1987. Citations on pages 33, 100, and 108.

DADPAY, A.; SOOFI, E. S.; SOYER, R. Information measures for generalized gamma family. **Journal of Econometrics**, Elsevier, v. 138, n. 2, p. 568–585, 2007. Citation on page 23.

DEY, S.; KUMAR, D.; RAMOS, P. L.; LOUZADA, F. Exponentiated chen distribution: Properties and estimation. **Communications in Statistics - Simulation and Computation**, Taylor & Francis, v. 46, n. 10, p. 8118–8139, 2017. Citation on page 29.

FOLLAND, G. B. **Real analysis: modern techniques and their applications**. 2. ed. New York: Wiley, 1999. Citations on pages 43, 44, and 103.

FONSECA, T. C.; FERREIRA, M. A.; MIGON, H. S. Objective bayesian analysis for the student-t regression model. **Biometrika**, Oxford University Press, v. 95, n. 2, p. 325–333, 2008. Citations on pages 24, 87, and 97.

GAEDDERT, J.; ANNAMALAI, A. Further results on nakagami-m parameter estimation. In: IEEE. **Vehicular Technology Conference, 2004. VTC2004-Fall. 2004 IEEE 60th**. [S.l.], 2004. v. 6, p. 4255–4259. Citations on pages 37 and 99.

GAMERMAN, D.; LOPES, H. F. **Markov chain Monte Carlo: stochastic simulation for Bayesian inference**. [S.l.]: CRC Press, 2006. Citations on pages 49 and 93.

GELFAND, A. E.; SMITH, A. F. Sampling-based approaches to calculating marginal densities. **Journal of the American statistical association**, Taylor & Francis Group, v. 85, n. 410, p. 398–409, 1990. Citation on page 67.

GEWEKE, J. *et al.* **Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments**. [S.l.]: Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA, 1991. Citation on page 93.

GHITANY, M.; AL-MUTAIRI, D.; BALAKRISHNAN, N.; AL-ENEZI, L. Power lindley distribution and associated inference. **Computational Statistics & Data Analysis**, Elsevier, v. 64, p. 20–33, 2013. Citations on pages 111 and 115.

GHITANY, M.; ALQALLAF, F.; AL-MUTAIRI, D.; HUSAIN, H. A two-parameter weighted lindley distribution and its applications to survival data. **Mathematics and Computers in Simulation**, Elsevier, v. 81, n. 6, p. 1190–1201, 2011. Citation on page 55.

____. A two-parameter weighted lindley distribution and its applications to survival data. **Mathematics and Computers in Simulation**, Elsevier, v. 81, n. 6, p. 1190–1201, 2011. Citations on pages 111, 124, and 126.

GHITANY, M.; ATIEH, B.; NADARAJAH, S. Lindley distribution and its application. **Mathematics and computers in simulation**, Elsevier, v. 78, n. 4, p. 493–506, 2008. Citation on page 111.

GLASER, R. E. Bathtub and related failure rate characterizations. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 75, n. 371, p. 667–672, 1980. Citation on page 28.

HAGER, H. W.; BAIN, L. J. Inferential procedures for the generalized gamma distribution. **Journal of the American Statistical Association**, Taylor & Francis, v. 65, n. 332, p. 1601–1609, 1970. Citations on pages 79 and 81.

HWANG, T.-Y.; HUANG, P.-H. On new moment estimation of parameters of the gamma distribution using its characterization. **Annals of the Institute of Statistical Mathematics**, Springer, v. 54, n. 4, p. 840–847, 2002. Citation on page 99.

JEFFREYS, H. An invariant form for the prior probability in estimation problems. In: THE ROYAL SOCIETY. **Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences**. [S.l.], 1946. v. 186, n. 1007, p. 453–461. Citations on pages 23, 30, 31, 80, and 87.

KASS, R. E.; WASSERMAN, L. The selection of prior distributions by formal rules. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 91, n. 435, p. 1343–1370, 1996. Citations on pages 30 and 80.

KIM, K.; LATCHMAN, H. A. Statistical traffic modeling of mpeg frame size: Experiments and analysis. **Statistical Methods in Radio Wave Propagation**, v. 29, n. 3, p. 456–465, 2009. Citation on page 37.

KUNDU, D.; RAQAB, M. Z. Bayesian inference and prediction of order statistics for a type-ii censored weibull distribution. **Journal of statistical planning and inference**, Elsevier, v. 142, n. 1, p. 41–47, 2012. Citation on page 126.

LAWLESS, J. F. **Statistical models and methods for lifetime data**. [S.l.]: John Wiley & Sons, 2011. Citation on page 54.

LI, H.-C.; HONG, W.; WU, Y.-R.; FAN, P.-Z. On the empirical-statistical modeling of sar images with generalized gamma distribution. **IEEE Journal of selected topics in signal processing**, IEEE, v. 5, n. 3, p. 386–397, 2011. Citation on page 23.

LISEO, B. Elimination of nuisance parameters with reference priors. **Biometrika**, Biometrika Trust, v. 80, n. 2, p. 295–304, 1993. Citation on page 69.

LORENZ, M. O. Methods of measuring the concentration of wealth. **Publications of the American statistical association**, Taylor & Francis, v. 9, n. 70, p. 209–219, 1905. Citation on page 118.

LOUZADA, F.; RAMOS, P. Efficient closed-form maximum a posteriori estimators for the gamma distribution. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 99, n. 0, p. 1–13, 2018. Citation on page 126.

LOUZADA, F.; RAMOS, P. L. Some comments on closed-form estimators for the gamma distribution derived from likelihood equations. **Submitted**, 2018. Citation on page 126.

LOUZADA, F.; RAMOS, P. L.; PERDONÁ, G. S. Different estimation procedures for the parameters of the extended exponential geometric distribution for medical data. **Computational and Mathematical Methods in Medicine**, Hindawi Publishing Corporation, v. 2016, 2016. Citation on page 29.

MARANI, A.; LAVAGNINI, I.; BUTTAZZONI, C. Statistical study of air pollutant concentrations via generalized gamma distributions. **Journal of the Air Pollution Control Association**, Taylor & Francis, v. 36, n. 11, p. 1250–1254, 1986. Citation on page 23.

MIGON, H. S.; GAMERMAN, D.; LOUZADA, F. **Statistical inference: an integrated approach**. [S.l.]: CRC press, 2014. Citation on page 30.

MILLER, R. B. Bayesian analysis of the two-parameter gamma distribution. **Technometrics**, Taylor & Francis, v. 22, n. 1, p. 65–69, 1980. Citations on pages 59, 66, and 67.

MOALA, F. A.; RAMOS, P. L.; ACHCAR, J. A. Bayesian inference for two-parameter gamma distribution assuming different noninformative priors. **Revista Colombiana de Estadística**, v. 36, n. 2, p. 321–338, 2013. Citations on pages 69 and 70.

MUDHOLKAR, G. S.; SRIVASTAVA, D. K.; FREIMER, M. The exponentiated weibull family: a reanalysis of the bus-motor-failure data. **Technometrics**, Taylor & Francis, v. 37, n. 4, p. 436–445, 1995. Citation on page 112.

MUDHOLKAR, G. S.; SRIVASTAVA, D. K.; KOLLIA, G. D. A generalization of the weibull distribution with application to the analysis of survival data. **Journal of the American Statistical Association**, Taylor & Francis Group, v. 91, n. 436, p. 1575–1583, 1996. Citation on page 112.

MUKERJEE, R.; DEY, D. K. Frequentist validity of posterior quantiles in the presence of a nuisance parameter: higher order asymptotics. **Biometrika**, Biometrika Trust, v. 80, n. 3, p. 499–505, 1993. Citation on page 33.

NAKAGAMI, N. The m-distribution a general formulation of intensity distribution of rapid fading. **Statistical Methods in Radio Wave Propagation: Proceedings of a Symposium**, Pergamon Press, p. 3–36, 1960. Citations on pages 37, 41, and 99.

NAKAHARA, H.; CARCOLÉ, E. Maximum-likelihood method for estimating coda q and the nakagami-m parameter. **Bulletin of the Seismological Society of America**, Seismological Society of America, v. 100, n. 6, p. 3174–3182, 2010. Citation on page 37.

NOORTWIJK, J. M. V. Bayes estimates of flood quantiles using the generalised gamma distribution. **System and Bayesian Reliability**, p. 351–374, 2001. Citations on pages 23, 24, 80, 87, 94, 95, and 97.

NORTHROP, P.; ATTALIDES, N. Posterior propriety in bayesian extreme value analyses using reference priors. **Statistica Sinica**, Academia Sinica, Institute of Statistical Science, v. 26, n. 2, 2016. Citations on pages 24, 59, and 80.

O'HAGAN, A.; BUCK, C. E.; DANESHKHAH, A.; EISER, J. R.; GARTHWAITE, P. H.; JENKINSON, D. J.; OAKLEY, J. E.; RAKOW, T. **Uncertain judgements: eliciting experts' probabilities**. [S.l.]: John Wiley & Sons, 2006. Citations on pages 30 and 79.

OLCAY, A. H. Mean residual life function for certain types of non-monotonic ageing. **Communications in statistics. Stochastic models**, Taylor & Francis, v. 11, n. 1, p. 219–225, 1995. Citation on page 41.

PAL, M.; ALI, M. M.; WOO, J. Exponentiated weibull distribution. **Statistica**, v. 66, n. 2, p. 139–147, 2006. Citation on page 52.

POPESCU, T. D.; AIORDACHIOAIE, D. Signal segmentation in time-frequency plane using renyi entropy-application in seismic signal processing. In: IEEE. **2013 Conference on Control and Fault-Tolerant Systems (SysTol)**. [S.l.], 2013. p. 312–317. Citation on page 117.

PRENTICE, R. L. A log gamma model and its maximum likelihood estimation. **Biometrika**, Biometrika Trust, v. 61, n. 3, p. 539–544, 1974. Citations on pages 79, 81, and 93.

RAJU, B. I.; SRINIVASAN, M. A. Statistics of envelope of high-frequency ultrasonic backscatter from human skin in vivo. **IEEE transactions on ultrasonics, ferroelectrics, and frequency control**, IEEE, v. 49, n. 7, p. 871–882, 2002. Citation on page 23.

RAMOS, P.; LOUZADA, F. The generalized weighted lindley distribution: Properties, estimation, and applications. **Cogent Mathematics**, Taylor & Francis, v. 3, n. 1, p. 1256022, 2016. Citation on page 126.

RAMOS, P. L.; ACHCAR, J. A.; MOALA, F. A.; RAMOS, E.; LOUZADA, F. Bayesian analysis of the generalized gamma distribution using non-informative priors. **Statistics**, Taylor & Francis, v. 51, n. 4, p. 824–843, 2017. Citation on page 126.

RAMOS, P. L.; DEY, D. K.; LOUZADA, F.; RAMOS, E. On posterior properties of the two parameter gamma family of distributions. **Submitted**, 2018. Citation on page 126.

RAMOS, P. L.; LOUZADA, F. Bayesian reference analysis for the generalized gamma distribution. **Submitted**, 2018. Citation on page 126.

RAMOS, P. L.; LOUZADA, F.; RAMOS, E. An efficient, closed-form map estimator for nakagami-m fading parameter. **IEEE Communications Letters**, IEEE, v. 20, n. 11, p. 2328–2331, 2016. Citation on page 126.

____. Posterior properties of the nakagami-m distribution using non-informative priors and applications in reliability. **IEEE Transactions on reliability**, Taylor & Francis, v. 99, n. 1, p. 1–13, 2018. Citation on page 126.

RAMOS, P. L.; MOALA, F. A.; ACHCAR, J. A. Objective priors for estimation of extended exponential geometric distribution. **Journal of Modern Applied Statistical Methods**, v. 13, n. 2, p. 12, 2014. Citations on pages 15 and 29.

RAMOS, P. L.; RAMOS, E.; DEY, D. K.; LOUZADA, F. Posterior properties of the generalized gamma family of distributions and its higher moments. **Submitted**, 2018. Citation on page 126.

RENYI, A. On measures of entropy and information. In: **Fourth Berkeley symposium on mathematical statistics and probability**. [S.l.: s.n.], 1961. v. 1, p. 547–561. Citation on page 117.

RODRIGUES, G. C.; LOUZADA, F.; RAMOS, P. L. Poisson-exponential distribution: different methods of estimation. **Journal of Applied Statistics**, Taylor & Francis, v. 45, n. 1, p. 128–144, 2018. Citation on page 29.

ROY, V.; DEY, D. K. Propriety of posterior distributions arising in categorical and survival models under generalized extreme value distribution. **Statistica Sinica**, JSTOR, p. 699–722, 2014. Citation on page 126.

RUDIN, W. *et al.* **Principles of mathematical analysis**. [S.l.]: McGraw-hill New York, 1964. Citation on page 34.

SARKAR, S.; GOEL, N.; MATHUR, B. Adequacy of nakagami-m distribution function to derive giuh. **Journal of Hydrologic Engineering**, American Society of Civil Engineers, v. 14, n. 10, p. 1070–1079, 2009. Citation on page 37.

____. Performance investigation of nakagami-m distribution to derive flood hydrograph by genetic algorithm optimization approach. **Journal of Hydrologic Engineering**, American Society of Civil Engineers, v. 15, n. 8, p. 658–666, 2010. Citation on page 37.

SCHWARTZ, J.; GODWIN, R. T.; GILES, D. E. Improved maximum-likelihood estimation of the shape parameter in the nakagami distribution. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 83, n. 3, p. 434–445, 2013. Citation on page 108.

SHANKAR, P. M.; DUMANE, V.; REID, J. M.; GENIS, V.; FORSBERG, F.; PICCOLI, C. W.; GOLDBERG, B. B. Classification of ultrasonic b-mode images of breast masses using nakagami distribution. **IEEE transactions on ultrasonics, ferroelectrics, and frequency control**, IEEE, v. 48, n. 2, p. 569–580, 2001. Citation on page 37.

SHANNON, C. E. A mathematical theory of communication. **ACM SIGMOBILE Mobile Computing and Communications Review**, ACM, v. 5, n. 1, p. 3–55, 2001. Citation on page 116.

SON, Y.-S.; OH, M.-R. Bayesian estimation of the nakagami-m fading parameter. **Communications for Statistical Applications and Methods**, The Korean Statistical Society, v. 14, n. 2, p. 345–353, 2007. Citation on page 38.

SONG, K.-S. Globally convergent algorithms for estimating generalized gamma distributions in fast signal and image processing. **IEEE Transactions on Image Processing**, IEEE, v. 17, n. 8, p. 1233–1250, 2008. Citations on pages 25 and 99.

STACY, E. W. A generalization of the gamma distribution. **The Annals of Mathematical Statistics**, JSTOR, p. 1187–1192, 1962. Citation on page 23.

STACY, E. W.; MIHRAM, G. A. Parameter estimation for a generalized gamma distribution. **Technometrics**, Taylor & Francis, v. 7, n. 3, p. 349–358, 1965. Citations on pages 79 and 81.

SUN, D.; YE, K. Frequentist validity of posterior quantiles for a two-parameter exponential family. **Biometrika**, Biometrika Trust, v. 83, n. 1, p. 55–65, 1996. Citations on pages 59, 69, and 71.

TAHAI, A.; MEYER, M. J. A revealed preference study of management journals' direct influences. **Strategic Management Journal**, JSTOR, v. 20, n. 3, p. 279–296, 1999. Citation on page 23.

TIBSHIRANI, R. Noninformative priors for one parameter of many. **Biometrika**, Biometrika Trust, v. 76, n. 3, p. 604–608, 1989. Citations on pages 32 and 33.

TSUI, P.; HUANG, C.; WANG, S. Use of nakagami distribution and logarithmic compression in ultrasonic tissue characterization. **Journal of Medical and Biological Engineering**, WALTER H CHANG, v. 26, n. 2, p. 69, 2006. Citation on page 37.

WANG, N.; SONG, X.; CHENG, J. Generalized method of moments estimation of the nakagami-m fading parameter. **IEEE Transactions on Wireless Communications**, IEEE, v. 11, n. 9, p. 3316–3325, 2012. Citations on pages 38, 99, 101, and 102.

YE, Z.-S.; CHEN, N. Closed-form estimators for the gamma distribution derived from likelihood equations. **The American Statistician**, Taylor & Francis, v. 71, n. 2, p. 177–181, 2017. Citations on pages 107, 108, and 109.

ZAKERZADEH, H.; DOLATI, A. Generalized lindley distribution. **Journal of Mathematical Extension**, 2009. Citation on page 111.

ZELLNER, A. Maximal data information prior distributions. **In New Methods in the applications of Bayesian Methods**, North- Holland, Amsterdam., 1977. Citations on pages 23, 30, 31, and 80.

____. Maximal data information prior distributions. **Basic Issues in Econometrics**, U. of Chicago Press, 1984. Citations on pages 23, 30, 80, and 86.

ZHANG, Q. A note on the estimation of nakagami-m fading parameter. **IEEE communications letters**, IEEE, v. 6, n. 6, p. 237–238, 2002. Citations on pages 37 and 99.

CHAPTER

A

USEFUL PROPORTIONALITIES

The following proportionalities are also useful to prove results related to the posterior distribution.

Proposition A.0.1. The following results hold

$$\sqrt{\phi^2\psi'(\phi)^2-\psi'(\phi)-1} \underset{\phi\to 0^+}{\propto} 1 \quad\text{and}\quad \sqrt{\phi^2\psi'(\phi)^2-\psi'(\phi)-1} \underset{\phi\to\infty}{\propto} \frac{1}{\phi}.$$

Proof. Let us present the proof for the first case. Considering the recurrence relation $\psi^{'}(\phi) = \dfrac{1}{\phi^2}+\psi^{'}(\phi+1)$, we have

$$\begin{aligned}\sqrt{\phi^2\psi'(\phi)^2-\psi'(\phi)-1} =& \sqrt{\frac{1}{\phi^2}+2\psi'(\phi+1)+\phi^2\psi'(\phi+1)^2-\frac{1}{\phi^2}-\psi'(\phi)-1}\\ =& \sqrt{\psi'(\phi+1)-1+\phi^2\psi'(\phi+1)^2}.\end{aligned}$$

Therefore

$$\lim_{\phi\to 0^+}\sqrt{\phi^2\psi'(\phi)^2-\psi'(\phi)-1} = \sqrt{\psi'(1)-1} = \sqrt{\frac{\pi^2}{6}-1} > 0. \tag{A.1}$$

For the second case, note that (see Abramowitz (ABRAMOWITZ; STEGUN, 1972), pg 260)

$$\psi'(\phi) = \frac{1}{\phi}+\frac{1}{2\phi^2}+\frac{1}{6\phi^3}+o\left(\frac{1}{\phi^3}\right).$$

Hence, it follows that

$$\begin{aligned}&\phi^2\left(\frac{1}{\phi}+\frac{1}{2\phi^2}+\frac{1}{6\phi^3}+o\left(\frac{1}{\phi^3}\right)\right)^2-\frac{1}{\phi}-\frac{1}{2\phi^2}-\frac{1}{6\phi^3}-o\left(\frac{1}{\phi^3}\right)-1=\\ &\phi^2\left(\frac{1}{\phi^2}+\frac{1}{\phi^3}+\frac{7}{12\phi^4}+o\left(\frac{1}{\phi^4}\right)\right)-\frac{1}{\phi}-\frac{1}{2\phi^2}+o\left(\frac{1}{\phi^2}\right)-1=\\ &1+\frac{1}{\phi}+\frac{7}{12\phi^2}+o\left(\frac{1}{\phi^2}\right)-\frac{1}{\phi}-\frac{1}{2\phi^2}+o\left(\frac{1}{\phi^2}\right)-1=\frac{1}{12\phi^2}+o\left(\frac{1}{\phi^2}\right).\end{aligned}$$

Therefore,

$$\lim_{\phi\to\infty}\frac{\sqrt{\phi^2\psi'(\phi)^2-\psi'(\phi)-1}}{\phi^{-1}}=\lim_{\phi\to\infty}\sqrt{\frac{1}{12}+o(1)}=\sqrt{\frac{1}{12}}.$$

□

Proposition A.0.2. The following results hold

$$\frac{\Gamma(n\phi)}{\Gamma(\phi)^n}\underset{\phi\to 0^+}{\propto}\phi^{n-1}\quad\text{and}\quad\frac{\Gamma(n\phi)}{\Gamma(\phi)^n}\underset{\phi\to\infty}{\propto}n^{n\phi}\phi^{\frac{n-1}{2}}.$$

Proof. Considering the recurrence relation $\Gamma(z)=\frac{1}{z}\Gamma(z+1)$ it follows that $\lim_{z\to 0^+}\frac{\Gamma(z)}{\frac{1}{z}}=\Gamma(1)=1$. Therefore

$$\Gamma(z)\underset{\phi\to 0^+}{\propto}\frac{1}{z}\quad\text{and}\tag{A.2}$$

$$\frac{\Gamma(n\phi)}{\Gamma(\phi)^n}\underset{\phi\to 0^+}{\propto}\frac{\frac{1}{n\phi}}{\frac{1}{\phi^n}}\underset{\phi\to 0^+}{\propto}\phi^{n-1}.\tag{A.3}$$

Now, considering Stirling's approximation for gamma function

$$\lim_{z\to\infty}\frac{\Gamma(z)}{\sqrt{2\pi}z^{z-\frac{1}{2}}e^{-z}}=1\Rightarrow\lim_{z\to\infty}\frac{\Gamma(z)}{z^{z-\frac{1}{2}}e^{-z}}=\sqrt{2\pi}\Rightarrow\Gamma(z)\underset{z\to\infty}{\propto}z^{z-\frac{1}{2}}e^{-z}.$$

Then, by Proposition **??** we have

$$\frac{\Gamma(n\phi)}{\Gamma(\phi)^n}\underset{\phi\to\infty}{\propto}\frac{(n\phi)^{n\phi-\frac{1}{2}}e^{-n\phi}}{\left(\phi^{\phi-\frac{1}{2}}e^{-\phi}\right)^n}=n^{-\frac{1}{2}}n^{n\phi}\phi^{\frac{n-1}{2}}\underset{\phi\to\infty}{\propto}n^{n\phi}\phi^{\frac{n-1}{2}}.$$

□

Proposition A.0.3. Let $\mathrm{p}(\alpha)=\log\left(\frac{\frac{1}{n}\sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}\right)$, for $t_1,t_2,\ldots,t_n$ positive and not all equal, then $\mathrm{p}(\alpha)>0$ and the following hold

$$\mathrm{p}(\alpha)\underset{\alpha\to 0^+}{\propto}\alpha^2\quad\text{and}\quad\mathrm{p}(\alpha)\underset{\alpha\to\infty}{\propto}\alpha.$$

Proof. Note that $\frac{\frac{1}{n}\sum_{i=1}^n t_i^\alpha}{\sqrt[n]{\prod_{i=1}^n t_i^\alpha}}>1\Rightarrow\mathrm{p}(\alpha)>0$ by the arithmetic-geometric inequality.

Now, let $u_i=\frac{t_i}{\sqrt[n]{\prod_{i=1}^n t_i}}$, $i=1,\ldots,n$ and $u_m=\max\{u_1,\ldots,u_n\}$. Since $t_1,\ldots,t_n$ are not all equal then $u_m>1$ and

$$\lim_{\alpha\to\infty}\frac{\mathrm{p}(\alpha)}{\alpha}=\lim_{\alpha\to\infty}\frac{1}{\alpha}\log\left(\sum_{i=1}^n u_i^\alpha\right)-\frac{\log(n)}{\alpha}=\lim_{\alpha\to\infty}\log\left\{\left(\sum_{i=1}^n u_i^\alpha\right)^{\frac{1}{\alpha}}\right\}$$

$$=\lim_{\alpha\to\infty}\log\left(u_m\left(\sum_{i=1}^n(u_i/u_m)^\alpha\right)^{\frac{1}{\alpha}}\right).$$

Moreover, $1 \leq \sum_{i=1}^{n} (u_i/u_m)^{\alpha} \leq n$, then $1 \leq \left(\sum_{i=1}^{n} (u_i/u_m)^{\alpha}\right)^{\frac{1}{\alpha}} \leq n^{\frac{1}{\alpha}}$, which implies

$$\lim_{\alpha\to\infty} \left(\sum_{i=1}^{n} (u_i/u_m)^{\alpha}\right)^{\frac{1}{\alpha}} = 1 \text{ and}$$
$$\lim_{\alpha\to\infty} \frac{\mathrm{p}(\alpha)}{\alpha} = \lim_{\alpha\to\infty} \log\left(u_m \left(\sum_{i=1}^{n} (u_i/u_m)^{\alpha}\right)^{\frac{1}{\alpha}}\right) = \log(u_m) > 0 \tag{A.4}$$

which proves the first result.

Now, $\sum_{i=1}^{n} \log(u_i) = \log\left(\frac{\prod_{i=1}^{n} t_i}{\prod_{i=1}^{n} t_i}\right) = \log(1) = 0$ and

$$\begin{aligned}
\lim_{a\to 0^+} \frac{\mathrm{p}(\alpha)}{\alpha^2} &= \frac{\log\left(\frac{1}{n}\sum_{i=1}^{n} u_i^{\alpha}\right)}{\alpha^2} \overset{L'h}{=} \lim_{\alpha\to 0^+} \frac{\frac{\sum_{i=1}^{n}\log(u_i)u_i^{\alpha}}{\sum_{i=1}^{n} u_i^{\alpha}}}{2\alpha} \\
&\overset{L'h}{=} \frac{1}{2}\lim_{\alpha\to 0^+} \frac{(\sum_{i=1}^{n}\log(u_i)^2 u_i^{\alpha})(\sum_{i=1}^{n} u_i^{\alpha}) - (\sum_{i=1}^{n}\log(u_i)u_i^{\alpha})^2}{(\sum_{i=1}^{n} u_i^{\alpha})^2} \\
&= \frac{1}{2}\frac{n\sum_{i=1}^{n}\log(u_i)^2 - (\sum_{i=1}^{n}\log(u_i))^2}{n^2} = \frac{1}{2}\frac{\sum_{i=1}^{n}\log(u_i)^2}{n} > 0.
\end{aligned}$$

Note that $\frac{1}{2}\frac{\sum_{i=1}^{n}\log(u_i)^2}{n} \neq 0$ since, otherwise, we would have $\log(u_i) = 0 \Leftrightarrow u_i = 1, \forall i$ and would imply that t_i are all equal, which contradicts the hypothesis. Hence $\mathrm{p}(\alpha) \underset{\alpha\to 0^+}{\propto} \alpha^2$ which proves the second result. $\square$

Proposition A.0.4. Let $\mathrm{q}(\alpha) = \log\left(\frac{\sum_{i=1}^{n} t_i^{\alpha}}{\sqrt[n]{\prod_{i=1}^{n} t_i^{\alpha}}}\right)$, for $t_1, t_2, \ldots, t_n$ positive and not all equal, then $\mathrm{q}(\alpha) > 0$ and the following hold

$$\mathrm{q}(\alpha) \underset{\alpha\to 0^+}{\propto} 1 \quad \text{and} \quad \mathrm{q}(\alpha) \underset{\alpha\to\infty}{\propto} \alpha.$$

Proof. Note that $\frac{\sum_{i=1}^{n} t_i^{\alpha}}{\sqrt[n]{\prod_{i=1}^{n} t_i^{\alpha}}} > n \Rightarrow q(\alpha) > 0$ by the arithmetic-geometric inequality. Since $q(\alpha) = \log(n) + p(\alpha)$ and by Proposition A.0.3 $\lim_{\alpha\to 0^+} p(\alpha) = 0$ it follows that

$$\lim_{\alpha\to 0^+} q(\alpha) = \log(n) > 0. \tag{A.5}$$

which proves the first proportionality.

Analogously, from $q(\alpha) = \log(n) + p(\alpha)$ and Proposition A.0.3 it follows that $\mathrm{q}(\alpha) \underset{\alpha\to\infty}{\propto} \alpha$, hence the second proportionality is proved. $\square$

Proposition A.0.5. Let $\mathrm{q}(\alpha)$ be the same defined in Proposition A.0.4, then the following results are valid for $k \in \mathbb{R}^+$ and $r \in \mathbb{R}^+$.

$$\gamma(k, r\mathrm{q}(\alpha)) \underset{\alpha\to 0^+}{\propto} 1 \quad \text{and} \quad \gamma(k, r\mathrm{q}(\alpha)) \underset{\alpha\to\infty}{\propto} 1. \tag{A.6}$$

Proof. From (A.5) and the continuity of incomplete gamma function in $\mathbb{R}^+ \times \mathbb{R}^+$ we have

$$\lim_{\alpha \to 0^+} \gamma(k, r\mathrm{q}(\alpha)) = \gamma(k, r\log(n)) \Rightarrow \gamma(k, r\mathrm{q}(\alpha)) \underset{\alpha \to 0^+}{\propto} 1. \tag{A.7}$$

Now, from the definition of lower incomplete gamma function, it follows directly that $\lim_{x\to\infty} \gamma(y,x) = \Gamma(y)$ for $y > 0$. But, since $\mathrm{q}(\alpha) \underset{\alpha\to\infty}{\propto} \alpha$, we have $\lim_{\alpha\to\infty} \mathrm{q}(\alpha) = \infty$. Therefore

$$\lim_{\alpha \to \infty} \gamma(k, r\mathrm{q}(\alpha)) = \Gamma(k) \Rightarrow \gamma(k, r\mathrm{q}(\alpha)) \underset{\alpha \to 0^+}{\propto} 1. \tag{A.8}$$

□

Proposition A.0.6. Let $\mathrm{p}(\alpha)$ be the same defined in Proposition A.0.3 and let $t_m = \max\{t_1, \ldots, t_n\}$. Then the following results are valid for $k \in \mathbb{R}^+$ and $r \in \mathbb{R}^+$,

$$\Gamma(k, r\mathrm{p}(\alpha)) \underset{\alpha \to 0^+}{\propto} 1 \quad \text{and} \quad \Gamma(k, r\mathrm{p}(\alpha)) \underset{\alpha \to \infty}{\propto} \alpha^{k-1} e^{-r\log\left(\frac{t_m}{\sqrt[n]{\prod_{i=1}^n t_i}}\right)\alpha} \tag{A.9}$$

where $\Gamma(y,x) = \int_x^\infty w^{y-1} e^{-w} dw$ is the upper incomplete gamma function.

Proof. From the definition of upper incomplete gamma function, it follows directly that $\lim_{x\to 0^+} \Gamma(y,x) = \Gamma(y)$ for $y > 0$. However, as $\mathrm{p}(\alpha) \underset{\alpha\to 0^+}{\propto} \alpha^2$, we have $\lim_{\alpha\to 0^+} \mathrm{p}(\alpha) = 0$. Therefore

$$\lim_{\alpha \to 0^+} \Gamma(k, r\mathrm{p}(\alpha)) = \Gamma(k) \Rightarrow \Gamma(k, r\mathrm{p}(\alpha)) \underset{\alpha \to 0^+}{\propto} 1. \tag{A.10}$$

Now, by L'hospital rule and the definition of upper incomplete gamma function,

$$\lim_{x \to \infty} \frac{\Gamma(s,x)}{x^{s-1} e^{-x}} = 1.$$

However, $\mathrm{p}(\alpha) \underset{\alpha\to\infty}{\propto} \alpha^2$ which implies $\lim_{\alpha\to\infty} \mathrm{p}(\alpha) = \infty$ and

$$\lim_{\alpha \to \infty} \frac{\Gamma(k, r\mathrm{p}(\alpha))}{(r\mathrm{p}(\alpha))^{k-1} e^{-r\mathrm{p}(\alpha)}} = 1 \Rightarrow \Gamma(k, r\mathrm{p}(\alpha)) \underset{\alpha \to \infty}{\propto} \mathrm{p}(\alpha)^{k-1} e^{-r\mathrm{p}(\alpha)}.$$

Moreover, let c be the number of t_i equal to t_m for $i = 1, \ldots, n$, then

$$\lim_{\alpha \to \infty} \mathrm{p}(\alpha) - \log\left(\frac{t_m}{\sqrt[n]{\prod_{i=1}^n t_i}}\right)\alpha = \lim_{\alpha \to \infty} \log\left(\left(\frac{t_1}{t_m}\right)^\alpha + \cdots + \left(\frac{t_n}{t_m}\right)^\alpha\right) = \log(c).$$

Therefore

$$\begin{aligned}
\lim_{\alpha \to \infty} \frac{\mathrm{p}(\alpha)^{k-1} e^{-r\mathrm{p}(\alpha)}}{\alpha^{k-1} e^{-r\log\left(\frac{t_m}{\sqrt[n]{\prod_{i=1}^n t_i}}\right)\alpha}} &= \lim_{\alpha \to \infty} \left(\frac{\mathrm{p}(\alpha)}{\alpha}\right)^{k-1} e^{-r\left(\mathrm{p}(\alpha) - \log\left(\frac{t_m}{\sqrt[n]{\prod_{i=1}^n t_i}}\right)\alpha\right)} \\
&= 1^{k-1} \times e^{-r\log(c)} = c^{-r} > 0.
\end{aligned}$$

This implies that $\mathrm{p}(\alpha)^{k-1}e^{-r\mathrm{p}(\alpha)} \underset{\alpha\to\infty}{\propto} \alpha^{k-1}e^{-r\log\left(\frac{t_m}{\sqrt[n]{\prod_{i=1}^{n} t_i}}\right)\alpha}$. Finally

$$\Gamma(k, r\mathrm{p}(\alpha)) \underset{\alpha\to\infty}{\propto} \alpha^{k-1}e^{-r\log\left(\frac{t_m}{\sqrt[n]{\prod_{i=1}^{n} t_i}}\right)\alpha}.$$

$\square$

ICMC USP
SÃO CARLOS