



Modelos de fração de cura de mistura e não-mistura na distribuição Weibull Modificada Generalizada

Eduardo Cardoso de Oliveira

Orientador: Prof. Dr. Jorge Alberto Achcar

Coorientador: Prof. Dr. Mario Hissamitsu Tarumoto

Programa: Matemática Aplicada e Computacional

UNIVERSIDADE ESTADUAL PAULISTA

Faculdade de Ciências e Tecnologia de Presidente Prudente

Programa de Pós-Graduação em Matemática Aplicada e Computacional

**Modelos de fração de cura de mistura e
não-mistura na distribuição Weibull
Modificada Generalizada**

Eduardo Cardoso de Oliveira

Orientador: Prof. Dr. Jorge Alberto Achcar

Coorientador: Prof. Dr. Mario Hissamitsu Tarumoto

Dissertação apresentada ao Programa de Pós-Graduação em Matemática Aplicada e Computacional da Faculdade de Ciências e Tecnologia da UNESP para obtenção do título de Mestre em Matemática Aplicada e Computacional.

Presidente Prudente, Setembro de 2015

FICHA CATALOGRÁFICA

O46m Oliveira, Eduardo Cardoso de.
Modelos de fração de cura de mistura e não-mistura na distribuição Weibull Modificada Generalizada / Eduardo Cardoso de Oliveira. - Presidente Prudente : [s.n.], 2015
xiii, 169 f. : il.

Orientador: Jorge Alberto Achcar
Dissertação (mestrado) - Universidade Estadual Paulista, Faculdade de Ciências e Tecnologia
Inclui bibliografia

1. Fração de cura. 2. Análise de sobrevivência. 3. Simulação. I. Achcar, Jorge Alberto. II. Universidade Estadual Paulista. Faculdade de Ciências e Tecnologia. III. Título.

BANCA EXAMINADORA



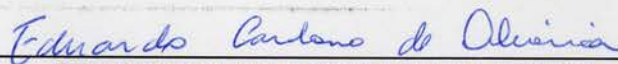
Prof. Dr. JORGE ALBERTO ACHCAR
ORIENTADOR



Prof. Dr. FERNANDO ANTONIO MOALA
UNESP/FCT



Prof. Dr. EMILIO AUGUSTO COELHO BARROS
UTFPR



EDUARDO CARDOSO DE OLIVEIRA

Presidente Prudente (SP), 21 de setembro de 2015.

Resultado: APROVADO.

À toda minha família, em especial a minha esposa e filhos.

Agradecimentos

Agradeço a minha família pela paciência e apoio, em especial a minha esposa Priscila e meus filhos Maria Eduarda e Pedro.

Ao meu orientador Prof. Dr. Jorge Alberto Achcar e ao Prof. Dr. Mario Hissamitsu Tarumoto, meu coorientador, pela confiança e paciência durante o mestrado.

Aos professores do programa pela dedicação, presteza e apoio.

Aos professores membros da banca pelas sugestões que enriqueceram essa dissertação.

Aos funcionários da Seção de Pós-Graduação por todo o auxílio prestado.

Ao Fernando Pacanelli Martins pelo apoio e presteza para a utilização do cluster.

Ao Conselho de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro, que possibilitou a realização do mestrado.

À Deus por estar presente em todos os momentos de minha vida.

*“ Os dois dias mais importantes da sua vida são:
o dia em que você nasceu,
e o dia em que você descobre o porquê ”.*

Mark Twain

Resumo

A análise de sobrevivência é uma das áreas da estatística com grande aplicabilidade em diversas áreas, especialmente em medicina, que é aplicada quando se pretende analisar um fenômeno em relação ao tempo transcorrido entre um evento inicial, no qual o indivíduo entra em um estado particular, e um evento final que modifica este estado. Uma particularidade na análise de dados de sobrevivência é a presença de dados censurados. Em muitas aplicações pode-se observar um grande número de indivíduos censurados, em geral à direita. Esse fato pode indicar que, em parte dos indivíduos em estudo, o evento de interesse nunca acontecerá, ou seja, pode se ter uma fração da população que é imune (ou que se cura). Neste caso, é importante estudar a proporção de pacientes que permanecem curados e analisar os efeitos das covariáveis, não só sobre o tempo de sobrevivência, como também sobre a fração de cura. Modelos para análise de dados com proporção de curados são frequentemente chamados de modelos com fração de cura. Em um estudo experimental que exista uma proporção substancial de indivíduos que não falhem ao final do experimento, não faz sentido considerar tais indivíduos como observações censuradas vindas da mesma população daqueles que haviam morrido. Sendo assim, é natural considerar um modelo construído a partir da mistura de duas distribuições: uma representando a distribuição dos indivíduos não curados e outra correspondendo a uma distribuição dos indivíduos curados. Várias distribuições de probabilidade podem ser usadas na análise de dados de sobrevivência na presença ou não de fração de cura. Neste trabalho foi analisado o modelo de mistura e de não-mistura para a fração de cura considerando em especial a distribuição Weibull Modificada Generalizada (WMG), comparando com vários modelos de sobrevivência existentes, sob os enfoques bayesiano e clássico. Para ilustrar o uso da modelagem com fração de cura é apresentado um estudo de simulação com geração de dados considerando indivíduos curados e comparando os resultados estimados pelos métodos clássico e bayesiano. É apresentado também duas aplicações com dados reais de pacientes com câncer, onde é realizado um estudo semi-paramétrico e um estudo de modelagem utilizando os principais modelos de sobrevivência comparados ao modelo WMG, sem fração de cura, com fração de cura de mistura e com fração de cura de não-mistura. Finalizando as aplicações foram incluídas as covariáveis identificadas no estudo semi-paramétrico nos modelos escolhidos através dos critérios BIC e CAIC sob o enfoque clássico e EBIC e ECAIC sob o enfoque bayesiano. Os resultados obtidos demonstraram que os modelos com fração de cura podem apresentar soluções mais adequadas aos dados com presença de indivíduos curados, tanto sob o enfoque clássico quanto sob o enfoque bayesiano.

Palavras-Chave: *Fração de cura, Análise de Sobrevivência, Simulação.*

Abstract

Survival analysis is one of the areas of statistics with great applicability in various fields, especially in medicine, which is applied when you want to analyze a phenomenon in relation to the time between an initial event in which the individual enters a particular state, and a final event which changes this state. A particularity in the survival data analysis is the presence of censored data. In many applications it can be observed a large number of subjects censored generally right. This may indicate that part of the study subjects, the event of interest will never happen, or may have a fraction of the population that is immune (or heals). In this case, it is important to study the proportion of patients who remain healed and analyze the effects of covariates, not only about survival time, but also about healing fraction. Models for data analysis proportion of cured are often called models with healing fraction. In an experimental study that there is a substantial proportion of individuals who do not fail at the end of the experiment, it makes sense to consider such individuals as censored observations coming from the same population of those who had died. Therefore, it is natural to consider a model built from the two distributions mix: one representing the distribution of uncured and another corresponding to a distribution of healed individuals individuals. Several probability distributions can be used to analyze survival data on the presence or absence of fraction of cure. In this study we analyzed the mixture model and non-mixture to cure fraction considering in particular the distribution Weibull Modified Generalized (WMG), comparing with various existing survival models, under the approaches Bayesian and classic. To illustrate the use of modelagem curing fraction is presented a simulation study to generate data regarding individuals cured and comparing the results estimated by classical and Bayesian methods. It also presented two applications with real data of cancer patients, which is performed a semi-parametric study and a modeling study using the main survival models compared to WMG model without fraction of cure, with a mixture and of non-mix. Finishing applications were included covariates identified in the semi-parametric study on the models chosen by the BIC and CAIC criteria under the classical approach and EBIC and ECAIC under the Bayesian approach. The results showed that the models with fraction of cure can present best solutions to data with the presence of healed individuals, in both approach.

Keywords: *Rate of Cure, Survival Analysis, Simulation.*

Lista de Figuras

1.1	Dados simulados para uma amostra de tamanho $n=100$ a partir da distribuição Weibull Modificada Generalizada com fração de cura de 15% com censura de 35% (a) e sem censura (b).	24
2.1	Estimativa de Kaplan-Meier para os dados de pacientes com câncer internados na UTI do INCA.	31
2.2	(a) Funções densidade, (b) de distribuição, (c) de sobrevivência e (d) de risco da distribuição WMG.	38
4.1	Curvas de sobrevivência estimadas pelo método de Kaplan-Meier para as covariáveis dos dados de pacientes com câncer internados na UTI do INCA.	56
4.2	$\text{Log}(\widehat{\Lambda}_{0j}(t))$ versus tempo para as covariáveis <i>gptumor</i> , <i>leucopenia</i> , <i>comorbi</i> e <i>desnut.</i>	59
4.3	Resíduos padronizados de Schoenfeld associados às covariáveis.	60
4.4	Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) versus a curva de sobrevivência estimada por Kaplan-Meier, sem fração de cura, sob enfoque clássico.	63
4.5	Gráficos das sobrevivências estimadas por Kaplan-Meier versus as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), sem fração de cura, sob enfoque clássico.	63
4.6	Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) versus a curva de sobrevivência estimada por Kaplan-Meier, sem fração de cura, sob enfoque Bayesiano.	65
4.7	Gráficos das sobrevivências estimadas por Kaplan-Meier versus as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), sem fração de cura, sob enfoque Bayesiano.	65
4.8	Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) versus a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de mistura, sob enfoque clássico.	67

4.9	Gráficos das sobrevivências estimadas por Kaplan-Meier <i>versus</i> as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de mistura, sob enfoque clássico.	67
4.10	Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) <i>versus</i> a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de mistura, sob enfoque Bayesiano.	69
4.11	Gráficos das sobrevivências estimadas por Kaplan-Meier <i>versus</i> as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de mistura, sob enfoque Bayesiano.	69
4.12	Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) <i>versus</i> a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de não-mistura, sob enfoque clássico.	71
4.13	Gráficos das sobrevivências estimadas por Kaplan-Meier <i>versus</i> as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de não-mistura, sob enfoque clássico.	71
4.14	Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) <i>versus</i> a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de não-mistura, sob enfoque Bayesiano.	73
4.15	Gráficos das sobrevivências estimadas por Kaplan-Meier <i>versus</i> as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de não-mistura, sob enfoque Bayesiano.	73
4.16	Curvas de sobrevivência (esquerda) e de risco (direita) estimadas pelos modelos WG e WMG de fração de cura de mistura e de não-mistura para os dados de pacientes internados na UTI do INCA, sob o enfoque clássico.	74
4.17	Curvas de sobrevivência (esquerda) e de risco (direita) estimadas pelos modelos WG e WMG de fração de cura de mistura e de não-mistura para os dados de pacientes internados na UTI do INCA, sob o enfoque Bayesiano.	75
4.18	Curvas de sobrevivência estimadas pelo método de Kaplan-Meier para as covariáveis dos dados de pacientes com câncer gástrico.	79
4.19	$\text{Log} \left(\widehat{\Lambda}_{0j}(t) \right)$ <i>versus</i> tempo para a covariável <i>trat.</i>	80
4.20	Resíduos padronizados de Schoenfeld associados à covariável <i>trat.</i>	80
4.21	Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) <i>versus</i> a curva de sobrevivência estimada por Kaplan-Meier, sem fração de cura, sob enfoque clássico.	83

4.22	Gráficos das sobrevivências estimadas por Kaplan-Meier <i>versus</i> as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), sem fração de cura, sob enfoque clássico.	83
4.23	Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) <i>versus</i> a curva de sobrevivência estimada por Kaplan-Meier, sem fração de cura, sob enfoque Bayesiano.	85
4.24	Gráficos das sobrevivências estimadas por Kaplan-Meier <i>versus</i> as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), sem fração de cura, sob enfoque Bayesiano.	85
4.25	Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) <i>versus</i> a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de mistura, sob enfoque clássico.	87
4.26	Gráficos das sobrevivências estimadas por Kaplan-Meier <i>versus</i> as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de mistura, sob enfoque clássico.	87
4.27	Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) <i>versus</i> a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de mistura, sob enfoque Bayesiano.	89
4.28	Gráficos das sobrevivências estimadas por Kaplan-Meier <i>versus</i> as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de mistura, sob enfoque Bayesiano.	89
4.29	Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) <i>versus</i> a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de não-mistura, sob enfoque clássico.	91
4.30	Gráficos das sobrevivências estimadas por Kaplan-Meier <i>versus</i> as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de não-mistura, sob enfoque clássico.	91
4.31	Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) <i>versus</i> a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de não-mistura, sob enfoque Bayesiano.	93

4.32	Gráficos das sobrevivências estimadas por Kaplan-Meier <i>versus</i> as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de não-mistura, sob enfoque Bayesiano.	93
4.33	Curvas de sobrevivência (esquerda) e de risco (direita) estimadas pelo modelo WMG de fração de cura de mistura e de não-mistura para os dados de pacientes com câncer gástrico do Hospital do Câncer de Barretos, sob o enfoque clássico.	94
4.34	Curvas de sobrevivência (esquerda) e de risco (direita) estimadas pelos modelos EE sem fração de cura, WG e WMG de fração de cura de mistura e EE e WMG de fração de cura de não-mistura para os dados de pacientes com câncer gástrico do Hospital do Câncer de Barretos, sob o enfoque Bayesiano.	95
C.1	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição GG sem fração de cura.	145
C.2	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição G sem fração de cura.	146
C.3	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição LN sem fração de cura.	146
C.4	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição W sem fração de cura.	146
C.5	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição EE sem fração de cura.	147
C.6	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WG sem fração de cura.	147
C.7	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WMG sem fração de cura.	148
C.8	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição GG de fração de cura de mistura.	148
C.9	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição G de fração de cura de mistura.	149
C.10	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição LN de fração de cura de mistura.	149
C.11	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição W de fração de cura de mistura.	150
C.12	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição EE de fração de cura de mistura.	150
C.13	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WG de fração de cura de mistura.	151
C.14	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WMG de fração de cura de mistura.	151

C.15	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição GG de fração de cura de não-mistura.	152
C.16	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição G de fração de cura de não-mistura.	152
C.17	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição LN de fração de cura de não-mistura.	153
C.18	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição W de fração de cura de não-mistura.	153
C.19	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição EE de fração de cura de não-mistura.	154
C.20	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WG de fração de cura de não-mistura.	154
C.21	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WMG de fração de cura de não-mistura.	155
C.22	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição GG sem fração de cura.	156
C.23	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição G sem fração de cura.	156
C.24	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição LN sem fração de cura.	157
C.25	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição W sem fração de cura.	157
C.26	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição EE sem fração de cura.	157
C.27	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WG sem fração de cura.	158
C.28	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WMG sem fração de cura.	158
C.29	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição GG de fração de cura de mistura.	159
C.30	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição G de fração de cura de mistura.	159
C.31	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição LN de fração de cura de mistura.	160
C.32	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição W de fração de cura de mistura.	160
C.33	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição EE de fração de cura de mistura.	161

C.34	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WG de fração de cura de mistura.	161
C.35	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WMG de fração de cura de mistura.	162
C.36	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição GG de fração de cura de não-mistura.	162
C.37	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição G de fração de cura de não-mistura.	163
C.38	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição LN de fração de cura de não-mistura.	163
C.39	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição W de fração de cura de não-mistura.	164
C.40	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição EE de fração de cura de não-mistura.	164
C.41	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WG de fração de cura de não-mistura.	165
C.42	Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WMG de fração de cura de não-mistura.	165

Lista de Tabelas

3.1	Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=20$, sem censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$	46
3.2	Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=20$, com 35% de censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$	47
3.3	Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=50$, sem censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$	48
3.4	Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=50$, com 35% de censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$	49
3.5	Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=100$, sem censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$	50
3.6	Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=100$, com 35% de censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$	51
3.7	Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=500$, sem censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$	52
3.8	Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=500$, com 35% de censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$	53
3.9	CV médio, PC média e Proporção de parâmetros dentro do IC das estimações obtidas pelos métodos clássico e Bayesianos, para todos os tamanhos amostras considerando dados sem censura e com 35% de censura.	54
4.1	Variáveis utilizadas presentes nos dados de pacientes com câncer internados na UTI do INCA.	55
4.2	Testes <i>logrank</i> e de Wilcoxon-Gehan utilizados para as covariáveis consideradas no estudo de pacientes com câncer.	57
4.3	Testes <i>logrank</i> e de Wilcoxon-Gehan utilizados para a covariável <i>gptumor</i>	57
4.4	Seleção de covariáveis usando o modelo gama generalizado.	58
4.5	Testes da proporcionalidade das taxas de falha no modelo ajustado.	59
4.6	Resultados do ajuste do modelo de Cox para os dados do estudo e as razões de taxas de falhas (RTF).	60

4.7	Logaritmo da função de verossimilhança e resultados do Erro Padrão (EP), teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos sem fração de cura.	62
4.8	Resultados das posteriores assumindo o modelo sem fração de cura.	64
4.9	Logaritmo da função de verossimilhança e resultados do teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos de fração de cura de mistura.	66
4.10	Resultados das posteriores assumindo o modelo de fração de cura de mistura.	68
4.11	Logaritmo da função de verossimilhança e resultados do teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos de fração de cura de não-mistura.	70
4.12	Resultados das posteriores assumindo o modelo de fração de cura de não-mistura.	72
4.13	Resumo da análise sob o enfoque clássico segundo os critérios Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC).	74
4.14	Resumo da análise sob o enfoque Bayesiano segundo os critérios EBIC e ECAIC.	75
4.15	Logaritmo da função de verossimilhança e resultados do teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos sem e com covariáveis.	76
4.16	Resumo da análise sob o enfoque clássico segundo os critérios Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos sem e com covariáveis.	77
4.17	Resultados das posteriores assumindo o modelo sem e com covariáveis.	77
4.18	Resumo da análise sob o enfoque Bayesiano segundo os critérios EBIC e ECAIC, para os modelos sem e com covariáveis.	78
4.19	Variáveis presentes nos dados de pacientes com câncer gástrico do Hospital do Câncer de Barretos.	78
4.20	Testes <i>logrank</i> e de Wilcoxon-Gehan utilizados para as covariáveis consideradas no estudo de pacientes com câncer.	79
4.21	Seleção de covariáveis usando o modelo gama generalizado.	79
4.22	Testes da proporcionalidade das taxas de falha no modelo ajustado para pacientes com câncer gástrico.	80
4.23	Resultados do ajuste do modelo de Cox para os dados do estudo e as razões de taxas de falhas (RTF), para pacientes com câncer gástrico.	81
4.24	Logaritmo da função de verossimilhança e resultados do Erro Padrão (EP), teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos sem fração de cura.	82
4.25	Resultados das posteriores assumindo o modelo sem fração de cura.	84
4.26	Logaritmo da função de verossimilhança e resultados do teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos de fração de cura de mistura.	86
4.27	Resultados das posteriores assumindo o modelo de fração de cura de mistura.	88

4.28	Logaritmo da função de verossimilhança e resultados do teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos de fração de cura de não-mistura.	90
4.29	Resultados das posteriores assumindo o modelo de fração de cura de não-mistura.	92
4.30	Resumo da análise sob o enfoque clássico segundo os critérios Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC).	94
4.31	Resumo da análise sob o enfoque Bayesiano segundo os critérios EBIC e ECAIC.	95
4.32	Logaritmo da função de verossimilhança e resultados do teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos sem e com a covariável <i>tratamento</i>	96
4.33	Resumo da análise sob o enfoque clássico segundo os critérios Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos sem e com a covariável <i>tratamento</i>	96
4.34	Resultados das posteriores assumindo, para os modelos sem e com a covariável <i>tratamento</i>	97
4.35	Resumo da análise sob o enfoque Bayesiano segundo os critérios EBIC e ECAIC, para os modelos sem e com a covariável <i>tratamento</i>	97

Lista de Siglas

CV: Coeficiente de Variação.
DP: Desvio-Padrão.
EE: Exponencial Exponenciada.
EP: Erro Padrão.
EQM: Erro Quadrático Médio.
FCT: Faculdade de Ciências e Tecnologia.
GG: Gama Generalizada.
IC: Intervalo de Confiança/Credibilidade.
INCA: Instituto Nacional de Câncer.
KM: Estimador de Kaplan-Meier.
LN: Log-Normal.
MCMC: métodos de Monte Carlo em Cadeias de Markov.
PC: Probabilidade de Cobertura.
RTF: Razões de Taxas de Falhas.
SANN: Simulated annealing.
TRV: Teste de Razão de Verossimilhanças.
UTI: Unidade de Terapia Intensiva.
WG: Weibull Generalizada.
WMG: Weibull Modificada Generalizada.

Sumário

Resumo	5
Abstract	7
Lista de Figuras	8
Lista de Tabelas	14
Lista de Siglas	19
Capítulos	
1 Introdução	23
2 Desenvolvimento teórico	27
2.1 Introdução à Análise de Sobrevivência	27
2.1.1 Tipos de Censura	27
2.1.2 Funções de Sobrevivência de Interesse	29
2.1.3 Estimador de Kaplan-Meier (KM)	30
2.2 Análise de Regressão	31
2.2.1 Modelo de Regressão de Cox	32
2.3 Seleção de modelos	34
2.4 Modelo com Fração de Cura	35
2.5 Distribuição de Weibull Modificada Generalizada (WMG)	36
2.5.1 Modelos Especiais	37
2.5.2 Modelo Padrão de fração de cura para a distribuição WMG	39
2.6 Análise Bayesiana	40
2.6.1 Estimação pontual e intervalar	40
2.6.2 Análise Bayesiana de modelos considerando fração de cura	42
3 Simulações	43
4 Aplicações	55
4.1 Pacientes com câncer internados na UTI do INCA	55
4.1.1 Análise semi-paramétrica e Seleção de Covariáveis	57
4.1.2 Modelos sem fração de cura	62
4.1.3 Modelos de fração de cura de mistura	66
4.1.4 Modelos de fração de cura de não-mistura	70
4.1.5 Resumo da aplicação com Pacientes com Câncer na UTI do INCA	74
4.1.6 Acrescentando covariáveis ao modelo selecionado para a aplicação com Pacientes com Câncer na UTI do INCA	76

4.2	Dados de câncer gástrico do Hospital do Câncer de Barretos	78
4.2.1	Análise semi-paramétrica e Seleção de Covariáveis	79
4.2.2	Modelos sem fração de cura	82
4.2.3	Modelos de fração de cura de mistura	86
4.2.4	Modelos de fração de cura de não-mistura	90
4.2.5	Resumo da aplicação com Pacientes com Câncer Gástrico do Hospital do Câncer de Barretos	94
4.2.6	Acrescentando covariáveis ao modelo selecionado para a aplicação com Pacientes com Câncer Gástrico do Hospital do Câncer de Barretos	96
5	Conclusão	99
	Referências	100
	Apêndices	
A	Parametrização das distribuições utilizadas	107
A.1	Gama Generalizada (GG)	107
A.2	Gama	108
A.3	Log-Normal (LN)	108
A.4	Weibull	109
A.5	Weibull Generalizada (WG)	109
A.6	Exponencial Exponenciada (EE)	110
B	Programas	111
B.1	Simulação	111
B.1.1	Geração dos dados	111
B.1.2	Estimações	114
B.2	Aplicações	126
C	Gráficos da densidade a posteriori, autocorrelação e histórico de séries temporais das amostras simuladas	145
C.1	Pacientes com câncer internados na UTI do INCA	145
C.2	Dados de câncer gástrico do Hospital do Câncer de Barretos	156

Introdução

A análise de sobrevivência é uma das áreas da estatística com grande aplicabilidade em diversas áreas especialmente em medicina e engenharia. Em análise de sobrevivência, a resposta é por natureza longitudinal. Nos estudos clínicos, particularmente no estudo de doenças crônicas, a medida da sobrevida dos pacientes torna-se um critério importante para avaliar a eficácia dos tratamentos adaptados a determinadas doenças crônicas. Isto se dá através do tempo de sobrevida dos doentes, desde a aplicação do tratamento até a ocorrência do evento de interesse. Esse evento de interesse pode ser o óbito do paciente, o tempo de alta hospitalar ou o tempo da ocorrência de uma recidiva.

A análise de sobrevivência é uma técnica estatística aplicada quando se pretende analisar um fenômeno em relação a um período de tempo, isto é, ao tempo transcorrido entre um evento inicial, no qual o indivíduo entra em um estado particular, e um evento final, que modifica este estado. Define-se sobrevida como o intervalo de tempo desde a entrada do indivíduo no estudo (data do diagnóstico) até a ocorrência do evento de interesse (falha ou óbito), ou até ao que se denomina censura.

Em estudos de análise de sobrevivência, avalia-se o tempo até que certo evento ocorra. Entretanto, existem casos em que parte dos indivíduos pode não apresentar o evento de interesse. Por exemplo, se um grupo de pacientes que passou por um tratamento de um mesmo tipo de câncer é acompanhado para se conhecer o tempo até o reaparecimento da doença, uma parcela deles poderá ter recidivas e outra não apresentar a recorrência do evento. Nesse caso temos a presença de fração de cura. Isso motiva o desenvolvimento de modelos estatísticos para análise de sobrevivência na presença de fração de cura, isto é, incluindo os pacientes imunes no estudo.

Uma situação comum na presença de fração de cura pode ser visualizada no gráfico do estimador não-paramétrico de Kaplan-Meier para a função de sobrevivência onde a curva permanece horizontal após um determinado tempo. Este fato traduz-se graficamente na estabilização da curva da estimativa de Kaplan-Meier da função de sobrevivência, como ilustra a Figura 1.1.

Muitas vezes, ao se analisar os dados de sobrevivência, percebe-se que há um grande número de indivíduos censurados à direita, ou seja, indivíduos que apresentam o evento de interesse após o início do estudo. Esse fato pode ser uma indicação de que, em parte dos indivíduos em estudo, o evento de interesse nunca acontecerá, ou seja, pode se ter uma fração da população que é imune (ou que se cura). A literatura apresenta muitos artigos com diferentes modelos de sobrevivência na presença de fração de cura, nesse trabalho utiliza-se [54] como principal referência para os modelos de mistura e não mistura de fração de cura baseados na distribuição Weibull Modificado Generalizada.

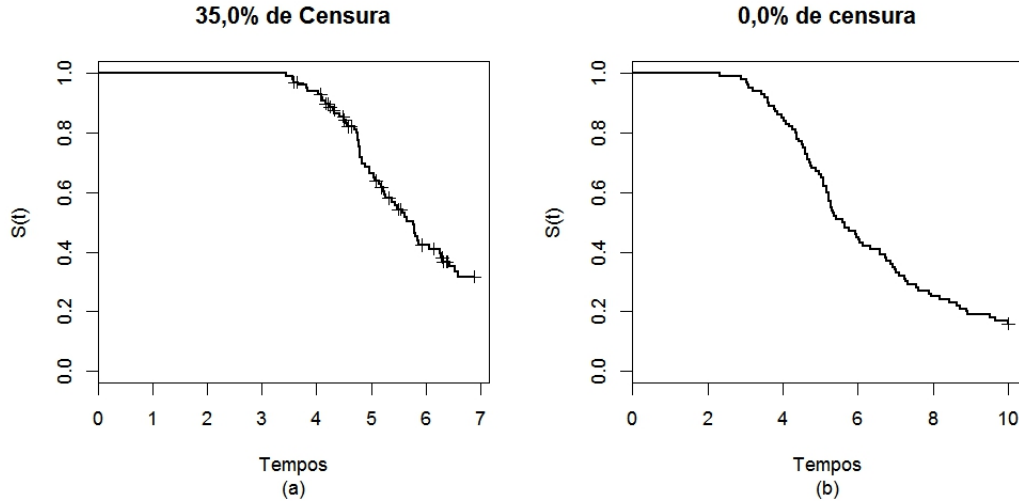


Figura 1.1: Dados simulados para uma amostra de tamanho $n=100$ a partir da distribuição Weibull Modificada Generalizada com fração de cura de 15% com censura de 35% (a) e sem censura (b).

Outra referência importante nesta área é o artigo introduzido por [19] em que é apresentado um modelo de mistura para modelar os casos em que há uma proporção π de indivíduos curados. Em [2], modelos de sobrevivência com fração de cura são construídos a partir de interpretações médicas baseadas no número não observado de células suscetíveis a ter câncer, que um indivíduo possui.

O aumento nos últimos anos do número de pessoas com doenças crônico-degenerativas tem-se constituído em um grande desafio para os serviços de saúde e para a sociedade. Isso motiva estudos envolvendo novos modelos para análise de sobrevivência pois essas técnicas estatísticas têm grande importância na análise de dados associados a essas doenças, na descoberta de fatores de riscos e para a realização de previsões.

A aplicação de novos modelos estatísticos nos estudos, que buscam conhecer as variáveis relevantes que explicam este alto número de casos envolvendo doenças crônicas, pode levar a conclusões mais claras e confiáveis, a fim de buscar gestão de saúde mais eficientes.

A proposta deste trabalho é estudar o modelo de mistura padrão e modelos de não-mistura, supondo a distribuição Weibull Modificada Generalizada para os tempos de vida dos indivíduos em risco, introduzidos por [54]. Além disso, tem-se como foco, o estudo de diferentes técnicas clássicas e Bayesianas existentes na literatura para verificar a adequabilidade do modelo proposto para a análise dos dados.

Neste trabalho é analisado o modelo de mistura e de não-mistura para a fração de cura para a distribuição Weibull Modificada Generalizada, comparando com os modelos de sobrevivência, sob os enfoques Bayesiano e clássico. Na abordagem Bayesiana realizou-se um estudo da sensibilidade nas inferências obtidas, assumindo diferentes distribuições a priori ou diferentes escolhas dos valores dos hiperparâmetros destas distribuições, para os modelos de mistura e de não-mistura da fração de cura além do modelo sem a presença da fração de cura. Explorando a literatura Bayesiana propõe-se distribuições a priori informativas e não-informativas, para a estimação dos parâmetro utilizando-se os métodos MCMC (Monte Carlo em cadeias de Markov) através do algoritmo de Metropolis-Hastings. Para ilustração da metodologia proposta, apresentou-se um estudo de simulação e realizaram-se duas aplicações com dados reais de pacientes com câncer.

Para o desenvolvimento do trabalho desenvolveu-se programas computacionais para a análise estatística dos dados assumindo os modelos propostos, utilizando principalmente, o *software* R e o *software* OpenBugs.

Desenvolvimento teórico

Nessa seção é apresentado a teoria necessária para a realização do trabalho, tais como introdução a análise de sobrevivência, tipos de censura, funções de sobrevivência de interesse, estimador de Kaplan-Meier (KM), análise de regressão, modelo de regressão de Cox, seleção de modelos, modelo padrão de fração de cura, distribuição de Weibull Modificada Generalizada (WMG), modelo Padrão de fração de cura para a distribuição WMG e análise Bayesiana.

2.1 Introdução à Análise de Sobrevivência

Em modelos de sobrevivência, o objetivo principal é o tempo até a ocorrência de um determinado evento de interesse, em geral denominado tempo de sobrevivência, vida ou de falha. O tempo até a ocorrência da morte de um paciente com câncer, a recidiva de um tumor, ou de um indivíduo quitar uma dívida com uma instituição financeira, são todos fenômenos que podem ser objetos de estudo em análise de sobrevivência.

No desenvolvimento deste texto, consideramos que os indivíduos são as unidades em estudo e o conjunto de dados observados é referente ao tempo de vida dos indivíduos.

A principal característica relacionada a estes dados, diz respeito à presença de observações incompletas, denominado como censura, e esta ocorre por vários motivos, dentre elas, o abandono do tratamento, a saída do estudo por outros motivos além do estudado ou a não ocorrência da falha até o término do experimento. É importante observar, que mesmo sendo observações parciais, elas fornecem alguma informação sobre o tempo até a falha e não devem ser omitidas na análise do problema.

2.1.1 Tipos de Censura

Existem vários tipos de censuras, dentre elas citam-se:

- Censura tipo I: Ocorre quando o experimento é conduzido até um período de tempo pré-definido.
- Censura tipo II: É aquela em que o estudo termina após ocorrer o evento de interesse em um número pré-estabelecido de indivíduos.
- Censura aleatória: Acontece quando há uma falha por outros motivos no ocorrer do estudo sem ter relação ao evento de interesse.

O mecanismo de censura pode ser informativo e não informativo. Diz-se que a censura é não informativa, se a ocorrência independe do mecanismo que provoca a falha. Suponha que um laboratório está interessado na eficácia de uma nova droga e no decorrer do experimento os responsáveis pelo estudo retiram aqueles indivíduos que apresentaram alguma complicação de saúde devido ao tratamento com esta nova droga; neste caso, se esta diante de uma censura informativa. Neste trabalho consideram-se que o mecanismo de censura é não informativa.

Esses mecanismos de censura apresentados são conhecidos por *censura à direita*, pois o tempo de ocorrência do evento de interesse está à direita do tempo observado, esta é a situação mais comum em dados de sobrevivência. Mas existem outras duas formas de censura que podem ocorrer, a censura à esquerda e censura intervalar.

A *censura à esquerda* ocorre quando o tempo observado é maior do que o tempo de ocorrência do evento de interesse, o tempo de falha, ou seja, o evento de interesse já aconteceu quando o indivíduo foi observado. Em [31] é citado o exemplo de um estudo para determinar a idade em que as crianças aprendem a ler em uma determinada comunidade. Quando os pesquisadores começaram a pesquisa algumas crianças já sabiam ler e não lembravam quando isso ocorreu. Neste estudo pode ocorrer o chamado tempos de vida duplamente censurados, que significa que no estudo pode ocorrer tanto a censura à direita, pessoas que não sabiam ler quando iniciou-se o estudo, quanto a censura à esquerda, pessoas que já sabiam ler quando iniciou-se o estudo.

A *censura intervalar*, segundo [31], é um tipo de censura mais geral, pois em estudos em que os pacientes são acompanhados em visitas periódicas e é conhecido somente que o evento de interesse ocorreu em um intervalo de tempo, ou seja, o tempo de falha não é conhecido exatamente, sabe-se apenas que ocorreu em um intervalo. Usualmente esses dados são chamados de dados de censura intervalar. As censuras à direita e à esquerda podem ser consideradas como sendo um caso especial da censura intervalar, pois nesses casos o limite inferior e o limite superior do intervalo são iguais, zero para à esquerda e infinito para a censura à direita.

Existem ainda, em alguns estudos de sobrevivência o truncamento que é em muitas vezes confundido com a censura. O truncamento é quando um indivíduo é retirado do estudo por alguma condição. Os pacientes não são acompanhados a partir do tempo inicial, mas sim somente após um certo evento. O truncamento à esquerda é quando se deseja fazer um estudo em que os indivíduos experimentam o evento de interesse à partir de um determinado evento. Por exemplo, um estudo do tempo de vida dos moradores de uma certa localidade em informações provenientes de um banco de dados da previdência, ou seja, somente moradores aposentados fazem parte da amostra.

Para o truncamento à direita, tem-se indivíduos com uma condição que ainda não manifestaram um certo evento e são retirados da amostra. Por exemplo, pacientes de AIDS que estão infectados mas não manifestaram a doença até o momento de início do estudo são retirados da amostra. Sendo assim apenas pacientes que comprovadamente têm a doença fazem parte da amostra.

O conjunto de dados que temos interesse em analisar consiste de n observações dadas por dados bivariados na forma $(T_1, \delta_1), (T_2, \delta_2), \dots, (T_n, \delta_n)$, sendo que T_i é o tempo realmente observado para a i -ésima unidade, sendo que esta pode ter sido uma falha ou uma censura, ou seja, definimos T_i^0 o tempo de sobrevivência da i -ésima unidade e X_i o tempo até a censura da i -ésima unidade em experimento. Assim,

$$T_i = \min\{T_i^0, X_i\}, i = 1, \dots, n,$$

sendo $T_i^0, \dots, T_n^0$ variáveis aleatórias i.i.d. independentes de $X_1, \dots, X_n$. A variável indicadora δ_i é tal que $\delta_i = 0$, se a i -ésima observação é censurada e $\delta_i = 1$, se a i -ésima

observação está associada a uma falha, a qual é representada da seguinte forma,

$$\delta_i = \begin{cases} 1, & \text{se } T_i^0 \leq X_i \\ 0, & \text{se } T_i^0 > X_i \end{cases}$$

2.1.2 Funções de Sobrevivência de Interesse

Seja T uma variável aleatória não negativa, contínua, representando o tempo de falha. Usualmente o comportamento de T é especificada através da função densidade de probabilidade, $f(t)$, função de sobrevivência, $S(t)$, ou função de risco, $h(t)$.

Como o tempo de vida é não negativo, temos que $f(t) \geq 0$ para todo $t \geq 0$ e

$$\int_0^{\infty} f(t) dt = 1,$$

com função de distribuição acumulada de T definida por:

$$F(t) = P(T \leq t) = \int_0^t f(u) du \quad (2.1)$$

A função de sobrevivência é definida como a probabilidade de uma unidade sobreviver pelo menos até um instante t pré-estabelecido, isto é,

$$S(t) = P(T \geq t) = \int_t^{\infty} f(u) du = 1 - F(t) \quad (2.2)$$

Em decorrência de (2.1), a função de sobrevivência $S(t)$ apresenta as seguintes propriedades:

- i. É uma função não crescente;
- ii. $S(0) = 1$;
- iii. $\lim_{t \rightarrow \infty} S(t) = 0$

Como consequência da propriedade (iii), a função de sobrevivência definida em (2.2) é denominada função de sobrevivência própria.

A função de risco ou taxa de falha, é definida como o limite da probabilidade de um indivíduo falhar no intervalo de tempo $[t, t + \Delta t]$, dado que o indivíduo tenha sobrevivido até o instante t , e é dada por:

$$h(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t}.$$

Esta função pode ser interpretada como sendo a taxa de falha instantânea no tempo t . A função de risco descreve a distribuição do tempo de vida, ou seja, representa a forma como a taxa instantânea de falha muda com o tempo. Graficamente a função de risco pode apresentar comportamento constante, crescente, decrescente e até mesmo formas não monótonas.

Algumas funções são usuais para descrever os tempos de vida, como por exemplo, a distribuição exponencial que apresenta forma de risco constante, a distribuição de Weibull com formas crescente, decrescente e constante. No entanto, estas distribuições não são

capazes de modelar taxas de falha não monótonas, como em forma de banheira ou unimodal, o que são bem comuns em análise de sobrevivência. Com a proposta de obter uma distribuição que comporta as várias formas mencionadas, [1] propôs a distribuição Weibull Modificada Generalizada (WMG), que é capaz de modelar funções de risco monótonas e também as não monótonas.

A partir das funções $f(t)$, $S(t)$ e $h(t)$, algumas relações matemáticas podem ser definidas:

$$\begin{aligned} f(t) &= -\frac{dS(t)}{dt} \\ h(t) &= \frac{f(t)}{S(t)} \end{aligned} \quad (2.3)$$

$$H(t) = \int_0^t h(u) \, du \quad (2.4)$$

$$S(t) = \exp\{-H(t)\}$$

Sendo $H(t)$, denotada por função de risco acumulada. Desta forma, no conhecimento de uma das funções é possível encontrar as demais.

2.1.3 Estimador de Kaplan-Meier (KM)

Uma maneira de estimar a função de sobrevivência considerando um conjunto de dados com censuras, foi proposto por [4]. Este método, não paramétrico, é o mais utilizado em estudos clínicos, e é também conhecido na literatura como estimador produto-limite. Este é uma adaptação da função de sobrevivência empírica, que considera tantos intervalos de tempo quantos forem o número de falhas distintas. Os limites dos intervalos de tempo são os tempos de falha da amostra. Considere:

- $t_{(1)} < t_{(2)} < \dots < t_{(k)}$, os k tempos distintos e ordenados de falha;
- d_j o número de falhas em $t_{(j)}$, $j = 1, \dots, k$;
- n_j o número de indivíduos em risco em $t_{(j)}$, ou seja, os indivíduos que não falharam e não censuraram até o instante imediatamente anterior a $t_{(j)}$.

O estimador de Kaplan-Meier (KM) é definido por:

$$\hat{S}(t) = \prod_{j:t_j < t} \left(\frac{n_j - d_j}{n_j} \right) = \prod_{j:t_j < t} \left(1 - \left(\frac{d_j}{n_j} \right) \right). \quad (2.5)$$

A expressão (2.5) é uma função escada com degraus nos tempos observados de falha de tamanho $\frac{1}{n}$, em que n é o tamanho da amostra.

A Figura 2.1 ilustra a curva de sobrevivência estimada através de KM, em que os dados são provenientes de um estudo de coorte com 862 pacientes com câncer internados na UTI do INCA, publicado por [26], cujo objetivo foi avaliar fatores associados à sobrevivência.

Observa-se na Figura 2.1 que após um determinado tempo a curva tende a se estabilizar não havendo mais falhas. Isto sugere que os indivíduos censurados no final do experimento possam ser imunes ao risco em questão ou foram curados durante o experimento. As técnicas usuais de sobrevivência não incorporam a fração de cura.

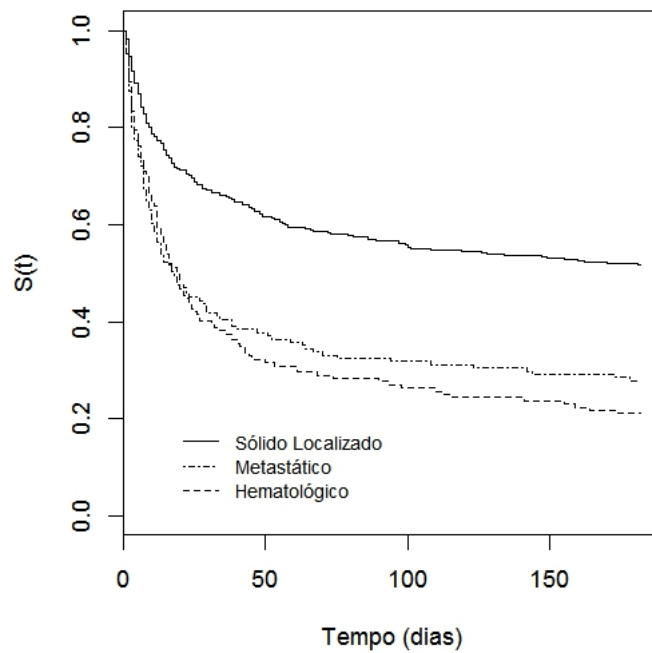


Figura 2.1: Estimativa de Kaplan-Meier para os dados de pacientes com câncer internados na UTI do INCA.

2.2 Análise de Regressão

A Análise de Regressão possibilita encontrar uma relação razoável entre as variáveis de entrada e saída, por meio de relações empíricas. A utilização desta abordagem necessita de coleta de dados e do uso de métodos estatísticos de Análise de Regressão.

Se estamos interessados na relação de uma variável de entrada com a variável resposta temos o caso de Regressão Linear Simples. Mas se queremos relacionar a variável resposta com mais de uma variável regressora, a Regressão Linear Múltipla é utilizada.

Modelos de Regressão são construídos com os objetivos:

i) Predição - Uma vez que esperamos que grande parte da variação da variável de saída seja explicada pelas variáveis de entrada, podemos utilizar o modelo para obter valores de Y correspondentes a valores de X que não estavam entre os dados. Esse procedimento é chamado de predição e, em geral, usamos valores de X que estão dentro do intervalo de variação estudado. A utilização de valores fora desse intervalo recebe o nome de extrapolação e deve ser usada com muito cuidado, pois, o modelo adotado pode não ser correto fora do intervalo estudado;

ii) Seleção de variáveis - Frequentemente, não se tem ideia de quais são as variáveis que afetam significativamente a variação de Y. Para responder a esse tipo de questão, estudos são realizados com um grande número de variáveis. A análise de regressão pode auxiliar no processo de seleção de variáveis eliminando aquelas cuja contribuição não seja importante;

iii) Estimação de parâmetros - Dado um modelo e um conjunto de dados referente às variáveis resposta e preditoras, estimar parâmetros ou ajustar um modelo aos dados significa obter valores ou estimativas para os parâmetros, por algum processo, tendo por base o modelo e os dados observados;

iv) Inferência - O ajuste de um modelo de regressão em geral tem por objetivos básicos, além de estimar os parâmetros, realizar inferências sobre eles, tais como, testes de hipóteses e intervalos de confiança.

2.2.1 Modelo de Regressão de Cox

Na prática, os dados de sobrevivência frequentemente apresentam dados censurados e presença de covariáveis que podem estar relacionadas com os tempos de sobrevivência. Dessa forma os modelos de regressão linear simples e múltiplo usuais introduzidos, não são adequados para a análise de dados de sobrevivência. Essas covariáveis devem ser incluídas na análise estatística dos dados para explicar o seu possível efeito no tempo de sobrevivência. Uma das alternativas metodológicas existentes para a análise de dados de sobrevivência que incorpora a introdução de covariáveis é o popular modelo de riscos proporcionais, usualmente considerado na análise de dados de sobrevivência médicos.

Uma escolha comum para fazer esta relação entre a variável resposta e as covariáveis é dada pelo modelo de riscos proporcionais, introduzido por [34]. A importância do modelo de Cox deve-se ao fato deste ser um modelo semi-paramétrico, de fácil interpretação e no qual a ocorrência de censura é facilmente acomodada. Além disso encontra-se disponível na maioria dos programas estatísticos. No entanto, apesar destas vantagens, é necessário algum cuidado na sua utilização. Assim ao longo, desta seção, não só fazemos uma descrição deste modelo, como também apresentamos alguns métodos gráficos e testes de hipóteses para a sua validação.

O modelo de Cox assume que a função de risco pode ser escrita na forma,

$$h(t; z) = h_0(t) e^{\theta' z},$$

onde $h_0(t)$ é uma função de risco suporte arbitrária não negativa, θ' um vetor de coeficientes de regressão transposto e $\mathbf{Z}$ um vetor de covariáveis.

Note-se que a razão das funções de risco para dois indivíduos com vetores de covariáveis fixos é constante no tempo - este é o motivo pelo qual o modelo é também conhecido por modelo de riscos proporcionais. Esta razão não depende do tempo, isto é, o risco de falha de um indivíduo em relação ao outro é constante para todos os tempos de acompanhamento. Os dois componentes multiplicativos do modelo são de naturezas distintas, um não-paramétrico e o outro paramétrico sendo este o motivo do modelo ser do tipo semi-paramétrico o que o torna flexível.

Assim, duas características deste modelo são:

- as funções de risco correspondentes a dois indivíduos diferentes com vetores de covariáveis Z_1 e Z_2 são proporcionais, pelo fato de que a razão entre estas funções em qualquer instante t

$$\frac{h(t; z_1)}{h(t; z_2)} = \frac{h_0(t) e^{\theta' z_1}}{h_0(t) e^{\theta' z_2}} = e^{\theta'(z_1 - z_2)}$$

não depende de t ;

- as covariáveis afetam a função de risco de modo multiplicativo de acordo com o fator $e^{\theta' z}$, que é designado por risco relativo.

O modelo de regressão de Cox é caracterizado pelos coeficientes θ que medem o efeito das covariáveis sobre a função de risco. Dessa maneira é necessário um método de estimação para se fazer inferência no modelo. [34] construiu uma função de verossimilhança que não depende de $h_0(t)$, permitindo assim a realização de inferências sobre θ sem que seja necessário especificar $h_0(t)$.

Considerando n indivíduos em estudo e k instantes distintos de observação de eventos, tal que $t_1 < t_2 < \dots < t_k$, em que $k \leq n$, o conjunto de risco no instante t_i , designado por R_i é definido por

$$R_i = R(t_i) = \{j : t_j \geq t_i\},$$

sendo o conjunto de índices associados aos indivíduos em observação imediatamente antes do instante t_i .

A função de verossimilhança, proposta por [34] para a realização de inferência sobre θ , é dada por

$$L(\theta) = \prod_{i=1}^t \frac{e^{\theta' z_i}}{\sum_{l \in R_i} e^{\theta' z_l}},$$

em que z_i é o vetor de covariáveis associadas ao indivíduo com evento no instante t_i . [35] designou a função $L(\theta)$ por função de verossimilhança parcial. Assim, mostrou que, embora não se trate de uma função no sentido usual, dado não permitir a obtenção de um estimador do vetor de parâmetros θ , e sob certas condições bastante gerais, verifica as propriedades usuais dos estimadores de máxima verossimilhança.

Note-se que a função de verossimilhança proposta por [34] não depende de $h_0(t)$, o que permite a realização de inferência sobre o vetor de parâmetros θ , sem que seja necessário fazer restrições à forma de $h_0(t)$.

Este fato, faz com que o método de máxima verossimilhança usual não seja utilizado, pois a presença da componente não-paramétrica $h_0(t)$ na função de verossimilhança torna este método inapropriado. Frente a tal dificuldade, [35] propôs o método de máxima verossimilhança parcial que condiciona à verossimilhança a história dos tempos de sobrevivência e censuras anteriores e desta forma elimina a função de base desconhecida [41]. Tal como [35] mostrou, o método de construção de verossimilhança parcial, sob condições de regularidade bastante fracas, leva às propriedades assintóticas usuais de inferência baseada na verossimilhança.

Caso ocorra falha em mais que um indivíduo em simultâneo dando origem a valores iguais, a função de verossimilhança parcial não poderá ser aplicada. Nesta situação, para os n indivíduos em estudo, suponha que foram observados eventos nos instantes $t_1 < t_2 < \dots < t_k$. Seja d_i o número de eventos ocorridos no instante t_i , e Z_{ij} o vetor de covariáveis explicativas associadas ao indivíduo j , cujo evento ocorre em t_i , $j = 1, \dots, d_i$, $i = 1, \dots, k$. Na comparação com o número de indivíduos pertencentes ao conjunto de risco R_i , $i = 1, \dots, k$, se o número d_i de indivíduos com evento em t_i for pequeno, então a função de verossimilhança parcial pode ser aproximada pela função, proposta por [23],

$$L(\theta) = \prod_{i=1}^k \frac{e^{\theta' s_i}}{[\sum_{l \in R_i} e^{\theta' z_l}]^{d_i}}.$$

onde $s_i = \sum_{j=1}^{d_i} Z_{ij}$, para $i = 1, \dots, k$. Esta função de verossimilhança é habitualmente usada nos programas estatísticos. Caso sejam observados um número substancial de tempos de sobrevivência iguais, então será aconselhável optar por um modelo discreto.

Segundo [39], o modelo de Cox tem como desvantagem a vulnerabilidade a certos problemas de consistência, sendo assim é necessário o desenvolvimento de modelos alternativos a serem aplicados em situações que as hipóteses do modelo de Cox não são válidas. O modelo de Cox é válido quando se considera um vetor com p covariáveis. No entanto, a proporcionalidade entre as funções de risco correspondentes aos indivíduos diferentes não se mantém, em geral, quando é omissa alguma covariável, ainda que esta seja independente das restantes. O mesmo problema se põe quando se verifica alterações na precisão com que são medidas as covariáveis. Portanto, se uma covariável é medida com erro ou é mal classificada, o estimador de verossimilhança parcial não será a melhor opção, visto que conduzirá a resultados enviesados.

2.3 Seleção de modelos

Alguns critérios comuns na literatura podem ser utilizados para seleção de modelos. Esses critérios levam em consideração a complexidade do modelo no critério de seleção. São critérios que, essencialmente, penalizam a verossimilhança, utilizando o número de variáveis do modelo e, eventualmente, o tamanho da amostra. Essa penalização é feita subtraindo-se do valor da verossimilhança uma determinada quantidade do quão complexo é o modelo.

É proposto em [12] a utilização da informação de Kullback-Leibler para a seleção de modelos. Ele estabeleceu uma relação entre a máxima verossimilhança e a informação de Kullback-Leibler desenvolvendo então um critério para estimar a informação de Kullback-Leibler, posteriormente chamado critério de informação de Akaike (AIC). Em 2007 foi proposto uma correção por [5] conhecido como AIC corrigido (CAIC).

O critério de informação Bayesiano (BIC), também chamado de critério de Schwarz, é um critério de avaliação de modelos definido em termos da probabilidade *a posteriori*, sendo assim chamado porque [68] forneceu um argumento Bayesiano para prová-lo. Segundo esses critérios, o melhor modelo será aquele que apresentar menor valor para AIC, BIC e CAIC.

Considerando uma amostra aleatória $X_1, \dots, X_n$ de tamanho n e o vetor de parâmetros θ , as estatísticas AIC, BIC e CAIC podem ser calculadas por:

$$AIC = -2l(\theta) + 2d,$$

$$BIC = -2l(\theta) + d \log(n),$$

$$CAIC = AIC + \frac{2(d+2)(d+3)}{(n-d-3)},$$

em que $l(\theta)$ representa o máximo da função de log-verossimilhança, d é o número de parâmetros do modelo e n é o número de observações em estudo.

Para o caso da metodologia Bayesiana em que as amostras das distribuições *a posteriori* para os parâmetros do modelo são obtidos usando métodos MCMC, o critério mais utilizado é o Deviance Information Criterion (DIC), mas no caso de assimetria extrema, bimodalidade ou modelos de mistura, o DIC pode não ser adequado, então optou-se pela utilização dos critérios EAIC, EBIC e ECAIC.

Uma extensão natural dos critérios AIC, BIC e CAIC, para o contexto Bayesiano, é considerar a densidade *a posteriori* dos parâmetros do modelo, baseando-se na média *a posteriori* do *deviance*, $E[D(\theta)] = \bar{D}(\theta)$, sendo que $D(\theta)$ é o *deviance*.

$$EAIC = \bar{D}(\theta) + 2d,$$

$$EBIC = \bar{D}(\theta) + d \log(n),$$

$$ECAIC = EAIC + \frac{2(d+2)(d+3)}{(n-d-3)},$$

Os critérios EAIC, EBIC e ECAIC indicam os melhores modelos quanto menor for o valor observado.

2.4 Modelo com Fração de Cura

Segundo [58], em muitos estudos de acompanhamento, pacientes que respondem bem ao tratamento podem tornar-se imunes aos sinais e sintomas da doença e assim serem considerados curados, ou seja, não suscetíveis ao evento de interesse. Neste caso, é importante estudar a proporção de pacientes que permanecem curados e analisar os efeitos das covariáveis, não só sobre o tempo de sobrevivência, como também sobre a taxa (ou fração) de cura.

Modelos para análise de dados com proporção de curados são frequentemente chamados de modelos com fração de cura.

Considere uma população heterogênea que pode ser dividida em duas subpopulações: imunes (curados) e suscetíveis (não curados). Em um estudo experimental analisado por [63], por exemplo, os autores observaram que havia uma proporção substancial de animais expostos a toxinas que continuavam vivos ao final do experimento. Segundo eles, não fazia sentido considerar tais animais como observações censuradas vindas da mesma população daqueles que haviam morrido.

Devido a essa heterogeneidade, é natural considerar um modelo construído a partir da mistura de duas distribuições: uma representando a distribuição dos tempos de falha ou sobrevivência dos não curados e outra correspondendo a uma distribuição degenerada que permita tempos de sobrevida (em princípio, infinitos) para pacientes curados [53].

O modelo de mistura de fração de cura, também conhecido como modelo padrão de fração de cura, assume que a população estudada é uma mistura de indivíduos suscetíveis, que experimentam o evento de interesse e indivíduos não suscetíveis que nunca o experimentam. Os indivíduos que não correm o risco são considerados imunes, não-suscetíveis ou curados (ver por exemplo, [53]). Diferentes formulações, paramétricas e não paramétricas, têm sido consideradas para o modelo de fração de cura.

Seguindo [53], suponha que a população é dividida em dois grupos de indivíduos: um grupo de indivíduos curados com probabilidade π e um grupo de indivíduos suscetíveis com função de sobrevivência $S_0(t) = P(T > t)$, onde T indica o tempo de vida do indivíduo com probabilidade $1 - \pi$.

Neste *modelo de mistura*, temos

$$S(t) = \pi + (1 - \pi) S_0(t), \quad (2.6)$$

em que $\pi \in (0, 1)$ e $S_0(t)$ é a função de sobrevivência para os indivíduos susceptíveis (observações não-censuradas).

Considerando uma amostra aleatória (T_i, δ_i) , $i = 1, \dots, n$, a contribuição do i -ésimo indivíduo para a função de probabilidade é dada por:

$$L_i = [f(t_i)]^{\delta_i} [S(t_i)]^{1-\delta_i}, \quad (2.7)$$

onde δ_i é uma variável indicadora de censura. Isto é, $\delta_i = 1$ para um tempo de vida observada e $\delta_i = 0$ para um tempo de vida censurado.

Assumindo o “modelo de mistura” (2.6), a função densidade de probabilidade (fdp) para o tempo de vida T (partindo de $f(t) = \frac{dF(t)}{dt}$), é dada por:

$$f(t) = (1 - \pi)f_0(t), \quad (2.8)$$

em que $F(t) = 1 - S(t)$ e $f_0(t)$ é a função densidade de probabilidade dos indivíduos susceptíveis. A substituição da densidade de mistura e da função de sobrevivência na função de verossimilhança padrão (2.7), produz a verossimilhança do modelo de mistura de fração de cura,

$$L_i = (1 - \pi)^{\delta_i} [f_0(t_i)]^{\delta_i} [\pi + (1 - \pi) S_0(t_i)]^{1-\delta_i}. \quad (2.9)$$

Portanto, o log da verossimilhança para as n observações é dada por,

$$\ln L = \sum_{i=1}^n \delta_i \ln(1 - \pi) + \sum_{i=1}^n \delta_i \ln(f_0(t_i)) + \sum_{i=1}^n (1 - \delta_i) \ln[\pi + (1 - \pi) S_0(t_i)], \quad (2.10)$$

em que $\sum_{i=1}^n \delta_i$ é o número de observações sem censura. Parametricamente, as distribuições Exponencial e Weibull são escolhas usuais para $S_0(t)$ (ver por exemplo, [78]).

Em [51] foi sugerida uma formulação de modelagem alternativa de não-mistura que define uma assíntota para o risco acumulado e, conseqüentemente, para a fração de cura. Neste caso, a função de sobrevivência para o modelo de não-mistura da fração de cura é dada por,

$$S(t) = \pi^{F_0(t)}, \quad (2.11)$$

onde $0 < \pi < 1$ representa a probabilidade dos indivíduos curados e $F_0(t) = 1 - S_0(t)$ é a função de distribuição para os indivíduos susceptíveis.

De (2.11), a função de sobrevivência e de risco para o modelo de não-mistura da fração de cura pode ser escrita, respectivamente, da seguinte forma

$$S(t) = \exp[F_0(t) \ln \pi], \quad h(t) = -(\ln \pi) f_0(t). \quad (2.12)$$

Uma vez que $h(t) = \frac{f(t)}{S(t)}$, a contribuição do i -ésimo indivíduo para a função de verossimilhança (2.7) é dada por

$$L_i = [h(t_i)]^{\delta_i} S(t_i), \quad (2.13)$$

isto é,

$$L_i = [-(\ln \pi) f_0(t_i)]^{\delta_i} \exp[F_0(t_i) \ln \pi]. \quad (2.14)$$

Supondo uma amostra aleatória de tamanho n , a função de log-verossimilhança é dada por

$$\ln L = \sum_{i=1}^n \delta_i \ln(-(\ln \pi)) + \sum_{i=1}^n \delta_i \ln(f_0(t_i)) + (\ln \pi) \sum_{i=1}^n [1 - S_0(t_i)]. \quad (2.15)$$

2.5 Distribuição de Weibull Modificada Generalizada (WMG)

Segundo [1], várias distribuições têm sido utilizadas para descrever tempos de sobrevivência, sendo a distribuição de Weibull a mais amplamente usada para modelar dados de sobrevivência para fenômenos com taxas de falha (funções de risco) monótonas. Entretanto, esta distribuição não apresenta flexibilidade para modelar funções de risco não monótonas, dessa forma novas classes de distribuições foram propostas baseadas em modificações da distribuição de Weibull com a finalidade de acomodar funções de risco não monótonas.

A distribuição Weibull Modificada Generalizada (WMG), proposta por [1], apresenta flexibilidade em acomodar diversas formas para a funções de risco, podendo assim ser utilizada em uma variedade de problemas para modelagem de dados de sobrevivência.

Outra característica desta distribuição é apresentar vários sub modelos, como as distribuições: exponencial, exponencial exponenciada, Weibull, Weibull exponenciada, Weibull modificada, dentre outras.

Uma variável aleatória T , não negativa ($t > 0$) e contínua, tem distribuição WMG com parâmetros $\alpha > 0$, $\gamma \geq 0$, $\lambda \geq 0$ e $\beta > 0$, se sua função densidade de probabilidade é dada por

$$f_0(t) = \frac{\lambda \beta t^{\gamma-1} (\gamma + \alpha t) e^{\alpha t - \lambda t^\gamma e^{\alpha t}}}{(1 - e^{-\lambda t^\gamma e^{\alpha t}})^{1-\beta}}, \quad (2.16)$$

em que λ é o parâmetro de escala, γ e β são parâmetros de forma e α funciona como um fator de aceleração no tempo. Sendo assim, o termo $e^{\alpha t}$ pode ser visto como um fator acelerador na sobrevivência, significando que conforme o tempo aumenta, o parâmetro α funciona como um fator de fragilidade na sobrevivência do indivíduo.

As funções de sobrevivência e de risco são expressas respectivamente por

$$S_0(t) = 1 - \left(1 - e^{-\lambda t^\gamma e^{\alpha t}}\right)^\beta, \quad (2.17)$$

e

$$h_0(t) = \frac{\lambda \beta t^{\gamma-1} (\gamma + \alpha t) e^{\alpha t - \lambda t^\gamma e^{\alpha t}} (1 - e^{-\lambda t^\gamma e^{\alpha t}})^{\beta-1}}{1 - (1 - e^{-\lambda t^\gamma e^{\alpha t}})^\beta}, \quad (2.18)$$

Uma característica da distribuição WMG é que sua função de risco acomoda diversas formas dependendo basicamente dos parâmetros de forma γ e β com as seguintes propriedades:

- i. Para $\gamma \geq 1$, $0 < \beta < 1$ e $\forall t > 0$, $h'(t) > 0$, $h(t)$ é crescente;
- ii. Para $0 < \gamma < 1$, $\beta > 1$ e $\forall t > 0$, $h'(t) < 0$, $h(t)$ é decrescente;
- iii. Para $0 < \gamma < 1$ e $\beta \rightarrow \infty$, $h(t)$ é unimodal;
- iv. Se $\alpha = 0$, $\gamma > 1$ e $\gamma\beta < 1$, $h(t)$ tem forma de banheira;
- v. Quando $\beta = 1$, $h(t)$ tem forma de banheira.

A geração dos valores aleatórios da distribuição WMG, pode-se obter através da expressão¹

$$t = \frac{\gamma W \left(\frac{\alpha \left(-\frac{\log(1-u^{1/\beta})}{\lambda} \right)^{1/\gamma}}{\gamma} \right)}{\alpha} \quad (2.19)$$

em que $u \sim U[0, 1]$, $\lambda > 0$, $\gamma \geq 0$, $\alpha \geq 0$ e $\beta > 0$, W é conhecido como a função Ômega de Lambert que pode ser obtida pela a função inversa da expressão $f(w) = we^w$.

Na Figura 2.2 apresenta-se o comportamento das funções: a) densidade, b) distribuição c) de sobrevivência e d) de risco, da distribuição WMG, com os parâmetros $\alpha = 0.5$, $\lambda = 0.5$, $\gamma = 0.5$ e 2.0 e $\beta = 0.5$, 1.5 e 10.0 .

2.5.1 Modelos Especiais

A distribuição WMG apresenta como casos particulares alguns modelos listados abaixo, em que apresentamos a função de sobrevivência.

¹Expressão obtida através do programa *Mathematica*®, sendo a função inversa da função de distribuição acumulada da distribuição WMG.

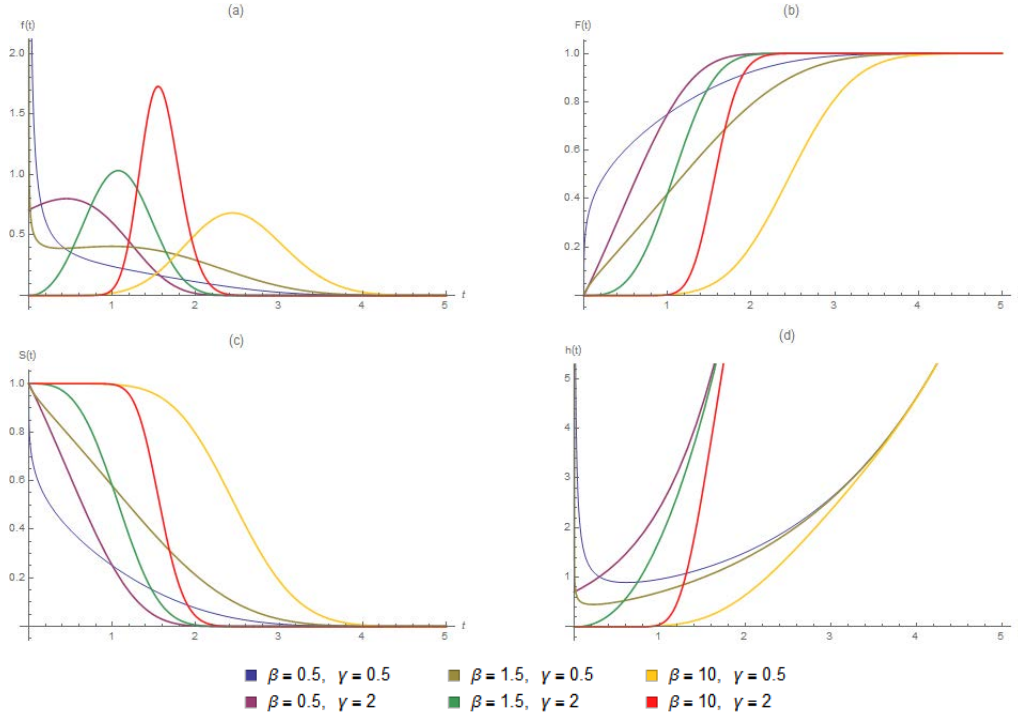


Figura 2.2: (a) Funções densidade, (b) de distribuição, (c) de sobrevivência e (d) de risco da distribuição WMG.

- *Modelo exponencial.*

O modelo WMG com $\alpha = 0$, $\beta = 1$ e $\gamma = 1$ se reduz a $S(t) = e^{-\lambda t}$.

- *Modelo Rayleigh*

Para $\alpha = 0$, $\beta = 1$ e $\gamma = 2$, obtemos $S(t) = e^{-\lambda t^2}$.

- *Modelo Rayleigh generalizada*

Fixando $\alpha = 0$ e $\gamma = 2$, temos $S(t) = 1 - \left(1 - e^{-\lambda t^2}\right)^\beta$.

- *Modelo Weibull*

Para $\alpha = 0$ e $\beta = 1$, $S(t)$ para o modelo WMG se reduz a $S(t) = e^{-\lambda t^\gamma}$.

- *Modelo exponencial exponenciada*

Se $\alpha = 0$ e $\gamma = 1$, temos $S(t) = 1 - (1 - e^{-\lambda t})^\beta$.

- *Modelo Weibull exponenciada*

Se apenas $\alpha = 0$, obtemos $S(t) = 1 - (1 - e^{-\lambda t^\gamma})^\beta$.

- *Modelo Weibull modificada*

$S(t)$ para a distribuição WMG com $\beta = 1$ se reduz a $S(t) = e^{-\lambda t^\gamma e^{\alpha t}}$.

2.5.2 Modelo Padrão de fração de cura para a distribuição WMG

Assumindo o modelo de mistura (2.6), a função de log-verossimilhança (2.10) para π , α , β , λ e γ é dada por

$$l(\pi, \alpha, \beta, \lambda, \gamma) = r \ln(1 - \pi) + r \ln \beta + r \ln \lambda + (\beta - 1) \nu + (\gamma - 1) \eta + \omega - \lambda A_1(\alpha, \gamma) + \alpha \varphi + A_2(\pi, \alpha, \beta, \lambda, \gamma) \quad (2.20)$$

em que

$$r = \sum_{i=1}^n \delta_i, \quad \nu = \sum_{i=1}^n \delta_i \ln [1 - e^{-\lambda t_i^\gamma e^{\alpha t_i}}], \quad \varphi = \sum_{i=1}^n \delta_i t_i, \\ \eta = \sum_{i=1}^n \delta_i \ln t_i, \quad \omega = \sum_{i=1}^n \delta_i \ln [\gamma + \alpha t_i], \quad A_1(\alpha, \gamma) = \sum_{i=1}^n \delta_i t_i^\gamma e^{\alpha t_i},$$

e

$$A_2(\pi, \alpha, \beta, \lambda, \gamma) = \sum_{i=1}^n (1 - \delta_i) \ln \left[\pi + (1 - \pi) \left(1 - (1 - e^{-\lambda t_i^\gamma e^{\alpha t_i}})^\beta \right) \right].$$

Assumindo o modelo de não-mistura (2.11), a função de log-verossimilhança (2.15) para π , α , β , λ e γ é dada por

$$l(\pi, \alpha, \beta, \lambda, \gamma) = r \ln(-(\ln \pi)) + r \ln \beta + r \ln \lambda + (\beta - 1) \nu + (\gamma - 1) \eta + \omega - \lambda A_1(\alpha, \gamma) + \alpha \varphi + (\ln \pi) A_3(\alpha, \beta, \lambda, \gamma) \quad (2.21)$$

em que

$$r = \sum_{i=1}^n \delta_i, \quad \nu = \sum_{i=1}^n \delta_i \ln [1 - e^{-\lambda t_i^\gamma e^{\alpha t_i}}], \quad \varphi = \sum_{i=1}^n \delta_i t_i, \\ \eta = \sum_{i=1}^n \delta_i \ln t_i, \quad \omega = \sum_{i=1}^n \delta_i \ln [\gamma + \alpha t_i], \quad A_1(\alpha, \gamma) = \sum_{i=1}^n \delta_i t_i^\gamma e^{\alpha t_i},$$

e

$$A_3(\alpha, \beta, \lambda, \gamma) = \sum_{i=1}^n \left[1 - \left(1 - (1 - e^{-\lambda t_i^\gamma e^{\alpha t_i}})^\beta \right) \right].$$

2.6 Análise Bayesiana

Considere uma quantidade de interesse desconhecida θ (tipicamente não observável). A informação de que dispomos sobre θ , resumida probabilisticamente através de $p(\theta)$, pode ser aumentada observando-se uma quantidade aleatória X relacionada a θ , com distribuição amostral $p(x|\theta)$ definindo esta relação. A ideia de que após observar $X = x$ a quantidade de informação sobre θ aumenta é intuitiva e o teorema de Bayes é a regra de atualização utilizada para quantificar este aumento de informação, da seguinte forma:

$$p(\theta|x) = \frac{p(\theta, x)}{p(x)} = \frac{p(x|\theta)p(\theta)}{p(x)} = \frac{p(x|\theta)p(\theta)}{\int p(\theta, x) d\theta} = \frac{L(\theta; x)\pi(\theta)}{\int L(\theta; x)\pi(\theta) d\theta}$$

A expressão $\frac{1}{p(x)}$ não depende de θ , funcionando como uma constante normalizadora de $p(\theta|x)$.

Para um valor fixo de x , a função $l(\theta; x) = p(x|\theta)$ fornece a verossimilhança de cada um dos possíveis valores de θ , enquanto $p(\theta)$ é chamada *distribuição a priori* de θ . Estas duas fontes de informação, *priori* e verossimilhança, são combinadas levando à *distribuição a posteriori* de θ , $p(\theta|x)$. Assim, a forma usual do teorema de Bayes é

$$p(\theta|x) \propto l(\theta; x)p(\theta),$$

ou seja, a *distribuição a posteriori* é proporcional a função de verossimilhança multiplicada pela *distribuição a priori*.

Esta forma simplificada do teorema de Bayes é útil em problemas que envolvam estimação de parâmetros já que o denominador é apenas uma constante normalizadora. Em outras situações, como seleção de modelos, este termo tem um papel crucial.

É intuitivo também que a probabilidade *a posteriori* de um particular conjunto de valores de θ será pequena se $p(\theta)$ ou $l(\theta; x)$ for pequena para este conjunto. Em particular, se atribuímos probabilidade *a priori* igual a zero para um conjunto de valores de θ então a probabilidade *a posteriori* será zero qualquer que seja a amostra observada.

A constante normalizadora da distribuição a posteriori pode ser facilmente recuperada pois

$$p(\theta|x) = k p(x|\theta) p(\theta)$$

em que

$$k^{-1} = \int p(x|\theta) p(\theta) d\theta = E_{\theta} [p(X|\theta)] = p(x)$$

é chamada distribuição preditiva. Esta é a distribuição esperada para observação x dado θ . Assim antes de observar X podemos checar a adequação da *priori* fazendo previsões via $p(x)$. Se X observado recebia pouca probabilidade preditiva então o modelo deve ser questionado.

2.6.1 Estimação pontual e intervalar

Partindo da necessidade de resumir a informação contida *a posteriori* através de poucos valores numéricos, tem-se como forma mais simples a estimação pontual. Nela, a distribuição *a posteriori* é resumida através de um único número, $\hat{\theta}$.

Da perspectiva de *Teoria da Decisão*, para escolher uma estimativa pontual $\hat{\theta}$ de alguma quantidade deve-se trabalhar como se $\hat{\theta}$ fosse θ . Para isso, é necessário especificar uma função de perda $l(\hat{\theta}|\theta)$, medindo as consequências de trabalhar como se o verdadeiro valor da quantidade de interesse θ fosse $\hat{\theta}$. A perda esperada *a posteriori* se $\hat{\theta}$ for usado é:

$$l(\hat{\theta}|x) = \int_{\Theta} l(\hat{\theta}|\theta) p(\theta|x) d\theta$$

e a estimativa de Bayes é o valor de $\hat{\theta}$ que minimiza $l(\hat{\theta}|\theta)$ em Θ . O estimador de Bayes é a função dos dados $\theta^*(x) = \operatorname{argmin}_{\theta \in \Theta} l(\hat{\theta}|x)$ [40].

A estimativa de Bayes depende da função de perda escolhida. A função de perda geralmente deve ser escolhida tendo como base as utilizações previstas da estimativa. No entanto, algumas funções de perda convencionais têm sido sugeridas para as situações em que não estão previstas utilizações particulares. Estas funções de perda produzem estimativas que podem ser consideradas como simples descrições da localização da distribuição posterior.

Se a função de perda é quadrática, então a estimativa de Bayes é a média *a posteriori*, $\hat{\theta} = E[\theta|x]$ com:

$$E[\theta_i|x] = \int_{\Theta} \theta_i p(\theta|x) d\theta, \quad i = 1, \dots, k$$

Se a função de perda é a função 0-1 definida como:

$$l(\hat{\theta}|\theta) = \begin{cases} 1, & \text{se } |\hat{\theta} - \theta| > \varepsilon \\ 0, & \text{se } |\hat{\theta} - \theta| < \varepsilon \end{cases}, \quad \forall \varepsilon > 0$$

então, neste caso pode-se mostrar que a estimativa de Bayes é a moda *a posteriori*, tal que

$$\begin{aligned} p(\hat{\theta}|x) &= \max_{\theta \in \Theta} p(\theta|x) \\ &= \max_{\theta \in \Theta} \{p(\theta) l(\theta|x)\} \end{aligned}$$

Se θ é univariada e a função de perda é linear, a estimativa de Bayes é a mediana *a posteriori* tal que o vetor de medianas *a posteriori*: $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_k)$ é definido como:

$$\begin{aligned} p\{\theta_i \geq \hat{\theta}_i|x\} &\geq \frac{1}{2} \\ p\{\theta_i \leq \hat{\theta}_i|x\} &\geq \frac{1}{2}, \quad i = 1, \dots, k \end{aligned}$$

Segundo [37], a principal restrição da estimação pontual é que quando estimamos um parâmetro através de um único valor numérico, toda a informação presente na distribuição *a posteriori* é resumida através desse número. É importante também associar alguma informação sobre o quão precisa é a especificação desse número.

O intervalo de confiança Bayesiano, ou intervalo de credibilidade, é obtido de uma região de Θ que contenha uma parte substancial da massa probabilística *a posteriori* [60].

Então, *IC* é um intervalo de credibilidade de $100(1 - \alpha)\%$ para θ se:

$$P[\theta \in IC] \geq 1 - \alpha$$

Intervalos de credibilidade são invariantes. Assim, para qualquer intervalo de credibilidade $100(1 - \alpha)\%$ $IC = [a, b]$ de θ , então $[\varphi(a), \varphi(b)]$ também é um intervalo de credibilidade $100(1 - \alpha)\%$ para $\varphi(\theta)$, onde $\varphi(\cdot)$ é uma transformação um a um.

Dados os infinitos intervalos de credibilidade $100(1 - \alpha)\%$ que podem ser construídos, é interessante apenas aquele com o menor tamanho (comprimento, área, volume) possível.

Obtidos em regiões onde todos os pontos da região tem maior densidade de probabilidade (HPD) a *posteriori* que os pontos exteriores à região.

Os intervalos HPD são invariantes a transformações lineares $\phi(\cdot)$, mas não a transformações um a um $\varphi(\cdot)$. Assim, para qualquer HPD $100(1 - \alpha)\%$ $IC = [a, b]$ de θ , então $[\varphi(a), \varphi(b)]$ também é um intervalo de HPD $100(1 - \alpha)\%$ para $\phi(\theta)$, mas $[\varphi(a), \varphi(b)]$ não será um HPD $100(1 - \alpha)\%$ para $\varphi(\theta)$. No entanto, mesmo perdendo a característica de ser HPD, $[\varphi(a), \varphi(b)]$ ainda será um intervalo de credibilidade $100(1 - \alpha)\%$ para $\varphi(\theta)$.

2.6.2 Análise Bayesiana de modelos considerando fração de cura

Para uma análise Bayesiana dos modelos de mistura e de não-mistura para a distribuição WMG, introduzido na seção 2.5.2, podemos assumir *prioris* não informativas, tais como uma *distribuição a priori* $U(0, 1)$ para a probabilidade de cura π e a *distribuição a priori* $Gama(0.01, 0.01)$ para os parâmetros de escala (λ), do fator de aceleração (α) e de forma (β e γ), onde $Gama(a, b)$ indica a distribuição gama com média $\frac{a}{b}$ e variância $\frac{a}{b^2}$. Assume-se também a independência *a priori* entre π , α , β , λ e γ .

Sendo assim, a *distribuição a posteriori* conjunta para o modelo padrão de fração de cura pode ser obtida por

$$p(\theta|x) = L(\pi, \alpha, \beta, \lambda, \gamma; x) p(\theta) \quad (2.22)$$

em que $p(\theta|x)$ é a distribuição a posteriori conjunta, $p(\theta) = p(\pi) p(\alpha) p(\beta) p(\lambda) p(\gamma)$ é a distribuição a priori de $\theta = (\pi, \alpha, \beta, \lambda, \gamma)$ e $L(\pi, \alpha, \beta, \lambda, \gamma; x)$ é a função de verossimilhança da distribuição WMG.

Na presença de um vetor de covariáveis $x = (x_1, x_2, \dots, x_k)'$ que afetam os parâmetros π , α e λ , mas que não afetam os parâmetros de forma β e γ , pode-se supor o seguinte modelo de regressão

$$\begin{aligned} \lambda_i &= v_0 \exp(v_1 x_{1i} + \dots + v_k x_{ki}), \\ \alpha_i &= \phi_0 \exp(\phi_1 x_{1i} + \dots + \phi_k x_{ki}). \end{aligned}$$

e

$$\ln\left(\frac{\pi_i}{1-\pi_i}\right) = \tau_0 + \tau_1 x_{1i} + \dots + \tau_k x_{ki}. \quad (2.23)$$

Sendo assim, para os modelos de mistura e de não-mistura introduzidos na seção 2.5.2, pode-se considerar uma *priori* com distribuição $Gamma(0.01, 0.01)$ para os parâmetros de regressão v_0 , ϕ_0 e τ_0 e uma *priori* com distribuição $N(0, 100)$ para os parâmetros de regressão v_l , ϕ_l e τ_l , $l = 1, \dots, k$. Também pode-se assumir a independência *a priori* entre os parâmetros.

Sumários *a posteriori* de interesse são obtidos a partir de amostras simuladas para a *distribuição a posteriori conjunta* usando métodos de Monte Carlo em Cadeias de Markov (MCMC) como o algoritmo amostrador de Gibbs (ver maiores detalhes em [42]) ou o algoritmo Metropolis-Hastings (ver maiores detalhes em [29]).

Neste trabalho, foram utilizados os programas Mathematica, R, SAS e OpenBUGS para a obtenção dos resultados sob o enfoque clássico e Bayesiano.

Simulações

Neste capítulo são descritas as simulações computacionais realizadas via software R (versão 3.0.1), juntamente com os resultados numéricos obtidos, os programas utilizados para essa seção são apresentados no apêndice B. As simulações têm por objetivo gerar dados de sobrevivência com fração de curados a fim de testar a significância dos parâmetros do modelo de fração de cura.

Foram geradas 1000 amostras para os tamanhos amostrais $n = (20, 50, 100, 500)$, da distribuição Weibull Modificada Generalizada com fração de cura, sem censura e com 35% de censura, considerando os seguintes valores para os parâmetros: $\alpha = 0.10$, $\beta = 9.00$, $\gamma = 0.80$ e $\lambda = 0.40$ e da fração de cura $\pi = 0.15$. A geração dos dados simulados foi realizada, basicamente, em 4 passos:

- 1º: calculou-se aleatoriamente o número de indivíduos não curados (nc), através da distribuição binomial com parâmetros, n e π ;
- 2º: gerou-se os tempos através da função geradora de valores aleatórias (2.19), para os indivíduos não curados;
- 3º: definiu-se as censuras aleatórias novamente através da distribuição binomial com parâmetros nc e $1 - \text{censura}$;
- 4º: definiu-se o tempo dos indivíduos curados sendo o máximo tempo observado na geração do 2º passo.

Foram consideradas as estimativas clássicas e Bayesianas para os parâmetros do modelo. Para tanto, foram adotados 3 conjuntos de *prioris* não informativas, no primeiro conjunto de *prioris* utilizou-se *distribuição à priori* $U(0, 1)$ para todos os parâmetros, no segundo conjunto de *prioris* utilizou-se a *distribuição à priori* $Gamma(1; 0, 1)$ todos os parâmetros e no terceiro conjunto de *prioris* utilizou-se para alguns parâmetros diferentes *prioris*, sendo que para o parâmetro $\pi \sim Beta(1; 1)$, para o parâmetro $\alpha \sim Gamma(1; 0, 01)$, para o parâmetro $\beta \sim Gamma(0, 01; 0, 01)$, para o parâmetro $\gamma \sim Gamma(1; 0, 01)$ e para o parâmetro $\lambda \sim Gamma(1; 0, 01)$.

Adotou-se a estratégia de utilização da *priori* uniforme para todos os parâmetros no primeiro conjunto para demonstrar que a utilização de *prioris* que não contenham todo o espaço paramétrico pode gerar estimações muito ruins, como pode ser visto nas Tabelas 3.1 a 3.8.

Toda inferência Bayesiana dos parâmetros foi realizada via MCMC - Markov Chain Monte Carlo - através do pacote MCMCPack do R (R Core Team, 2013), usando o método de *Simulated Annealing* (SANN), considerando 100.000 iterações, 50.000 de burning

e espaçamento igual 10, sendo assim, para as estimativas de cada parâmetro foram consideradas 5.000 iterações.

O método SANN é, por padrão, uma variante do recozimento simulado dada em [17]. *Simulated Annealing* pertence à classe de métodos estocásticos de otimização global. Ele usa apenas valores da função, mas é relativamente lento. Ele também trabalha para funções não-diferenciáveis. Esta implementação utiliza a função Metrópolis para a probabilidade de aceitação. Não é um método de uso geral, mas pode ser muito útil na obtenção de um bom valor para dados com mais de uma moda.

Esta técnica começa sua busca a partir de uma solução inicial qualquer. O procedimento principal consiste em um loop ou laço que a cada iteração, gera aleatoriamente um único vizinho s' da solução corrente s , e é testada a variação Δ do valor da função objetivo, isto é, $\Delta = f(s') - f(s)$, onde temos as seguintes situações:

- $\Delta < 0$: O método aceita a solução e s' passa a ser a nova solução.
- $\Delta = 0$: A aceitação da solução é indiferente.
- $\Delta > 0$: Para calcular a probabilidade de se aceitar a nova solução, uma função conhecida por fator de Boltzmann, que é dada por $e^{-\Delta/T}$, onde T é um parâmetro do método, chamado de temperatura e que regula a probabilidade de soluções ruins, que poderá ser:
 - Gera-se um número aleatório u retirado de uma distribuição uniforme no intervalo $[0, 1]$.
 - Se u for menor ou igual a $e^{-\Delta/T}$, aceita-se a solução.
 - Se u for maior que $e^{-\Delta/T}$, rejeita-se a solução.

O procedimento é finalizado quando a temperatura chega a um valor próximo de zero e nenhuma solução que piore o valor da melhor solução seja mais aceita, ou seja, quando o sistema estiver estável. A solução obtida quando o sistema encontra-se nesta situação evidencia o encontro de um mínimo local.

Para todas as simulações foram realizados os diagnósticos de convergência de Geweke e Heidelberg & Walch, e em todos os casos os dois testes apresentaram convergência. Foi analisado também a taxa de aceitação das simulações e considerou-se como aceitável quando a taxa de aceitação encontrava-se entre 20% e 40%, como é recomendado em diversas bibliografias estudadas.

A seguir, são apresentados nas Tabelas 3.1 a 3.8 as estatísticas para a comparação de qual método de estimação, Clássico ou Bayesiano, apresentou os melhores resultados para os modelos WMG de fração de cura de mistura, de fração de cura de não-mistura e sem fração de cura, nos tamanhos amostrais $n = 20, 50, 100, 500$, sem a presença de censura e na presença de censura de 35%. Para o método de estimação Bayesiano foi considerado as prioris Uniforme (U), Gama (G) e Híbrida (H).

As estatísticas apresentadas nas tabelas 3.1 a 3.8 são a Média, Desvio-Padrão (DP), Coeficiente de Variação (CV), Erro Quadrático Médio (EQM), Probabilidade de Cobertura (PC) e Intervalo de confiança para o método clássico ou Intervalo de Credibilidade para os métodos Bayesianos, ambos os intervalos considerando 95% de confiabilidade/-credibilidade (IC).

Como critério para a obtenção da melhor estimação observou-se os menores valores de EQM e CV, maior PC e se o verdadeiro valor de θ_i está dentro do IC.

Para os dados simulados sem a presença de censura para todos os tamanhos amostrais observa-se que o método Bayesiano com priori Gama apresentaram melhores resultados

para a estimação dos parâmetros nos modelos WMG em estudo (de fração de cura de mistura, de fração de cura de não-mistura e sem fração de cura). O mesmo ocorreu para as simulações com presença de censura de 35%, nos tamanhos menores $n = 20$, 50 e 100 , exceto para o modelo WMG sem fração de cura no tamanho amostral $n = 100$.

Pela tabela 3.9, pode-se observar que apenas o modelo WMG de fração de cura de não-mistura sem a presença de censura para o tamanho amostral $n = 500$, estimado pelo método clássico apresentou menor coeficiente de variação médio, mas no que diz respeito à probabilidade de cobertura média e na proporção de Intervalos de Confiança/Credibilidade que contém o verdadeiro valor do parâmetro θ_i , o método Bayesiano com priori Gama apresentou os melhores resultados. Sendo assim o método Bayesiano com priori Gama também foi considerado como o melhor método de estimação para o modelo WMG de fração de cura de não-mistura sem presença de censura para o tamanho amostral $n = 500$.

Ainda pela tabela 3.9 observa-se que o tamanho amostral $n = 100$ para o modelo WMG sem fração de cura com a presença de censura de 35% os métodos Bayesianos apresentaram melhores resultados sendo que no que diz respeito ao CV médio a priori Uniforme apresentou melhores resultados, para a PC média apresentou melhor resultado com a priori Gama e na proporção de IC em que os verdadeiros valores de θ_i estão contidos a priori Híbrida apresentou melhor resultado. Como a priori Híbrida foi a única que os verdadeiros valores de todos θ_i estavam contidos no IC, então essa foi a priori selecionada.

Para o tamanho $n = 500$ com a presença de censura de 35%, o método clássico apresentou menor CV médio para os modelos WMG de fração de cura de mistura e de fração de cura de não-mistura, já para o modelo sem a fração de cura o método Bayesiano com priori Gama apresentou menor CV médio. No que diz respeito a PC média todos os métodos apresentaram valores abaixo de 30% de cobertura. Quanto à proporção de IC que contém o verdadeiro valor de θ_i , a maior proporção foi de 60%, sendo que para o modelo WMG sem fração de cura cada método conseguiu estimar apenas um parâmetro, considerando o IC. Para esse tamanho amostral definiu-se então a escolha do método Bayesiano com priori Híbrida para o modelo WMG de fração de cura de mistura e a priori Gama para os modelos WMG de fração de cura de não-mistura e sem fração de cura.

Pelos resultados obtidos das simulações pode-se observar que, conforme aumenta-se o tamanho amostral para os dados com presença de censura, os parâmetros mais afetados são α e π , que correspondem ao fator de aceleração do tempo e à fração de cura, respectivamente. Os resultados obtidos mostram também que o modelo WMG apresenta um problema de identificabilidade muito grande.

Tabela 3.1: Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=20$, sem censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$.

Modelos WMG	θ_i	Métodos / Prioris	Média	DP	CV	EQM	PC	IC 95%	
								Inferior	Superior
Com mistura	α	Máxima Verossimilhança	0,0740	0,0864	1,1679	0,0081	0,9950	-0,0954	0,2434
		Uniforme	0,0943	0,0318	0,3368	0,0010	0,9990	0,0429	0,1573
		Gama	0,1385	0,0460	0,3319	0,0036	1,0000	0,0595	0,2245
		Híbrida	0,1363	0,0498	0,3651	0,0038	1,0000	0,0548	0,2269
	β	Máxima Verossimilhança	5,7882	2,7313	0,4719	17,7684	0,8460	5,6188	5,9576
		Uniforme	59,8222	33,9377	0,5673	3733,5129	0,6160	10,6546	127,7531
		Gama	12,3362	3,1024	0,2515	20,7456	1,0000	6,9850	18,6161
		Híbrida	19,8389	14,0353	0,7075	314,2724	0,9890	4,7583	47,0323
	γ	Máxima Verossimilhança	1,0143	0,5050	0,4979	0,3007	0,9610	0,8449	1,1837
		Uniforme	0,4274	0,1324	0,3098	0,1564	0,8220	0,1828	0,6913
		Gama	0,6523	0,1309	0,2007	0,0389	1,0000	0,4243	0,8923
		Híbrida	0,6407	0,1423	0,2222	0,0456	0,9980	0,4143	0,9124
	λ	Máxima Verossimilhança	0,2330	0,1622	0,6960	0,0541	0,7920	0,0636	0,4024
		Uniforme	1,4020	0,4848	0,3458	1,2387	0,6300	0,3802	2,2896
		Gama	0,5392	0,1452	0,2693	0,0405	0,9940	0,2672	0,8254
		Híbrida	0,6596	0,2693	0,4083	0,1399	0,9860	0,2185	1,1906
	π	Máxima Verossimilhança	0,0789	0,0773	0,9799	0,0110	0,9880	-0,0905	0,2483
		Uniforme	0,1371	0,0575	0,4192	0,0035	1,0000	0,0446	0,2459
		Gama	0,1398	0,0575	0,4113	0,0034	1,0000	0,0441	0,2433
		Híbrida	0,1402	0,0580	0,4137	0,0035	1,0000	0,0432	0,2454
Sem mistura	α	Máxima Verossimilhança	0,0799	0,0795	0,9954	0,0067	0,9950	-0,0760	0,2358
		Uniforme	0,0959	0,0322	0,3353	0,0010	0,9990	0,0473	0,1603
		Gama	0,1442	0,0441	0,3056	0,0039	0,9990	0,0670	0,2256
		Híbrida	0,1425	0,0461	0,3237	0,0039	0,9990	0,0635	0,2262
	β	Máxima Verossimilhança	5,7347	2,5744	0,4489	17,2835	0,7900	5,5788	5,8906
		Uniforme	43,2901	25,8670	0,5975	1844,2492	0,7300	5,8517	95,5596
		Gama	10,9405	3,1531	0,2882	13,6976	0,9900	4,9284	16,8922
		Híbrida	15,3473	9,9060	0,6455	138,3189	0,9890	3,3833	36,5801
	γ	Máxima Verossimilhança	0,9033	0,4787	0,5299	0,2396	0,9810	0,7474	1,0592
		Uniforme	0,4089	0,1336	0,3267	0,1708	0,8260	0,1798	0,6788
		Gama	0,6070	0,1431	0,2357	0,0577	0,9950	0,3704	0,8938
		Híbrida	0,6112	0,1756	0,2873	0,0664	0,9970	0,3246	0,9044
	λ	Máxima Verossimilhança	0,2005	0,1398	0,6973	0,0593	0,6860	0,0446	0,3564
		Uniforme	1,1597	0,4543	0,3917	0,7833	0,7710	0,2117	2,0099
		Gama	0,4375	0,1347	0,3080	0,0195	0,9790	0,1694	0,6853
		Híbrida	0,5239	0,2264	0,4322	0,0666	0,9810	0,1416	0,9714
	π	Máxima Verossimilhança	0,0787	0,0731	0,9291	0,0104	0,7650	-0,0772	0,2346
		Uniforme	0,1482	0,0565	0,3811	0,0032	0,9960	0,0505	0,2534
		Gama	0,1473	0,0570	0,3871	0,0033	0,9970	0,0502	0,2498
		Híbrida	0,1504	0,0568	0,3777	0,0032	0,9950	0,0537	0,2578
Sem fração de cura	α	Máxima Verossimilhança	0,0633	0,0851	1,3450	0,0086	0,9900	-0,1036	0,2302
		Uniforme	0,0789	0,0344	0,4362	0,0016	0,9480	0,0244	0,1477
		Gama	0,1050	0,0459	0,4368	0,0021	0,9940	0,0286	0,1951
		Híbrida	0,1065	0,0489	0,4592	0,0024	0,9910	0,0256	0,2009
	β	Máxima Verossimilhança	4,6354	1,5028	0,3242	21,3060	0,6640	4,4685	4,8023
		Uniforme	36,0478	15,8085	0,4385	981,2440	0,9290	12,5715	67,7324
		Gama	11,2818	2,1025	0,1864	9,6226	0,9990	7,2188	15,1409
		Híbrida	14,3696	5,7240	0,3983	61,5641	0,9990	4,3730	24,8655
	γ	Máxima Verossimilhança	1,1821	0,3647	0,3085	0,2789	0,9470	1,0152	1,3490
		Uniforme	0,4516	0,1195	0,2647	0,1357	0,9100	0,2574	0,7230
		Gama	0,6294	0,1027	0,1632	0,0396	0,9970	0,4463	0,8554
		Híbrida	0,6301	0,1073	0,1702	0,0403	0,9950	0,4229	0,8304
	λ	Máxima Verossimilhança	0,1956	0,1204	0,6157	0,0563	0,6560	0,0287	0,3625
		Uniforme	1,2364	0,3878	0,3136	0,8498	0,8680	0,4793	2,0110
		Gama	0,5849	0,1509	0,2580	0,0569	0,9950	0,3050	0,8928
		Híbrida	0,6515	0,2144	0,3291	0,1091	0,9920	0,2152	1,0360

Tabela 3.2: Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=20$, com 35% de censura. Verdadeiros valores: $\alpha = 0, 10$, $\beta = 9, 00$, $\gamma = 0, 80$, $\lambda = 0, 40$ e $\pi = 0, 15$.

Modelos WMG	θ_i	Métodos / Prioris	Média	DP	CV	EQM	PC	IC 95%	
								Inferior	Superior
Com mistura	α	Máxima Verossimilhança	0,3910	0,2633	0,6734	0,1540	0,9920	-0,1250	0,9070
		Uniforme	0,2049	0,0507	0,2473	0,0136	1,0000	0,1108	0,2977
		Gama	0,3051	0,0643	0,2109	0,0462	1,0000	0,1994	0,4459
		Híbrida	0,2791	0,0621	0,2223	0,0359	1,0000	0,1667	0,4006
	β	Máxima Verossimilhança	6,4285	2,4974	0,3885	12,8436	0,9420	5,9125	6,9445
		Uniforme	74,6061	47,4357	0,6358	6552,0517	0,3730	11,7574	173,5936
		Gama	15,7361	5,7500	0,3654	78,4042	0,9760	7,0891	28,8643
		Híbrida	29,4907	23,4794	0,7962	970,5968	0,8750	4,9503	84,7787
	γ	Máxima Verossimilhança	1,0907	0,7315	0,6707	0,6190	0,9940	0,5747	1,6067
		Uniforme	0,7074	0,1927	0,2724	0,0457	0,9970	0,3679	1,0918
		Gama	0,8885	0,1865	0,2100	0,0426	1,0000	0,5107	1,2319
		Híbrida	0,8580	0,1741	0,2029	0,0336	1,0000	0,5328	1,1884
	λ	Máxima Verossimilhança	0,0633	0,0644	1,0172	0,1175	0,1900	-0,4527	0,5793
		Uniforme	0,6058	0,3180	0,5249	0,1434	0,7930	0,1332	1,2427
		Gama	0,2005	0,0917	0,4575	0,0482	0,5620	0,0552	0,3744
		Híbrida	0,2835	0,1480	0,5220	0,0355	0,7600	0,0705	0,5713
	π	Máxima Verossimilhança	0,1797	0,1571	0,8745	0,0256	0,9200	-0,3363	0,6957
		Uniforme	0,2505	0,1042	0,4161	0,0210	0,9310	0,0985	0,4817
		Gama	0,2448	0,0993	0,4055	0,0188	0,9530	0,0851	0,4475
		Híbrida	0,2470	0,1004	0,4066	0,0195	0,9470	0,0802	0,4439
Sem mistura	α	Máxima Verossimilhança	0,3746	0,2565	0,6847	0,1411	0,9950	-0,1281	0,8773
		Uniforme	0,2097	0,0653	0,3115	0,0163	0,9780	0,0960	0,3328
		Gama	0,2874	0,0586	0,2040	0,0385	1,0000	0,1855	0,4077
		Híbrida	0,2641	0,0572	0,2166	0,0302	1,0000	0,1584	0,3751
	β	Máxima Verossimilhança	6,2498	2,5448	0,4072	14,0329	0,9120	5,7471	6,7525
		Uniforme	60,2856	48,2002	0,7995	4951,1488	0,4090	8,4341	162,9037
		Gama	15,8453	6,1084	0,3855	84,1342	0,9450	6,6916	28,1179
		Híbrida	27,4997	20,4944	0,7453	761,8396	0,8420	5,2573	67,7028
	γ	Máxima Verossimilhança	1,0333	0,6903	0,6680	0,5304	0,9910	0,5306	1,5360
		Uniforme	0,6618	0,1798	0,2717	0,0514	0,9830	0,3135	0,9891
		Gama	0,8288	0,2032	0,2452	0,0421	1,0000	0,4177	1,1701
		Híbrida	0,8080	0,1919	0,2375	0,0369	0,9990	0,4641	1,1629
	λ	Máxima Verossimilhança	0,0580	0,0621	1,0711	0,1208	0,1580	-0,4447	0,5607
		Uniforme	0,5331	0,3109	0,5832	0,1143	0,7160	0,0952	1,2227
		Gama	0,2024	0,0943	0,4658	0,0479	0,4800	0,0572	0,3739
		Híbrida	0,2851	0,1585	0,5559	0,0383	0,7040	0,0368	0,5978
	π	Máxima Verossimilhança	0,1814	0,1548	0,8534	0,0249	0,8020	-0,3213	0,6841
		Uniforme	0,2652	0,1035	0,3904	0,0240	0,9250	0,0707	0,4676
		Gama	0,2533	0,0996	0,3934	0,0206	0,9510	0,0851	0,4498
		Híbrida	0,2582	0,1016	0,3937	0,0220	0,9340	0,0901	0,4576
Sem fração de cura	α	Máxima Verossimilhança	0,2474	0,1990	0,8043	0,0613	0,9970	-0,1426	0,6374
		Uniforme	0,1813	0,0522	0,2882	0,0093	1,0000	0,0841	0,2829
		Gama	0,2231	0,0616	0,2760	0,0190	0,9980	0,1091	0,3376
		Híbrida	0,1997	0,0578	0,2894	0,0133	0,9990	0,0979	0,3128
	β	Máxima Verossimilhança	5,9556	1,7040	0,2861	12,1690	0,9280	5,5656	6,3456
		Uniforme	29,6438	14,3187	0,4830	630,9872	0,8430	10,5332	55,0457
		Gama	14,2353	3,8837	0,2728	42,4766	0,9740	7,5608	22,2619
		Híbrida	27,4538	20,4204	0,7438	757,1182	0,8790	5,6779	65,0768
	γ	Máxima Verossimilhança	1,3005	0,5884	0,4524	0,5963	0,9830	0,9105	1,6905
		Uniforme	0,6461	0,1705	0,2638	0,0527	0,9810	0,3476	1,0115
		Gama	0,7244	0,1587	0,2191	0,0309	0,9960	0,3970	1,0349
		Híbrida	0,6707	0,1713	0,2555	0,0460	0,9900	0,3470	0,9956
	λ	Máxima Verossimilhança	0,0738	0,0526	0,7133	0,1091	0,2010	-0,3162	0,4638
		Uniforme	0,5324	0,2457	0,4616	0,0779	0,8910	0,1617	1,0005
		Gama	0,3301	0,1375	0,4166	0,0238	0,8740	0,1084	0,6018
		Híbrida	0,4816	0,2826	0,5868	0,0864	0,8910	0,1230	1,0464

Tabela 3.3: Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=50$, sem censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$.

Modelos WMG	θ_i	Métodos / Prioris	Média	DP	CV	EQM	PC	IC 95%	
								Inferior	Superior
Com mistura	α	Máxima Verossimilhança	0,0520	0,0548	1,0542	0,0053	0,9940	-0,0554	0,1594
		Uniforme	0,0805	0,0269	0,3341	0,0011	0,9960	0,0403	0,1357
		Gama	0,1019	0,0351	0,3443	0,0012	0,9970	0,0463	0,1718
		Híbrida	0,0971	0,0361	0,3720	0,0013	0,9970	0,0406	0,1690
	β	Máxima Verossimilhança	5,5855	2,0876	0,3738	16,0124	0,6870	5,4781	5,6929
		Uniforme	49,5744	23,5524	0,4751	2200,4456	0,5760	11,5521	94,7248
		Gama	13,7302	3,7993	0,2767	36,7953	1,0000	6,7480	21,0455
		Híbrida	22,9183	13,6884	0,5973	380,9042	0,9790	5,2233	51,2223
	γ	Máxima Verossimilhança	1,1742	0,4215	0,3590	0,3175	0,8620	1,0668	1,2816
		Uniforme	0,4706	0,1350	0,2869	0,1267	0,8280	0,2300	0,7515
		Gama	0,7225	0,1288	0,1783	0,0226	1,0000	0,4700	0,9330
		Híbrida	0,6733	0,1541	0,2289	0,0398	0,9930	0,4234	0,9836
	λ	Máxima Verossimilhança	0,2143	0,1203	0,5612	0,0489	0,6700	0,1069	0,3217
		Uniforme	1,3221	0,3992	0,3020	1,0095	0,5680	0,6096	2,1539
		Gama	0,5767	0,1207	0,2093	0,0458	0,9980	0,3463	0,8161
		Híbrida	0,7673	0,2813	0,3666	0,2140	0,9820	0,2682	1,3309
	π	Máxima Verossimilhança	0,1087	0,0632	0,5815	0,0057	0,9460	0,0013	0,2161
		Uniforme	0,1245	0,0469	0,3765	0,0028	0,9510	0,0413	0,2232
		Gama	0,1298	0,0472	0,3637	0,0026	0,9520	0,0429	0,2274
		Híbrida	0,1290	0,0471	0,3653	0,0027	0,9520	0,0481	0,2305
Sem mistura	α	Máxima Verossimilhança	0,0689	0,0647	0,9391	0,0051	0,9950	-0,0579	0,1957
		Uniforme	0,0850	0,0258	0,3032	0,0009	0,9980	0,0403	0,1373
		Gama	0,1105	0,0343	0,3108	0,0013	0,9990	0,0564	0,1769
		Híbrida	0,1089	0,0376	0,3450	0,0015	0,9990	0,0456	0,1816
	β	Máxima Verossimilhança	5,4490	1,7968	0,3298	15,8348	0,6360	5,3222	5,5758
		Uniforme	36,1510	18,7663	0,5191	1088,9991	0,7510	8,5934	75,9703
		Gama	13,0255	4,2059	0,3229	33,8766	0,9980	5,7126	21,5560
		Híbrida	17,5696	10,8925	0,6200	191,9646	0,9760	3,5828	35,7490
	γ	Máxima Verossimilhança	1,0403	0,3855	0,3705	0,2062	0,9120	0,9135	1,1671
		Uniforme	0,4238	0,1271	0,2998	0,1577	0,8070	0,1916	0,6756
		Gama	0,6235	0,1310	0,2101	0,0483	0,9990	0,3821	0,8365
		Híbrida	0,6150	0,1625	0,2642	0,0606	0,9870	0,3182	0,8797
	λ	Máxima Verossimilhança	0,1799	0,0985	0,5474	0,0581	0,5770	0,0531	0,3067
		Uniforme	1,1045	0,3704	0,3354	0,6334	0,7780	0,3943	1,8423
		Gama	0,5209	0,1352	0,2595	0,0329	0,9920	0,2319	0,7648
		Híbrida	0,6219	0,2581	0,4151	0,1158	0,9780	0,1448	1,0821
	π	Máxima Verossimilhança	0,1033	0,0599	0,5803	0,0058	0,8720	-0,0235	0,2301
		Uniforme	0,1256	0,0449	0,3574	0,0026	0,9520	0,0427	0,2147
		Gama	0,1296	0,0452	0,3488	0,0025	0,9600	0,0481	0,2242
		Híbrida	0,1295	0,0450	0,3474	0,0024	0,9620	0,0417	0,2168
Sem fração de cura	α	Máxima Verossimilhança	0,0167	0,0279	1,6714	0,0077	0,9540	-0,0380	0,0714
		Uniforme	0,0424	0,0247	0,5837	0,0039	0,5420	0,0079	0,0925
		Gama	0,0479	0,0285	0,5956	0,0035	0,6820	0,0089	0,1048
		Híbrida	0,0472	0,0286	0,6051	0,0036	0,6540	0,0089	0,1049
	β	Máxima Verossimilhança	4,5842	1,2053	0,2629	20,9509	0,4860	4,5295	4,6389
		Uniforme	33,6471	12,3174	0,3661	759,0436	0,8230	15,6998	61,0117
		Gama	14,3393	2,9276	0,2042	37,0700	0,9990	9,5149	20,2897
		Híbrida	19,0785	7,0733	0,3707	151,5576	0,9870	8,3853	32,6271
	γ	Máxima Verossimilhança	1,2548	0,2602	0,2074	0,2745	0,8600	1,2001	1,3095
		Uniforme	0,5430	0,1110	0,2045	0,0784	0,8860	0,3270	0,7595
		Gama	0,7154	0,0932	0,1302	0,0158	1,0000	0,5464	0,9045
		Híbrida	0,6857	0,1025	0,1496	0,0236	0,9860	0,4868	0,8883
	λ	Máxima Verossimilhança	0,2130	0,1062	0,4988	0,0462	0,6200	0,1583	0,2677
		Uniforme	1,2163	0,2962	0,2436	0,7540	0,6700	0,7097	1,8190
		Gama	0,7018	0,1403	0,2000	0,1108	0,9960	0,4479	0,9979
		Híbrida	0,8256	0,2207	0,2674	0,2298	0,9710	0,4147	1,2566

Tabela 3.4: Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=50$, com 35% de censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$.

Modelos WMG	θ_i	Métodos / Prioris	Média	DP	CV	EQM	PC	IC 95%	
								Inferior	Superior
Com mistura	α	Máxima Verossimilhança	0,0833	0,0967	1,1606	0,0096	0,9950	-0,1062	0,2728
		Uniforme	0,0885	0,0327	0,3694	0,0012	0,9990	0,0412	0,1553
		Gama	0,1215	0,0447	0,3679	0,0025	0,9990	0,0525	0,2081
		Híbrida	0,1186	0,0489	0,4123	0,0027	0,9990	0,0481	0,2184
	β	Máxima Verossimilhança	5,9394	2,4180	0,4071	15,2080	0,7670	5,7499	6,1289
		Uniforme	46,6452	24,6457	0,5284	2023,9642	0,7240	8,2277	93,6976
		Gama	12,4484	3,4384	0,2762	23,7028	0,9990	6,2296	19,0440
		Híbrida	19,8513	12,8456	0,6471	282,5949	0,9850	3,8487	45,4799
	γ	Máxima Verossimilhança	1,0182	0,4483	0,4403	0,2484	0,9460	0,8287	1,2077
		Uniforme	0,4380	0,1309	0,2989	0,1482	0,8580	0,1940	0,6911
		Gama	0,6557	0,1284	0,1958	0,0373	1,0000	0,4108	0,9017
		Híbrida	0,6318	0,1423	0,2253	0,0485	0,9960	0,3989	0,8966
	λ	Máxima Verossimilhança	0,2315	0,1403	0,6060	0,0480	0,7370	0,0420	0,4210
		Uniforme	1,2704	0,4098	0,3225	0,9252	0,7280	0,5553	2,2005
		Gama	0,5357	0,1372	0,2561	0,0372	0,9990	0,2720	0,7818
		Híbrida	0,6820	0,2775	0,4069	0,1564	0,9790	0,1297	1,1756
	π	Máxima Verossimilhança	0,2266	0,1187	0,5236	0,0199	0,8020	0,0371	0,4161
		Uniforme	0,2469	0,0825	0,3341	0,0162	0,8700	0,1005	0,4095
		Gama	0,2535	0,0815	0,3214	0,0173	0,8560	0,0970	0,4018
		Híbrida	0,2534	0,0822	0,3245	0,0174	0,8570	0,1048	0,4119
Sem mistura	α	Máxima Verossimilhança	0,0885	0,0966	1,0915	0,0095	0,9930	-0,1008	0,2778
		Uniforme	0,0901	0,0351	0,3893	0,0013	0,9980	0,0393	0,1578
		Gama	0,1255	0,0427	0,3399	0,0025	0,9990	0,0532	0,2093
		Híbrida	0,1252	0,0477	0,3813	0,0029	1,0000	0,0526	0,2203
	β	Máxima Verossimilhança	5,7471	2,0605	0,3585	14,8229	0,7350	5,5578	5,9364
		Uniforme	40,2980	22,5399	0,5593	1487,1053	0,7420	6,6732	88,5721
		Gama	12,0669	3,5330	0,2928	21,8754	0,9960	5,8613	19,1202
		Híbrida	16,6798	11,2601	0,6751	185,6434	0,9790	4,1871	37,9042
	γ	Máxima Verossimilhança	0,9663	0,4443	0,4598	0,2249	0,9490	0,7770	1,1556
		Uniforme	0,4141	0,1488	0,3593	0,1710	0,7900	0,1912	0,7193
		Gama	0,6113	0,1397	0,2286	0,0551	1,0000	0,3726	0,8604
		Híbrida	0,6083	0,1733	0,2849	0,0668	0,9910	0,3297	0,8784
	λ	Máxima Verossimilhança	0,1977	0,1192	0,6028	0,0551	0,6650	0,0084	0,3870
		Uniforme	1,1520	0,4263	0,3701	0,7472	0,7650	0,3423	2,0498
		Gama	0,4886	0,1330	0,2722	0,0255	0,9930	0,2210	0,7247
		Híbrida	0,5799	0,2598	0,4481	0,0998	0,9810	0,1479	1,0745
	π	Máxima Verossimilhança	0,2252	0,1161	0,5154	0,0191	0,7930	0,0359	0,4145
		Uniforme	0,2530	0,0803	0,3175	0,0171	0,8550	0,1033	0,4086
		Gama	0,2577	0,0795	0,3083	0,0179	0,8420	0,1117	0,4141
		Híbrida	0,2585	0,0796	0,3079	0,0181	0,8460	0,1107	0,4131
Sem fração de cura	α	Máxima Verossimilhança	0,0188	0,0317	1,6851	0,0076	0,9700	-0,0433	0,0809
		Uniforme	0,0430	0,0223	0,5183	0,0037	0,6030	0,0107	0,0873
		Gama	0,0523	0,0275	0,5257	0,0030	0,7840	0,0131	0,1065
		Híbrida	0,0531	0,0295	0,5548	0,0031	0,7910	0,0111	0,1100
	β	Máxima Verossimilhança	4,6828	1,1739	0,2507	20,0146	0,6010	4,6207	4,7449
		Uniforme	30,3962	12,2637	0,4035	608,0483	0,8850	12,5187	56,7697
		Gama	13,6756	3,4019	0,2488	33,4227	1,0000	8,5447	20,5416
		Híbrida	15,8184	6,5065	0,4113	88,7827	0,9910	6,7199	28,5069
	γ	Máxima Verossimilhança	1,1126	0,2479	0,2228	0,1591	0,9460	1,0505	1,1747
		Uniforme	0,4683	0,1051	0,2244	0,1210	0,8360	0,2590	0,6616
		Gama	0,6075	0,0839	0,1381	0,0441	0,9960	0,4476	0,7805
		Híbrida	0,6208	0,1093	0,1761	0,0441	0,9890	0,4041	0,8461
	λ	Máxima Verossimilhança	0,2204	0,1075	0,4877	0,0438	0,6810	0,1583	0,2825
		Uniforme	1,2043	0,3105	0,2578	0,7432	0,8000	0,6141	1,7885
		Gama	0,7133	0,1442	0,2022	0,1189	0,9980	0,4447	0,9913
		Híbrida	0,7544	0,2240	0,2969	0,1757	0,9880	0,3343	1,2202

Tabela 3.5: Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=100$, sem censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$.

Modelos WMG	θ_i	Métodos / Prioris	Média	DP	CV	EQM	PC	IC 95%	
								Inferior	Superior
Com mistura	α	Máxima Verossimilhança	0,0584	0,0546	0,9345	0,0047	0,9750	-0,0486	0,1654
		Uniforme	0,0820	0,0299	0,3652	0,0012	0,9820	0,0335	0,1431
		Gama	0,0905	0,0343	0,3795	0,0013	0,9940	0,0356	0,1598
		Híbrida	0,0885	0,0339	0,3826	0,0013	0,9870	0,0346	0,1580
	β	Máxima Verossimilhança	5,3396	1,5482	0,2899	15,7932	0,5600	5,2326	5,4466
		Uniforme	41,1635	19,4194	0,4718	1411,2236	0,6430	11,2899	79,7116
		Gama	15,2558	4,6143	0,3025	60,4048	0,9990	7,4072	24,7314
		Híbrida	23,2033	13,8446	0,5967	393,2156	0,9510	6,5394	49,9129
	γ	Máxima Verossimilhança	1,2125	0,3763	0,3104	0,3117	0,8010	1,1055	1,3195
		Uniforme	0,5010	0,1522	0,3038	0,1125	0,8540	0,2056	0,7708
		Gama	0,7407	0,1382	0,1866	0,0226	0,9990	0,4799	0,9822
		Híbrida	0,6910	0,1649	0,2386	0,0390	0,9840	0,3675	0,9895
	λ	Máxima Verossimilhança	0,1969	0,0926	0,4705	0,0498	0,5560	0,0899	0,3039
		Uniforme	1,1968	0,3608	0,3015	0,7650	0,6360	0,5998	2,0139
		Gama	0,6131	0,1320	0,2154	0,0628	0,9990	0,3513	0,8689
		Híbrida	0,7799	0,2895	0,3712	0,2280	0,9610	0,2416	1,3423
	π	Máxima Verossimilhança	0,1307	0,0473	0,3618	0,0026	0,9350	0,0237	0,2377
		Uniforme	0,1332	0,0382	0,2867	0,0017	0,9460	0,0577	0,2051
		Gama	0,1374	0,0377	0,2746	0,0016	0,9530	0,0615	0,2070
		Híbrida	0,1364	0,0377	0,2763	0,0016	0,9530	0,0621	0,2072
Sem mistura	α	Máxima Verossimilhança	0,0800	0,0646	0,8074	0,0046	0,9880	-0,0466	0,2066
		Uniforme	0,0881	0,0286	0,3248	0,0010	0,9950	0,0393	0,1475
		Gama	0,1033	0,0343	0,3319	0,0012	0,9980	0,0463	0,1746
		Híbrida	0,1030	0,0377	0,3655	0,0014	0,9960	0,0420	0,1769
	β	Máxima Verossimilhança	5,3229	1,4271	0,2681	15,5560	0,5120	5,1963	5,4495
		Uniforme	30,9504	14,7846	0,4777	700,1874	0,7800	6,4125	60,5317
		Gama	13,3491	4,2471	0,3182	36,9351	0,9980	6,4359	21,5071
		Híbrida	16,2420	10,9115	0,6718	171,3893	0,9680	4,0272	36,2227
	γ	Máxima Verossimilhança	1,0510	0,3645	0,3468	0,1957	0,8630	0,9244	1,1776
		Uniforme	0,4389	0,1344	0,3061	0,1484	0,8240	0,1884	0,6894
		Gama	0,6390	0,1350	0,2113	0,0441	1,0000	0,3970	0,9220
		Híbrida	0,6604	0,1886	0,2855	0,0550	0,9780	0,3429	1,0129
	λ	Máxima Verossimilhança	0,1739	0,0830	0,4773	0,0580	0,4400	0,0473	0,3005
		Uniforme	1,0173	0,3235	0,3180	0,4856	0,8050	0,4327	1,7143
		Gama	0,5318	0,1389	0,2611	0,0366	0,9950	0,2925	0,8227
		Híbrida	0,5804	0,2681	0,4620	0,1044	0,9690	0,1549	1,1125
	π	Máxima Verossimilhança	0,1235	0,0472	0,3825	0,0029	0,9240	-0,0031	0,2501
		Uniforme	0,1291	0,0371	0,2875	0,0018	0,9400	0,0561	0,1989
		Gama	0,1331	0,0369	0,2774	0,0016	0,9530	0,0582	0,2010
		Híbrida	0,1329	0,0367	0,2761	0,0016	0,9530	0,0635	0,2044
Sem fração de cura	α	Máxima Verossimilhança	0,0065	0,0104	1,5988	0,0089	0,7920	-0,0139	0,0269
		Uniforme	0,0213	0,0160	0,7518	0,0065	0,1180	0,0032	0,0537
		Gama	0,0219	0,0160	0,7293	0,0064	0,1520	0,0042	0,0541
		Híbrida	0,0217	0,0158	0,7303	0,0064	0,1320	0,0042	0,0547
	β	Máxima Verossimilhança	4,9245	1,0651	0,2163	17,7428	0,4420	4,9041	4,9449
		Uniforme	32,3799	10,7651	0,3325	662,3914	0,5780	15,0845	55,3554
		Gama	18,9391	5,4394	0,2872	128,3424	0,9420	9,8623	29,2492
		Híbrida	27,5091	12,0963	0,4397	488,7623	0,7970	11,0869	53,8463
	γ	Máxima Verossimilhança	1,1969	0,1987	0,1660	0,1970	0,8180	1,1765	1,2173
		Uniforme	0,6037	0,0945	0,1565	0,0474	0,7830	0,4162	0,7893
		Gama	0,7182	0,0817	0,1138	0,0134	0,9770	0,5528	0,8767
		Híbrida	0,6572	0,1015	0,1545	0,0307	0,8800	0,4628	0,8502
	λ	Máxima Verossimilhança	0,2449	0,0972	0,3968	0,0335	0,6620	0,2245	0,2653
		Uniforme	1,1917	0,2583	0,2167	0,6934	0,4130	0,7256	1,7523
		Gama	0,8485	0,1779	0,2097	0,2328	0,8220	0,5057	1,1921
		Híbrida	1,0534	0,2834	0,2691	0,5073	0,6900	0,5698	1,6843

Tabela 3.6: Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=100$, com 35% de censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$.

Modelos WMG	θ_i	Métodos / Prioris	Média	DP	CV	EQM	PC	IC 95%	
								Inferior	Superior
Com mistura	α	Máxima Verossimilhança	0,4105	0,1878	0,4574	0,1316	0,9910	0,0425	0,7785
		Uniforme	0,1988	0,0673	0,3387	0,0143	0,9920	0,0896	0,3355
		Gama	0,2796	0,0690	0,2469	0,0370	0,9600	0,1559	0,4187
		Híbrida	0,2573	0,0791	0,3074	0,0310	0,9630	0,1120	0,4003
	β	Máxima Verossimilhança	5,7308	1,3637	0,2380	12,5453	0,7610	5,3628	6,0988
		Uniforme	49,0869	31,1471	0,6345	2576,1338	0,3880	6,4697	106,6866
		Gama	16,6663	6,3603	0,3816	99,1838	0,8540	7,0993	29,9538
		Híbrida	24,5733	15,7130	0,6394	489,1801	0,7520	5,4536	54,6817
	γ	Máxima Verossimilhança	1,1287	0,6599	0,5846	0,5430	0,9800	0,7607	1,4967
		Uniforme	0,7812	0,2398	0,3070	0,0578	0,9930	0,3478	1,2257
		Gama	0,8898	0,2662	0,2992	0,0789	0,9970	0,3591	1,3631
		Híbrida	0,8761	0,2573	0,2937	0,0720	0,9980	0,4009	1,3340
	λ	Máxima Verossimilhança	0,0487	0,0315	0,6478	0,1244	0,0310	-0,3193	0,4167
		Uniforme	0,5002	0,2773	0,5544	0,0869	0,6750	0,0906	1,0286
		Gama	0,2211	0,1006	0,4549	0,0421	0,5730	0,0656	0,4191
		Híbrida	0,2835	0,1482	0,5229	0,0355	0,7020	0,0589	0,6014
	π	Máxima Verossimilhança	0,2655	0,0716	0,2698	0,0185	0,5140	-0,1025	0,6335
		Uniforme	0,2442	0,0676	0,2768	0,0134	0,6880	0,1202	0,3710
		Gama	0,2507	0,0660	0,2633	0,0145	0,6710	0,1195	0,3650
		Híbrida	0,2497	0,0669	0,2678	0,0144	0,6650	0,1212	0,3700
Sem mistura	α	Máxima Verossimilhança	0,4236	0,1781	0,4204	0,1364	0,9900	0,0745	0,7727
		Uniforme	0,1997	0,0653	0,3270	0,0142	0,9900	0,0782	0,3230
		Gama	0,2655	0,0612	0,2304	0,0311	0,9870	0,1632	0,3871
		Híbrida	0,2425	0,0743	0,3065	0,0258	0,9920	0,1158	0,3944
	β	Máxima Verossimilhança	5,6292	1,2804	0,2275	13,0000	0,7180	5,2801	5,9783
		Uniforme	36,8094	23,0499	0,6262	1304,1256	0,3580	8,7305	86,0295
		Gama	16,8398	7,1446	0,4243	112,4569	0,8480	6,6493	30,5251
		Híbrida	25,3852	18,7670	0,7393	620,3229	0,7040	5,4510	70,5113
	γ	Máxima Verossimilhança	0,9971	0,6294	0,6312	0,4346	0,9860	0,6480	1,3462
		Uniforme	0,7056	0,1979	0,2805	0,0481	0,9900	0,3613	1,1158
		Gama	0,8262	0,2173	0,2630	0,0478	1,0000	0,3760	1,2285
		Híbrida	0,8111	0,2095	0,2583	0,0440	0,9980	0,4288	1,2114
	λ	Máxima Verossimilhança	0,0459	0,0304	0,6620	0,1263	0,0330	-0,3032	0,3950
		Uniforme	0,4356	0,2125	0,4878	0,0464	0,6850	0,1044	0,8852
		Gama	0,2229	0,0997	0,4473	0,0413	0,4730	0,0691	0,4086
		Híbrida	0,2998	0,1838	0,6130	0,0438	0,6010	0,0591	0,6976
	π	Máxima Verossimilhança	0,2581	0,0736	0,2850	0,0171	0,5830	-0,0910	0,6072
		Uniforme	0,2425	0,0663	0,2732	0,0129	0,7320	0,1139	0,3619
		Gama	0,2460	0,0663	0,2694	0,0136	0,6920	0,1207	0,3743
		Híbrida	0,2452	0,0658	0,2685	0,0134	0,6850	0,1122	0,3609
Sem fração de cura	α	Máxima Verossimilhança	0,0722	0,0820	1,1352	0,0075	0,9950	-0,0884	0,2328
		Uniforme	0,0946	0,0446	0,4718	0,0020	0,9780	0,0196	0,1826
		Gama	0,1028	0,0473	0,4598	0,0022	0,9890	0,0289	0,1967
		Híbrida	0,0937	0,0464	0,4947	0,0022	0,9780	0,0246	0,1851
	β	Máxima Verossimilhança	5,5469	1,1738	0,2116	13,3005	0,6440	5,3863	5,7075
		Uniforme	25,7538	8,2295	0,3195	348,3447	0,6080	12,3550	40,0449
		Gama	19,8501	7,6069	0,3832	175,5315	0,8620	9,8377	35,4347
		Híbrida	33,8853	20,4945	0,6048	1038,8800	0,6050	8,9352	81,2711
	γ	Máxima Verossimilhança	1,7793	0,3662	0,2058	1,0929	0,7510	1,6187	1,9399
		Uniforme	0,8727	0,1979	0,2268	0,0444	0,9800	0,4759	1,2490
		Gama	0,9661	0,1932	0,2000	0,0649	0,9970	0,5375	1,2707
		Híbrida	0,8587	0,2159	0,2514	0,0500	0,9740	0,4510	1,2634
	λ	Máxima Verossimilhança	0,0701	0,0337	0,4802	0,1100	0,0280	-0,0905	0,2307
		Uniforme	0,5115	0,1696	0,3315	0,0412	0,8630	0,2181	0,8143
		Gama	0,4035	0,1479	0,3665	0,0219	0,9180	0,1711	0,7422

Tabela 3.7: Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=500$, sem censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$.

Modelos WMG	θ_i	Métodos / Priors	Média	DP	CV	EQM	PC	IC 95%	
								Inferior	Superior
Com mistura	α	Máxima Verossimilhança	0,0727	0,0421	0,5796	0,0025	0,8340	-0,0099	0,1553
		Uniforme	0,0987	0,0245	0,2482	0,0006	0,9600	0,0508	0,1449
		Gama	0,0886	0,0265	0,2988	0,0008	0,9730	0,0425	0,1394
		Híbrida	0,0881	0,0257	0,2916	0,0008	0,9690	0,0412	0,1364
	β	Máxima Verossimilhança	5,1539	0,8262	0,1603	15,4740	0,1360	5,0713	5,2365
		Uniforme	28,6097	13,7653	0,4811	573,8359	0,7390	6,3522	54,0249
		Gama	14,3946	5,4138	0,3761	58,3816	0,9870	6,6139	25,7599
		Híbrida	17,4891	10,8042	0,6178	188,6788	0,9360	5,0550	41,6809
	γ	Máxima Verossimilhança	1,2177	0,2515	0,2066	0,2377	0,4250	1,1351	1,3003
		Uniforme	0,5262	0,2127	0,4043	0,1201	0,8140	0,1540	0,9272
		Gama	0,7858	0,1806	0,2298	0,0328	0,9920	0,4163	1,1160
		Híbrida	0,7808	0,2528	0,3237	0,0642	0,9600	0,3008	1,2341
	λ	Máxima Verossimilhança	0,1848	0,0506	0,2740	0,0489	0,1810	0,1022	0,2674
		Uniforme	0,9963	0,3539	0,3552	0,4807	0,7500	0,4035	1,7268
		Gama	0,5803	0,1805	0,3111	0,0651	0,9860	0,2624	0,9307
		Híbrida	0,6509	0,3109	0,4777	0,1595	0,9390	0,1972	1,3494
	π	Máxima Verossimilhança	0,1469	0,0223	0,1515	0,0005	0,8750	0,0643	0,2295
		Uniforme	0,1477	0,0167	0,1133	0,0003	0,9390	0,1169	0,1834
		Gama	0,1479	0,0167	0,1129	0,0003	0,9400	0,1172	0,1833
		Híbrida	0,1479	0,0167	0,1129	0,0003	0,9390	0,1174	0,1838
Sem mistura	α	Máxima Verossimilhança	0,1223	0,0474	0,3873	0,0027	0,8780	0,0295	0,2151
		Uniforme	0,1170	0,0271	0,2320	0,0010	0,9400	0,0632	0,1704
		Gama	0,1171	0,0299	0,2553	0,0012	0,9600	0,0533	0,1685
		Híbrida	0,1153	0,0300	0,2603	0,0011	0,9620	0,0517	0,1677
	β	Máxima Verossimilhança	5,1358	0,6808	0,1326	15,3948	0,1060	5,0430	5,2286
		Uniforme	19,7796	9,9044	0,5007	214,2000	0,8610	5,2359	39,9665
		Gama	12,0133	4,2767	0,3560	27,3520	0,9840	4,7963	19,6946
		Híbrida	12,9317	6,1937	0,4790	53,7811	0,9670	4,1491	25,2629
	γ	Máxima Verossimilhança	0,9348	0,2526	0,2702	0,0819	0,7260	0,8420	1,0276
		Uniforme	0,4703	0,2038	0,4333	0,1502	0,8270	0,1383	0,8545
		Gama	0,6449	0,1998	0,3098	0,0639	0,9830	0,3079	1,0158
		Híbrida	0,6728	0,2429	0,3611	0,0751	0,9780	0,2873	1,1621
	λ	Máxima Verossimilhança	0,1715	0,0457	0,2666	0,0543	0,1190	0,0787	0,2643
		Uniforme	0,7642	0,3094	0,4049	0,2283	0,8680	0,1879	1,3465
		Gama	0,4932	0,1699	0,3445	0,0375	0,9820	0,1764	0,7878
		Híbrida	0,5135	0,2274	0,4429	0,0646	0,9620	0,1497	0,9644
	π	Máxima Verossimilhança	0,1465	0,0209	0,1430	0,0005	0,8680	0,0537	0,2393
		Uniforme	0,1458	0,0174	0,1197	0,0003	0,9310	0,1129	0,1805
		Gama	0,1462	0,0173	0,1185	0,0003	0,9300	0,1136	0,1819
		Híbrida	0,1462	0,0173	0,1185	0,0003	0,9380	0,1140	0,1814
Sem fração de cura	α	Máxima Verossimilhança	0,0024	0,0020	0,8397	0,0095	0,0330	-0,0015	0,0063
		Uniforme	0,0022	0,0010	0,4644	0,0096	0,0000	0,0007	0,0041
		Gama	0,0022	0,0010	0,4359	0,0096	0,0000	0,0009	0,0041
		Híbrida	0,0022	0,0010	0,4481	0,0096	0,0000	0,0007	0,0040
	β	Máxima Verossimilhança	4,8742	0,5684	0,1166	17,3454	0,0880	4,8703	4,8781
		Uniforme	27,0408	4,7465	0,1755	347,9780	0,0010	17,8280	36,1423
		Gama	22,9722	3,9001	0,1698	210,4176	0,0200	15,9339	30,8630
		Híbrida	24,8985	4,6209	0,1856	274,0942	0,0080	14,2907	32,8568
	γ	Máxima Verossimilhança	1,0953	0,1033	0,0943	0,0979	0,4610	1,0914	1,0992
		Uniforme	0,6509	0,0482	0,0741	0,0246	0,1420	0,5607	0,7485
		Gama	0,6822	0,0516	0,0757	0,0165	0,3830	0,5757	0,7773
		Híbrida	0,6673	0,0509	0,0763	0,0202	0,2630	0,5623	0,7601
	λ	Máxima Verossimilhança	0,2791	0,0624	0,2238	0,0185	0,4940	0,2752	0,2830
		Uniforme	1,1350	0,1341	0,1181	0,5581	0,0000	0,8651	1,3823
		Gama	1,0266	0,1279	0,1246	0,4089	0,0030	0,7643	1,2729
		Híbrida	1,0780	0,1368	0,1269	0,4784	0,0010	0,8036	1,3316

Tabela 3.8: Estatísticas das estimações obtidas pelos métodos clássico e Bayesianos, para amostra de tamanho $n=500$, com 35% de censura. Verdadeiros valores: $\alpha = 0,10$, $\beta = 9,00$, $\gamma = 0,80$, $\lambda = 0,40$ e $\pi = 0,15$.

Modelos WMG	θ_i	Métodos / Prioris	Média	DP	CV	EQM	PC	IC 95%	
								Inferior	Superior
Com mistura	α	Máxima Verossimilhança	0,5119	0,1103	0,2155	0,1818	0,4090	0,2957	0,7281
		Uniforme	0,3101	0,0858	0,2768	0,0515	0,4010	0,1497	0,4748
		Gama	0,3515	0,0731	0,2080	0,0686	0,3570	0,2027	0,5023
		Híbrida	0,3358	0,0879	0,2618	0,0633	0,4110	0,1646	0,5151
	β	Máxima Verossimilhança	5,2548	0,8099	0,1541	14,6817	0,2230	5,0386	5,4710
		Uniforme	24,9926	15,6346	0,6256	499,9588	0,4690	5,4849	59,2601
		Gama	15,8620	5,8025	0,3658	80,7212	0,6480	6,3064	26,0424
		Híbrida	18,6722	10,5083	0,5628	203,8649	0,6500	4,6164	39,2887
	γ	Máxima Verossimilhança	0,7938	0,4596	0,5790	0,2111	0,9570	0,5776	1,0100
		Uniforme	0,6091	0,2631	0,4320	0,1056	0,8720	0,1554	1,0728
		Gama	0,6292	0,3076	0,4889	0,1237	0,8440	0,1462	1,1927
		Híbrida	0,6681	0,2642	0,3955	0,0871	0,9240	0,2237	1,2246
	λ	Máxima Verossimilhança	0,0454	0,0191	0,4207	0,1261	0,0020	-0,1708	0,2616
		Uniforme	0,3154	0,1688	0,5353	0,0356	0,4540	0,0751	0,6948
		Gama	0,2250	0,0953	0,4236	0,0397	0,3090	0,0711	0,3999
		Híbrida	0,2418	0,1258	0,5201	0,0408	0,3440	0,0569	0,4815
	π	Máxima Verossimilhança	0,2900	0,0290	0,0999	0,0204	0,0010	0,0738	0,5062
		Uniforme	0,2836	0,0264	0,0930	0,0186	0,0060	0,2287	0,3325
		Gama	0,2854	0,0258	0,0904	0,0190	0,0010	0,2337	0,3356
		Híbrida	0,2848	0,0262	0,0919	0,0189	0,0040	0,2297	0,3331
Sem mistura	α	Máxima Verossimilhança	0,5258	0,1094	0,2081	0,1933	0,5150	0,3114	0,7402
		Uniforme	0,3150	0,0771	0,2449	0,0522	0,3650	0,1548	0,4672
		Gama	0,3527	0,0771	0,2187	0,0698	0,3060	0,2019	0,5067
		Híbrida	0,3422	0,0790	0,2309	0,0649	0,2980	0,1948	0,5186
	β	Máxima Verossimilhança	5,2108	0,7312	0,1403	14,8924	0,2470	4,9964	5,4252
		Uniforme	21,8946	9,4356	0,4310	255,2131	0,3310	6,8081	41,0708
		Gama	15,9143	6,2322	0,3916	86,6093	0,5940	5,6039	27,4370
		Híbrida	18,2632	7,4674	0,4089	141,5135	0,4350	6,0661	33,2337
	γ	Máxima Verossimilhança	0,7090	0,4567	0,6441	0,2166	0,9790	0,4946	0,9234
		Uniforme	0,4541	0,2010	0,4426	0,1600	0,7560	0,1251	0,8077
		Gama	0,5108	0,2351	0,4602	0,1388	0,8190	0,1335	0,9649
		Híbrida	0,4614	0,2426	0,5258	0,1734	0,7370	0,1176	0,9025
	λ	Máxima Verossimilhança	0,0401	0,0170	0,4239	0,1298	0,0060	-0,1743	0,2545
		Uniforme	0,3249	0,1274	0,3920	0,0219	0,5090	0,0858	0,5734
		Gama	0,2332	0,0992	0,4254	0,0377	0,3140	0,0651	0,4013
		Híbrida	0,2786	0,1204	0,4322	0,0292	0,4260	0,0607	0,4911
	π	Máxima Verossimilhança	0,2865	0,0290	0,1011	0,0195	0,0050	0,0721	0,5009
		Uniforme	0,2785	0,0287	0,1030	0,0173	0,0260	0,2196	0,3351
		Gama	0,2811	0,0276	0,0983	0,0179	0,0170	0,2246	0,3344
		Híbrida	0,2799	0,0280	0,1002	0,0177	0,0170	0,2257	0,3356
Sem fração de cura	α	Máxima Verossimilhança	0,0129	0,0164	1,2726	0,0079	0,9690	-0,0193	0,0451
		Uniforme	0,0130	0,0073	0,5637	0,0076	0,0100	0,0042	0,0225
		Gama	0,0131	0,0063	0,4781	0,0076	0,0110	0,0052	0,0258
		Híbrida	0,0131	0,0062	0,4703	0,0076	0,0140	0,0042	0,0239
	β	Máxima Verossimilhança	5,5930	0,5989	0,1071	11,9657	0,2350	5,5608	5,6252
		Uniforme	24,2129	5,1176	0,2114	257,5954	0,0200	14,3177	34,1922
		Gama	21,0790	4,1225	0,1956	162,8801	0,0610	13,1695	29,0410
		Híbrida	22,3580	4,9697	0,2223	203,1095	0,0610	13,0682	32,3569
	γ	Máxima Verossimilhança	1,8715	0,1579	0,0844	1,1730	0,2990	1,8393	1,9037
		Uniforme	1,1437	0,0996	0,0870	0,1280	0,0300	0,9500	1,3333
		Gama	1,1970	0,1002	0,0837	0,1677	0,0050	1,0262	1,3931
		Híbrida	1,1771	0,1026	0,0872	0,1527	0,0310	0,9771	1,3591
	λ	Máxima Verossimilhança	0,0763	0,0202	0,2642	0,1052	0,0050	0,0441	0,1085
		Uniforme	0,4580	0,0952	0,2078	0,0124	0,7590	0,2811	0,6532
		Gama	0,4028	0,0811	0,2014	0,0066	0,8860	0,2499	0,5560
		Híbrida	0,4245	0,0922	0,2173	0,0091	0,8560	0,2589	0,6199

Tabela 3.9: CV médio, PC média e Proporção de parâmetros dentro do IC das estimações obtidas pelos métodos clássico e Bayesianos, para todos os tamanhos amostras considerando dados sem censura e com 35% de censura.

n	Censura	Modelos WMG	CV médio				PC média				proporção de IC			
			MV	U	G	H	MV	U	G	H	MV	U	G	H
20	0%	Com mistura	70,8	39,0	26,3	42,6	91,6	81,3	99,9	99,5	100,0	60,0	100,0	100,0
		Sem mistura	66,8	41,3	28,4	42,2	84,3	86,4	99,2	99,2	100,0	80,0	100,0	100,0
		Sem fração de cura	64,8	36,3	26,1	33,9	81,4	91,4	99,6	99,4	100,0	25,0	100,0	100,0
	35%	Com mistura	68,7	42,0	31,1	43,6	80,8	81,9	89,8	91,6	80,0	60,0	60,0	80,0
		Sem mistura	70,8	49,1	32,5	43,9	77,2	80,2	87,5	89,6	80,0	100,0	60,0	80,0
		Sem fração de cura	56,4	37,4	29,6	46,9	77,7	92,9	96,1	94,0	75,0	75,0	75,0	100,0
50	0%	Com mistura	58,7	35,0	25,2	39,1	83,2	78,4	98,9	98,1	100,0	40,0	100,0	100,0
		Sem mistura	54,7	36,4	27,6	41,1	79,8	85,7	99,0	98,0	100,0	80,0	100,0	100,0
		Sem fração de cura	66,0	34,9	28,2	34,8	73,0	73,0	91,9	90,0	100,0	0,0	50,0	75,0
	35%	Com mistura	65,4	38,0	27,4	42,3	84,9	83,6	97,1	96,3	100,0	60,0	100,0	100,0
		Sem mistura	62,8	42,0	28,3	44,7	82,7	83,0	96,6	95,9	100,0	80,0	100,0	100,0
		Sem fração de cura	66,2	35,1	27,9	36,0	80,0	78,1	94,5	94,0	100,0	0,0	50,0	100,0
100	0%	Com mistura	50,1	36,0	27,1	39,7	76,5	81,2	98,9	96,7	100,0	40,0	100,0	100,0
		Sem mistura	47,5	35,7	28,1	44,6	74,5	86,9	98,9	97,3	100,0	60,0	100,0	100,0
		Sem fração de cura	59,5	36,4	33,5	39,8	67,9	47,3	72,3	62,5	100,0	0,0	25,0	25,0
	35%	Com mistura	48,2	45,9	34,6	44,1	65,5	74,7	81,1	81,6	80,0	100,0	80,0	80,0
		Sem mistura	48,5	43,0	34,1	47,9	66,2	75,1	80,0	79,6	80,0	100,0	80,0	80,0
		Sem fração de cura	50,8	33,7	35,2	45,7	60,5	85,7	94,2	82,5	75,0	75,0	75,0	100,0
500	0%	Com mistura	30,5	37,2	30,4	42,8	49,0	84,0	97,6	94,9	60,0	80,0	100,0	100,0
		Sem mistura	26,4	39,3	31,6	38,6	53,9	88,5	96,8	96,1	60,0	100,0	100,0	100,0
		Sem fração de cura	31,7	20,6	20,6	21,1	26,9	3,6	10,2	6,8	50,0	0,0	0,0	0,0
	35%	Com mistura	34,2	46,7	37,2	43,5	31,8	44,0	43,2	46,7	40,0	60,0	40,0	60,0
		Sem mistura	35,4	37,8	37,4	39,9	35,0	39,7	41,0	38,3	40,0	60,0	60,0	60,0
		Sem fração de cura	43,2	26,7	24,0	25,0	37,7	20,5	24,1	24,1	25,0	25,0	25,0	25,0

Aplicações

Neste capítulo é apresentado duas aplicações, pacientes com câncer internados na UTI do INCA e dados de câncer gástrico de pacientes do Hospital do Câncer de Barretos, com um estudo exploratório e comparando os modelos Gama Generalizado (GG), Gama, Log-normal (LN), Weibull, Exponencial Exponenciada (EE), Weibull Generalizada (WG) e Weibull Modificada Generalizada (WMG), sob o enfoque Clássico e Bayesiano, as parametrizações dos modelos podem ser encontradas no apêndice A. Os programas utilizados para essa seção são apresentados no apêndice B.

4.1 Pacientes com câncer internados na UTI do INCA

Para a obtenção dos resultados dos modelos de regressão utilizou-se os dados provenientes de estudo de coorte com 862 pacientes com câncer internados na UTI do INCA, publicado por [31], cujo objetivo foi avaliar fatores associados à sobrevivência. A tabela 4.1 apresenta as variáveis que compõem os dados utilizados.

Tabela 4.1: Variáveis utilizadas presentes nos dados de pacientes com câncer internados na UTI do INCA.

Variáveis	Descrição
t	Tempo a partir do primeiro caso acompanhado em dias.
$status$	Óbito = 1, censura = 0
$sexo$	Male = masculino, Fem = feminino
$fxetaria$	Idade até 60 anos, inclusive = 1, Mais de 60 anos = 0
$gptumor$	Tipo de tumor: Loco: sólido localizado; Mtx: metastático; Hemato: hematológico
$desnut$	1 = perda de peso recente acima de >10% ou IMC <18; 0 = não
$comorbi$	1 = comorbidades ¹ severas presentes; 0 = ausente
$leucopenia$	1 = leucopenia ² presente; 0 = ausente

¹Comorbidade é um termo usado para descrever a ocorrência simultânea de dois ou mais problemas de saúde em um mesmo indivíduo. Esse é um fenômeno freqüente na prática clínica, e sua identificação é um fator importante que afeta tanto o prognóstico dos pacientes como a conduta terapêutica do médico.

²Leucopenia é quando o número de leucócitos, que são as células de defesa do sangue, está baixo. O tratamento vai depender da causa e da intensidade. A leucopenia pode aparecer normalmente em algumas situações da vida, e não exigir tratamento algum; ou pode ser sinal de infecções, inflamações, doenças da medula óssea, doenças autoimunes, doenças da tireoide e do baço, além de consequência ao uso de algumas medicações e a tratamentos como quimioterapia e radioterapia.

Como análise exploratória dos dados construiu-se as estimativas de Kaplan-Meier para as variáveis que compõem o banco de dados. Na Figura 4.1 apresenta-se as curvas de sobrevivência estimadas pelo método de Kaplan-Meier para cada variável. Nos dados coletados não existe a variável faixa etária, nos dados é oferecida a idade em anos, mas para esse trabalho optou-se pela utilização de uma categorização da variável idade em duas categorias considerando como delimitador da primeira categoria a mediana (correspondente a 60 anos), ficando então uma categoria como “até 60 anos” e outra como “acima de 60 anos”.

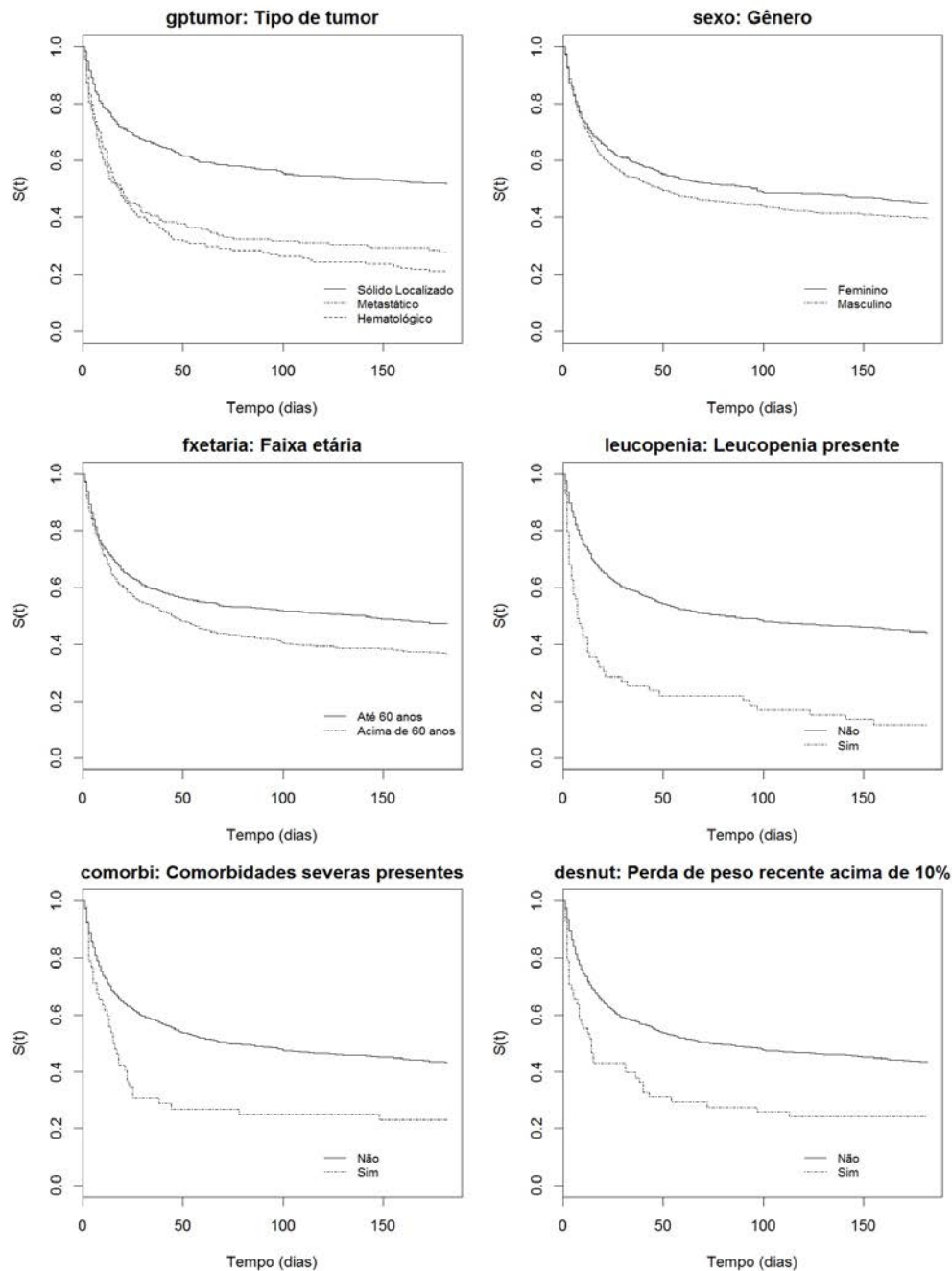


Figura 4.1: Curvas de sobrevivência estimadas pelo método de Kaplan-Meier para as covariáveis dos dados de pacientes com câncer internados na UTI do INCA.

No que diz respeito ao tipo de tumor, observa-se pela Tabela 4.2 que os pacientes com tumor sólido localizado têm um tempo maior até o óbito, do que os pacientes com tumor

do tipo Metastático e Hematológico, sendo que não aparentam apresentar ter grande diferença de tempo até o óbito. Quanto as variáveis faixa etária e gênero, ambas aparentam não fazer diferença no tempo até o óbito entre as categorias de cada variável. Quanto as variáveis presença de leucopenia, presença de comorbidades severas e perda de peso recente acima de 10%, o caso afirmativo para essas variáveis, aparenta diminuir o tempo até o óbito.

A Tabela 4.2 mostra os valores dos testes *logrank* e a modificação de Peto & Peto do teste de Wilcoxon-Gehan utilizados para testar a igualdade das curvas de sobrevivência obtidas para as covariáveis.

Tabela 4.2: Testes *logrank* e de Wilcoxon-Gehan utilizados para as covariáveis consideradas no estudo de pacientes com câncer.

Covariável	Testes (valor p)	
	<i>logrank</i>	Wilcoxon-Gehan
<i>gptumor</i> : Tipo do tumor	75,2 (< 0,001)	69,3 (< 0,001)
<i>sexo</i> : Gênero dos pacientes	2,6 (0,1080)	2,0 (0,1540)
<i>fxetaria</i> : Faixa etária dos pacientes	8,6 (0,0034)	6,6 (0,0101)
<i>leucopenia</i> : Presença de leucopenia	48,6 (< 0,001)	50,0 (< 0,001)
<i>comorbi</i> : Presença de comorbidades severas	13,4 (< 0,001)	12,6 (< 0,001)
<i>desnut</i> : Perda de peso recente acima de 10%	15,4 (< 0,001)	18,1 (< 0,001)

Segundo as recomendações em [31] os valores *p* menores que 0,25 em pelo menos um dos testes de comparação das curvas de sobrevivência devem permanecer no estudo, sendo assim, com base na Tabela 4.2, todas as covariáveis devem permanecer no estudo.

Na covariável *gptumor* existem três categorias, sendo necessário testar as categorias dois a dois para identificar quais das curvas de sobrevivência estimadas diferem. A Tabela 4.3 apresenta os valores dos testes *logrank* e a modificação de Peto & Peto do teste de Wilcoxon-Gehan utilizados para testar a igualdade das curvas de sobrevivência obtidas para a covariável *gptumor*.

Tabela 4.3: Testes *logrank* e de Wilcoxon-Gehan utilizados para a covariável *gptumor*.

Tipos de tumor	Testes (valor p)	
	<i>logrank</i>	Wilcoxon-Gehan
1: Sólido localizado vs 2: Metastático	39,1 (< 0,001)	40,2 (< 0,001)
1: Sólido localizado vs 3: Hematológico	58,3 (< 0,001)	51,0 (< 0,001)
2: Metastático vs 3: Hematológico	0,6 (0,4280)	0,0 (0,8830)

Conforme as recomendações em [31] já citada nos comentários da Tabela 4.2, com base na Tabela 4.3 não há evidências na diferença entre o tempo até o óbito para os tipos de tumor Metastático e Hematológico.

4.1.1 Análise semi-paramétrica e Seleção de Covariáveis

Para a seleção das covariáveis existem rotinas automáticas que podem ser utilizadas, tais como os métodos *forward*, *backward* e *stepwise*, porém como esses métodos têm a desvantagem de tenderem a identificar um particular conjunto de covariáveis, em vez de possíveis conjuntos igualmente adequados para explicar a variável resposta. As rotinas automáticas não possibilitam a inclusão de variáveis sem significância estatística mas com significância prática de acordo com a importância clínica. Sendo assim optou-se por

utilizar uma estratégia de seleção de modelos derivada da proposta de [30]. Para a seleção dos modelos adotou-se o nível de significância de referência de 0,10.

Os valores do ajuste do modelo de regressão paramétrico foram obtidos pelos pacotes *Survival* e *flexsurv* do programa R. A Tabela 4.4 apresenta o valor p do teste da razão das verossimilhanças com uma distribuição qui-quadrado de referência com graus de liberdade igual ao número de termos excluídos (diferença entre o número de parâmetros dos dois modelos a serem comparados).

Tabela 4.4: Seleção de covariáveis usando o modelo gama generalizado.

Passos	Modelo	$-2\log L(\theta)$	valor p
Passo 1	Nulo	6386,228	-
	gptumor	6319,662	<0,0001
	sexo	6383,687	0,1109
	fxetaria	6377,839	0,0038
	leucopenia	6351,976	<0,0001
	comorbi	6375,419	0,0010
	desnut	6373,789	0,0004
Passo 2	gptumor + fxetaria + leucopenia + comorbi + desnut	6259,331	-
	fxetaria + leucopenia + comorbi + desnut	6314,968	1,0000
	gptumor + leucopenia + comorbi + desnut	6284,577	1,0000
	gptumor + fxetaria + comorbi + desnut	6276,289	1,0000
	gptumor + fxetaria + leucopenia + desnut	6268,714	1,0000
	gptumor + fxetaria + leucopenia + comorbi	6270,181	1,0000
Passo 3	gptumor + fxetaria + leucopenia + comorbi + desnut	6259,331	-
	gptumor + fxetaria + leucopenia + comorbi + desnut + sexo	6258,124	0,2720
Passo 4	gptumor + fxetaria + leucopenia + comorbi + desnut	6259,331	-
	gptumor + fxetaria + leucopenia + comorbi + desnut + gptumor*leucopenia	6259,170	0,6885
	gptumor + fxetaria + leucopenia + comorbi + desnut + gptumor*fxetaria	6258,171	0,2816
	gptumor + fxetaria + leucopenia + comorbi + desnut + gptumor*comorbi	6258,308	0,3120
	gptumor + fxetaria + leucopenia + comorbi + desnut + gptumor*desnut	6258,377	0,3287
	gptumor + fxetaria + leucopenia + comorbi + desnut + leucopenia*fxetaria	6256,980	0,1252
	gptumor + fxetaria + leucopenia + comorbi + desnut + leucopenia*comorbi	6259,328	0,9555
	gptumor + fxetaria + leucopenia + comorbi + desnut + leucopenia*desnut	6251,939	0,0066
	gptumor + fxetaria + leucopenia + comorbi + desnut + fxetaria*comorbi	6259,062	0,6044
	gptumor + fxetaria + leucopenia + comorbi + desnut + fxetaria*desnut	6259,246	0,7716
	gptumor + fxetaria + leucopenia + comorbi + desnut + comorbi*desnut	6259,313	0,8946

Portanto há evidências de que o modelo mais adequado é aquele que contém as covariáveis *gptumor*, *fxetaria*, *leucopenia*, *comorbi*, *desnut* e a interação entre *leucopenia*desnut*. Sendo assim, essas covariáveis são importantes para descrever o comportamento do tempo até o óbito dos pacientes com câncer internados na UTI do INCA.

Na Figura 4.2, encontram-se os gráficos envolvendo o logaritmo da função de taxa de falha acumulada das covariáveis *gptumor*, *leucopenia*, *comorbi* e *desnut*. Pela figura observa-se que as curvas não indicam violação da suposição de taxas de falha proporcionais. A situação mais extrema de violação é caracterizada por curvas que se cruzam. Para o tipo de tumor considerou-se apenas duas categorias, pois pelos resultados obtidos através da Tabela 4.3 não há evidências de que a diferença entre o tempo até o óbito para os tipos de tumor Metastático e Hematológico seja estatisticamente significativa.

A Figura 4.3 apresenta os “resíduos padronizados de Schoenfeld” associados as covariáveis *gptumor*, *leucopenia*, *comorbi* e *desnut* e a interação entre *leucopenia*desnut* para averiguar a suposição de proporcionalidade de interação entre variáveis.

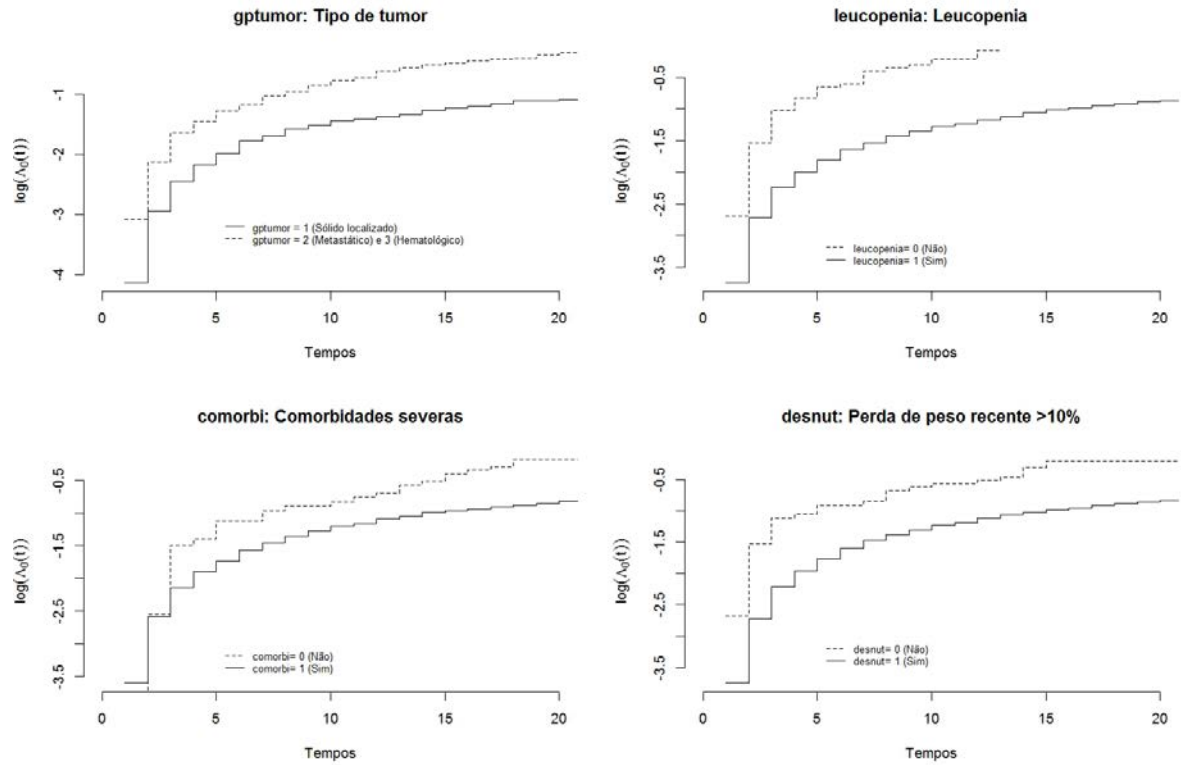


Figura 4.2: $\log(\hat{\Lambda}_{0j}(t))$ versus tempo para as covariáveis *gptumor*, *leucopenia*, *comorbi* e *desnut*.

Pela Figura 4.3 pode-se observar a ausência de tendências acentuadas para as covariáveis introduzidas no modelo, o mesmo pode ser confirmado pela Tabela 4.5. Desse modo, a análise de resíduos mostra também não haver evidências de violação da suposição de taxas de falhas proporcionais.

Tabela 4.5: Testes da proporcionalidade das taxas de falha no modelo ajustado.

Covariável	$\rho(\rho)$	χ^2	valor p
factor(tumor)1	0,00978	0,04774	0,827
leucopenia	0,02158	0,22956	0,632
fxetaria	0,02692	0,35095	0,554
comorbi	-0,02974	0,44443	0,505
desnut	-0,02415	0,28911	0,591
leucopenia:desnut	0,00273	0,00375	0,951

A Tabela 4.6 apresenta os resultados obtidos do ajuste do modelo de taxas de falha proporcionais de Cox com as covariáveis selecionadas e suas respectivas razões de taxas de falhas (RTF).

Pela Tabela 4.6 pode-se observar que todas as covariáveis presentes no modelo apresentam diferença significativa, ao nível de significância de 5%.

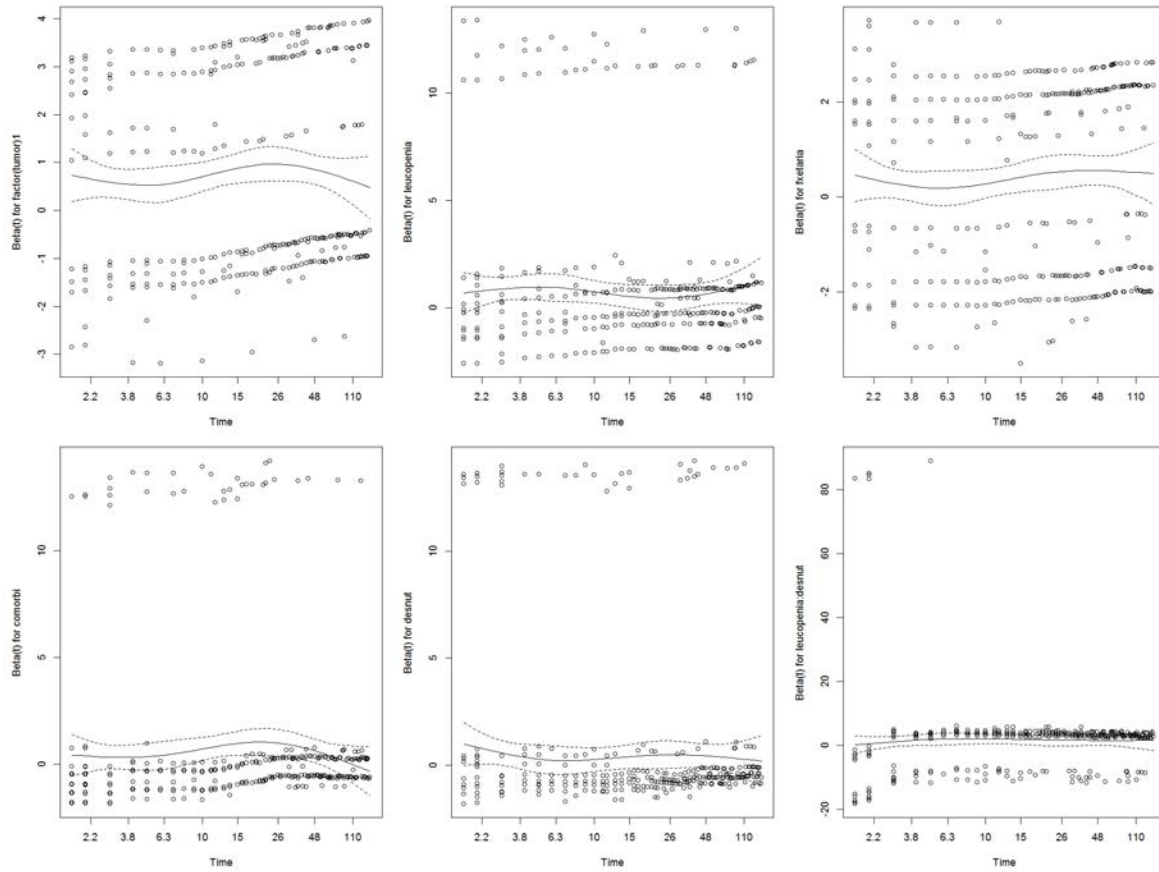


Figura 4.3: Resíduos padronizados de Schoenfeld associados às covariáveis.

Tabela 4.6: Resultados do ajuste do modelo de Cox para os dados do estudo e as razões de taxas de falhas (RTF).

Covariável	Estimativa	Erro-Padrão	Valor p	RTF	IC95% (RTF)
factor(gptumor)1	0,726	0,094	<0,0001	2,066	(1,719; 2,483)
leucopenia	0,753	0,163	<0,0001	2,123	(1,543; 2,921)
fxetaria	0,388	0,093	<0,0001	1,473	(1,227; 1,768)
comorbi	0,561	0,167	0,0008	1,752	(1,263; 2,431)
desnut	0,447	0,170	0,0086	1,564	(1,120; 2,183)
leucopenia:desnut	1,486	0,476	0,0018	4,419	(1,738; 11,235)

Os resultados da Tabela 4.6 podem ser interpretados da seguinte forma:

- a taxa de óbitos precoce de pacientes com câncer que apresentam tumor dos tipos Metastático e Hematológico são estimadas em 2,1 vezes maior de ocorrer do que em pacientes com câncer do tipo Sólido Localizado, com 95% de confiança que essa estimativa está entre 1,719 e 2,483.
- a taxa de óbitos precoce de pacientes com presença de leucopenia é estimada em 2,1 vezes maior de ocorrer do que os pacientes sem a presença da leucopenia, com 95% de confiança que essa estimativa está entre 1,543 e 2,921.
- a taxa de óbitos precoce de pacientes por faixa etária é estimada em 1,5 vezes maior para pacientes com idade acima de 60 anos, essa estimativa está entre 1,227 e 1,768, com 95% de confiança.

- a taxa de óbitos precoce de pacientes com presença de comorbidades severas é estimada em 1,8 vezes maior de ocorrer do que em pacientes sem a presença da comorbidades severas, com 95% de confiança que essa estimativa está entre 1,263 e 2,431.
- a taxa de óbitos precoce de pacientes com perda de peso recente acima de 10% é estimada em 1,6 vezes maior de ocorrer do que em pacientes sem perda de peso recente acima de 10%, com 95% de confiança que essa estimativa está entre 1,120 e 2,183.
- a taxa de óbitos precoce de pacientes com presença de leucopenia e com perda de peso recente acima de 10% é estimada em 4,4 vezes maior de ocorrer do que em pacientes sem presença de leucopenia e sem perda de peso recente acima de 10%, com 95% de confiança que essa estimativa está entre 1,738 e 11,235, ou seja, a interação entre a presença de leucopenia e a perda peso aumenta a taxa de óbitos precoce.

4.1.2 Modelos sem fração de cura

Nessa seção é realizado o estudo dos modelos paramétricos Gama Generalizado (GG), Gama, Log-normal (LN), Weibull, Exponencial Exponenciada (EE), Weibull Generalizada (WG) e Weibull Modificada Generalizada (WMG), sem fração de cura sob os enfoques Clássico e Bayesiano. As tabelas com as estimativas da função de sobrevivência são apresentadas no Apêndice A.

4.1.2.1 Estimação Clássica

Para a estimação sob o enfoque clássico utilizando o pacote *maxLik* do *software* R, realizou-se a comparação entre os modelos paramétricos Gama Generalizado (GG), Gama, Log-normal (LN), Weibull, Exponencial Exponenciada (EE), Weibull Generalizada (WG) e Weibull Modificada Generalizada (WMG), esse último o modelo alvo do estudo.

A Tabela 4.7 apresenta os testes da razão de verossimilhanças (TRV) para a identificação do modelo paramétrico com evidências de ser o mais adequado aos dados, comparando-os o modelo WMG.

Tabela 4.7: Logaritmo da função de verossimilhança e resultados do Erro Padrão (EP), teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos sem fração de cura.

Modelo	θ_i	Estimativa	EP	$-\log(L(\theta))$	TRV	Valor p	BIC	CAIC
GG	α	4,46233	0,26047	2767,407	71,67	<0,0001	5555,09	5540,88
	β	0,04875	0,02570					
	γ	0,18948	0,00737					
Gama	α	0,41113	0,02010	2823,743	184,35	<0,0001	5661,00	5651,53
	β	0,00140	0,00018					
LN	μ	4,42364	0,10208	2742,369	21,60	<0,0001	5498,26	5488,79
	σ	2,62351	0,09255					
Weibull	γ	0,47816	0,18880	2797,776	132,41	<0,0001	5609,07	5599,60
	λ	0,07786	0,00763					
EE	α	0,00184	0,00022	2824,554	185,97	<0,0001	5662,63	5653,15
	β	0,40907	0,01997					
WG	α	3,83162	0,04307	2724,338	14,46	<0,0001	5468,95	5454,75
	β	149,50230	1,55000					
	γ	0,07845	0,00276					
WMG	α	2,76274	0,06269	2687,600	-	-	5490,18	5471,24
	β	51,87790	2,78100					
	γ	0,10021	0,00497					
	λ	$1,66 \times 10^7$	0,00011					

Com base na Tabela 4.7, há indicação da adequação do modelo WG para a análise dos dados em estudo, tanto pelo TRV quanto pelos critérios BIC e CAIC. A partir da Figura 4.4 é possível observar que o modelo WMG parece ser mais adequado aos dados, pois a curva localiza-se mais próxima da curva de KM, e como os valores obtidos nos critérios entre os modelos WG e WMG, então há indícios que o modelo WMG seja o mais adequado dentre os modelos sem fração de cura.

Pela Figura 4.5, que apresenta os gráficos de sobrevivência estimada por Kaplan-Meier e observa-se que o modelo WMG apresenta o ajuste mais satisfatório.

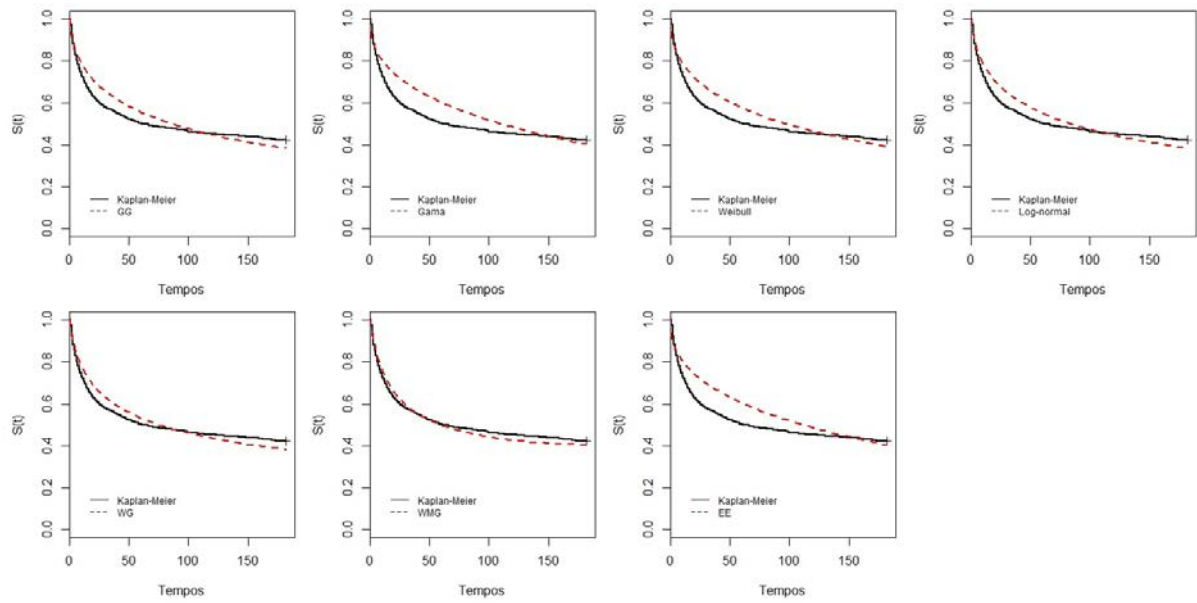


Figura 4.4: Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) *versus* a curva de sobrevivência estimada por Kaplan-Meier, sem fração de cura, sob enfoque clássico.

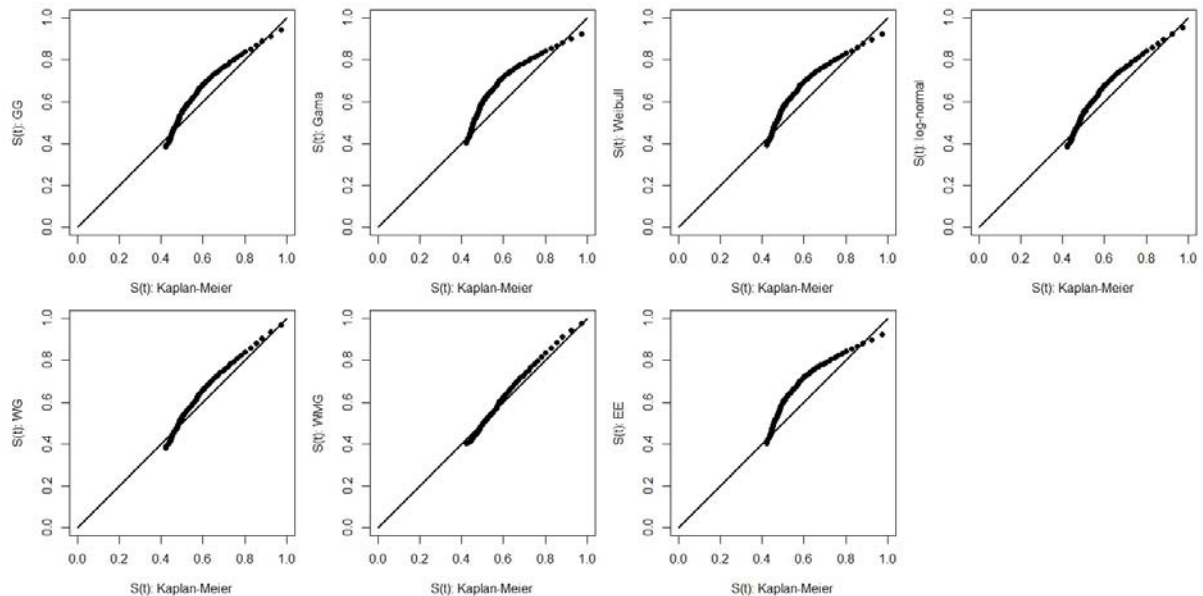


Figura 4.5: Gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), sem fração de cura, sob enfoque clássico.

4.1.2.2 Estimação Bayesiana

Agora utilizando a Inferência Bayesiana reproduziu-se as mesmas análises da Inferência Clássica para os modelos citados na seção anterior. Para a estimação de todos os parâmetros de todos os modelos utilizou-se priores não informativas. A convergência do algoritmo foi verificada a partir de gráficos da densidade a posteriori, autocorrelação e histórico de séries temporais das amostras simuladas dos modelos citados, que são apresentados no Apêndice C. A Tabela 4.8 apresenta os resultados das posteriores. Para a obtenção dos resultados utilizou-se o *software Openbugs*, considerando *prioris* não informativas, que podem ser consultadas no apêndice B.

Tabela 4.8: Resultados das posteriores assumindo o modelo sem fração de cura.

Modelo	θ_i	Média	Mediana	DP	IC (95%)		EBIC	ECAIC
					Inferior	Superior		
GG	α	0,0165	0,0163	0,0021	0,0127	0,0210	8905,91	8891,71
	β	0,0050	0,0050	0,0001	0,0049	0,0051		
	γ	37,9800	37,7200	4,8410	29,3400	48,2000		
Gama	α	0,6886	0,6877	0,0282	0,6347	0,7438	9461,16	9455,28
	β	0,0074	0,0074	0,0004	0,0066	0,0083		
LN	μ	3,6400	3,6400	0,0569	3,5300	3,7520	9646,73	9637,92
	σ	1,6700	1,6700	0,0400	1,5940	1,7510		
Weibull	γ	0,7981	0,7980	0,0242	0,7512	0,8470	9490,81	9484,94
	λ	0,0294	0,0293	0,0036	0,0228	0,0371		
EE	α	0,0019	0,0019	0,0002	0,0014	0,0023	5664,66	5658,78
	β	0,4102	0,4089	0,0200	0,3730	0,4508		
WG	α	1,5170	1,5180	0,2049	1,1100	1,9070	5511,43	5502,61
	β	12,5900	12,3300	3,0300	7,4960	19,1700		
	γ	0,1481	0,1470	0,0132	0,1259	0,1772		
WMG	α	1,4110	1,4190	0,1695	1,0330	1,7050	5523,56	5511,81
	β	10,9100	10,9200	2,0360	6,7640	14,8400		
	γ	0,1537	0,1526	0,0126	0,1336	0,1847		
	λ	$1,44 \times 10^5$	$1,06 \times 10^5$	$1,42 \times 10^5$	$5,25 \times 10^7$	$5,31 \times 10^5$		

Esses resultados sugerem que o modelo WG é o mais apropriado para análise destes dados, já que todos os parâmetros são significantes e apresenta o menor valor dos critérios EBIC e ECAIC. A partir da Figura 4.6 é possível observar que os modelos WG e WMG parecem ser mais adequados aos dados, pois a curva se apresenta mais próxima da curva de KM, mas como o parâmetro λ é muito próximo de zero, o modelo mais simples, WG, deve ser utilizado para a estimação dos dados sem fração de cura.

De acordo com a Figura 4.7, que apresenta as curvas de sobrevivência empírica estimada por Kaplan-Meier, observa-se que os modelos WG e WMG apresenta o ajuste mais satisfatório.

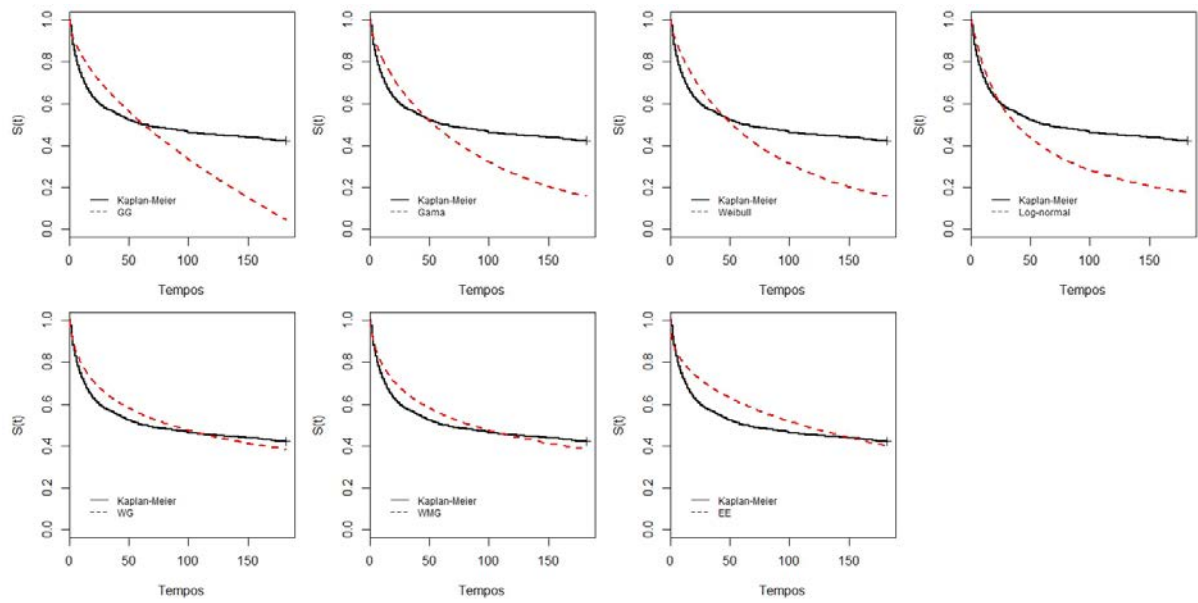


Figura 4.6: Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) *versus* a curva de sobrevivência estimada por Kaplan-Meier, sem fração de cura, sob enfoque Bayesiano.

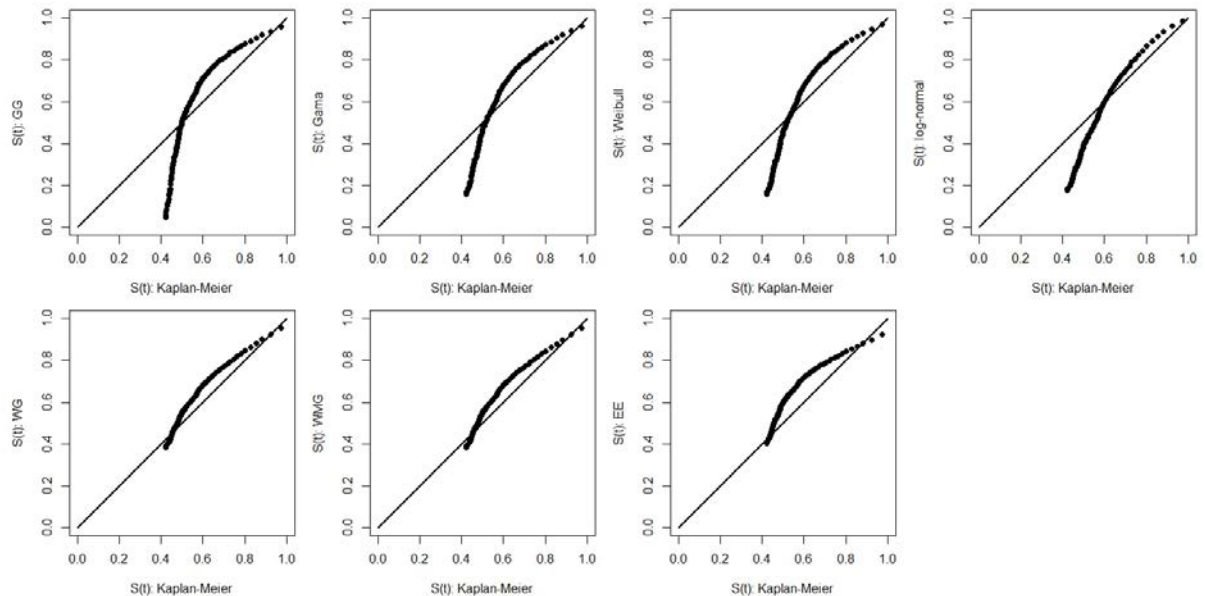


Figura 4.7: Gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), sem fração de cura, sob enfoque Bayesiano.

4.1.3 Modelos de fração de cura de mistura

Aqui é realizado novamente o estudo dos modelos paramétricos Gama Generalizado (GG), Gama, Log-normal (LN), Weibull, Exponencial Exponenciada (EE), Weibull Generalizada (WG) e Weibull Modificada Generalizada (WMG), mas agora considerando a fração de cura de mistura, sob os enfoques Clássico e Bayesiano. As tabelas com as estimativas da função de sobrevivência são apresentadas no Apêndice A.

4.1.3.1 Estimação Clássica

A Tabela 4.9, apresenta os testes da razão de verossimilhanças para a identificação do modelo paramétrico utilizando fração de cura de mistura com evidências de ser o mais adequado aos dados, sendo assim, há indicação da adequação do modelo WG para a análise dos dados em estudo, pelos critérios BIC e CAIC. O teste TRV mostra que os modelos WG e WMG podem ser considerados como adequados aos dados.

Tabela 4.9: Logaritmo da função de verossimilhança e resultados do teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos de fração de cura de mistura.

Modelo	θ_i	Estimativa	EP	$-\log(L(\theta))$	TRV	Valor p	BIC	CAIC
GG	α	15,0556	0,2955	2817,893	288,21	<0,0001	5662,82	5643,88
	β	0,0012	0,0001					
	γ	0,2895	0,0034					
	π	0,4531	0,0171					
Gama	α	0,7143	0,0401	2733,220	118,87	<0,0001	5486,72	5472,51
	β	0,0243	0,0021					
	π	0,4177	0,0170					
LN	μ	2,6560	0,0772	2682,762	17,95	<0,0001	5385,80	5371,60
	σ	1,4634	0,0610					
	π	0,3965	0,0185					
Weibull	γ	25,5920	1,49128	2721,673	95,77	<0,0001	5463,62	5449,42
	λ	0,7646	0,02856					
	π	0,4145	0,01714					
EE	α	0,0269	0,0020	2734,982	122,39	<0,0001	5490,24	5476,03
	β	0,7125	0,0415					
	π	0,4180	0,0169					
WG	α	4,3796	0,0531	2674,371	1,17	<0,0001	5375,78	5356,84
	β	295,5185	1,4910					
	γ	0,1220	0,0058					
	π	0,3753	0,0203					
WMG	α	4,8960	0,0559	2673,786	-	-	5374,61	5355,67
	β	500,0496	1,4880					
	γ	0,1114	0,0053					
	λ	$4,43 \times 10^5$	0,0002					
	π	0,3765	0,0283					

A partir da Figura 4.8 é possível observar que todos os modelos de fração de cura de mistura apresentaram um ajuste mais próximo da curva de KM, destacando-se os modelos LN, WG e WMG. A Figura 4.9, apresenta os gráficos de sobrevivência estimada por Kaplan-Meier e as curvas de sobrevivências dos modelos estudados, com a mesma indicação da Figura 4.8. Assim, considerando os resultados, os modelos WG e WMG são os mais apropriados para o caso.

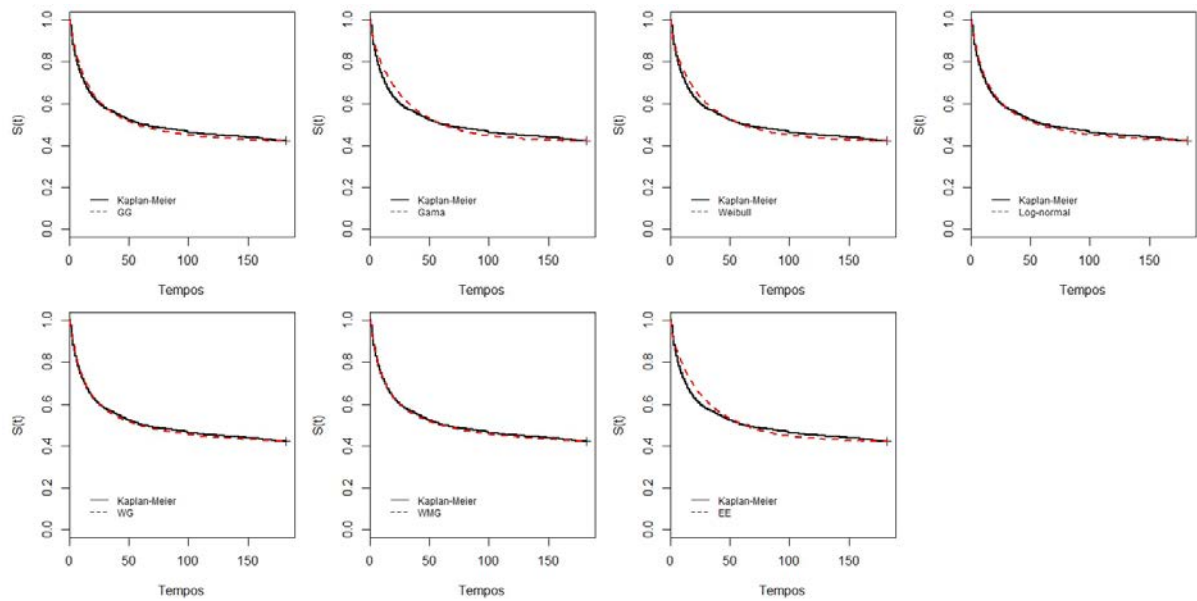


Figura 4.8: Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) *versus* a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de mistura, sob enfoque clássico.

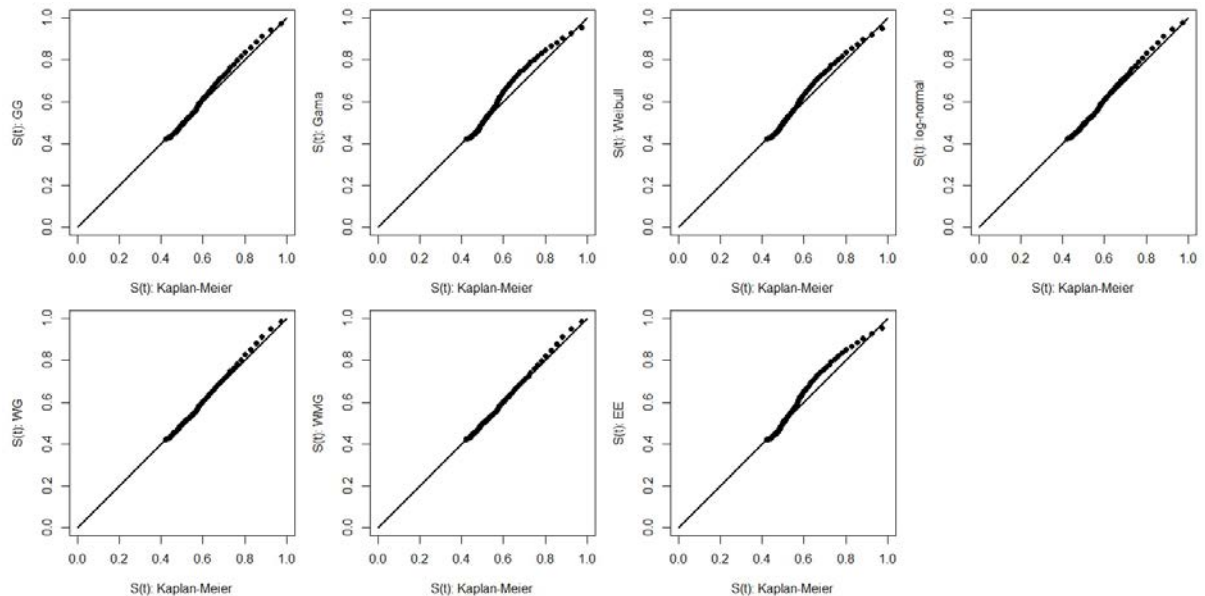


Figura 4.9: Gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de mistura, sob enfoque clássico.

4.1.3.2 Estimação Bayesiana

Nesta seção reproduziu-se as análises da seção anterior agora sob o enfoque Bayesiano para os modelos de fração de cura de mistura citados na seção anterior, utilizando para todos os parâmetros de todos os modelos utilizou-se priores não informativas. A Tabela 4.10, observa-se que resultados sugerem o modelo WG como o mais apropriado para análise destes dados, já que todos os parâmetros são significantes e os critérios EBIC e ECAIC apresentam o menor valor.

Tabela 4.10: Resultados das posteriores assumindo o modelo de fração de cura de mistura.

Modelo	θ_i	Média	Mediana	DP	IC (95%)		EBIC	ECAIC
					Inferior	Superior		
GG	α	0,0167	0,0165	0,0023	0,0126	0,0216	7525,58	7513,84
	β	0,0050	0,0050	0,0001	0,0049	0,0051		
	γ	37,5100	37,2800	5,1030	28,6200	48,5300		
	π	0,3941	0,3942	0,0181	0,3583	0,4295		
Gama	α	0,6880	0,6876	0,0278	0,6343	0,7438	7956,67	7947,85
	β	0,0074	0,0074	0,0004	0,0066	0,0083		
	π	0,3147	0,3150	0,0229	0,2691	0,3597		
LN	μ	3,6390	3,6390	0,0563	3,5280	3,7480	8119,68	8107,93
	σ	1,6700	1,6700	0,0405	1,5930	1,7510		
	π	0,2993	0,2996	0,0237	0,2522	0,3451		
Weibull	γ	0,7983	0,7978	0,0228	0,7547	0,8436	7985,78	7976,97
	λ	0,0294	0,0292	0,0034	0,0232	0,0365		
	π	0,3143	0,3142	0,0227	0,2697	0,3592		
EE	α	0,0270	0,0269	0,0020	0,0231	0,0309	5493,22	5484,41
	β	0,7133	0,7125	0,0409	0,6350	0,7954		
	π	0,4183	0,4180	0,0170	0,3854	0,4514		
WG	α	1,0980	1,0960	0,2005	0,7207	1,4870	5404,04	5392,29
	β	8,1160	7,9000	2,1120	4,5920	12,7400		
	γ	0,3094	0,3063	0,0346	0,2543	0,3788		
	π	0,3996	0,3999	0,0184	0,3627	0,4346		
WMG	α	1,1170	1,1130	0,1959	0,7447	1,5030	5412,00	5397,31
	β	8,2700	8,0220	2,1200	4,7300	12,9100		
	γ	0,3048	0,3025	0,0313	0,2517	0,3717		
	λ	0,0002	0,0002	0,0002	0,0000	0,0008		
	π	0,4036	0,4036	0,0184	0,3669	0,4390		

A Figura 4.10 mostra que os modelos WG e WMG, parecem ser mais adequados aos dados, pois a curva de sobrevivência estimada pelos modelos se apresentam mais próximas da curva de sobrevivência estimada por KM. Os mesmo ocorre na Figura 4.11 pelos gráficos da sobrevivências estimadas por KM *versus* os modelos estudados.

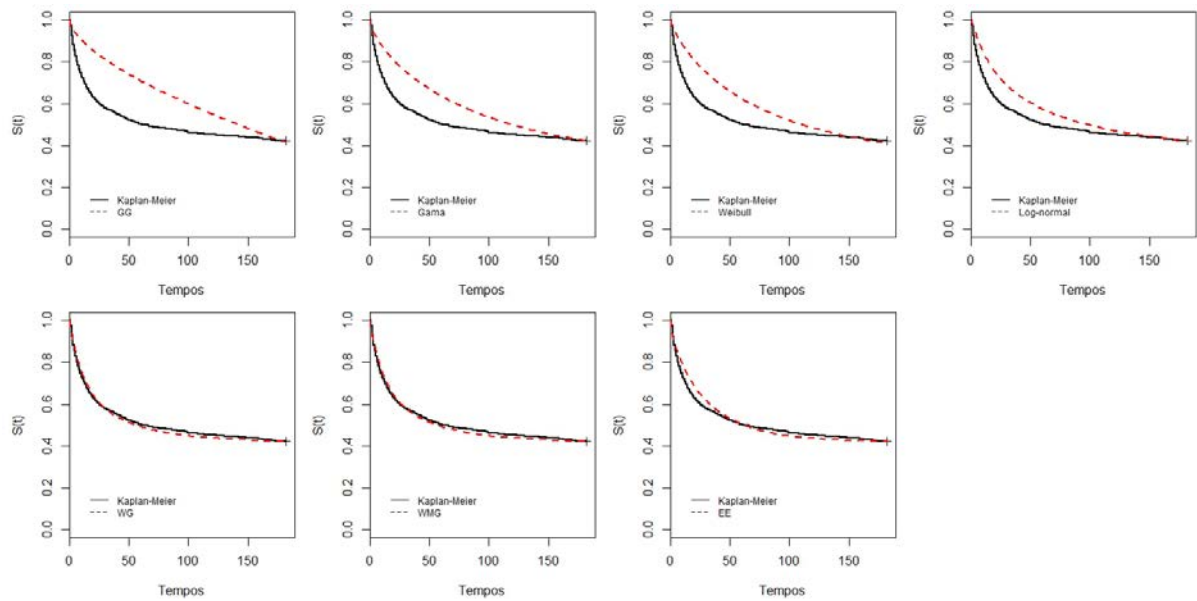


Figura 4.10: Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) *versus* a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de mistura, sob enfoque Bayesiano.

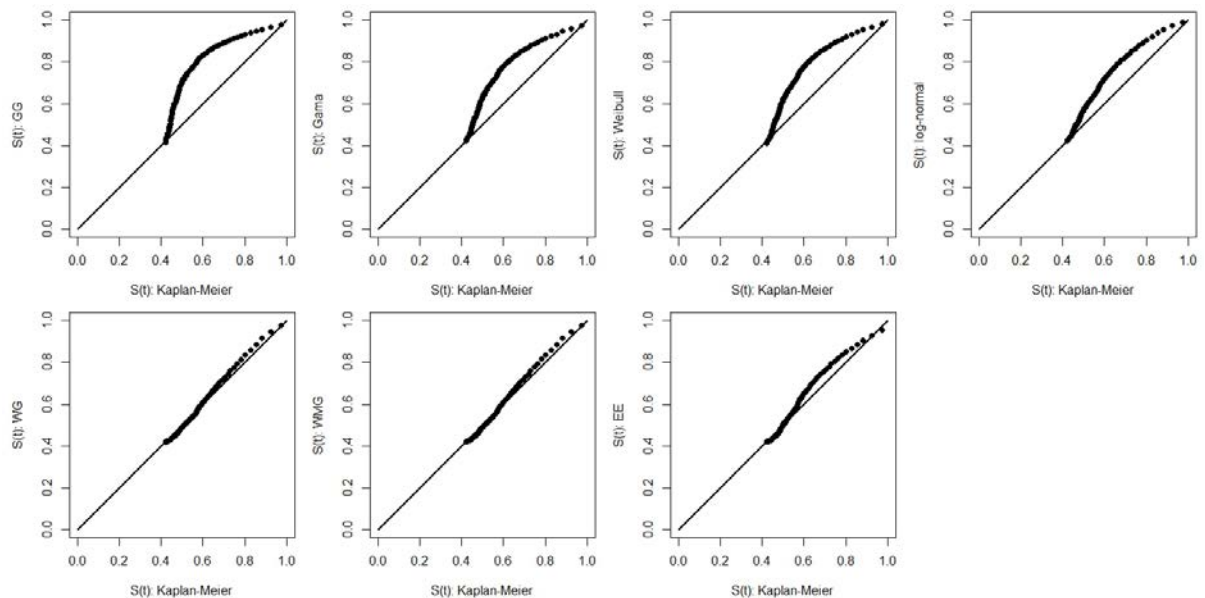


Figura 4.11: Gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de mistura, sob enfoque Bayesiano.

4.1.4 Modelos de fração de cura de não-mistura

Apresenta-se nessa seção os modelos estudados nas seções anteriores, mas dessa vez considerando fração de cura de não-mistura, sob os enfoques Clássico e Bayesiano. As tabelas com as estimativas da função de sobrevivência são apresentadas no Apêndice A.

4.1.4.1 Estimação Clássica

A Tabela 4.11 apresenta os testes da razão de verossimilhanças para a identificação do modelo paramétrico utilizando fração de cura de não-mistura com evidências de ser o mais adequado aos dados. Com base na Tabela 4.11 há indícios da adequação do modelo WG para a análise dos dados em estudo, pelos critérios BIC e CAIC. O teste TRV mostra que os modelos WG e WMG podem ser considerados como adequados aos dados.

Tabela 4.11: Logaritmo da função de verossimilhança e resultados do teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos de fração de cura de não-mistura.

Modelo	θ_i	Estimativa	EP	$-\log(L(\theta))$	TRV	Valor p	BIC	CAIC
GG	α	4,4551	0,4468	2692,885	2941,17	<0,0001	5412,81	5393,87
	β	0,3434	0,1833					
	γ	0,3449	0,0216					
	π	0,4121	0,0179					
Gama	α	0,8159	0,0426	2720,477	96,35	<0,0001	5461,23	5447,03
	β	0,0216	0,0021					
	π	0,4083	0,0171					
LN	μ	3,0798	0,1021	2680,493	16,38	0,0003	5381,26	5367,06
	σ	1,5182	0,0688					
	π	0,3896	0,0193					
Weibull	γ	35,7638	0,0306	2714,376	84,15	<0,0001	5449,03	5434,82
	λ	0,8327	0,5239					
	π	0,4084	0,0167					
EE	α	0,0232	0,0020	2721,240	97,88	<0,0001	5462,76	5448,55
	β	0,8177	0,0440					
	π	0,4081	0,0170					
WG	α	3,9880	0,0455	2674,480	4,36	0,0368	5376,00	5357,06
	β	223,2305	1,5010					
	γ	0,1168	0,0061					
	π	0,3676	0,0214					
WMG	α	7,0633	0,0461	2672,301	-	-	5378,40	5354,73
	β	5000,0190	1,4830					
	γ	0,0719	0,0037					
	λ	0,0001	0,0002					
	π	0,3726	0,0314					

A Figura 4.12 apresenta as curvas de sobrevivência estimadas e pode-se observar que todos os modelos de fração de cura de não-mistura apresentaram um ajuste mais próximo da curva de sobrevivência estimada por KM, destacando-se os modelos WG e WMG. Isso sendo confirmado pelos gráficos das estimativas de sobrevivências apresentadas pela Figura 4.13.

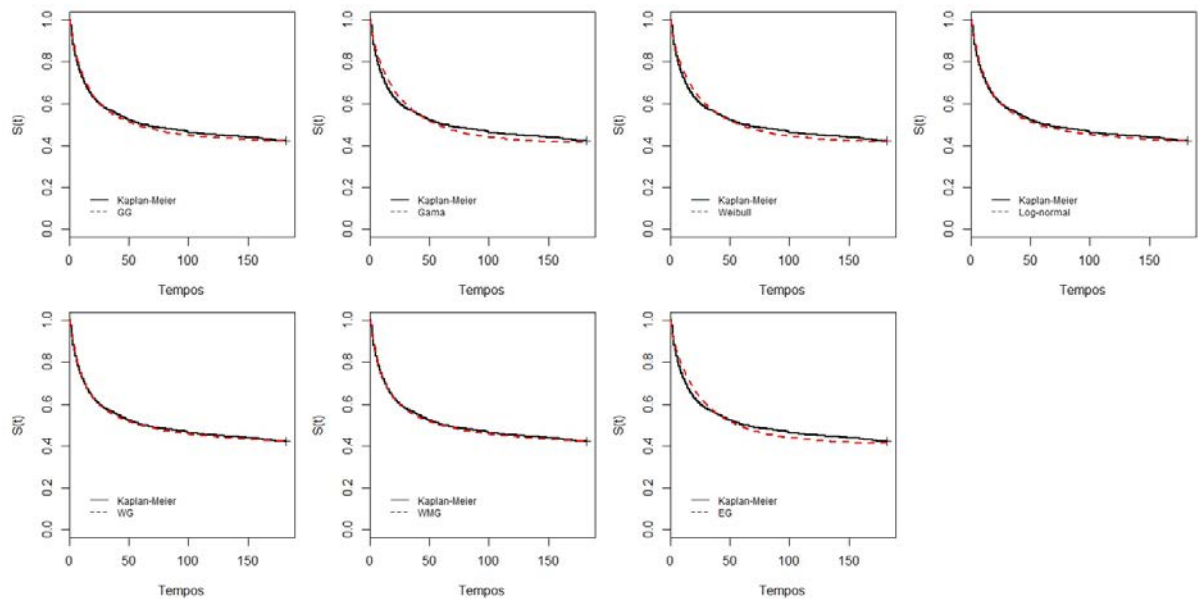


Figura 4.12: Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) *versus* a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de não-mistura, sob enfoque clássico.

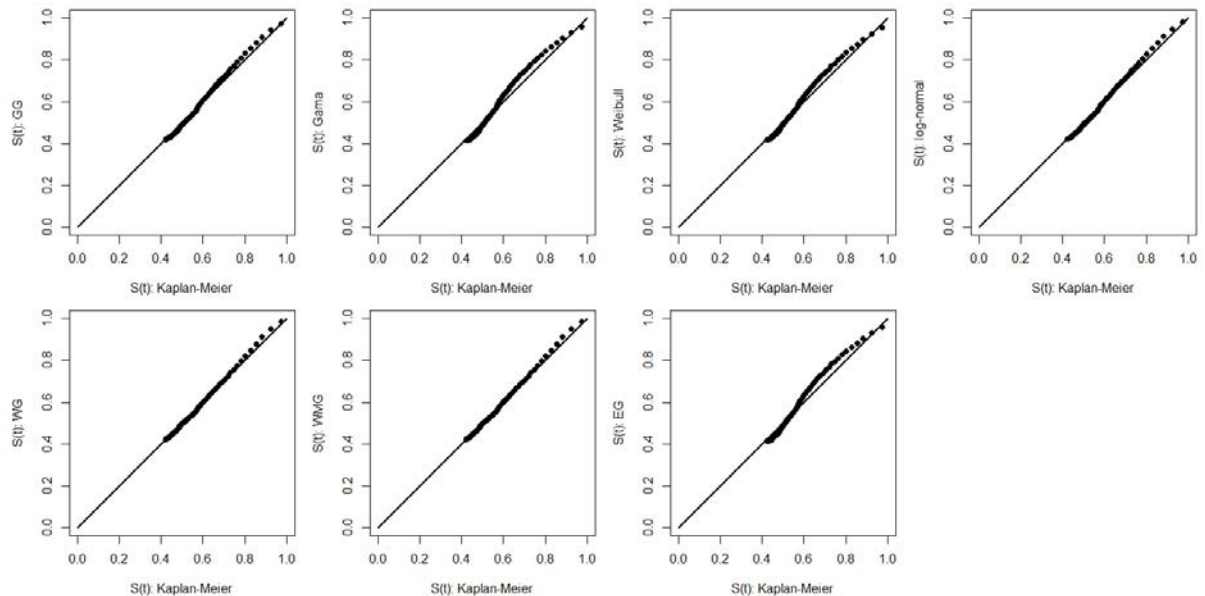


Figura 4.13: Gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de não-mistura, sob enfoque clássico.

4.1.4.2 Estimação Bayesiana

Os resultados dessa seção foram obtidos considerando a fração de cura de não-mistura para os modelos utilizados nas seções anteriores, sob o enfoque Bayesiano, utilizando-se prioris não informativas. Pela Tabela 4.12, os resultados sugerem que os modelos WG e WMG são os mais apropriados para análise destes dados, já que todos os parâmetros são significantes e apresentam o menor valor dos critérios EBIC e ECAIC.

Tabela 4.12: Resultados das posteriores assumindo o modelo de fração de cura de não-mistura.

Modelo	θ_i	Média	Mediana	DP	IC (95%)		EBIC	ECAIC
					Inferior	Superior		
GG	α	0,0167	0,0165	0,0022	0,0129	0,0213	7344,35	7332,60
	β	0,0050	0,0050	0,0001	0,0049	0,0051		
	γ	37,5700	37,3700	4,8010	28,9100	47,6800		
	π	0,3492	0,3493	0,0171	0,3159	0,3833		
Gama	α	0,6884	0,6876	0,0280	0,6348	0,7439	7830,06	7821,25
	β	0,0074	0,0074	0,0004	0,0066	0,0083		
	π	0,3241	0,3238	0,0186	0,2880	0,3617		
LN	μ	3,6400	3,6400	0,0572	3,5280	3,7520	8034,07	8022,32
	σ	1,6700	1,6690	0,0398	1,5940	1,7490		
	π	0,3313	0,3309	0,0187	0,2955	0,3683		
Weibull	γ	0,7980	0,7980	0,0224	0,7545	0,8425	7851,79	7842,98
	λ	0,0294	0,0293	0,0034	0,0233	0,0365		
	π	0,3212	0,3210	0,0180	0,2867	0,3580		
EE	α	0,0233	0,0232	0,0019	0,0195	0,0271	5465,71	5456,90
	β	0,8181	0,8170	0,0434	0,7347	0,9061		
	π	0,4081	0,4082	0,0169	0,3751	0,4410		
WG	α	0,9396	0,9332	0,1975	0,5720	1,3350	5399,85	5388,10
	β	7,6360	7,3610	2,1030	4,2760	12,4500		
	γ	0,3123	0,3092	0,0374	0,2490	0,3937		
	π	0,3934	0,3936	0,0188	0,3564	0,4294		
WMG	α	0,9638	0,9598	0,1987	0,5824	1,3530	5407,58	5383,91
	β	7,8190	7,5770	2,1250	4,2980	12,4200		
	γ	0,3068	0,3038	0,0363	0,2455	0,3870		
	λ	0,0003	0,0002	0,0003	0,0000	0,0010		
	π	0,3994	0,3995	0,0183	0,3634	0,4348		

Na Figura 4.14 é possível observar que os modelos WG e WMG, parecem ser mais adequados aos dados, pois a curva de sobrevivência estimadas pelos modelos em estudo, se apresenta mais próxima da curva de sobrevivência estimada por KM, resultado confirmado pela Figura 4.15.

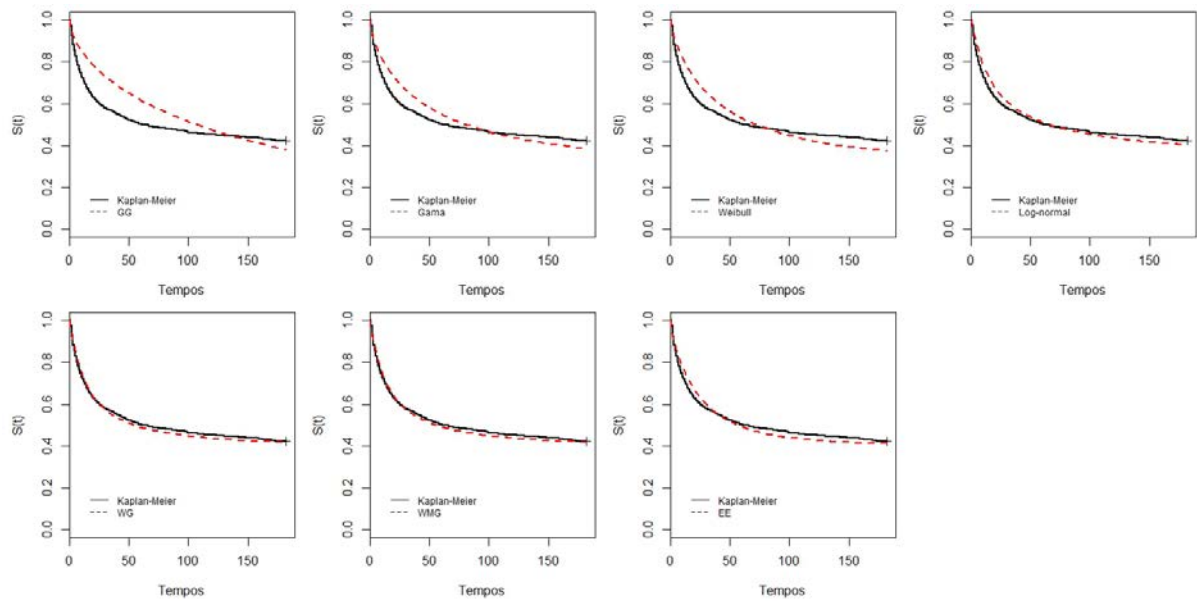


Figura 4.14: Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) *versus* a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de não-mistura, sob enfoque Bayesiano.

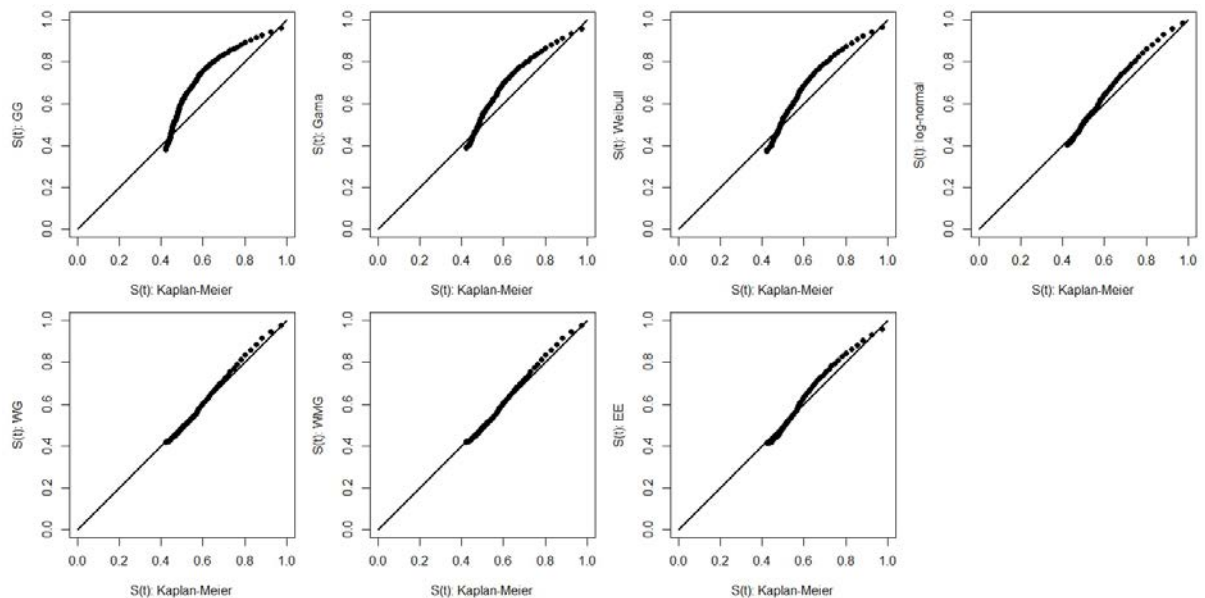


Figura 4.15: Gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de não-mistura, sob enfoque Bayesiano.

4.1.5 Resumo da aplicação com Pacientes com Câncer na UTI do INCA

De acordo com os resultados observados nas seções 4.1.2, 4.1.3 e 4.1.4, sob o enfoque clássico as melhores indicações de ajustes aos dados foram para os modelos WG e WMG, sendo que quando consideramos o modelo com fração de cura apresenta-se um melhor ajuste segundo os critérios BIC e CAIC, de mistura ou de não-mistura, pois a diferença entre as estatísticas dos critérios é muito pequena, como segue na Tabela 4.13.

Tabela 4.13: Resumo da análise sob o enfoque clássico segundo os critérios Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC).

	Modelo	BIC	CAIC
Sem fração de cura	WMG	5490,18	5471,24
Mistura	WG	5375,78	5356,84
	WMG	5374,61	5355,67
Não-mistura	WG	5376,00	5357,06
	WMG	5378,40	5354,73

Sendo assim, a Figura 4.16, apresenta a curva de sobrevivência (esquerda) e de risco (direita) estimadas pelos modelos WG e WMG, considerando fração de cura de mistura e de não mistura, sob enfoque clássico.

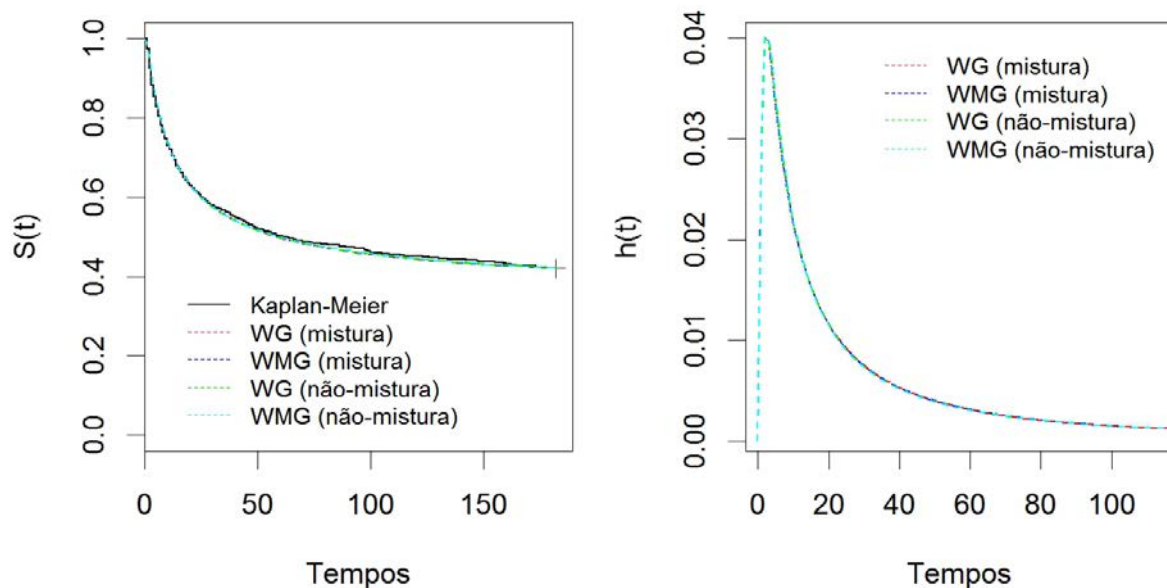


Figura 4.16: Curvas de sobrevivência (esquerda) e de risco (direita) estimadas pelos modelos WG e WMG de fração de cura de mistura e de não-mistura para os dados de pacientes internados na UTI do INCA, sob o enfoque clássico.

Sob o enfoque Bayesiano, observa-se que segundo os critérios EBIC e ECAIC mais uma vez os modelos WG e WMG foram os que indicaram melhor ajuste aos dados e pela Tabela 4.14, observa-se que os modelos que consideram a fração de cura têm menor valor da estatística dos critérios EBIC e ECAIC.

A Figura 4.17, apresenta a curva de sobrevivência (esquerda) e de risco (direita) estimadas pelos modelos WG e WMG, considerando fração de cura de mistura e de não mistura, sob enfoque Bayesiano.

Tabela 4.14: Resumo da análise sob o enfoque Bayesiano segundo os critérios EBIC e ECAIC.

	Modelo	EBIC	ECAIC
Sem fração de cura	WG	5511,43	5502,61
Mistura	WG	5404,04	5392,29
	WMG	5412,00	5397,31
Não-mistura	WG	5399,85	5388,10
	WMG	5407,58	5383,91

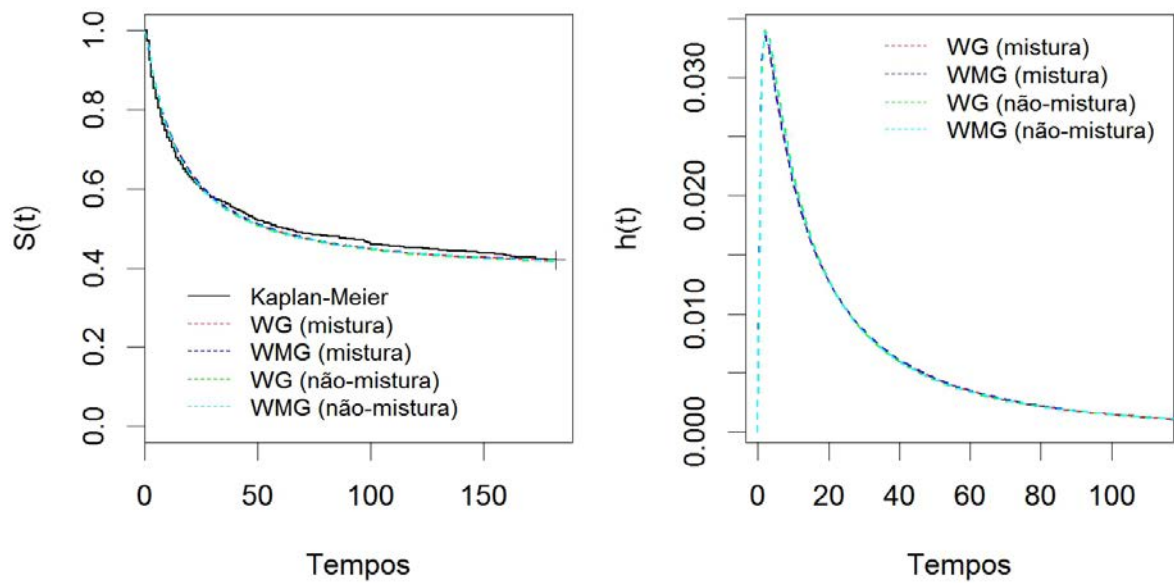


Figura 4.17: Curvas de sobrevivência (esquerda) e de risco (direita) estimadas pelos modelos WG e WMG de fração de cura de mistura e de não-mistura para os dados de pacientes internados na UTI do INCA, sob o enfoque Bayesiano.

Consequentemente, tanto o modelo Weibull Generalizado (WG) quanto o modelo Weibull Modificado Generalizado (WMG) considerando a fração de cura podem ser utilizados para os dados utilizados nessa aplicação. Vale lembrar que o modelo WG é um caso especial do modelo WMG.

4.1.6 Acrescentando covariáveis ao modelo selecionado para a aplicação com Pacientes com Câncer na UTI do INCA

Como verificado na seção anterior o modelo WG considerando a fração de cura, de mistura ou não mistura, são os mais adequados aos dados da aplicação, agora acrescenta-se as covariáveis (*gptumor*, *fxetaria*, *leucopenia*, *comorbi*, *desnut* e a interação entre *leucopenia*desnut*) consideradas significativas na seção 4.1.1 e verifica-se se há ganho com relação ao modelo sem covariáveis.

Na tabela 4.15 é apresentado os resultados do modelo WG de fração de cura de mistura e não mistura considerando covariáveis sob o enfoque clássico, também é apresentado os valores dos critérios TRV, BIC e CAIC para esses modelos.

Tabela 4.15: Logaritmo da função de verossimilhança e resultados do teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos sem e com covariáveis.

Modelo	θ_i	Estimativa	EP	$-\log(L(\theta))$	TRV	Valor p	BIC	CAIC
WG com mistura	β	4,32842	0,73886	2630,092	-	-	5368,33	5293,00
	γ	0,36911	0,03208					
	ϕ_0	0,12864	0,0037					
	ϕ_1	0,14001	0,06468					
	ϕ_2	-0,04999	0,06290					
	ϕ_3	0,15523	0,09823					
	ϕ_4	0,19961	0,09763					
	ϕ_5	0,28037	0,10225					
	ϕ_6	0,30677	0,26216					
	τ_0	0,63499	0,13577					
	τ_1	-1,51155	0,19528					
	τ_2	-0,88314	0,17195					
	τ_3	-1,11539	0,44862					
	τ_4	0,27195	0,33669					
	τ_5	0,97718	0,38790					
	τ_6	0,26385	1,54893					
WG sem mistura	β	3,53299	1,07773	2626,695	6,79	<0,0001	5361,54	5286,20
	γ	0,43946	0,05949					
	ϕ_0	0,44591	0,14561					
	ϕ_1	-0,05205	0,07510					
	ϕ_2	-0,08655	0,07578					
	ϕ_3	0,04180	0,12154					
	ϕ_4	0,22467	0,13040					
	ϕ_5	-0,10501	0,16982					
	ϕ_6	0,84988	0,41866					
	τ_0	0,80219	0,13965					
	τ_1	-1,35385	0,18751					
	τ_2	-0,94344	0,16897					
	τ_3	-1,49362	0,43507					
	τ_4	-0,90929	0,35760					
	τ_5	-1,36810	0,48825					
	τ_6	0,74355	1,37447					

Segundo a tabela 4.16, os modelos considerando covariáveis, no enfoque clássico, apresentam melhores resultados segundo os critérios BIC e CAIC.

Tabela 4.16: Resumo da análise sob o enfoque clássico segundo os critérios Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos sem e com covariáveis.

	Modelo	BIC	CAIC
Mistura	sem covariáveis	5375,78	5356,84
	com covariáveis	5368,33	5293,00
Não-mistura	sem covariáveis	5376,00	5357,06
	com covariáveis	5361,54	5286,20

Através da tabela 4.17 observa-se os resultados do modelo WG de fração de cura de mistura e não mistura considerando covariáveis sob o enfoque Bayesiano, também é apresentado os valores dos critérios EBIC e ECAIC para esses modelos.

Tabela 4.17: Resultados das posteriores assumindo o modelo sem e com covariáveis.

Modelo	θ_i	Média	Mediana	DP	IC (95%)		EBIC	ECAIC
					Inferior	Superior		
WG com mistura	β	8,099	7,866	2,187	4,584	12,970	5487,386	5440,387
	γ	0,322	0,319	0,033	0,267	0,394		
	ϕ_0	1,044	1,039	0,195	0,673	1,429		
	ϕ_1	0,045	0,044	0,026	-0,003	0,096		
	ϕ_2	0,004	0,004	0,026	-0,047	0,054		
	ϕ_3	0,031	0,031	0,030	-0,027	0,088		
	ϕ_4	0,019	0,019	0,030	-0,039	0,077		
	ϕ_5	0,033	0,033	0,028	-0,022	0,089		
	ϕ_6	0,016	0,016	0,032	-0,049	0,079		
	τ_0	0,013	0,009	0,013	0,000	0,048		
	τ_1	-0,074	-0,074	0,028	-0,130	-0,021		
	τ_2	-0,054	-0,055	0,029	-0,111	0,007		
	τ_3	-0,015	-0,015	0,032	-0,078	0,047		
	τ_4	-0,015	-0,015	0,031	-0,073	0,047		
	τ_5	-0,023	-0,024	0,032	-0,084	0,041		
	τ_6	-0,003	-0,003	0,032	-0,065	0,057		
WG sem mistura	β	7,488	7,219	2,100	4,146	12,230	5484,881	5437,882
	γ	0,331	0,328	0,037	0,268	0,413		
	ϕ_0	0,904	0,896	0,192	0,545	1,297		
	ϕ_1	0,042	0,042	0,025	-0,009	0,091		
	ϕ_2	0,002	0,002	0,026	-0,050	0,053		
	ϕ_3	0,031	0,031	0,030	-0,028	0,091		
	ϕ_4	0,018	0,018	0,029	-0,039	0,076		
	ϕ_5	0,034	0,034	0,029	-0,023	0,091		
	ϕ_6	0,015	0,015	0,032	-0,050	0,076		
	τ_0	0,014	0,010	0,013	0,000	0,050		
	τ_1	-0,077	-0,078	0,030	-0,134	-0,019		
	τ_2	-0,054	-0,053	0,030	-0,113	0,006		
	τ_3	-0,017	-0,017	0,031	-0,079	0,045		
	τ_4	-0,015	-0,016	0,031	-0,076	0,044		
	τ_5	-0,022	-0,022	0,032	-0,087	0,041		
	τ_6	-0,006	-0,005	0,032	-0,069	0,056		

Sob o enfoque Bayesiano, os critérios EBIC e ECAIC sugerem que os modelos WG de fração de cura de mistura e não-mistura sem considerar as covariáveis apresentam melhores resultados, segundo a tabela 4.18.

Tabela 4.18: Resumo da análise sob o enfoque Bayesiano segundo os critérios EBIC e ECAIC, para os modelos sem e com covariáveis.

	Modelo	EBIC	ECAIC
Mistura	sem covariáveis	5404,04	5392,29
	com covariáveis	5487,39	5440,39
Não-mistura	sem covariáveis	5399,85	5388,10
	com covariáveis	5484,88	5437,88

Considerando o custo computacional e os resultados obtidos a utilização do modelo WG de fração de cura de mistura ou de não-mistura, sem a consideração de covariáveis, podem ser utilizados, sob o enfoque Bayesiano. Já sob o enfoque clássico, há uma diferença significativa entre os resultados dos critérios BIC e CAIC, sendo que a utilização das covariáveis no modelo de fração de cura de mistura ou de não-mistura apresentam melhores resultados.

4.2 Dados de câncer gástrico do Hospital do Câncer de Barretos

Como uma segunda aplicação em dados reais, utilizou-se 201 dados do Hospital do Câncer de Barretos, coletados entre janeiro de 2002 e dezembro de 2007. O câncer gástrico é uma das principais causas de morte relacionada ao câncer [54] e a ressecção da mucosa é aceito como uma opção de tratamento para os casos precoces da doença. A taxa de sobrevida em 5 anos a seguir todos os tipos de ressecções aumentou significativamente, de 20,7% antes 1970 - 28,4% antes de 1990.

Além disso, a taxa de sobrevida em 5 anos seguintes, curados ou ressecção radical, aumentou 37,6 - 55,4% no mesmo período. Assim, as novas tecnologias para otimizar as decisões médicas e para o desenvolvimento de novas terapias são de grande importância para melhorar a sobrevida de câncer gástrico.

A quimioterapia (CRT) é o tratamento padrão para paciente com câncer gástrico nos países ocidentais como tratamento adjuvante aos pacientes submetidos à ressecção curativa. A Tabela 4.19 apresenta as variáveis presentes nos dados citados.

Tabela 4.19: Variáveis presentes nos dados de pacientes com câncer gástrico do Hospital do Câncer de Barretos.

Variáveis	Descrição
t	Tempo a partir da cirurgia em meses.
status	Óbito = 1, censura = 0
trat	CRT = quimioterapia. Surgery alone = somente cirurgia

Como análise exploratória dos dados construiu-se as estimativas de Kaplan-Meier para as variáveis que compõem o banco de dados. Na Figura 4.18 apresentam-se as curvas de sobrevivência estimadas pelo método de Kaplan-Meier para a variável tratamento.

Pela Figura 4.18, no que diz respeito ao tipo de tratamento, observa-se que os pacientes que fazem a quimioterapia têm um tempo de sobrevida maior até o óbito, do que os pacientes que fazem apenas a cirurgia.

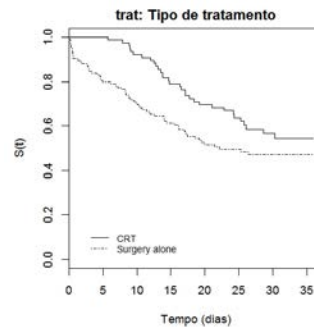


Figura 4.18: Curvas de sobrevivência estimadas pelo método de Kaplan-Meier para as covariáveis dos dados de pacientes com câncer gástrico.

A Tabela 6.20 mostra os valores dos testes *logrank* e a modificação de Peto & Peto do teste de Wilcoxon-Gehan utilizados para testar a igualdade das curvas de sobrevivência obtidas para as covariáveis.

Tabela 4.20: Testes *logrank* e de Wilcoxon-Gehan utilizados para as covariáveis consideradas no estudo de pacientes com câncer.

Covariável	Testes (valor p)	
	<i>logrank</i>	Wilcoxon-Gehan
trat: Tipo de tratamento	4,3 (0,0373)	7,4 (0,007)

Segundo as recomendações em [31], os valores p menores que 0,25 em pelo menos um dos testes de comparação das curvas de sobrevivência devem permanecer no estudo, sendo assim com base na Tabela 4.20, a covariável tratamento “*trat*” deve permanecer no estudo.

4.2.1 Análise semi-paramétrica e Seleção de Covariáveis

Utilizando as mesmas técnicas citadas na seção 4.1.1 segue agora a seleção das covariáveis, adotando-se novamente o nível de significância de referência de 0,10.

Para a determinação de qual distribuição de probabilidade melhor se ajusta aos dados, tomou-se como ponto de partida a distribuição gama generalizada (GG), esta distribuição assume uma variedade de formas, pois tem dois parâmetros de forma, além do parâmetro de escala. Foram utilizadas também as distribuições paramétricas gama, Weibull, log-normal e Weibull generalizada (WG), Weibull Modificada Generalizada (WMG) e seu caso particular Exponencial Exponenciada (EE) de fração de curade mistura e de não-mistura.

Os valores do ajuste do modelo de regressão paramétrico foram obtidos pelos pacotes *Survival* e *flexsurv* do programa R. A Tabela 4.21 apresenta o valor p do teste da razão das verossimilhanças na comparação com a distribuição qui-quadrado de referência com graus de liberdade igual ao número de termos excluídos (diferença entre o número de parâmetros dos dois modelos a serem comparados).

Tabela 4.21: Seleção de covariáveis usando o modelo gama generalizado.

Passos	Modelo	$-2\log L(\theta)$	valor p
Passo 1	Nulo	930,5128	-
	trat	926,0759	0,0352

Portanto há evidências de que o modelo mais adequado é o modelo que contém a covariável *trat*. Sendo assim, essa covariável é importante para descrever o comportamento do tempo até o óbito dos pacientes com câncer gástrico.

Na Figura 4.19, encontram-se os gráficos envolvendo o logaritmo da função de taxa de falha acumulada da covariável *trat*. Pela figura observa-se que as curvas não indicam violação da suposição de taxas de falha proporcionais. A situação mais extrema de violação é caracterizada por curvas que se cruzam.

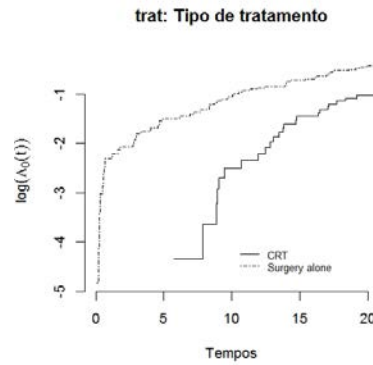


Figura 4.19: $\text{Log}(\widehat{\Lambda}_{0j}(t))$ versus tempo para a covariável *trat*.

A Figura 4.20 apresenta os resíduos padronizados de Schoenfeld associados a covariável *trat* para averiguar a suposição de proporcionalidade, mostrando que há indicação de tendências para a covariável no modelo, o mesmo pode ser confirmado pela Tabela 4.22. Mas analisando novamente a Figura 4.19, situações extremas dessa violação, que são caracterizadas por curvas que se cruzam, não foram observadas nessa covariável. Desse modo podemos concluir não haver evidências de séria violação da suposição de taxas de falhas proporcionais.

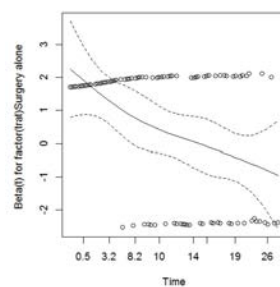


Figura 4.20: Resíduos padronizados de Schoenfeld associados à covariável *trat*.

Tabela 4.22: Testes da proporcionalidade das taxas de falha no modelo ajustado para pacientes com câncer gástrico.

Covariável	ρ	χ^2	valor p
fator(trat)Surgery alone	-0,425	16,2	< 0,001

A Tabela 4.23 apresenta os resultados obtidos do ajuste do modelo de taxas de falha proporcionais de Cox com a covariável selecionada.

Tabela 4.23: Resultados do ajuste do modelo de Cox para os dados do estudo e as razões de taxas de falhas (RTF), para pacientes com câncer gástrico.

Covariável	Estimativa	Erro-Padrão	Valor p	RTF	IC95% (RTF)
fator(trat)Surgery alone	0,450	0,218	0,039	1,568	(1,023; 2,404)

Pelos resultados pode-se interpretar da seguinte forma:

- a taxa de óbitos precoce de pacientes com câncer gástrico que não faz quimioterapia é estimada em 1,6 vezes maior de ocorrer do que os pacientes que fazem a quimioterapia após a cirurgia, com 95% de confiança que essa estimativa está entre 1,023 e 2,404.

4.2.2 Modelos sem fração de cura

Nessa seção é realizado o estudo dos modelos paramétricos Gama Generalizado (GG), Gama, Log-normal (LN), Weibull, Exponencial Exponenciada (EE), Weibull Generalizada (WG) e Weibull Modificada Generalizada (WMG), sem fração de cura sob os enfoques Clássico e Bayesiano. As tabelas com as estimativas da função de sobrevivência são apresentadas no Apêndice A.

4.2.2.1 Estimação Clássica

Para a estimação sob o enfoque clássico utilizando o pacote *maxLik* do *software* R, realizou-se a comparação entre os modelos paramétricos Gama Generalizado (GG), Gama, Log-normal (LN), Weibull, Exponencial Exponenciada (EE), Weibull Generalizada (WG) e Weibull Modificada Generalizada (WMG), esse último o modelo alvo do estudo.

A Tabela 4.24 apresenta os testes da razão de verossimilhanças (TRV) para a identificação do modelo paramétrico com evidências de ser o mais adequado aos dados.

Tabela 4.24: Logaritmo da função de verossimilhança e resultados do Erro Padrão (EP), teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos sem fração de cura.

Modelo	θ_i	Estimativa	EP	$-\log(L(\theta))$	TRV	Valor p	BIC	CAIC
GG	α	0,9391	0,1357	447,3176	2,3647	0,1241	910,54	900,94
	β	54,0835	5,3870					
	γ	0,8635	0,1525					
Gama	α	0,7828	0,0911	447,3268	2,3462	0,3094	905,26	898,86
	β	0,0152	0,0035					
LN	μ	3,5795	0,1814	450,5821	4,1643	0,1247	911,77	905,37
	σ	2,0085	0,1614					
Weibull	γ	0,8207	0,0779	447,3247	2,3505	0,3087	905,26	898,85
	λ	0,0403	0,0108					
EE	α	0,0167	0,0033	447,3369	2,3262	0,3125	905,28	898,88
	β	0,7814	0,0915					
WG	α	0,0301	0,0858	447,3184	2,3630	0,1242	910,55	900,95
	β	0,9146	0,7565					
	γ	0,8795	0,5779					
WMG	α	0,5174	0,1630	448,4999	-	-	918,21	905,43
	β	3,3780	0,9832					
	γ	0,2658	0,0592					
	λ	0,0068	0,0042					

Com base na Tabela 4.24, não há indicação de que algum modelo seja inadequado para a análise dos dados em estudo, tanto pelo TRV quanto pelos critérios BIC e CAIC. A partir da Figura 4.21 é possível observar que todos os modelos estudados parecem ser adequados aos dados, pois suas curvas de sobrevivência estimadas se apresentam próxima da curva de sobrevivência estimada por KM. Sendo assim, pode-se escolher os modelos mais simples, tais como Gama, Weibull e EE para a análise desses dados.

Pela Figura 4.22, que apresenta os gráficos de sobrevivência estimada por Kaplan-Meier, observa-se que todos os modelos apresentam bom ajuste aos dados, como já havia sido observado pela Tabela 4.24. Mas, também pode-se observar que conforme o tempo aumenta a curva de sobrevivência estimada tende a “fugir” da curva de KM, tendendo a zero.

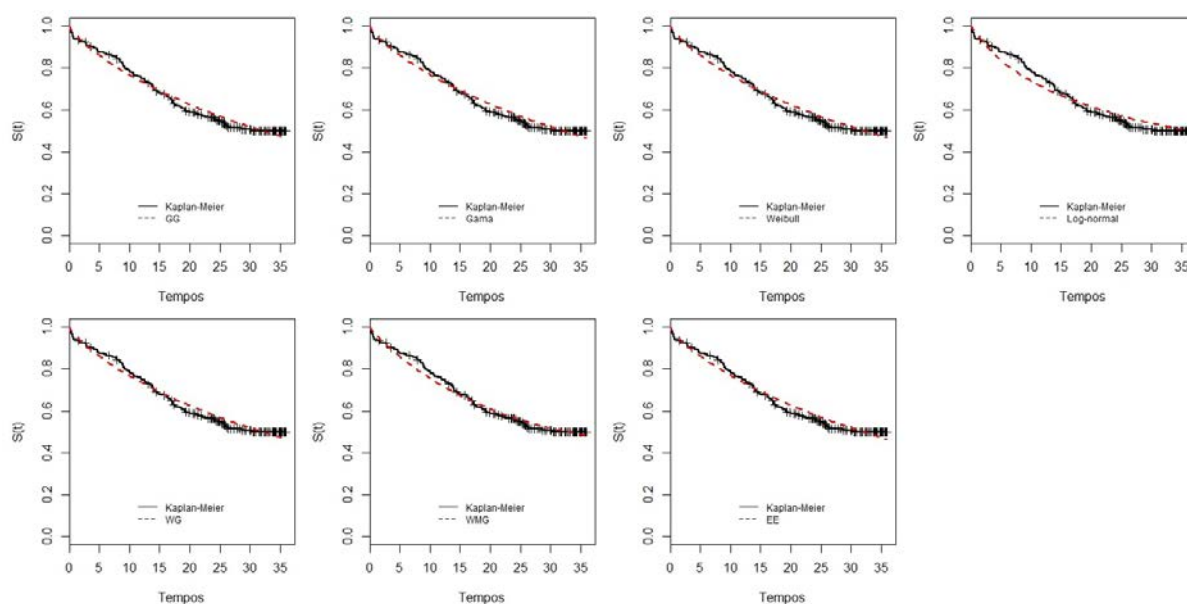


Figura 4.21: Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) *versus* a curva de sobrevivência estimada por Kaplan-Meier, sem fração de cura, sob enfoque clássico.

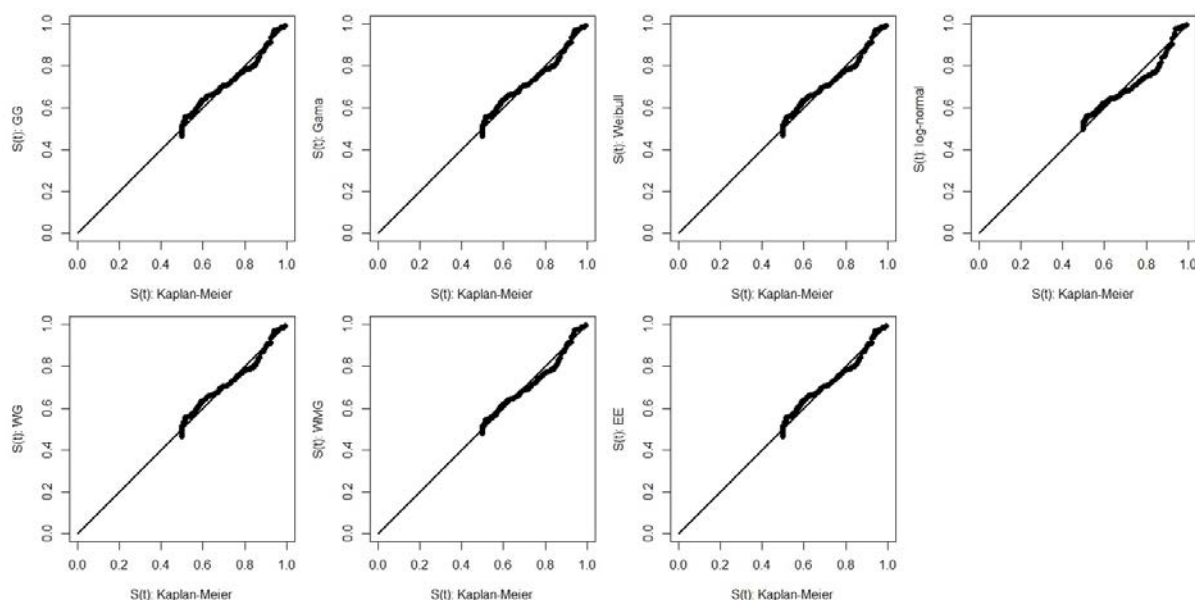


Figura 4.22: Gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), sem fração de cura, sob enfoque clássico.

4.2.2.2 Estimação Bayesiana

Agora utilizando a Inferência Bayesiana reproduziu-se as mesmas análises da Inferência Clássica para os modelos citados na seção anterior. Para a estimação de todos os parâmetros de todos os modelos utilizou-se priores não informativas. A convergência do algoritmo foi verificada a partir de gráficos da densidade a posteriori, autocorrelação e histórico de séries temporais das amostras simuladas dos modelos citados, que são apresentados no Apêndice C. A Tabela 4.25 apresenta os resultados das posteriores. Para a obtenção dos resultados utilizou-se o *software Openbugs*, considerando *prioris* não informativas, que podem ser consultadas no apêndice B.

Tabela 4.25: Resultados das posteriores assumindo o modelo sem fração de cura.

Modelo	θ_i	Média	Mediana	DP	IC (95%)		EBIC	ECAIC
					Inferior	Superior		
GG	α	0,0891	0,0860	0,0221	0,0548	0,1401	1503,34	1493,74
	β	0,0259	0,0260	0,0006	0,0247	0,0270		
	γ	12,3600	12,1000	2,8260	7,7040	18,5000		
Gama	α	1,4990	1,4960	0,1313	1,2530	1,7620	1612,19	1608,26
	β	0,0731	0,0728	0,0075	0,0588	0,0884		
LN	μ	2,6360	2,6370	0,0833	2,4730	2,7980	1721,59	1715,69
	σ	1,1850	1,1830	0,0590	1,0760	1,3080		
Weibull	γ	1,4990	1,4990	0,0936	1,3200	1,6750	1589,52	1585,58
	λ	0,0101	0,0096	0,0033	0,0053	0,0175		
EE	α	0,0172	0,0171	0,0031	0,0114	0,0237	907,08	903,14
	β	0,7898	0,7857	0,0839	0,6392	0,9653		
WG	α	0,1574	0,1196	0,1365	0,0070	0,5103	913,12	907,22
	β	1,6390	1,4840	0,7615	0,6218	3,5340		
	γ	0,6429	0,6041	0,2101	0,3406	1,1830		
WMG	α	0,1946	0,1533	0,1564	0,0109	0,5904	919,41	911,55
	β	1,8060	1,6230	0,8592	0,6463	3,8320		
	γ	0,5310	0,4997	0,1862	0,2519	0,9711		
	λ	0,0066	0,0056	0,0049	0,0003	0,0183		

Esses resultados sugerem que o modelo EE é o mais apropriado para análise destes dados, já que todos os parâmetros são significantes e apresenta o menor valor dos critérios EBIC e ECAIC.

A partir da Figura 4.23 é possível observar que o modelo EE parece ser mais adequado aos dados, pois a curva se apresenta mais próxima da curva de KM, sendo assim esse modelo deve ser utilizado para a estimação dos dados sem fração de cura.

De acordo com a Figura 4.24, que apresenta as curvas de sobrevivência empírica estimada por Kaplan-Meier, observa-se que o modelo EE apresenta o ajuste mais satisfatório.

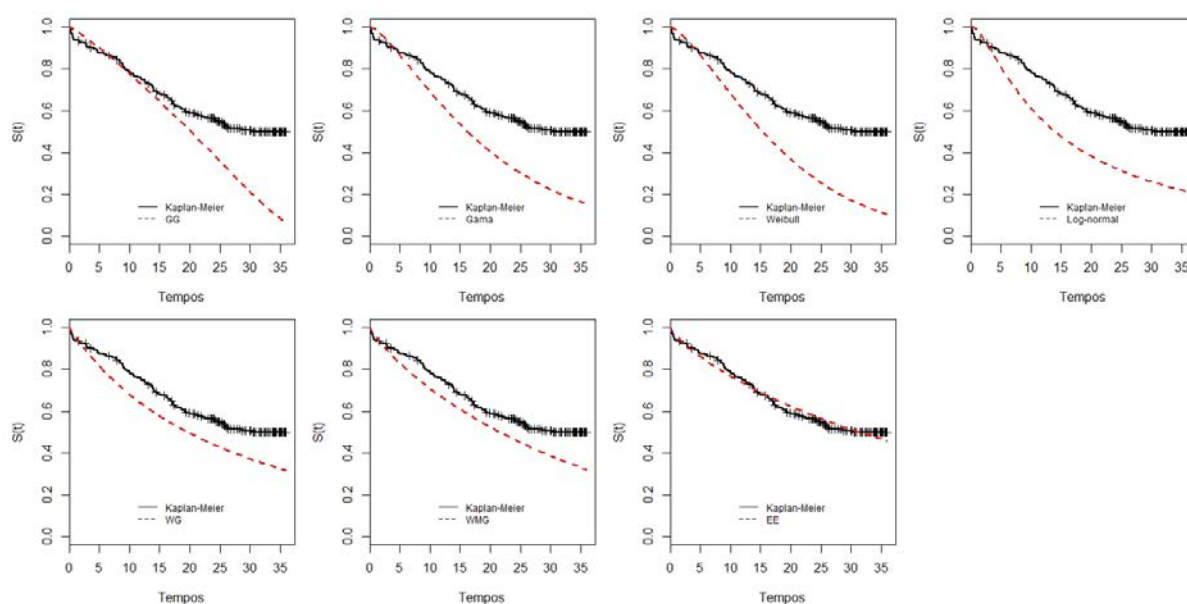


Figura 4.23: Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) *versus* a curva de sobrevivência estimada por Kaplan-Meier, sem fração de cura, sob enfoque Bayesiano.

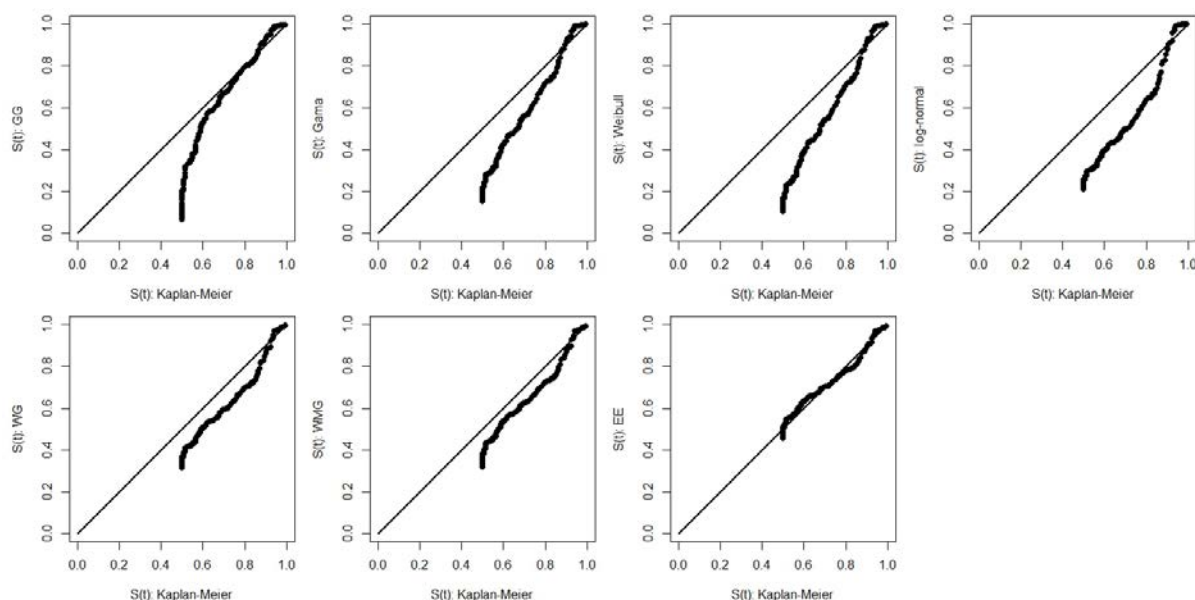


Figura 4.24: Gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), sem fração de cura, sob enfoque Bayesiano.

4.2.3 Modelos de fração de cura de mistura

Aqui é realizado novamente o estudo dos modelos paramétricos Gama Generalizado (GG), Gama, Log-normal (LN), Weibull, Exponencial Exponenciada (EE), Weibull Generalizada (WG) e Weibull Modificada Generalizada (WMG), mas agora considerando a fração de cura de mistura, sob os enfoques Clássico e Bayesiano. As tabelas com as estimativas da função de sobrevivência são apresentadas no Apêndice A.

4.2.3.1 Estimação Clássica

A Tabela 4.26, apresenta os testes da razão de verossimilhanças para a identificação do modelo paramétrico utilizando fração de cura de mistura com evidências de ser o mais adequado aos dados, sendo assim, há indicação da adequação do modelo WG para a análise dos dados em estudo, pelos critérios BIC e CAIC. O teste TRV mostra que o modelo WMG pode ser considerado como mais adequado aos dados.

Tabela 4.26: Logaritmo da função de verossimilhança e resultados do teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos de fração de cura de mistura.

Modelo	θ_i	Estimativa	EP	$-\log(L(\theta))$	TRV	Valor p	BIC	CAIC
GG	α	1,9250	0,6395	447,7243	10,6614	0,0011	916,66	903,88
	β	10,8722	6,8190					
	γ	0,5104	0,1770					
	π	0,0870	0,3816					
Gama	α	0,8891	0,1333	446,9285	9,0696	0,0107	909,77	900,16
	β	0,0336	0,0171					
	π	0,3015	0,1411					
LN	μ	3,5832	0,5136	450,5878	16,3882	0,0003	917,09	907,48
	σ	2,0063	0,2359					
	π	0,0008	0,2172					
Weibull	γ	25,3602	4,9245	447,0938	9,4004	0,0091	910,10	900,50
	λ	0,9250	0,0948					
	π	0,3060	0,0840					
EE	α	0,0357	0,0184	446,9072	9,0270	0,0110	909,72	900,12
	β	0,8676	0,1361					
	π	0,3080	0,1508					
WG	α	0,0656	0,0298	447,5567	10,3260	0,0013	916,33	903,55
	β	1,1632	0,2449					
	γ	0,7996	0,1357					
	π	0,2182	0,1818					
WMG	α	0,4903	0,2323	442,3936	-	-	911,30	895,37
	β	2,9195	1,2709					
	γ	0,2468	0,0867					
	λ	0,0423	0,0114					
	π	0,4977	0,0415					

A partir da Figura 4.25 é possível observar que todos os modelos de fração de cura de mistura apresentaram um ajuste mais próximo da curva de KM, destacando-se o modelo WMG. A Figura 4.26, apresenta os gráficos de sobrevivência estimada por Kaplan-Meier e as curvas de sobrevivências dos modelos estudados, indicando que o modelo WMG apresenta um ajuste um pouco melhor aos dados. Assim, considerando os resultados o modelo WMG é o mais apropriado.

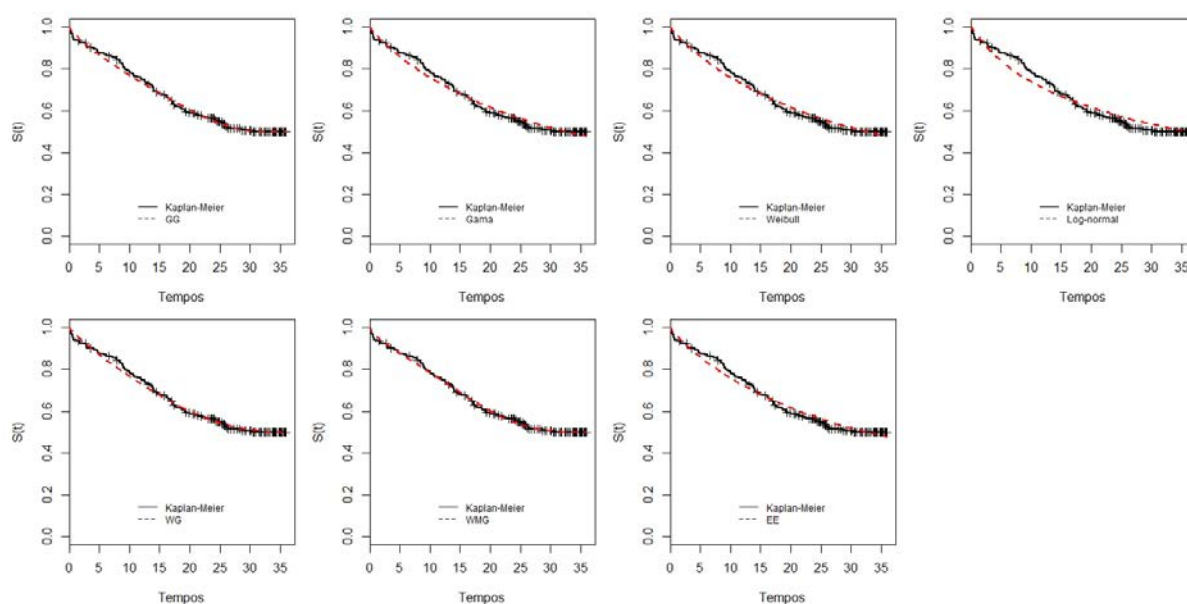


Figura 4.25: Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) *versus* a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de mistura, sob enfoque clássico.

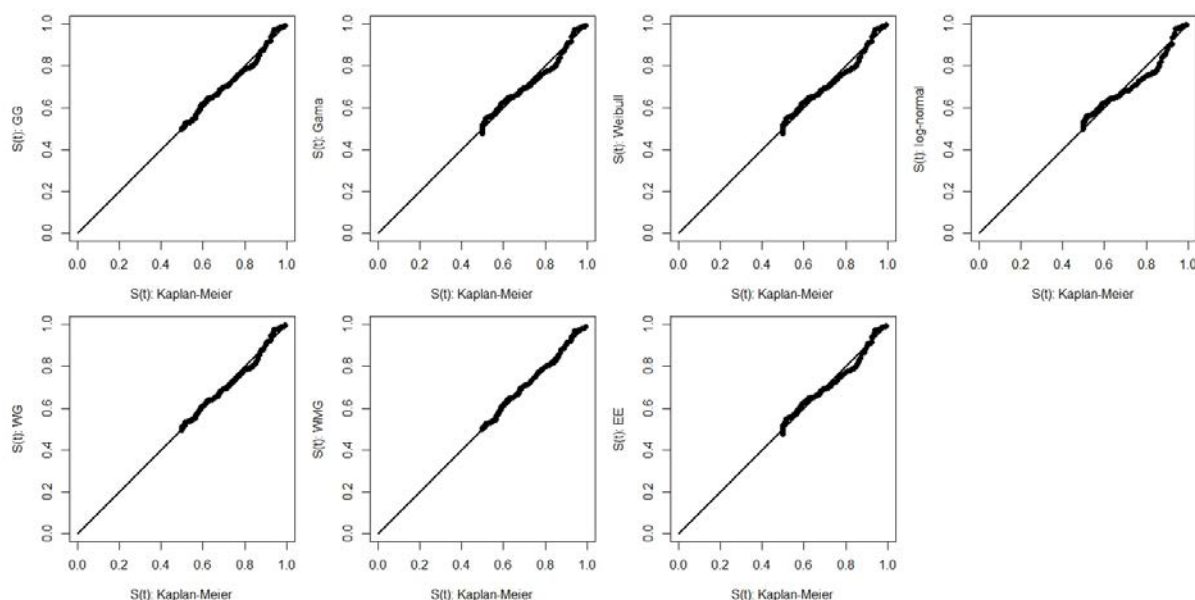


Figura 4.26: Gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de mistura, sob enfoque clássico.

4.2.3.2 Estimação Bayesiana

Nesta seção reproduziu-se as análises da seção anterior agora sob o enfoque Bayesiano para os modelos de fração de cura de mistura citados na seção anterior, utilizando para todos os parâmetros de todos os modelos utilizou-se priores não informativas. A Tabela 4.27, observa-se que resultados sugerem os modelos EE e WMG como os modelos mais apropriados para análise desses dados, já que todos os parâmetros são significantes e os critérios EBIC e ECAIC apresentam o menor valor.

Tabela 4.27: Resultados das posteriores assumindo o modelo de fração de cura de mistura.

Modelo	θ_i	Média	Mediana	DP	IC (95%)		EBIC	ECAIC
					Inferior	Superior		
GG	α	0,0934	0,0902	0,0252	0,0546	0,1545	1371,85	1363,98
	β	0,0259	0,0260	0,0006	0,0248	0,0270		
	γ	11,9300	11,5700	2,9360	7,0470	18,5900		
	π	0,4060	0,4059	0,0456	0,3177	0,4950		
Gama	α	1,5010	1,5000	0,1371	1,2400	1,7720	1471,87	1465,97
	β	0,0732	0,0731	0,0079	0,0584	0,0894		
	π	0,3743	0,3751	0,0527	0,2669	0,4757		
LN	μ	2,6350	2,6360	0,0827	2,4720	2,8020	1573,50	1565,63
	σ	1,1850	1,1830	0,0590	1,0760	1,3070		
	π	0,3461	0,3467	0,0570	0,2320	0,4593		
Weibull	γ	1,5080	1,5060	0,0903	1,3370	1,6890	1451,49	1445,59
	λ	0,0097	0,0094	0,0030	0,0051	0,0165		
	π	0,3842	0,3844	0,0506	0,2829	0,4831		
EE	α	0,0378	0,0358	0,0141	0,0161	0,0691	912,26	906,36
	β	0,8925	0,8851	0,1226	0,6712	1,1450		
	π	0,2875	0,3024	0,1146	0,0449	0,4691		
WG	α	0,0987	0,0653	0,1021	0,0052	0,3924	918,16	910,30
	β	1,2290	1,0930	0,5790	0,5450	2,7480		
	γ	0,8773	0,8279	0,2984	0,4235	1,5940		
	π	0,2478	0,2499	0,1383	0,0144	0,4900		
WMG	α	0,2162	0,1666	0,1750	0,0215	0,6777	914,38	904,55
	β	1,5890	1,3920	0,7833	0,6324	3,6360		
	γ	0,4471	0,4135	0,1917	0,1681	0,8978		
	λ	0,0576	0,0562	0,0174	0,0269	0,0965		
	π	0,4908	0,4916	0,0427	0,4063	0,5688		

A Figura 4.27 mostra que os modelos EE e WMG, parecem ser mais adequados aos dados, pois a curva de sobrevivência estimada pelos modelos se apresentam mais próximas da curva de sobrevivência estimada por KM. O mesmo ocorre na Figura 4.28 pelos gráficos da sobrevivências estimadas por KM *versus* os modelos estudados.

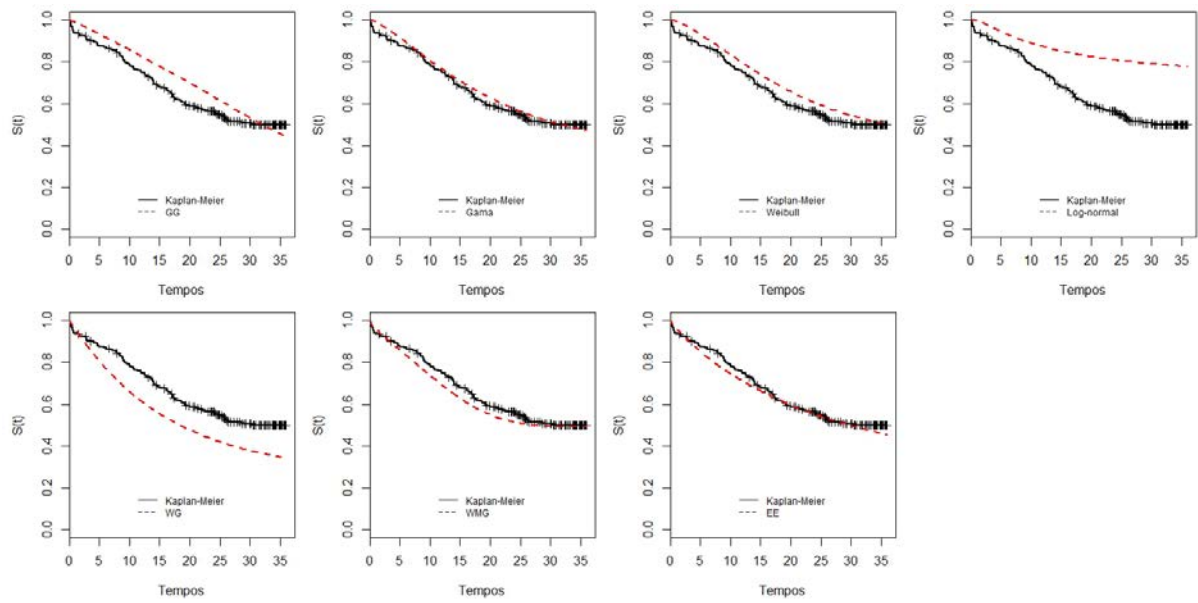


Figura 4.27: Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) *versus* a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de mistura, sob enfoque Bayesiano.

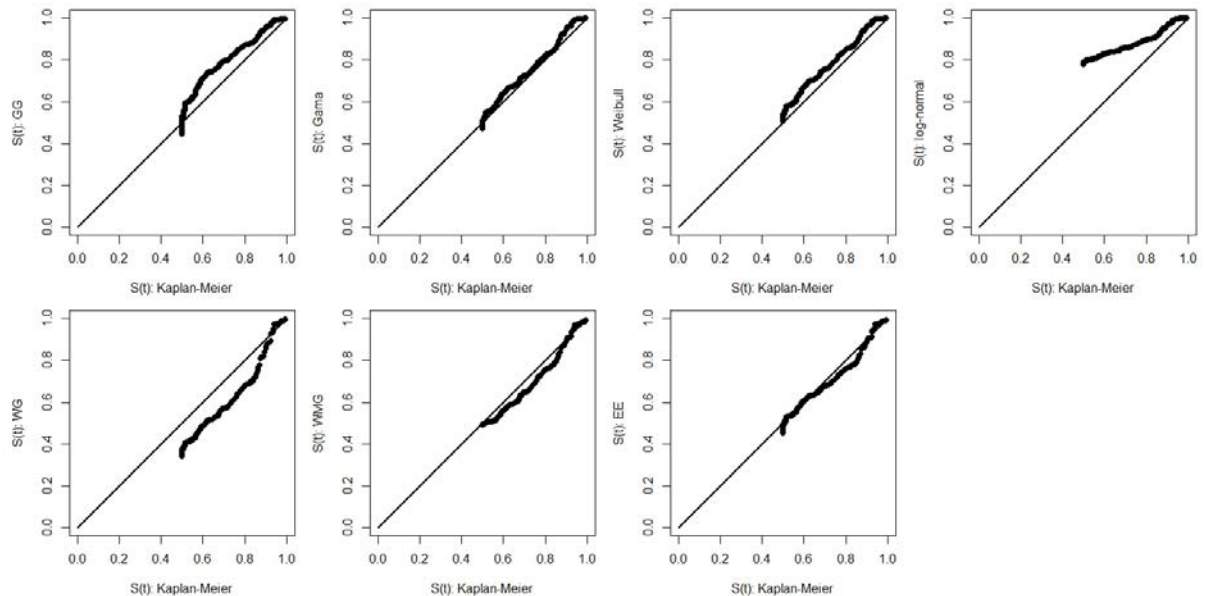


Figura 4.28: Gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de mistura, sob enfoque Bayesiano.

4.2.4 Modelos de fração de cura de não-mistura

Apresenta-se nessa seção os modelos estudados nas seções anteriores, mas dessa vez considerando fração de cura de não-mistura, sob os enfoques Clássico e Bayesiano. As tabelas com as estimativas da função de sobrevivência são apresentadas no Apêndice A.

4.2.4.1 Estimação Clássica

A Tabela 4.28 apresenta os testes da razão de verossimilhanças para a identificação do modelo paramétrico utilizando fração de cura de não-mistura com evidências de ser o mais adequado aos dados. Com base na Tabela 4.28 há indícios da adequação do modelo WMG para a análise dos dados em estudo, pelos critérios BIC e CAIC. O teste TRV mostra que o modelo WMG pode ser considerado como adequado aos dados.

Tabela 4.28: Logaritmo da função de verossimilhança e resultados do teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos de fração de cura de não-mistura.

Modelo	θ_i	Estimativa	EP	$-\log(L(\theta))$	TRV	Valor p	BIC	CAIC
GG	α	3,8091	0,5234	448,1407	12,8847	0,0003	917,49	904,71
	β	2,5469	1,6872					
	γ	0,3021	0,0488					
	π	0,0303	0,0416					
Gama	α	0,8768	0,1410	447,2254	11,0542	0,0040	910,36	900,76
	β	0,0098	0,0201					
	π	0,1243	0,3520					
LN	μ	7,5329	1,0479	449,1506	14,9045	0,0006	914,21	904,61
	σ	2,9863	0,3371					
	π	0,0005	0,0015					
Weibull	γ	75,5510	4,6026	447,3159	11,2353	0,0036	910,54	900,94
	λ	0,9086	0,0765					
	π	0,1551	0,0386					
EE	α	0,0099	0,0211	447,2235	11,0504	0,0040	910,36	900,75
	β	0,8736	0,1423					
	π	0,1172	0,3667					
WG	α	0,0270	0,0206	447,4904	11,5841	0,0007	916,19	903,41
	β	1,0209	0,3089					
	γ	0,8909	0,2889					
	π	0,2235	0,2648					
WMG	α	0,0065	0,1676	441,6983	-	-	909,91	893,98
	β	0,5705	2,9841					
	γ	1,0214	5,9189					
	λ	0,0920	0,1717					
	π	0,4978	0,0381					

A Figura 4.29 apresenta as curvas de sobrevivência estimadas e pode-se observar que todos os modelos de fração de cura de não-mistura apresentaram um ajuste mais próximo da curva de sobrevivência estimada por KM, destacando-se o modelo WMG. Isso sendo confirmado pelos gráficos das estimativas de sobrevivências apresentadas pela Figura 4.30.

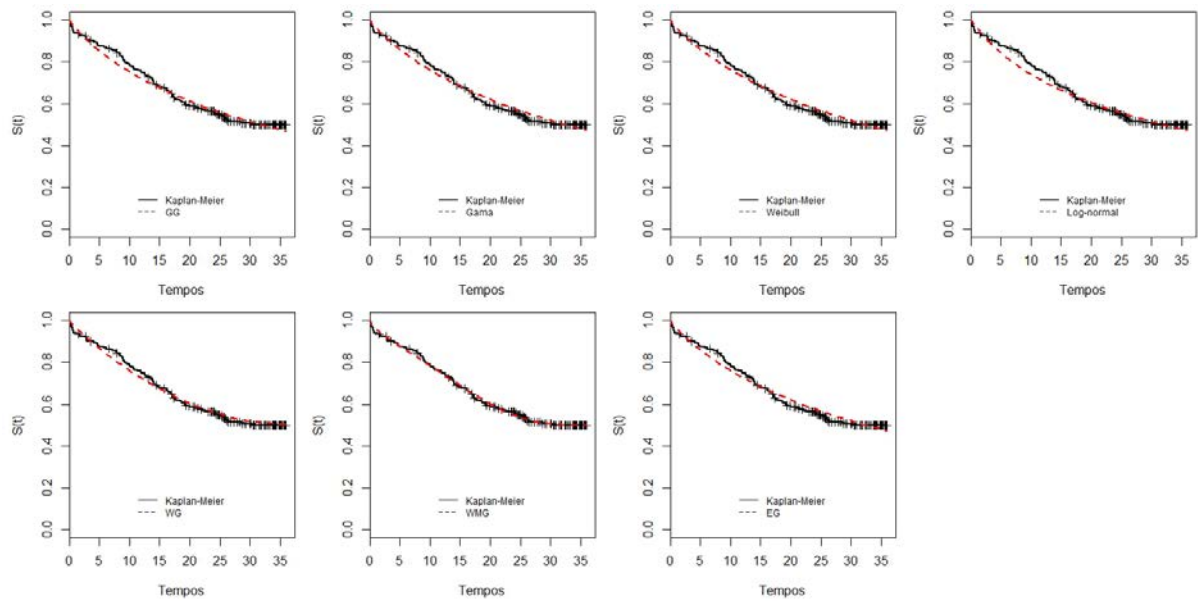


Figura 4.29: Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) *versus* a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de não-mistura, sob enfoque clássico.

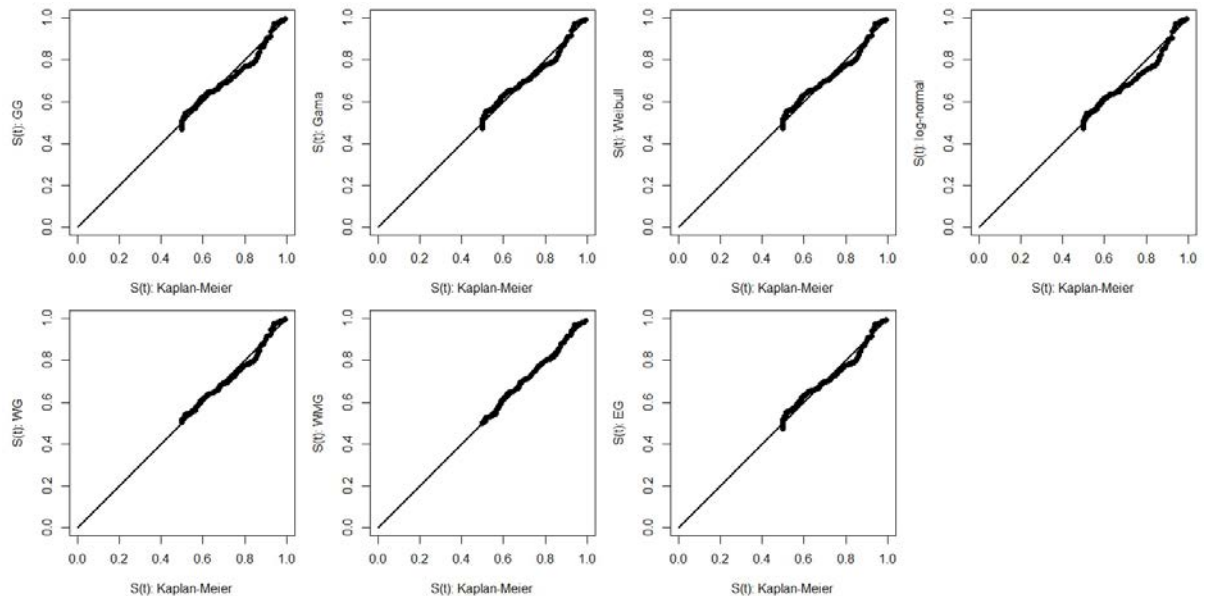


Figura 4.30: Gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de não-mistura, sob enfoque clássico.

4.2.4.2 Estimação Bayesiana

Os resultados dessa seção foram obtidos considerando a fração de cura de não-mistura para os modelos utilizados nas seções anteriores, sob o enfoque Bayesiano, utilizando-se prioris não informativas. Pela Tabela 4.29, os resultados sugerem que os modelos EE, WG e WMG são os mais apropriados para análise destes dados, já que todos os parâmetros são significantes e apresentam o menor valor dos critérios EBIC e ECAIC.

Tabela 4.29: Resultados das posteriores assumindo o modelo de fração de cura de não-mistura.

Modelo	θ_i	Média	Mediana	DP	IC (95%)		EBIC	ECAIC
					Inferior	Superior		
GG	α	0,0913	0,0881	0,0242	0,0537	0,1468	1357,89	1350,02
	β	0,0259	0,0260	0,0006	0,0248	0,0270		
	γ	12,1500	11,8300	2,9480	7,2670	18,7800		
	π	0,4225	0,4221	0,0405	0,3451	0,5056		
Gama	α	1,4950	1,4910	0,1349	1,2370	1,7780	1468,26	1462,36
	β	0,0729	0,0726	0,0078	0,0582	0,0888		
	π	0,4261	0,4255	0,0419	0,3439	0,5086		
LN	μ	2,6360	2,6360	0,0831	2,4720	2,7970	1579,30	1571,43
	σ	1,1830	1,1810	0,0587	1,0740	1,3040		
	π	0,4295	0,4291	0,0426	0,3455	0,5128		
Weibull	γ	1,4990	1,4960	0,0947	1,3160	1,6890	1442,38	1436,48
	λ	0,0101	0,0097	0,0032	0,0051	0,0178		
	π	0,4206	0,4207	0,0417	0,3384	0,5040		
EE	α	0,0222	0,0209	0,0110	0,0051	0,0475	913,01	907,11
	β	0,9354	0,9306	0,1126	0,7217	1,1800		
	π	0,2578	0,2636	0,1088	0,0310	0,4463		
WG	α	0,0902	0,0616	0,0844	0,0066	0,3201	919,20	911,33
	β	1,5190	1,3770	0,6407	0,6755	3,1340		
	γ	0,7383	0,7011	0,2478	0,3590	1,2970		
	π	0,1997	0,1875	0,1221	0,0116	0,4430		
WMG	α	0,1602	0,1133	0,1560	0,0056	0,5550	914,53	898,59
	β	1,5160	1,3140	0,8576	0,5264	3,6300		
	γ	0,4830	0,4348	0,2297	0,1670	1,0480		
	λ	0,0654	0,0632	0,0214	0,0299	0,1179		
	π	0,4902	0,4919	0,0428	0,4045	0,5682		

Na Figura 4.31 é possível observar que os modelos EE e WMG, parecem ser mais adequados aos dados, pois a curva de sobrevivência estimadas pelos modelos em estudo, se apresenta mais próxima da curva de sobrevivência estimada por KM, resultado confirmado pela Figura 4.32.

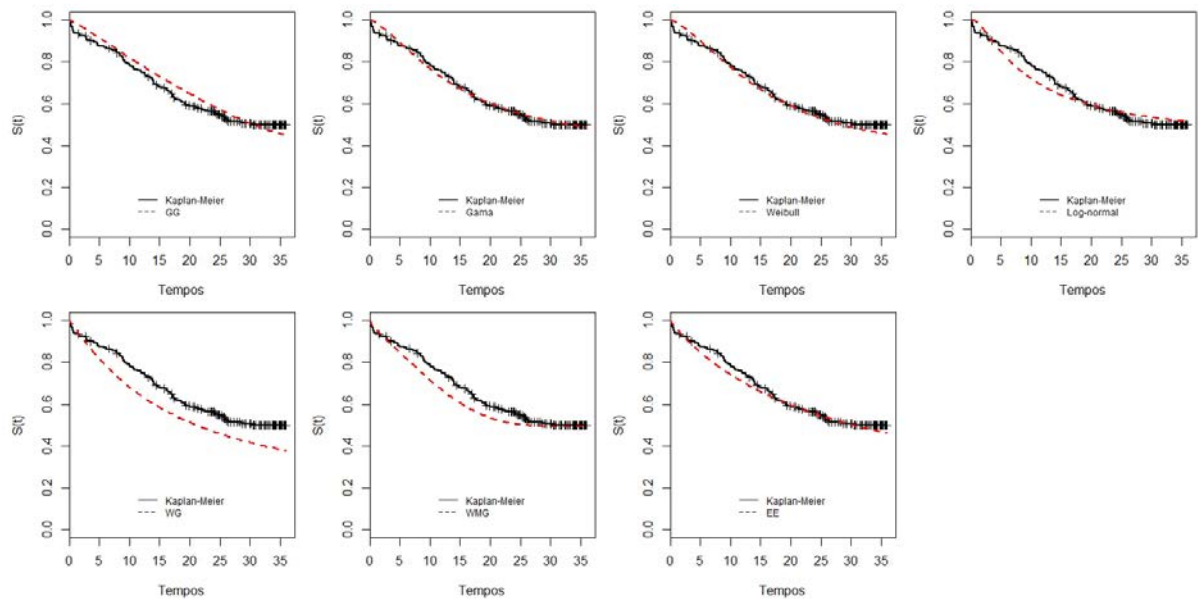


Figura 4.31: Curvas de sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG) *versus* a curva de sobrevivência estimada por Kaplan-Meier, de fração de cura de não-mistura, sob enfoque Bayesiano.

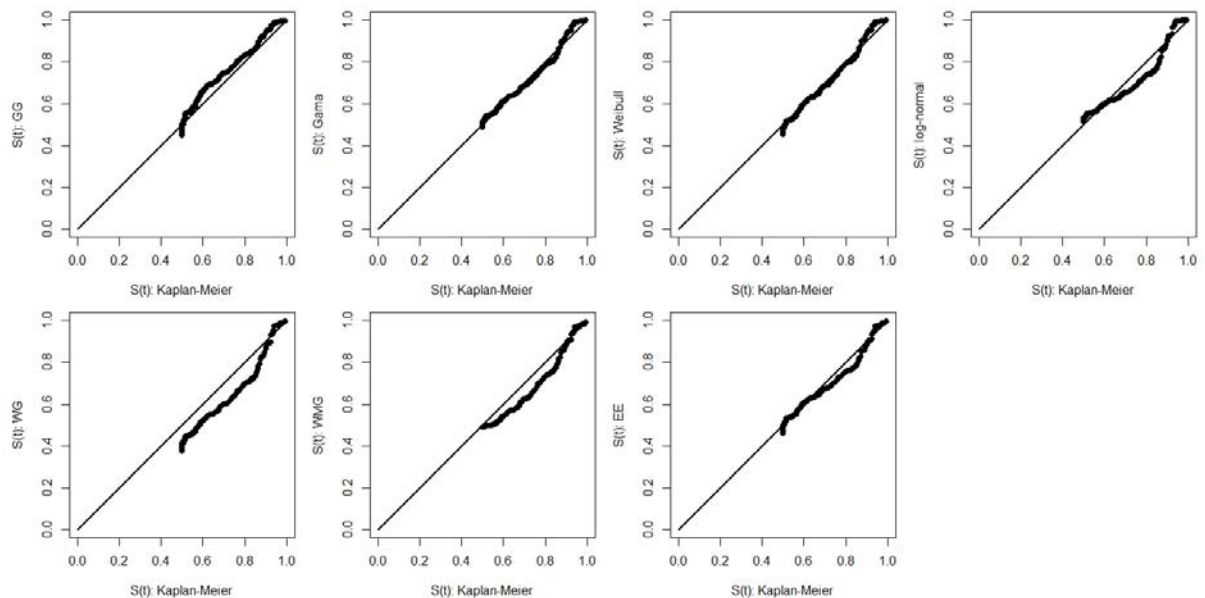


Figura 4.32: Gráficos das sobrevivências estimadas por Kaplan-Meier *versus* as sobrevivências estimadas pelos modelos Gama Generalizado (GG), Gama, Weibull, Log-Normal (LN), Weibull Generalizada (WG) Exponencial Exponenciada (EE) e Weibull Modificada Generalizada (WMG), de fração de cura de não-mistura, sob enfoque Bayesiano.

4.2.5 Resumo da aplicação com Pacientes com Câncer Gástrico do Hospital do Câncer de Barretos

De acordo com os resultados observados nas seções 4.2.2, 4.2.3 e 4.2.4, sob o enfoque clássico as melhores indicações de ajustes aos dados foram para os modelos GG, Gama, Weibul, EE, WG e WMG. Todos os modelos apresentados na Tabela 4.30 têm valores muito próximo, mas ao analisarmos mais criteriosamente as Figuras da seção 4.2.2 sob o enfoque clássico sem fração de cura observa-se que a curva de sobrevivência estimada pelos modelos Gama, Weibull e EE tendem a se afastar da curva de sobrevivência estimada por KM conforme o tempo aumenta, sendo assim consideramos os modelos de fração de cura, de mistura ou de não-mistura, como o melhor ajuste.

Tabela 4.30: Resumo da análise sob o enfoque clássico segundo os critérios Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC).

	Modelo	BIC	CAIC
Sem fração de cura	Gama	905,26	898,86
	Weibull	905,26	898,85
	EE	905,28	898,88
Mistura	WMG	911,30	895,37
Não-mistura	WMG	909,91	893,98

Sendo assim, a Figura 4.33, apresenta a curva de sobrevivência (esquerda) e de risco (direita) estimadas pelo modelo WMG de fração de cura de mistura e de não mistura, sob enfoque clássico. Observando as curvas de sobrevivência estimadas plotadas na Figura 4.33 que o modelo WMG de fração de cura tanto de mistura quanto de não-mistura se aproximam mais da curva de sobrevivência estimada por KM.

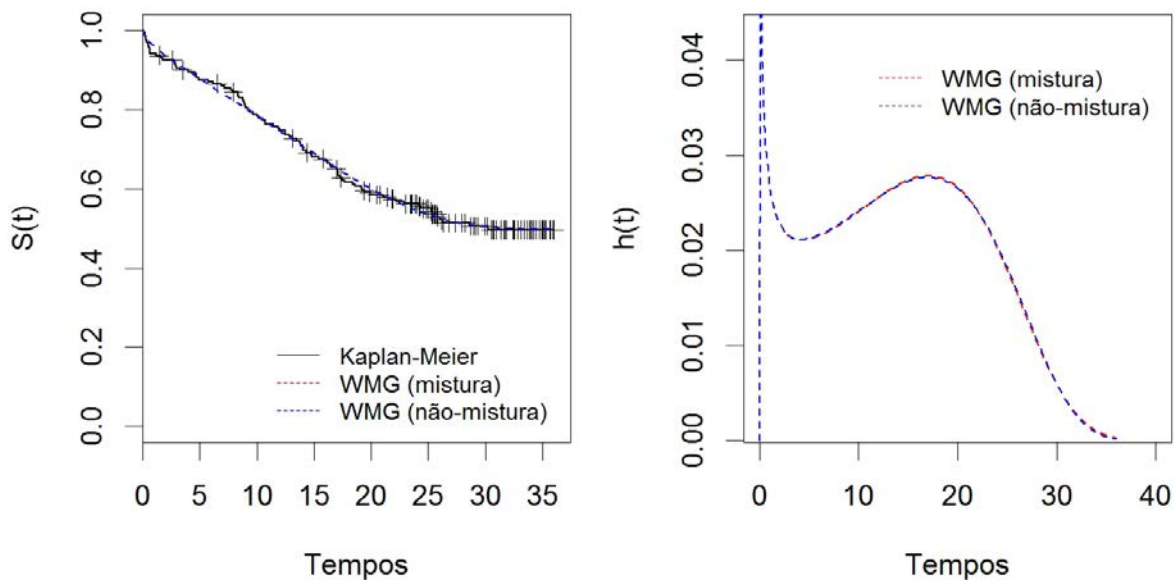


Figura 4.33: Curvas de sobrevivência (esquerda) e de risco (direita) estimadas pelo modelo WMG de fração de cura de mistura e de não-mistura para os dados de pacientes com câncer gástrico do Hospital do Câncer de Barretos, sob o enfoque clássico.

Sob o enfoque Bayesiano, observa-se pela Tabela 4.31, que segundo os critérios EBIC e ECAIC os modelos EE sem fração de cura e WMG de fração de cura de não-mistura, respectivamente, foram os que indicaram melhor ajuste aos dados.

Tabela 4.31: Resumo da análise sob o enfoque Bayesiano segundo os critérios EBIC e ECAIC.

	Modelo	EBIC	ECAIC
Sem fração de cura	EE	907,08	903,14
Mistura	WG	912,26	906,36
	WMG	914,38	904,55
Não-mistura	EE	913,01	907,11
	WMG	914,53	898,59

A Figura 4.34, apresenta a curva de sobrevivência (esquerda) e de risco (direita) estimadas pelos modelos da Tabela 4.31, sob enfoque Bayesiano. Pela Figura 4.34 o modelo WMG de fração de cura de mistura apresenta curva de sobrevivência estimada mais próxima da curva de sobrevivência estimada por KM.

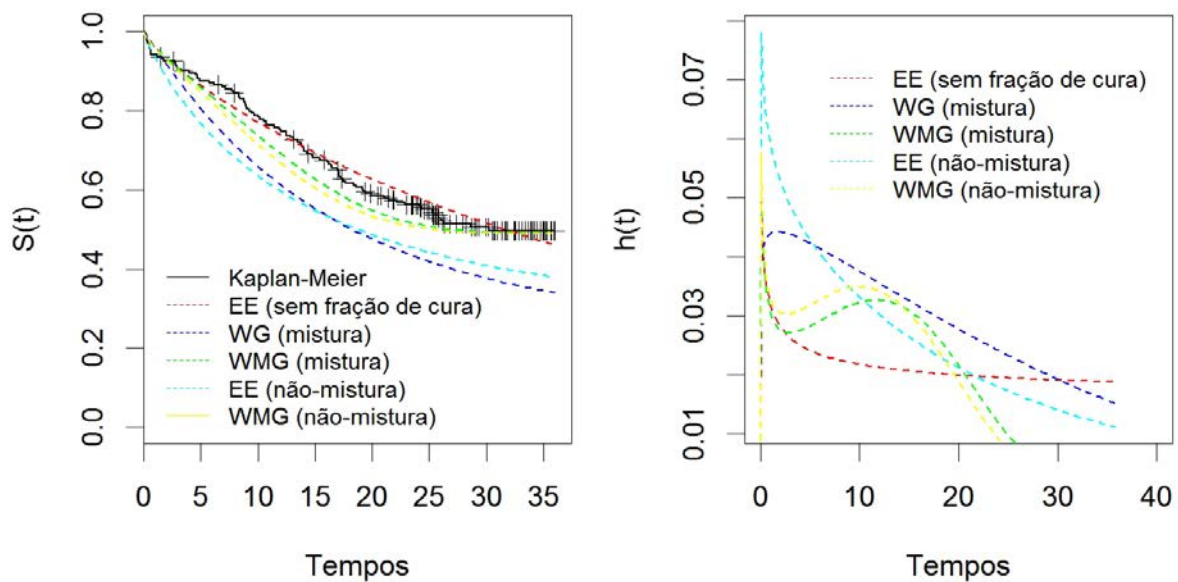


Figura 4.34: Curvas de sobrevivência (esquerda) e de risco (direita) estimadas pelos modelos EE sem fração de cura, WG e WMG de fração de cura de mistura e EE e WMG de fração de cura de não-mistura para os dados de pacientes com câncer gástrico do Hospital do Câncer de Barretos, sob o enfoque Bayesiano.

Consequentemente, o modelo Weibull Modificado Generalizado (WMG) de fração de cura de mistura podem ser utilizados para os dados dessa aplicação sob os enfoques clássico e Bayesiano.

4.2.6 Acrescentando covariáveis ao modelo selecionado para a aplicação com Pacientes com Câncer Gástrico do Hospital do Câncer de Barretos

Como verificado na seção anterior o modelo WG considerando a fração de cura, de mistura ou não mistura, são os mais adequados aos dados da aplicação, agora acrescenta-se a covariável *tratamento* considerada significativa na seção 4.2.1 e verifica-se se há ganho com relação ao modelo sem a covariável.

Na tabela 4.32 é apresentado os resultados do modelo WMG de fração de cura de mistura e não mistura considerando a covariável *tratamento* sob o enfoque clássico, também é apresentado os valores dos critérios TRV, BIC e CAIC para esses modelos.

Tabela 4.32: Logaritmo da função de verossimilhança e resultados do teste da razão de verossimilhanças (TRV), Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos sem e com a covariável *tratamento*.

Modelo	θ_i	Estimativa	EP	$-\log(L(\theta))$	TRV	Valor p	BIC	CAIC
WMG com mistura	β	2,8659	1,3493	430,5190	-	-	903,46	878,20
	γ	0,2309	0,0690					
	ϕ_0	0,1834	0,0799					
	ϕ_1	1,1983	0,3002					
	v_0	0,0838	0,0145					
	v_1	-0,6798	0,6102					
	τ_0	0,1387	0,2457					
	τ_1	-0,3040	0,3317					
WMG sem mistura	β	81,4174	5,0034	428,5261	6,79	<0,0001	899,48	874,21
	γ	0,0650	0,0164					
	ϕ_0	2,3165	0,2033					
	ϕ_1	0,4025	0,0913					
	v_0	0,0291	0,0059					
	v_1	-0,6394	0,3688					
	τ_0	0,1628	0,2594					
	τ_1	-0,3698	0,3456					

Segundo a tabela 4.33, os modelos considerando covariáveis, no enfoque clássico, apresentam melhores resultados segundo os critérios BIC e CAIC. O modelo WMG de fração de cura de não-mistura considerando a covariável *tratamento*, apresentou o valor dos critérios um pouco menor, então há indícios de que esse seria o modelo mais indicado aos dados.

Tabela 4.33: Resumo da análise sob o enfoque clássico segundo os critérios Bayesian Information Criterion (BIC) e Akaike Information Criterion Corrigido (CAIC), para os modelos sem e com a covariável *tratamento*.

	Modelo	BIC	CAIC
Mistura	sem covariáveis	911,30	895,37
	com covariáveis	903,46	878,20
Não-mistura	sem covariáveis	909,91	893,98
	com covariáveis	899,48	874,21

Através da tabela 4.34 observa-se os resultados do modelo WMG de fração de cura de mistura e não mistura considerando a covariável *tratamento*, sob o enfoque Bayesiano, também é apresentado os valores dos critérios EBIC e ECAIC para esses modelos.

Tabela 4.34: Resultados das posteriores assumindo, para os modelos sem e com a covariável *tratamento*.

Modelo	θ_i	Média	Mediana	DP	IC (95%)		EBIC	ECAIC
					Inferior	Superior		
WMG com mistura	β	1,529	1,329	0,768	0,574	3,465	928,80	928,80
	γ	0,464	0,425	0,213	0,172	0,979		
	ϕ_0	0,199	0,154	0,164	0,013	0,626		
	ϕ_1	0,018	0,017	0,033	-0,042	0,085		
	ν_0	0,062	0,061	0,016	0,033	0,097		
	ν_1	0,012	0,012	0,030	-0,046	0,075		
	τ_0	0,099	0,078	0,081	0,004	0,301		
	τ_1	-0,006	-0,006	0,032	-0,068	0,057		
WMG sem mistura	β	0,991	0,787	0,679	0,360	2,826	913,07	903,53
	γ	0,806	0,746	0,407	0,224	1,687		
	ϕ_0	0,073	0,026	0,118	$5,39 \times 10^4$	0,402		
	ϕ_1	0,013	0,013	0,031	-0,050	0,072		
	ν_0	0,077	0,075	0,026	0,034	0,132		
	ν_1	0,010	0,011	0,031	-0,051	0,071		
	τ_0	0,021	$9,09 \times 10^4$	0,048	$3,82 \times 10^8$	0,181		
	τ_1	-0,007	-0,008	0,033	-0,075	0,064		

Sob o enfoque Bayesiano, o critério EBIC sugere que os modelos WMG de fração de cura de não-mistura sem considerar a covariável *tratamento* apresenta melhor resultado, segundo a tabela 4.35.

Tabela 4.35: Resumo da análise sob o enfoque Bayesiano segundo os critérios EBIC e ECAIC, para os modelos sem e com a covariável *tratamento*.

	Modelo	EBIC	ECAIC
Mistura	sem covariáveis	914,38	904,55
	com covariáveis	928,80	928,80
Não-mistura	sem covariáveis	914,53	898,59
	com covariáveis	913,07	903,53

Considerando o custo computacional e os resultados obtidos a utilização do modelo WMG de fração de cura de não-mistura, sem a consideração da covariável *tratamento*, pode ser utilizado, sob o enfoque Bayesiano. Já sob o enfoque clássico há uma diferença significativa entre os resultados dos critérios BIC e CAIC, sendo que a utilização da covariável *tratamento* no modelo de fração de cura de não-mistura apresenta melhor resultado.

Conclusão

Nesse trabalho, buscou-se mostrar a utilidade da inclusão da fração de cura em modelos de análise de sobrevivência sobretudo no modelo Weibull Modificado Generalizado, proposto por [1], que é um modelo flexível, tendo em vista suas possibilidades em relação as curvas de taxa de falhas. Todo o desenvolvimento foi motivado por [54].

Para se alcançar o objetivo realizou-se um estudo de simulação para diferentes tamanhos amostrais considerando a presença de censura e sem a presença de censura para o modelo WMG sem a fração de cura e de fração de cura de mistura e de não-mistura. Os dados simulados consideraram uma fração de cura de 15%, ou seja, da população estimada 15% era imune ou se curaram. Os resultados apresentados pela simulação mostraram que a presença de censura afeta mais as estimativas dos parâmetros π e α , que correspondem à fração de cura e fator de aceleração do tempo de falha, respectivamente, e essa influência aumenta acompanhando o tamanho amostral.

Para os resultados da simulação utilizou-se o pacote MCMCPACK do programa R, com o método SANN, *Simulated Annealing*, pois conforme o tamanho amostral aumenta, consequentemente a quantidade de indivíduos imunes ou curados também aumenta, e isso faz com que os dados simulados tendam a ser bi-modal. Para as simulações o tempo computacional foi muito grande, exigindo a utilização de um computador de alto rendimento, sendo assim utilizou-se o *cluster* disponibilizado pela FCT para que as análises fossem realizadas. Os modelos que melhor se ajustaram aos dados simulados foram os que utilizaram os métodos Bayesianos para a estimação dos parâmetros. A estimação utilizando métodos Bayesianos com priori Gama produziram melhores resultados do que os do método Clássico, para tamanhos amostrais menores.

Nas simulações Bayesianas a utilização de diferentes prioris para cada parâmetro não surtiu o efeito esperado e os melhores resultados foram obtido considerando a priori Gama para todos os parâmetros.

Além da simulação procurou-se fazer a ilustração da utilidade da inclusão do parâmetro de fração de cura através de duas aplicações, ambas de pacientes com câncer, a primeira de pacientes com câncer internados na UTI do INCA e a segunda aplicação em pacientes com câncer gástrico do Hospital do Câncer de Barretos.

Foi realizado um estudo exploratório inicial para a identificação de covariáveis significativas para a estimação através do modelo de COX, esse estudo foi importante para o entendimento de cada aplicação. Mas o modelo de Cox não se ajusta a qualquer situação clínica.

Nas aplicações observa-se que a utilização do parâmetro de fração de cura melhora o ajuste das curvas de sobrevivência estimadas pelos modelos paramétricos, e observa-se ainda que o modelo WMG descreve bem a curva de sobrevivência empírica. Para

a estimação pelos métodos Bayesianos utilizou-se o programa OPENBUGS e para os métodos Clássicos utilizou-se o programa R através do pacote *Survival*.

Os modelos paramétricos acrescentando o parâmetro de fração de cura de mistura ou de não-mistura sem a presença de covariáveis já estima a curva de sobrevivência de forma satisfatória sem a necessidade da inclusão de covariáveis no modelo.

O modelo WMG sem a fração de cura, tende a se afastar da curva estimada por KM conforme o tempo aumenta, pois a função de sobrevivência do modelo tende a zero, já os modelos WMG de fração de cura de mistura e de não-mistura, conseguem captar essa porcentagem de pessoas imunes ou curadas, o que não acontece com o modelo sem fração de cura.

Como perspectiva futura de trabalho, pretende-se realizar um estudo em modelos discretos de fração de cura e também a utilização de outras prioris na simulação de dados com a presença de indivíduos imunes ou curados, além da utilização de covariáveis nos modelos estudados.

Referências

- [1] J.M.F. Carrasco; Ortega, E.M.M.; Cordeiro, G.M. A generalized modified weibull distribution for lifetime modelling. *Computational Statistics and Data Analysis*, 53(2):450–462, 2008.
- [2] M.H. Chen; Ibrahim, J.G.; Sinha, D. A new bayesian model for survival data with a surviving fraction. *Journal of the American Statistical Association*, 94(447):909–919, 1999.
- [3] H.W. Hager; Bain, L.J. Inferential procedures for the generalized gamma distribution. *Journal of the American Statistical Association*, 65:1601–1609, 1970.
- [4] E.L. Kaplan; Meier, P. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53(282):457–481, 1958.
- [5] H. Liang; Zou, G. Improved aic selection strategy for survival analysis. *Computational Statistics & Data Analysis*, 52(5):2538–2548, 1974.
- [6] E.W. Stacy; Mihram, G.A. Parameter estimation for a generalized gamma distribution. *Technometrics*, 7:349–358, 1965.
- [7] A.D. Tsodikov; Ibrahim, J.G.; Yakovlev, A.Y. Estimating cure rates from survival data: an alternative to two-component mixture models. *J. Amer. Statist. Assoc.*, 98:1063–1078, 2003.
- [8] Zellner; A. *Maximal Data Information Prior Distributions*. U. of Chicago Press, 1984.
- [9] A.M. Abreu. Análise de sobrevivência ? modelos de cura. *Boletim da Sociedade Portuguesa de Estatística*, Outubro de 2011.
- [10] J.A. Achcar; Barros, E.A.; Mazucheli, J. . Cure fraction models using mixture and non-mixture models. *Tatra Mountains Mathematical Publications*, 51:1–9, 2012.
- [11] J.A. Achcar; Barros, E.A.; Mazucheli, J. . Block and basu bivariate lifetime distribution in the presence of cure fraction. *Journal of Applied Statistics*, pages 1–11, 2013.
- [12] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19:716–723, 1974.
- [13] T. Ando. Bayesian predictive information criterion for the evaluation of hierarchical bayesian and empirical bayes models. *Biometrika*, 94(2):443–458, 2007. doi:10.1093/biomet/asm017.

- [14] T. Ando. Bayesian model averaging and bayesian predictive information criterion for model selection. *Journal of the Japan Statistical Society*, 38(2):243–257, 2008.
- [15] T. Ando. *Bayesian Model Selection and Statistical Modeling*. CRC Press, Chapter 7, 2010.
- [16] T. Ando. Predictive bayesian model selection. *American Journal of Mathematical and Management Sciences*, 31(1-2):13–38, 2011. doi:10.1080/01966324.2011.10737798.
- [17] C.J.P. Belisle. Convergence theorems for a class of simulated annealing algorithms on r^d . *Journal of Applied Probability*, 29(4):885–895, 1992.
- [18] J.O. Berger; Bernardo, J.M. . On the development of the reference prior method. *Fourth Valencia International Meeting on Bayesian Statistics, Spain*, 1992.
- [19] J. Berkson; R. Gage. Survival curve for cancer patients following treatment. *Journal of the American Statistical Association*, 47(259):501–515, 1952.
- [20] J.M. Bernardo. Reference posterior distributions for bayesian inference. *Journal Royal Statistical Society*, 41(2):113–147, 1979.
- [21] P. Borges. Novos modelos de sobrevivência com fração de cura baseados no processo da carcinogênese. Dissertação de mestrado, Universidade Federal de São Carlos, UFSCar, São Carlos-SP, 2012. 92 f.
- [22] G.E.P. Box; Tiao, G.C. *Bayesian Inference in Statistical Analysis*. Reading: Addison-Wesley, p.588, 1973.
- [23] N. Breslow; Crowley, J. A large sample study of the life table and product limit estimates under random censorship. *Annals of Statistics*, 2:437–453, 1974.
- [24] V.F. Calsavara. Modelos de sobrevivência com fração de cura usando um termo de fragilidade e tempo de vida weibull modificada generalizada. Dissertação de mestrado, Universidade Federal de São Carlos, UFSCar, São Carlos-SP, 2011. 72 f.
- [25] V.G. Cancho; Ortega, E.M.M.; Paula, G.A. On estimation and influence diagnostics for log-birnbaum-saunders student-t regression models: Full bayesian analysis. *Journal of Statistical Planning and Inference*, 140:2486–2496, 2010.
- [26] M.S. Carvalho; Andreozzi, V.L.; Codeço, C.T.; Campos, D.P.; Barbosa, M.T.S.; Shimakura, S. E. *Análise de sobrevivência: teoria e aplicações em saúde*. Rio de Janeiro: Fiocruz, 2011.
- [27] G. Celeux; Forbes, F.; Robert, C.; Titterington, D. Deviance information criteria for missing data models. *Bayesian Analysis*, 1(4):651–706, 2006.
- [28] M.H. Chen; Shao, Q.M.; Ibrahim, J.G. *Monte Carlo Methods in Bayesian Computation*. Springer Series in Statistics, 2000.
- [29] S. Chib; Greenberg, E. Understanding the metropolis-hasting algorithm. *The American Statistician*, 49(4):327–335, 1995.
- [30] D. Collett. *Modelling Survival Data in Medical Research*. 2ed., Chapman and Hall, London, 2003.

- [31] E.A. Colosimo; S.R. Giolo. *Análise de sobrevivência aplicada*. São Paulo: Edgard Blücher, 2006.
- [32] M. DeIorio; Robert, C.P. Discussion of spiegelhalter et al. *The Royal Statistical Society*, Series B(64):629–630, 2002.
- [33] D.M. Diez. Survival analysis in r. *openintro.org*, 2013. Disponível em: [https://www.openintro.org/stat/Survival Analysis in R.pdf](https://www.openintro.org/stat/Survival%20Analysis%20in%20R.pdf).
- [34] Cox; D.R. Regression models and life-tables. *Journal of the Royal Statistical Society*, 34(2):187–220, 1972.
- [35] Cox; D.R. Partial likelihood. *Biometrika*, 62(2):269–276, 1975.
- [36] R.S. Ehlers. *Introdução à Inferência Bayesiana*. 5ed., UFPR., 2007.
- [37] R.S. Ehlers. *Inferência Bayesiana*. São Paulo: USP, 2011.
- [38] S. Faria Júnior. Um ambiente computacional para um teste de significância bayesiano. Dissertação de mestrado, Universidade de São Paulo - Instituto de Matemática e Estatística, São Paulo-SP, 2006.
- [39] A.M.R. Fernandes. Análise de dados em modelos multiestado. Tese de mestrado - 2º ciclo de estudos de mestrado em estatística, Universidade do Minho - Escola de Ciências, Braga-Portugal, Dezembro 2010.
- [40] L.M. Fernandes. Inferência bayesiana em modelos discretos com fração de cura. Dissertação de mestrado, Universidade de Brasília, Brasília-DF, 2013.
- [41] J.M. Ferreira. Análise de sobrevivência: uma visão de risco comportamental na utilização de cartão de crédito. Dissertação (mestrado em biometria e estatística aplicada), Universidade Federal Rural de Pernambuco, Recife-PE, Março 2007.
- [42] A.E. Gelfand; Smith, F.M. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85(410):398–409, 1990.
- [43] R.D. Gupta; Kundu, D. . Generalized exponential distributions. *Aust. N.Z. J. Stat*, 41:173–188, 1999.
- [44] R.D. Gupta; Kundu, D. . Generalized exponential distributions: different methods of estimation. *J. Stat. Comp. Simul.*, 69:315–338, 2001.
- [45] M. Hamada; Wilson, A.G.; Reese, C.S.; Martz, H.F. . *Bayesian Reliability*. Springer, New York, 2008.
- [46] J.G. Ibrahim; Chen, M.H.; Sinha, D.; Campos, D.P.; Barbosa, M.T.S.; Shimakura, S. E. *Bayesian Survival Analysis*. Springer Series in Statistics, 2001.
- [47] E.T. Jaynes. Discussion, in j.m. bernardo, m.h. degroot, d.v. lindley, and a.f.m. smith, editors, bayesian statistics. *Valencia, Spain: University Press*, pages 218–229, 1980.
- [48] H. Jeffreys. *Theory of probability*. 3rd rev. ed., London: Oxford U. Press, 1967.
- [49] J.P. Klein; Moeschberger, M.L. . *Survival Analysis: Techniques for Censored and Truncated Data*. Springer-Verlag, New York, NY, 1997.

- [50] P.C. Lambert. Modeling of the cure fraction in survival studies. *The Stata Journal*, 7(3):351–375, 2007.
- [51] P.C. Lambert; Dickman, P.W.; Weston, C.L.; Thompson, J.R. Estimating the cure fraction in population-based cancer studies by using finite mixture models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 59(1):35–55, 2010.
- [52] J.F. Lawless. *Statistical models and methods for lifetime data*. New York: John Wiley and Sons, 580 p, 1982.
- [53] R.A. Maller; Zhou, X. *Survival analysis with long-term survivors*. John Wiley & Sons Ltd., Chichester, United Kingdom, 1996.
- [54] E.Z. Martinez; J.A. Achcar; A.A.A. Jácome; J.S. Santos. Mixture and non-mixture cure fraction models based on the generalized modified weibull distribution with an application to gastric cancer data. *Computer Methods and Programs in Biomedicine*, 112(3):343–355, 2013.
- [55] H.S. Migon; Gamerman, D. . *Inferência Estatística : Uma abordagem integrada*. Rio de Janeiro, 1999.
- [56] C.A. Niiyama. Análise clássica e bayesiana do modelo weibull modificado generalizado. Dissertação de mestrado, Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, FCT-UNESP, Presidente Prudente-SP, Setembro 2013. 130.
- [57] S.N. Oliveira. Interação genótipo x ambiente para peso vivo e modelagem estatística para seleção genética em tilápias do nilo (*oreochromis niloticus*). Tese de doutorado, Universidade Estadual de Maringá, Maringá-PR, Outubro 2013. 87 p.: il.
- [58] A.T. Paes. Uso de modelos com fração de cura na análise de dados de sobrevivência com omissão nas covariáveis. Tese de doutorado, Universidade de São Paulo, São Paulo-SP, Junho 2007. 129 f.
- [59] M.A.R. Pascoa. Extensões da distribuição gama generalizada: propriedades e aplicações. Tese de doutorado, Escola Superior de Agricultura “Luiz de Queiroz”, Piracicaba-SP, 2012. 143 p.: il.
- [60] C.D.M. Paulino; Turkman, M.A.A.; Murteira, B. . *Estatística bayesiana*. Fundação Calouste Gulbenkian, 2003.
- [61] C.A. Pereira; Stern, J.M. Evidence and credibility: full bayesian significance teste of precise hypothesis. *Entropy*, 1:99–110, 1999.
- [62] T. Pham-Gia. Sample size determination in bayesian statistics-a commentary. *The Statistician*, 44(2):163–166, 1995.
- [63] D.A. Pierce; Stewart, W.H.; Kopecky, K. Distribution free regression analysis of grouped survival data. *Biometrics*, 35(4):785–793, 1979.
- [64] P.L. Ramos. Introdução aos métodos de monte carlo via cadeias de markov e aplicações em inferência bayesiana. Trabalho de estágio não obrigatório do curso de estatística, Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, FCT-UNESP, Presidente Prudente-SP, 2009.

- [65] P.L. Ramos. Aspectos computacionais para inferência na distribuição gama generalizada. Dissertação de mestrado, Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, FCT-UNESP, Presidente Prudente-SP, Fevereiro 2014. 164.
- [66] A.C. Santos. Análise bayesiana de interferência e toxidez em ensaios de diluição serial. Tese de doutorado, Universidade Federal de Lavras, Lavras-MG, 2008. 97 p.: il.
- [67] SAS. The mcmc procedure, sas/statr user's guide, version 9.22. Cary, NC: SAS Institute Inc., pages 4102–4326, 2010.
- [68] G.E. Schwarz. Estimating the dimension of a model. *Annals of Statistics*, 6(2):461–464, 1978.
- [69] T.K.O. Shimizu. Análise bayesiana de dados composicionais na presença de covariáveis. Dissertação de mestrado, Faculdade de Ciências e Tecnologia, Universidade Estadual Paulista, FCT-UNESP, Presidente Prudente-SP, Fevereiro 2014. 112.
- [70] A.C. Sibim. Método para a estimação de quantis de enchentes extremas com o emprego conjunto de análise bayesiana, de informações não sistemáticas e de distribuições limitadas superiormente. Tese de doutorado, Universidade Federal de Minas Gerais, Belo Horizonte-MG, 2009. 202 p.
- [71] A.C. Sibim. Estimação e diagnóstico na distribuição exponencial por partes em análise de sobrevivência com fração de cura. Dissertação de mestrado programa de pós-graduação em ciências de computação e matemática computacional, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos-SP, 2011. 63 p.
- [72] R.A. Silva. Modelos de fração de cura com fatores latentes competitivos e fragilidade. Dissertação de mestrado, Universidade de São Paulo, São Paulo-SP, Junho 2011. 87 f.
- [73] D. Spiegelhalter; Best, N.; Carlin, B.; Van Der Linde, A. Bayesian measures of model complexity and fit. *American Journal of Mathematical and Management Sciences*, 64:583–639, 2002.
- [74] E.W. Stacy. A generalization of the gamma distribution. *Annals of Mathematical Statistics*, 28:1187–1192, 1962.
- [75] M.H. Tarumoto. Um modelo weibull bivariado para riscos competitivos. Tese de doutorado, Universidade Estadual de Campinas, Campinas-SP, 2001. 156 p.
- [76] N. Tsuchiya. Utilização de técnicas bayesianas em modelos de regressão de poisson para dados de contagem longitudinais e dados de contagem com medidas repetidas apresentando excesso de zeros. Dissertação de mestrado, Universidade Federal de São Carlos, São Carlos-SP, Agosto 2008. 129 f.
- [77] Nelson; W. *Applied Life Data Analysis*. John Wiley & Sons, New York, 662 p, 1982.
- [78] A. Wienke; Locatello, I.; Yashin, A.I. The modelling of a cure fraction in bivariate time-to-event data. *Austrian J. Statist*, 35:67–76, 2006.
- [79] A. Yakovlev; Tsodikov, A.D. . *Stochastic Models of Tumor Latency and Their Biostatistical Applications*. World Scientific, Singapore, 1996.

-
- [80] M. Zhou. Nonparametric bayesian dictionary learning and count and mixture modeling. Dissertation submitted in partial fulfillment of the requirements for the degree of doctor of philosophy, Duke University, Durham NC, USA, 2013. 187 p.

Parametrização das distribuições utilizadas

Aqui são apresentadas as funções densidade de probabilidade (f_0), distribuição (F_0), de sobrevivência (S_0) e de risco (h_0) de todas as distribuições utilizadas.

A.1 Gama Generalizada (GG)

Uma variável aleatória T , não negativa ($t > 0$) e contínua, tem distribuição GG com parâmetros $\alpha > 0$, $\beta > 0$ e $\gamma > 0$, se sua função densidade de probabilidade é dada por:

$$f_0(t) = \frac{\gamma}{\beta \Gamma(\alpha)} e^{-\left(\frac{t}{\beta}\right)^\gamma} \left(\frac{t}{\beta}\right)^{\alpha\gamma-1}$$

As funções de distribuição, de sobrevivência e de risco são expressas respectivamente por

$$F_0(t) = 1 - \frac{\Gamma\left(\alpha, \left(\frac{t}{\beta}\right)^\gamma\right)}{\Gamma(\alpha)}$$

$$S_0(t) = \frac{\Gamma\left(\alpha, \left(\frac{t}{\beta}\right)^\gamma\right)}{\Gamma(\alpha)}$$

$$h_0(t) = \frac{\gamma t^{\alpha\gamma-1} e^{-\left(\frac{t}{\beta}\right)^\gamma}}{\beta \Gamma(\alpha)}$$

A.2 Gama

Uma variável aleatória T , não negativa ($t > 0$) e contínua, tem distribuição Gama com parâmetros $\alpha > 0$ e $\beta > 0$, se sua função densidade de probabilidade é dada por:

$$f_0(t) = \frac{1}{\beta \Gamma(\alpha)} e^{-\left(\frac{t}{\beta}\right)} \left(\frac{t}{\beta}\right)^{\alpha-1}$$

As funções de distribuição, de sobrevivência e de risco são expressas respectivamente por

$$F_0(t) = 1 - \frac{\Gamma\left(\alpha, \frac{t}{\beta}\right)}{\Gamma(\alpha)}$$

$$S_0(t) = \frac{\Gamma\left(\alpha, \frac{t}{\beta}\right)}{\Gamma(\alpha)}$$

$$h_0(t) = \frac{t^{\alpha-1} e^{-\frac{t}{\beta}}}{\beta^\alpha \Gamma\left(\alpha, \frac{t}{\beta}\right)}$$

A.3 Log-Normal (LN)

Uma variável aleatória T , não negativa ($t > 0$) e contínua, tem distribuição LN com parâmetros $\mu \in \mathbb{R}$ e $\sigma > 0$, se sua função densidade de probabilidade é dada por:

$$f_0(t) = \frac{1}{\sqrt{2\pi\sigma t}} e^{-\frac{\log(t)-\mu}{2\sigma^2}}$$

As funções de distribuição, de sobrevivência e de risco são expressas respectivamente por

$$F_0(t) = \frac{1}{2} \left(\sqrt{\frac{1}{\sigma^2}} \sigma - \operatorname{erf} \left(\frac{\mu - \log(t)}{\sqrt{2}\sigma} \right) \right)$$

$$S_0(t) = \frac{1}{2} \left(\operatorname{erf} \left(\frac{\mu - \log(t)}{\sqrt{2}\sigma} \right) - \sqrt{\frac{1}{\sigma^2}} \sigma \right) + 1$$

$$h_0(t) = \frac{\sqrt{\frac{2}{\pi}} e^{-\frac{(\mu - \log(t))^2}{2\sigma^2}}}{\sigma t \left(\operatorname{erf}\left(\frac{\mu - \log(t)}{\sqrt{2}\sigma}\right) + 1 \right)}$$

em que, $\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$.

A.4 Weibull

Uma variável aleatória T , não negativa ($t > 0$) e contínua, tem distribuição Weibull com parâmetros $\lambda > 0$ e $\gamma > 0$, se sua função densidade de probabilidade é dada por:

$$f_0(t) = \frac{\lambda e^{-\left(\frac{t}{\gamma}\right)^\lambda} \left(\frac{t}{\gamma}\right)^{\lambda-1}}{\gamma}$$

As funções de distribuição, de sobrevivência e de risco são expressas respectivamente por

$$F_0(t) = 1 - e^{-\left(\frac{t}{\gamma}\right)^\lambda}$$

$$S_0(t) = e^{-\left(\frac{t}{\gamma}\right)^\lambda}$$

$$h_0(t) = \lambda \gamma^{-\lambda} t^{\lambda-1}$$

A.5 Weibull Generalizada (WG)

Uma variável aleatória T , não negativa ($t > 0$) e contínua, tem distribuição WG com parâmetros $\alpha > 0$, $\beta > 0$, $\gamma > 0$, se sua função densidade de probabilidade é dada por:

$$f_0(t) = \alpha \beta \gamma t^{\gamma-1} e^{\alpha(-t^\gamma)} (1 - e^{\alpha(-t^\gamma)})^{\beta-1}$$

As funções de distribuição, de sobrevivência e de risco são expressas respectivamente por

$$F_0(t) = (1 - e^{\alpha(-t^\gamma)})^\beta$$

$$S_0(t) = 1 - (1 - e^{\alpha(-t^\gamma)})^\beta$$

$$h_0(t) = \frac{\alpha\beta\gamma t^{\gamma-1} e^{\alpha(-t^\gamma)} (1 - e^{\alpha(-t^\gamma)})^{\beta-1}}{1 - (1 - e^{\alpha(-t^\gamma)})^\beta}$$

A.6 Exponencial Exponenciada (EE)

Uma variável aleatória T , não negativa ($t > 0$) e contínua, tem distribuição GG com parâmetros $\alpha > 0$ e $\beta > 0$, se sua função densidade de probabilidade é dada por:

$$f_0(t) = \alpha\beta e^{\alpha(-t)} (1 - e^{-\alpha t})^{\beta-1}$$

As funções de distribuição, de sobrevivência e de risco são expressas respectivamente por

$$F_0(t) = (1 - e^{-\alpha t})^\beta$$

$$S_0(t) = 1 - (1 - e^{-\alpha t})^\beta$$

$$h_0(t) = \frac{\alpha\beta e^{-\alpha t} (1 - e^{-\alpha t})^{\beta-1}}{1 - (1 - e^{-\alpha t})^\beta}$$

Programas

A seguir, apresentamos os programas elaborados para a obtenção dos resultados apresentados.

B.1 Simulação

Nessa seção apresenta-se os programas elaborados para a obtenção dos resultados das simulações.

B.1.1 Geração dos dados

```
## FUNÇÃO ÔMEGA DE LAMBERT
ProductLog = function(z,b=0,maxiter=10,eps=.Machine$double.eps,
  min.imag=1e-9) {
  if (any(round(Re(b)) != b))
    stop("branch number for W must be an integer")
  if (!is.complex(z) && any(z<0)) z=as.complex(z)
  ## series expansion about -1/e
  ##
  ## p = (1 - 2*abs(b)).*sqrt(2*e*z + 2);
  ## w = (11/72)*p;
  ## w = (w - 1/3).*p;
  ## w = (w + 1).*p - 1
  ##
  ## first-order version suffices:
  ##
  w = (1 - 2*abs(b))*sqrt(2*exp(1)*z + 2) - 1
  ## asymptotic expansion at 0 and Inf
  ##
  v = log(z + as.numeric(z==0 & b==0)) + 2*pi*b*1i;
  v = v - log(v + as.numeric(v==0))
  ## choose strategy for initial guess
  ##
  c = abs(z + exp(-1));
  c = (c > 1.45 - 1.1*abs(b));
  c = c | (b*Im(z) > 0) | (!Im(z) & (b == 1))
  w = (1 - c)*w + c*v
```

```

## Halley iteration
##
for (n in 1:maxiter) {
  p = exp(w)
  t = w*p - z
  f = (w != -1)
  t = f*t/(p*(w + f) - 0.5*(w + 2.0)*t/(w + f))
  w = w - t
  if (abs(Re(t)) < (2.48*eps)*(1.0 + abs(Re(w)))
      && abs(Im(t)) < (2.48*eps)*(1.0 + abs(Im(w))))
    break
}
if (n==maxiter) warning(paste("iteration limit (",maxiter,
  ") reached, result of W may be inaccurate",sep=""))
if (all(Im(w)<min.imag)) w = as.numeric(w)
return(w)
}

ns<-c(20,50,100,500)          ## TAMANHOS AMOSTRAIS
amostra=1000                  ## REPETIÇÕES DOS TAMANHOS AMOSTRAIS
alpha=0.1
beta=9
gama=0.8
lambda=0.4
p.cura<-0.15                  ## proporção de curados

## GERAÇÃO COM 35% DE CENSURA
p.cens<-0.35                  ## proporção de censura (dos não curados)

for (l in 1:length(ns)) {
  n=ns[l]

  X=NULL
  X<-matrix(nrow=amostra,ncol=n)
  cens<-matrix(nrow=amostra,ncol=n)
  dim(cens)

  for (i in 1:amostra) {
    n.suscept[i]<-n-rbinom(1,n,p.cura) # gerando a quantidade de NÃO-CURADOS
    for (j in 1:n.suscept[i]) {
      u=runif(1,0,1)
      X[i,j]=(gama*ProductLog((alpha*(-(log(1 - (u**(1/beta)))/lambda))
        *(1/gama))/gama))/alpha
    }
    cens[i,1:n.suscept[i]]=rbinom(n.suscept[i],1,1-p.cens)
    if (n>n.suscept[i]) {
      X[i,(n.suscept[i]+1):n]<-rep(max(X[i,],na.rm=T),n-n.suscept[i])
      cens[i,(n.suscept[i]+1):n]<-rep(0,n-n.suscept[i])
    }
  }
}
(x<-X)
dim(X)

```

```

dadosx=as.data.frame(round(X,2), 1000,n)
colnames(dadosx)=1:n
write.table(dadosx,
  paste("C:/Users/EduardoCardosode/Dropbox/Mestrado/Dissertação/Dados e
    Programas/Simulação/dadosWMGx",n,".txt",sep=""),
  sep=";", col.names=TRUE, row.names=F, quote=TRUE, na="NA")

dadosc=as.data.frame(cens, 1000,n)
colnames(dadosc)=1:n
write.table(dadosc,
  paste("C:/Users/EduardoCardosode/Dropbox/Mestrado/Dissertação/Dados e
    Programas/Simulação/dadosWMGc",n,".txt",sep=""),
  sep=";", col.names=TRUE, row.names=F, quote=TRUE, na="NA")
}

## GERAÇÃO COM 0% DE CENSURA
p.cens<-0                                ## proporção de censura (dos não curados)

for (l in 1:length(ns)) {
  n=ns[l]
  X=NULL
  X<-matrix(nrow=amostra,ncol=n)
  cens<-matrix(1,nrow=amostra,ncol=n)
  dim(cens)

  for (i in 1:amostra) {
    n.suscept[i]<-n-rbinom(1,n,p.cura) # gerando a quantidade de NÃO-CURADOS
    for (j in 1:n.suscept[i]) {
      u=runif(1,0,1)
      f = function(t) {
        log(1-u^(1/beta))+(1/lambda)*(t^gama)*exp(alpha*t)
      }
      X[i,j] = uniroot(f=f,interval=c(0,150))$root
      #X[i,j] = (gama*ProductLog((alpha*(-(log(1 - (u**(1/beta)))/lambda
        ))**(1/gama))/gama))/alpha
    }
    cens[i,1:n.suscept[i]]=rbinom(n.suscept[i],1,1-p.cens)
    if (n>n.suscept[i]) {
      X[i,(n.suscept[i]+1):n]<-rep(max(X[i,],na.rm=T),n-n.suscept[i])
      cens[i,(n.suscept[i]+1):n]<-rep(0,n-n.suscept[i])
    }
  }
  (x<-X)
  dim(X)

  dadosx=as.data.frame(round(X,2), 1000,n)
  colnames(dadosx)=1:n
  write.table(dadosx,
    paste("C:/Users/EduardoCardosode/Dropbox/Mestrado/Dissertação/Dados e
      Programas/Simulação/dados2WMG0x",n,".txt",sep=""),
    sep=";", col.names=TRUE, row.names=F, quote=TRUE, na="NA")

  dadosc=as.data.frame(cens, 1000,n)

```

```
colnames(dadosc)=1:n
write.table(dadosc,
  paste("C:/Users/EduardoCardosode/Dropbox/Mestrado/Dissertação/Dados e
    Programas/Simulação/dados2WMG0c",n,".txt",sep=""),
  sep=";", col.names=TRUE, row.names=F, quote=TRUE, na="NA")
}
}
```

B.1.2 Estimações

```
set.seed(1234567890)
#setwd('C:/Users/EduardoCardosode/Dropbox/Mestrado/Dissertação/Dados e Programas
  /Simulação/MCMCPACK')
rm(list = ls())

library("coda")
library(MCMCpack)
library(msm)
library(maxLik)
require(TeachingDemos)
library(xtable)

print("Weibull Modificada Generalizada")
## dist = WMG para dados com 35% de censura, WMG0 para dados com 0% de censura
dist='WMG'

## tp = CM para mistura, SM para não-mistura, SFC para sem fração de cura
tp='CM'

VERO <- function(p,tempo,censura) {
  a<-p[1]
  b<-p[2]
  g<-p[3]
  l<-p[4]
  f<-p[5]
  f0 = 1*b*(tempo**(g-1))*(g+(a*tempo))*exp((a*tempo)-l*(tempo**g)*exp(a*
    tempo))*((1-exp(-l*(tempo**g)*exp(a*tempo)))*(b-1))
  S0 = 1-(1-exp(-l*(tempo**g)*exp(a*tempo)))*b
  return (
    sum( censura*log(1-f) )
    + sum( censura * log( f0 ) )
    + sum( (1-censura) * log( f + ( (1-f) * S0 ) ) )
  )
}

n<-20      ## n = 20, 50, 100 ou 500
tune.U=0.2  # ajuste para priori Uniforme
#tune.G=0.3 # ajuste para priori Gama
#tune.H=0.25 # ajuste para priori Híbrida

# verdadeiros valores dos parâmetros
amostra=1000
alpha=0.1
beta=9
```

```

gama=0.8
lambda=0.4
cura=0.15 ## cura = pi

alfa.verd=alpha;
beta.verd=beta;
gama.verd=gama;
lambda.verd=lambda;
cura.verd=cura;

alfa.hat <- double(amostra)
beta.hat <- double(amostra)
gama.hat <- double(amostra)
lambda.hat <- double(amostra)
cura.hat <- double(amostra)
sd.alfa.hat <- double(amostra)
sd.beta.hat <- double(amostra)
sd.gama.hat <- double(amostra)
sd.lambda.hat <- double(amostra)
sd.cura.hat <- double(amostra)
ll.hat=rep(NA,times=amostra)
ICONFai <- double(amostra)
ICONFas <- double(amostra)
ICONFbi <- double(amostra)
ICONFbs <- double(amostra)
ICONFgi <- double(amostra)
ICONFgs <- double(amostra)
ICONfli <- double(amostra)
ICONFls <- double(amostra)
ICONFci <- double(amostra)
ICONFcs <- double(amostra)
prob_cob_a_mv <- rep(1,times=amostra)
prob_cob_b_mv <- rep(1,times=amostra)
prob_cob_g_mv <- rep(1,times=amostra)
prob_cob_l_mv <- rep(1,times=amostra)
prob_cob_c_mv <- rep(1,times=amostra)

X<-as.matrix(read.table(paste("dados",dist,"x",n,".txt",sep=""), header=TRUE,
  fill=TRUE,sep=';'))
x<-X;dim(X)
cens<-as.matrix(read.table( paste("dados",dist,"c",n,".txt",sep=""), header=TRUE
  , fill=TRUE,sep=';'))

#=====
# Estimação clássica - parâmetros -
#=====
#-----
loglik <- function(p,tempo,censura) {
  a<-p[1]
  b<-p[2]
  g<-p[3]
  l<-p[4]
  f<-p[5]

```

```

    if ( (a>0) && (b>0) && (g>0) && (l>0) && (f>0) && (f<1) ) {
        return ((-1)*(VERO(p,tempo,censura)))
    } else return (-Inf)
}

for (v in 1:amostra) {
  options(show.error.messages = FALSE)
  ml = try(optim(c(0.05,5,0.6,0.3,0.1), loglik, tempo=x[v,],
    censura=cens[v,], method="SANN", hessian=T), TRUE)
  options(show.error.messages = TRUE, warn=10)
  if (is.list(ml[1])) {
    fit=ml$par;
    sd=try(sqrt(abs(diag(solve(ml$hessian)))), TRUE)
    alfa.hat[v]=fit[1];
    beta.hat[v]=fit[2];
    gama.hat[v]=fit[3];
    lambda.hat[v]=fit[4]
    cura.hat[v]=fit[5]
    sd.alfa.hat[v]=sd[1];
    sd.beta.hat[v]=sd[2];
    sd.gama.hat[v]=sd[3];
    sd.lambda.hat[v]=sd[4]
    sd.cura.hat[v]=sd[5]
    ll.hat[v] = -ml$value

    #Estatísticas de Interesse - Clássico
    a1=alfa.hat[v]; a2=sd.alfa.hat[v];
    b1=beta.hat[v]; b2=sd.beta.hat[v];
    g1=gama.hat[v]; g2=sd.gama.hat[v];
    l1=lambda.hat[v]; l2=sd.lambda.hat[v];
    c1=cura.hat[v]; c2=sd.cura.hat[v];

    ICONFai[v]=a1-1.96*a2; ICONFas[v]=a1+1.96*a2
    ICONFbi[v]=b1-1.96*b2; ICONFbs[v]=b1+1.96*b2
    ICONFgi[v]=g1-1.96*g2; ICONFgs[v]=g1+1.96*g2
    ICONFli[v]=l1-1.96*l2; ICONFls[v]=l1+1.96*l2
    ICONFci[v]=c1-1.96*c2; ICONFcs[v]=c1+1.96*c2

    if(is.na(ICONFai[v])|is.na(ICONFas[v])) prob_cob_a_mv[v]=0 else if
      (alpha <ICONFai[v]|alpha >ICONFas[v]) prob_cob_a_mv[v]=0 else
      prob_cob_a_mv[v]=1
    if(is.na(ICONFbi[v])|is.na(ICONFbs[v])) prob_cob_b_mv[v]=0 else if
      (beta <ICONFbi[v]|beta >ICONFbs[v]) prob_cob_b_mv[v]=0 else
      prob_cob_b_mv[v]=1
    if(is.na(ICONFgi[v])|is.na(ICONFgs[v])) prob_cob_g_mv[v]=0 else if
      (gama <ICONFgi[v]|gama >ICONFgs[v]) prob_cob_g_mv[v]=0 else
      prob_cob_g_mv[v]=1
    if(is.na(ICONFli[v])|is.na(ICONFls[v])) prob_cob_l_mv[v]=0 else if
      (lambda<ICONFli[v]|lambda>ICONFls[v]) prob_cob_l_mv[v]=0 else
      prob_cob_l_mv[v]=1
    if(is.na(ICONFci[v])|is.na(ICONFcs[v])) prob_cob_c_mv[v]=0 else if
      (cura <ICONFci[v]|cura >ICONFcs[v]) prob_cob_c_mv[v]=0 else
      prob_cob_c_mv[v]=1
  }
}

```

```

        guarda=cbind(a1,a2,ICONFai[v], ICONFas[v],
                      b1, b2,ICONFbi[v],ICONFbs[v],
                      g1, g2,ICONFgi[v],ICONFgs[v],
                      l1, l2,ICONFl1i[v],ICONFl1s[v],
                      c1, c2,ICONFci[v],ICONFcs[v],
                      prob_cob_a_mv, prob_cob_b_mv, prob_cob_g_mv,
                      prob_cob_l_mv, prob_cob_c_mv)
    }
}
(sah=summary(alfa.hat))
(sbh=summary(beta.hat))
(sgh=summary(gama.hat))
(slh=summary(lambda.hat))
(sch=summary(cura.hat))

nahamostra = amostra
nbhamostra = amostra
nghamostra = amostra
nlhamostra = amostra
nchamostra = amostra

if (!is.na(sah["NA's"])) { nahamostra = amostra-sah["NA's"]}
if (!is.na(sbh["NA's"])) { nbhamostra = amostra-sbh["NA's"]}
if (!is.na(sgh["NA's"])) { nghamostra = amostra-sgh["NA's"]}
if (!is.na(slh["NA's"])) { nlhamostra = amostra-slh["NA's"]}
if (!is.na(sch["NA's"])) { nchamostra = amostra-sch["NA's"]}

(ll.hat)
(pars.mu.hat=c(mean(alfa.hat,na.rm=T),mean(beta.hat,na.rm=T),mean(gama.hat,na.rm=T),mean(lambda.hat,na.rm=T),mean(cura.hat,na.rm=T)))
(pars.sd.hat=c(mean(sd.alfa.hat,na.rm=T),mean(sd.beta.hat,na.rm=T),mean(sd.gama.hat,na.rm=T),mean(sd.lambda.hat,na.rm=T),mean(sd.cura.hat,na.rm=T)))
(pars.ll.hat=mean(ll.hat,na.rm=T))
(EQM.hat=c(mean((alfa.hat - alfa.verd)^2,na.rm=T),mean((beta.hat - beta.verd)^2,na.rm=T),mean((gama.hat - gama.verd)^2,na.rm=T),mean((lambda.hat - lambda.verd)^2,na.rm=T),mean((cura.hat - cura.verd)^2,na.rm=T)))
(Efic.hat=c(NA,NA,NA,NA,NA))
(prob.hat=c(sum(prob_cob_a_mv,na.rm=T)/(nahamostra),sum(prob_cob_b_mv,na.rm=T)/(nbhamostra),sum(prob_cob_g_mv,na.rm=T)/(nghamostra),sum(prob_cob_l_mv,na.rm=T)/(nlhamostra),sum(prob_cob_c_mv,na.rm=T)/(nchamostra)))

#####
# Posteriori UNIFORME - U
#####
posteriori<-function(p,tempo,censura){
a<-p[1]
b<-p[2]
g<-p[3]
l<-p[4]
f<-p[5]
if ( (a>0) && (b>0) && (g>0) && (l>0) && (f>0) && (f<1) ) {

```

```

        return ((1)*(VERO(p,tempo,censura)
            + log(1)
            + log(1)
            + log(1)
            + log(1)
            + log(1)
            )
    )
} else return (-Inf)
}

#=====
# Posteriori GAMMA - G
#=====
#posteriori<-function(p,tempo,censura){
#a<-p[1]
#b<-p[2]
#g<-p[3]
#l<-p[4]
#f<-p[5]
#if ( (a>0) && (b>0) && (g>0) && (l>0) && (f>0) && (f<1) ) {
#    f0 = 1*b*(tempo**(g-1))*(g+(a*tempo))*exp((a*tempo)-1*(tempo**g)*exp(a*
#        tempo))*((1-exp(-1*(tempo**g)*exp(a*tempo)))*(b-1))
#    S0 = 1-(1-exp(-1*(tempo**g)*exp(a*tempo)))*b
#    return ((1)*(VERO(p,tempo,censura)
#        + dgamma(a,1, .1,log=T)
#        + dgamma(b,1, .1,log=T)
#        + dgamma(g,1, .1,log=T)
#        + dgamma(l,1, .1,log=T)
#        + dgamma(f,1, .1,log=T)
#    )
#    )
#} else return (-Inf)
#}

#=====
# Posteriori HÍBRIDA - H
#=====
#posteriori<-function(p,tempo,censura){
#a<-p[1]
#b<-p[2]
#g<-p[3]
#l<-p[4]
#f<-p[5]
#if ( (a>0) && (b>0) && (g>0) && (l>0) && (f>0) && (f<1) ) {
#    return ((1)*(VERO(p,tempo,censura)
#        + dgamma(a,1, 0.01,log=T)
#        + dgamma(b,0.01, 0.01,log=T)
#        + dgamma(g,1, 0.01,log=T)
#        + dgamma(l,1, 0.01,log=T)
#        + dbeta(f,1,1,log=T)
#    )
#    )
#}

```

```

#} else return (-Inf)
#}

#=====
# Estimação Bayesiana
#=====
mediasMDL=matrix(NA,amostra,5)
dpMDL=matrix(NA,amostra,5)
NaiveSEMDL=matrix(NA,amostra,5)
TimeSEMDL=matrix(NA,amostra,5)
ICsupMDL=matrix(NA,amostra,5)
ICinfMDL=matrix(NA,amostra,5)
llMDL=rep(NA,times=amostra)
prob_cob_bayes_MDL <- matrix(NA,amostra,5)
erro=0

for (v in 1:amostra) {
  theta=0
  contador=0
  tempo=x[v,]
  censura=cens[v,]
  cat("UNIFORME - ",v,"\n"); # UNIFORME, GAMA OU HIBRIDA, para monitorar se
    houve algum erro
  if (v==1) { sink(paste('saida','_',dist,'_',tp,'_',n,'.R',sep=""), append
    =F) } else { sink(paste('saida','_',dist,'_',tp,'_',n,'.R',sep=""),
    append=T) }
  options(show.error.messages = FALSE)
  while (!is.matrix(theta) & contador < 100) {
    theta<-try(MCMCmetrop1R(posteriori, c(0.05,5,0.6,0.3,0.1),
      verbose=0, thin=10, burnin=50000,
      mcmc=100000,
      tempo=tempo, censura=censura,
      tune=tune.U, ### Depende da
      posteriori utilizada
      force.samp=T, optim.method="SANN"))
    contador=contador+1
  }
  options(show.error.messages = TRUE)
  sink()
  if (is.matrix(theta)) {
    #(params[v,]=colMeans(theta))
    res=summary(theta)

    mediasMDL[v,]=(res$statistics[,1])
    dpMDL[v,]=t(res$statistics[,2])
    NaiveSEMDL[v,]=t(res$statistics[,3])
    TimeSEMDL[v,]=t(res$statistics[,4])
    ICinfMDL[v,]=t(res$quantiles[,1])
    ICsupMDL[v,]=t(res$quantiles[,5])
    llMDL[v]=posteriori(mediasMDL[v,],tempo,censura)

    if(alfa.verd <ICinfMDL[v,1]|alfa.verd >ICsupMDL[v,1])
      prob_cob_bayes_MDL[v,1]=0 else prob_cob_bayes_MDL[v,1]=1
  }
}

```

```

        if(beta.verd < ICinfMDL[v,2] | beta.verd > ICsupMDL[v,2])
            prob_cob_bayes_MDL[v,2]=0 else prob_cob_bayes_MDL[v,2]=1
        if(gama.verd < ICinfMDL[v,3] | gama.verd > ICsupMDL[v,3])
            prob_cob_bayes_MDL[v,3]=0 else prob_cob_bayes_MDL[v,3]=1
        if(lambda.verd < ICinfMDL[v,4] | lambda.verd > ICsupMDL[v,4])
            prob_cob_bayes_MDL[v,4]=0 else prob_cob_bayes_MDL[v,4]=1
        if(cura.verd < ICinfMDL[v,5] | cura.verd > ICsupMDL[v,5])
            prob_cob_bayes_MDL[v,5]=0 else prob_cob_bayes_MDL[v,5]=1
    } else {
        if (v==1) {
            mediasMDL[v,]= c(0.05,5,0.6,0.3,0.1)
        } else {
            mediasMDL[v,]=mediasMDL[v-1,]
            dpMDL[v,]=dpMDL[v-1,]
            NaiveSEMDL[v,]=NaiveSEMDL[v-1,]
            TimeSEMDL[v,]=TimeSEMDL[v-1,]
            ICinfMDL[v,]=ICinfMDL[v-1,]
            ICsupMDL[v,]=ICsupMDL[v-1,]
            llMDL[v]=llMDL[v-1]
            prob_cob_bayes_MDL[v,]=prob_cob_bayes_MDL[v-1,]
        }
        erro=erro+1
    }
}
erro

alfa.bayes = mediasMDL[,1]
beta.bayes = mediasMDL[,2]
gama.bayes = mediasMDL[,3]
lambda.bayes = mediasMDL[,4]
cura.bayes = mediasMDL[,5]

sd.alfa.bayes = dpMDL[,1]
sd.beta.bayes = dpMDL[,2]
sd.gama.bayes = dpMDL[,3]
sd.lambda.bayes = dpMDL[,4]
sd.cura.bayes = dpMDL[,5]

(sa=summary(alfa.bayes))
(sb=summary(beta.bayes))
(sg=summary(gama.bayes))
(sl=summary(lambda.bayes))
(sc=summary(cura.bayes))

naamostra = amostra
nbamostra = amostra
ngamostra = amostra
nlamostra = amostra
ncamostra = amostra

if (!is.na(sa["NA's"])) { naamostra = amostra-sa["NA's"]}
if (!is.na(sb["NA's"])) { nbamostra = amostra-sb["NA's"]}
if (!is.na(sg["NA's"])) { ngamostra = amostra-sg["NA's"]}

```

```

if (!is.na(sl["NA's"])) { nlamostra = amostra-sl["NA's"]}
if (!is.na(sc["NA's"])) { ncamostra = amostra-sc["NA's"]}

POST = 'unif';          ## unif para Uniforme, gama para Gama e hib para híbrida

jpeg(filename = paste(POST, '_', dist, '_', tp, '_', n, '.jpg', sep=""), width = 1000,
      height = 1000, pointsize = 12, quality = 100, bg = 'white')
  par(mfrow=c(6,4))
  par(mai=c(0.77,0.77,0.0, 0.07))
  ts.plot(alfa.bayes,ylab=expression(alpha.bayes))
  par(mai=c(0.77,0.77,0.45, 0.07))
  acf(alfa.bayes,main="expression(alpha.bayes)",lag=30)
  plot(density(alfa.bayes))
  plot(density(alfa.hat))

  par(mai=c(0.77,0.77,0.0, 0.07))
  ts.plot(beta.bayes,ylab=expression(beta.bayes))
  par(mai=c(0.77,0.77,0.45, 0.07))
  acf(beta.bayes,main="expression(beta.bayes)")
  plot(density(beta.bayes))
  plot(density(beta.hat))

  par(mai=c(0.77,0.77,0.0, 0.07))
  ts.plot(gama.bayes,ylab=expression(gama.bayes))
  par(mai=c(0.77,0.77,0.45, 0.07))
  acf(gama.bayes,main="expression(gama.bayes)")
  plot(density(gama.bayes))
  plot(density(gama.hat))

  par(mai=c(0.77,0.77,0.0, 0.07))
  ts.plot(lambda.bayes,ylab=expression(lambda.bayes))
  par(mai=c(0.77,0.77,0.45, 0.07))
  acf(lambda.bayes,main="expression(lambda.bayes)")
  plot(density(lambda.bayes))
  plot(density(lambda.hat))

  par(mai=c(0.77,0.77,0.0, 0.07))
  ts.plot(cura.bayes,ylab=expression(cura.bayes))
  par(mai=c(0.77,0.77,0.45, 0.07))
  acf(cura.bayes,main="expression(cura.bayes)")
  plot(density(cura.bayes))
  plot(density(cura.hat))

  hist(x,freq=F,ylim=c(0,1))
  lines(density(x),col="red")
dev.off()

## Aqui muda onde está U para H ou G conforme muda-se a posteriori de interesse
chainalphaU=mcmc(alfa.bayes)
summary(chainalphaU)
gw=geweke.diag(chainalphaU, frac1=0.1, frac2=0.5)
cbind(gw$z,if(abs(gw$z)<1.959964){'S'}else{'N'})
(gw.alpha.U=cbind(gw$z,if(abs(gw$z)<1.959964){'*'}else{''},sep=''))

```

```

(heidel.alpha.U=heidel.diag(chainalphaU, eps=0.1, pvalue=0.05))
hpd.alpha.U=emp.hpd(chainalphaU)

chainbetaU=mcmc(beta.bayes)
summary(chainbetaU)
gw=geweke.diag(chainbetaU, frac1=0.1, frac2=0.5)
cbind(gw$z,if(abs(gw$z)<1.959964){'S'}else{'N'})
(gw.beta.U=cbind(gw$z,if(abs(gw$z)<1.959964){'*'}else{''},sep=''))
(heidel.beta.U=heidel.diag(chainbetaU, eps=0.1, pvalue=0.05))
hpd.beta.U=emp.hpd(chainbetaU)

chaingamaU=mcmc(gama.bayes)
summary(chaingamaU)
gw=geweke.diag(chaingamaU, frac1=0.1, frac2=0.5)
cbind(gw$z,if(abs(gw$z)<1.959964){'S'}else{'N'})
(gw.gama.U=cbind(gw$z,if(abs(gw$z)<1.959964){'*'}else{''},sep=''))
(heidel.gama.U=heidel.diag(chaingamaU, eps=0.1, pvalue=0.05))
hpd.gama.U=emp.hpd(chaingamaU)

chainlambdaU=mcmc(lambda.bayes)
summary(chainlambdaU)
gw=geweke.diag(chainlambdaU, frac1=0.1, frac2=0.5)
cbind(gw$z,if(abs(gw$z)<1.959964){'S'}else{'N'})
(gw.lambda.U=cbind(gw$z,if(abs(gw$z)<1.959964){'*'}else{''},sep=''))
(heidel.lambda.U=heidel.diag(chainlambdaU, eps=0.1, pvalue=0.05))
hpd.lambda.U=emp.hpd(chainlambdaU)

chaincuraU=mcmc(cura.bayes)
summary(chaincuraU)
gw=geweke.diag(chaincuraU, frac1=0.1, frac2=0.5)
cbind(gw$z,if(abs(gw$z)<1.959964){'S'}else{'N'})
(gw.cura.U=cbind(gw$z,if(abs(gw$z)<1.959964){'*'}else{''},sep=''))
(heidel.cura.U=heidel.diag(chaincuraU, eps=0.1, pvalue=0.05))
hpd.cura.U=emp.hpd(chaincuraU)

txt=readLines(paste('saida', '_ ', dist, '_ ', tp, '_ ', n, '.R', sep=""))
txt=txt[seq(4,length(txt),5)]
tx=rep(0,length(txt))
for (i in 1:length(txt)) {#
    tx[i]=as.double(substring(txt[i],36))
}

(11.U=11MDL)
(pars.mu.U=c(mean(alfa.bayes,na.rm=T),mean(beta.bayes,na.rm=T),mean(gama.bayes,
    na.rm=T),mean(lambda.bayes,na.rm=T),mean(cura.bayes,na.rm=T)))
(pars.sd.U=c(mean(sd.alfa.bayes,na.rm=T),mean(sd.beta.bayes,na.rm=T),mean(sd.
    gama.bayes,na.rm=T),mean(sd.lambda.bayes,na.rm=T),mean(sd.cura.bayes,na.rm=T
    )))
(pars.11.U=mean(11MDL,na.rm=T))
(EQM.U=c(mean((alfa.bayes - alfa.verd)^2,na.rm=T),mean((beta.bayes - beta.verd)
    ^2,na.rm=T),mean((gama.bayes - gama.verd)^2,na.rm=T),mean((lambda.bayes -
    lambda.verd)^2,na.rm=T),mean((cura.bayes - cura.verd)^2,na.rm=T)))

```

```

(Efic.U=c(mean((alfa.bayes - alfa.verd)^2,na.rm=T)/mean((alfa.hat - alfa.verd)
^2,na.rm=T),mean((beta.bayes - beta.verd)^2,na.rm=T)/mean((beta.hat - beta.
verd)^2,na.rm=T),mean((gama.bayes - gama.verd)^2,na.rm=T)/mean((gama.hat -
gama.verd)^2,na.rm=T),mean((lambda.bayes - lambda.verd)^2,na.rm=T)/mean((
lambda.hat - lambda.verd)^2,na.rm=T),mean((cura.bayes - cura.verd)^2,na.rm=T
)/mean((cura.hat - cura.verd)^2,na.rm=T)))
(prob.U=c(sum(prob_cob_bayes_MDL[,1],na.rm=T)/(nahamostra),sum(
prob_cob_bayes_MDL[,2],na.rm=T)/(nbhamostra),sum(prob_cob_bayes_MDL[,3],na.
rm=T)/(nghamostra),sum(prob_cob_bayes_MDL[,4],na.rm=T)/(nlhamostra),sum(
prob_cob_bayes_MDL[,5],na.rm=T)/(nchamostra)))
(tx.U=mean(tx,na.rm=T))
(txsd.U=sd(tx,na.rm=T))
(hpd.inf.U=c(hpd.alpha.U[1],hpd.beta.U[1],hpd.gama.U[1],hpd.lambda.U[1],hpd.cura
.U[1]))
(hpd.sup.U=c(hpd.alpha.U[2],hpd.beta.U[2],hpd.gama.U[2],hpd.lambda.U[2],hpd.cura
.U[2]))
(gw.U=c(gw.alpha.U[1],gw.beta.U[1],gw.gama.U[1],gw.lambda.U[1],gw.cura.U[1]))
(gws.U=c(gw.alpha.U[2],gw.beta.U[2],gw.gama.U[2],gw.lambda.U[2],gw.cura.U[2]))
if(heidel.alpha.U[1]==1){hda='S'}else{hda='N'}
if(hda=='S'){if(heidel.alpha.U[4]==1){hda=paste(hda,' - S',sep="")}else{hda=
paste(hda,' - N',sep=")}}
if(heidel.beta.U[1]==1){hdb='S'}else{hdb='N'}
if(hdb=='S'){if(heidel.beta.U[4]==1){hdb=paste(hdb,' - S',sep="")}else{hdb=paste
(hdb,' - N',sep=")}}
if(heidel.gama.U[1]==1){hdg='S'}else{hdg='N'}
if(hdg=='S'){if(heidel.gama.U[4]==1){hdg=paste(hdg,' - S',sep="")}else{hdg=paste
(hdg,' - N',sep=")}}
if(heidel.lambda.U[1]==1){hdl='S'}else{hdl='N'}
if(hdl=='S'){if(heidel.lambda.U[4]==1){hdl=paste(hdl,' - S',sep="")}else{hdl=
paste(hdl,' - N',sep=")}}
if(heidel.cura.U[1]==1){hdc='S'}else{hdc='N'}
if(hdc=='S'){if(heidel.cura.U[4]==1){hdc=paste(hdc,' - S',sep="")}else{hdc=paste
(hdc,' - N',sep=")}}
(hd.U=c(hda,hdb,hdg,hdl,hdc))

# GERAÇÃO DE TABELAS PARA RESULTADOS
pars.mu.hat=round(pars.mu.hat,4)
pars.mu.U=round(pars.mu.U,4)
pars.mu.G=round(pars.mu.G,4)
pars.mu.H=round(pars.mu.H,4)
pars.sd.hat=round(pars.sd.hat,4)
pars.sd.U=round(pars.sd.U,4)
pars.sd.G=round(pars.sd.G,4)
pars.sd.H=round(pars.sd.H,4)
pars.ll.hat=round(pars.ll.hat,4)
pars.ll.U=round(pars.ll.U,4)
pars.ll.G=round(pars.ll.G,4)
pars.ll.H=round(pars.ll.H,4)
EQM.hat=round(EQM.hat,4)
EQM.U=round(EQM.U,4)
EQM.G=round(EQM.G,4)
EQM.H=round(EQM.H,4)
Efic.hat=round(Efic.hat,4)

```

```

Efic.U=round(Efic.U,4)
Efic.G=round(Efic.G,4)
Efic.H=round(Efic.H,4)
prob.hat=round(prob.hat,4)
prob.U=round(prob.U,4)
prob.G=round(prob.G,4)
prob.H=round(prob.H,4)
tx.U=round(tx.U,4)
tx.G=round(tx.G,4)
tx.H=round(tx.H,4)
hpd.inf.U=round(hpd.inf.U,4)
hpd.inf.G=round(hpd.inf.G,4)
hpd.inf.H=round(hpd.inf.H,4)
hpd.sup.U=round(hpd.sup.U,4)
hpd.sup.G=round(hpd.sup.G,4)
hpd.sup.H=round(hpd.sup.H,4)
gw.U=round(as.double(gw.U),4)
gw.G=round(as.double(gw.G),4)
gw.H=round(as.double(gw.H),4)

res=matrix(NA,8,20)
(res=
  rbind(
    cbind(
      pars.mu.hat[1],pars.mu.U[1],pars.mu.G[1],pars.mu.H[1],
      pars.mu.hat[2],pars.mu.U[2],pars.mu.G[2],pars.mu.H[2],
      pars.mu.hat[3],pars.mu.U[3],pars.mu.G[3],pars.mu.H[3],
      pars.mu.hat[4],pars.mu.U[4],pars.mu.G[4],pars.mu.H[4],
      pars.mu.hat[5],pars.mu.U[5],pars.mu.G[5],pars.mu.H[5]
    ),
    cbind(
      pars.sd.hat[1],pars.sd.U[1],pars.sd.G[1],pars.sd.H[1],
      pars.sd.hat[2],pars.sd.U[2],pars.sd.G[2],pars.sd.H[2],
      pars.sd.hat[3],pars.sd.U[3],pars.sd.G[3],pars.sd.H[3],
      pars.sd.hat[4],pars.sd.U[4],pars.sd.G[4],pars.sd.H[4],
      pars.sd.hat[5],pars.sd.U[5],pars.sd.G[5],pars.sd.H[5]
    ),
    cbind(
      EQM.hat[1],EQM.U[1],EQM.G[1],EQM.H[1],
      EQM.hat[2],EQM.U[2],EQM.G[2],EQM.H[2],
      EQM.hat[3],EQM.U[3],EQM.G[3],EQM.H[3],
      EQM.hat[4],EQM.U[4],EQM.G[4],EQM.H[4],
      EQM.hat[5],EQM.U[5],EQM.G[5],EQM.H[5]
    ),
    cbind(
      Efic.hat[1],Efic.U[1],Efic.G[1],Efic.H[1],
      Efic.hat[2],Efic.U[2],Efic.G[2],Efic.H[2],
      Efic.hat[3],Efic.U[3],Efic.G[3],Efic.H[3],
      Efic.hat[4],Efic.U[4],Efic.G[4],Efic.H[4],
      Efic.hat[5],Efic.U[5],Efic.G[5],Efic.H[5]
    ),
    cbind(
      prob.hat[1],prob.U[1],prob.G[1],prob.H[1],

```

```

        prob.hat[2],prob.U[2],prob.G[2],prob.H[2],
        prob.hat[3],prob.U[3],prob.G[3],prob.H[3],
        prob.hat[4],prob.U[4],prob.G[4],prob.H[4],
        prob.hat[5],prob.U[5],prob.G[5],prob.H[5]
    ),
    cbind(
        NA,tx.U,tx.G,tx.H,
        NA,tx.U,tx.G,tx.H,
        NA,tx.U,tx.G,tx.H,
        NA,tx.U,tx.G,tx.H,
        NA,tx.U,tx.G,tx.H
    ),
    cbind(
        NA,hpd.inf.U[1],hpd.inf.G[1],hpd.inf.H[1],
        NA,hpd.inf.U[2],hpd.inf.G[2],hpd.inf.H[2],
        NA,hpd.inf.U[3],hpd.inf.G[3],hpd.inf.H[3],
        NA,hpd.inf.U[4],hpd.inf.G[4],hpd.inf.H[4],
        NA,hpd.inf.U[5],hpd.inf.G[5],hpd.inf.H[5]
    ),
    cbind(
        NA,hpd.sup.U[1],hpd.sup.G[1],hpd.sup.H[1],
        NA,hpd.sup.U[2],hpd.sup.G[2],hpd.sup.H[2],
        NA,hpd.sup.U[3],hpd.sup.G[3],hpd.sup.H[3],
        NA,hpd.sup.U[4],hpd.sup.G[4],hpd.sup.H[4],
        NA,hpd.sup.U[5],hpd.sup.G[5],hpd.sup.H[5]
    )
)
)
rownames(res)=c('Mu','SD','EQM','Efic','Prob','Tx','HPD.Inf','HPD.Sup')
colnames(res)=c('a.MV','a.U','a.G','a.H','b.MV','b.U','b.G','b.H','g.MV','g.U','
    g.G','g.H','l.MV','l.U','l.G','l.H','c.MV','c.U','c.G','c.H')
res1=matrix(NA,4,20)
(res1=
    rbind(
        cbind(
            pars.ll.hat,pars.ll.U,pars.ll.G,pars.ll.H,
            pars.ll.hat,pars.ll.U,pars.ll.G,pars.ll.H,
            pars.ll.hat,pars.ll.U,pars.ll.G,pars.ll.H,
            pars.ll.hat,pars.ll.U,pars.ll.G,pars.ll.H,
            pars.ll.hat,pars.ll.U,pars.ll.G,pars.ll.H
        ),
        cbind(
            NA,gw.U[1],gw.G[1],gw.H[1],
            NA,gw.U[2],gw.G[2],gw.H[2],
            NA,gw.U[3],gw.G[3],gw.H[3],
            NA,gw.U[4],gw.G[4],gw.H[4],
            NA,gw.U[5],gw.G[5],gw.H[5]
        ),
        cbind(
            NA,gws.U[1],gws.G[1],gws.H[1],
            NA,gws.U[2],gws.G[2],gws.H[2],
            NA,gws.U[3],gws.G[3],gws.H[3],
            NA,gws.U[4],gws.G[4],gws.H[4],

```

```

        NA,gws.U[5],gws.G[5],gws.H[5]
    ),
    cbind(
        NA,hd.U[1],hd.G[1],hd.H[1],
        NA,hd.U[2],hd.G[2],hd.H[2],
        NA,hd.U[3],hd.G[3],hd.H[3],
        NA,hd.U[4],hd.G[4],hd.H[4],
        NA,hd.U[5],hd.G[5],hd.H[5]
    )
)
)
rownames(res1)=c('l1','Geweke','GewekeS','Heidel')
colnames(res1)=c('a.MV','a.U','a.G','a.H','b.MV','b.U','b.G','b.H','g.MV','g.U',
    'g.G','g.H','l.MV','l.U','l.G','l.H','c.MV','c.U','c.G','c.H')
print(xtable(t(res),digits=4),floating=T)
print(xtable(t(res1)),floating=T)

unlink(paste('saida','_',dist,'_',tp,'_',n,'.R',sep=""))

save.image(file=paste(dist,'_',tp,'_',n,'.RData'))

```

B.2 Aplicações

Nessa seção apresenta-se os programas elaboradas para a obtenção dos resultados das aplicações: Pacientes com câncer internados na UTI do INCA (ctinca ou ct) e Dados de câncer gástrico do Hospital do Câncer de Barretos (gastriccancer ou gc).

B.2.0.1 Análise semi-paramétrica

```

rm(list = ls())
set.seed(1234)
library(survival)
library(flexsurv)
setwd('C:/Users/EduardoCardosode/Dropbox/Mestrado/Dissertação/Dados e Programas/
    Aplicação R')
bd='ctinca.csv' # ou 'gastriccancer.csv'
dados<-read.table(bd, header=TRUE, fill=TRUE,sep=';')
setwd('C:/Users/EduardoCardosode/Dropbox/Mestrado/Dissertação/Resultados')

median(dados[, "idade"])

# AJUSTES PARA UTILIZAÇÃO DE COVARIÁVEIS
for(i in 1:nrow(dados)) {
    if(dados[i,"idade"]<=60) dados$fxetaria[i]=0
    else if(dados[i,"idade"]>60) dados$fxetaria[i]=1
}

### ANÁLISE SEMI-PARAMÉTRICA
jpeg(filename = 'ct-AE-f00.jpg', width = 1500, height = 2000, pointsize = 12,
    quality = 100, bg = 'white')
par(mar=c(4.5,4.5,2,1),mfrow=c(3,2),cex=2)
V<- survfit(Surv(tempo,status)~factor(gptumor), data = dados)
summary(V)

```

```

survdif(Surv(tempo,status)~factor(gptumor), data = dados,rho=0)
plot(V,lty=c(1,4,2),mark.time=F,xlab="Tempo (dias)",ylab="S(t)")
title("gptumor: Tipo de tumor")
legend(115,0.2,lty=c(1),c("Sólido Localizado"),bty="n",cex=0.8)
legend(115,0.15,lty=c(4),c("Metastático"),bty="n",cex=0.8)
legend(115,0.1,lty=c(2),c("Hematológico"),bty="n",cex=0.8)

V<- survfit(Surv(tempo,status)~factor(sexo), data = dados)
summary(V)
survdif(Surv(tempo,status)~factor(sexo), data = dados,rho=0)
plot(V,lty=c(1,4),mark.time=F,xlab="Tempo (dias)",ylab="S(t)")
title("sexo: Gênero")
legend(115,0.2,lty=c(1),c("Feminino"),bty="n",cex=0.8)
legend(115,0.15,lty=c(4),c("Masculino"),bty="n",cex=0.8)

V<- survfit(Surv(tempo,status)~factor(fxetaria), data = dados)
summary(V)
survdif(Surv(tempo,status)~factor(fxetaria), data = dados,rho=0)
plot(V,lty=c(1,4),mark.time=F,xlab="Tempo (dias)",ylab="S(t)")
title("fxetaria: Faixa etária")
legend(115,0.2,lty=c(1),c("Até 60 anos"),bty="n",cex=0.8)
legend(115,0.15,lty=c(4),c("Acima de 60 anos"),bty="n",cex=0.8)

V<- survfit(Surv(tempo,status)~leucopenia, data = dados)
summary(V)
survdif(Surv(tempo,status)~leucopenia, data = dados,rho=0)
plot(V,lty=c(1,4),mark.time=F,xlab="Tempo (dias)",ylab="S(t)")
title("leucopenia: Leucopenia presente")
legend(115,0.15,lty=c(1),c("Não"),bty="n",cex=0.8)
legend(115,0.10,lty=c(4),c("Sim"),bty="n",cex=0.8)

V<- survfit(Surv(tempo,status)~comorbi, data = dados)
summary(V)
survdif(Surv(tempo,status)~comorbi, data = dados,rho=0)
plot(V,lty=c(1,4),mark.time=F,xlab="Tempo (dias)",ylab="S(t)")
title("comorbi: Comorbidades severas presentes")
legend(115,0.15,lty=c(1),c("Não"),bty="n",cex=0.8)
legend(115,0.10,lty=c(4),c("Sim"),bty="n",cex=0.8)

V<- survfit(Surv(tempo,status)~desnut, data = dados)
summary(V)
survdif(Surv(tempo,status)~desnut, data = dados,rho=0)
plot(V,lty=c(1,4),mark.time=F,xlab="Tempo (dias)",ylab="S(t)")
title("desnut: Perda de peso recente acima de 10%")
legend(115,0.15,lty=c(1),c("Não"),bty="n",cex=0.8)
legend(115,0.10,lty=c(4),c("Sim"),bty="n",cex=0.8)
dev.off()

#logrank
survdif(Surv(tempo,status)~factor(gptumor), data = dados,rho=0)
survdif(Surv(tempo,status)~factor(sexo), data = dados,rho=0)
survdif(Surv(tempo,status)~factor(fxetaria), data = dados,rho=0)
survdif(Surv(tempo,status)~leucopenia, data = dados,rho=0)

```

```

survdif(Surv(tempo,status)~comorbi, data = dados,rho=0)
survdif(Surv(tempo,status)~desnut, data = dados,rho=0)

#Wilcoxon
survdif(Surv(tempo,status)~factor(gptumor), data = dados,rho=1)
survdif(Surv(tempo,status)~factor(sexo), data = dados,rho=1)
survdif(Surv(tempo,status)~factor(fxetaria), data = dados,rho=1)
survdif(Surv(tempo,status)~leucopenia, data = dados,rho=1)
survdif(Surv(tempo,status)~comorbi, data = dados,rho=1)
survdif(Surv(tempo,status)~desnut, data = dados,rho=1)

### gptumor
#logrank
survdif(Surv(tempo,status)~gptumor, data=dados[dados[, "gptumor"]%in%c('1','2')
,], rho=0) ## Loco vs Mtx
survdif(Surv(tempo,status)~gptumor, data=dados[dados[, "gptumor"]%in%c('1','3')
,], rho=0) ## Loco vs Hemato
survdif(Surv(tempo,status)~gptumor, data=dados[dados[, "gptumor"]%in%c('2','3')
,], rho=0) ## Mtx vs Hemato

#Wilcoxon
survdif(Surv(tempo,status)~gptumor, data=dados[dados[, "gptumor"]%in%c('1','2')
,], rho=1) ## Loco vs Mtx
survdif(Surv(tempo,status)~gptumor, data=dados[dados[, "gptumor"]%in%c('1','3')
,], rho=1) ## Loco vs Hemato
survdif(Surv(tempo,status)~gptumor, data=dados[dados[, "gptumor"]%in%c('2','3')
,], rho=1) ## Mtx vs Hemato

# AJUSTES PARA UTILIZAÇÃO DE COVARIÁVEIS
for(i in 1:nrow(dados)) {
  if(dados[i,"gptumor"]==1) dados$tumor[i]=0
  else if(dados[i,"gptumor"]>1) dados$tumor[i]=1
}

#Passo 1
(fit<-coxph(Surv(tempo,status)~1,x = T,method="breslow", data = dados))
(llg=2*fit$loglik)
(fit1<-coxph(Surv(tempo,status)~factor(tumor),x = T,method="breslow", data =
dados))
(llm=2*fit1$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fit2<-coxph(Surv(tempo,status)~factor(sexo),x = T,method="breslow", data =
dados))
(llm=2*fit2$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fit3<-coxph(Surv(tempo,status)~fxetaria,x = T,method="breslow", data = dados))
(llm=2*fit3$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fit4<-coxph(Surv(tempo,status)~leucopenia,x = T,method="breslow", data = dados)
)
(llm=2*fit4$loglik[2])
round(1-pchisq(llm-llg,1),4)

```

```

(fit5<-coxph(Surv(tempo,status)~comorbi,x = T,method="breslow", data = dados))
(llm=2*fit5$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fit6<-coxph(Surv(tempo,status)~desnut,x = T,method="breslow", data = dados))
(llm=2*fit6$loglik[2])
round(1-pchisq(llm-llg,1),4)

#Passo 2
(fit1<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
,x = T,method="breslow", data = dados))
(llg=2*fit1$loglik[2])
(fit2<-coxph(Surv(tempo,status)~leucopenia+fxetaria+comorbi+desnut,x = T,method
="breslow", data = dados))
(llm=2*fit2$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fit3<-coxph(Surv(tempo,status)~factor(tumor)+fxetaria+comorbi+desnut,x = T,
method="breslow", data = dados))
(llm=2*fit3$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fit4<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+comorbi+desnut,x = T,
method="breslow", data = dados))
(llm=2*fit4$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fit5<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+desnut,x = T,
method="breslow", data = dados))
(llm=2*fit5$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fit6<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi,x = T,
method="breslow", data = dados))
(llm=2*fit6$loglik[2])
round(1-pchisq(llm-llg,1),4)

#Passo 3
(fit1<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
,x = T,method="breslow", data = dados))
(llg=2*fit1$loglik[2])
(fit2<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
+factor(sexo),x = T,method="breslow", data = dados))
(llm=2*fit2$loglik[2])
round(1-pchisq(llm-llg,1),4)

#Passo 4
(fit1<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
,x = T,method="breslow", data = dados))
(llg=2*fit1$loglik[2])
(fit2<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
+factor(tumor)*leucopenia,x = T,method="breslow", data = dados))
(llm=2*fit2$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fit3<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
+factor(tumor)*fxetaria,x = T,method="breslow", data = dados))
(llm=2*fit3$loglik[2])
round(1-pchisq(llm-llg,1),4)

```

```

(fit4<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
  +factor(tumor)*comorbi,x = T,method="breslow", data = dados))
(llm=2*fit4$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fit5<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
  +factor(tumor)*desnut,x = T,method="breslow", data = dados))
(llm=2*fit5$loglik[2])
round(1-pchisq(llm-llg,1),4)#
(fit6<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
  +leucopenia*fxetaria,x = T,method="breslow", data = dados))
(llm=2*fit6$loglik[2])
round(1-pchisq(llm-llg,1),4)#
(fit7<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
  +leucopenia*comorbi,x = T,method="breslow", data = dados))
(llm=2*fit7$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fit8<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
  +leucopenia*desnut,x = T,method="breslow", data = dados))
(llm=2*fit8$loglik[2])
round(1-pchisq(llm-llg,1),4)#
(fit9<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
  +fxetaria*comorbi,x = T,method="breslow", data = dados))
(llm=2*fit9$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fita<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
  +fxetaria*desnut,x = T,method="breslow", data = dados))
(llm=2*fita$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fitb<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
  +comorbi*desnut,x = T,method="breslow", data = dados))
(llm=2*fitb$loglik[2])
round(1-pchisq(llm-llg,1),4)
fits=fit1

```

#Passo 5

```

(fit1<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
  +leucopenia*desnut,x = T,method="breslow", data = dados))
(llg=2*fit1$loglik[2])
(fit2<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
  +factor(tumor)*desnut+leucopenia*fxetaria,x = T,method="breslow", data =
  dados))
(llm=2*fit2$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fit3<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
  +factor(tumor)*desnut+leucopenia*desnut,x = T,method="breslow", data = dados
  ))
(llm=2*fit3$loglik[2])
round(1-pchisq(llm-llg,1),4)
(fit4<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
  +leucopenia*desnut+leucopenia*fxetaria,x = T,method="breslow", data = dados
  )
(llm=2*fit4$loglik[2])
round(1-pchisq(llm-llg,1),4)

```

```

fit4=fit8

jpeg(filename = 'ct-AE-f01.jpg', width = 1500, height = 1000, pointsize = 12,
      quality = 100, bg = 'white')
par(mfrow=c(2,2),cex=1.5)
fitk<-coxph(Surv(tempo[tumor==1],status[tumor==1])~1,data=dados,x=T,method="
      breslow")
ss<- survfit(fitk)
s0<-round(ss$surv,digits=5)
H0<- -log(s0)
plot(ss$time,log(H0),xlim=range(c(0,20)),xlab="Tempos",
      ylab=expression(log(Lambda[0]*(t))), bty="n",type="s")
fitl<-coxph(Surv(tempo[tumor!=1],status[tumor!=1])~1,data=dados,x=T,method="
      breslow")
ss<- survfit(fitl)
s0<-round(ss$surv,digits=5)
H0<- -log(s0)
lines(ss$time,log(H0),type="s",lty=2)
legend(5,-3,lty=c(1,2,4),c("tumor = 1 (Sólido localizado)","tumor = 2 (
      Metastático) e 3 (Hematológico)"),lwd=1,bty="n",cex=0.7)
title("tumor: Tipo de tumor")

fitk<-coxph(Surv(tempo[leucopenia==0],status[leucopenia==0])~1,data=dados,x=T,
      method="breslow")
ss<- survfit(fitk)
s0<-round(ss$surv,digits=5)
H0<- -log(s0)
plot(ss$time,log(H0),xlim=range(c(0,20)),xlab="Tempos",
      ylab=expression(log(Lambda[0]*(t))), bty="n",type="s")
fitl<-coxph(Surv(tempo[leucopenia==1],status[leucopenia==1])~1,data=dados,x=T,
      method="breslow")
ss<- survfit(fitl)
s0<-round(ss$surv,digits=5)
H0<- -log(s0)
lines(ss$time,log(H0),type="s",lty=2)
legend(5,-3,lty=c(2,1),c("leucopenia= 0 (Não)","leucopenia= 1 (Sim)"),lwd=1,bty
      ="n",cex=0.7)
title("leucopenia: Leucopenia")

fitk<-coxph(Surv(tempo[comorbi==0],status[comorbi==0])~1,data=dados,x=T,method
      ="breslow")
ss<- survfit(fitk)
s0<-round(ss$surv,digits=5)
H0<- -log(s0)
plot(ss$time,log(H0),xlim=range(c(0,20)),xlab="Tempos",
      ylab=expression(log(Lambda[0]*(t))), bty="n",type="s")
fitl<-coxph(Surv(tempo[comorbi==1],status[comorbi==1])~1,data=dados,x=T,method
      ="breslow")
ss<- survfit(fitl)
s0<-round(ss$surv,digits=5)
H0<- -log(s0)
lines(ss$time,log(H0),type="s",lty=2)

```

```

legend(5,-3,lty=c(2,1),c("comorbi= 0 (Não)","comorbi= 1 (Sim)"),lwd=1,bty="n",
      cex=0.7)
title("comorbi: Comorbidades severas")

fitk<-coxph(Surv(tempo[desnut==0],status[desnut==0])~1,data=dados,x=T,method="
      breslow")
ss<- survfit(fitk)
s0<-round(ss$surv,digits=5)
H0<- -log(s0)
plot(ss$time,log(H0),xlim=range(c(0,20)),xlab="Tempos",
      ylab=expression(log(Lambda[0]*(t))), bty="n",type="s")
fitl<-coxph(Surv(tempo[desnut==1],status[desnut==1])~1,data=dados,x=T,method="
      breslow")
ss<- survfit(fitl)
s0<-round(ss$surv,digits=5)
H0<- -log(s0)
lines(ss$time,log(H0),type="s",lty=2)
legend(5,-3,lty=c(2,1),c("desnut= 0 (Não)","desnut= 1 (Sim)"),lwd=1,bty="n",cex
      =0.7)
title("desnut: Perda de peso recente >10%")
dev.off()

```

```

resid(fit4,type="scaledsch")
cox.zph(fit4, transform="identity")    ### g(t) = t
jpeg(filename = 'ct-AE-f02.jpg', width = 2000, height = 1500, pointsize = 12,
      quality = 100, bg = 'white')
par(mar=c(4.5,4.5,1,1),mfrow=c(3,4),cex=1.5)
plot(cox.zph(fit4))
dev.off()

```

```

resid(fits,type="scaledsch")
cox.zph(fits, transform="identity")    ### g(t) = t
jpeg(filename = 'ct-AE-f03.jpg', width = 1500, height = 1000, pointsize = 12,
      quality = 100, bg = 'white')
par(mar=c(4.5,4.5,1,1),mfrow=c(2,3),cex=1.5)
plot(cox.zph(fits))
dev.off()

```

```

Ht<-basehaz(fits,centered=F)
tempos<-Ht$time
H0<-Ht$hazard
S0<- exp(-H0)
round(cbind(tempos, S0,H0),digits=5)

```

```

(fitk<-coxph(Surv(tempo,status)~factor(tumor)+leucopenia+fxetaria+comorbi+desnut
      +leucopenia*desnut,x = T,method="breslow", data = dados))
summary(fitk)
Ht<-basehaz(fitk,centered=F)
tempos<-Ht$time
H0<-Ht$hazard
S0<- exp(-H0)
round(cbind(tempos,S0,H0),digits=5)

```

```

tt<-sort(tempos)
aux1<-as.matrix(tt)
n<-nrow(aux1)
aux2<-as.matrix(cbind(tempos,S0))
S00<-rep(max(aux2[,2]),n)
  for(i in 1:n){
    if(tt[i]> min(aux2[,1])){
      i1<- aux2[,1]<= tt[i]
      S00[i]<-min(aux2[i1,2])}]
ts0<-cbind(tt,S00)
ts0
(b<-fitk$coefficients)
id<-50
st1<- S00^(exp(b[4]))          # S(t|x) Loco
st2<- S00^(exp(b[1]+((b[4]+b[7])))) # S(t|x) Mtx e Hemato
jpeg(filename = 'ct-AE-f04.jpg', width = 500, height = 500, pointsize = 12,
      quality = 100, bg = 'white')
par(mar=c(4.5,4.5,3,1),mfrow=c(1,1),cex=1.5)
plot(tt,st1,type="s",ylim=range(c(0,1)),xlab="Tempos",ylab="S(t|x)",lty=1)
lines(tt,st2,type="s",lty=2)
legend(0,0.2,lty=c(1,2,3,4),c("Sólido localizado","Metastático e Hematológico")
      , lwd=1,bty="n",cex=0.7)
title("tumor: Tipo de tumor")
dev.off()

jpeg(filename = 'ct-AE-f05.jpg', width = 500, height = 500, pointsize = 12,
      quality = 100, bg = 'white')
par(mar=c(4.5,4.5,1,1),mfrow=c(1,1),cex=1.5)
fitx<-coxph(Surv(tempo[tumor==1],status[tumor==1])~1,data=dados,x=T,method="
      breslow")
ss<- survfit(fitx)
s0<-round(ss$surv,digits=5)
H0<- -log(s0)
plot(ss$time,log(H0),xlim=range(c(0,20)),xlab="Tempos",
      ylab=expression(log(Lambda[0]*(t))), bty="n",type="s")
fity<-coxph(Surv(tempo[tumor==2],status[tumor==2])~1,data=dados,x=T,method="
      breslow")
ss<- survfit(fity)
s0<-round(ss$surv,digits=5)
H0<- -log(s0)
lines(ss$time,log(H0),type="s",lty=2)
fitz<-coxph(Surv(tempo[tumor==3],status[tumor==3])~1,data=dados,x=T,method="
      breslow")
ss<- survfit(fitz)
s0<-round(ss$surv,digits=5)
H0<- -log(s0)
lines(ss$time,log(H0),type="s",lty=3)
legend(5,-3,lty=c(2,1),c("Sólido localizado","Metastático","Hematológico"),lwd
      =1,bty="n",cex=0.7)
title("tumor: Tipo de tumor")
dev.off()

fit<-coxph(Surv(tempo,cens)~V1+V3+V4+V6,data=desmame,x = T,method="breslow")

```

```

summary(fitk)
fit$loglik

(ekm<-survfit(fit4, data=dados))
summary(ekm)

jpeg(filename = 'ct-AE-f06.jpg', width = 500, height = 500, pointsize = 12,
      quality = 100, bg = 'white')
par(mar=c(4.5,4.5,1,1),mfrow=c(1,1),cex=1.5)
plot(ekm,xlab="Tempo (dias)", conf.int=F, ylab="S(t) estimada")
legend(1,0.2,c("Kaplan-Meier"),lwd=1,bty="n")
dev.off()

jpeg(filename = 'ct-AE-f07.jpg', width =2000, height = 500, pointsize = 12,
      quality = 100, bg = 'white')
par(mar=c(4.5,4.5,1,1),mfrow=c(1,3),cex=1.5)
invst<-qnorm(st)
plot(time, -log(st),pch=16,xlab="tempos",ylab="-log(S(t))")
plot(log(time),log(-log(st)),pch=16,xlab="log(tempos)",ylab="log(-log(S(t)))")
plot(log(time),invst, pch=16,xlab="log(tempos)", ylab=expression(Phi^-1 * (S(t)
)))
dev.off()

```

B.2.0.2 Enfoque Clássico

```

rm(list = ls())
set.seed(1234567890)
library(survival)
library(flexsurv)
library(zipfR)
library(maxLik)
setwd('C:/Users/EduardoCardosode/Dropbox/Mestrado/Dissertação/Dados e Programas/
      Aplicação R')
bd='ctinca.csv' # ou 'gastriccancer.csv'
dados<-read.table(bd, header=TRUE, fill=TRUE,sep=';')
setwd('C:/Users/EduardoCardosode/Dropbox/Mestrado/Dissertação/Resultados')

# AJUSTES PARA UTILIZAÇÃO DE COVARIÁVEIS
for(i in 1:nrow(dados)) {
  if(dados[i,"idade"]<=60) dados$fxetaria[i]=1
  else if(dados[i,"idade"]>60) dados$fxetaria[i]=2
  if(dados[i,"gptumor"]==1) dados$tumor[i]=0
  else if(dados[i,"gptumor"]>1) dados$tumor[i]=1
}
n=dim(dados)[1]

sf=Surv(tempo,status)~1
(ekm<-survfit(sf, data=dados))
time<-ekm$time
st<-ekm$surv

loglike=c()
AIC=c()

```

```

BIC=c()
CAIC=c()
d=c()

# NAS FUNÇÕES LOGLIK DEVE-SE MUDAR CONFORME A NECESSIDADE
# PARA MODELOS DE MISTURA
#           f = (1-pi)*f0;
#           S = pi+(1-pi)*S0
#           L = (f**status)*(S**(1-status));
# PARA MODELOS DE NÃO-MISTURA
#           h = -log(pi)*f0;
#           S = exp(F0*log(pi));
#           L = (h**status)*S;
# PARA MODELOS SEM FRAÇÃO DE CURA
#           h = f0/S0;
#           L = (h**status)*S0;

print("Gamma Generalizada")
dist='GG'
loglik <- function(param) {
  alpha <- param[1]
  beta <- param[2]
  gamma <- param[3]
  pi <- param[4]
  mi=0
  status<-dados$status
  tempo=(dados$tempo)
  if ( (alpha>0) && (beta>0) && (gamma>0) && (pi>0) && (pi<1) ) {
    f0 = exp(-((tempo-mi)/beta)**gamma)*(((tempo-mi)/beta)**(-1 + (
      alpha*gamma)))**gamma/(beta*gamma(alpha))
    F0 = (Rgamma(alpha,((tempo-mi)/beta)**gamma))
    S0 = 1-F0;
    f = (1-pi)*f0;
    S = pi+(1-pi)*S0
    L = (f**status)*(S**(1-status));
    ll<-log(L)
  } else return (ll <- -Inf)
  ll
}

(res <- maxLik(loglik, start=c(16,0.01,.2,.5), method='BFGS')) # use 'wrong'
start values
summary(res)
(pars=coef(res))
(loglike[dist]=logLik(res))
stgg<-pars[4]+(1-pars[4])*(1-(Rgamma(pars[1],((time)/pars[2] )**pars[3])))
d[dist]=length(pars)
AIC[dist]=-(2*loglike[dist])+(2*d[dist])
BIC[dist]=-(2*loglike[dist])+(d[dist]*log(n))
CAIC[dist]=AIC[dist]+(2*(d[dist]+2)*(d[dist]+3)/(n-d[dist]-3))

print("Gamma")
dist='Gamma'
loglik <- function(param) {

```

```

mi <- param[1]
fi <- param[2]
pi<- param[3]
status<-dados$status
tempo=dados$tempo
if ( (mi>0) && (fi>0) && (pi>0) && (pi<1) ) {
  f0 = dgamma(tempo,mi,fi);
  F0 = pgamma(tempo,mi,fi);
  S0 = 1-F0;
  f = (1-pi)*f0;
  S = pi+(1-pi)*S0
  L = (f**status)*(S**(1-status));
  ll<-log(L)
} else return (ll <- -Inf)
ll
}
(res <- maxLik(loglik, start=c(.5,.01,.5), method='BFGS')) # use 'wrong' start
values
summary(res)
(pars=coef(res))
(loglike[dist]=logLik(res))
stg<-pars[3]+(1-pars[3])*(1-pgamma(time,pars[1],pars[2]))
d[dist]=length(pars)
AIC[dist]=-(2*loglike[dist])+(2*d[dist])
BIC[dist]=-(2*loglike[dist])+(d[dist]*log(n))
CAIC[dist]=AIC[dist]+(2*(d[dist]+2)*(d[dist]+3)/(n-d[dist]-3))

print("Log-Normal")
dist='LN'
loglik <- function(param) {
  mi <- param[1]
  sigma <- param[2]
  pi<- param[3]
  status<-dados$status
  tempo=dados$tempo
  if ( (sigma>0) && (pi>0) && (pi<1) ) {
    f0 = dlnorm(tempo,mi,sigma);
    F0 = plnorm(tempo,mi,sigma);
    S0 = 1-F0;
    f = (1-pi)*f0;
    S = pi+(1-pi)*S0
    L = (f**status)*(S**(1-status));
    ll<-log(L)
  } else return (ll <- -Inf)
  ll
}
(res <- maxLik(loglik, start=c(3,2,.5), method='BFGS')) # use 'wrong' start
values
summary(res)
(pars=coef(res))
(loglike[dist]=logLik(res))
stln<-pars[3]+(1-pars[3])*(1-plnorm(time,pars[1],pars[2]))
d[dist]=length(pars)

```

```

AIC[dist]=-(2*loglike[dist])+(2*d[dist])
BIC[dist]=-(2*loglike[dist])+(d[dist]*log(n))
CAIC[dist]=AIC[dist]+(2*(d[dist]+2)*(d[dist]+3)/(n-d[dist]-3))

print("Weibull")
dist='W'
loglik <- function(param) {
  lambda <- param[1]
  gamma <- param[2]
  pi<- param[3]
  status<-dados$status
  tempo=dados$tempo
  if ( (lambda >0) && (gamma >0) && (pi>0) && (pi<1) ) {
    f0 = dweibull(tempo,lambda,gamma);
    F0 = pweibull(tempo,lambda,gamma);
    S0 = 1-F0;
    f = (1-pi)*f0;
    S = pi+(1-pi)*S0
    L = (f**status)*(S**(1-status));
    ll<-log(L)
  } else return (ll <- -Inf)
  ll
}

(res <- maxLik(loglik, start=c(.9,26,.5), method='BFGS')) # use 'wrong' start
values
summary(res)
(pars=coef(res))
(loglike[dist]=logLik(res))
F0 = pweibull(time,pars[1],pars[2]);
stw<-pars[3]+(1-pars[3])*(1-F0)
d[dist]=length(pars)
AIC[dist]=-(2*loglike[dist])+(2*d[dist])
BIC[dist]=-(2*loglike[dist])+(d[dist]*log(n))
CAIC[dist]=AIC[dist]+(2*(d[dist]+2)*(d[dist]+3)/(n-d[dist]-3))

print("Exponencial Exponenciada")
dist='EE'
loglik <- function(param) {
  alpha <- param[1]
  beta <- param[2]
  pi<- param[3]
  status<-dados$status
  tempo=dados$tempo
  if ( (alpha >0) && (beta >0) && (pi>0) && (pi<1) ) {
    f0 = alpha*beta*exp(-alpha*(tempo))*((1-exp(-alpha*(tempo)))**
      (beta-1));
    S0 = 1-(1-exp(-alpha*(tempo)))**beta;
    F0 = (1-exp(-alpha*(tempo)))**beta;
    f = (1-pi)*f0;
    S = pi+(1-pi)*S0
    L = (f**status)*(S**(1-status));
    ll<-log(L)
  } else return (ll <- -Inf)
}

```

```

11
}
(res <- maxLik(loglik, start=c(.2,1,.5), method='BFGS')) # use 'wrong' start
  values
summary(res)
(pars=coef(res))
(loglike[dist]=logLik(res))
steg<-pars[3]+(1-pars[3])*(1-(1-exp(-pars[1]*(time))))**pars[2])
d[dist]=length(pars)
AIC[dist]=- (2*loglike[dist])+(2*d[dist])
BIC[dist]=- (2*loglike[dist])+(d[dist]*log(n))
CAIC[dist]=AIC[dist]+(2*(d[dist]+2)*(d[dist]+3)/(n-d[dist]-3))

print("Weibull Generalizada")
dist='WG'
loglik <- function(param) {
  alpha <- param[1]
  beta <- param[2]
  gama <- param[3]
  pi<- param[4]
  status<-dados$status
  tempo=dados$tempo
  if ( (alpha >0) && (beta >0) && (gama >0) && (pi>0) && (pi<1) ) {
    f0 = alpha*beta*gama*(tempo**(gama-1))*exp(-alpha*(tempo**gama))
      *((1-exp(-alpha*(tempo**gama)))**(beta-1));
    S0 = 1-(1-exp(-alpha*(tempo**gama)))**beta;
    F0 = (1-exp(-alpha*(tempo**gama)))**beta;
    f = (1-pi)*f0;
    S = pi+(1-pi)*S0
    L = (f**status)*(S**(1-status));
    ll<-log(L)
  } else return (ll <- -Inf)
  ll
}
(res <- maxLik(loglik, start=c(2,300,.85,.5), method='BFGS', iterlim=5000)) #
  use 'wrong' start values
summary(res)
(pars=coef(res))
(loglike[dist]=logLik(res))
stwg<-pars[4]+(1-pars[4])*(1-(1-exp(-pars[1]*(time**pars[3]))))**pars[2])
d[dist]=length(pars)
AIC[dist]=- (2*loglike[dist])+(2*d[dist])
BIC[dist]=- (2*loglike[dist])+(d[dist]*log(n))
CAIC[dist]=AIC[dist]+(2*(d[dist]+2)*(d[dist]+3)/(n-d[dist]-3))

print("Weibull Modificada Generalizada")
dist='WMG'
loglik <- function(param) {
  alpha <- param[1]
  beta <- param[2]
  gama <- param[3]
  lambda<- param[4]
  pi<- param[5]

```

```

status<-dados$status
tempo=dados$tempo
if ( (alpha >0) && (beta >0) && (gama >0) && (lambda>0) && (pi>0) && (pi
  <1) ) {
  f0 = alpha*beta*(gama+(lambda*tempo))*(tempo**(gama-1))*exp((
    lambda*tempo)-alpha*(tempo**gama)*exp(lambda*tempo))*((1-exp(-
    alpha*(tempo**gama)*exp(lambda*tempo)))**(beta-1));
  S0 = 1-(1-exp(-alpha*(tempo**gama)*exp(lambda*tempo)))**beta;
  F0 = (1-exp(-alpha*(tempo**gama)*exp(lambda*tempo)))**beta;
  f = (1-pi)*f0;
  S = pi+(1-pi)*S0
  L = (f**status)*(S**(1-status));
  ll<-log(L)
} else return (ll <- -Inf)
ll
}
(res <- maxLik(loglik, start=c(10, 500, .05, .001, .15), method='BFGS', iterlim
  =10000)) # use 'wrong' start values(pars=coef(res))
summary(res)
(loglike[dist]=logLik(res))
stwmg<-pars[5]+(1-pars[5])*(1-(1-exp(-pars[1]*(time**pars[3]))*exp(pars[4]*time))
  )**pars[2])
d[dist]=length(pars)
(AIC[dist]=-(2*loglike[dist])+(2*d[dist]))
(BIC[dist]=-(2*loglike[dist])+(d[dist]*log(n)))
(CAIC[dist]=AIC[dist]+(2*(d[dist]+2)*(d[dist]+3)/(n-d[dist]-3)))

round(cbind(time,st,stgg,stg,stln,stw,steg,stwg,stwmg),3)

TRV=c()
pvalor=c()
for(i in 1:(length(loglike)-1)){
  TRV[i]=2*(loglike[i]-loglike['WMG'])
  pvalor[i]=round(1-pchisq(abs(TRV[i]),abs(d[i]-d['WMG'])),4)
}
TRV[i+1]=0
pvalor[i+1]=0
TAB=cbind(as.matrix(loglike),as.matrix(TRV),as.matrix(pvalor),as.matrix(AIC),as.
  matrix(BIC),as.matrix(CAIC))
colnames(TAB)=c(expression(log(L(Theta))), 'TRV', 'valor p', 'AIC', 'BIC', 'CAIC')
TAB

## gráficos kaplan-meyer
jpeg(filename = 'ct-CM-sCV-f00.jpg', width = 1500, height = 750, pointsize = 12,
  quality = 100, bg = 'white')
par(mar=c(4.5,4.5,0.5,0.5), mfrow=c(2,4), cex=1.3, cex.sub=1.3)
plot(ekm, conf.int=F, xlab="Tempos", lwd=2, ylab="S(t)")
lines(c(0,time),c(1,stgg), col='red', lwd=2, lty=2)
legend(10,0.2,lty=c(1,2),c("Kaplan-Meier", "GG"),bty="n",cex=0.8)
plot(ekm, conf.int=F, xlab="Tempos", lwd=2, ylab="S(t)")
lines(c(0,time),c(1,stg), col='red', lwd=2, lty=2)
legend(10,0.2,lty=c(1,2),c("Kaplan-Meier", "Gama"),bty="n",cex=0.8)
plot(ekm, conf.int=F, xlab="Tempos", lwd=2, ylab="S(t)")

```

```

lines(c(0,time),c(1,stw), col='red', lwd=2, lty=2)
legend(10,0.2,lty=c(1,2),c("Kaplan-Meier", "Weibull"),bty="n",cex=0.8)
plot(ekm, conf.int=F, xlab="Tempos", lwd=2, ylab="S(t)")
lines(c(0,time),c(1,stln), col='red', lwd=2, lty=2)
legend(10,0.2,lty=c(1,2),c("Kaplan-Meier", "Log-normal"),bty="n",cex=0.8)
plot(ekm, conf.int=F, xlab="Tempos", lwd=2, ylab="S(t)")
lines(c(0,time),c(1,stw), col='red', lwd=2, lty=2)
legend(10,0.2,lty=c(1,2),c("Kaplan-Meier", "WG"),bty="n",cex=0.8)
plot(ekm, conf.int=F, xlab="Tempos", lwd=2, ylab="S(t)")
lines(c(0,time),c(1,stwmg), col='red', lwd=2, lty=2)
legend(10,0.2,lty=c(1,2),c("Kaplan-Meier", "WMG"),bty="n",cex=0.8)
plot(ekm, conf.int=F, xlab="Tempos", lwd=2, ylab="S(t)")
lines(c(0,time),c(1,steg), col='red', lwd=2, lty=2)
legend(10,0.2,lty=c(1,2),c("Kaplan-Meier", "EE"),bty="n",cex=0.8)
dev.off()

## gráficos qq-plot
jpeg(filename = 'ct-CM-sCV-f01.jpg', width = 1500, height = 750, pointsize = 12,
      quality = 100, bg = 'white')
par(mar=c(4.5,4.5,0.5,0.5), mfrow=c(2,4), cex=1.3, cex.sub=1.3)
plot(st,stgg,pch=16,ylim=range(c(0.0,1)), xlim=range(c(0,1)), xlab = "S(t):
      Kaplan-Meier", ylab="S(t): GG")
lines(c(0,1), c(0,1), type="l", lty=1, lwd=2)
plot(st,stg,pch=16,ylim=range(c(0.0,1)), xlim=range(c(0,1)), xlab = "S(t):
      Kaplan-Meier", ylab="S(t): Gama")
lines(c(0,1), c(0,1), type="l", lty=1, lwd=2)
plot(st,stw,pch=16,ylim=range(c(0.0,1)), xlim=range(c(0,1)), xlab = "S(t):
      Kaplan-Meier", ylab="S(t): Weibull")
lines(c(0,1), c(0,1), type="l", lty=1, lwd=2)
plot(st,stln,pch=16,ylim=range(c(0.0,1)), xlim=range(c(0,1)), xlab = "S(t):
      Kaplan-Meier", ylab="S(t): log-normal")
lines(c(0,1), c(0,1), type="l", lty=1, lwd=2)
plot(st,stw,pch=16,ylim=range(c(0.0,1)), xlim=range(c(0,1)), xlab = "S(t):
      Kaplan-Meier", ylab="S(t): WG")
lines(c(0,1), c(0,1), type="l", lty=1, lwd=2)
plot(st,stwmg,pch=16,ylim=range(c(0.0,1)), xlim=range(c(0,1)), xlab = "S(t):
      Kaplan-Meier", ylab="S(t): WMG")
lines(c(0,1), c(0,1), type="l", lty=1, lwd=2)
plot(st,steg,pch=16,ylim=range(c(0.0,1)), xlim=range(c(0,1)), xlab = "S(t):
      Kaplan-Meier", ylab="S(t): EE")
lines(c(0,1), c(0,1), type="l", lty=1, lwd=2)
dev.off()

```

B.2.0.3 Enfoque Bayesiano

```

# NAS FUNÇÕES LOGLIK DEVE-SE MUDAR CONFORME A NECESSIDADE
# PARA MODELOS DE MISTURA
#           L[i] <- pow(1-p,d[i])*pow(f0[i],d[i])*pow(p+(1-p)*S0[i],1-d[i])
# PARA MODELOS DE NÃO-MISTURA
#           F0[i] <- pow(1-exp(-alpha*(t[i])),beta)
#           h[i] <- -(log(p))*f0[i]
#           L[i] <- pow(h[i],d[i])*exp(F0[i]*log(p))
# PARA MODELOS SEM FRAÇÃO DE CURA

```

```

#           L[i] <- pow(f0[i],d[i])*pow(S0[i],1-d[i])

# GAMA GENERALIZADA
model{
  for (i in 1:N) {
    f0[i] <- t[i]
    f0[i] ~ dggamma(alpha,beta,gamma)
    S0[i] <- R(f0[i],f0[i])
    L[i] <- pow(f0[i],d[i])*pow(S0[i],1-d[i])
    logL[i] <- log(L[i])
    zeros[i] <-0
    zeros[i] ~ dloglik(logL[i])
    CP0inv[i]<-pow(f0[i],-1)
  }
  # Prior distributions
  alpha ~ dgamma(1,1)
  beta ~ dgamma(1,1)
  gamma ~ dgamma(1,1)
}

# GAMA
model{
  for (i in 1:N) {
    f0[i] <- t[i]
    f0[i] ~ dgamma(alpha,beta)
    S0[i] <- R(f0[i],f0[i])
    L[i] <- pow(f0[i],d[i])*pow(S0[i],1-d[i])
    logL[i] <- log(L[i])
    zeros[i] <-0
    zeros[i] ~ dloglik(logL[i])
    CP0inv[i]<-pow(f0[i],-1)
  }
  # Prior distributions
  alpha ~ dgamma(1,1)
  beta ~ dgamma(1,1)
}

# LOG-NORMAL
model{
  for (i in 1:N) {
    f0[i] <- t[i]
    f0[i] ~ dlnorm(mu,tau)
    S0[i] <- R(f0[i],f0[i])
    L[i] <- pow(f0[i],d[i])*pow(S0[i],1-d[i])
    logL[i] <- log(L[i])
    zeros[i] <-0
    zeros[i] ~ dloglik(logL[i])
    CP0inv[i]<-pow(f0[i],-1)
  }
  # Prior distributions
  mu ~ dnorm(0,1)
  tau ~ dgamma(1,1)
  sigma <- 1/sqrt(tau)
}

```

```

}

# WEIBULL
model{
  for (i in 1:N) {
    f0[i] <- t[i]
    f0[i] ~ dweib(gamma,lambda)
    S0[i] <- R(f0[i],f0[i])
    L[i] <- pow(f0[i],d[i])*pow(S0[i],1-d[i])
    logL[i] <- log(L[i])
    zeros[i] <-0
    zeros[i] ~ dloglik(logL[i])
    CP0inv[i]<-pow(f0[i],-1)
  }
  # Prior distributions
  gamma ~ dgamma(1,1)
  lambda ~ dgamma(1,1)
}

# EXPONENCIAL EXPONENCIADA
model{
  for (i in 1:N) {
    f0[i] <- alpha*beta*exp(-alpha*(t[i]))*pow((1-exp(-alpha*(t[i]))),
      ,(beta-1))
    S0[i] <- 1-pow(1-exp(-alpha*(t[i])),beta)
    L[i] <- pow(f0[i],d[i])*pow(S0[i],1-d[i])
    logL[i] <- log(L[i])
    zeros[i] <-0
    zeros[i] ~ dloglik(logL[i])
    CP0inv[i]<-pow(f0[i],-1)
  }
  # Prior distributions
  alpha ~ dgamma(1,1)
  beta ~ dgamma(1,1)
}

# WEIBULL GENERALIZADA
model{
  for (i in 1:N) {
    f0[i] <- alpha*beta*gama*pow(t[i],(gama-1))*exp(-alpha*pow(t[i],
      gama))*pow((1-exp(-alpha*pow(t[i],gama))), (beta-1))
    S0[i] <- 1-pow((1-exp(-alpha*pow(t[i],gama))),beta)
    L[i] <- pow(f0[i],d[i])*pow(S0[i],1-d[i])
    logL[i] <- log(L[i])
    zeros[i] <-0
    zeros[i] ~ dloglik(logL[i])
    CP0inv[i]<-pow(f0[i],-1)
  }
  # Prior distributions
  alpha ~ dgamma(1,1)
  beta ~ dgamma(1,1)
  gama ~ dgamma(1,1)
}

```

```
# WEIBULL MODIFICADA GENERALIZADA
model{
  for (i in 1:N) {
    f0A[i] <- alpha*beta*pow(t[i],gamma-1)*(gamma+lambda*t[i])*exp(
      lambda*t[i]-alpha*pow(t[i],gamma)*exp(lambda*t[i]))
    f0B[i] <- pow(1-exp(-alpha*pow(t[i],gamma)*exp(lambda*t[i])),1-
      beta)
    f0[i] <-f0A[i]/f0B[i]
    S0[i] <- 1-pow(1-exp(-alpha*pow(t[i],gamma)*exp(lambda*t[i])),beta
      )
    L[i] <- pow(f0[i],d[i])*pow(S0[i],1-d[i])
    logL[i] <- log(L[i])
    zeros[i] <-0
    zeros[i] ~ dloglik(logL[i])
    CP0inv[i]<-pow(f0[i],-1)
  }
  # Prior distributions
  alpha ~ dgamma(1,1)
  beta ~ dgamma(1,1)
  gamma ~ dgamma(1,1)
  lambda ~ dgamma(1,1)
}
```

B.2.0.4 Utilizando covariáveis

```
### Enfoque clássico
phi0 <- param[3]
for(i in 1:6){ phi[i] <- param[i+3] }
upsulon0 <- param[10]
for(i in 1:6){ upsulon[i] <- param[i+10] }
tau0 <- param[17]
for(i in 1:6){ tau[i] <- param[i+17] }
alpha <- (phi0*exp((phi[1]*dados$tumor)+(phi[2]*dados$fxetaria)+(phi[3]*
  dados$leucopenia)+(phi[4]*dados$comorbi)+(phi[5]*dados$desnut)+(phi
  [6]*dados$leucop_desnut)))
lambda <- (upsulon0*exp((upsulon[1]*dados$tumor)+(upsulon[2]*
  dados$fxetaria)+(upsulon[3]*dados$leucopenia)+(upsulon[4]*
  dados$comorbi)+(upsulon[5]*dados$desnut)+(upsulon[6]*
  dados$leucop_desnut)))
pi1 <- (tau0+((tau[1]*dados$tumor)+(tau[2]*dados$fxetaria)+(tau[3]*
  dados$leucopenia)+(tau[4]*dados$comorbi)+(tau[5]*dados$desnut)+(tau
  [6]*dados$leucop_desnut)))
pi=exp(pi1)/(1+exp(pi1))

### Enfoque Bayesiano
alpha[i] <- phi0*exp((phi1*tumor[i])+(phi2*fxetaria[i])+(phi3*
  desnut[i])+(phi4*comorbi[i])+(phi5*leucopenia[i])+(phi6*
  leucop_desnut[i]))
lambda[i] <- upsulon0*exp((upsulon1*tumor[i])+(upsulon2*fxetaria[i]
  )+(upsulon3*desnut[i])+(upsulon4*comorbi[i])+(upsulon5*
  leucopenia[i])+(upsulon6*leucop_desnut[i]))
```

```
pi1[i] <- tau0+((tau1*tumor[i])+(tau2*fxetaria[i])+(tau3*desnut[i]
  ])+(tau4*comorbi[i])+(tau5*leucopenia[i])+(tau6*leucop_desnut[
  i]))
p[i] <- exp(pi1[i])/(1+exp(pi1[i]))
# Prior distributions
tau0 ~ dgamma(1,1)
tau1 ~ dnorm(0,1000)
tau2 ~ dnorm(0,1000)
tau3 ~ dnorm(0,1000)
tau4 ~ dnorm(0,1000)
tau5 ~ dnorm(0,1000)
tau6 ~ dnorm(0,1000)
phi0 ~ dgamma(1,1)
phi1 ~ dnorm(0,1000)
phi2 ~ dnorm(0,1000)
phi3 ~ dnorm(0,1000)
phi4 ~ dnorm(0,1000)
phi5 ~ dnorm(0,1000)
phi6 ~ dnorm(0,1000)
upsulon0 ~ dgamma(1,1)
upsulon1 ~ dnorm(0,1000)
upsulon2 ~ dnorm(0,1000)
upsulon3 ~ dnorm(0,1000)
upsulon4 ~ dnorm(0,1000)
upsulon5 ~ dnorm(0,1000)
upsulon6 ~ dnorm(0,1000)
beta ~ dgamma(1,1)
gama ~ dgamma(1,1)
```

Gráficos da densidade a posteriori, autocorrelação e histórico de séries temporais das amostras simuladas

Neste apêndice é apresentado os gráficos gerados pelo programa OPENBUGS, utilizados para a verificação de convergência das estimações dos modelos utilizados sob o enfoque Bayesiano.

C.1 Pacientes com câncer internados na UTI do INCA

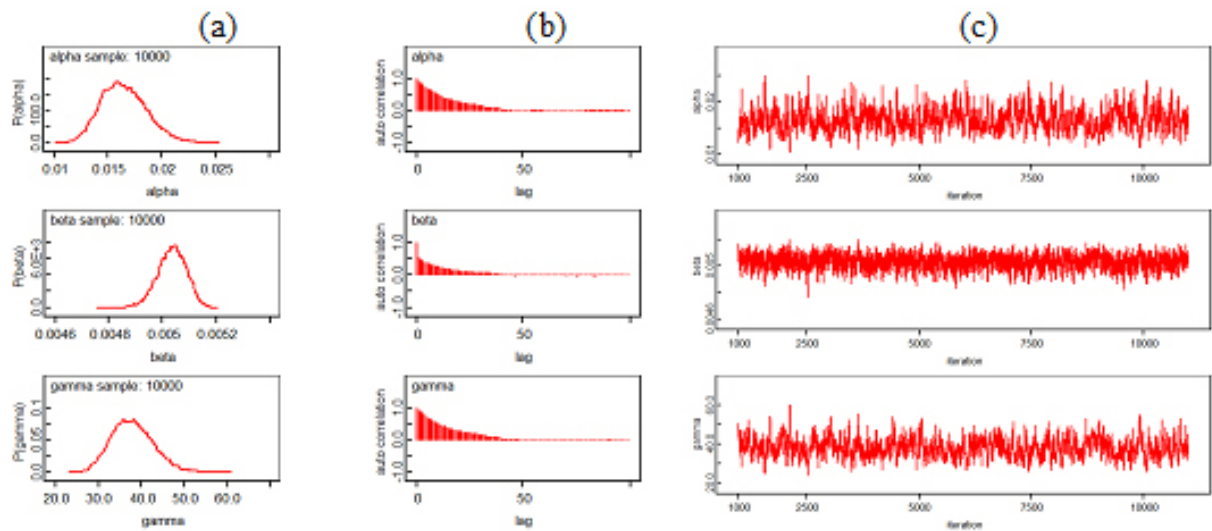


Figura C.1: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição GG sem fração de cura.

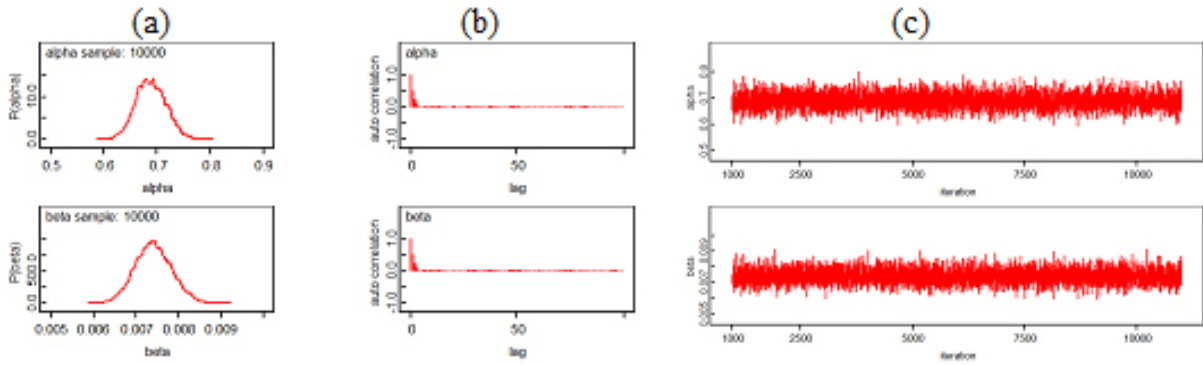


Figura C.2: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição G sem fração de cura.

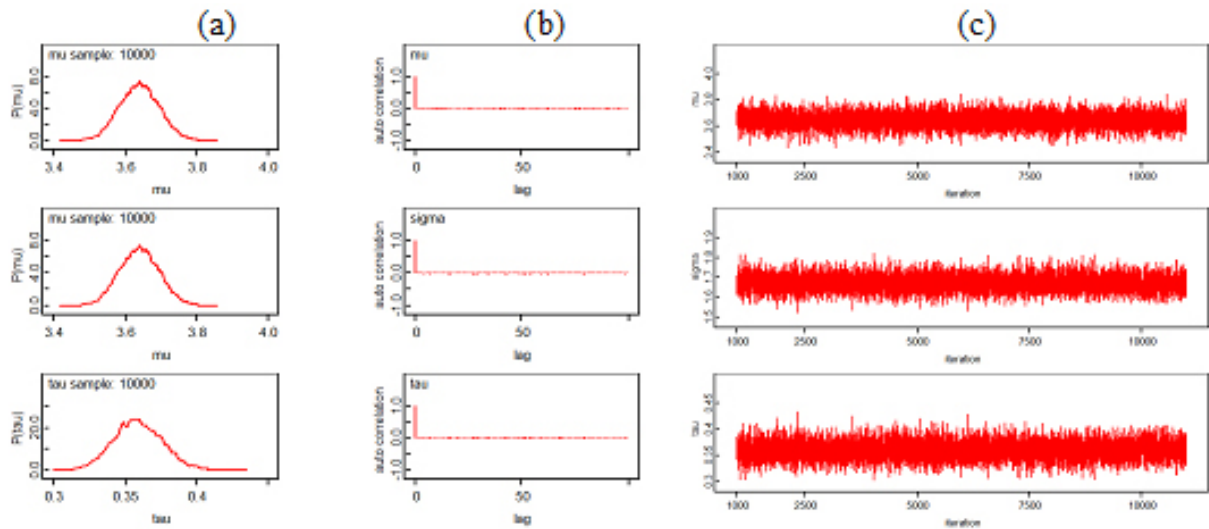


Figura C.3: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição LN sem fração de cura.

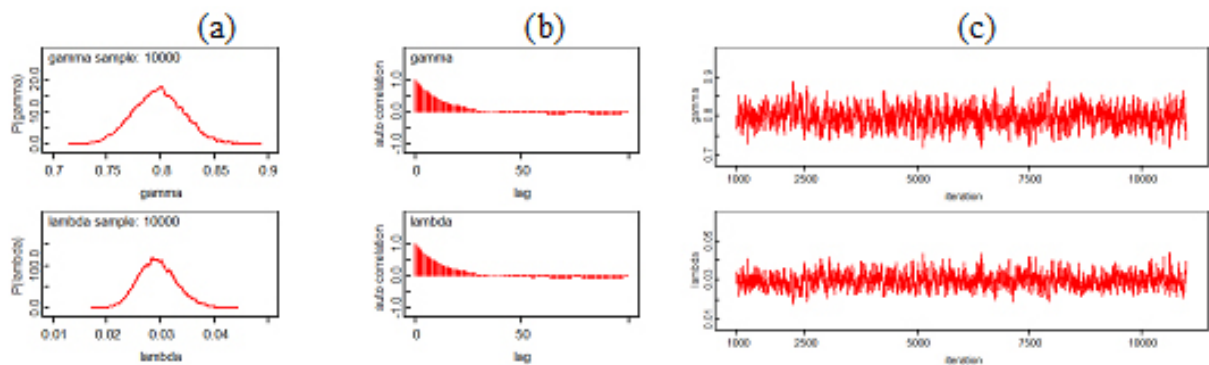


Figura C.4: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição W sem fração de cura.

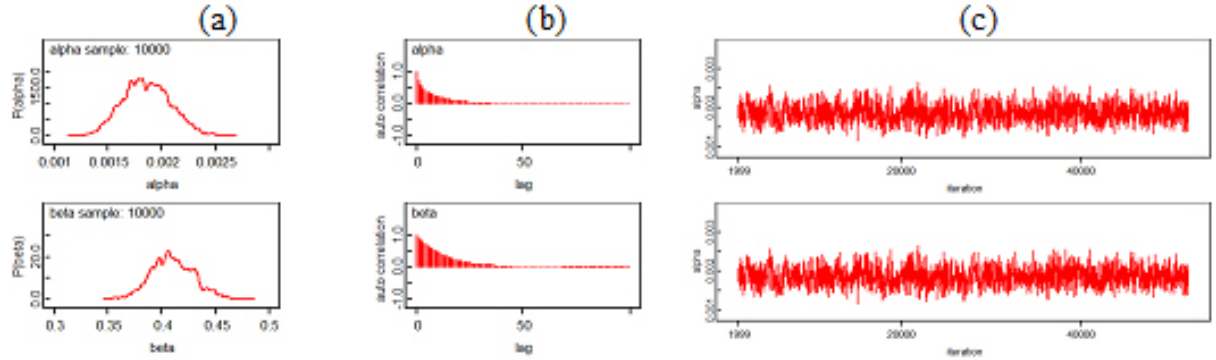


Figura C.5: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição EE sem fração de cura.

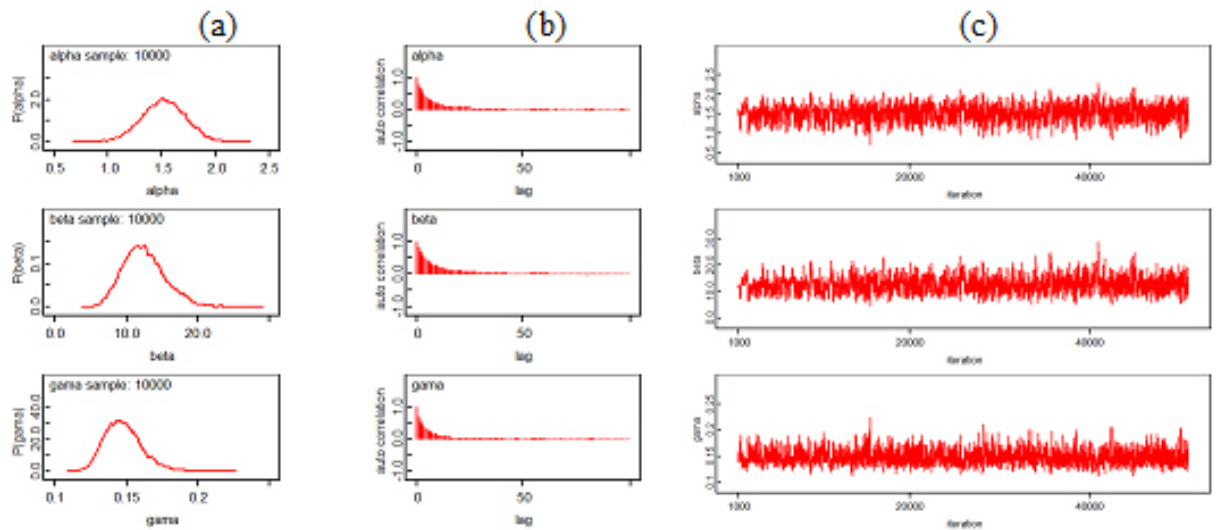


Figura C.6: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WG sem fração de cura.

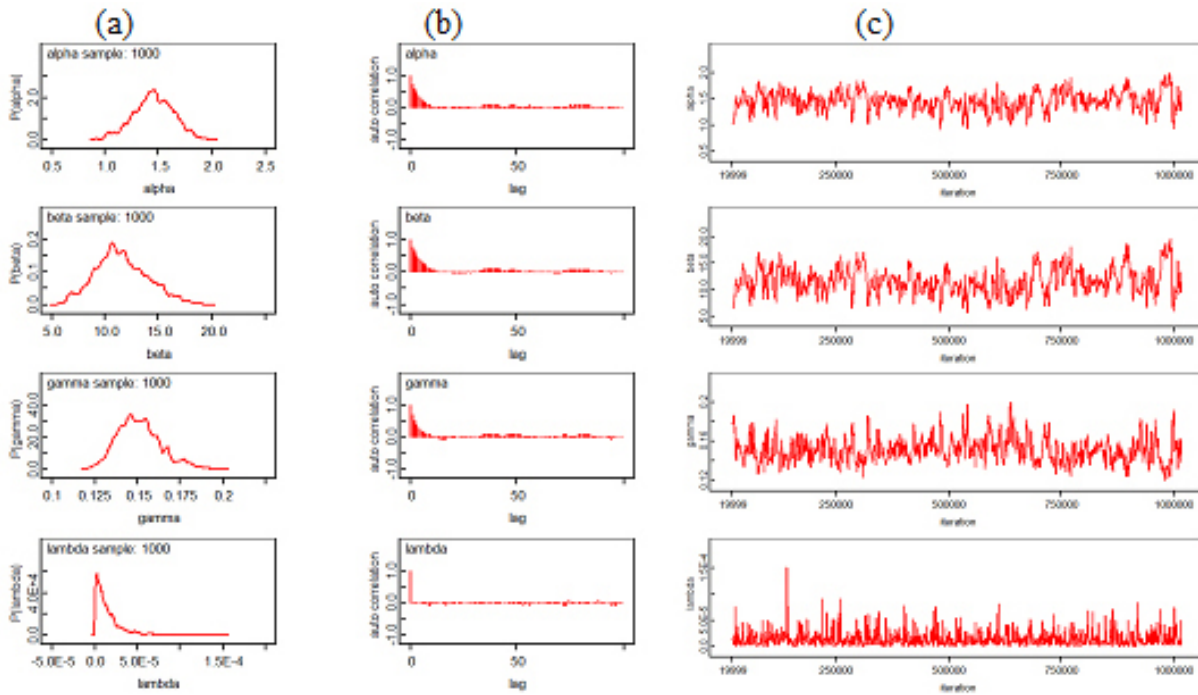


Figura C.7: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WMG sem fração de cura.

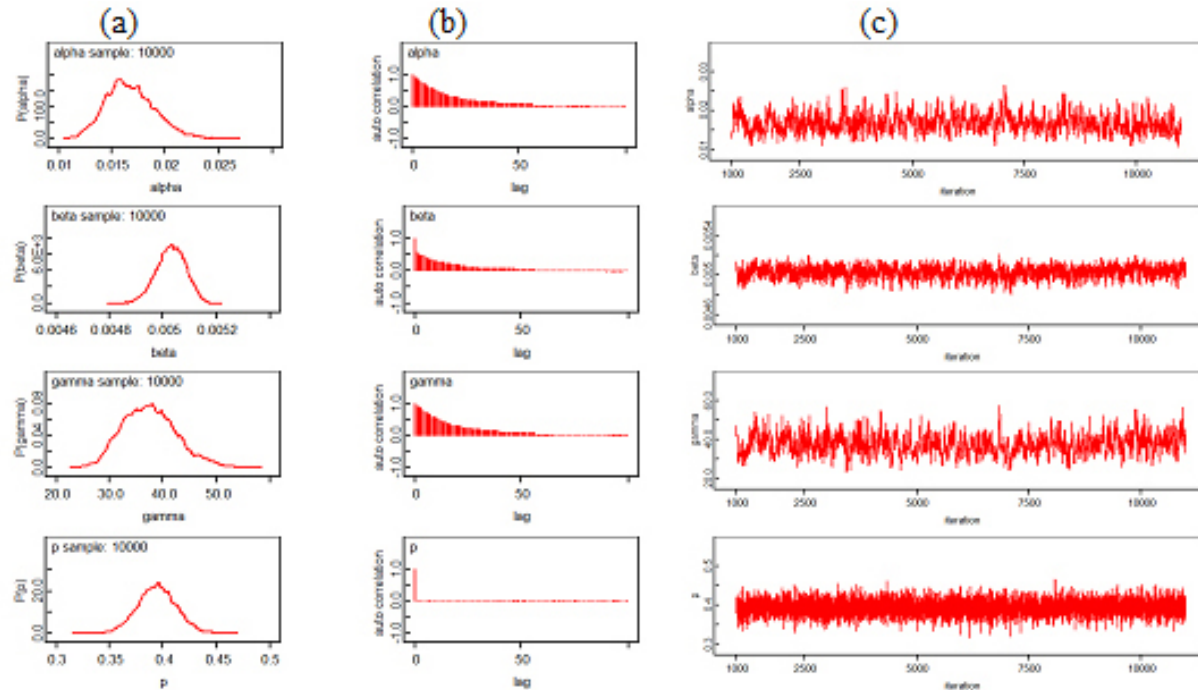


Figura C.8: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição GG de fração de cura de mistura.

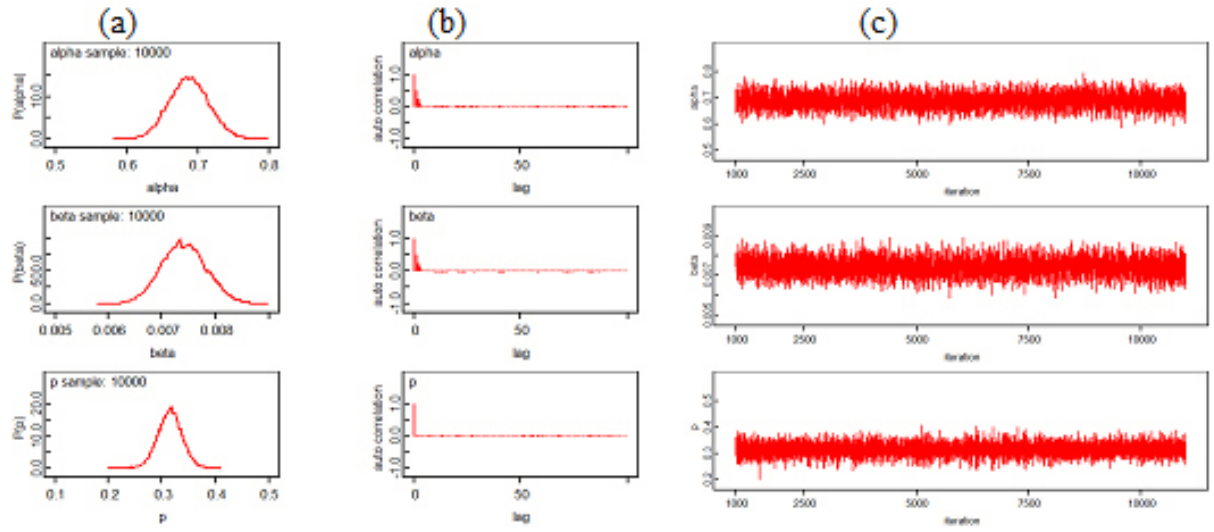


Figura C.9: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição G de fração de cura de mistura.

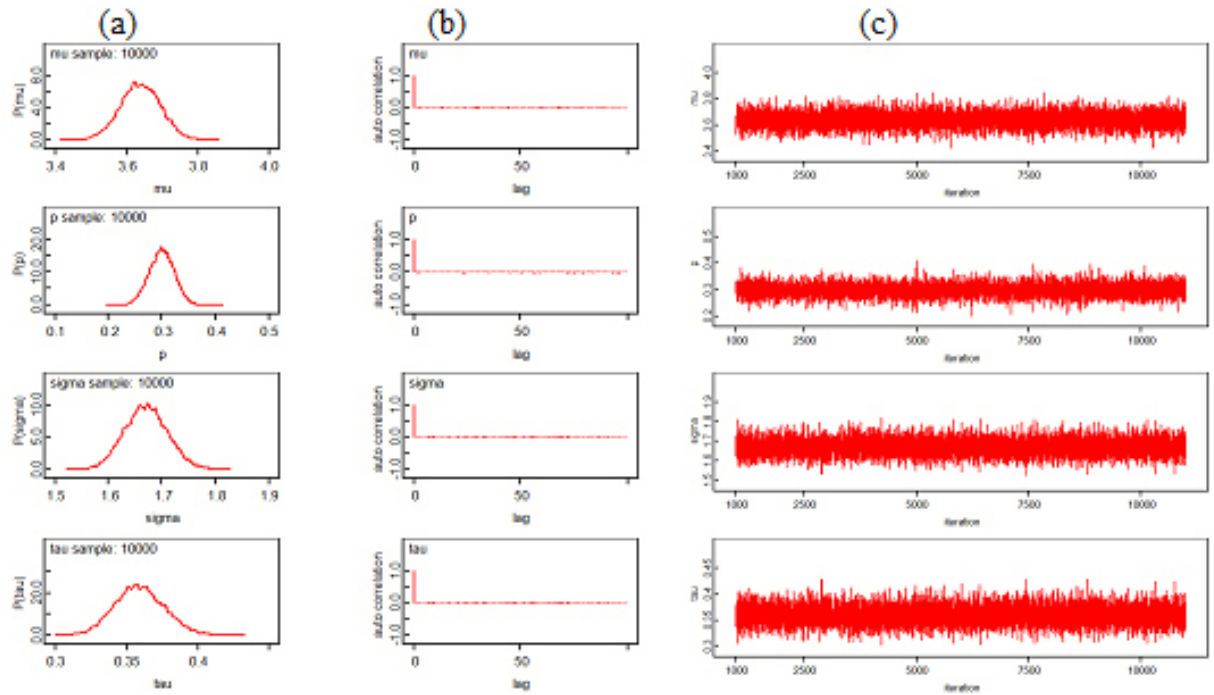


Figura C.10: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição LN de fração de cura de mistura.

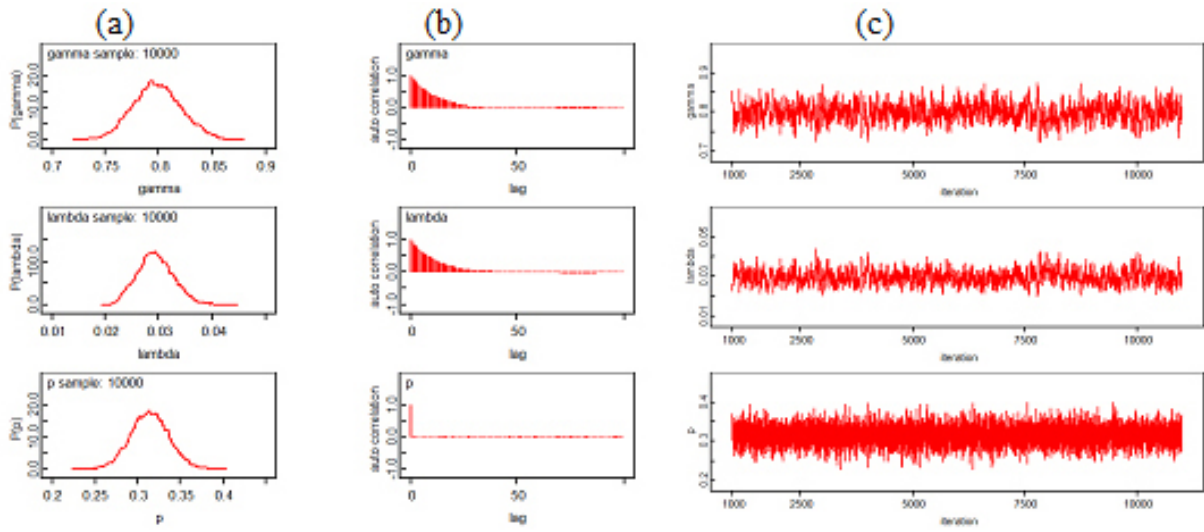


Figura C.11: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição W de fração de cura de mistura.

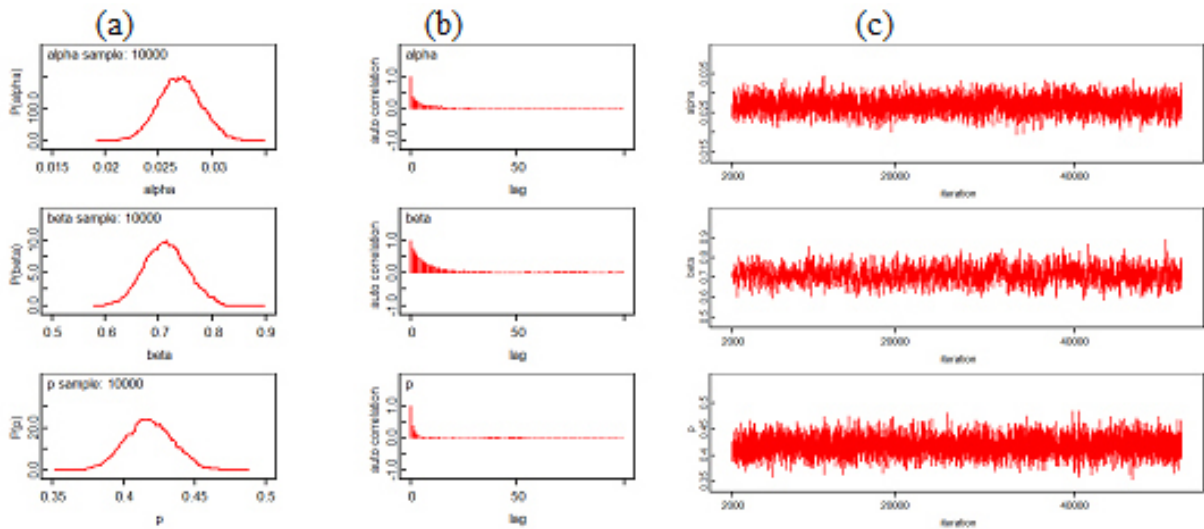


Figura C.12: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição EE de fração de cura de mistura.

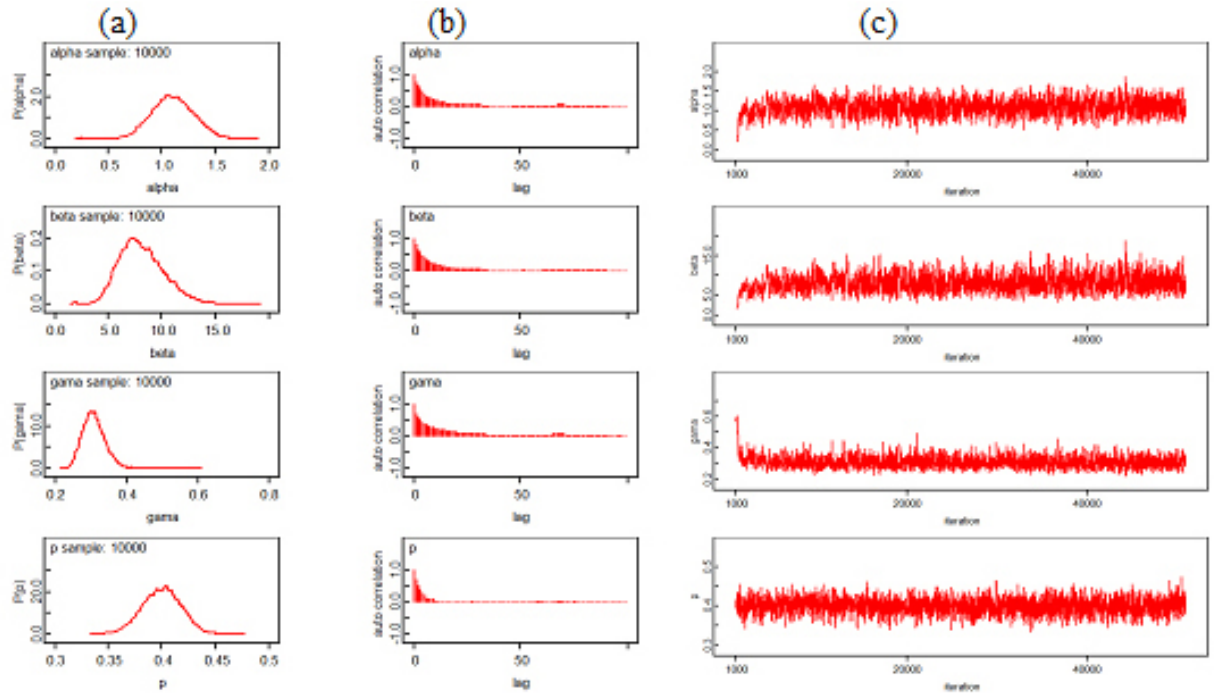


Figura C.13: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WG de fração de cura de mistura.

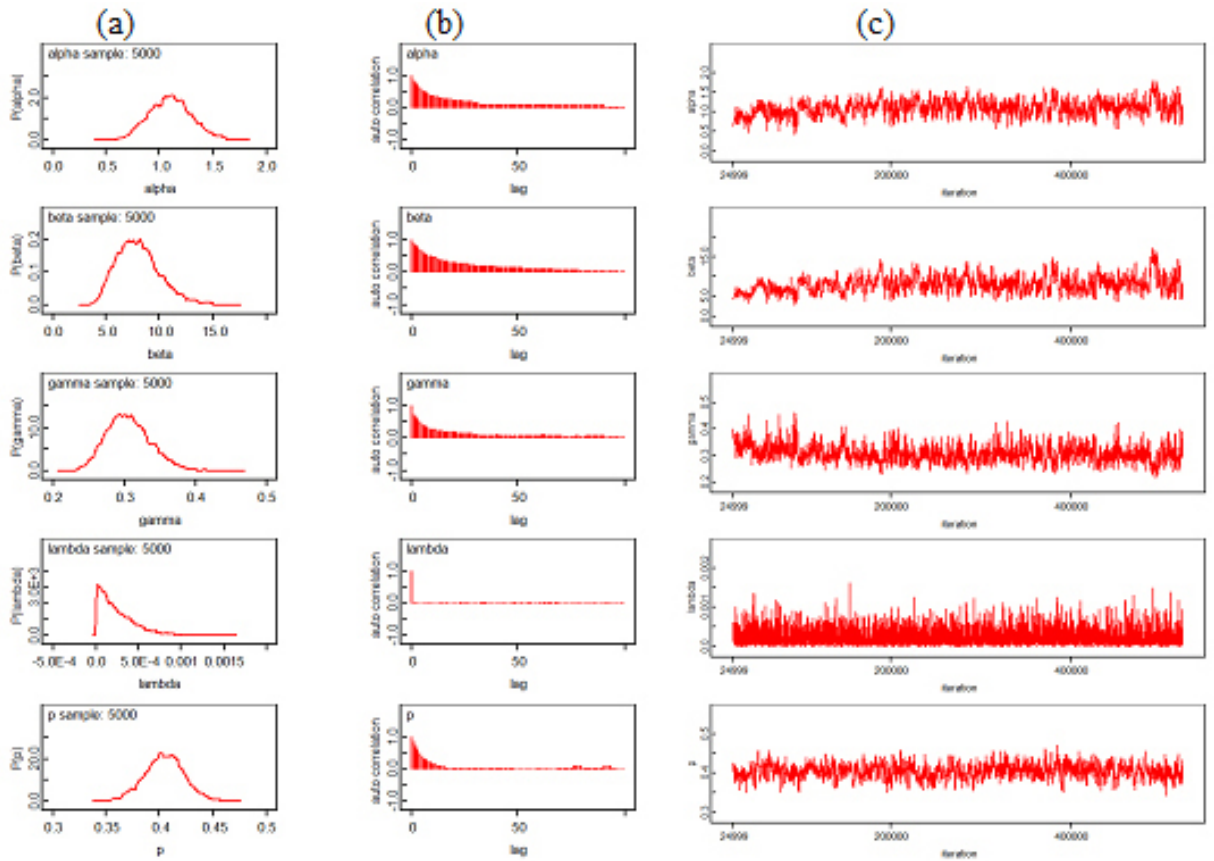


Figura C.14: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WMG de fração de cura de mistura.

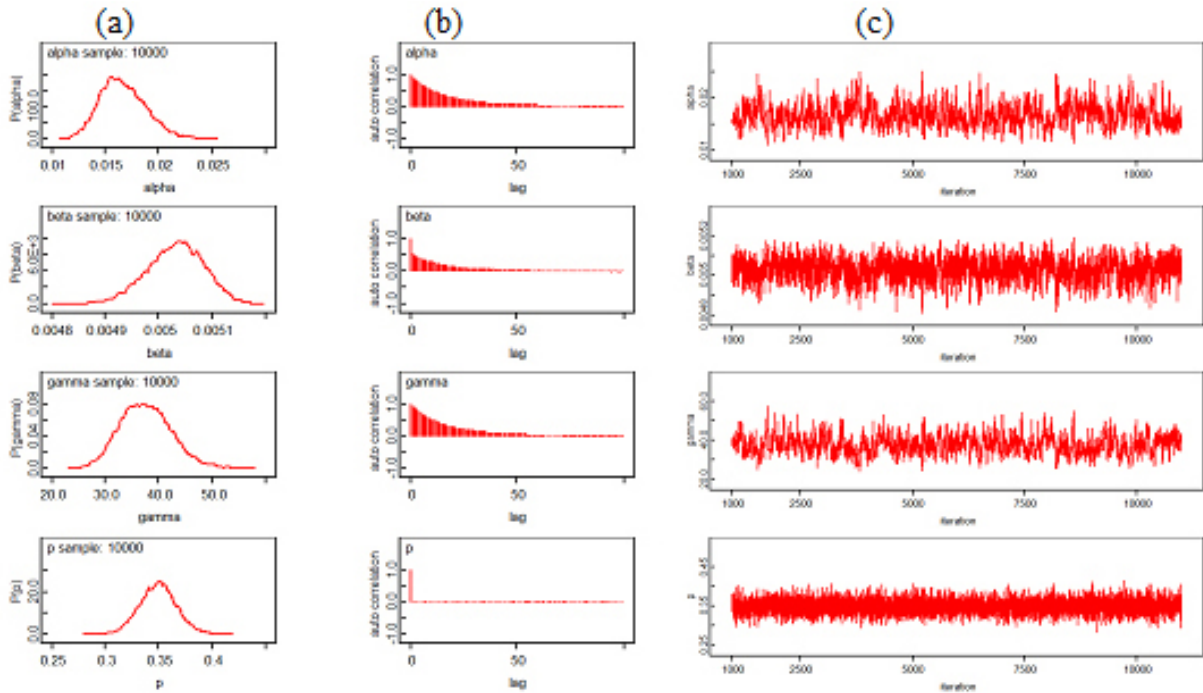


Figura C.15: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição GG de fração de cura de não-mistura.

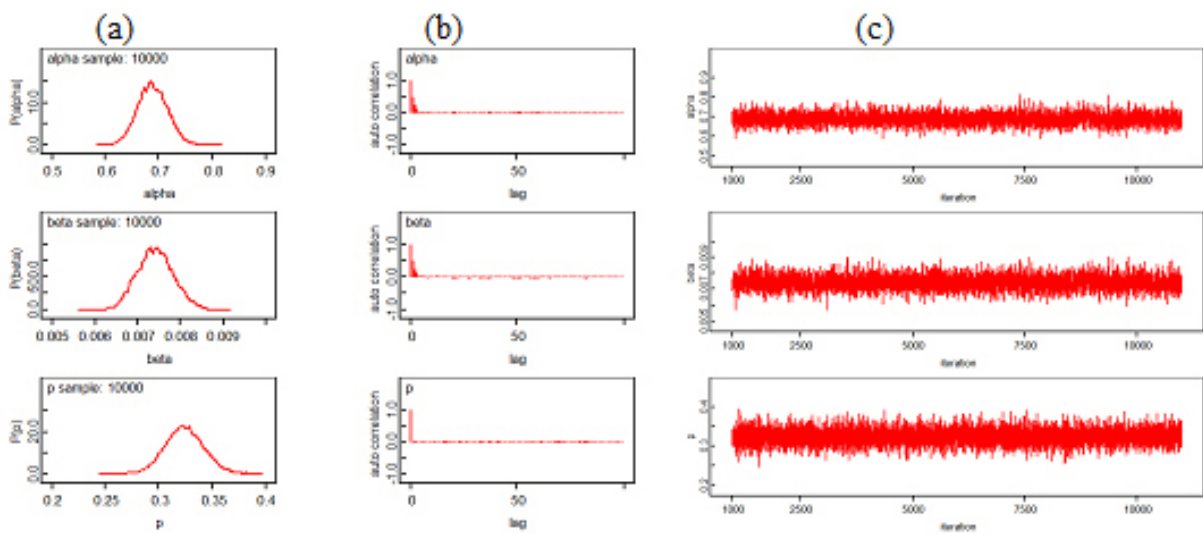


Figura C.16: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição G de fração de cura de não-mistura.

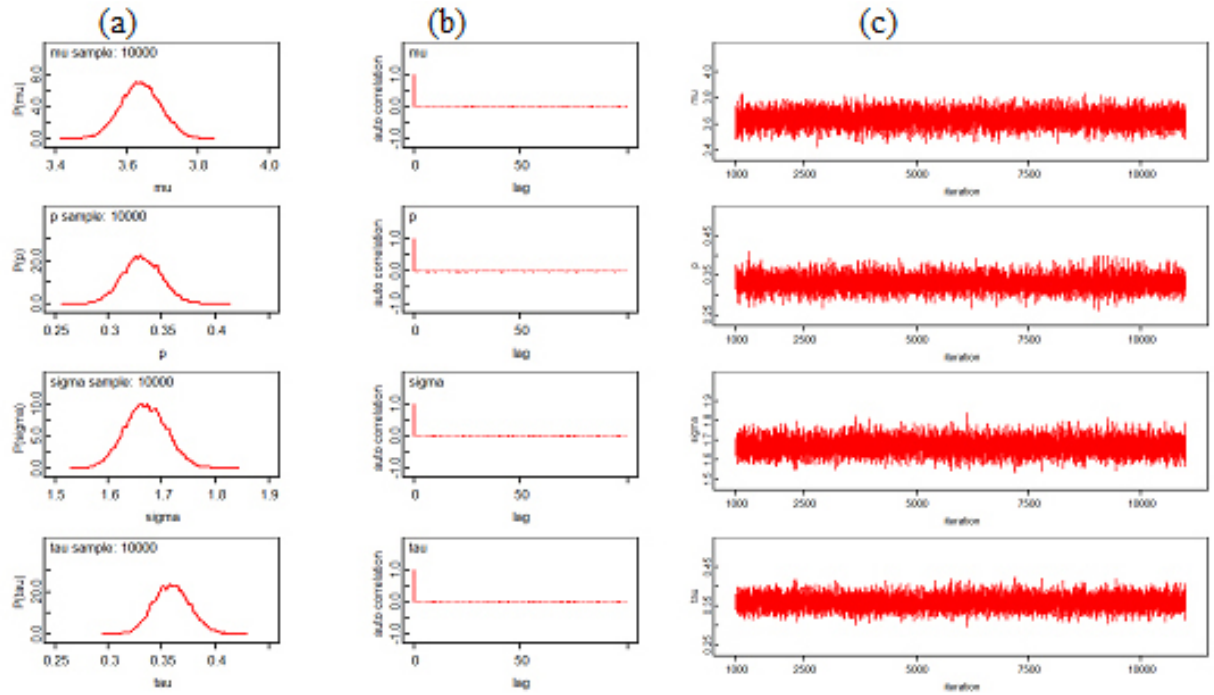


Figura C.17: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição LN de fração de cura de não-mistura.

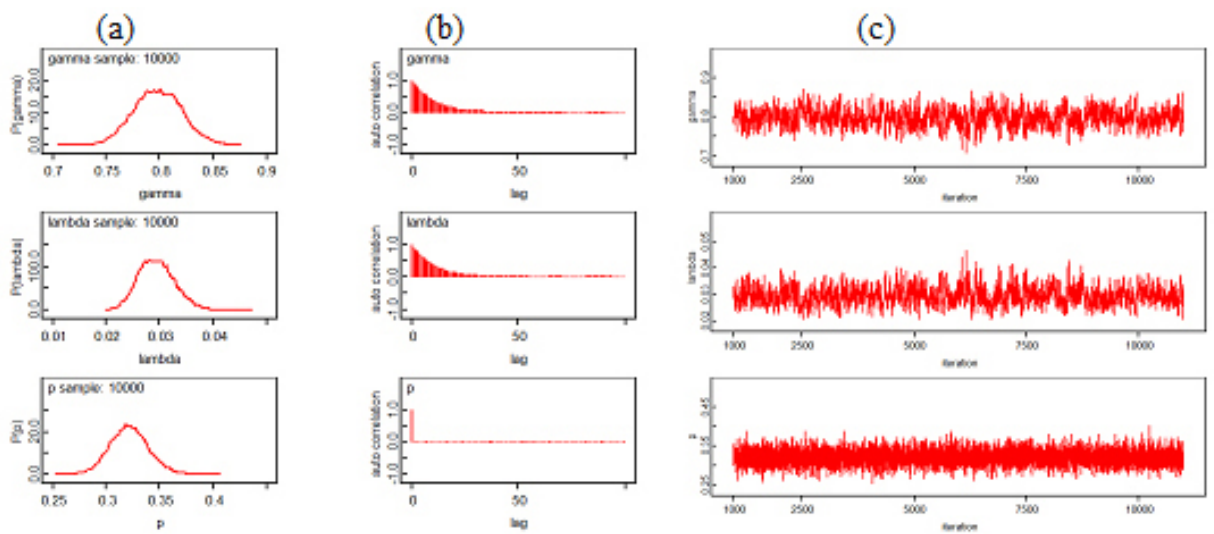


Figura C.18: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição W de fração de cura de não-mistura.

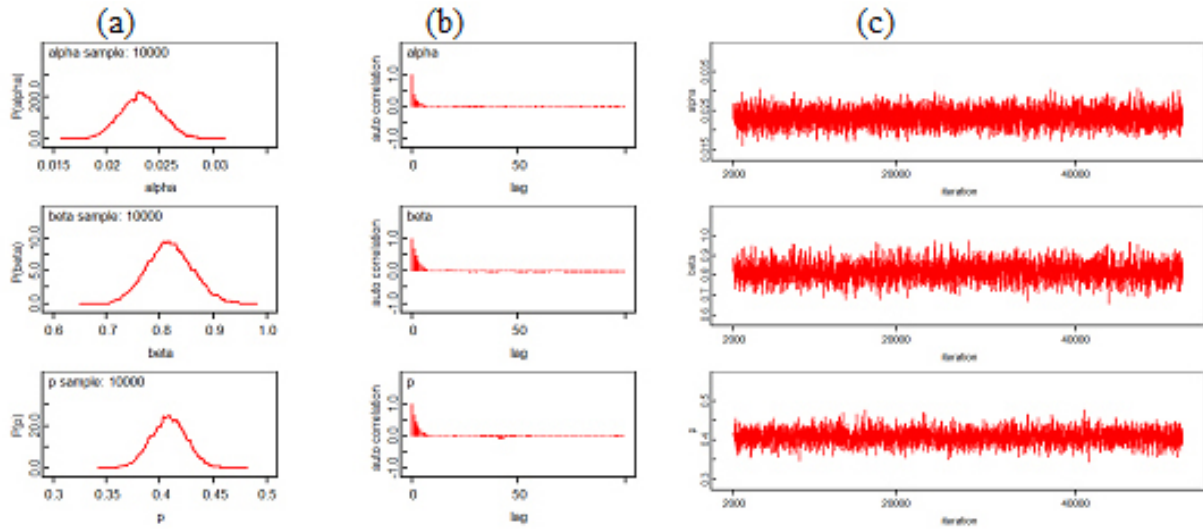


Figura C.19: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição EE de fração de cura de não-mistura.

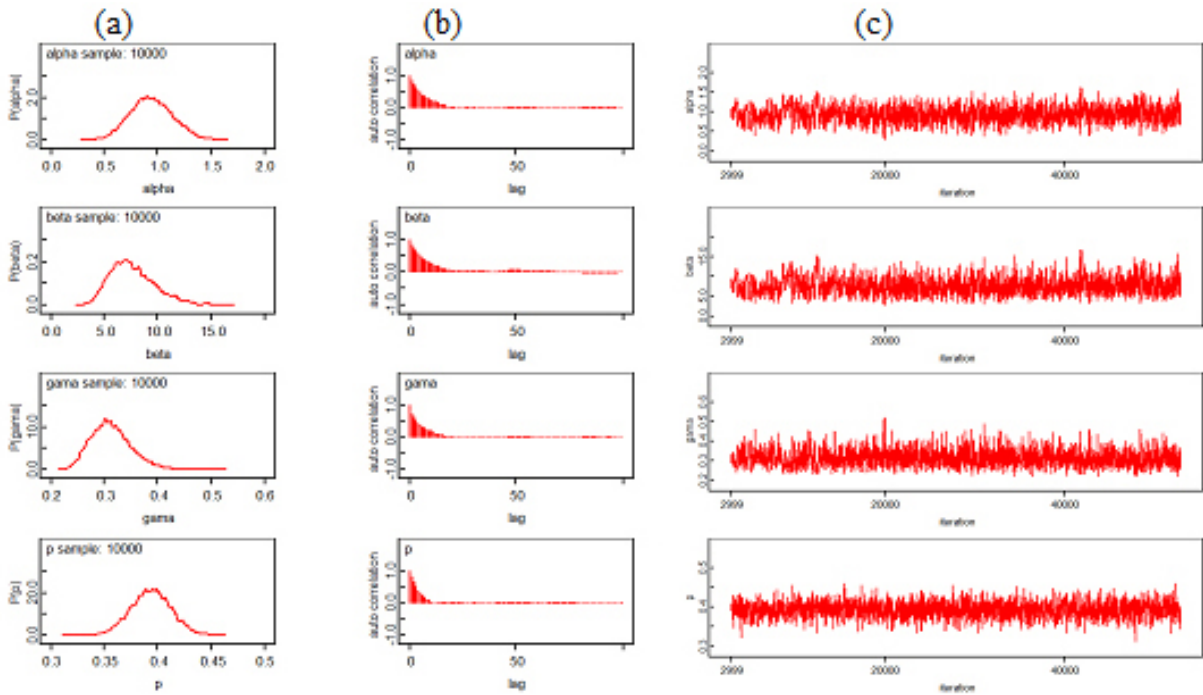


Figura C.20: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WG de fração de cura de não-mistura.

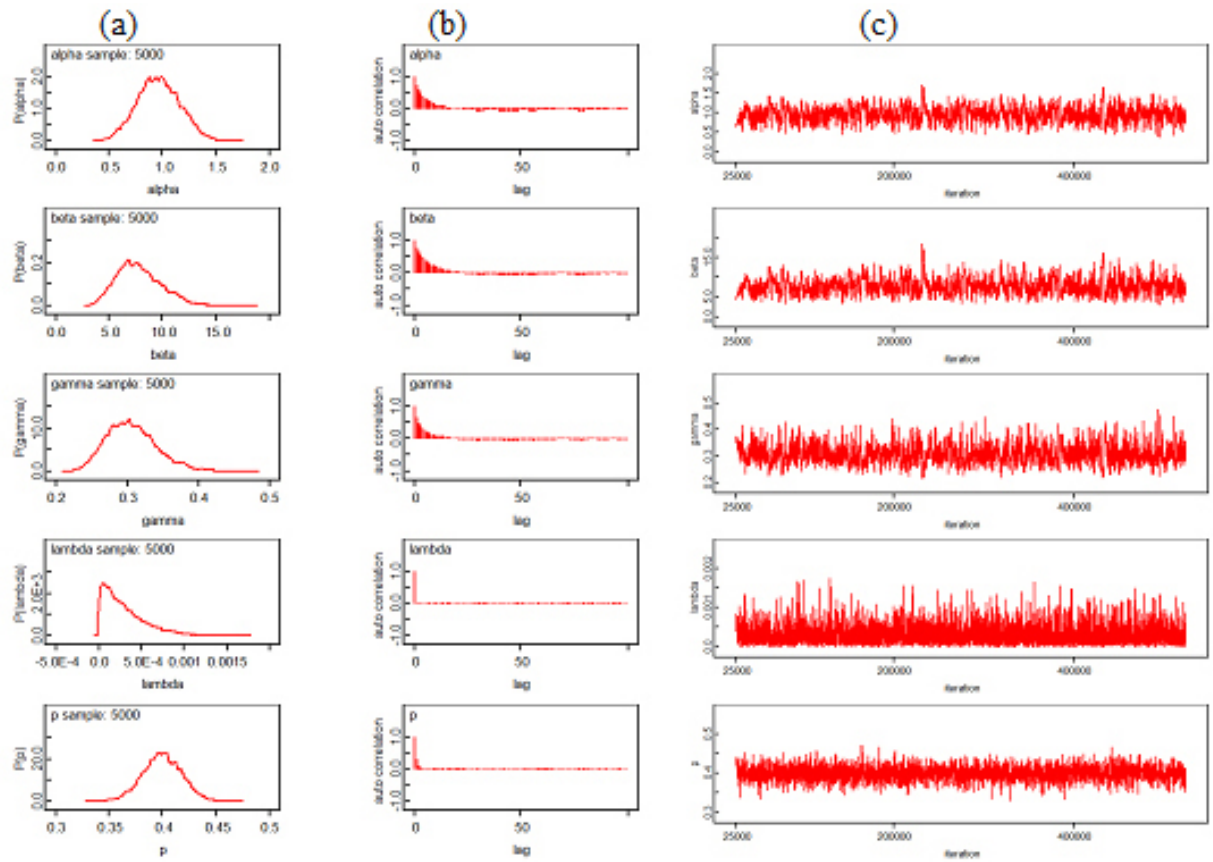


Figura C.21: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WMG de fração de cura de não-mistura.

C.2 Dados de câncer gástrico do Hospital do Câncer de Barretos

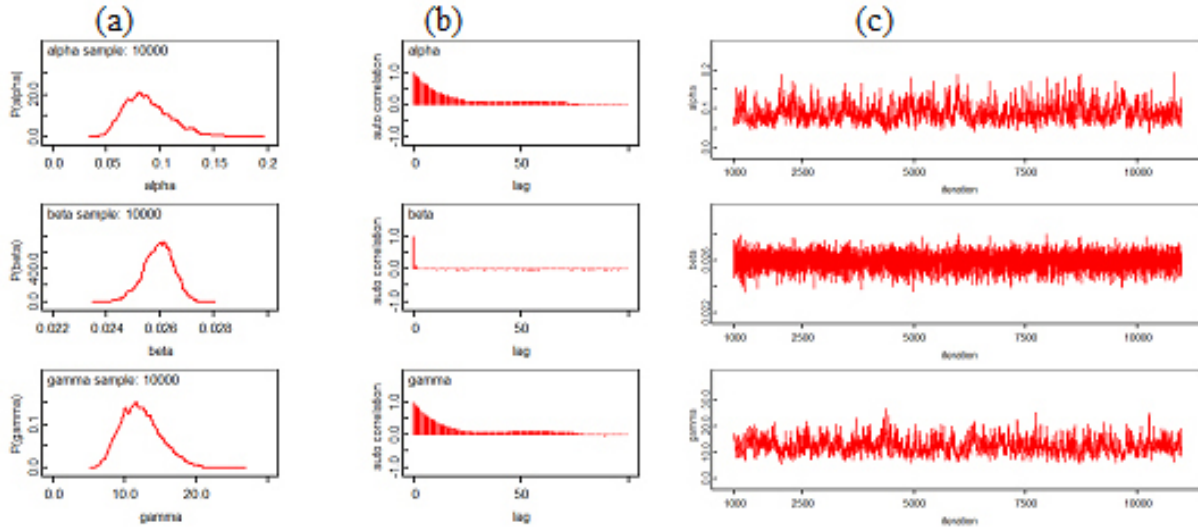


Figura C.22: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição GG sem fração de cura.

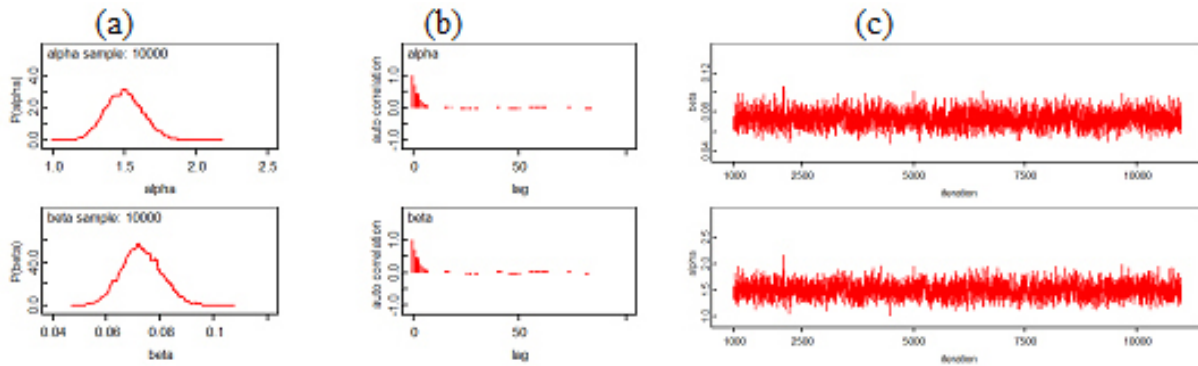


Figura C.23: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição G sem fração de cura.

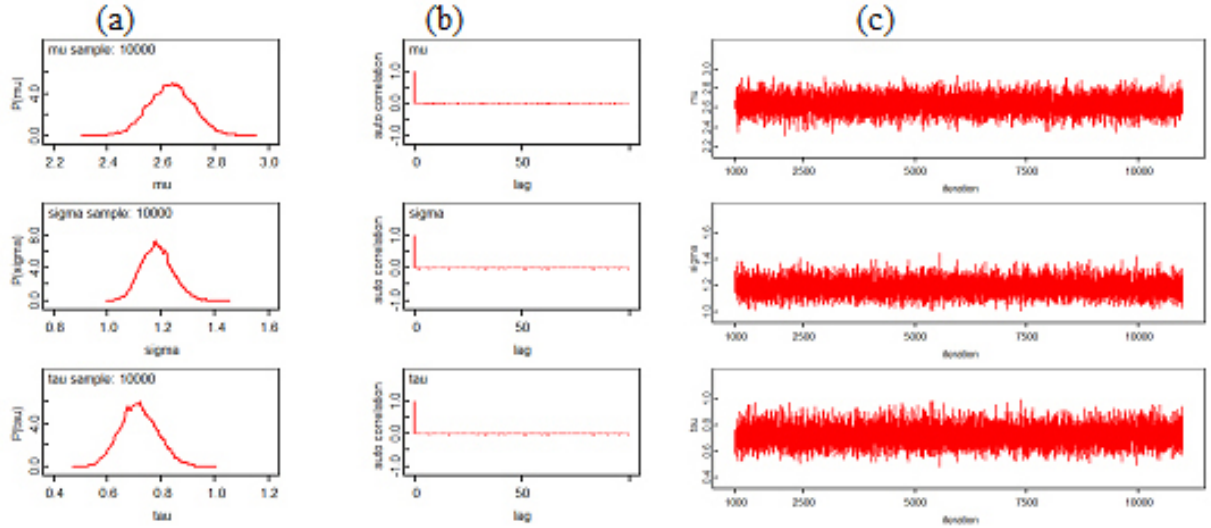


Figura C.24: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição LN sem fração de cura.

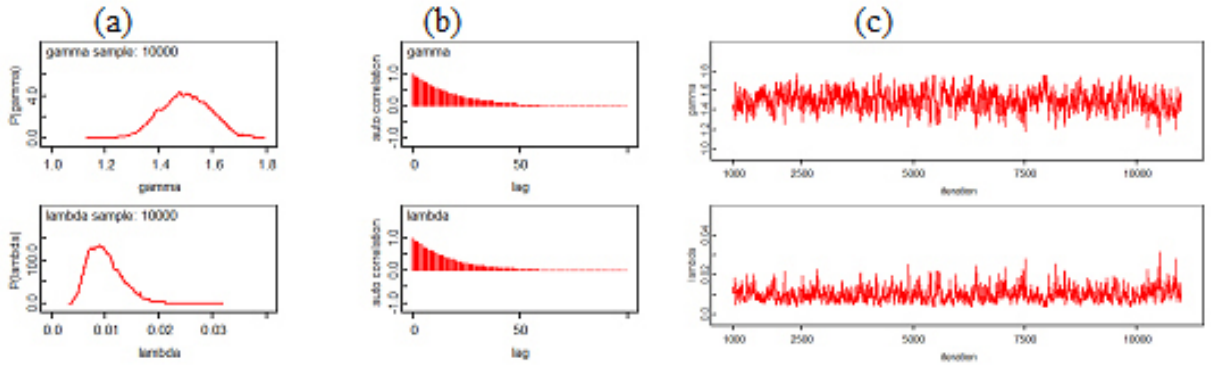


Figura C.25: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição W sem fração de cura.

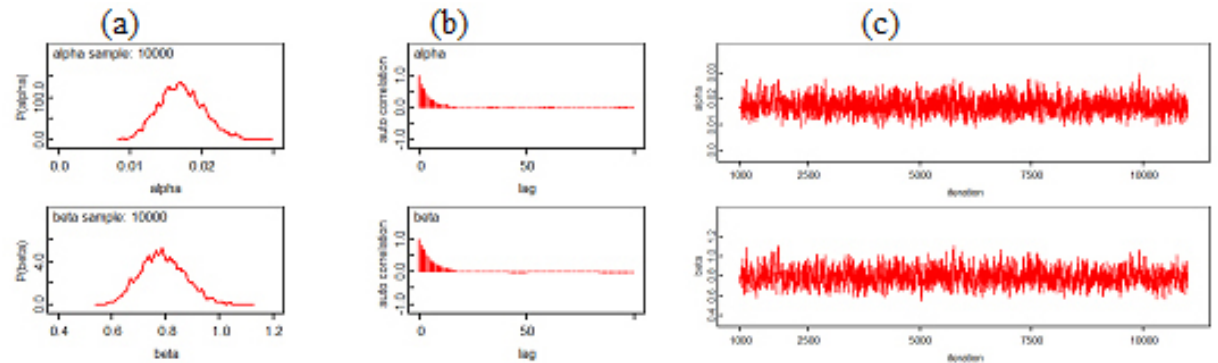


Figura C.26: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição EE sem fração de cura.

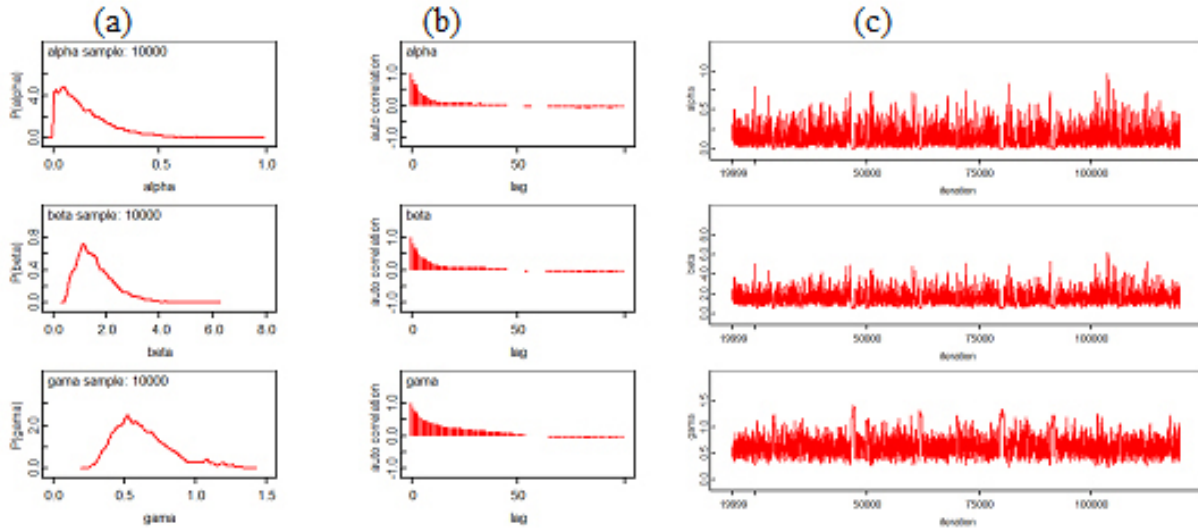


Figura C.27: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WG sem fração de cura.

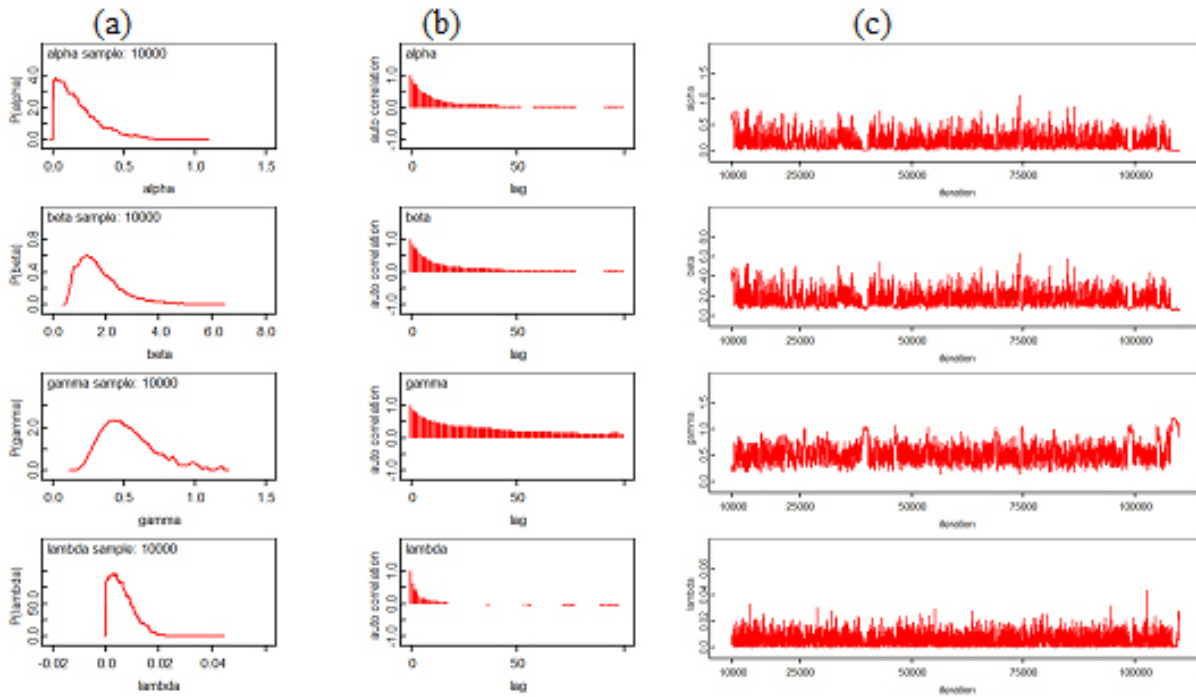


Figura C.28: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WMG sem fração de cura.

[79] [8] [9] [10] [11] [18] [20] [22] [43] [44] [47] [3] [67] [74] [6] [7] [21] [24] [59] [64] [71] [80]
[33] [66] [75] [70] [69] [65] [56] [57] [76] [72] [62] [68] [50] [27] [32] [58]

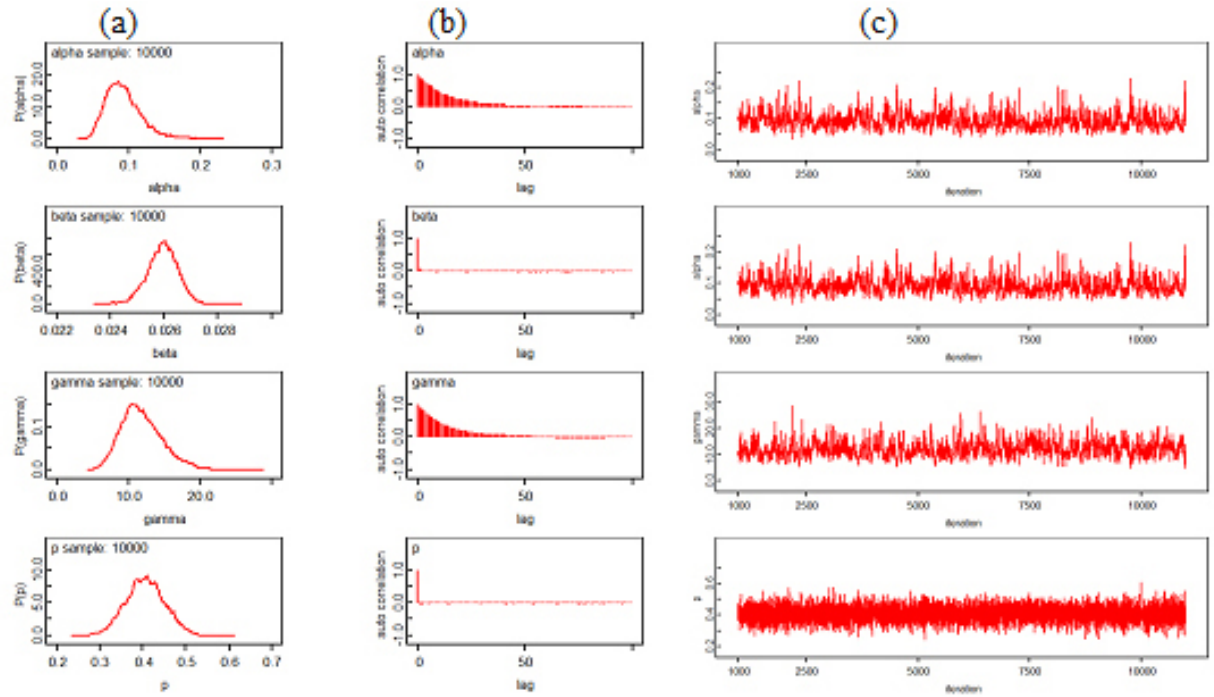


Figura C.29: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição GG de fração de cura de mistura.

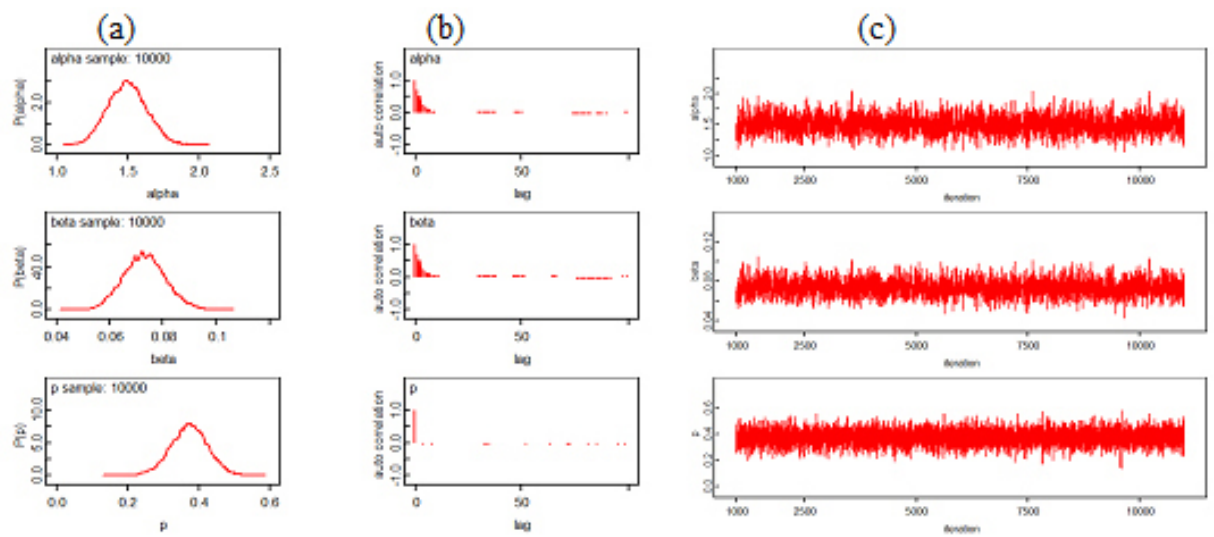


Figura C.30: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição G de fração de cura de mistura.

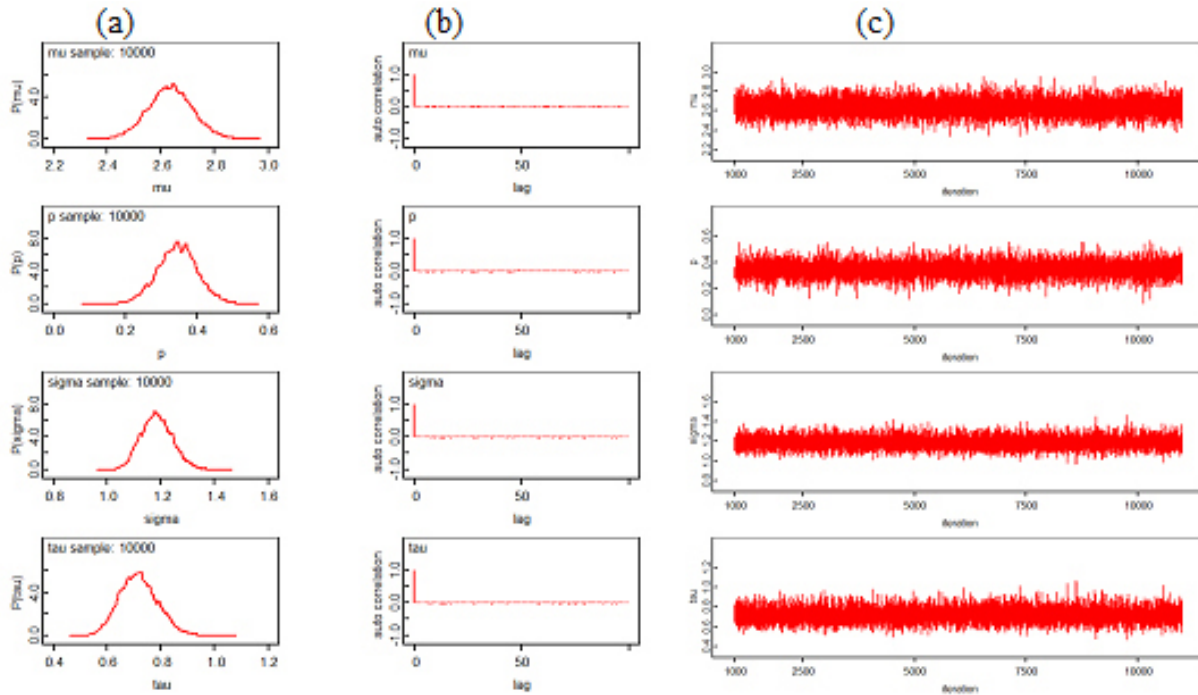


Figura C.31: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição LN de fração de cura de mistura.

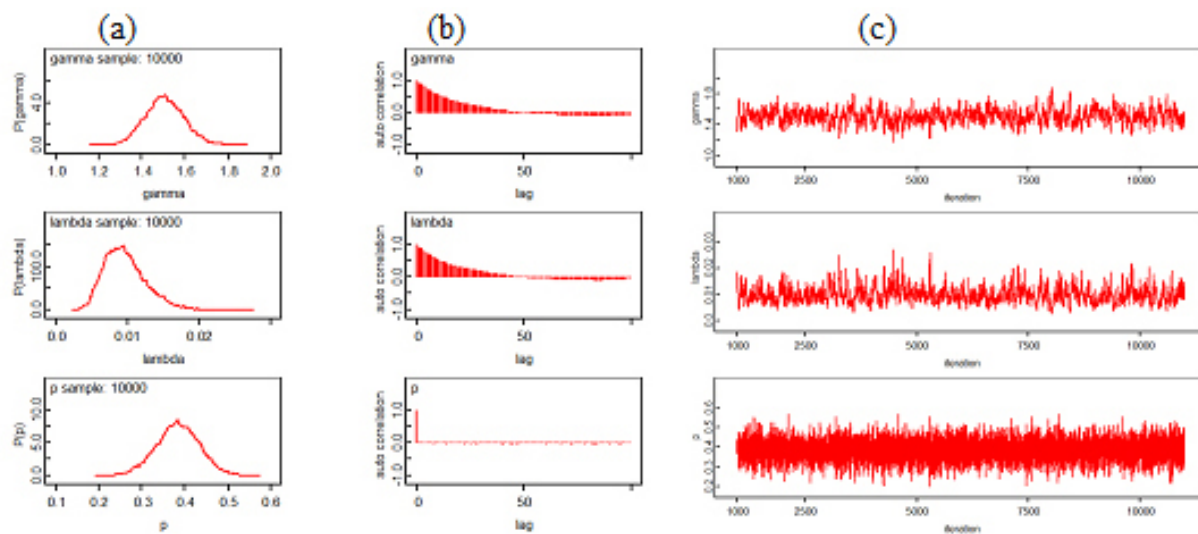


Figura C.32: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição W de fração de cura de mistura.

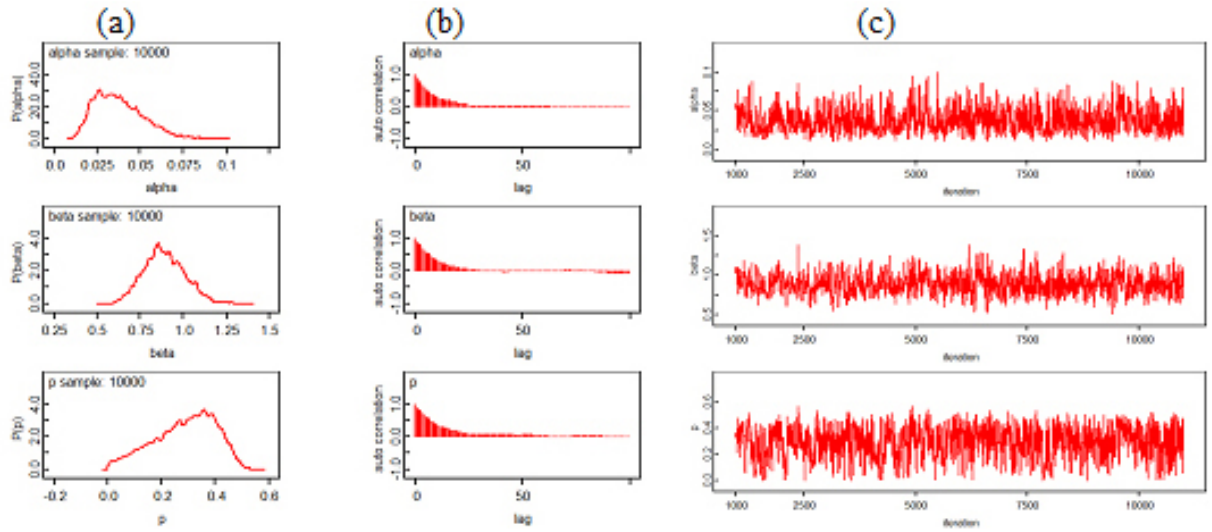


Figura C.33: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição EE de fração de cura de mistura.

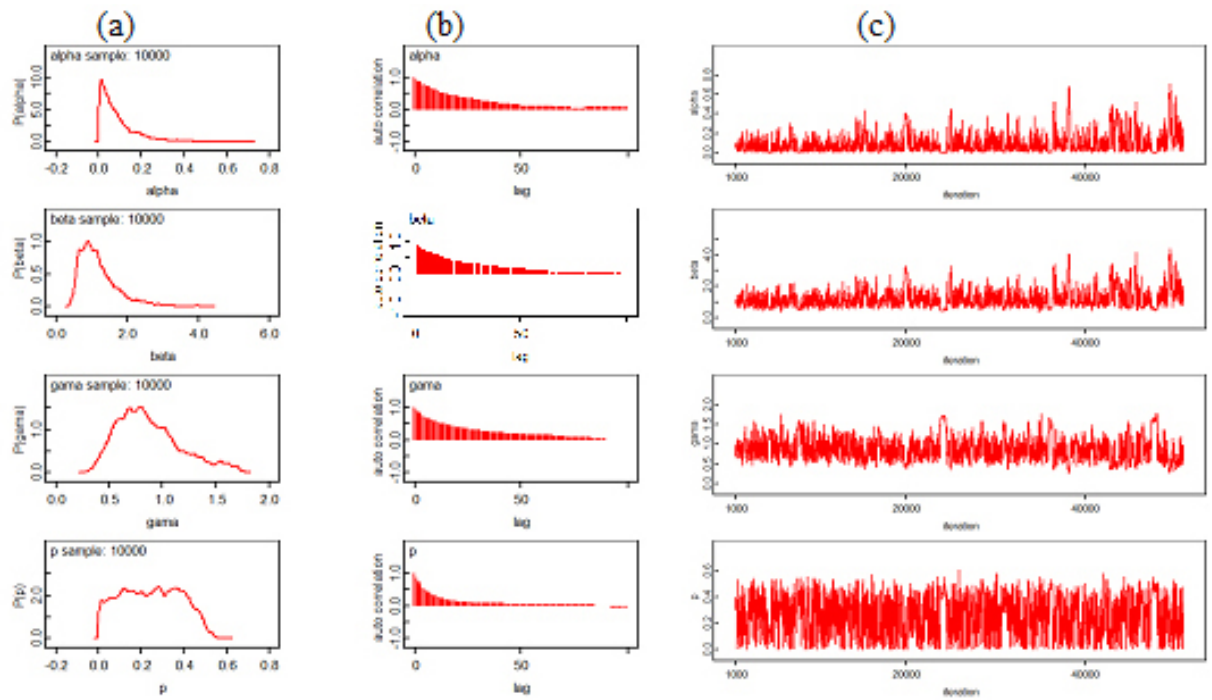


Figura C.34: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WG de fração de cura de mistura.

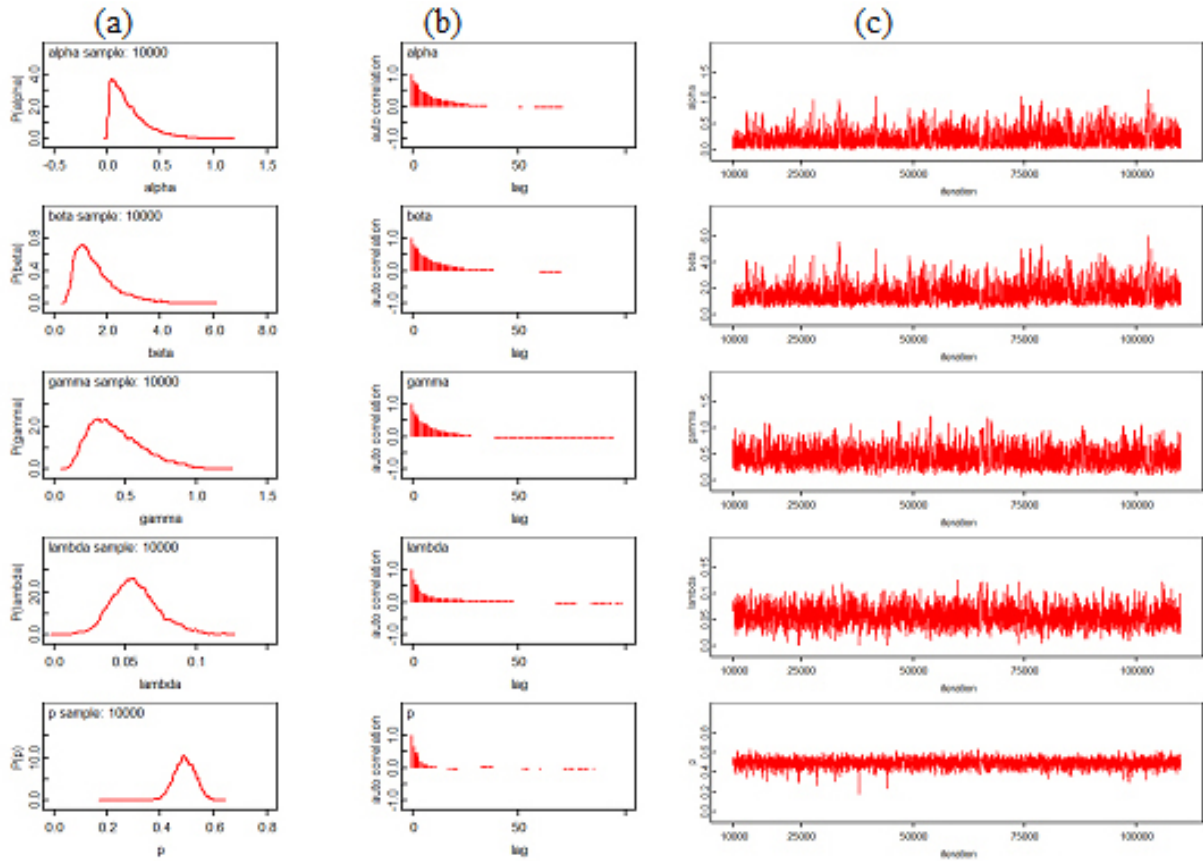


Figura C.35: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WMG de fração de cura de mistura.

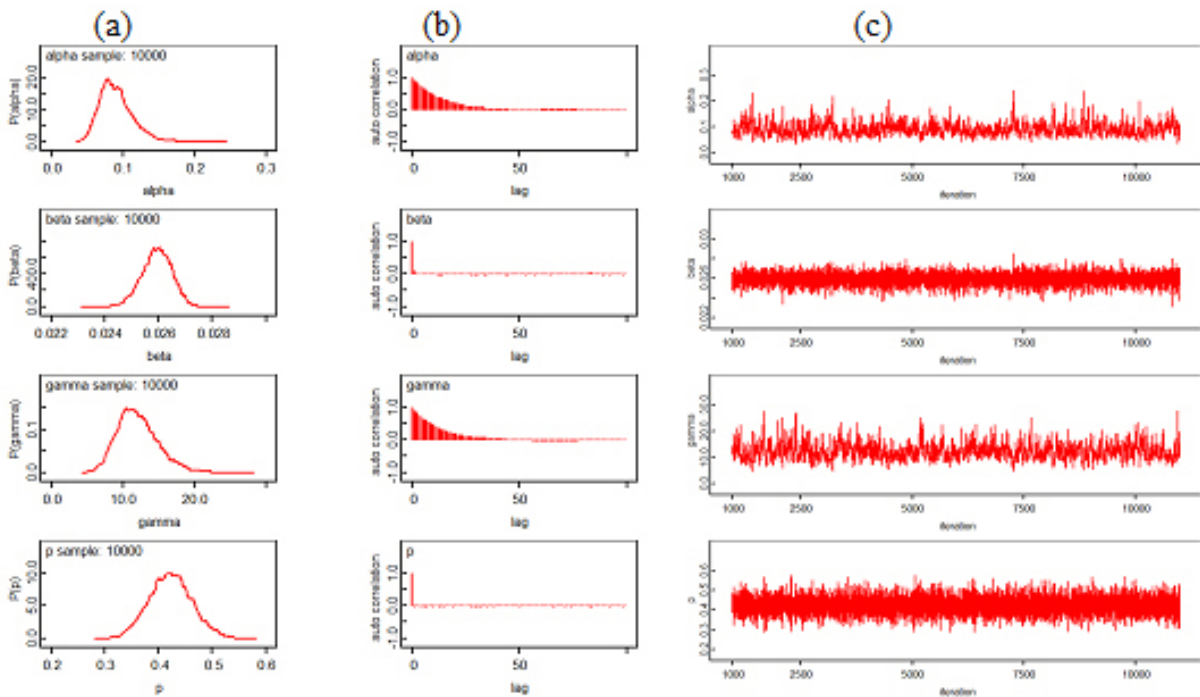


Figura C.36: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição GG de fração de cura de não-mistura.

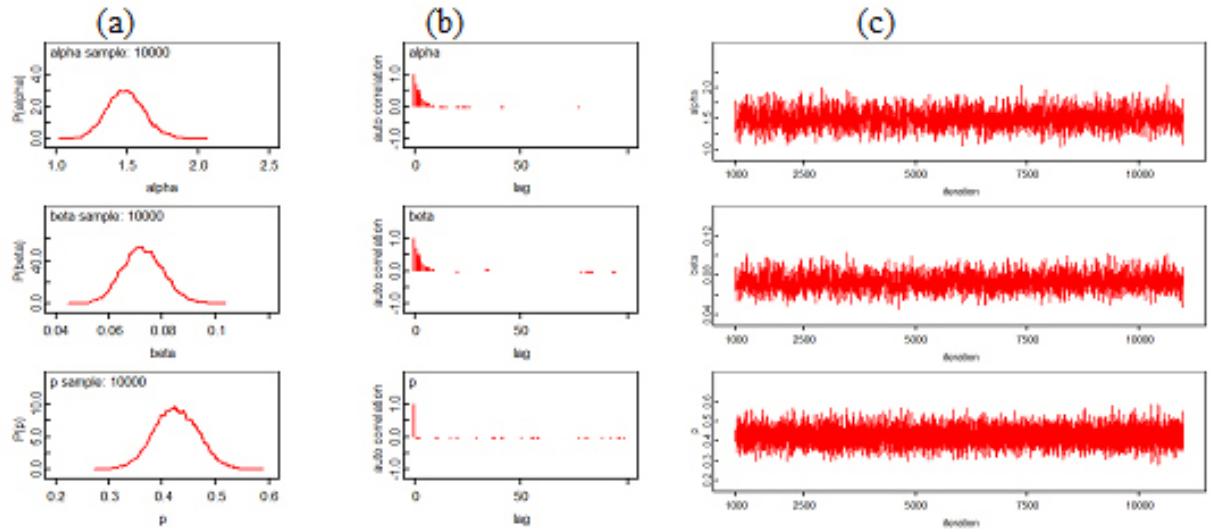


Figura C.37: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição G de fração de cura de não-mistura.

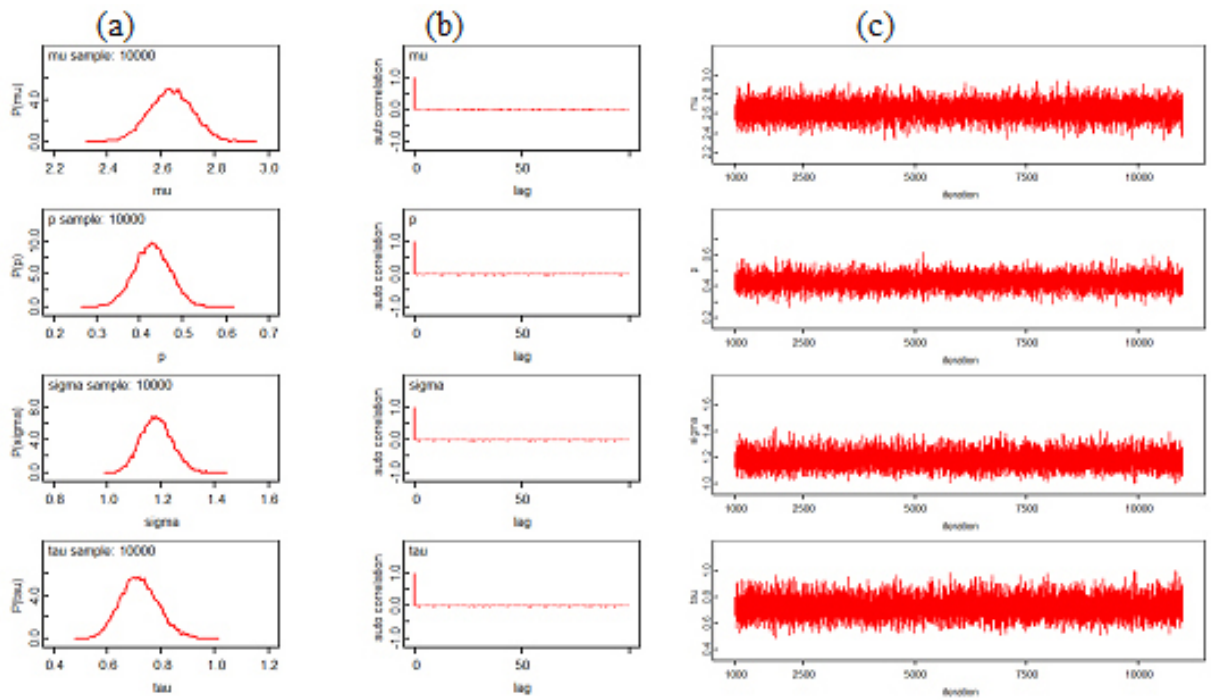


Figura C.38: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição LN de fração de cura de não-mistura.

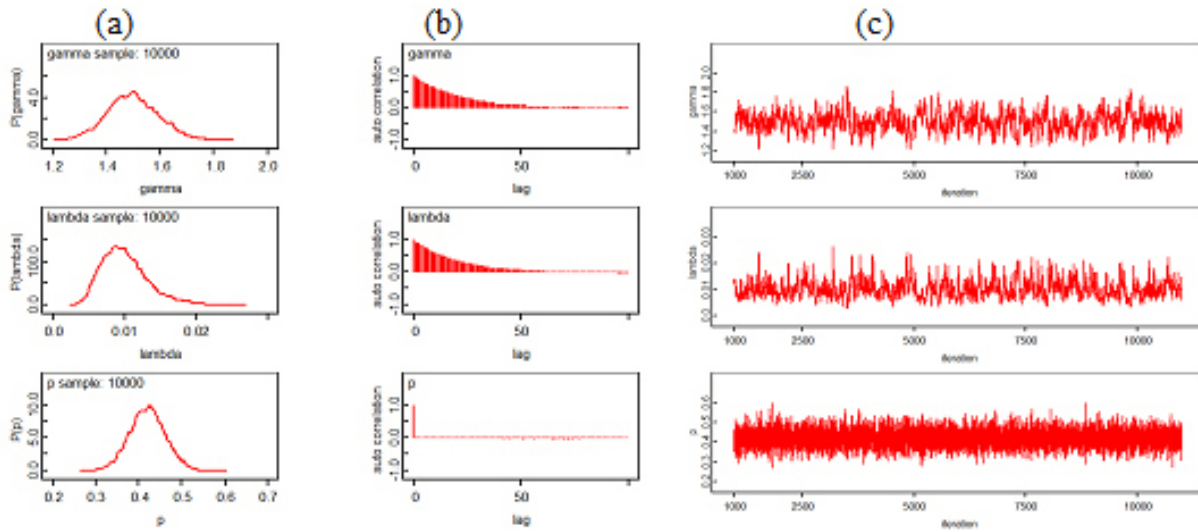


Figura C.39: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição W de fração de cura de não-mistura.

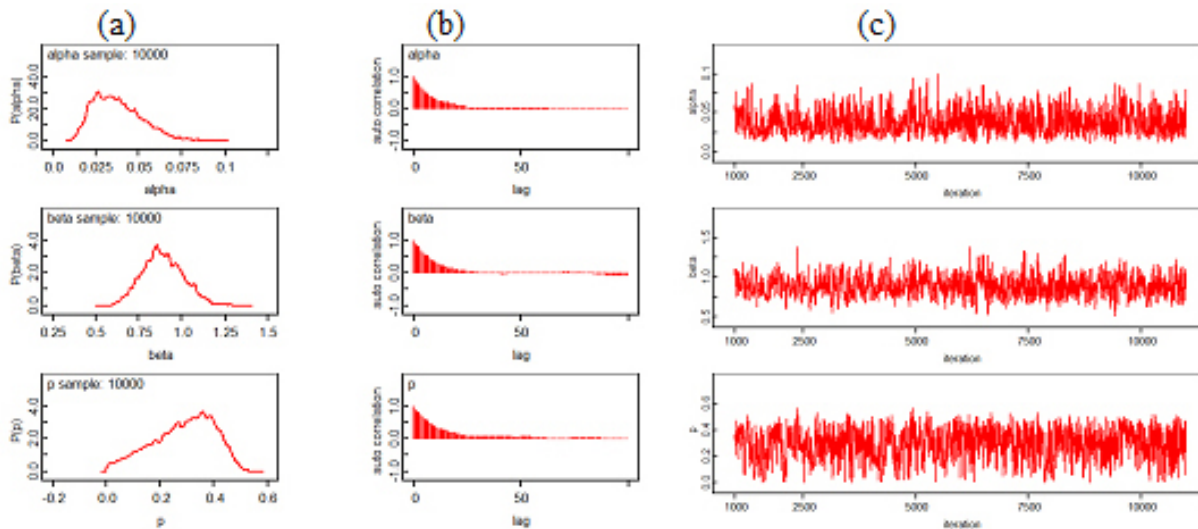


Figura C.40: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição EE de fração de cura de não-mistura.

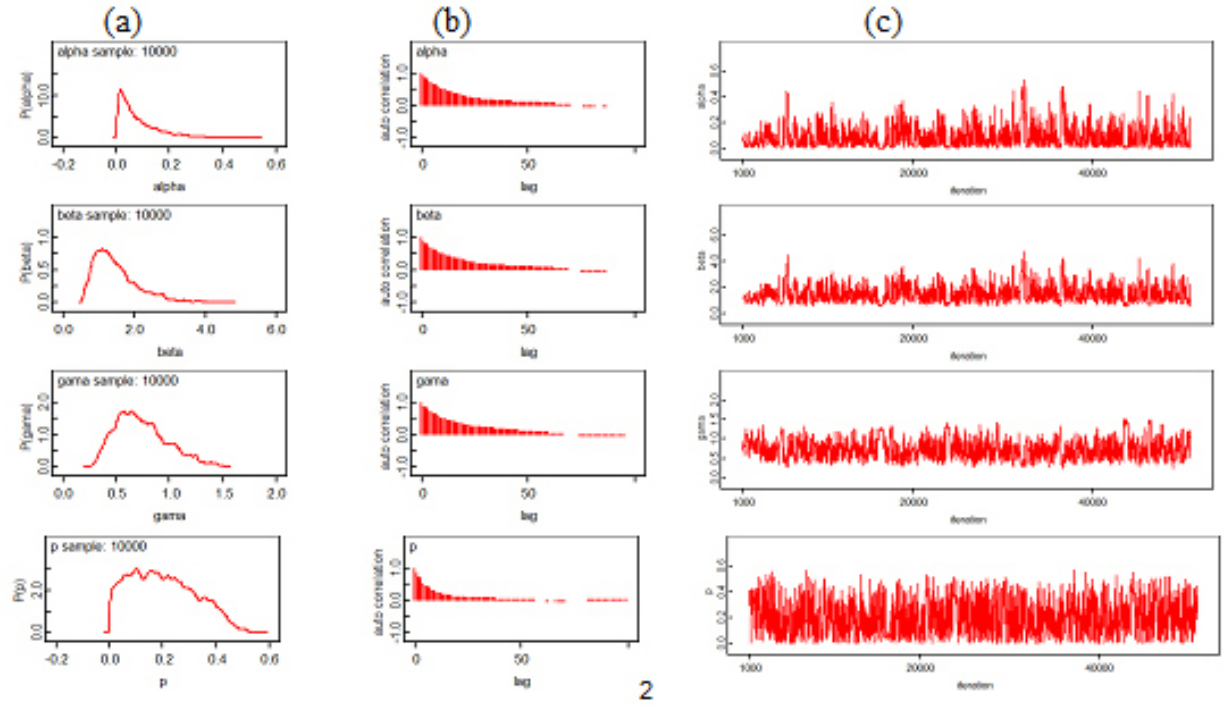


Figura C.41: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WG de fração de cura de não-mistura.

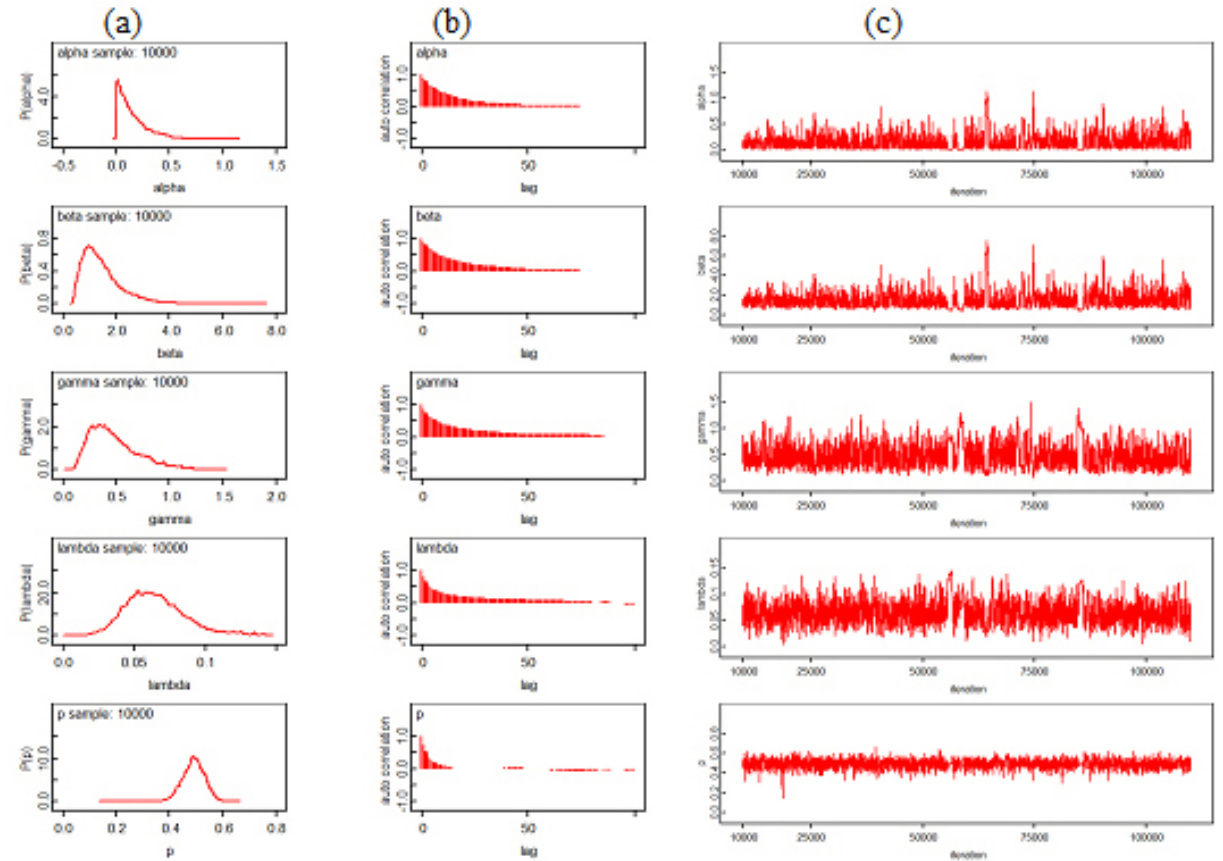


Figura C.42: Gráficos da densidade a posteriori (a), autocorrelação (b) e histórico (c) das cadeias geradas para o modelo da distribuição WMG de fração de cura de não-mistura.